HAL open science

Cliométrie de l'histoire globale des transitions anthropologique, cognitive, institutionnelle, démographique et économique

Kué Gilles Gaba

▶ **To cite this version:**

Kué Gilles Gaba. Cliométrie de l'histoire globale des transitions anthropologique, cognitive, institutionnelle, démographique et économique. Sociologie. Université de Lyon, 2021. Français. NNT : 2021LYSE1010 . tel-03603068

HAL Id: tel-03603068

https://theses.hal.science/tel-03603068

Submitted on 9 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N°d'ordre NNT : 2021LYSE1010

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale ED 486
SCIENCES ÉCONOMIQUES ET DE GESTION

Spécialité de doctorat : Sciences Economiques

Soutenue publiquement le 03/02/2021, par :
Kué Gilles Gaba

Cliométrie de l'histoire globale des transitions anthropologique, cognitive, institutionnelle, démographique et économique

---

Devant le jury composé de :

Cécile Bastidon, Maître de Conférences à l'Université de Toulon	Rapporteure
Etienne Farvaque, Professeur à l'Université de Lille	Rapporteur
Jean Belin, Maître de Conférences à l'Université de Bordeaux	Examinateur
Nathalie Havet, Maître de Conférences à l'Univ. Claude Bernard Lyon 1	Examinatrice
Caroline Hillairet, Professeure à l'ENSAE ParisTech	Examinatrice
Stéphane Loisel, Professeur à l'Université Claude Bernard Lyon 1	Directeur de thèse
Antoine Parent, Professeur à l'Université Paris 8	Co-Directeur de thèse

Remerciements

Mes premiers remerciements s'adressent à mes deux directeurs de thèse Stéphane Loisel et Antoine Parent : Stéphane Loisel pour son exigence, sa disponibilité et sa générosité par tout temps et en tout lieu, malgré ses lourds engagements ; Antoine Parent pour sa rigueur, sa culture économique, mais aussi ses encouragements efficaces et son enthousiasme communicatif. Les avoir côtoyés a été un bonheur et une bénédiction.

Cécile Bastidon et Etienne Farvaque m'ont fait l'honneur d'avoir rapporté cette thèse : qu'ils en soient remerciés pour la lourde tâche de lecture et leurs remarques enrichissantes qui m'ont aidé à aller plus loin.

Je suis reconnaissant à Jean Belin, Nathalie Havet et Caroline Hillairet pour avoir accepté de faire partie du jury de thèse et également pour leur lecture attentive, questions et remarques.

Anani Olympio m'a incité à entreprendre cette thèse et a été mon compagnon de route jusqu'à la toute fin : je lui exprime ici toute ma gratitude.

Toute ma reconnaissance à Karim Barigou pour son accompagnement technique et méthodologique, et également sa générosité et sa gentillesse naturelle dans les derniers moments.

Yahia Salhi et Xavier Milhaud dont j'ai parasité le bureau durant plusieurs années, ont été d'une gentillesse exquise et d'une disponibilité totale pour tous types de questions. Un grand merci à eux.

L'accueil et l'intégration au laboratoire SAF a été très agréable : j'en remercie Alexis Bienvenüe, Denis Clot, Anne Eyraud-Loisel, Nicolas Leboisne, Christian Robert, Didier Rullière, Pierre-Emmanuel Thérond, Julien Vedani.

J'ai également une pensée à mes camarades Pierre, Sarah, Romain, Pierrick, Claire, Tachfine.

Je n'oublie pas mes amis de Lomé pour leur aide et assistance durant ces années éprouvantes.

Enfin, merci de tout cœur à Harold Gaba pour ses encouragements, Joanita Gaba pour son enthousiasme et Ayite Camille Gaba mon *sparring-partner* intellectuel et compagnon de tous les jours.

A mon père Helmut Mawutodji Gaba,
et à ma mère Jeannette de Medeiros

Résumé

Une des originalités de notre approche cliométrique de l'histoire globale, a été de proposer une grille de lecture quantitative, multidisciplinaire et systémique. Notre grille est composée de cinq domaines interdépendants, désignés par l'acronyme ACIDE : Anthropologie, Cognition, Institutions, Démographie et Economie. Cette grille ACIDE est destinée à l'analyse multidisciplinaire des phénomènes économiques et en histoire globale.

Dans le domaine institutionnel, nous avons défini deux indices originaux : l'un mesurant l'inclusivité économique et l'autre la non-inclusivité économique. Cela nous a permis de montrer de manière inattendue que la Chine a été plus inclusive économiquement que la moyenne de l'OCDE et la moyenne mondiale de 1995 à 2010. Même si l'inclusivité économique chinoise a chutée depuis lors, elle reste relativement proche de la moyenne mondiale.

Une autre originalité de notre approche institutionnelle consiste à caractériser les institutions via, d'une part, des indicateurs « externes » liés à la souveraineté et la puissance, et d'autre part, des indicateurs « internes » de gouvernance et de respect des droits de l'individu. Nous définissons ainsi une matrice 2x2 de mesure des institutions regroupées en quatre catégories : les institutions politiques internes, les institutions politiques externes, les institutions économiques internes et les institutions économiques externes.

L'analyse des dynamiques transitionnelles ACIDE, nécessite également de développer un temps transitionnel qui évolue au rythme de la transition étudiée et non du temps calendaire. Ce temps transitionnel permet de faire une analyse comparative de trajectoires de pays dont les transitions se déroulent à des époques, à des rythmes et à des intensités hétérogènes.

Dans un domaine ACIDE donné, il est également nécessaire de détecter les phases historiques des trajectoires nationales et d'utiliser ces phases pour faire une analyse séquentielle des trajectoires. Notre apport original a été d'adapter l'analyse séquentielle, usuellement biographique, à l'analyse macroéconomique globale.

Une des limites des analyses longitudinales est la disponibilité de données quantitatives riches sur le temps long. Notre solution originale à cette limite a été d'inférer les trajectoires moyennes sur le temps long, via une coupe transversale à une date donnée. Pour cela nous avons proposé une méthode permettant de définir et tester statistiquement un temps transitionnel pour le domaine étudié et qui permet alors de pallier l'absence du temps calendaire dans la coupe transversale. Avec le temps transitionnel ainsi créé, il est possible de caractériser une trajectoire transitionnelle moyenne en l'absence de temps calendaire.

Il est intéressant d'utiliser l'aspect universel et propagatoire de la transition dans un domaine disciplinaire ACIDE donné, pour effectuer des prévisions actuarielles de long terme pour les pays à transition plus tardive, en prenant en compte, non seulement leurs trajectoires propres, mais également l'historique des trajectoires des pays à transition précoce. Notre solution originale à ce besoin passe par la création de variables exogènes dites « variables cliométriques » qui permettent de faire le lien statistique entre les deux types de pays. Nous appliquons notre technique dans le domaine démographique pour les prévisions actuarielles à long terme des taux de mortalité par âge et sexe.

Les modèles généralisés et optimisés que nous avons introduits ont permis d'améliorer les modèles actuariels classiques à l'horizon de 50 ans. Pour le pays émergent étudié (Inde), la prise en compte des « variables cliométriques » a abouti à des prévisions beaucoup plus précises à l'horizon de 40 ans.

Enfin, nous appliquons la grille cliométrique multidisciplinaire ACIDE à l'étude des déterminants multidisciplinaires de la mortalité dans l'infanterie française, au cours de la première Guerre mondiale.

Nous avons développé une mesure originale de la mortalité en conflit armé, qui tient compte des différences d'effectifs entre unités, mais aussi de la durée de mobilisation des régiments et de l'hétérogénéité du risque de mortalité auquel sont soumises les différentes unités.

Nous avons ainsi pu montrer que durant le Grande Guerre de 14-18, la part de mortalité dans l'infanterie expliquée par les facteurs anthropologiques, cognitifs, et socio-économiques est approximativement égale à la moitié de la part de mortalité dus aux causes purement militaires.

Mots-clés : cliométrie, histoire globale, temps transitionnel, transition économique, transition démographique, transition anthropologique, transition cognitive, transition institutionnelle, prospective actuarielle, mortalité dans l'infanterie

Abstract

One of the originalities of our cliometric approach to global history was to propose a quantitative, multidisciplinary and systemic reading grid. Our grid is composed of five interdependent domains, designated by the acronym ACIDE : Anthropology, Cognition, Institutions, Demography and Economics. This ACIDE grid is intended for the multidisciplinary analysis of economic phenomena and global history.
In the institutional domain, we defined two original indices: one measuring economic inclusiveness and the other measuring economic non-inclusiveness. This allowed us to unexpectedly show that China was more economically inclusive than the OECD and world average from 1995 to 2010. Although Chinese economic inclusiveness has fallen since then, it remains relatively close to the world average.
Another originality of our institutional approach consists in characterizing institutions through, on the one hand, "external" indicators related to sovereignty and power, and on the other hand, "internal" indicators of governance and respect for individual rights. We thus define a 2x2 matrix for measuring institutions grouped into four categories: internal political institutions, external political institutions, internal economic institutions and external economic institutions.
The analysis of ACIDE's transitional dynamics also requires the development of a transitional time that evolves at the pace of the transition under study and not at the rate of calendar time. This transitional time makes it possible to make a comparative analysis of the trajectories of countries whose transitions take place at heterogeneous times, rhythms and intensities.
In a given ACIDE domain, it is also necessary to detect the historical phases of national trajectories and to use these phases to make a sequential analysis of the trajectories. Our original contribution has been to adapt the sequential analysis, usually biographical, to the global macroeconomic analysis.
One of the limitations of longitudinal analyses is the availability of rich quantitative data over time. Our original solution to this limitation was to infer the average trajectories over time, via a cross-section at a given date. To do so, we proposed a method to define and statistically test a transitional time for the domain under study, which then allows us to compensate for the absence of calendar time in the cross-section. With the transitional time thus created, it is possible to characterize an average transitional trajectory in the absence of calendar time.
It is interesting to use the universal and propagatory aspect of transition in a given ACIDE disciplinary domain to make long-term actuarial forecasts for countries with later transition, taking into account not only their own trajectories, but also the historical trajectories of countries with early transition. Our original solution to this need involves the creation of exogenous variables known as "cliometric variables" that allow us to make the statistical link between the two types of countries. We apply our technique in the demographic field for long-term actuarial forecasts of mortality rates by age and sex.
The generalized and optimized models that we have introduced have improved the classical actuarial models with a 50-year horizon. For the emerging country under study (India), the inclusion of "cliometric variables" has led to much more accurate predictions with a 40-year horizon.
Finally, we apply the ACIDE multidisciplinary cliometric grid to the study of the multidisciplinary determinants of mortality in the French infantry during the First World War.
We have developed an original measure of mortality in armed conflict, which takes into account the differences in size between units, but also the duration of mobilization of regiments and the heterogeneity of the mortality risk to which the different units are subjected.
We have thus been able to show that during the Great War of 14-18, the share of mortality in the infantry explained by anthropological, cognitive, and socio-economic factors was approximately equal to half of the share of mortality due to purely military causes.

Keywords: cliometrics, global history, transitional time, economic transition, demographic transition, anthropological transition, cognitive transition, institutional transition, actuarial foresight, infantry mortality

Table des matières

Remerciements *2*

Résumé *4*

Abstract *6*

Table des matières *8*

Introduction générale *11*

1. Introduction 12
2. Contributions et principaux résultats 15

Chapitre 1. Approche cliométrique conceptuelle de l'histoire globale *24*

1. Table des matières 25
2. Introduction 26
3. La transition économique 29
4. La transition démographique 36
5. Anthropologie culturelle 44
6. Institutions politiques et économiques 63
7. Cognition, éducation et capital humain 84
8. Temps transitionnel 94
9. Conclusion 97

Chapitre 2. Analyse séquentielle des trajectoires macro-économiques, dans l'histoire globale récente *100*

1. Table des matières 101
2. Introduction 102
3. Détection de phases 104
4. Analyse des séquences d'états *(State sequences Analysis STS)* 108
5. Analyse des séquences d'évènements *(Time stamped Events analysis TSE)* 125
6. Contributions des composantes structurelles macroéconomiques, à la variation du PIB par habitant 130
7. Conclusion 142
8. Bibliographie 145

Chapitre 3. Analyse cliométrique transversale des transitions ACIDE *147*

1. Table des matières 148

2. Introduction 149
3. Rappels 151
4. La transition économique 152
5. La transition démographique 181
6. Anthropologie culturelle 206
7. Institutions politiques et économiques 225
8. Cognition, éducation et capital humain 248
9. Analyse multi-transitionnelle 279
10. Conclusion 297

Chapitre 4. Prospective actuarielle et cliométrique des taux de mortalité 302

1. Table des matières 303
2. Introduction 304
3. Relation entre le modèle Lee-Carter et la régression PCR 307
4. Modèle CBD 311
5. Régression sur composantes principales 312
6. Modèle Logit-PCR 315
7. Amélioration du modèle Logit-PCR 316
8. Régression PLS 318
9. Améliorations du modèle Logit-PCR en utilisant la régression PLS 320
10. Variable et modèle PCR cliométrique 321
11. Résultats empiriques 329
14 Conclusion 355
15 Bibliographie 359

Chapitre 5. Analyse cliométrique de la Mortalité dans l'infanterie pendant la Grande Guerre 361

1. Table des matières 362
2. Introduction 363
3. Base de données cliométriques de la Grande Guerre 365
4. Géomatique des mouvements des RI 370
5. Définition, mesure de la mortalité journalière dans les RI 376
6. Analyse multidimensionnelle ACIDE de la mortalité dans l'infanterie 378
7. Bilan des facteurs de la mortalité des régiments 388
8. Conclusion 390

9. Bibliographie 391

Conclusion 392

Bibliographie 396

Annexes 410

1. Annexes 411
2. Courbes longitudinales des variables passives 415

Table des figures 432

Liste des tableaux 445

Introduction générale

1. Introduction

Cette thèse est au confluent de plusieurs disciplines : cliométrie, histoire globale, anthropologie, cognition, institutions, démographie, économie et mathématiques appliquées. Elle s'inscrit pleinement dans l'objectif de l'équipe pluridisciplinaire CAC, *Cliometrics and complexity* (2020), qui est de renouveler l'approche de l'histoire à partir de méthodes nouvelles pour cette discipline, c'est à dire empruntés à d'autres sciences. Cette thèse entend illustrer par l'exemple, cet objectif d'enrichir et réinterpréter l'histoire économique.

Elle s'inscrit donc pleinement dans la tradition cliométrique (qui est d'analyser l'histoire à l'aide de modèles économiques enrichis d'applications économétriques), mais préfigure ce que sera le second temps de cette « révolution cliométrique », en recourant de façon plus générale aux outils de la complexité et des mathématiques appliquées pour analyser l'histoire globale (transdisciplinaire, de temps long, de dimension mondiale mais basée sur les diversités des trajectoires nationales connectées).
Cette thèse a donc une ambition forte : celle de faire partie de l'avant-garde d'un champ de recherche en plein devenir, celui de la cliométrie et complexité (multidisciplinarité, non-linéarité des associations).

Dans cette perspective novatrice, cette thèse embrasse un sujet immense et fondateur de la littérature en macroéconomie et en histoire économique, qui est la question des étapes du développement économique et des phases de transition. Ce thème structurant élaboré par Thompson (1929), Adolphe Landry (1934), Frank Notestein (1945), Alan Fisher (1939), Colin Clark (1940), Fourastié (1949), Rostow (1959), Sauvy (1980) est ici étendu et enrichi de double manière : par l'analyse statistique systématique des dynamiques historiques des transitions démographiques et économiques ; par l'adjonction, la définition et l'analyse de trois autres types de transition définies dans la thèse : les transitions anthropologique, institutionnelle, et cognitive.
Cette approche statistique des différentes transitions, inédite dans la littérature, est par essence pluridisciplinaire : l'un des objets de la thèse est de s'interroger sur les séquences de transformations macroéconomiques des pays dans l'histoire.
Cette thèse, en repartant de la décomposition sectorielle de l'économie, non plus dans une perspective statique comme chez les premiers auteurs, mais dynamique, vient aussi enrichir l'approche de la théorie de la croissance unifiée de Galor (2011).

Le premier objectif de cette thèse est la définition, la mesure et l'analyse des transitions économiques, démographiques, anthropologiques, institutionnelles, cognitives et de leurs interactions dans une approche de l'histoire globale par, d'une part, la mise en évidence des profils spécifiques de chaque transition, et d'autre part, l'étude des relations complexes que ces transitions entretiennent entre elles.

Le second objectif est l'application des connaissances acquises sur la grille multidisciplinaire et les transitions associées, pour les prévisions de mortalité (et donc d'espérances de vies) dans le domaine actuariel, et pour l'analyse multidisciplinaire des déterminants de la mortalité durant une période historique de guerre.

Originale par son objet, cette thèse l'est aussi par sa construction et son ambition théorique : la combinaison de l'analyse transversale contemporaine et de l'analyse longitudinale permet de produire une étude cliométrique aussi bien rétrospective (reconstitution historique), que prospective. Plusieurs champs disciplinaires (anthropologie, cognition, institutions, démographie, économie, ACIDE) sont analysés au travers du prisme d'une approche transversale et d'une approche longitudinale. Ce croisement disciplinaire rend notre approche systémique. Néanmoins, notre approche systémique de la

complexité des trajectoires historiques repose sur une unité de méthode : l'exploitation cliométrique de méthodes non-linéaires et complexes.

Les transitions étudiées dans la thèse mettent en évidence la présence généralisée de trajectoires historiques et d'associations non-linéaires. L'étude de ces non-linéarités nécessite le recours à des méthodologies non-linéaires, capables de gérer des phénomènes de multicolinéarité des variables, en particulier, des méthodes de classification, des méthodes factorielles (ACP, AFCM), des modèles PLS *(Partial Least Square)* ou d'arbre de segmentation, associées à l'analyse séquentielle d'états ou d'évènements historiques.

Notre apport par cette thèse est essentiellement d'enrichir l'approche cliométrique de la complexité par les méthodes ci-dessus énumérées.

Une ambition théorique importante de cette thèse est *in fine*, de savoir si, méthodologiquement, il est fondé de construire une méthode d'inférence longitudinale à partir d'une analyse en coupe transversale : la photographie à un instant donné d'un pays dans ses dimensions économiques, démographiques, anthropologiques, cognitives, institutionnelles, nous renseigne-t-elle sur son passé et sur son futur ?
Par ailleurs pouvons-nous capitaliser sur les phénomènes de diffusion globale de certaines transitions, pour faire de la prospective sur les trajectoires des nations à transition plus tardive, à l'aide d'une modélisation « déphasée » des trajectoires historiques de pays à transition plus précoce.
Et enfin, dans quelle mesure pouvons-nous « expliquer » et caractériser des faits historiques, par des variables exogènes à spectre large donc nécessairement multidisciplinaire ?

Cette thèse n'aurait pu être réalisée sans un travail considérable en amont de collecte des bases de données existantes dans le monde et le développement d'une application web dédiée à l'analyse transitionnelle et séquentielle des phases.

Afin de mener à bien une analyse pluridisciplinaire, les données récoltées couvrent les champs suivants : démographie, économie, santé, éducation, ressources naturelles, énergie, commerce international, emploi, gouvernance et institutions, environnement, infrastructures de transport et communications, armées et sécurité, religion, anthropologie culturelle. Au total cette base comporte 3213 variables remontant pour certaines au 18ème siècle.
Les données collectées ne sont pas le fruit d'un travail d'archives, ce sont toutes des données de seconde main, soit déjà mises en ligne, soit extraites d'articles ou d'ouvrages et reconstituées à partir d'eux.
En revanche, ces données ont toutes été retravaillées afin de les harmoniser dans une base commune. Le travail de programmation (codage) et d'harmonisation des données a été effectué entièrement par l'auteur en langage SAS. Les modules d'analyse ont également été conçus et développés par l'auteur, avec le langage R. La partie web de l'application est développée en langages PHP/MySQL et R, adossés aux *web frameworks* CodeIgniter et Shiny respectivement.
Cette base a vocation à tomber à terme dans le domaine public en étant accessible directement en ligne à l'usage des chercheurs et du grand public.

Une autre base de données a été constituée pour le chapitre 5 avec une application web dédiée à l'analyse cliométrique de la mortalité, des mouvements de régiments, et des batailles durant Grande Guerre. Contrairement à la première, cette base de données est une base de première main et a nécessité un travail de collecte, codification, correction, considérable. Cette base comporte 993 variables de description des régiments, de leurs mouvements, des combats et batailles, et d'informations du recensement de 1911 relatives aux départements d'origine des régiments. En outre, ces données ont été complétées par des données papiers tirées de Guillot et Parent (2018), Gilles, Guironnet et Parent (2014), Todd (1990, 1998, 1999), Le Bras et Todd (1981, 2013).

L’application est composée de deux modules, l’un consacré à l’analyse explicative de la mortalité dans les régiments, l’autre à une animation géomatique des troupes et des combats totalement inédite et qui est une contribution aux savoirs sur la première guerre mondiale. Les langages de programmation et de calcul sont identiques à ceux utilisés pour les chapitres précédents.

Les revues bibliographiques relatives aux thèmes abordés sont intégrées directement dans les chapitres concernés ; ce qui rend cette introduction générale relativement courte, et les chapitre d’autant plus denses.

2. Contributions et principaux résultats

La thèse se structure en cinq chapitres dont nous présentons les contributions et principaux résultats importants ci-après.
Dans le chapitre 1, nous traitons les questions de définition et d'indicateurs des trajectoires nationales dans les domaines ACIDE ; nous y abordons les aspects économiques des disciplines connexes (anthropologie, cognition, institutions, et démographie) dans notre grille multidisciplinaire. La notion de temps transitionnel (différent du temps calendaire) y est également développée, ainsi que le test statistique associé.

Le chapitre 2 est consacré à l'application des méthodes séquentielles à l'histoire économique récente, à travers la détection et l'analyse d'états (phases) structurelles de l'économie, ainsi que des évènements de changements structurels majeurs.

Dans le chapitre 3, nous inférons sur les trajectoires historiques nationales et complexes ACIDE de temps long, à partir de données internationales en coupe transversales. Nous y testons également l'hypothèse d'une théorie transitionnelle dans les domaines (anthropologie, cognition, institutions) où cette théorie n'existe pas.
Cela nécessite la définition d'un temps transitionnel pour le domaine étudié, qui pallie l'absence du temps calendaire en coupe transversale. Avec ce temps transitionnel, il est alors possible de définir et de caractériser une trajectoire transitionnelle moyenne en l'absence de temps calendaire.

Le chapitre 4 a consisté, à utiliser le phénomène de diffusion géographique de la transition démographique pour faire des prévisions actuarielles de long terme des taux de mortalité par âge et sexe dans un pays émergent, à partir des trajectoires des taux de mortalité dans les pays à transitions démographiques précoces (typiquement les pays européens) ; nous y avons également proposé des améliorations par généralisation et optimisation des méthodes actuarielles classiques s'appliquant à tous les pays.

Enfin, le chapitre 5, applique la grille multidisciplinaire ACIDE, pour analyser, rechercher, et quantifier les déterminants ACIDE et militaires de la mortalité dans l'infanterie durant la première guerre mondiale ; nous y développons également une mesure de la mortalité quotidienne par régiment qui tient compte simultanément de l'effectif initial du régiment, de sa durée de mobilisation, et du risque de mortalité durant sa période de mobilisation, ainsi qu'un logiciel internet permettant de suivre au jour le jour les concentrations et combats par département durant la Grande Guerre.

Chapitre 1 : Approche cliométrique conceptuelle de l'histoire globale

L'histoire globale est une méthode d'étude historique qui se distingue (Testot *et al.*, 2015, p.6) par quatre caractéristiques essentielles : la transdisciplinarité (anthropologie, démographie, science politique, archéologie, sociologie, par exemple), l'approche en temps long, la vision globale (pays interconnectés à l'échelle mondiale), et les échelles multiples d'analyse (jeu de va-et-vient contrastés, entre le global et le local, le long terme et le court terme, le niveau macro avec le niveau micro).

La cliométrie se définit comme l'application de modèles économiques enrichis d'applications économétriques à l'histoire quantitative et institutionnelle. Elle permet une lecture des évènements historiques dans une perspective de temps long, avec une capacité de prospective à long terme et multigénérationnelle dans les domaines étudiés.

Notre démarche cliométrique est basée sur une grille de lecture multidisciplinaire anthropologique, cognitive, institutionnelle, démographique et économique (ACIDE), appliquée à l'histoire globale.

Que ce soit dans les deux domaines ACIDE où un processus de transition moderne est déjà théorisé (i.e. économie et démographie), ou les trois autres domaines (i.e. anthropologie, cognition, institutions) pour lesquels il n'existe pas de théories transitionnelles quantitatives établies, la question qui se pose est celle d'établir une définition précise de ces domaines telles que nous souhaitons les étudier, ainsi que les indicateurs statistiques nationaux permettant de les mesurer sans ambiguïté.

La transition économique moderne est typiquement le passage d'une économie agraire, vers une économie de service. Avec au passage, une poussée industrielle transitoire, qui décline plus ou moins fortement selon la tradition industrielle des pays.
Cependant, nous avons constaté qu'il ne suffit pas de disposer des indicateurs de poids relatif (en parts d'emplois ou de valeur ajoutée) des trois grands secteurs économiques pour établir sans ambiguïté la phase transitionnelle dans laquelle se trouve un pays à une date donnée ; il a été nécessaire d'utiliser des indicateurs structurels complémentaires : le pourcentage de la population en âge de travailler (15-64 ans), le taux d'emploi de la population en âge de travailler, la productivité par emploi dans chaque secteur.

Le concept de transition démographique désigne le passage d'un régime avec fécondité et mortalité « élevées », à un régime à fécondité et mortalité nettement plus basses. Cette transition est initiée par une chute de la mortalité, suivie quelques décennies plus tard par celle de la fécondité.
Malgré l'intérêt certain du solde migratoire dans le parcours démographique d'un pays, les taux de mortalité et de natalité ont été suffisants pour l'analyse de la transition démographique.
Toutefois, nous avons observé que durant les transitions démographiques (notamment européennes), les périodes de plus fort accroissement naturel correspondent aux périodes de solde migratoire négatif, se manifestant par des sorties massives nettes de population. Il existe ainsi une corrélation forte et positive entre le solde naturel et l'émigration.

L'anthropologie culturelle couvre les habitudes acquises par les êtres humains en société. De nombreux modèles anthropologiques existent dans la littérature et nous avons pu en exposer les principaux, et étudier la disponibilité et la fiabilité des données nationales les concernant : le modèle d'Edward T. Hall (1950-1955, au *Foreign Service Institute FSI, USA*), le modèle Kluckhohn-Strodtbeck (1961, USA), le modèle Geert Hofstede (1970, université de Limburg, Maastricht, Pays-Bas), le modèle Fons Trompenaars et Hampden-Turner (1993), le modèles des traits d'histoire de vie (stratégies évolutives d'adaptation, *life history theory*), le modèle Le Play-Todd des systèmes familiaux préindustriels, et enfin le modèle Inglehart-Welzel issu des enquêtes de la *World Value Survey* (Inglehart et Welzel, 2005).
D'un point de vue économique, nous montrons que les politiques publiques de taxation et de redistributions des revenus sont très fortement déterminées par les profils anthropologiques des pays.
En ajoutant une dimension inégalitaire au modèle anthropologique Inglehart-Welzel (2005), nous proposons les variables suivantes pour caractériser la dynamique anthropologique : valeurs laïques-rationnelles (vs valeurs traditionnelles), valeurs d'expression personnelle (vs valeurs de survie), valeurs inégalitaires (mesurées par l'indice de Gini après taxes et redistribution, qui a une très forte corrélation avec les variables anthropologiques).

Les institutions économiques désignent l'ensemble des règles légales régissant l'activité économique.
En combinant les recommandations institutionnelles des principaux économistes « institutionnalistes », nous pouvons conclure que les bonnes institutions économiques sont des instruments stratégiques au service du développement économique, de l'indépendance et de la puissance du pays. Elles doivent, pour être opérationnelles, s'adapter à la culture du pays et à ses phases de développement, mais aussi à sa situation concurrentielle vis-à-vis de ses partenaires commerciaux.

Les institutions régissant les échanges externes servent de levier d'adaptation mondiale, tandis que les institutions économiques internes doivent renforcer l'inclusivité (pluralité des participants et des bénéficiaires), le respect des droits de propriété et des libertés économiques individuelles. Historiquement, l'activité qui procure les plus importants gains de productivité est l'industrie, qui mérite de ce fait un soin particulier.
Après une revue des indicateurs institutionnels existants, nous constatons une carence de mesures de l'inclusivité économique.
En utilisant la différence entre le rang économique (PIB par habitant) et le rang humain (indice de développement humain) d'un pays, nous avons défini deux indices originaux : l'un mesurant l'inclusivité et l'autre la non-inclusivité économiques. Il s'agit ici d'une des originalités de notre approche institutionnelle.
En calculant notre indice d'inclusivité des institutions économiques en 2015, nous constatons que contrairement à l'hypothèse avancée par Acemoglu et Robinson (2015, p.130) en se basant sur les institutions politiques, la Chine a été plus inclusive économiquement que la moyenne de l'OCDE et la moyenne mondiale de 1995 à 2010. Même si l'inclusivité économique chinoise a chutée depuis lors, elle reste relativement proche de la moyenne mondiale.

Une autre originalité de notre approche institutionnelle consiste à caractériser les institutions via, d'une part, des indicateurs « externes » liés à la souveraineté, et d'autre part, des indicateurs « internes » de gouvernance et de respect des droits de l'individu. Nous définissons ainsi une matrice 2x2 de mesure des institutions regroupées en quatre catégories : les institutions politiques internes, les institutions politiques externes, les institutions économiques internes et les institutions économiques externes.

Nous caractérisons la dynamique cognitive par les quatre variables suivantes : le taux d'adultes non scolarisés de 25-64 ans ; le taux d'adultes dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint ; les compétences cognitives numériques-littéraires-logiques mesurées par le test de QI (dont de nombreuses études que nous citons, ont établi la forte corrélation avec les tests de compétences cognitives numériques et littéraires qui malheureusement se limitent à quelques pays).

Une autre originalité de notre approche consiste à développer une notion de temps transitionnel différent du temps calendaire, qui s'écoule au rythme des phases et évènements transitionnels : ce temps transitionnel a pour fonction de comparer la dynamique d'entités (géographiques ou autre) différentes traversées par un même phénomène transitionnel, mais qui se déroule à des périodes historiques (donc calendaires) différentes.
Techniquement, nous définissons le temps transitionnel, comme une série temporelle qui a une dépendance monotone forte avec le temps. Nous mesurons cette dépendance monotone (concordante ou discordante) par le coefficient de corrélation des rangs de Kendall τ, permettant ainsi de tester l'hypothèse qu'une série temporelle puisse être utilisée comme temps transitionnel.
Nous définissons l'âge transitionnel d'un pays, comme différence entre le temps transitionnel et sa valeur minimale internationale dans la période de transition.

Chapitre 2 : Analyse séquentielle des trajectoires macroéconomiques, dans l'histoire globale récente

L'analyse séquentielle des parcours biographiques (*life course analysis, event history analysis*) est un champ de recherche très actif dans les sciences sociales qui consiste à réaliser des statistiques descriptives, des typologies, et des modèles explicatifs, de séquences d'états ou d'évènements biographiques.

L'originalité de notre approche cliométrique de l'histoire économique globale a consisté à appliquer l'analyse séquentielle biographique, aux données longitudinales macroéconomiques.

Nos contributions se situent à quatre principaux niveaux, que nous détaillons dans les paragraphes qui suivent.

Premièrement, la détection des phases économiques historiques des 27 dernières années, pour chacun des pays de la planète (à l'exception des *petits* Etats à économies spécifiques), par des méthodes de classification statistiques sur composantes principales.
Deuxièmement, l'analyse des séquences d'états (*State sequences Analysis STS*) nous permet de réaliser une analyse séquentielle des trajectoires macroéconomiques (encore appelées séquences) des pays à l'échelle planétaire sur la période étudiée.
Le troisième niveau de contribution utilise l'analyse des séquences d'évènements (*Time stamped Events analysis TSE*) pour étudier et possiblement anticiper les séquences d'évènements (changements d'état) économiques majeures.
Enfin, quatrièmement, toujours dans l'analyse longitudinale, nous proposons une méthode d'estimation séquentielle (puis par agrégation) des contributions à la variation pluriannuelle du PIB par habitant, des principaux secteurs économiques (agriculture, industrie, services), et de la dynamique démographique de la population en âge de travailler.

La détection des phases économiques historiques nous a permis de découvrir quatre phases (ou états) structurelles possibles de l'économie, sur la période étudiée.
Le nombre de pays dans l'état (ou phase) économique la moins avancée (State.1) au sens de la productivité agricole est en décroissance continue. L'état économique suivant (State.2) décroit légèrement, tandis que le nombre de pays en phase 3 (State.3) est en forte hausse. Le nombre de pays les plus développés (State.4) connait également une augmentation.
Il y a des progrès économiques (au sens de la hausse de la productivité agricole) sur tous les continents, caractérisés par des passages d'états moins développés à des états plus développés.
L'état le plus fréquent par an est State.2, mais c'est l'état State.3 (plus avancé que State.2) qui est le mode depuis 2016.
Globalement chaque pays semble avoir passé l'essentiel de son temps dans un état spécifique, avec une probabilité réduite (inférieure à 4%) d'évolution vers un autre état. Il y a donc une inertie importante, avec des évolutions relativement lentes.
La matrice de probabilités de transition montre l'absence de sauts de deux états ou plus, d'une année à l'autre. Le phénomène d'inertie est très visible avec des probabilités de rester dans le même état d'une année à l'autre, qui sont toutes supérieures à 96%.

Le nombre optimal de classes de trajectoires nationales est de quatre.
Fait notable, la classe de trajectoire qui a le plus faible effectif (24 pays) est la classe la moins développée économiquement. Et le groupe le plus large (43 pays) est celui des pays ayant une trajectoire de type 2 donc qui viennent de quitter la classe la moins développée.

Des tests du Khi-2 permettent d'identifier des sous-séquences bien spécifiques à chacune des quatre classes de trajectoires nationales.
La zone géographique des pays explique 23.4% (pseudo-R^2) de la variabilité des trajectoires. La zone de plus forte variabilité des trajectoires des pays est l'Europe. Avec la régression par arbre de segmentation, les covariables, *Region, High income (yes, no), Middle income (yes, no) et Low income (yes, no)* expliquent 34% de la discrépance (hétérogénéité des trajectoires).

Il est possible de calculer le nombre de réalisations d'une sous-séquence particulière d'évènements, à l'échelle mondiale sur la période étudiée.
Sur la période historique étudiée, l'évènement (State.1 vers State.2), s'est produit pour 14 (12%) trajectoires de pays, tandis que l'évènement (State.3 vers State.4) ne s'est produit que dans 5 pays (4%).

L'évènement intermédiaire (State.2 vers State.3) s'est réalisé dans 21 (18%) pays. Sur la période étudiée, aucun pays n'a pu faire un saut progressif de deux états économiques successifs.

Des algorithmes spécifiques (méthodes de fouilles des données séquentielles) permettent d'extraire ces règles d'association séquentielles en se basant sur le seuil de support exigé et un niveau de confiance minimal. Les principales règles observées indiquent que les pays progressent sans rupture, ni régression durable en terme de développement économique, sur la période étudiée.

Chapitre 3 : Analyse cliométrique transversale des transitions ACIDE

L'absence de données économiques structurelles sur le temps long a été une limite à l'exploration historique du chapitre 2, notamment en ce qui concerne le manque de données nationales de long terme sur le secteur industriel. Cela a une répercussion importante étant donné que le secteur industriel est une composante fondamentale de la structure économique d'un pays, et donc pour son suivi par rapport à la théorie de la transition économique.
Ayant défini la grille multidisciplinaire au chapitre 1, et ayant constaté les limites de l'approche longitudinale en termes de profondeur d'historique au chapitre 2, nous allons à présent tenter une inférence longitudinale à travers une coupe transversale grâce l'usage du temps transitionnel précédemment défini au chapitre 1.

Pour étudier les phénomènes transitionnels des domaines ACIDE, par une approche internationale en coupe transversale, nous-nous heurtons immédiatement à une question de temporalité : comment analyser la dimension temporelle dès lors que le temps est absent, par définition, de l'analyse transversale ?
Nous prenons en compte cette objection en introduisant les notions de temps transitionnel et d'âge transitionnel dans l'analyse transversale. Cette analyse est totalement inédite dans la mesure où dans les travaux d'économie et en cliométrie, cette question n'est jamais abordée.
Nous définissons le temps transitionnel comme une série temporelle qui a une dépendance monotone forte avec le temps. Nous mesurons cette dépendance monotone (concordante ou discordante) par le coefficient de corrélation des rangs de Kendall τ, permettant ainsi de tester l'hypothèse qu'une variable (série temporelle) puisse être utilisée comme temps transitionnel.
Nous définissons l'âge transitionnel d'un pays, comme différence entre le temps transitionnel et sa valeur minimale internationale dans la période de transition.
Dans la logique transversale, nous désignons par « phases », les classes d'âge transitionnel obtenues par classification statistique (sur les variables descriptives du domaine et du temps transitionnel choisi) des pays : les pays d'une même classe statistique sont homogènes en termes d'âge transitionnel. L'introduction de l'âge transitionnel dans une coupe transversale, est également celle du temps transitionnel, puisque l'âge transitionnel se compte en unités de temps transitionnel. Le contenu informationnel de l'approche transversale est ainsi rendu comparable et partiellement substituable à l'approche longitudinale : le temps transitionnel se substituant au temps calendaire longitudinal. Nous détaillons cette démarche par domaine ACIDE dans les paragraphes qui suivent.

Dans les deux domaines ACIDE où un processus de transition moderne est déjà théorisé (i.e. économie et démographie), la question qui se pose est celle de sa reproductibilité historique dans les pays émergents ou à faibles revenus.
Dans les trois autres domaines ACIDE (i.e. anthropologie, cognition, institutions) pour lesquels il n'existe pas de théories transitionnelles quantitatives établies, se posent les questions suivantes : quels temps transitionnels adopter dans une approche transversale ? Peut-on proposer une théorie transitionnelle moderne dans ces domaines ?

Du point de vue multi-transitionnel, les questions qui se posent concernent le timing croisé des phases entre les cinq transitions ACIDE : quelles phases sont concomitantes ou antagonistes ? Certaines phases

dans une transition spécifique peuvent-elles être considérées comme limitantes ou au contraire stimulantes, pour le déroulement du processus transitionnel dans d'autres domaines ?

La transition économique moderne est typiquement le passage d'une économie agraire, vers une économie de service, avec une poussée transitoire de l'emploi et la valeur ajoutée relatifs industriels.
Le taux d'emplois agricoles peut empiriquement être utilisé comme symétrique du temps transitionnel.
L'analyse transversale transitionnelle économique nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014. Par identification, ces 4 phases correspondent respectivement aux phases théoriques « agricole » ou parfois « démarrage », « démarrage » ou parfois « expansion », « achèvement », « achèvement » ou parfois « tertiaire ».
L'étude transversale confirme assez clairement la validité et la réalité contemporaine du processus de transition économique théorique de Fourastié.

Le concept de transition démographique désigne le passage d'un régime avec fécondité et mortalité « élevées », vers un régime à fécondité et mortalité nettement plus faibles. Cette transition est initiée par une chute de la mortalité, suivie quelques décennies plus tard par celle de la fécondité.
Le taux de mortalité infantile peut empiriquement être utilisé comme symétrique du temps transitionnel démographique, dans une étude en coupe transversale.
L'analyse transversale transitionnelle démographique, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014. Actuellement nous pouvons avancer qu'il n'y a plus de pays en période pré-transitoire, et seulement quelques rares pays en phase 1 théorique de la transition démographique.
L'étude réalisée confirme la validité et la réalité contemporaine du processus théorique de transition démographique.

Pour rappel, l'anthropologie culturelle couvre les habitudes acquises par les êtres humains en société.
En ajoutant une dimension inégalitaire au modèle anthropologique Inglehart-Welzel, nous proposons les variables suivantes pour caractériser la dynamique anthropologique : valeurs laïques-rationnelles (vs valeurs traditionnelles), valeurs d'expression personnelle (vs valeurs de survie), valeurs inégalitaires (mesurées par l'indice de Gini après taxes et redistribution, qui a très forte corrélation avec les variables anthropologiques).
Le ratio d'années d'études femmes *versus* hommes de 25-34 ans peut empiriquement être utilisé comme temps transitionnel anthropologique dans une étude en coupe transversale.
Nous désignons par transition anthropologique, une série de profondes et durables transformations anthropologiques d'un pays, pouvant s'étaler sur plusieurs générations.
L'analyse transversale transitionnelle anthropologique, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014.
Le schéma transitionnel théorique obtenu révèle en moyenne qu'après une baisse dans les trois premières phases, les valeurs d'« expression personnelle » (*versus* valeurs de « survie »), font un fort rebond vers un niveau sans précédent en phase 4.
Après une brève baisse « initiale », les valeurs « laïques-rationnelles » (*versus* valeurs « traditionnelles ») sont en progrès continu de la phase 2 à la phase 4. Nous notons toutefois une nette inflexion de la croissance entre les phases 3 et 4.
Quant aux valeurs « inégalitaires », elles suivent dans une première séquence, une évolution en cloche, suivi d'un léger rebond en phase 4.

Les institutions économiques désignent les règles légales régissant l'activité économique.
En nous restreignant au sous-domaine « politique interne » de notre matrice institutionnelle définie au chapitre 1, nous analysons la dynamique institutionnelle à travers les 2 variables suivantes : les libertés civiles et l'indice global de démocratie.
L'indice de développement humain peut empiriquement être utilisé comme temps transitionnel, dans une étude en coupe transversale. Cet indice mesure les résultats des dispositions institutionnelles dans

les domaines politiques, économiques, sanitaires et éducatives : il est donc légitime pour caractériser les institutions d'un pays à un moment donné.
Nous désignons par transition institutionnelle une série de profondes et durables transformations institutionnelles dans un pays, pouvant s'étaler sur plusieurs générations.
L'analyse transversale transitionnelle institutionnelle, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014.
Le schéma institutionnel transitionnel théorique obtenu révèle en moyenne que l'indice global de démocratie présente une forte baisse « initiale », suivi d'un fort rebond en phase 2, qui se termine par une hausse nettement plus faible entre les phases 3 et 4.
Enfin, l'indice de développement humain, qui synthétise le fonctionnement opérationnel des institutions sanitaires, économiques et éducatives, est en hausse continue sur les 4 phases.

Nous caractérisons la dynamique cognitive par les quatre variables suivantes : le taux d'adultes non scolarisés de 25-64 ans ; le taux d'adultes dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint ; les compétences cognitives numériques-littéraires-logiques sont mesurées par le test de QI.
Nous utilisons empiriquement le nombre moyen d'années de scolarisation des hommes de 25-64 ans comme temps transitionnel, dans l'étude en coupe transversale.
Nous désignons par transition cognitive une série de profondes et durables transformations cognitives de la population d'un pays, pouvant s'étaler sur plusieurs générations.
L'analyse transversale transitionnelle cognitive nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2012.
Selon le schéma théorique de transition cognitive que nous proposons, le score de compétences cognitives d'un pays augmente de manière continue en moyenne de 74 à 98.
Parallèlement, les adultes uniquement diplômés du primaire représentent la majorité des adultes alphabétisés diplômés durant les deux premières phases transitionnelles, avec un pic à 30% dans la deuxième phase. Cette proportion va décroître ensuite pour finir à 7% en phase 4.
A partir de la phase 3, les adultes uniquement « diplômés » du secondaire, représentent la majorité des adultes alphabétisés. Leur proportion passe de 11% en phase 1, à 44% en phase 4.
Quant aux diplômés du supérieur (tertiaire), leur proportion parmi les adultes augmente continument de la phase 1 (à 3%) à la phase 4 (à 21%).

Les associations sont d'intensité forte entre les cinq transitions ACIDE et l'ensemble de ces transitions se déroulent plus ou moins simultanément dans le même sens.
Par-contre il y a des phénomènes limitatifs : pour atteindre les phases « avancées » dans une transition, il peut être nécessaire d'avoir franchi les phases « initiales » ou « intermédiaires » dans d'autres transitions. Ainsi la phase 4 économique (la plus avancée actuellement) a un certain degré d'incompatibilité avec les phases 1 et 2 anthropologiques, les phases 1, 2 et 3 cognitives, et les phases 1 et 3 institutionnelles.

Chapitre 4 : Prospective actuarielle et cliométrique des taux de mortalité

La transition démographique moderne est le processus historique par lequel une population passe d'un régime démographique caractérisé par un taux de mortalité global et un taux de natalité élevés, à un nouveau régime caractérisé par des valeurs nettement plus faibles (division par 3, voire 4) de ces deux taux.
Nous remarquons empiriquement des décalages temporels et des différences d'allure entre les courbes des taux de mortalité selon les âges. Si une composante temporelle unique reflète la transition démographique du point de vue de la mortalité globale, la prise en compte de la transition sanitaire qui affecte les âges de manière différenciée, nécessite des composantes temporelles différenciées elles aussi par tranche d'âge.

La première originalité de notre approche consiste à utiliser le constat ci-dessus pour généraliser les modèles de mortalité classiques (notamment les modèles de Lee-Carter et CBD), en sélectionnant de manière optimale (minimisation du *MSE* de validation croisée), une ou plusieurs composantes principales (régression PCR) ou PLS (régression PLS), pour chaque âge modélisé. Nous appellerons ces extensions, des modèles respectivement, de type PCR-optimal et PLS-optimal.

L'autre constat historique majeur est celui d'une transition démographique universelle, se déroulant par vagues (ou cohortes) de pays. De plus, les rythmes des changements transitionnels sont assez hétérogènes d'une cohorte de pays à l'autre : les pays les plus récemment entrés en transition ont généralement des rythmes de changement plus rapides. Certains pays sont donc plus précoces dans la transition démographique que d'autres qui sont alors considérés comme étant en situation de rattrapage relatif.
La seconde originalité de ce chapitre consiste à utiliser l'histoire quantitative des mortalités des pays transitionnellement « plus âgés », pour tenter d'améliorer les prévisions à long terme de mortalité dans les pays transitionnellement « moins âgés » (cf. notion d'âge transitionnel défini au chapitre 1).
En pratique, ces améliorations pourraient s'appliquer aux études prospectives de la mortalité dans les pays émergents ou en développement.
La variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est une variable créée à partir des données historiques du pays A (transitionnellement « plus âgé » que le pays R), pour jouer un rôle de variable exogène dans les modèles de prévisions de mortalité du pays R.
Une telle variable sera ainsi utilisée comme variable explicative supplémentaire dans les modèles prospectifs de mortalité du pays R. Nous nommons les modèles utilisant la variable cliométrique, modèles cliométriques : nous en proposons deux types qui sont les modèles PCR-optimal-cliometric et PLS-optimal-cliometric.

Pour les travaux empiriques, nous utilisons dans ce chapitre, comme modèles classiques de référence, les modèles suivants : Lee-Carter, Lee-Carter Log-Poisson, CBD-logit-binomial.

L'utilisation du modèle généralisé et optimisé de type PCR-optimal a produit de meilleurs résultats que les modèles classiques de référence, en France et en Italie.
Pour la France, nous constatons un avantage systématique pour le modèle logit-PCR-optimal, quel que soit l'horizon de prévision étudié (jusqu'à 50 ans). Le modèle logit-PCR-optimal conserve son avantage relatif en présence de valeurs atypiques (guerres mondiales) ou d'un historique tronqué à 20 ans.
Pour l'Italie, nous constatons que, contrairement au cas Français où le modèle logit-PCR-optimal (sélection *stepwise*) est le plus performant jusqu'à 50 ans d'horizon de prévision, cet avantage n'excède pas l'horizon prédictif de 25 ans.
Nous en concluons que la supériorité empirique de performance du modèle logit-PCR-optimal par rapport aux modèles de référence, est relative au pays, à l'horizon de prévision et à l'âge dont la mortalité est modélisée (on peut conjecturer sur la période étudiée également). Ainsi, il revient à l'utilisateur de sélectionner les conditions d'utilisation du modèle logit-PCR-optimal, selon les résultats des prévisions testées.

Les modèles cliométriques ont permis d'améliorer de manière importante les prévisions de long terme dans le pays émergent testé, l'Inde.
Pour ce qui concerne l'Inde (en considérant les âges de manière globale), et quel que soit l'horizon (de 1 à 40 ans), les deux meilleurs modèles sont le modèle logit-PCR-Cliometric suivi de près par le modèle logit-PLS-Cliometric.
Avec ces deux modèles, les gains de précision obtenus par rapport aux modèles de référence sont assez importants, soit des réductions du MAPE *(mean absolute percentage error)* d'au moins -31% sur les modèles classiques de référence à un horizon de prévision de 40 ans (tous âges confondus).

Des travaux complémentaires pourront permettre de tester empiriquement une deuxième méthode d'estimation des variables cliométriques que nous avons également proposée, et comparer ainsi les performances avec la méthode qui a été mise en œuvre dans les travaux empiriques présentés.

Chapitre 5 : Analyse cliométrique de la mortalité dans l'infanterie pendant la Grande Guerre

Dans ce chapitre consacré à l'analyse cliométrique de la mortalité d'infanterie française pendant la Grande Guerre, nous appliquons la grille de lecture multidimensionnelle ACIDE (Anthropologie, Cognition et société, Institutions, Démographie, Economie) des origines géographiques des régiments, à l'analyse explicative de la mortalité de leurs soldats. A cette grille ACIDE, s'ajoute l'analyse selon des facteurs purement militaires de participation aux combats et d'appartenance à l'armée mobile.

Ce chapitre s'inscrit dans la continuité de la littérature existante consacrée aux morts de la Grande Guerre, complète et enrichit celle-ci.
Nous y présentons notre base de données cliométrique unique, historique, multidimensionnelle, pluridisciplinaire, ainsi que le module d'animation géomatique qui reproduit les mouvements quotidiens des régiments d'infanterie sur le front entre juillet 1914 et novembre 1918. Ce dernier a vocation à terme, à être mis à disposition de la communauté des chercheurs de l'histoire de la Grande Guerre.
Afin de mener à bien notre étude de mortalité dans l'infanterie, nous définissons une mesure de mortalité originale qui corrige des effets de taille, de durée d'existence, d'exposition au risque de mortalité des régiments. Cet indicateur inédit a vocation à être généralisé à tout type de conflit à travers l'histoire, ce qui est un apport à une cliométrie des conflits armés.

Afin de détecter les facteurs non-militaires pouvant affecter la mortalité dans l'infanterie, nous effectuons une l'analyse multidimensionnelle ACIDE de la mortalité selon les profils anthropologiques et socio-économiques des origines géographiques des régiments. Cette analyse ACIDE est complétée par l'adjonction de variables explicatives purement militaires : type de régiment (appartenant à l'armée mobile ou non), modalités de participation aux batailles et combats.

Pour effectuer les calculs répétitifs sans perte de temps, nous avons conçu et développé une application internet de mesure et de tests d'associations linéaires et non-linéaires, quel que soit le type de variable testée, qualitative ou quantitative. L'application est basée sur les technologies PHP/MySQL et R, adossées aux *web frameworks* CodeIgniter et Shiny respectivement.

Sur un plan méthodologique, afin d'estimer la part de variance de mortalité expliquée par un groupe de variables, nous avons systématiquement corrigé les effets de colinéarité des variables explicatives en utilisant des modèles de régression PLS, ou de régression sur axes factoriels issus d'ACP ou d'AFCM.

Sur un plan stratégique et organisationnel, l'analyse des valeurs anthropologiques d'une part, et des modalités associées au « timing » de l'engagement des régiments au combat d'autre part, est entièrement inédite et parvient à réduire la part traditionnellement dévolue au hasard dans la mort au front.

Sur le plan de la connaissance historique, nous avons pu mesurer et comparer les contributions à la mortalité des facteurs militaires et d'autres facteurs, au sein desquels, les aspects sociaux et anthropologiques l'emportent sur les facteurs économiques. En ce sens, cette thèse enrichit l'approche cliométrique existante de variables anthropologiques et sociétales.

Chapitre 1.
Approche cliométrique conceptuelle de l'histoire globale

1. Table des matières

Chapitre 1. Approche cliométrique conceptuelle de l'histoire globale 24

1. Table des matières 25

2. Introduction 26

3. La transition économique 29

3.1. Définitions 29

3.2. Mesures de la transition économique 35

4. La transition démographique 36

5. Anthropologie culturelle 44

5.1. Définitions 44

5.2. Modèles de l'anthropologie culturelle 46

5.3. Relations entre anthropologie et économie 57

6. Institutions politiques et économiques 63

6.1. Définition des institutions 63

6.2. Relations entre institutions et économie 64

6.3. Mesures des institutions 73

7. Cognition, éducation et capital humain 84

7.1. Définitions 84

7.2. Compétences cognitives et performances économiques 85

7.3. Mesures des compétences cognitives et du niveau d'éducation 86

8. Temps transitionnel 94

8.1. Propriété de concordance du subordinateur du temps 94

8.2. Coefficient de corrélation des rangs de Kendall et tendance temporelle monotone 94

8.3. Temps transitionnel 95

8.4. Age transitionnel 96

9. Conclusion 97

2. Introduction

Dans le cadre très actif des recherches sur thème de l'histoire global, Testot (Testot *et al.*, 2015, p.6) rappelle que l'histoire globale est avant tout une méthode, dont le *modus operandi* se décline sur quatre axes qu'il désigne par le sigle TLDE : T pour transdisciplinaire (anthropologie, démographie, science politique, archéologie, sociologie, sciences de l'environnement par exemple), L pour temps long, D pour longue distance (qui évoque le monde et ses connexions), E pour échelles au pluriel (jeu de va-et-vient contrastés, entre le global et local, le long terme et court terme, le macro et micro).
Plus récemment, Conrad (2017) prend en compte la connectivité historique croissante du monde et montre les limites intellectuelles à étudier les nations isolément. Cette perspective interconnectée du monde offre assurément une autre vue, une autre compréhension de l'histoire mondiale, différente des autres interprétations de l'histoire du monde. Vision qui fait de l'histoire globale, actuellement, un domaine des plus dynamiques et novateur de l'histoire.

Les liens entre l'économie et l'histoire globale sont réels et investigués par de nombreux auteurs.
Récemment, Baten *et al.* (2016) ont abordé dans un ouvrage collectif, l'histoire économique mondiale depuis le 16e siècle, en analysant les questions d'évolution, de divergences, de différences de revenus par habitant entre différentes régions du monde. Ces auteurs abordent également les relations entre les institutions, capital humain, le commerce international et la croissance économique.

La cliométrie utilise des données historiques quantitatives, applique des modèles économétriques, et adopte souvent une approche multidisciplinaire dans le cadre historique (Daudin, 2007).
Diebolt et Parent (2011) caractérisent la méthode cliométrique par « l'utilisation conjointe du discours historique, de la théorie économique, des outils statistiques et de la modélisation mathématique, pour tenter de transformer la vision moderne sur des débats historiques majeurs ». Ils explorent ainsi la cliométrie du capital humain, la cliométrie monétaire et financière, la cliométrie de la croissance et des cycles économiques.
La cliométrie permet donc une lecture quantitative des évènements historiques dans une perspective de temps long, et autorise la prospective à long terme multigénérationnelle des domaines étudiés.

En France, cette approche multidisciplinaire, quantitative et de long terme a été initiée par le courant de l'École des Annales dont les recherches sont publiées dans la revue « Annales d'histoire économique et sociale », fondée en 1929 par les historiens Lucien Febvre (1878-1956) et Marc Bloch (1886-1944), et dont Fernand Baudrel (1933 à 1985) prend la direction en 1956 après le décès de Lucien Febvre. Emmanuel Le Roy Ladurie chercheur de l'école des Annales publia une série remarquée d'ouvrage sur l'histoire quantitative du climat (Le Roy Ladurie [1997, 2004, 2006, 2009], Le Roy Ladurie et Séchet [2009]). Le Roy Ladurie fut une figure tutélaire pour Emmanuel Todd, historien dans la tradition des Annales.
Aux Etats-Unis, le courant de pensée cliométrique nait sous le nom de la *new economic history* et publie ses travaux dans la revue *Explorations in Economic History* dès 1949. La *Cliometrics society* est fondée en 1983, et en 1993, le « Prix Nobel » d'économie est attribué aux cliomètres Douglass North (1920-2015) et Robert Fogel (1926-2013).

A travers leurs définitions respectives, nous pouvons identifier les points communs entre l'histoire globale et la cliométrie : le temps long, la multidisciplinarité, la capacité à travailler sur plusieurs échelles d'espace ou de temps.
La cliométrie apporte à l'histoire globale son approche quantitative et ses modèles de mathématiques appliquées.
Pour Daudin (2007), l'utilisation des méthodes d'histoire globale et de la cliométrie en économie sont de nature non seulement à éclairer l'histoire économique et à tester les théories et hypothèses des

économistes, mais aussi à faire émerger de nouveaux faits stylisés et de nouvelles théories économiques, tel que l'on souhaité Kenneth Arrow et Robert Solow (cités par Daudin).
A titre d'exemple d'utilisation de l'histoire quantitative à des fin théoriques ou comptable, Daudin (2007) cite les concepts et le « développement de la comptabilité nationale dans les années 1930 et 1940, sous l'impulsion des travaux de Colin Clark et ceux de Simon Kuznets », auxquels nous pourrions rajouter la théorie de la transition économique (*structurale change*) par Colin Clark et Jean Fourastié, la « courbe de Kuznets » qui décrit la relation non-linéaire entre le revenu par habitant d'un pays et le niveau d'inégalité, et qui sont des construction de théories économiques à partir de l'histoire économique quantitative.

Pour ce qui est de la nécessité de l'approche multidisciplinaire en économie, l'apport de North (1994, 1996) a été décisif. Notamment à travers le rôle incontournable des institutions formelles ou informelles, mais aussi sur l'importance des aspects cognitifs et démographiques en économie.
North (1994) souligne l'intérêt d'une analyse de l'histoire économique, par une démarche multidisciplinaire, combinant les aspects institutionnels, cognitifs et démographiques : « En quoi une approche institutionnelle et cognitive peut-elle contribuer à améliorer notre compréhension du passé économique ? Tout d'abord, cela devrait avoir du sens à partir de la structure très inégale des performances économiques (…) Cette approche contribuera à notre compréhension de l'interaction complexe entre les institutions, la technologie et la démographie dans le processus global de changement économique. Une théorie complète de la performance économique impliquerait une telle approche intégrée de l'histoire économique. (…) Mais une des tâches majeures de l'histoire économique est d'intégrer ces différents domaines de Recherche »

Plus tard, North (1996) proposera un mécanisme dynamique d'interaction entre l'économie, les institutions, la démographie et la cognition : « Une compréhension globale de la performance économique dans le temps nécessite une fusion des théories sur les changements institutionnels, démographiques et relatifs au stock de connaissances afin d'avoir une approche globale des problèmes. Nous commençons seulement à explorer l'interaction entre ces trois sources de performance économique, mais je pense que nous pouvons aller loin dans le développement de modèles utiles de l'interaction entre elles, non seulement en termes d'institutions fournissant la structure d'incitation aux changements démographiques et technologiques, mais également dans la mesure où la démographie et la perception des 'impératifs' à travers le stock de connaissances, ont à leur tour façonné le changement des institutions. »

North (1996), illustre le mécanisme ci-dessus par l'exemple historique européenne : « Par exemple, la croissance économique moderne a pour origine la croissance du stock de connaissances associée à la révolution scientifique des XVIe et XVIIe siècles. Quelle est la source des attitudes, procédures et méthodes expérimentales qui ont caractérisé cette révolution ? C'était un phénomène occidental, lié évidemment aux développements institutionnels qui ont conduit à la montée du monde occidental d'une partie du monde relativement arriérée au Xe siècle à sa prééminence mondiale au XVIIIe siècle. Les développements institutionnels - politiques, économiques et sociaux - ont à leur tour reflété le développement d'un système de croyance.
D'où vient le système de croyance ? J'ai discuté ailleurs (North, in Davis 1995) qu'elle trouve son origine dans la façon dont les croyances religieuses (et la réaction à ces croyances) ont évolué dans l'Europe médiévale et proto-moderne, et la façon dont ces croyances étaient à leur tour fortement influencé par les expériences uniques qui ont caractérisé cette partie du monde ».

Pour aboutir à la croissance économique, North (1996) définit donc la chaine causale suivante : A partir de leurs « expériences », les humains mènent des « activités cognitives » qui génèrent des « croyances » et des « préférences ». Les « croyances » sont constituées d'un « stock de connaissances » scientifiques et de croyances « au-delà de la rationalité » composées de conjectures et de cosmogonies. Les « croyances » et « préférences » permettent ensuite aux acteurs économiques et politiques d'avoir une

« perception » de leurs « possibilités d'améliorer leurs positions concurrentielles ». Ces perceptions d'opportunités concurrentielles vont inciter les acteurs à effectuer des choix, notamment en innovations technologiques et/ou institutionnelles selon les incitations institutionnelles du moment.
La « structure institutionnelle » constitue la source des incitations des acteurs économiques et politiques. Lesquels incitations, lorsqu'elles sont favorables aux « activités productives », sont le vecteur d'investissements et de recherches à l'origine de la « révolution scientifique », aboutissant à la « révolution technologique » et à la « révolution démographique ».
Le « stock de connaissance » se retrouve donc augmenté par la « révolution scientifique », ce qui n'est pas sans conséquence sur les croyances et préférences.
Au final, les « développements institutionnels » incitant à la « révolution technologique » sont source de développement d'activités productives et de gains de productivités.

Et North (1996) de conclure : « Une théorie de la performance économique dans le temps impliquerait une intégration des théories du changement institutionnel, du changement démographique et de l'évolution du stock de connaissances ».

Il est important de noter que dans le schéma institutionnel de North, l'anthropologie culturelle, relève des catégories cognitive et institutionnelle « informelle ». Le terme « informel » étant utilisé par contraste avec les institutions « formelles » qui représentent le corpus juridique écrit.
Donc, la « théorie de la performance économique dans le temps » proposée par North (1996) relève d'une approche multidisciplinaire économique, institutionnelle, démographique, cognitive et implicitement anthropologique.
Dans notre approche, nous incluons explicitement les aspects quantitatifs d'anthropologie culturelle, qui sont implicites chez North.

Toujours dans une démarche quantitative multidisciplinaire (que nous qualifierions de cliométrique), Todd (1976) utilise des facteurs économiques et surtout démographiques (différentiel du taux de fécondités entre régions soviétiques, chute du taux de mortalité infantile) pour prédire la décomposition de l'Union soviétique.
Toujours dans la même veine, Todd (1983) proposa une explication des transformations politiques modernes au niveau mondial, notamment en relevant l'importance des facteurs anthropologiques (systèmes familiaux), éducatifs (donc cognitifs) et démographiques.
Todd (1984), utilisa également une grille cliométrique multidisciplinaire pour analyser les déterminants du processus de développement économique moderne.
Todd (1998) relève un lien fort entre l'anthropologie culturelle et l'activité industrielle.

L'originalité de notre démarche est de proposer une grille multidisciplinaire et quantitative, et donc systémique, basée sur des définitions concises et des indicateurs mesurables, dans une démarche cliométrique désignée par l'acronyme ACIDE (Anthropologique, Cognitive, Institutionnelle, Démographique et Economique). Cette grille multidisciplinaire et quantitative a vocation à être systématiquement utilisé pour l'analyse de phénomènes économiques, pour éviter les tautologies réductrices (expliquer l'économie par l'économie) et donner une profondeur historique, anthropologique, institutionnelle et culturelle à l'économie.
De manière générale, la grille cliométrique multidisciplinaire et quantitative proposée dans cette thèse, est destinée à un usage analytique de l'histoire globale.

L'originalité de cette thèse consiste aussi à développer une notion de temps transitionnel différent du temps calendaire, qui s'écoule au rythme des phases et évènements transitionnels : ce temps transitionnel a pour fonction de comparer la dynamique d'entités différentes (géographiques ou autres) traversés par un même phénomène transitionnel, mais qui se déroule à des périodes historiques (donc calendaires) différentes.

3. La transition économique

La transition économique au sens de Fourastié (1949) est une étude de la transformation, de l'activité
La transition économique au sens de Fourastié (1949) désigne la transformation de l'activité économique (produits, secteurs, ressources), de ses acteurs (démographie de l'emploi par secteur, formations et éducation, modes d'activité) et de son environnement (social, institutionnel, géographique).
Cette notion n'est pas assimilable à une théorie de la croissance économique ou des cycles économiques.
Le terme de « transition économique » ou « economic transition » n'existe pas dans la bibliographie internationale : Les termes équivalents trouvés dans la littérature sont « structural transformation » et surtout « structural change ».

3.1. Définitions

3.1.1. Définition de la transition économique

Une transition économique se définirait de manière générique comme une évolution de la vie économique (activités, acteurs, environnement), d'un état (ou configuration) plus ou moins stable sur plusieurs générations, vers un second état significativement différent et tout aussi relativement stable à une échelle multigénérationnelle.

Par rapport à cette définition, nous pouvons d'ores et déjà identifier historiquement (depuis la fin de la dernière glaciation vers -12 000 av. JC), deux très probables transitions économiques planétaires : la transition économique de la Révolution néolithique (dès -10 000 av. JC), et la transition économique de la Révolution industrielle (à partir du 18ème siècle)
Ces deux transitions ont touché par diffusion tous les continents à partir de foyers géographiques spécifiques. Les marques démographiques, sont visibles sur la *Figure 1.2*

Ainsi, la transition économique *moderne* (i.e. de la Révolution Industrielle) se définirait comme, le passage d'un régime économique relativement stable, où le taux de la population active agricole est « élevé » et dominant (plus de 70-80% de la population active, par exemple) et dans lequel le taux de la population active dans les services est « faible ». (10-20% par exemple), vers un autre régime relativement stable, où le taux de la population active agricole est significativement plus faible que lors de la phase initiale (moins de de 10-20% de la population active, par exemple), avec un taux de la population active des services est significativement plus élevé et dominant (plus de 70-80% de la population active, par exemple).

C'est donc typiquement le passage d'une économie agraire, vers une économie de service.
Avec au milieu, une poussée industrielle, croissante, puis décroissante de la population active, une hausse constante de la productivité des secteurs agricole et industriel, une baisse constante des prix réels des produits alimentaires et industriels, et enfin, une consommation de masse des produits alimentaires et industriels.

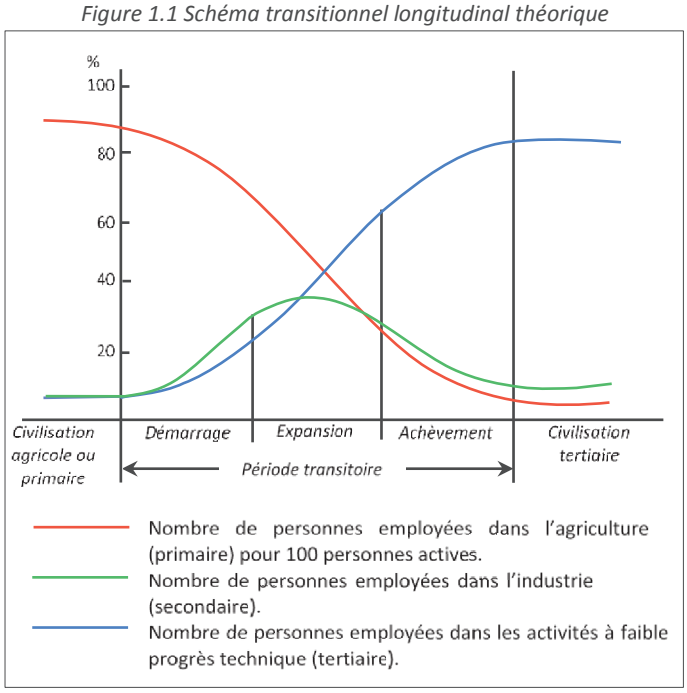
Dans ses travaux pionniers, Allan Fisher (1939), distingue des productions "primaires, secondaires et tertiaires" et classe les secteurs économiques selon une hiérarchie de besoins :
Le secteur primaire pour les besoins essentiels i.e. relatifs aux matières premières (agriculture, pêche, foresterie, mines), le secteur secondaire pour les produits finis (manufacture et construction), le secteur tertiaire pour des services comme les banques et les assurances.

Fisher (1939) indique que dans différentes étapes du développement économique, le secteur principal (en termes d'emploi et de production) varie : Le premier secteur initial est le primaire, puis le secondaire ; enfin, le tertiaire achève le processus.

Colin Clark (1940) dans les "Conditions du progrès économique", observe un transfert massif de main-d'œuvre de l'agriculture, vers les manufactures et les services, parallèlement à l'industrialisation et la croissance économique.

Ce phénomène de transition économique a été largement et profondément analysé par l'économiste et démographe Jean Fourastié (1949, 1965, 1975, 1977, 1980, 1982, 1986, 1987), aussi bien sous ses aspects économiques, que socio-démographiques.

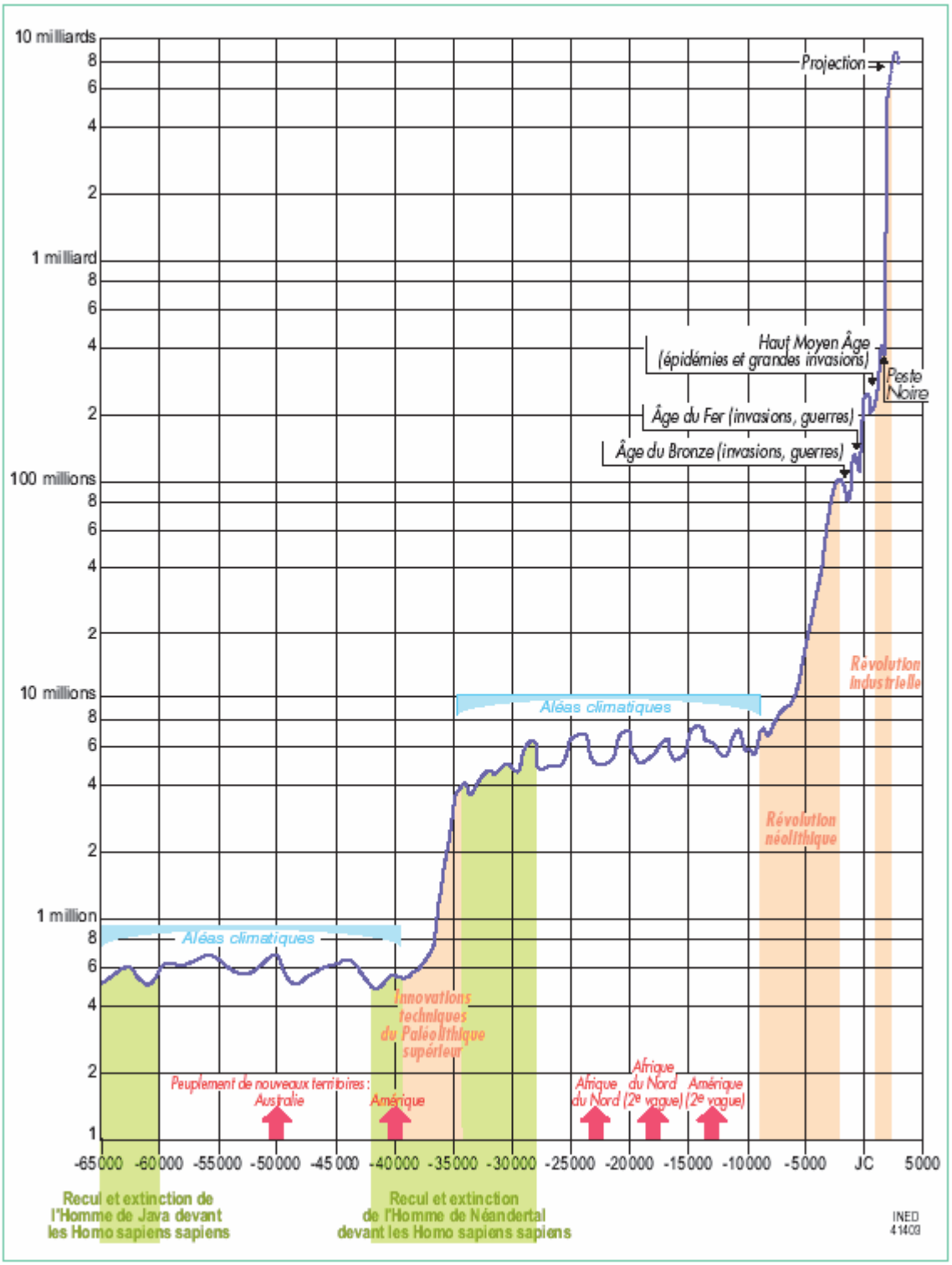
Figure 1.1 Schéma transitionnel longitudinal théorique



Source : Adapté de Fourastié (1989), p. 148.

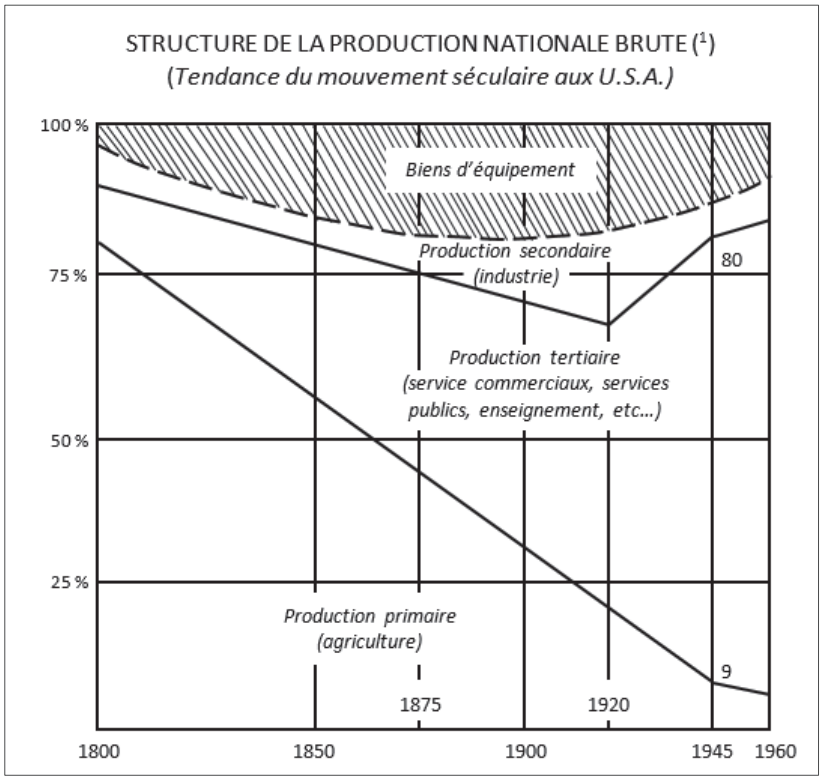
La transition économique, touche différents aspects dans chaque secteur économique : taux de la population active (ou d'emplois), production, producteurs, productivité, prix réels (exprimés en salaire horaire d'ouvrier non-qualifié), salaires réels (exprimés en salaire horaire d'ouvrier non-qualifié), durée de travail, lieux de travail (et de résidence des travailleurs), métiers (contenu des formations et durée des formations).

Figure 1.2 Evolution de la population mondiale



Source : Biraben (2003)

Figure 1.3 Exemple historique : Les USA, un cas avant-gardiste



Source : Jean Fourastié (1989), p. 105.

La figure ci-dessus illustrant la transition économique américaine indique pour le **secteur agricole**, un passage de 80% du PNB (1800) à 5% du PNB (1960). Soit une **division par 16** du poids du secteur primaire dans l'économie.

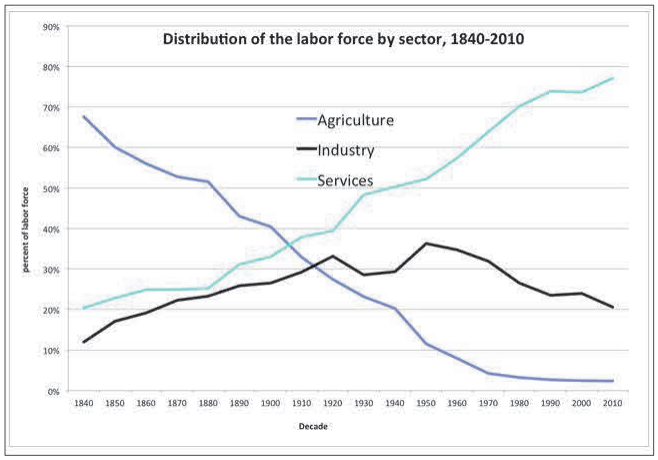
Pour le **secteur industriel**, deux phases successives se distinguent.
Une première phase de hausse où l'industrie passe de 6% du PNB (1800) à 31% du PNB (1920), soit une **multiplication par 5.2** du poids du secteur secondaire dans l'économie.
Une seconde phase de baisse, où l'industrie passe de 31% du PNB (1920) à 15% du PNB (1960), soit une **division par 2.1** du poids du secteur secondaire dans l'économie.
Le bilan des deux phases successives, est que l'industrie passe de 6% du PNB (1800) à 15% du PNB (1960), soit une **multiplication par 2.5** du poids du secteur secondaire dans l'économie.

Quant au secteur des services, il passe de 12-14% du PNB (1800) à 80% du PNB (1960), soit une **multiplication par 6** du poids du secteur tertiaire dans l'économie.
Le tertiaire est donc le secteur qui aura plus progressé en 1 siècle et demi de transition économique américaine, jusqu'à représenter l'essentiel du PNB, i.e. 80% : raison pour laquelle Fourastié parle de civilisation tertiaire, alors que d'autres y voyaient dans les premières phases l'amorce d'une civilisation industrielle.

L'agriculture est la grande perdante de la transition économique : elle occupe une place minime (5%) dans le PNB, alors qu'elle représentait l'essentiel (80%) de l'économie en 1800 avant la transition économique.

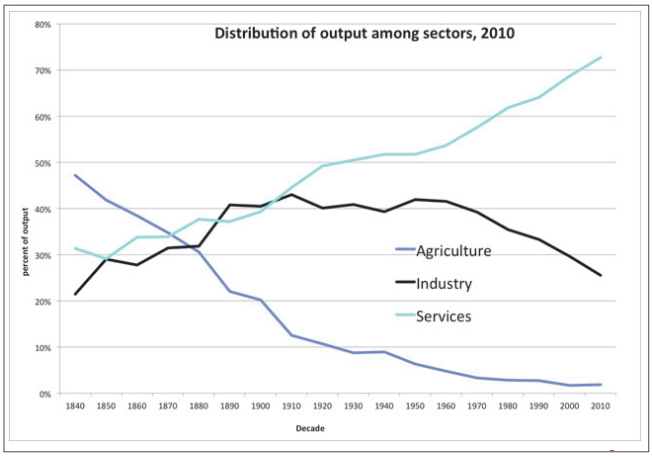
La transition économique américaine est illustrée par Johnston (2012), par les deux figures suivantes : la première (Figure 1.4) représente la transition économique selon les pourcentages de main d'œuvre sectoriels, tandis que la seconde (Figure 1.5) présente une approche par les parts de PIB sectoriels.

Figure 1.4 Répartition de la population active par secteur, 1840-2010



Source : Johnston (2012)

Figure 1.5 Répartition du produit entre les secteurs, 1840-2010

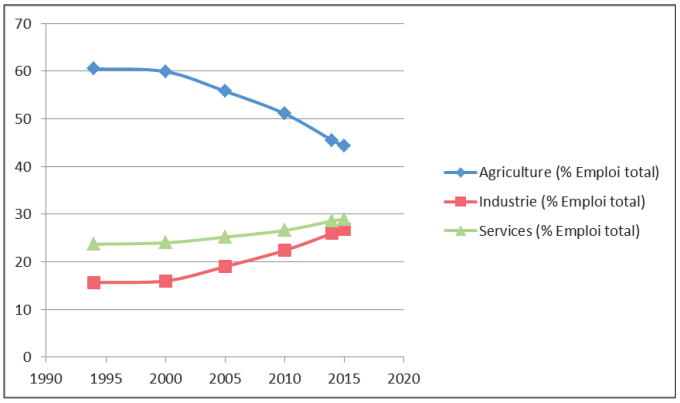


Source: Johnston (2012)

Sans être identiques, les allures des courbes de part de valeur ajoutée et de part d'emplois des secteurs, sont assez semblables.

Prenons à présent le cas d'un pays émergent, l'Inde :

Figure 1.6 Exemple historique : Le cas de l'Inde



Source : Perspective monde (2018). Données Banque Mondiale

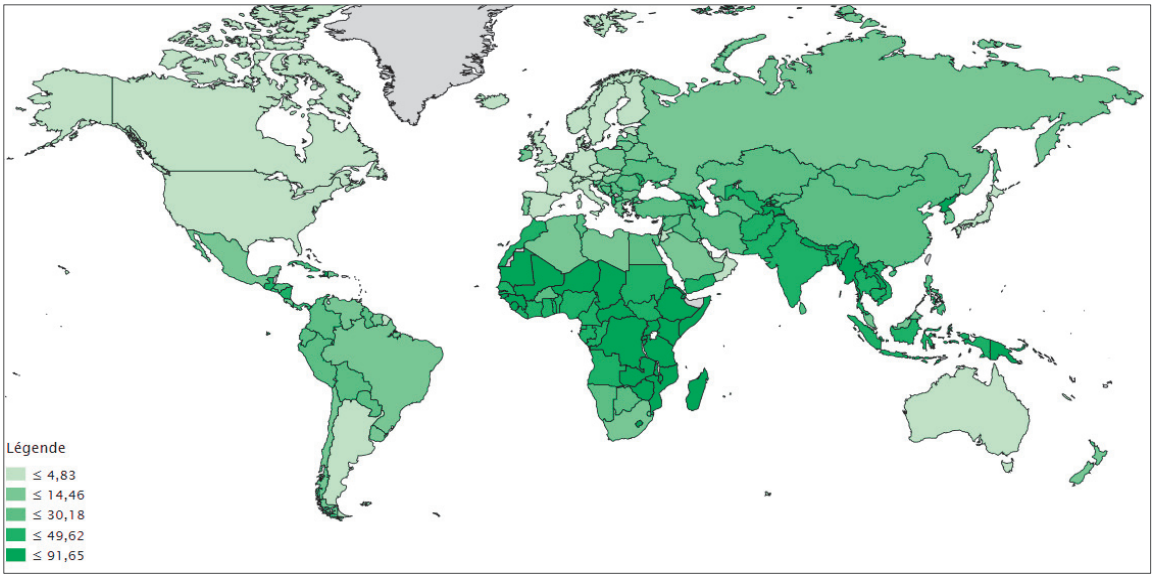
Avec sa répartition actuelle des emplois entre les 3 secteurs (29% services, 26% industrie, 46% agriculture), l'Inde se situe probablement au début de sa phase d'expansion, i.e. en début de phase 2 sur 3. Le pic industriel n'est

pas encore atteint, car l'agriculture représente encore 46% des emplois, bien loin des 5% attendus à la fin de la transition, et les services, ne représentent que 29% des emplois. Par conséquent, plusieurs décennies séparent probablement la situation actuelle du taux de 70-80% des emplois de service en fin de transition.

3.1.2. Positions transitionnelles par rapport au processus de transition économique

Dans la carte ci-dessous, il apparaît clairement qu'il existe plusieurs groupes de pays, par rapport à la part des emplois consacrée à l'agriculture (un indicateur clé de la transition économique).

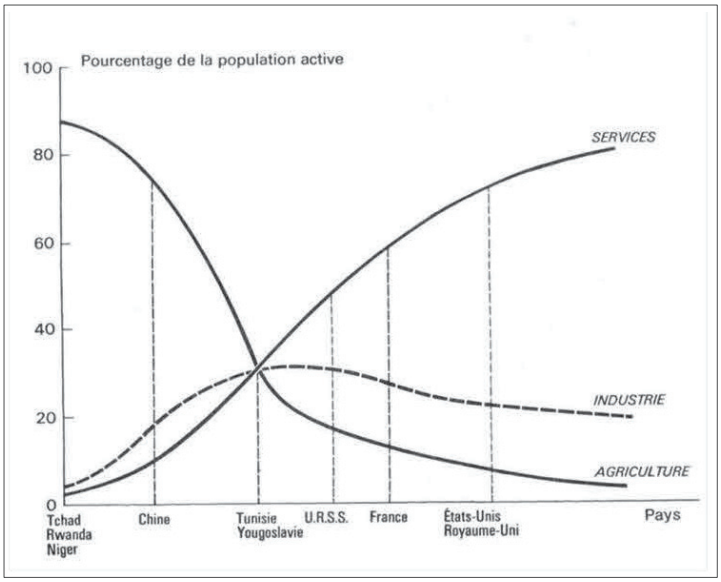
Figure 1.7 Carte mondiale indiquant la proportion d'emplois dans le secteur de l'agriculture (% de l'emploi total) 2015



Source : Perspective monde, date de consultation : 26/7/2019. Données Banque Mondiale.

De même, le graphique ci-dessous, positionne (en 1985) quelques pays, sur les trajectoires sectorielles de la transition économique.
Il est clair, également que les pays sont étalés sur les phases transitionnelles.
Ainsi, le décalage de la France par rapport aux USA (et aussi l'Angleterre) y apparait clairement.

Figure 1.8 Répartition de la population active dans quelques pays vers 1985



Source : Jean Fourastié (1989), p. 380.

3.2. Mesures de la transition économique

La transition économique moderne désigne typiquement le passage d'une économie agraire à une économie de services. Avec au passage, une poussée transitoire industrielle, qui décline plus ou moins fortement selon le pays.

Cependant, après une classification test des pays réalisée avec les pourcentages d'emplois dans les trois secteurs, nous avons constaté qu'il ne suffit pas de disposer des indicateurs de poids relatif (en parts d'emplois ou de valeur ajoutée) des trois grands secteurs économiques pour établir sans ambigüité la phase transitionnelle dans laquelle se trouve un pays à une date donnée ; il a été nécessaire d'utiliser des indicateurs structurels complémentaires : le pourcentage de la population en âge de travailler (15-64 ans), le taux d'emploi de la population en âge de travailler, la productivité par emploi dans chaque secteur.

En effet, une première classification empirique des pays selon les seules variables transitionnelles classiques (taux d'emplois respectifs des trois secteurs), fait apparaitre des classes hétérogènes en termes de revenu par habitant. Cela est incompatible avec l'idée d'assimiler ces classes à des pays se trouvant dans une même phase de transition économique.
La résolution du problème d'hétérogénéité du PIB par habitant dans les classes statistiques basées sur les variables transitionnelles classiques (i.e. pourcentage d'emploi dans les différents secteurs), consiste à créer des classes de pays semblables par rapport aux variables transitionnelles classiques, **mais** aussi semblables en termes de PIB par habitant.
Bairoch (2005, p.239) caractérise la structure d'une économie par la distribution des emplois dans les secteurs et le revenu par habitant.
En adoptant une décomposition comptable du PIB par habitant, régulièrement utilisée par les macroéconomistes (Cette et Strauss-Kahn [2003], Artus et Cette [2004], Bourdin [2007], Blanchet, Simon et Sylvander [2007], Heyer, Plane et Timbeau X. [2007]) pour comparer les évolutions structurelles entre pays, dont nous faisons une adaptation sectorielle, nous obtenons l'équation comptable suivante :

$$\frac{PIB}{Pop} = \sum_{i=1}^{3} \left(\frac{Pop_{15-64}}{Pop}\right) \cdot \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right) \cdot \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right) \cdot \left(\frac{PIB_i}{E_i}\right)$$

où : Pop est la population totale
Pop_{15-64} est la population âgée de 15 à 64 ans
E_i est le nombre d'emplois dans le secteur i
$\sum_{u=1}^{3} E_u$ est le nombre total d'emplois dans les trois secteurs
PIB_i est le PIB du secteur i

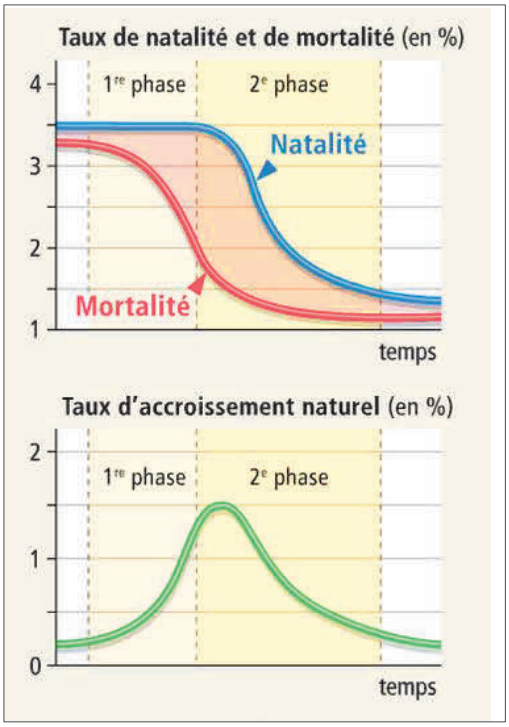
Les composants (entre parenthèses) de l'équation sont :
- Le pourcentage de la population en âge de travailler (15-64 ans)
- Taux d'emploi de la population en âge de travailler
- Variables transitionnelles classiques (i.e. taux d'emploi dans les différents secteurs)
- Productivité par emploi, dans chaque secteur

Nous utiliserons comme indicateurs structurels de l'économie d'un pays, les termes (entre parenthèses) de la partie droite de l'équation ci-dessus : en cas de classification statistique, cela permet d'atteindre par classe, le double objectif de similitude des pays en termes de variables transitionnelles classiques (taux d'emplois sectoriels), mais aussi en termes de PIB par habitant

4. La transition démographique

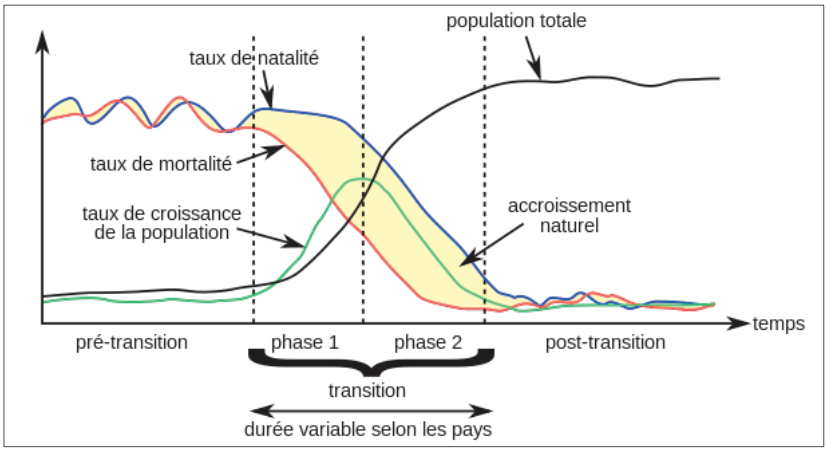
Le concept de transition démographique désigne le passage d'un régime démographique relativement stable, où la fécondité et la mortalité sont « élevées » vers un autre régime également relativement stable, où la fécondité et la mortalité sont significativement plus faibles que lors de la phase initiale, en passant d'abord par une chute de la mortalité, suivie quelques décennies plus tard par celle de la fécondité (Mazerolle, 2005).

Figure 1.9 Schéma transitionnel, phases et prospective



Source : Durand et al. (2008).

Figure 1.10 Schéma général de la transition démographique



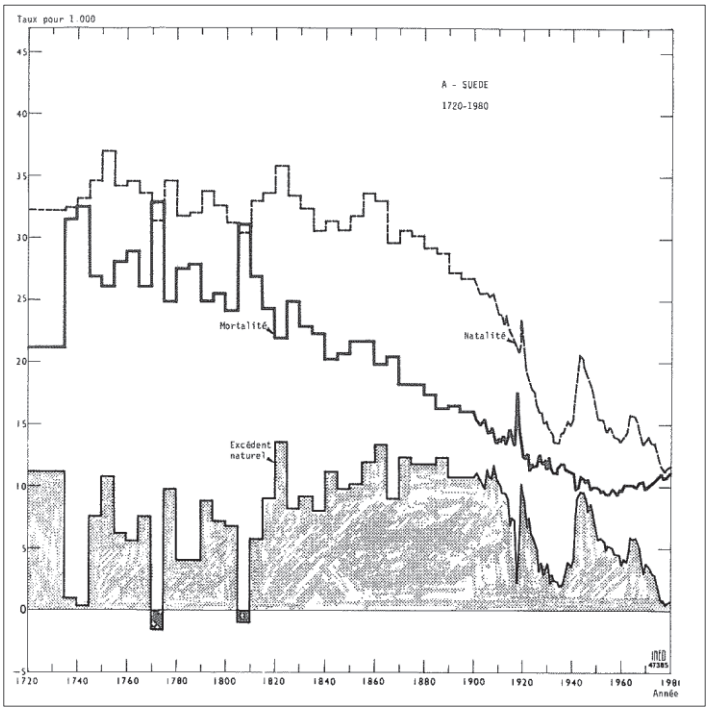
Source: Avdeev A., 2014.

L'expression « transition démographique » été introduite par le démographe américain Frank Notestein (1945).
Toutefois, Warren Thompson (1929), un autre démographe américain, avait très clairement identifié les deux principales phases de la transition démographique.

Le démographe Français, Adolphe Landry (1934), analysa également le phénomène de changement de régime démographique, dans son ouvrage « La révolution démographique ».
Nous illustrons ci-dessous à travers quelques exemples historiques emblématiques, la généralité des deux phases de cette transition démographique.

Exemples historiques : La Suède et l'Italie

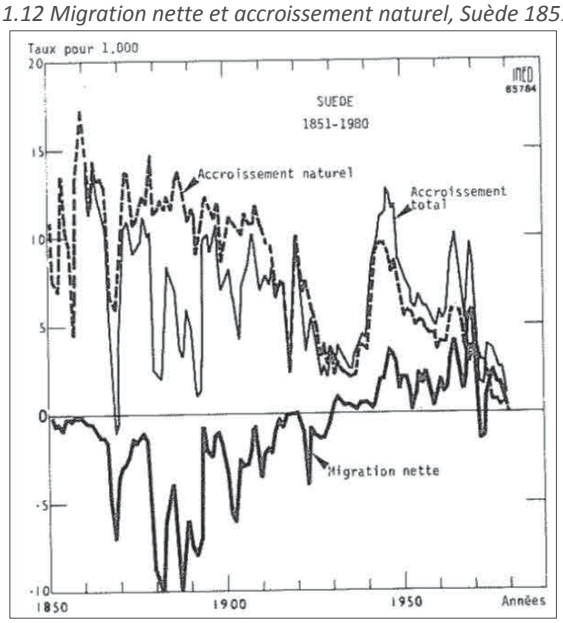
Figure 1.11 Mouvement naturel de la population, Suède 1720-1980



Source: Chesnais (1986) «The migratory transition», p. 224.

L'exemple suédois montre : une phase initiale très fluctuante, avec des niveaux élevés de mortalité et de natalité, puis une amorce de chute de la mortalité dès 1820-1830, suivie quelques décennies plus tard (dès 1850-1860) par une franche chute de la natalité. La phase de stabilisation finale n'est pas encore clairement discernable dans la figure ci-dessus.

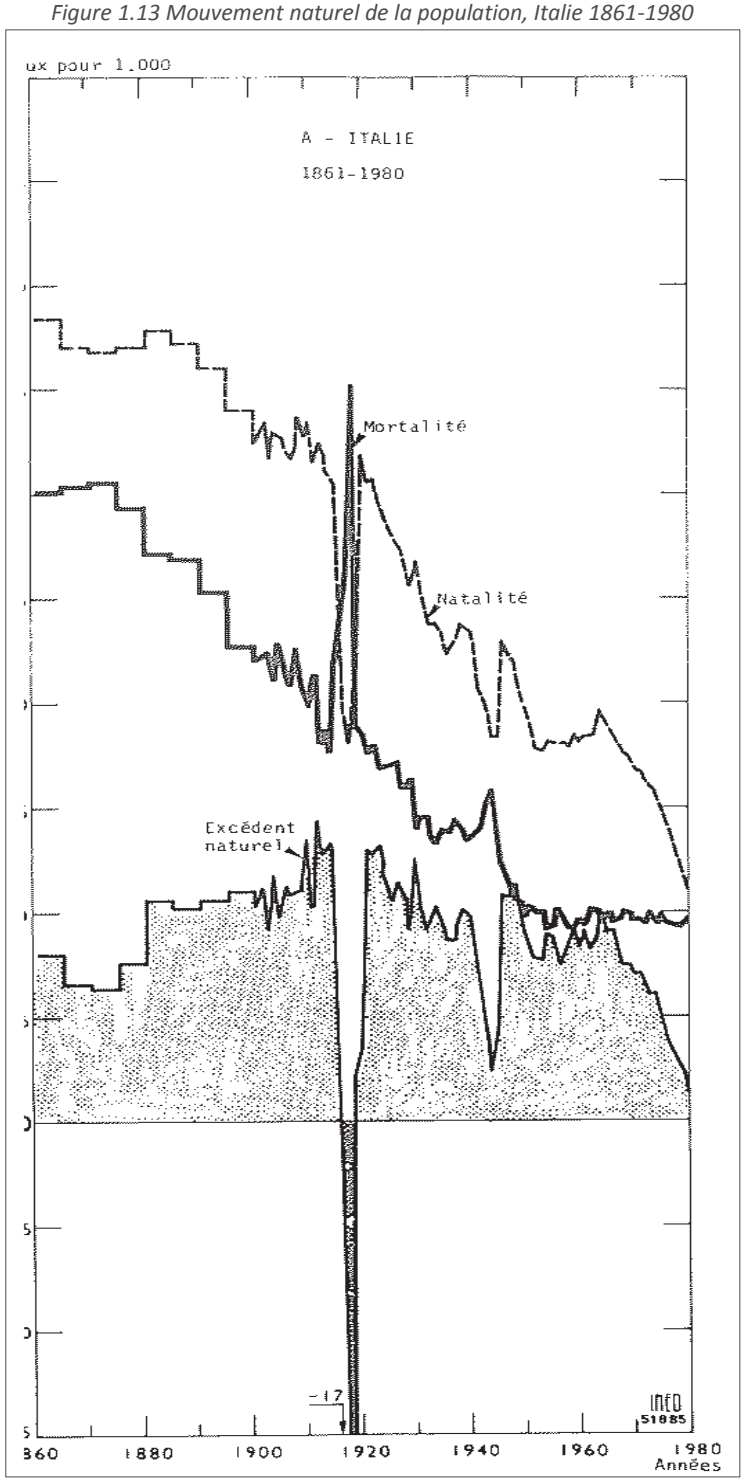
Figure 1.12 Migration nette et accroissement naturel, Suède 1851-1980



Source: Chesnais (1986) «The migratory transition», p. 179.

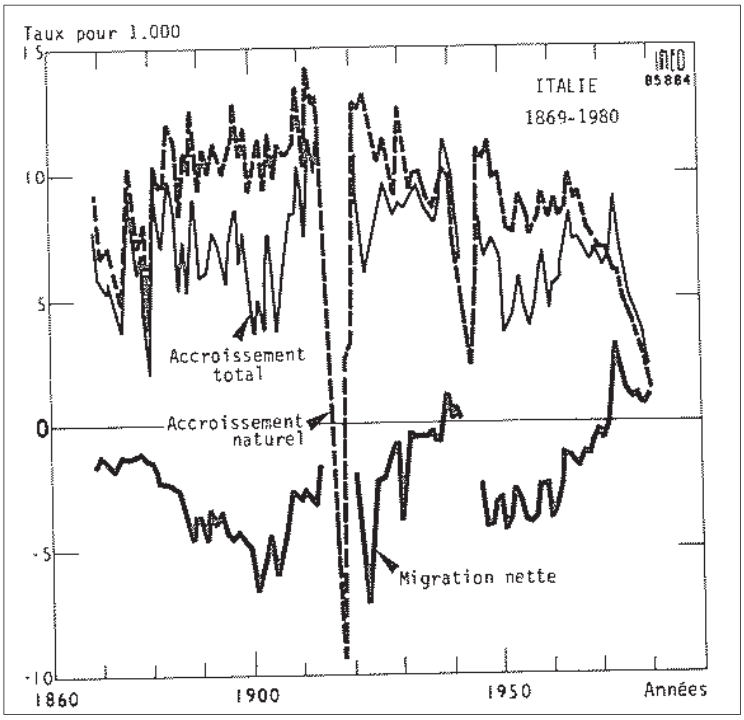
Nous observons que les périodes de plus fort accroissement naturel correspondent aux périodes de solde migratoire négatif, se manifestant par des sorties massives nettes de population.
Nous observons ainsi, un effet miroir entre la courbe de solde migratoire net, et celle de l'accroissement naturel : donc une corrélation forte et négative entre les deux indicateurs.
Cet effet miroir entre solde migratoire et accroissement naturel est concomitant du processus de transition démographique.
On retrouve le même phénomène en Italie (voir ci-après).

Figure 1.13 Mouvement naturel de la population, Italie 1861-1980



Source: Chesnais (1986) «The migratory transition», p. 237.

Figure 1.14 Migration nette et accroissement naturel, Italie 1869-1980

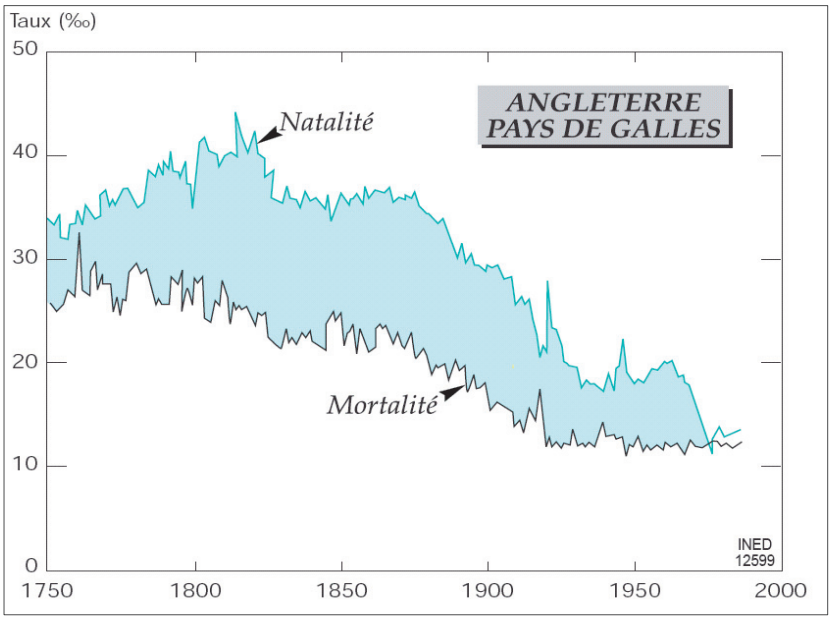


Source: Chesnais (1986) «The migratory transition», p. 179.

Les exemples historiques des vieux pays industriels comme l'Angleterre et l'Allemagne illustrent là encore, avec un décalage temporel par rapport à la Suède, le processus type de la transition démographique : une baisse continue du taux de mortalité, suivie une génération environ plus tard, d'une chute continue du taux de natalité. Après la transition, la mortalité et la stabilité se stabilisent à des niveaux significativement plus bas.

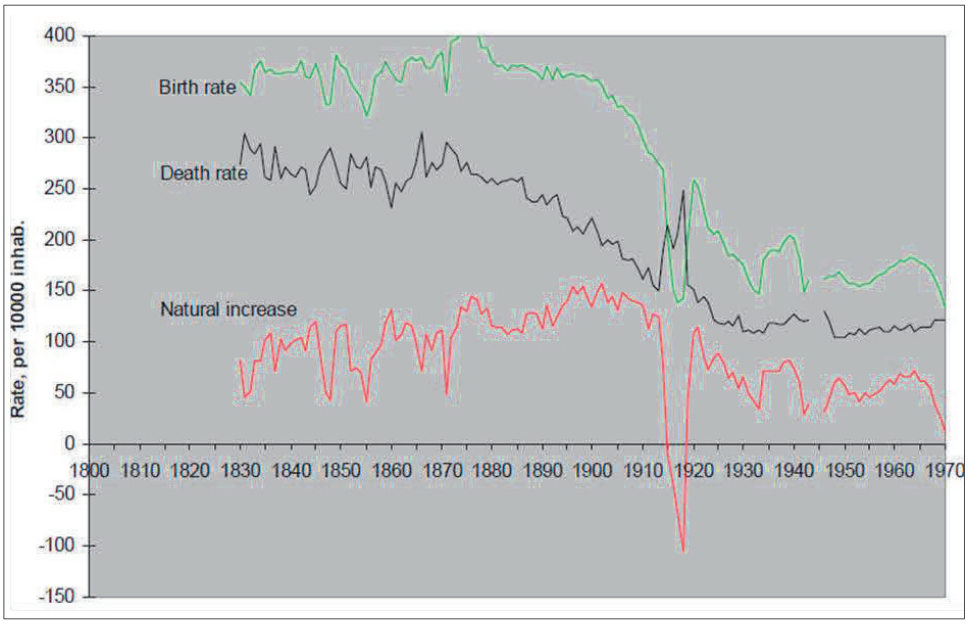
Exemples historiques : L'Angleterre & Pays de Galles

Figure 1.15 La transition démographique en Angleterre



Source : INED, Population et Sociétés n° 346, mai 1999.

Figure 1.16 La transition démographique en Allemagne

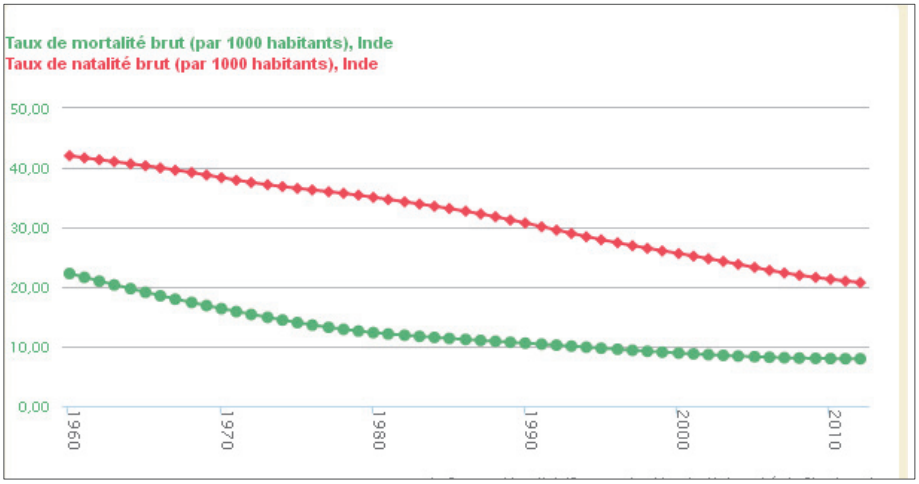


Source : Source : Mitchell, 2007.

Les jeunes nations connaissent également des tendances similaires, mais plus brèves et plus intenses. Les cas de l'Inde et de la Tunisie, choisis à dessein, illustrent parfaitement la généralité de la transition démographique.

Exemple historique : L'Inde

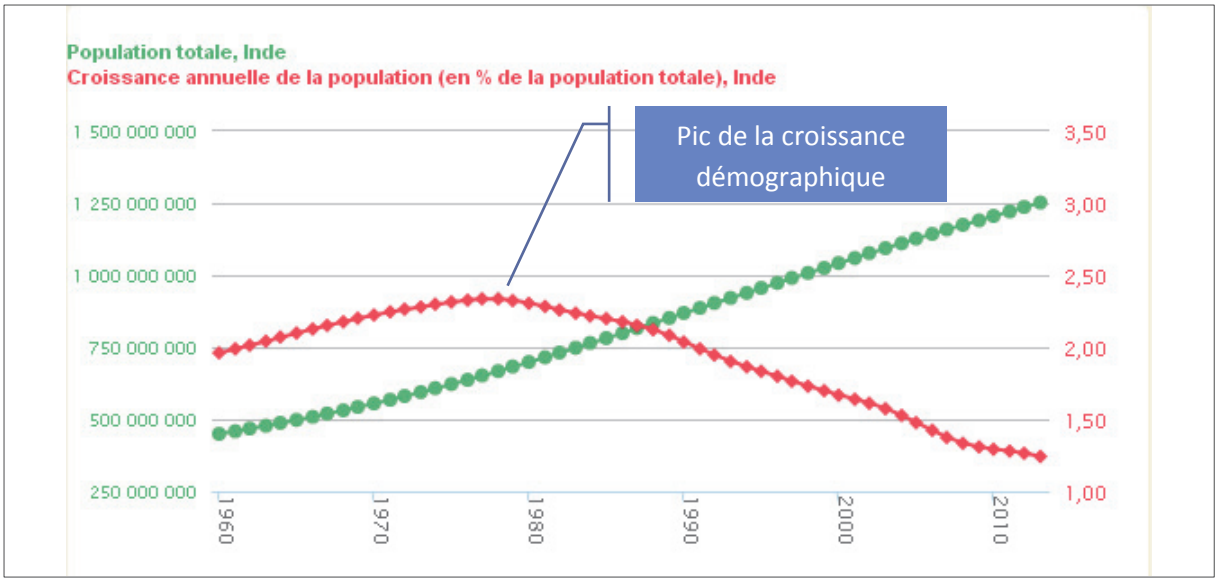
Figure 1.17 Mouvement naturel de la population, Inde 1960-2010



Source : La Banque Mondiale / Perspective Monde, Université de Sherbrooke.

L'Inde est bien dans une transition démographique (baisse continue et écarts importants entre les taux de mortalité et de natalité). La chute de la mortalité et de natalité ont commencés bien avant 1960. L'écart entre les 2 taux (mortalité et natalité) étant encore conséquent, la fin de la période transitoire est à prévoir plusieurs décennies après 2010.

Figure 1.18 Croissance annuelle de la population (en % de la population totale), Inde

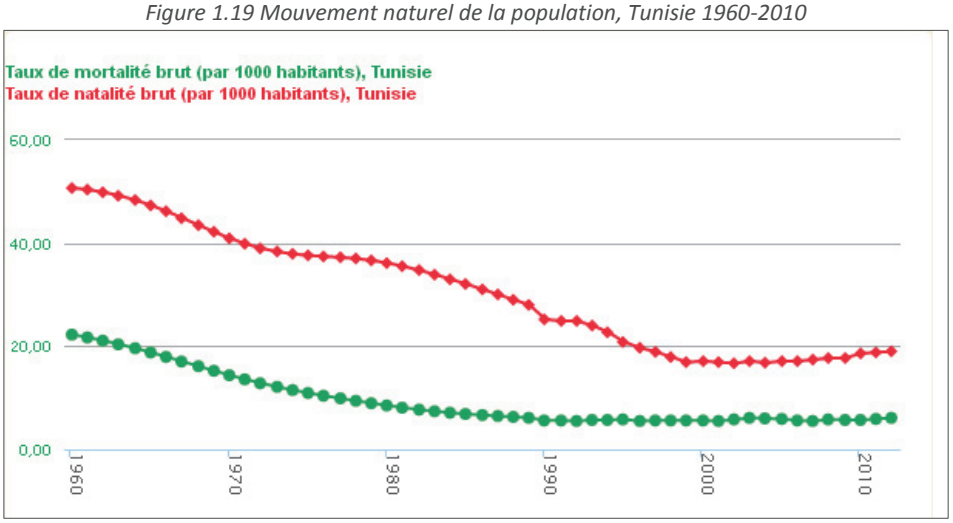


Source : La Banque Mondiale / Perspective Monde, Université de Sherbrooke.

Les données démographiques indiennes, indiquent que le pic de croissance démographique s'est situé en 1975 et que par conséquent, l'Inde est actuellement dans la phase 2 de sa transition démographique.

Exemple historique : La Tunisie

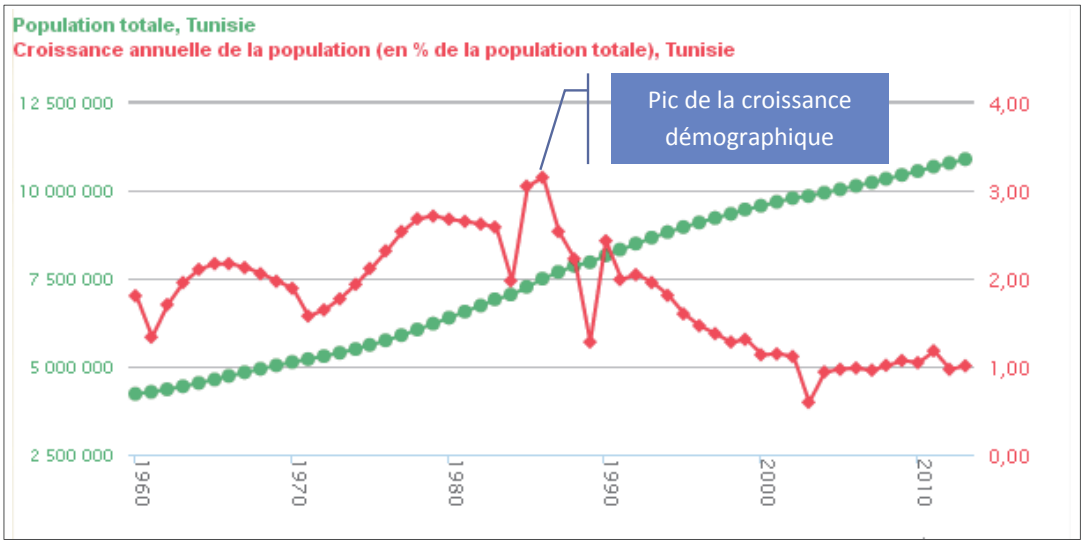
Figure 1.19 Mouvement naturel de la population, Tunisie 1960-2010



Source : La Banque Mondiale / Perspective Monde, Université de Sherbrooke.

Dans le cas tunisien, la transition démographique s'accélère après 1960. A partir de 1990, le taux de mortalité ne baisse plus et devient stable. Le taux de natalité connait le même processus à partir de 2000. Cependant, l'écart conséquent entre les deux taux indique que la hausse de la population va se poursuivre à un rythme stable mais relativement faible.

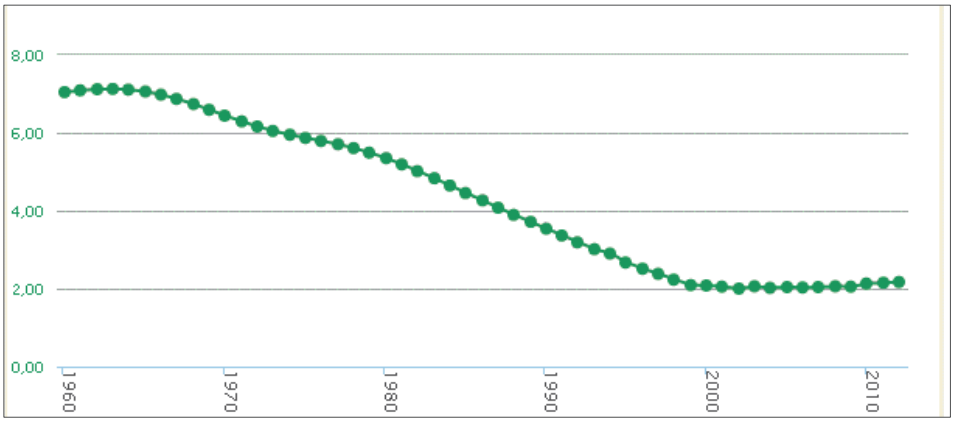
Figure 1.20 Croissance annuelle de la population (en % de la population totale), Tunisie



Source : La Banque Mondiale / Perspective Monde, Université de Sherbrooke.

Le pic de croissance démographique est atteint en 1985, ce qui marque la fin de la phase 1.
Nous observons également que le taux de croissance démographie est stabilisé depuis 2000, autour de 1% : ce qui suggère une fin proche de la phase 2.

Figure 1.21 Taux de fertilité (naissances par femme), Tunisie



Source : La Banque Mondiale / Perspective Monde, Université de Sherbrooke.

Depuis 2000, la fécondité tunisienne est à 2.1 enfants par femme (soit presque autant qu'en France) : Nous assistons à un achèvement probable de la transition, car ce niveau de fécondité correspond à un taux de simple renouvellement de population.
La Tunisie a donc achevé sa modernité démographique, qui d'ailleurs a été suivi par des changements institutionnels profonds à tendance démocratique. D'ores et déjà est entamé le progressif chemin d'accumulation du capital matériel et humain menant au développement économique.

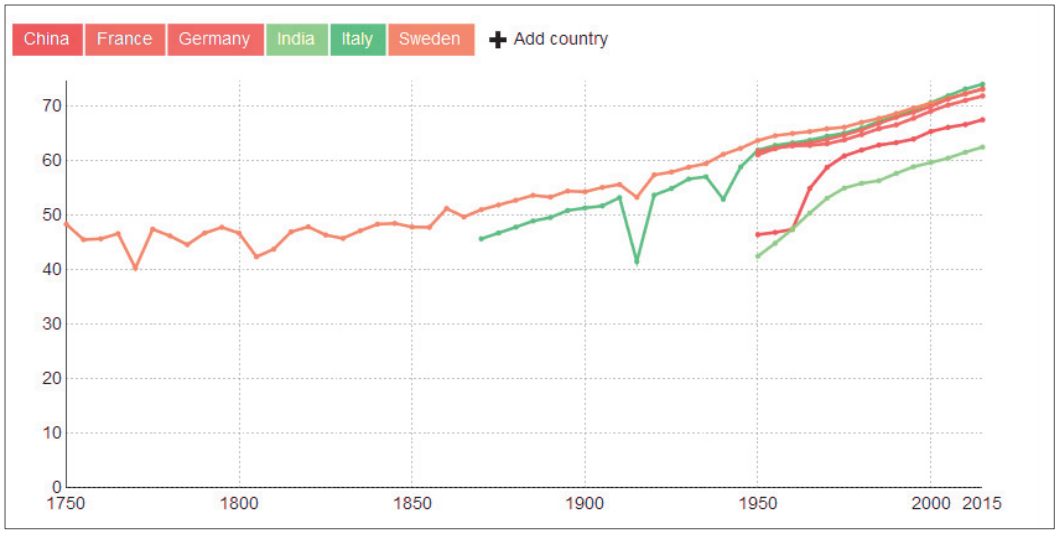
Longévité et vieillissement

Les deux autres conséquences majeures de la transition démographique sont la longévité et le vieillissement, corollaires logiques de la baisse continue de la mortalité.

Longévité

La hausse continue et générale de la longévité est illustrée par la Figure ci-après. Le nombre d'années restantes qu'un enfant de 10 ans devrait vivre est indiqué. À partir de 2015, les projections de la variante médiane de l'ONU sont présentées.

Figure 1.22 Espérance de vie à 10 ans, 1750 à 2015



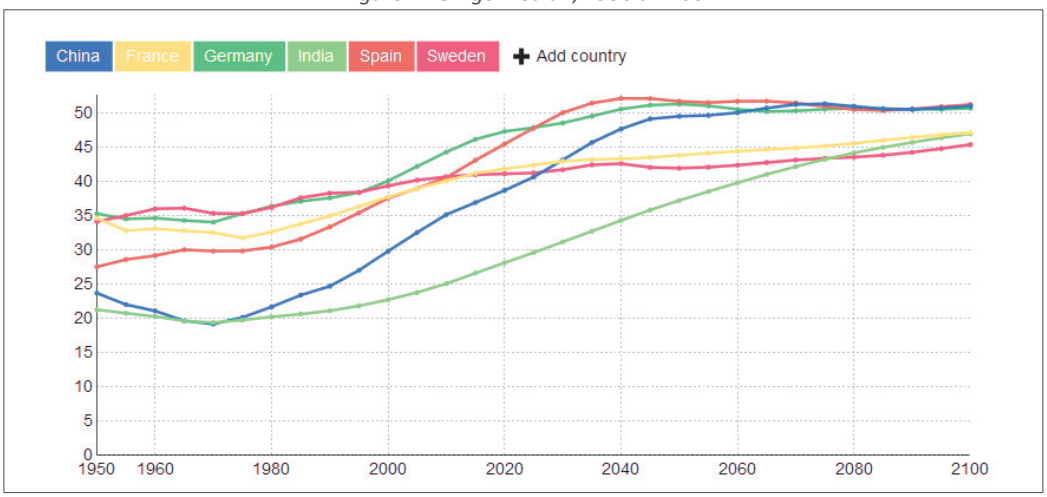
*Source: United Nations Population Division and Human Mortality Database/ OurWorldInData.org/life-expectancy/*CC BY-SA.*

Vieillissement

La hausse continue de l'âge médian illustre le fait connu du vieillissement des pays ayant terminé leur transition démographique depuis plusieurs générations. L'âge médian divise la population en deux parties de taille égale : il y a autant de personnes dont l'âge est supérieur à la médiane que d'âges inférieurs à la médiane.

Le graphique ci-dessous illustre le premier temps de rajeunissement, suivi d'une longue phase de vieillissement.

Figure 1.23 Âge médian, 1950 à 2100



*Source: United Nations Population Division (Median Age) 2015 Revision OurWorldInData.org/life-expectancy/*CC BY-SA.*

Note : Les années 1950 à 2015 montrent des estimations historiques. À partir de 2020, les projections de l'ONU (variante moyenne) sont présentées.

5. Anthropologie culturelle

5.1. Définitions

En anthropologie, on distingue l'***anthropologie physique***, qui étudie l'évolution et l'adaptation des humains en tant qu'êtres biologiques, de l'***anthropologie sociale et culturelle*** qui étudie la vie des hommes en société à travers leurs langues, leurs coutumes, leurs pratiques, leurs croyances, leurs mythes, leurs institutions.
Notre étude concerne l'anthropologie sociale et culturelle.

5.1.1. Signification spécifique du mot « culture » en anthropologie

Taylor (1871), anthropologue britannique, définit la ***culture*** comme l'« ensemble des habitudes acquises par les hommes en société ».
Pour Levi-Strauss : « Il ne faut pas confondre deux sens du mot culture. Dans son acceptation générale, **culture** désigne l'enrichissement éclairé du jugement et du goût. Dans le langage technique des anthropologues, (…) tout est objet d'étude : les productions qu'au premier sens du terme, on jugera les plus basses, comme les plus nobles ».

5.1.2. Anthropologie sociale et culturelle (Hofstede et Minkov, 2010)

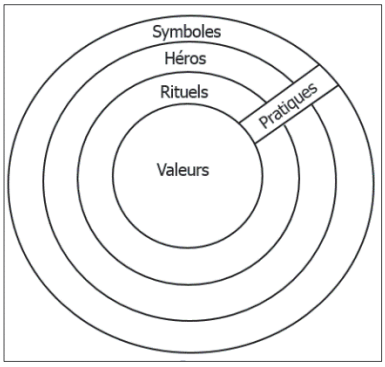
Nous ne pouvons pas définir l'anthropologie culturelle sans faire mention des études menées par Hofstede et Minkov (2010), qui enrichissent la définition de Taylor (1871).
Hofstede et Minkov (2010) étudient l'anthropologie sociale et culturelle comme un ensemble de « *valeurs* » sociales et leurs « *pratiques* » sociales associées.
Par ***valeur sociale***, Hofstede et Minkov (2010) entendent une préférence stable entre plusieurs solutions potentielles à une problématique sociale particulière. Ils expliquent une valeur comme un positionnement sur une dimension, celle d'une échelle de choix relative à une problématique générique (pouvant être rencontrée dans différents domaines sociaux).

Dans leur étude, Hofstede et Minkov (2010) identifient six dimensions dans le champ des valeurs : la ***distance hiérarchique*** longue ou courte (entre échelons), le ***confort face à l'incertitude*** (fort ou faible), les ***tendances « féminines »*** *ou* ***« masculines »*** (notons le caractère désuet de ces étiquettes), la ***relation au groupe de l'individu*** (communautarisme versus individualisme), la ***relation au temps*** (préférence pour le court terme, ou le long terme) et l'***indulgence*** (tolérance à l'hédonisme).
Les valeurs sociales vont servir de base à la mise en place des pratiques sociales.

Les ***pratiques sociales*** s'expriment dans divers domaines sociaux dont le domaine de la ***famille*** (sous les aspects nuptiaux et des relations entre membres), le domaine de l'***éducation scolaire***, ainsi que celui du ***travail*** (sur le plan du management de l'organisation et des motivations des travailleurs).
Ces pratiques se concrétisent également dans les domaines des ***institutions*** et dans **l'*économie*** (secteurs, acteurs, règlementations, fiscalité, par exemple).
Ces pratiques sociales se traduisent par des aspects et Figures « identitaires » (Hofstede et Minkov, 2010) : *rituels*, *héros* et *symboles*.
Les ***symboles*** sont des mots, actes, images, objets, porteurs d'une signification particulière pour les membres d'un groupe social.
Par ***rituels***, il faut entendre les pratiques sociales usuelles comme les salutations, les marques de respect, les cérémonies cultuelles ou culturelles.

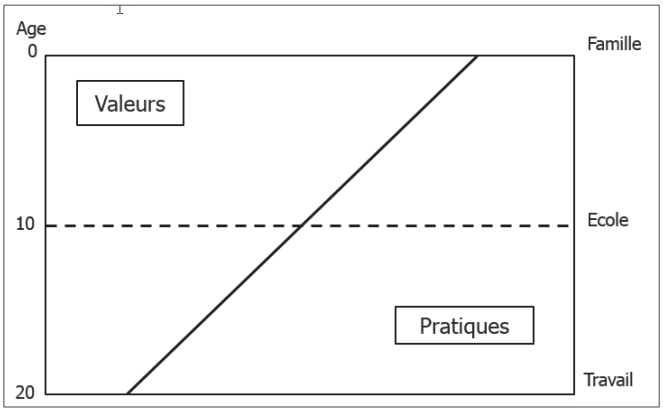
Les ***héros*** représentent des personnages historiques ou contemporains, réels ou fictifs porteurs des caractéristiques hautement valorisées par un groupe social.

Figure 1.24 Structure en pelure d'oignon (Hofstede et Minkov, 2010) : Valeurs et Pratiques « identitaires »



Hofstede et Minkov (2010)

Figure 1.25 Ages et lieux d'acquisitions des valeurs et pratiques sociales (Hofstede et Minkov, 2010)



Hofstede et Minkov (2010)

Les valeurs s'acquièrent aux âges les plus bas, tandis que les pratiques se confortent avec l'âge.

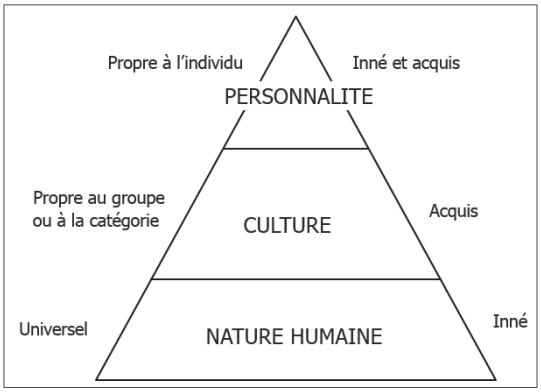
5.1.3. Différence entre anthropologie culturelle et études de personnalité

L'anthropologie culturelle a des objectifs différents des études de personnalité.
En effet, l'***anthropologie culturelle***, recherche des différences entre sociétés et est donc basée sur des données préalablement agrégées, alors que les ***études de personnalité*** recherchent des différences entre individus au sein d'une société donnée.

Ainsi, l'anthropologie culturelle va au-delà de l'étude d'une seule société avec une dimension comparatiste entre sociétés, alors que l'étude de personnalité se fait au sein d'une même société en y différenciant les individus.
Hofstede et Minkov (2010) présentent la situation en distinguant trois niveaux de programmation mentale : la *nature humaine*, la *culture* et la *personnalité*.

Le triangle ci-dessous illustre ces trois niveaux.

Figure 1.26 Les 3 niveaux de « programmation mentale humaine » (Hofstede et Minkov, 2010)



Hofstede et Minkov (2010)

En dehors des objectifs qui diffèrent entre l'anthropologie culturelle et les études de personnalité, il faut relever qu'elles ont des dimensions différentes.
Ces différences au niveau dimensionnel ressortent comme évidentes, en comparant les travaux de Hofstede et Minkov (2010) à ceux de McCrae et Costa (1990) qui font référence en psychologie de la personnalité.

Les six dimensions de l'anthropologie culturelle (Hofstede et Minkov, 2010), (la *distance hiérarchique* longue ou courte, le *confort face à l'incertitude* qui peut être fort ou faible, les *tendances* « féminines » ou « masculines », la *relation au groupe de l'individu*, la *relation au temps* et l'*indulgence*), se différencient clairement des catégories du modèle « **OCEAN »** (***O****penness*, ***C****onscientiousness*, ***E****xtraversion*, ***A****greeableness*, ***N****euroticism*) de McCrae et Costa (1990). Ces cinq traits de personnalité couramment dénommés « big five » sont : ***O****uverture* (curiosité et imagination), ***C****onscienciosité* (autodiscipline, respect des obligations, méthode), ***E****xtraversion* (stimulation et recherche de la compagnie des autres), ***A****gréabilité* (compatissant et coopératif) et ***N****euroticisme* ou névrosisme (instabilité émotionnelle).

5.2. Modèles de l'anthropologie culturelle

Des modèles ont été développés par de nombreux auteurs pour traiter de la question de l'anthropologie culturelle. Dans la suite, nous présenterons les modèles les plus représentatifs.

5.2.1. Modèle d'Edward T. Hall (1950-1955, au *Foreign Service Institute FSI, USA*)

Ce modèle décrit la perception et la gestion du temps en opposant les conceptions *monochrones* et *polychrones*.
La ***monochronie*** caractérise les sociétés où les individus font préférentiellement une seule tâche à la fois et ce de façon séquentielle, tandis que la ***polychronie*** caractérise les sociétés où les individus font plusieurs tâches à la fois et sont donc plus flexibles.
La *perception et la gestion de l'espace physique et de la distance entre individus* sont caractérisées par *l'**espace personnel restreint*** avec frontières marquées et **l'*espace personnel élargi***, avec frontières souples.
Le *mode de communication* est caractérisé par un axe opposant un ***contexte de communication riche*** qui se traduit par une forte communication informelle et implicite, à un ***contexte de communication pauvre*** qui se traduit par une faible communication informelle avec un niveau élevé de formalisation explicite.

5.2.2. Modèle Kluckhohn-Strodtbeck (1961, USA)

A travers ce modèle, Kluckhohn-Strodtbeck (1961) font ressortir cinq dimensions anthropologiques supposées caractériser toute société.
Ces cinq dimensions sont : la *relation homme/nature*, les *relations humaines*, la *gestion du temps*, les *croyances sur la nature humaine* ainsi que la *relation à l'activité*.
La ***relation homme/nature*** est caractérisée par la domination, l'harmonie ou la subjugation.
Les ***relations humaines*** font appel à l'interdépendance (décision collégiale) ou à l'individualisme (décision individuelle).
La ***gestion du temps*** passé, présent ou futur met en relief des degrés d'orientations dans la prise en compte du temps.
La ***croyance que la nature humaine*** (avec option modifiable/immuable) est soit « mauvaise », « neutre » (ou mixte), ou « bonne ».
Enfin la ***relation à l'activité*** décrit la conception des activités humaines : ces activités sont réalisées pour être (exister), se réaliser (être et devenir), ou faire (obtenir des résultats).

5.2.3. Modèle Geert Hofstede (1970, université de Limburg, Maastricht, Pays-Bas).

Hofstede (1970) puis Hofstede et Minkov (2010) par leur modèle, montrent empiriquement à l'aide d'axes factoriels, l'existence de six dimensions anthropologiques déjà évoquées plus haut.
La **d*istance hiérarchique longue/courte,*** indique le degré d'acceptation culturelle des inégalités de statuts et de pouvoir entre les individus. Cette dimension révèle le degré de respect dont font preuve les gens vis-à-vis de leur hiérarchie et de l'autorité. Les scores sont plus élevés au sein des cultures où le pouvoir central est « le pouvoir vénéré, mais où se développent des contre-pouvoirs, officiels ou occultes, s'opposant à une centralisation trop pesante ».

Le ***confort face à l'incertitude (fort ou faible)***, désigne le degré de tolérance d'une culture confrontée à l'inquiétude provoquée par les événements futurs. La relation à l'incertitude est considérée comme différente selon les sociétés, certaines utilisant des prévisions et des dispositifs pour gérer les événements (comme le principe de précaution), d'autres sociétés se montrant moins averses à l'incertitude.

Pour les ***tendances « féminines » ou « masculines »*** (étiquettes à caractère désuet), il s'agit de savoir si une société est, d'une part, caractérisée (de façon un peu datée), pas des aspects plutôt émotionnels (qualifié par les auteurs de « féminin ») ou factuels (qualifié par les auteurs de « masculin ») et, d'autre part, organisée autour d'une séparation marquée (ou non) du rôle des deux sexes dans les tâches de la vie quotidienne.

La ***relation au groupe de l'individu (communautarisme vs individualisme)***, exprime le degré de liberté et d'autonomie de l'individu par rapport à un groupe et aux normes sociales.

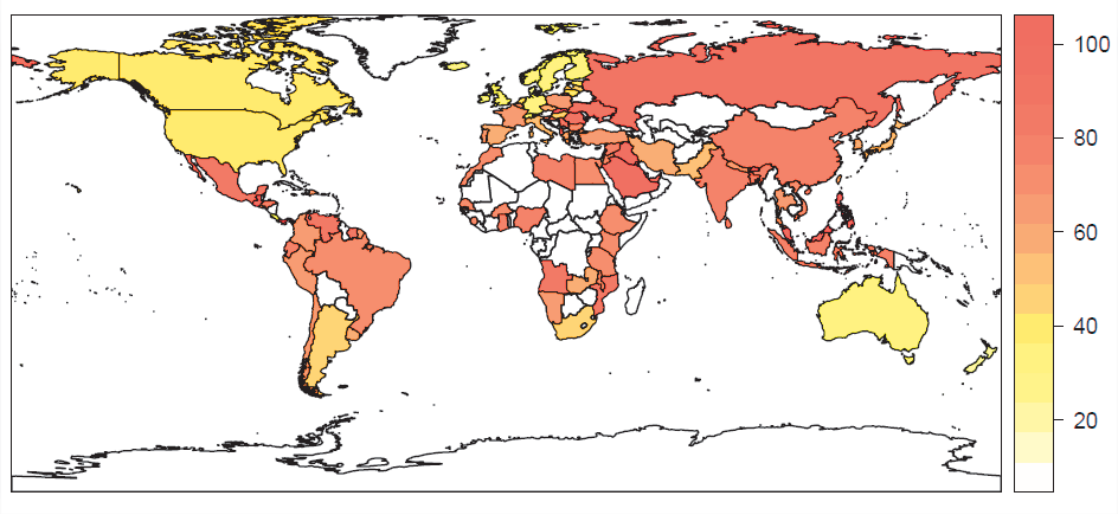
La ***relation au temps*** **(court terme vs long terme)**, caractérise la préférence sociale pour le long terme ou le court terme, notamment dans la prise de décision.

L'***indulgence*** est définie comme le degré d'autocontrôle du désir et des impulsions par les individus. Un contrôle relativement faible est appelé « indulgence » et un contrôle relativement fort est appelé « contrainte ». Les cultures y sont décrites comme « indulgentes » ou « restreintes ».

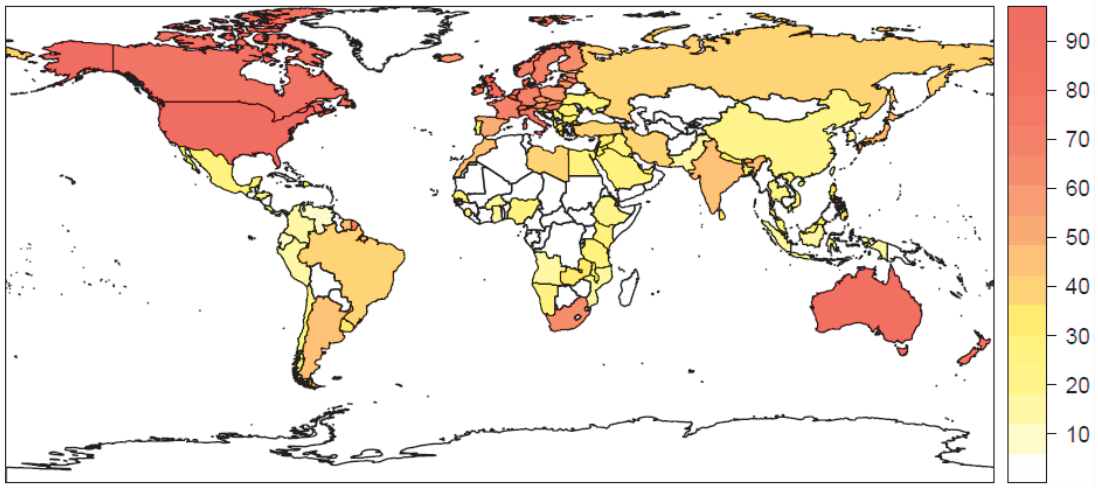
Nos cartographies ci-dessous illustrent que le niveau de chacune des six dimensions anthropologiques au sens de Hofstede et Minkov (2010), varie d'un pays à un autre :

Figure 1.27 Cartographies mondiales des dimensions anthropologiques au sens de Hofstede-Minkov (2010)

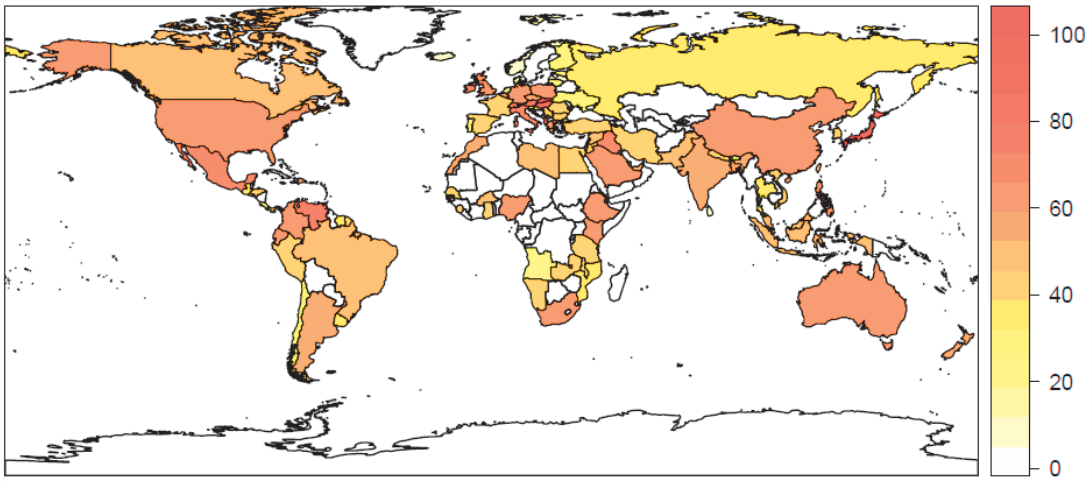
Distance (entre échelons) hiérarchique



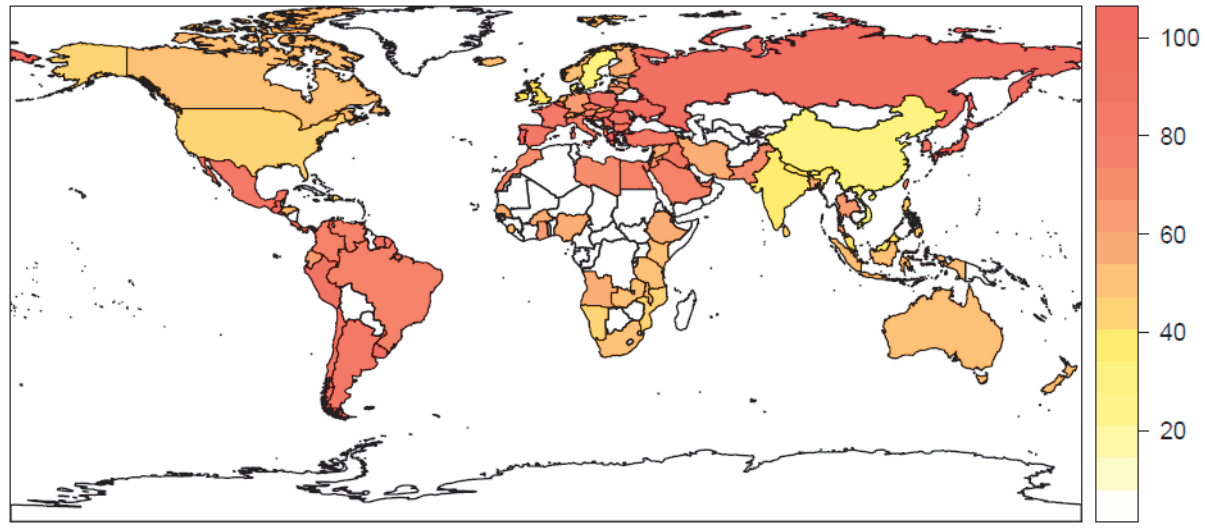
Individualisme (vs Communautarisme)



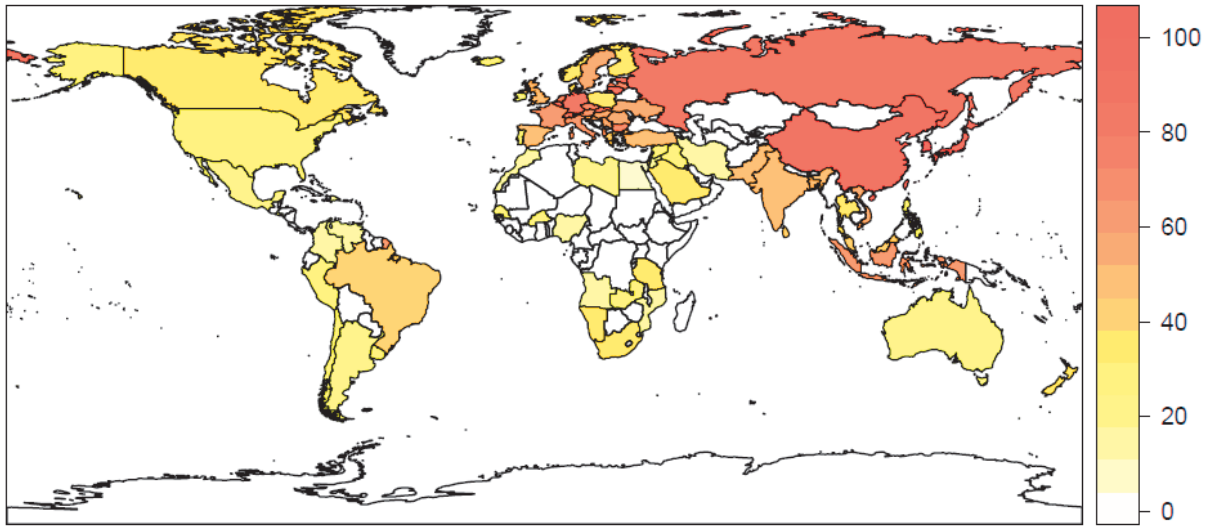
Tendance « masculine »



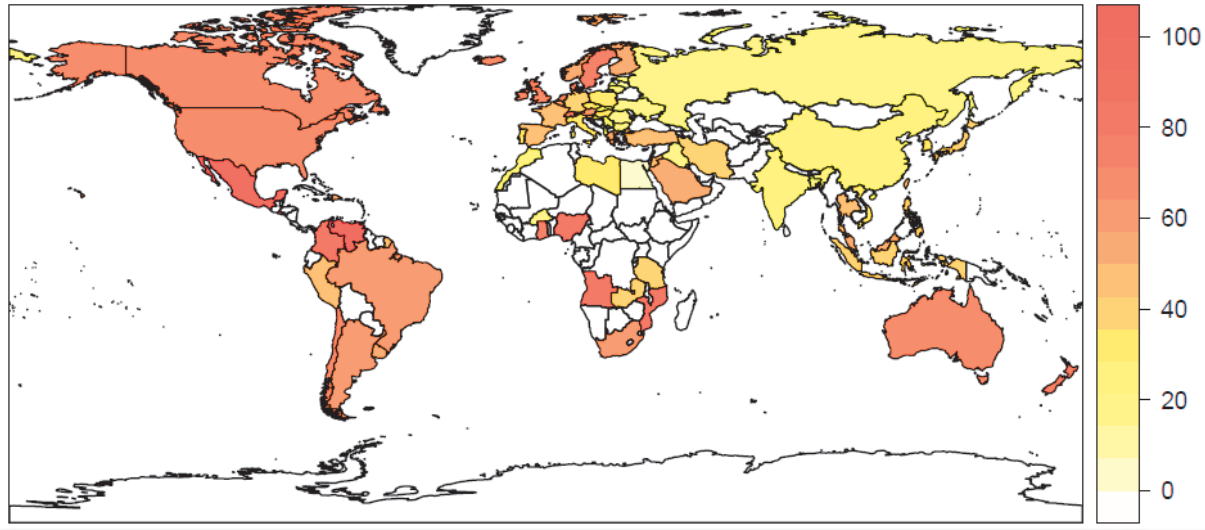
Aversion à l'incertitude



Orientation de long terme



Indulgence



5.2.4. Modèle Fons Trompenaars et Hampden-Turner (1993)

Ce modèle, développé au début des années 1990 par Trompenaars et Hampden-Turner (1993, 2008), permet d'analyser les différences culturelles entre sociétés.
A cet effet, six dimensions sont retenues : la *gestion des relations et des règles* qui peuvent être qualifiées d'universelles ou de particulières, la *relation groupe vs individu*, la *gestion et l'expression des sentiments*, l'importance du *statut social par rapport aux résultats concrets,* la *gestion du temps* (Passé/présent/futur, monochrone/multichrone), et enfin, le *rapport à la nature* exprimé par le volontarisme et le déterminisme.

5.2.5. Traits d'histoire de vie (stratégies évolutives d'adaptation, *life history theory*)

Pour caractériser les traits d'histoire de vie, le modèle *Life history theory* (traits d'histoire de vie) retient trois indicateurs : les indicateurs de *survie* (l'espérance de vie, les mesures de mortalité), les indicateurs de *reproduction* (fécondité, âge au 1er enfant) et les indicateurs *migratoires* (sédentarité, nomadisme, mobilité résidentielle), (Avdeev A., 2014)
Outre ces indicateurs, le modèle prend en compte les *contraintes* relatives à l'*allocation des ressources* (*trade-off entre survie et reproduction*) et celles liées à *l'instabilité et aux imprévisibilités* environnementales.
Le modèle *Life history theory* est basé sur des stratégies évolutives d'adaptation (où l'on met en œuvre des traits d'histoire de vie), telles que les stratégies « r/K » (r pour le taux de reproduction, et K pour la durée de survie) ou les stratégies de sélections par le groupe (Darwin, George Price).

5.2.6. Systèmes familiaux préindustriels : le modèle « Le Play-Todd »

Une classification mondiale des systèmes familiaux a été produite par Todd (1983) à partir des travaux précurseurs de Le Play (1871). La classification est effectuée à partir d'études anthropologiques des sociétés dans leur époque préindustrielle.

La valeur « autorité » (ou verticalité) se fonde sur une vision à long terme, qui suppose un projet « multigénérationnel » de collaboration forte entre générations. Cette forte collaboration parents/enfants est caractérisée par une transmission indivise du patrimoine économique productif et par une longue cohabitation parents/enfants.
Contrairement à l'autorité, la « liberté », prône une grande autonomie des enfants vis-à-vis des parents, qui suppose une faible collaboration entre générations. Le pôle liberté est caractérisé par une vision plutôt courtermiste de chaque génération independamment des générations en amont et en aval.
L'« égalité » impose une égalité de parts patrimoniales et de droits entre frères.
L'« inégalité » promeut le principe de différence de parts patrimoniales et de droits entre frères.

Le « *statut de la femme »* se définit par les droits relatifs des filles (par exemple la part d'héritage des filles par rapport aux garçons) et dans les relations conjugales à travers le rôle joué par la femme dans la filiation des enfants.
Le statut de la femme prend chez Todd (1983) trois modalités possibles : la « bilatéralité », la « patrilinéarité » ou la « matrilinéarité ».
La bilatéralité indique que les enfants appartiennent à la fois à la lignée de leur mère et de leur père.
La patrilinéarité correspond au cas où les enfants appartiennent plutôt à la lignée de leur père et la matrilinéarité se réfère aux situations où les enfants sont rattachés à la lignée de leur mère.
Le statut de la femme peut être mesuré par le *sex-ratio*, le taux d'alphabétisation des filles par rapport à celui des garçons.

La « *régulation de l'inceste* » s'illustre par l'existence d'une règle sociale qui interdit les unions matrimoniales entre des individus du fait d'un lien de parenté particulier.
Cette régulation se caractérise par les modalités ci-après : exogamie forte, faible régulation, mariage préférentiel parallèle patrilatéral ou mariage asymétrique.
L'exogamie forte est une coutume qui interdit formellement le mariage entre les individus d'une même famille en favorisant les mariages entre membres d'un autre groupe social.
Le mariage préférentiel parallèle patrilatéral stipule l'existence des règles incitatives favorisant l'union entre les enfants de deux frères.
Le mariage asymétrique interdit le mariage entre les enfants de deux frères, mais exprime une préférence pour le mariage des enfants d'un frère et d'une sœur. Les deux faces de ce mécanisme sont, une exogamie patrilatérale partielle et une endogamie matrilatérale.
Le mariage préférentiel parallèle patrilatéral et le mariage asymétrique sont de type endogamique.

Pour aboutir à la typologie culturelle des familles, Todd s'est prioritairement servi de deux dimensions à savoir : les *rapports entre parents et enfants* (Liberté/Autorité), et les *rapports dans la fratrie* (Egalité/Neutralité/Inégalité).
A ces deux principales dimensions se rajoutent trois autres dimensions complémentaires : le *statut de la femme* (Patrilinéarité/Bilatéralité/Matrilinéarité), la *régulation de l'inceste* (Endogamie/Exogamie) et le *nombre de conjoints* (polygamie sous la forme polygynie ou polyandrie, monogamie).

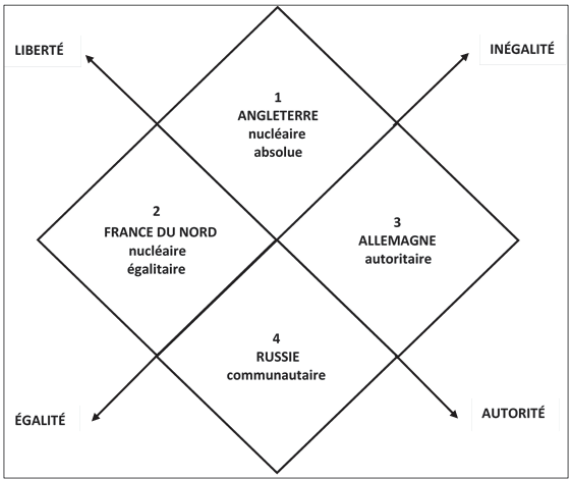
En utilisant les deux principales dimensions anthropologiques (liberté/autorité, égalité/inégalité), Todd (1983), obtient quatre types familiaux : le type ***nucléaire absolu*** (caractérisé par le principe de liberté et d'inégalité), le type ***nucléaire égalitaire*** (caractérisé par le principe de liberté et d'égalité), le type ***souche*** (caractérisé par le principe d'autorité et d'inégalité) et enfin le type ***communautaire*** (caractérisé par le principe d'autorité et d'égalité).

Tableau 1.1 Caractéristiques des types familiaux

Caractéristiques des types familiaux			
familial	***Principe d'égalité***	***Principe d'autorité***	***Formes typiques du groupe domestiq***
léaire solu	Non	Non	
léaire litaire	Oui	Non	
uche	Non	Oui	
unautaire	Oui	Oui	
omme	Filiation	Mariage	
emme	Fraternité		

Source: Todd (1983), p 28.

Figure 1.28 Parangon européen des types familiaux



Todd (1983) « La diversité du Monde », p.36.

Selon les études de Todd (1983), le type familial nucléaire absolu (respectivement « autoritaire », « communautaire », « nucléaire égalitaire ») a pour parangon européen l'Angleterre (respectivement l'Allemagne, la Russie, la France du Nord).

Todd(1984) en utilisant toutes les dimenssions anthropologiques,obtient sept types familliaux, hors systèmes familiaux africains :

Tableau 1.2 Types familiaux et phénomènes sociaux

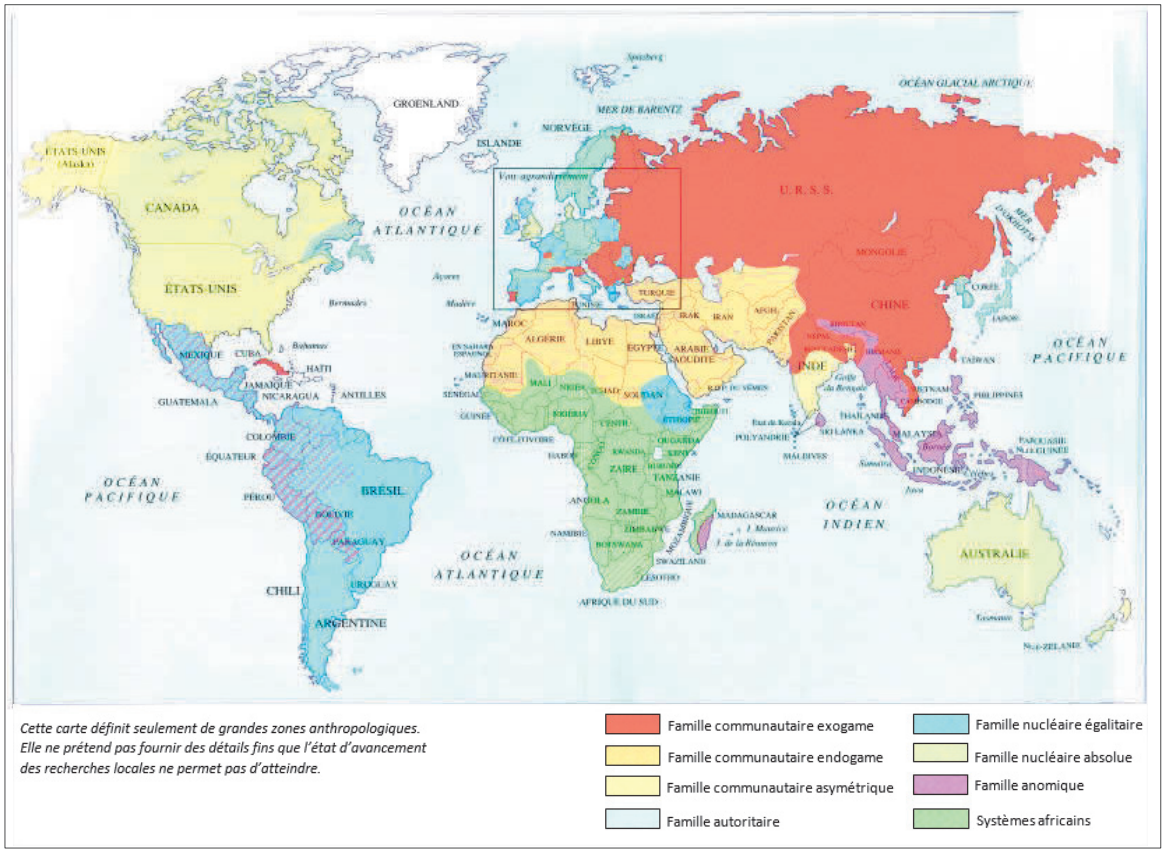
TYPES FAMILIAUX DEFINIS DANS « LA TROISIEME PLANETE »	EXEMPLE REGIONAL		Eléments de structure familiale			Implications sociales	
		STATUT DES FEMMES	RAPPORTS PARENTS / ENFANTS (AUTORITE)	RAPPORTS ENTRE FRERES (EGALITE)	REGULATION DE L'INCESTE	POTENTIEL CULTUREL	FORME IDEOLOGIQUE
Famille autoritaire	*Allemagne*	Bilatéralité	Verticalité	Inégalité	Exogamie forte	Très fort	Autoritarisme ethnocentrique
	Japon	Bilatéralité	Verticalité	Inégalité	Faible régulation	Très fort	
Famille nucléaire absolue	*Angleterre*	Bilatéralité (biais matrilinéaire)	Non-verticalité	Indifférence	Exogamie forte	Moyen (+)	Individualisme libéral
Famille nucléaire égalitaire	*France du Nord*	Bilatéralité (biais patrilinéaire)	Non-verticalité	Egalité	Exogamie forte	Moyen	Individualisme égalitaire
Famille communautaire exogame	*Russie*	Patrilinéarité (faible)	Verticalité	Egalité	Exogamie forte	Moyen	Communisme
	Chine	Patrilinéarité (moyenne)	Verticalité	Egalité	Exogamie forte	Moyen (-)	
	Inde du Nord	Patrilinéarité (forte)	Verticalité	Egalité	Exogamie forte	Faible	Incertitude
Famille anomique	*Thailande*	Bilatéralité (biais matrilinéaire)	Non-verticalité	Indifférence	Faible régulation	Moyen (+)	Flou Conceptuel (anomie idéologique)
	Amérique centrale	Bilatéralité (biais patrilinéaire)	Non-verticalité	Indifférence	Faible régulation	Moyen (-)	
Famille communautaire endogame	*Monde arabe*	Patrilinéarité (forte)	Non-verticalité	Egalité	Mariage préférentiel parallèle patrilatéral	Faible	Islam
Famille communautaire asymétrique	*Tamil Nadu (Inde du Sud)*	Patrilinéarité (faible)	Verticalité	Egalité	Mariage asymétrique	Moyen	Système des castes + communisme
	Kerala (Inde du Sud)	Matrilinéarité	Verticalité	Egalité	Mariage asymétrique	Fort	

Source : Todd (1984) « La diversité du Monde », p.388-389.

Chez Todd, le « *potentiel culturel* » est associé positivement au statut matrilinéaire ou bilatéral de la femme, ainsi qu'à la verticalité des relations parents/enfants. Il est négativement associé à une patrilinéarité forte (statut faible de la femme), ainsi qu'à la liberté dans les relations parents/enfants.
La « forme idéologique » indique l'orientation idéologique et politique de chaque société selon son type anthropologique (Todd, 1984).

Todd (1983), après étude et affectation des types familiaux à différentes régions et pays, publie la cartographie mondiale suivante :

Figure 1.29 . Cartographie mondiale des types familiaux

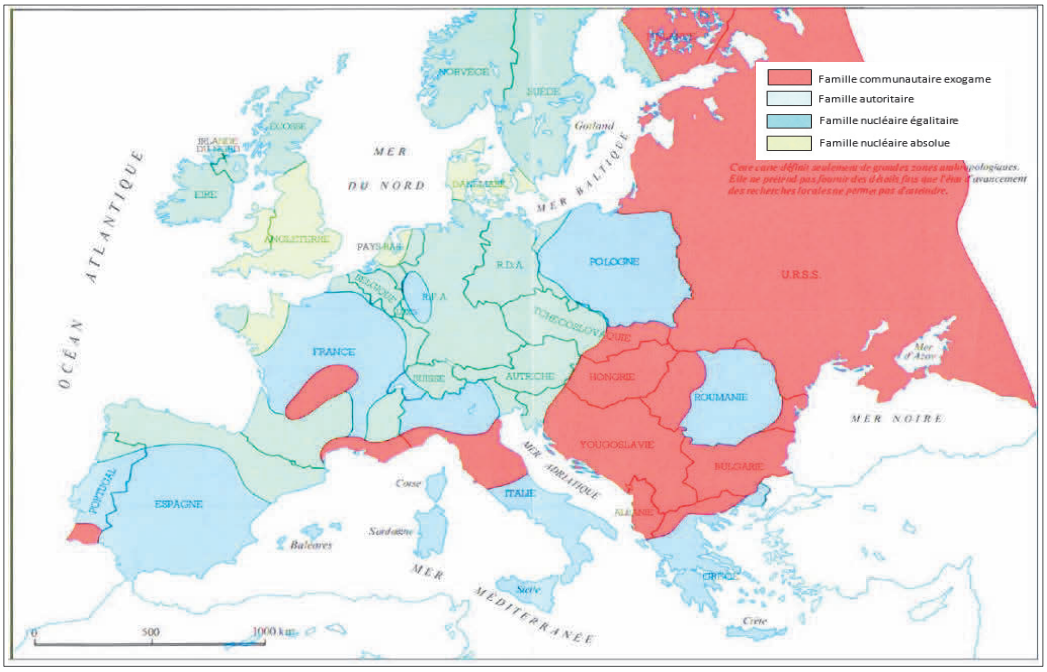


Cette carte définit seulement de grandes zones anthropologiques. Elle ne prétend pas fournir des détails fins que l'état d'avancement des recherches locales ne permet pas d'atteindre.

Source : Todd (1983).

Un zoom sur l'Europe permet à Todd (1984) d'obtenir la cartographie suivante :

Figure 1.30 Carte européenne des types familiaux



Source : Todd (1983) « la Troisième Planète. Structures familiales et systèmes idéologiques », p.260.

La cartographie mondiale des types familiaux permet de produire les Tableaux statistiques des types familiaux ci-après :

Tableau 1.3 Quelques statistiques des types familiaux

REPARTITION DES MASSES ANTHROPOLOGIQUES DANS LE MONDE (pourcentage de la population mondiale)	
Famille nucléaire absolue	8 %
Famille nucléaire égalitaire	11 %
Famille autoritaire	8 %
Famille communautaire exogame	41 %
Famille communautaire endogame	10 %
Famille communautaire asymétrique	7 %
Famille atomique	8 %
Systèmes africains	6 %

EN EUROPE, QUATRE TYPES FAMILIAUX SEULEMENT, TOUS EXOGAMES	
Famille nucléaire absolue	12 %
Famille nucléaire égalitaire	25 %
Famille autoritaire	26 %
Famille communautaire exogame	37 %

Todd (1980) « La diversité du Monde », p.50.

La classification de Todd étant basée sur l'époque préindustrielle de chaque pays, elle est de ce fait statique.
Pour adopter une approche dynamique, Todd (1996, 1998, 1999) utilise régulièrement des mesures corollaires des dimensions anthropologiques de Todd-Le Play, à défaut de mesures directes.
L'avantage de ces mesures corollaires est de permettre un suivi chronologique de la dynamique anthropologique, grâce aux statistiques internationales disponibles.
Nous synthétisons ci-après les diverses mesures corollaires des dimensions anthropologiques, utilisées par Todd (1996, 1998, 1999).

Verticalité (ou Autorité) versus Liberté
Dans une population, la verticalité est d'autant plus élevée que :

Tableau 1.4 Verticalité (ou Autorité) versus Liberté

Variable	**Niveau associé à la verticalité**
Fécondité	Faible
Education	Elevée
Industrie	Elevée
Solde commercial	Elevé
Suicides	Elevés
Homicides (et agressions)	Faibles
Niveau d'étatisation de l'économie (si verticalité et égalité)	Elevé
Exportations/PIB	Elevé
Exportions/Habitant	Elevé
Importations/Habitant	Faible
Age au 1er enfant	Elevé
Taux de célibat (surtout s'il y a inégalité)	Elevé
Nombre de frères et/ou sœurs (ou nombre de générations) cohabitant dans une même habitation	Elevé
Taux de femmes (ou d'adultes) sans enfant (si verticalité et inégalité)	Elevé
Mortalité infantile	Faible
Indice des libertés	Faible
Espérance de vie	Elevée
Consommation dans le PIB privé	Faible
Taux d'épargne	Elevé
Nombre d'entreprises leaders mondiaux (*hidden champions*)	Elevé
Taux de métier commun Parent/Enfants, i.e. stabilité emploi/secteur (services vs Agriculture + Industrie), sur 2 à plusieurs générations	Elevé

Endogamie versus Exogamie

La dimension endogamique est d'autant plus élevée que :

Tableau 1.5 Endogamie versus Exogamie

Variable	**Niveau associé à l'endogamie**
Taux d'ouverture économique = (Imports +Exports) / PIB	Faible
Etrangers (ou nés à l'étranger) résidents sur le territoire	Faible
Immigration	Faible
Taux d'union avec Etrangers	Faible
Nombre d'enfants issus des parents mixtes (nationaux et étrangers)	Faible
Pourcentage d'étrangers résidents	Faible
Taux de résidents nés à l'étranger	Faible
Taux de naturalisation (la différence entre les 2 points précédents est corrélée avec le taux de naturalisation)	Faible

Statut des femmes

Le statut des femmes est d'autant plus élevé que :

Tableau 1.6 Statut des femmes

Variable	**Niveau associé au statut élevé de la femme**
Education femmes/hommes	Elevé
Espérance de vie femmes/hommes	Elevé
Sex-ratio hommes/femmes	Faible
Emplois femmes/hommes	Elevé
Revenus femmes/hommes (à emploi ou éducation identiques)	Elevé
Taux de divorce	Elevé
Ecart d'âge mari/épouse (ou entre conjoints)	Faible
Indice de Développement Humain H/F	Elevé

Egalité versus Inégalité

Dans une population, la dimension anthropologique d'égalité est d'autant plus élevée que :

Tableau 1.7 Egalité versus Inégalité

Variable	**Niveau associé à l'égalité anthropologique**
Indice de Gini après redistribution	Faible
Ratio de redistribution : Gini après redistribution / Gini avant redistribution	Elevé
Transferts et subsides	Elevé
Poids de l'état	Elevé
Prélèvements obligatoires	Elevé
Espérance de vie (à PIB/habitant identique)	Elevé

5.2.7. Modèle Inglehart-Welzel

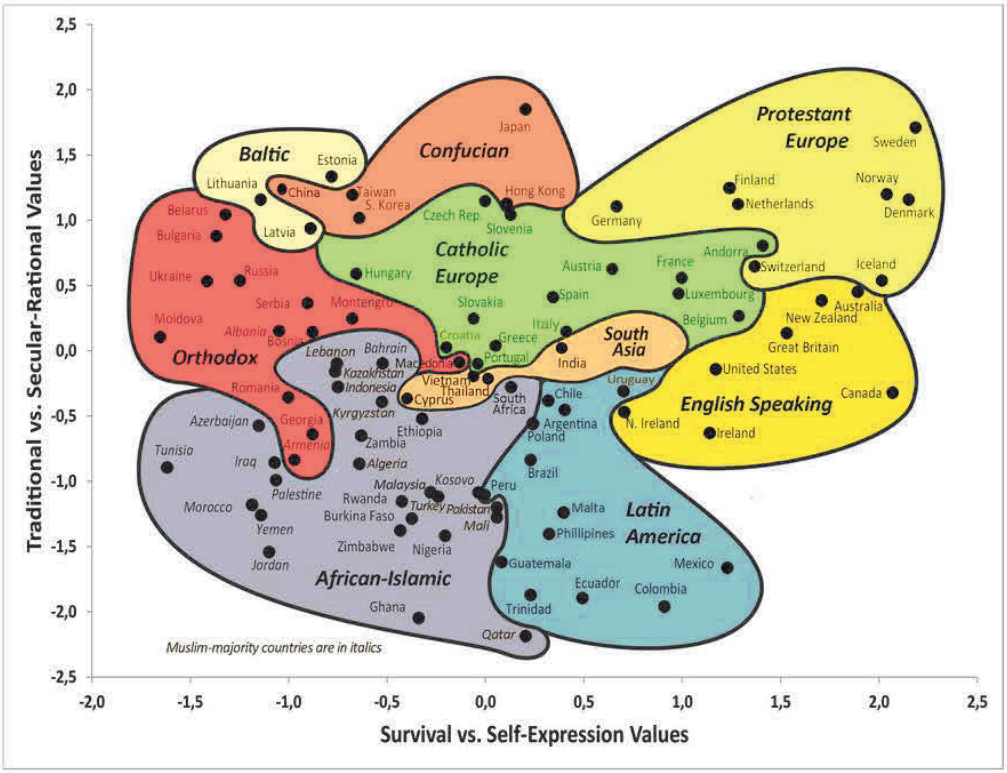
Le modèle (Inglehart et Welzel, 2005) issue de l'analyse statistique des données de l'enquête mondiale pluriannuelle *World Value Survey,* (Inglehart et al, 2014) s'articule autour de deux dimensions : la première dimension oppose les valeurs « *traditionnelles* » aux valeurs « *laïques-rationnelles* », tandis que la seconde dimension oppose les valeurs de « *survie* » aux valeurs « *d'expression personnelle* ».

Dimension 1 : Valeurs « traditionnelles » vs valeurs « laïques-rationnelles »
Les valeurs « traditionnelles » sont propres aux sociétés dans lesquelles la religion joue un rôle important, et au sein desquelles sont valorisées l'importance des liens parent-enfant et la déférence envers l'autorité. Elles rejettent le divorce, l'avortement, l'euthanasie et le suicide. Ces sociétés ont une grande fierté nationale et une vision nationaliste.
A contrario, les ***valeurs « laïques-rationnelles »*** caractérisent les sociétés ayant des préférences opposées aux valeurs « traditionnelles ».

Dimension 2 : Valeurs de « survie » vs valeurs « d'expression personnelle »
Par ***valeurs de « survie »***, il faut entendre l'ensemble des valeurs qui mettent prioritairement l'accent sur la sécurité économique et physique ; tandis que les ***valeurs « d'expression personnelle »*** mettent l'accent sur le bien-être subjectif, l'expression de soi et la qualité de vie.

La carte culturelle mondiale ci-dessous, produite par les politologues Ronald Inglehart et Christian Welzel (2019), classe les pays selon ces deux axes : les valeurs traditionnelles vs valeurs laïques-rationnelles et les valeurs de survie vs valeurs d'expression personnelle.

Figure 1.31 Carte culturelle mondiale



Source: Cultural map – World Value Survey wave 6 (2010-2014), Inglehart et Welzel (2019)

Dans la représentation culturelle mondiale ci-dessus Inglehart et Welzel (2019) positionnent la plupart des pays du monde par rapport aux deux dimensions culturelles qui sont les valeurs « traditionnelles vs valeurs laïques-rationnelles » et les valeurs de « survie vs valeurs d'expression personnelle ». Les auteurs

identifient neuf catégories culturelles dénommées : baltique, confucéenne, protestante européenne, orthodoxe, catholique européenne, anglo-saxonne, sud asiatique, latine américaine, africaine-islamique. Ces catégories culturelles sont positionnées sur la figure précédente par rapport aux deux dimensions culturelles.

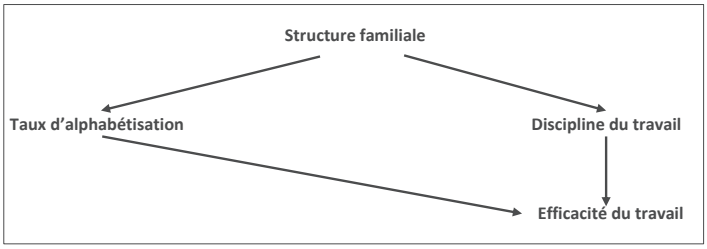
5.3. Relations entre anthropologie et économie

L'anthropologie est définie comme l'ensemble des sciences qui étudie la vie des hommes en société à travers leur langue, leurs coutumes, leurs pratiques, leurs croyances, leurs mythes, leurs institutions. Autant de domaines sociaux interagissant avec l'économie d'une société.

5.3.1. Anthropologie et efficacité du travail

Au début des années 80, Todd (1980, p.383), postulait une relation entre la structure familiale (systèmes familiaux) et l'efficacité du travail (productivité), en passant par la médiation de l'alphabétisation et de la « discipline du travail » ; la « discipline du travail » étant induite par le caractère autoritaire (ou vertical) du système familial. Ce que l'auteur illustre par la figure ci-après :

Figure 1.32 Structure familiale et efficacité du travail

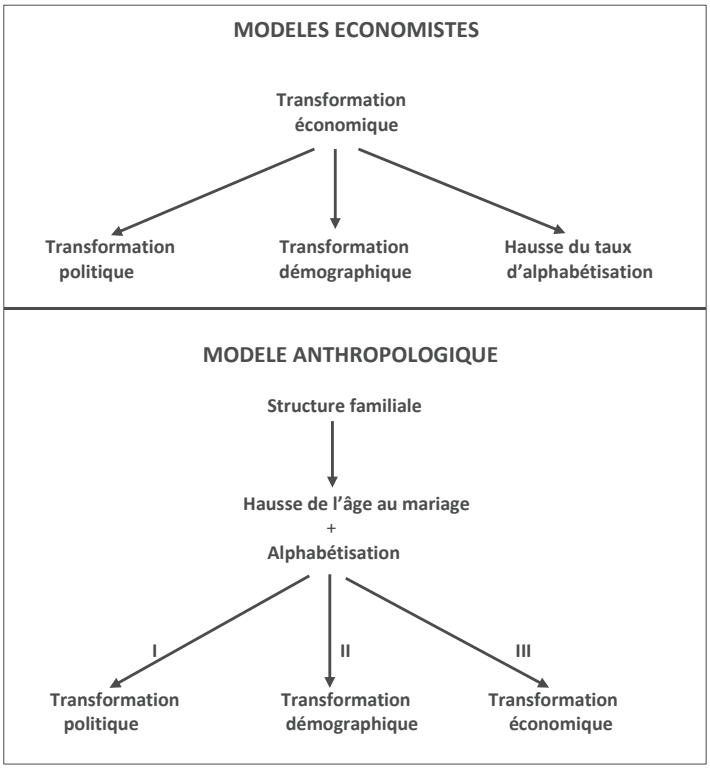


Source : Todd (1980), p.383.

5.3.2. Dynamiques historiques : modèle « anthropologique » contre modèle « économique » chez Todd (1984)

Todd (1984) définit deux modèles de transformation sociale, qu'il oppose entre eux : l'un qu'il qualifie d'« anthropologique » et l'autre d'« économiste ».
Le modèle « économiste » postule que la dynamique sociale (politique, démographique, éducative) provient essentiellement des transformations économiques. Tandis que le modèle « anthropologique » pose la structure familiale et ses effets (alphabétisation, verticalité) comme source des dynamiques démographiques, politiques, et économiques, qui s'enchainent entre elles dans cet ordre.
Todd (1984), après analyse historique, conclut à la non pertinence du modèle « économiste » et à l'effectivité historique du modèle « anthropologique ».

Figure 1.33 Relation entre structure familiale et efficacité du travail



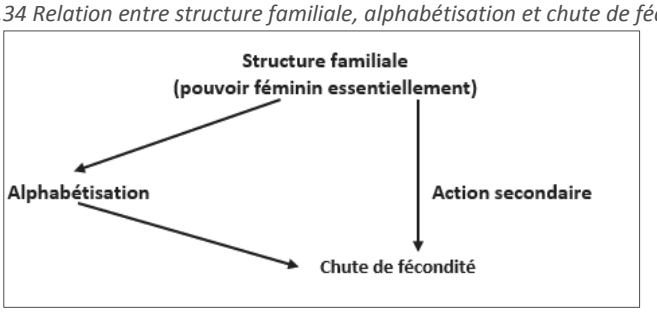
Source : Todd (1984), p.342.

Tableau 1.8 Quelques corrélations concernant le tiers monde (d'après les données statistiques de l'annexe 1)

Espérance de vie et alphabétisation	+ 0,85 (66 Etats)
Taux de natalité et alphabétisation	– 0,73 (62 Etats)
Espérance de vie et revenu par tête	+ 0,69 (65 Etats)
Espérance de vie et urbanisation	+ 0,70 (65 Etats)
Espérance de vie et taux de natalité	– 0,75 (67 Etats)
Revenu par tête et taux de natalité	– 0,58 (65 Etats)
Revenu par tête et alphabétisation	+ 0,56 (64 Etats)

Source : Todd (1984), p.364.

Figure 1.34 Relation entre structure familiale, alphabétisation et chute de fécondité



Source : Todd (1984), p.365.

Dans le modèle anthropologique, tout part de la structure familiale qui par son impulsion, favorise la hausse du taux d'alphabétisation et le recul de l'âge au mariage, facteurs qui à leur tour déclencheront une transformation démographique (*Figure 1.33*).

Cette transformation démographique consiste en une baisse de la mortalité, suivie d'une chute du taux de fécondité.

Le *Tableau 1.8* montre la forte corrélation associant positivement l'alphabétisation à la hausse de l'espérance de vie donc à la baisse de la mortalité. Cette hausse de l'espérance de vie est-elle même

fortement corrélée à l'augmentation du revenu par tête, à l'expansion urbaine, à la chute du taux de natalité.
La jeunesse alphabétisée et massifiée sera le fer de lance d'une révolution (transformation) politique (*Figure 1.33*), caractérisée par exemple par la chute des « anciens » régime en Europe à partir du 17éme siècle.
La principale cause de cette révolution politique évoquée par Todd (1984, p.346) est l'alphabétisation de masse (au moins 50% des hommes alphabétisés, selon le seuil empirique défini par l'historien Lawrence Stone [1969]) d'une jeunesse nombreuse issue de la transition démographique.
Les transformations politiques entrainent la fin des anciens régimes royaux et féodaux pour faire place à des régimes politiques tantôt libéraux, sociaux-démocrates, caractérisés par la fin de la royauté de droit divin, le partage du pouvoir et les libertés économiques (suppression de nombreux monopoles, meilleure garantie de la propriété), ou par des régimes communistes.

L'alphabétisation massive de la jeunesse ayant conduit à la transformation politique, va favoriser la transformation économique caractérisée par le décollage industriel.
Cette transformation économique encore appelé Révolution Industrielle, s'est répandue en Europe vers les 17è et 18è siècles. Elle est caractérisée par des régimes d'orientation libérale garantissant un meilleur respect des propriétés privées mais aussi celui des libertés politiques**,** économiques et religieux**.**
Notons que ce mouvement qui va des institutions politiques vers les transformations économiques, est cohérent avec l'analyse de North (1994), qui montre que les institutions politiques créent des structures incitatives pour l'économie.

5.3.3. Anthropologie et taux d'inflation

Todd (1983) étudie dans le Tableau comparatif ci-dessous, le lien entre taux d'inflation et types familiaux au sein de l'OCDE.
Ici, l'inflation apparait non plus comme un produit exclusif de la gestion monétaire et budgétaire de l'Etat, mais de façon inédite et originale comme une expression des systèmes familiaux.

Tableau 1.9 Structures familiales dominantes et taux d'inflation dans les pays de l'OCDE en 1980

Famille nucléaire absolue		Famille nucléaire égalitaire		Famille autoritaire	
Etats-Unis	13,5 %	France	13,6 %	Japon	8,0 %
Royaume-Uni	18,3 %	Italie	21,2 %	Allemagne	5,5 %
Canada	10,1 %	Espagne	15,5 %	Autriche	6,4 %
Danemark	12,3 %	Grève	24,9 %	Belgique	6,6 %
Pays-Bas	6,5 %	Portugal	16,6 %	Suède	13,7 %
Australie	10,2 %			Norvège	10,9 %
Nouvelle-Zélande	17,1 %			Irlande	18,2 %
Moyennes		*Moyennes*		*Moyennes*	
Simple :	*13,2 %*	Simple :	*18,4 %*	Simple :	*9,9 %*
Pondérée par les populations :	*13,6 %*	Pondérée :	*17,3 %*	Pondérée :	*7,6 %*

Source : Todd (1983), p.123.

Aux deux extrémités se trouvent la famille « autoritaire » avec les taux d'inflation les plus bas (9.9% en moyenne), et famille « nucléaire égalitaire » avec les taux les plus hauts (18.4% en moyenne) : soit un rapport du simple au double.
La famille « nucléaire absolue » occupe une position intermédiaire avec un taux d'inflation moyen de 13.9%.

Ces statistiques montrent les différences importantes et claires du taux d'inflation moyen selon les différents types familiaux. Ceci pourrait s'expliquer par les différences de valeurs anthropologiques, avec notamment les valeurs autoritaires (ou verticales) et inégalitaires qui joueraient un rôle favorable au ralentissement de l'inflation.

5.3.4. Anthropologie et industrie manufacturière

En étudiant le produit manufacturier par habitant des principaux pays de l'OCDE, Todd (1998) conclut que les principaux pays industriels en termes de produit manufacturier par habitant, sont de type familial souche (Japon, Allemagne) ou du moins en ont des composantes significatives (France, Italie).
La valeur autoritaire (ou verticale) est statistiquement le premier critère anthropologique expliquant le niveau de produit manufacturier par habitant.
La supériorité du Japon par rapport à l'Allemagne (tous deux souches) s'explique d'après Todd (1998) par le biais endogamique de la culture japonaise qui leur assurerait une fermeture et un protectionnisme « naturels ».
Todd (1998) trouve un lien fort entre l'anthropologie culturelle et l'activité industrielle au sein de l'OCDE.

5.3.5. Anthropologie et type de capitalisme

Porter (1990), avait étudié dans un ouvrage fouillé, les types de capitalismes nationaux, en se basant sur les critères discriminants suivants : niveau d'intervention de l'Etat, dotations en ressources naturelles, niveau de concurrence interne, rôle joué par les entreprises familiales, et concentrations industrielles régionales.
Albert (1991) a distingué dans *Capitalisme contre Capitalisme*, les types anglo-saxons et rhénans.
Hampden-Turner et Trompenaars (1993) établissent dans « *The Seven Cultures of Capitalism »* un système de valeurs structurant les capitalismes américain, japonais, allemand, britannique suédois et néerlandais.

Todd (1998) évoque l'hypothèse de « deux capitalismes » en comparant le capitalisme « souche » de l'Allemagne et du Japon, au capitalisme « nucléaire absolu » de la sphère anglo-saxonne (Angleterre, USA, Australie) par exemple.
« Les valeurs d'intégration hiérarchique et de continuité héritées de la famille-souche rendaient compte, de leur côté, de la résistance des appareils industriels de l'Allemagne et du Japon, et de la préférence de leurs économies pour le long terme », Todd (2017).
Quant au capitalisme « nucléaire absolu » ou « anglo-américain », il serait caractérisé par une préférence pour le court terme, une recherche de taux de profit élevés, un attachement moindre aux capacités industrielles, une financiarisation très poussée, et une relative indifférence aux inégalités (Todd, 2017).

Todd (1998), souligne toutefois, une distinction notable au sein du capitalisme « souche », entre l'Allemagne et le Japon : « L'Allemagne cependant, au contraire du Japon, tolère les investissements étrangers sur son sol. » (…) « Au Japon, le taux d'ouverture aux investissements extérieurs est négligeable, de 7 %, et l'on doit définir le pays comme fermé. » (…) « L'Allemagne semble en fait capable de gérer l'asymétrie de la circulation du capital en situation d'ouverture, tandis que le Japon ne peut concevoir qu'une fermeture radicale. Comment ne pas rapporter cette différence de comportement aux traits respectivement exogame et endogame des deux systèmes anthropologiques ? ».
Telle est l'interprétation de Todd (1998) que nous souhaitons intégrer dans notre approche cliométrique d'une histoire globale.

5.3.6. Anthropologie et inégalités de revenu

Le tableau ci-après mesure les inégalités de revenu à partir du coefficient de Gini en 2013, au sein des pays de l'OCDE à revenu élevé (définition d'après *World Bank* [2018])
Le tableau mesure les inégalités avant et après taxes et redistributions.

Tableau 1.10 Inégalités de revenu mesurées par le coefficient de Gini

Country	Inequality of Incomes before taxes and transfers	Inequality of Incomes after taxes and transfers
Chile	0,50	0,47
USA	0,51	0,40
Estonia	0,51	0,36
Israel	0,45	0,36
United Kingdom	0,53	0,36
Spain	0,53	0,35
Greece	0,57	0,34
Portugal	0,56	0,34
Italy	0,52	0,33
Canada	0,44	0,32
Ireland	0,58	0,31
Poland	0,47	0,30
Switzerland	0,39	0,30
France	0,50	0,29
Germany	0,51	0,29
Luxembourg	0,48	0,28
Sweden	0,44	0,28
Austria	0,50	0,28
Netherlands	0,43	0,28
Slovakia	0,43	0,27
Belgium	0,49	0,27
Czechia	0,46	0,26
Finland	0,50	0,26
Slovenia	0,47	0,26
Denmark	0,44	0,25
Norway	0,41	0,25
Iceland	0,39	0,24

Source : Roser (2015)

Le coefficient de Gini est en moyenne de 0,48 **avant** taxes et redistributions, et de 0,31 **après** taxes et redistributions : l'action publique (taxes et redistributions) a donc contribué à baisser les inégalités de revenus de 36%.
Le coefficient de variation de l'indice de Gini est de 7% **avant** taxes et redistributions, et de 14% **après** taxes et redistributions : l'action publique (taxes et redistributions), même si elle fait baisser en moyenne les inégalités, aboutit à un paysage international encore plus disparate en termes d'inégalités au sein de l'OCDE.

L'hypothèse avancée par Todd (1999, p. 157) est que cette disparité des inégalités de revenu au sein de l'OCDE pourrait s'expliquer par la diversité des types familiaux et des dimensions anthropologiques entre les pays de l'OCDE.

Nous avons donc testé l'hypothèse de Todd, en mesurant au sein de l'OCDE, les associations entre variables anthropologiques (types familiaux, dimensions) d'une part et inégalités de revenu (coefficient de Gini 2013, en moyenne mobile des 10 dernières années) d'autre part.

Les inégalités de revenus **avant** taxes et redistribution, s'expliquent peu par les variables anthropologiques : 19% de variance expliquée (R^2 ajusté), via un modèle *stepwise* de régression sur axes factoriels AFCM (analyse factorielle des correspondances multiples)
Par contre, les inégalités de revenus **après** taxes et redistribution, s'expliquent fortement par les variables anthropologiques : 63% de variance expliquée (R^2 ajusté), via un modèle *stepwise* de régression sur axes factoriels AFCM.
Notre conclusion est que les politiques publiques de taxation et de redistributions des revenus paraissent fortement reliées aux profils anthropologiques dans différents pays.

Tableau 1.11 Inégalités de revenu : Corrélations et variance expliquée par les variables anthropologiques

Types familiaux et dimensions anthropologiques (en % du territoire)	**Corrélation de Pearson avec indice de Gini**
Famille Nucléaire absolue	NS
Famille Anomique	0,46
Famille Autoritaire	-0,47
Famille Nucléaire égalitaire	NS
Famille Communautaire endogame	NS
Famille Communautaire exogame	NS
Régulation de l'inceste : Mariage asymétrique	NS
Régulation de l'inceste : Faible régulation	-0,41
Régulation de l'inceste : Mariage préférentiel parallèle patrilatéral	NS
Régulation de l'inceste : Exogamie forte	-0,41
Rapports entre frères (Egalité) : Egalité	NS
Rapports entre frères (Egalité) : Indifférence	NS
Rapports entre frères (Egalité) : Inégalité	-0,47
Rapport Parents /Enfants (Autorité) : Non Verticalité	0,58
Rapport Parents /Enfants (Autorité) : Verticalité	-0,58
Statut des femmes : Patrilinéarité, moyenne	NS
Statut des femmes : Bilatéralité	NS
Statut des femmes : Bilatéralité biais Patrilinéaire	NS
Statut des femmes : Patrilinéarité, faible	NS
Statut des femmes : Bilatéralité biais Matrilinéaire	NS
Statut des femmes : Matrilinéarité	NS
Statut des femmes : Patrilinéarité	NS
Statut des femmes : Patrilinéarité, forte	NS
R^2 multiple (modèle stepwise de régression sur axes factoriels AFCM)	0,7
R^2 ajusté multiple (modèle stepwise de régression sur axes factoriels AFCM)	0,63

NS : Valeur non-significative, au seuil de 5%

6. Institutions politiques et économiques

6.1. Définition des institutions

Etymologiquement (Boumard, 1996), le terme « institution » provient du nom latin *institutio,* formé à partir du verbe latin *instituo*. Le verbe *instituo* est composé d'une part du préfixe *in* qui signifie "dans, sur", et d'autre part du verbe *statuo* qui signifie « établir, fonder, organiser, régler ».
A partir du verbe *statuo*, le nom *institutio* est formé par adjonction du suffixe –*tio* (i.e. –tion en français).

Dans le droit romain, « *institutio* désignait à la fois la science des dispositions juridiques et ces dispositions elles-mêmes » (Boumard, 1996)

Le mot « institution » est utilisé pour la première fois en Français en 1190, dans les Sermons de Saint Bernard, en référence à la règle religieuse (Boumard, 1996).

D'un point de vue sociologique, Émile Durkheim (1894), indique dans une définition fondatrice, que le terme « institutions » désigne « toutes les croyances et tous les modes de conduite institués par la collectivité. La sociologie peut être alors définie comme la science des institutions, de leur genèse et de leur fonctionnement ».
Cette définition identifie les institutions aux règles sociales : c'est donc une définition relative à la chose instituée, en l'occurrence les règles ici instituées par la collectivité.

Pour Max Weber (1921), « l'institution se rapproche de l'idée d'association, c'est un groupement dont les règlements statutaires sont octroyés avec un succès relatif à l'intérieur d'une zone d'action délimitante à tous ceux qui agissent d'une manière définissable selon les critères déterminés ».
Cette définition peut être considérée comme extension de celle de Durkheim, dans la mesure où elle identifie les institutions aux organismes fonctionnant selon des règles précises.

Nous retrouvons en Durkheim et Weber, la dualité de l'institution, considérée à la fois comme chose instituée (les règles pour Durkheim) et organisme qui institue (Weber).

D'après Douglas North (1996), « Les institutions sont créées par des humains pour structurer leurs interactions afin de réduire les incertitudes liées à la réalisation de leurs objectifs (de ceux qui établissent les règles) dans les échanges sociaux, politiques et économiques. Je définis les institutions comme des règles formelles (constitutions, lois et lois communes, règlements, etc.), les contraintes informelles (normes de comportement, conventions et codes de conduite imposés de manière interne) et les caractéristiques de mise en œuvre de chacun. Parce qu'ils constituent la structure d'incitation d'une société, ils définissent la façon dont le jeu est joué et la façon dont ils évoluent déterminent la façon dont le jeu est joué à travers le temps. »

Cette définition large de North (1996) ajoute le caractère formel ou informel des institutions, à la définition classique de Durkheim (1894).
Mais, North crée une nette et essentielle distinction entre institutions et organisations en précisant que « si les institutions sont les règles du jeu, les organisations et leurs entrepreneurs en sont les joueurs » (North, 1994, p. 361).

Une conséquence importante de la définition de North est le caractère incitatif des institutions pour l'ensemble de la société : « les institutions constituent la structure d'incitation d'une société et les

institutions politiques et économiques, en conséquence sont le déterminant fondamental de la performance économique » North (1994).

Par ailleurs, nous-nous rendons compte que la définition de North, considère de fait l'anthropologie culturelle, comme une institution. Et il le dit même assez clairement en affirmant que « l'histoire démontre que les idées, les idéologies, les mythes, les dogmes et les préjugés comptent » (North, 1990).

Il nous faut donc préciser les domaines institutionnels (autres que culturels) que nous étudions dans cette partie institutionnelle de notre propos : Dans le cadre de cette thèse, les institutions sont analysées au sens d'institutions politiques et économiques.
L'aspect culturel (donc le plus souvent informel, car majoritairement inconscient d'après Hofstede et Minkov, 2010) étant étudié dans le domaine de l'anthropologie culturelle.
Notre distinction entre culture et institution s'appuie sur la discrimination introduite par North, que nous-nous proposons de rendre opérationnelle.

6.2. Relations entre institutions et économie

Les institutions économiques relèvent du domaine économique par définition : Elles englobent les règles et organismes formels ou informelles relevant de l'activité économique.
Le lien est donc immédiat entre institutions économiques et activité économique.
Pour ce qui est des institutions politiques, (ou politico-économiques), de nombreux auteurs ont mené réflexions et études sur leurs liens avec l'activité économique.

Notons ici l'existence de tout un champ de l'économie intitulé *law and finance* qui étudie les liens entre l'institutionnel et l'économique. Il existe notamment un champ de recherche entre les institutions héritées de la colonisation européenne et les trajectoires économiques post-coloniales des pays colonisés.
Parent (2017) identifie trois écoles de pensée ayant une influence sur l'étude des relations entre les institutions coloniales et leur impact sur la croissance économique : d'abord les travaux de Daron Acemoglu et de ses collègues (Acemoglu *et al.* [2001b], Acemoglu et Robinson [2015]) sur le rôle des institutions "extractives" par rapport aux institutions "inclusives", ensuite les travaux de Rafael La Porta et de ses collègues (La Porta *et al.*, 1998) sur le rôle des différentes formes de système juridique colonial sur le développement économonique ultérieur, et enfin les travaux sur l'influence d'autres facteurs, tels que la géographie (Gallup *et al.,* 1998).
Parent (2017), après une large revue de la littérature sur les institutions coloniales et les trajectoires de développement en Afrique, conclue que le poids et l'influence des institutions héritées de la colonisation s'amenuise avec le temps, au profit du progrès démocratique qui joue un rôle important dans leur développement économique.

Douglass North : incitations, coûts de transactions et rapports de forces sociales
North (1994) indique que « la structure créée par les humains pour organiser leur environnement politico-économique est le déterminant essentiel des performances d'une économie » et cela par le biais des incitations.
« Les organisations qui se constituent vont refléter les opportunités créées par la matrice institutionnelle. Si le cadre institutionnel récompense la piraterie, des organisations pirates seront créées ; s'il récompense les activités productives, des organisations — des firmes — seront créées afin de s'engager dans des activités productives » North (1994).

North (2005, p. 20) mentionne le cas des institutions garantissant les droits de propriété, qui ont pour conséquence de faire baisser les coûts de transaction. La baisse des coûts de transactions, entrainant à son tour, la hausse des transactions économiques de consommation et surtout d'investissement.

Dahlman (1979), distingue trois types de coûts de transaction : les « coûts de recherche et d'information », les « coûts de négociation et de décision » et enfin les « coûts de surveillance et d'exécution ».
Coase (1960) établit que les coûts de transactions sont fortement impactés par les institutions.

A titre illustratif, nous pouvons avancer que, les normes et formalités administratives, les barrières règlementaires à l'entrée de marchés, l'opacité sur les opportunités officielles ou sur les diverses règlementations incitatives, l'état des infrastructures de communication et de transport, l'incertitude sur le droit de propriété et les contrats, constituent des contextes institutionnels qui créent des coûts de transaction pour les acteurs économiques.

Dans les faits, la structure institutionnelle est « un mélange d'institutions qui promeuvent les activités qui accroissent la productivité et d'institutions qui créent des barrières à l'entrée, encouragent les restrictions monopolistes et empêchent la transmission à faible coût de l'information » North (1990).

North (1994, p. 361) pose que « c'est l'interaction entre les institutions et les organisations qui détermine l'évolution institutionnelle d'une économie », car non-seulement, les institutions déterminent les incitations qui influencent les décisions et l'évolution des acteurs économiques, mais également, se met simultanément en place un « processus de rétroaction » North (1990).
Dans ce processus de rétroaction, les organisations cherchent à modifier à leur tour, le cadre institutionnel dans lequel ils évoluent.
La distinction entre institutions et organisations est donc essentielle chez North, pour comprendre les mécanismes d'incitation et de rétroaction.

Les tentatives des entrepreneurs et organisations pour exercer quelque influence sur les institutions se feront dans l'objectif d'améliorer la position concurrentielle des acteurs. Il s'agira donc selon le cas de tentatives de « restriction à l'entrée, de la création d'un monopole ou d'une redistribution des revenus » (North, 1996).
L'alternative à la rétroaction institutionnelle comme moyen d'amélioration de position concurrentielle pour les organismes, est l'adoption de meilleures technologies pour une meilleure productivité.

Aussi, pour les entrepreneurs et organisations, le niveau d'incitation à influer par rétroaction sur les institutions, est lié positivement au degré de concurrence instauré par les institutions du moment, mais lié a contrario négativement aux incitations institutionnelles vers des activités productives et d'innovations technologiques aux fins d'amélioration de productivité.

Sur un plan social, « les institutions ne sont pas nécessairement ou même habituellement créées pour être socialement efficaces ; elles sont plutôt créées - ou tout au moins les règles formelles - pour servir les intérêts de ceux qui possèdent le pouvoir de négociation en vue de créer de nouvelles règles » (North, 1994, pp. 360-361). D'après North donc, les institutions ne profitent pas forcément à l'ensemble de la société, puisqu'elles sont le résultat d'un rapport de forces sociales.
Raison pour laquelle « alors que la technologie peut déterminer la limite supérieure de la performance économique, (…) cette limite est rarement atteinte, voire jamais atteinte » (North, 1996).

Cette vision sociale est très importante car cela peut permettre d'éclairer, au moins partiellement, la dynamique des institutions par le jeu des rapports de forces sociales.

Au-delà de l'influence des rapports de forces sociales, et des mécanismes de « rétroaction » des acteurs, North (1996), pousse encore plus loin la recherche des causes de la dynamique des institutions. Il va donc élargir le champ de recherche : « Une compréhension globale de la performance économique dans le temps nécessite une fusion des théories sur les changements institutionnels, démographiques et relatifs au stock de connaissances afin d'avoir une approche globale des problèmes. »

Acemoglu et Robinson : la nécessité d'inclusivité des institutions
Acemoglu (2012) adopte la même définition des institutions que North, en soulignant trois caractéristiques : D'abord, les institutions sont conçues par les hommes, ensuite, elles fixent les règles du jeu dans la société, et enfin, elles créent des incitations.

Comme pour North (1994), Acemoglu et Robinson (2015, p.110) considèrent les institutions à un moment donné comme une résultante des rapports de forces antagonistes au sein de la société : « En cas de conflit au sujet des institutions, la suite dépend des individus ou des catégories qui l'auront emporté sur un plan politique - ceux qui auront obtenu des soutiens plus nombreux ou des ressources supplémentaires, ou ceux qui auront su former des alliances plus efficaces. En somme, la discorde sera tranchée en fonction de la répartition du pouvoir politique dans une société donnée. »

Acemoglu et Robinson (2015, p.112) définissent deux formes antagonistes d'institutions, selon leur impact économique : « Quand des institutions politiques sont suffisamment centralisées et pluralistes, elles seront dites institutions politiques inclusives. Dans le cas contraire, on parlera d'institutions politiques extractives. »

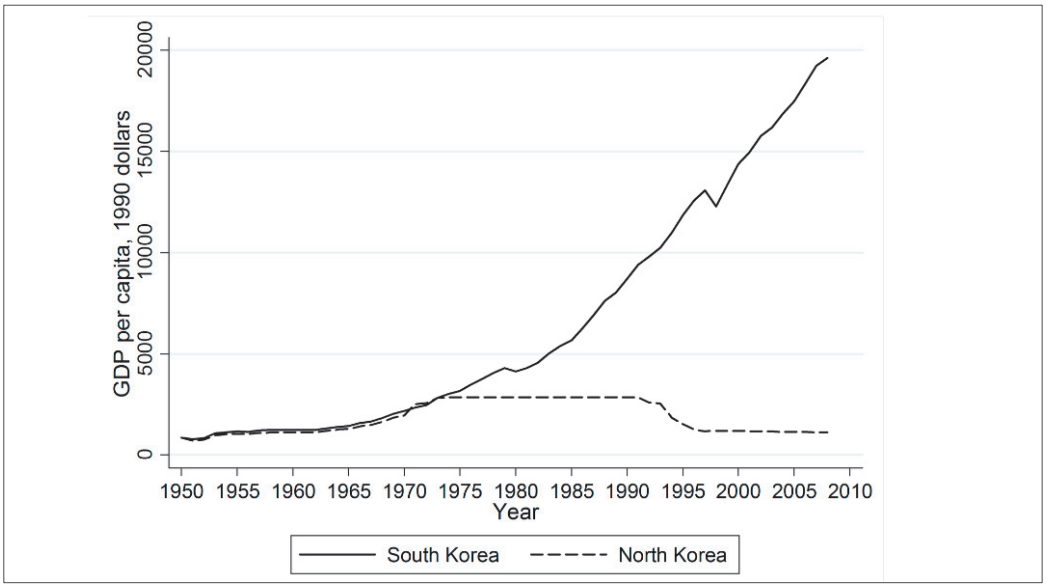
Nous observons que le pluralisme politique ici évoqué par Acemoglu et Robinson (2015), permet d'augmenter l'efficacité sociale des institutions : En effet, d'après la définition même des institutions North (1996), les participants à l'élaboration des institutions agissent pour leurs propres intérêts. Donc l'inclusion, augmente la pluralité des participants, et donc la pluralité des intérêts pris en compte par les institutions.
La centralisation, qui est la deuxième condition de l'inclusivité des institutions, signifie l'existence d'un pouvoir unifié sur un territoire, source d'une loi uniforme au territoire et d'une capacité à faire respecter cette loi.

Les institutions politiques inclusives, auront tendance à produire des institutions économiques également inclusives : « Des institutions politiques inclusives, parce qu'elles répartissent largement le pouvoir, auraient tendance à supprimer des institutions économiques qui dépouillent la majorité, érigent des barrières et font en sorte que les marchés ne fonctionnent qu'au profit d'une minorité », Acemoglu et Robinson (2015, p.112).
A contrario, les institutions politiques extractives, engendreront des institutions économiques extractives : « Les institutions politiques extractives concentrent le pouvoir entre les mains d'une élite réduite et n'imposent que peu de contraintes à l'exercice de ce pouvoir. Les institutions économiques sont souvent structurées par cette élite, qui s'efforce d'extraire les ressources de la société à son seul profit » (*ibid*, p.112).

Les mêmes auteurs citent la Glorieuse Révolution anglaise de 1688, qui instaure le pluralisme politique et centralise le pouvoir au Parlement, comme sources des premières institutions politiques inclusives de l'Histoire. S'en suivent des mesures économiques inclusives comme une garantie renforcée du droit de propriété matérielle et intellectuelle, l'application d'une même loi pour tous, une taxation plus égalitaire, la quasi-suppression des monopoles, la construction des grandes infrastructures de transport (routes, canaux, chemins de fer).
Pour Acemoglu et Robinson (2015, p.141), ce n'est donc point un hasard que « la révolution industrielle ait commencé en Angleterre quelques décennies après la Glorieuse Révolution ».

Un autre exemple d'après-guerre frappant sur l'importance primordiale des institutions pour l'économie est le cas grandeur nature de la divergence entre les deux Corée dès lors qu'elles ont suivi des voies institutionnelles radicalement distinctes à l'issue de la Guerre de Corée (1950 à 1953).

Figure 1.35 Expériences naturelles de l'histoire : par exemple la division de la Corée en Nord et Sud



Source: Acemoglu (2012) "Natural experiments of history : e.g. the division of Korea into North and South".

Acemoglu et Robinson (2015, p.117) affirment donc que les institutions économiques inclusives, produisent des transformations technologiques et de la croissance économique, qui par le jeu schumpetérien de la destruction créatrice, engendreront dynamiquement des groupes sociaux et des organismes perdants ou gagnants « dans le champ politique et sur les marchés économiques ».
D'après les auteurs, la crainte de ce changement, est un frein au mouvement vers des institutions politiques et économiques plus inclusives.

Nous pensons par ailleurs, que ces mêmes craintes de perte de position favorable peuvent, pousser certains groupes sociaux et organismes, à saper les fondements inclusifs d'institutions déjà inclusives, en poussant vers des réformes extractives par le biais du « processus de rétroaction » de North (1990).
En effet, les innovations de rupture rendent obsolètes une partie du capital matériel et du savoir-faire des personnes ou organismes « établis » (c'est nous qui nommons). Il serait donc difficile pour un « établi » (a fortiori protégé par des barrières à l'entrée) d'adopter de lui-même des innovations de rupture, à moins d'y être forcé par des motifs commerciaux ou légaux. Car cela reviendrait à d'une part, réaliser d'importants investissements vers de nouvelles technologies qu'il ne maitrise pas forcément, et d'autre part frapper d'obsolescence et perdre une partie significative de son capital matériel antécédent, de son savoir-faire durement acquis et de ses ressources humaines (le cas échéant, en cas de licenciements).

Acemoglu et Robinson (2015) notent toutefois que les institutions extractives peuvent créer temporairement des périodes de croissance économique, par exploitation de technologies déjà existantes associée à une allocation autoritaire de ressources nationales vers l'industrie.
L'exemple majeur cité est celui de l'URSS dont l'économie connut une forte croissance (+6% par an) entre 1928 et 1960, mais subit un coup d'arrêt dès 1970. Ce cas correspond à celui d'un régime extractif fortement centralisé, mais sans pluralisme.
Cependant, cette croissance extractive est jugée non soutenable par les auteurs, car les atteintes aux personnes et aux biens ne peuvent créer les incitations à l'entrepreneuriat et à l'investissement et donc aux transformations technologiques continues par destruction créatrice. Le groupe social au pouvoir, en situation monopolistique, n'investira pas non-plus dans de nouvelles technologies au point de menacer sa position dominante.

North (1994, p.367) est également du même avis sur la non-soutenabilité à terme d'une croissance économique dans un régime politique autocratique : « Si la croissance économique peut se produire à

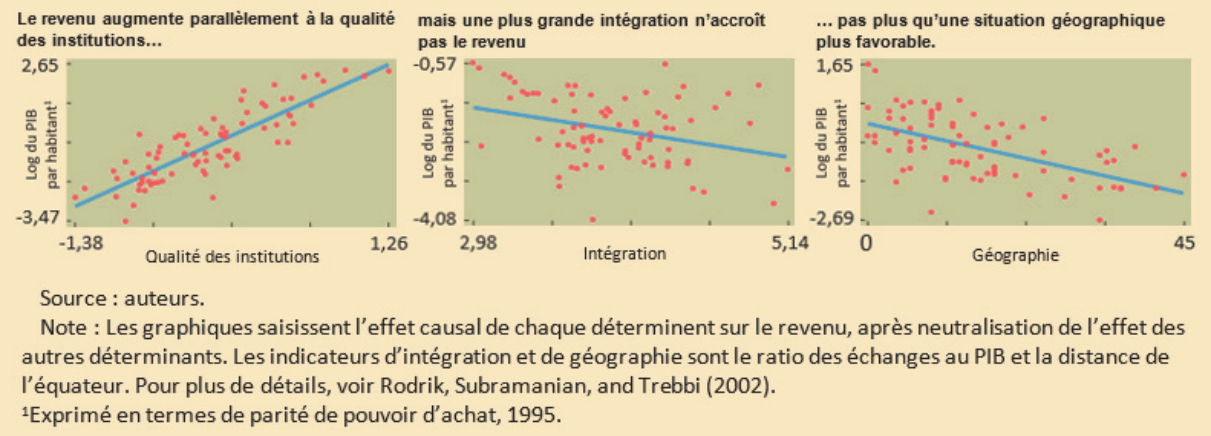
court terme avec des régimes autocratiques, la croissance économique à long terme implique le développement de l'Etat de droit »

Rodrik : « La primauté des institutions »

Rodrik et Subramanian (2003) ont dans une étude empirique, montré que la « qualité » des institutions est très positivement et significativement associé au revenu par habitant (en logarithme) ; tandis que l'ouverture au commerce international est positivement associé à la « qualité » des institutions, mais négativement (quoique faiblement) associé au revenu par habitant (en logarithme).
La « qualité » des institutions étant ici mesurée par un indicateur composite de variables mesurant la garantie du droit de propriété, ainsi que le respect de l'Etat de droit.

Il semble cohérent de dire que l'ouverture au commerce international est positivement associée à la « qualité » des institutions car les pays qui ont, au moment de l'étude, un niveau élevé de « qualité » des institutions, sont également ceux ont un revenu par habitant élevé et qui se conforment par ailleurs aux recommandations du Consensus de Washington (notamment en ce qui concerne le libre-échange). Mais les résultats montrent par ailleurs qu'une fois la corrélation avec la « qualité » institutionnelle neutralisée, le libre échange est négativement (quoique faiblement) associé au revenu par habitant (en logarithme).

Figure 1.36 Primauté des institutions



Source : Rodrik, Subramanian (2003)

Notons que le niveau d'ouverture au commerce international (libre échange) devrait être intégré aux institutions au sens de North (i.e. en tant qu'élément majeur de la règlementation économique). Cela n'a pas été le cas chez Rodrik et Subramanian (2003) qui se sont limités aux instituions que l'on pourrait appeler « internes » qui ne régissent pas les relations commerciales externes.
Rodrik (2001), indique que pour un pays en voie de développement (PVD), le retour sur investissement d'une intégration aux normes de l'OMC (mesures en douane, normes sanitaires et phytosanitaires, droits de propriété intellectuelle) est loin d'offrir le meilleur rendement en termes de développement, par rapport aux autres allocations alternatives (éducation des filles par exemple) des rares ressources dont dispose ce type de pays. Et il faudrait rajouter à cela l'exposition accrue aux risques financiers internationaux.

Parente et Prescott : approche ricardienne contemporaine et héritage mono-téléologique de l'économie politique smithienne

Pour Parente et Prescott (2002), les bonnes institutions économiques doivent reposer sur les principes de « concurrence, libre échange et privatisation ». Pour ces auteurs, l'entrave à ces politiques économiques est principalement le fait de groupes d'intérêts et corporatismes divers (patronaux ou salariés) ; ce faisant, ils empêchent les entreprises d'utiliser les meilleures technologies disponibles au

niveau mondial et donc leur pays d'accéder au meilleur niveau de productivité et de richesse permis par la technologie disponible.
Parente et Prescott (2002), soulignent également l'intérêt du fédéralisme avec une certaine autonomie locale sur la législation économique associée à un marché unique à l'échelle nationale : la compétition induite entre les entreprises des territoires fédérés, aurait alors un effet modérateur sur les tentatives d'empêchement des entreprises à utiliser les meilleures technologies disponibles au niveau mondial.

Nous constatons que Parente, Prescott (2002), sont des héritiers de la pensée libérale et libre-échangiste ricardienne (et même physiocratique). En comparant cette pensée économique avec celle du précurseur de l'école classique qu'est Smith (1776), nous constatons qu'il leur manque la seconde branche de l'économie politique telle que définie par Smith (1776, tome 2, P.11) et adoptée implicitement par List (1841) : A savoir que « l'économie politique, considérée comme une branche des connaissances du législateur et de l'homme d'État, se propose deux objets distincts : le premier, de procurer au peuple un revenu ou une subsistance abondante, ou, pour mieux dire, de le mettre en état de se procurer lui-même ce revenu, ou cette subsistance abondante ; le second, de fournir à l'Etat ou à la communauté un revenu suffisant pour le service public ; elle se propose d'enrichir à la fois le peuple et le souverain ».
Ces deux aspects (interne et externe) de l'économie politique smithienne sont rappelés récemment par Sabéran (2008).

De fait, la pensée économique libérale et libre-échangiste inconditionnelle ricardienne est une pensée incomplète où manquent les aspects géostratégiques économiques nécessaires à la protection (et même au développement) de la prospérité et des capacités intérieures. Cette optique ricardienne qui pense la prospérité (et la productivité) sans penser la gestion des risques géostratégiques, est aussi non-équilibrée que son double opposé qui est la pensée mercantiliste qui ne s'occupe principalement que des problématiques de géostratégiques de puissance sans s'inquiéter du pouvoir d'achat et de la productivité internes (Jessua, 1991, p.27).

A notre sens donc, c'est Smith (1776) qui unifia la pensée physiocratique et la pensée mercantiliste, en posant comme norme la recherche de la prospérité de la population (productivité, pouvoir d'achat) sans pour autant sacrifier la puissance et la souveraineté géostratégique du pays. Ce qui l'amena à poser régulièrement des exceptions aux principes libéraux et libre-échangistes à chaque fois que leur application pouvait nuire aux intérêts géostratégiques nationaux.

Ainsi, Smith (1776, tome 2, page 50) approuve l'Acte de navigation du 9 octobre 1651, qui réserve aux navires et aux équipages nationaux le droit d'entrer dans les ports de Grande-Bretagne. Motif : la navigation est « une espèce particulière d'industrie *[qui]* est nécessaire à la défense du pays. Par exemple, la défense de la Grande-Bretagne dépend beaucoup du nombre de ses vaisseaux et de ses matelots ».
Et pour bien se faire comprendre, Smith affirme sans ambages que « comme la sûreté de l'Etat est d'une plus grande importance que sa richesse, l'Acte de navigation est peut-être le plus sage de tous les règlements de commerce de l'Angleterre » (*ibid*, p.52).

Deuxièmement, Smith recommande la réciprocité des droits de douane pour protéger l'industrie manufacturière nationale : Dans ce cas donc, « il sera avantageux, en général, de mettre quelque charge sur l'industrie étrangère pour encourager l'industrie nationale, c'est quand le produit de celle-ci est chargé lui-même de quelque impôt dans l'intérieur. Dans ce cas, il paraît raisonnable d'établir un pareil impôt sur le produit du même genre, venu de fabrique étrangère » (*ibid*, p.52).

Toujours au nom de la défense nationale et d'une certaine souveraineté, Smith recommande la subvention pure et simple des industries ayant un intérêt géostratégique, si celles-ci venaient à être fragiles économiquement : « A la vérité, si quelque fabrique particulière était nécessaire à la défense nationale, il pourrait bien n'être pas très sage de rester en tout temps dans la dépendance de ses voisins

pour l'approvisionnement ; et si une fabrique de ce genre ne pouvait pas se soutenir chez nous sans protection, il serait assez raisonnable que toutes les autres branches d'industrie fussent imposées pour l'encourager. Peut-être pourrait-on justifier, d'après ce principe, les primes à l'exportation des voiles de marine et de la poudre de fabrique anglaise » (*ibid*, p.119).

Enfin, Smith, au nom de l'intérêt général (moralité de la population, défense nationale, compétences cognitives de la main d'œuvre, compréhension politique du citoyen) recommande le développement de l'éducation publique de masse (*ibid*, tome 2, p.404-413).
La conscription est vivement encouragée pour la défense nationale et les risques de « coup d'Etat » (*ibid*, tome 2, p.411).
Il en va de même pour le développement par l'Etat des infrastructures de communication et de transport, pour les externalités positives de développement du marché intérieur (*ibid*, tome 2, p.444).

Autre fait notable, Smith (1776) donna une importance particulière à l'industrie manufacturière dans le processus de gain de productivité par l'aptitude élevée à la division du travail et la mécanisation induite (*ibid*, tome 1, p.73-75) et même de la puissance et la souveraineté nationale. Ce dernier point fit dire à Smith que « Il faut observer qu'un grand pays n'a jamais subsisté ni pu subsister, sans qu'il ait eu chez lui quelque espèce de manufacture ; et quand on dit d'un pays qu'il n'avait point de manufactures, cela doit toujours s'entendre des fabriques d'ouvrages finis et recherchés, ou de ceux qui sont destinés à être vendus au loin » (*ibid*, tome 1, p.197).

Smith (1776) (tout comme List [1841]) souligna également le rôle bénéfique qu'avaient les manufactures sur les libertés civiles : « le commerce et les manufactures introduisirent par degrés un gouvernement régulier et le bon ordre, et avec eux la liberté et la sûreté individuelle, parmi les habitants de la campagne qui avaient vécu jusqu'alors dans un état de guerre presque continuel avec leurs voisins, et dans une dépendance servile de leurs supérieurs. De tous les effets du commerce et des manufactures, c'est sans comparaison le plus important, quoiqu'il ait été le moins observé » (*ibid*, tome 1, p.502). Il s'agit donc d'une externalité institutionnelle positive de l'industrie manufacturière.

Cela dit, Smith, s'il définit bien les deux aspects interne et externe de l'économie politique, ne les adapte pas pour autant stratégiquement aux phases de développements d'un pays.
Ce que fit List (1841).

List : le précurseur revisité par North, Acemoglu, Robinson et Rodrik
List (1841) distingue cinq phases successives de développement économique : l'état « sauvage », l'état « pastoral », l'état « agricole », l'état « agricole et manufacturier », et enfin, l'état « agricole, manufacturier et commercial ».
L'état agricole correspond à une économie essentiellement agricole, en termes de population active et de revenu.
L'état agricole-manufacturier décrit une économie où l'agriculture est certes importante, mais où l'industrie a pris son essor de manière significative.
Quant à l'état agricole-manufacturier-commercial, il correspond à une phase où l'agriculture joue un moindre rôle, mais avec une industrie forte qui sert de base à un commerce international axé sur les exportations.

List préconise des institutions économiques distinctes à chaque phase du développement économique, afin de permettre la transition vers les phases supérieures.

Ainsi, List montre l'utilité du « système douanier » pour le passage de l'état agricole vers l'état agricole-industriel, en se basant sur des nombreux cas historiques du 12e au 19e siècles, concernant les Italiens, les Hanséates, les Flamands, les Hollandais, les Anglais, les Espagnols, les Portugais, les Français, les Allemands, les Russes, et les Américains du Nord.

Ce « système douanier » est un ensemble de mesures restrictives douanières, vis-à-vis des pays plus avancés économiquement.
Le « système douanier » inclut des mesures de « prohibitions d'entrée, droits d'importation, restrictions maritimes, primes de sortie » ; L'objectif étant et permettre le développement de l'industrie manufacturière locale, en le protégeant de la concurrence des industries plus fortes des pays plus avancés dans les phases du développement économique.

Les institutions préconisées par List pour sortir un territoire de la phase agricole, ne se limitent pas au « système douanier ».

Ces institutions incluent également des institutions politiques centralisées (effectives et uniformes dans l'ensemble du pays), « des lois qui garantissent au citoyen sécurité pour sa personne et pour ses propriétés, et libre exercice de ses facultés morales et physiques, des institutions qui règlent et facilitent le commerce, en même temps que la suppression de celles qui oppriment l'industrie, la liberté, l'intelligence et la moralité, la suppression des institutions féodales, par exemple » (List, 1841, p. 101).
Ces institutions politiques d'orientations libérales sont une des composantes majeures des « forces productives » du pays.
Ainsi, List intègre directement et explicitement les institutions parmi les « forces productives » du pays, donc dans le mécanisme de production économique.

Par « forces productives » du pays, List (ibid., pp. 96,220) désigne d'une part, des « forces individuelles » dont le travail, « l'esprit d'invention », « l'esprit d'entreprise », l'épargne, la « moralité », « l'intelligence », les « capitaux matériels » ;
De plus, ces « forces productives » du pays, possèdent une composante sociale constituée par les « forces sociales » dont les « institutions », « l'indépendance » et la « puissance » du pays. (*ibid*, pp. 96, 221). Ces institutions sont bonnes économiquement dans la mesure où elles promeuvent la liberté, la protection de la propriété, et un système douanier adapté.
Et enfin, List compte les ressources naturelles (« fonds naturels »), comme « force productive » surtout si elles sont mises en valeur par l'industrie.

En plus du « système douanier », et des institutions internes garantissant droit de propriété et libertés, List montre l'utilité historique d'une troisième catégorie d'institutions visant à mettre en œuvre la « maxime » du « préférer constamment l'importation des forces productives à celle des marchandises ». List cite les mesures d'Edouard IV (15ème siècle) qui importa en Angleterre des milliers de moutons d'Espagne, transforma l'Angleterre en pays exportateur de laine tout en prohibant plus tard l'importation de laine d'Espagne.

Parmi les activités productives, l'industrie manufacturière occupe chez List, la part centrale, car non seulement elle met en valeur les ressources naturelles et les produits de l'agriculture locale et étrangère, mais c'est également cette activité productive manufacturière qui permettra au pays de devenir un pays commerçant en accédant grâce aux exportations, à la phase agricole-manufacturière-commerciale.
Au point que List finit par énoncer comme principe, « qu'une nation est d'autant plus riche et d'autant plus puissante qu'elle exporte plus de produits manufacturés, qu'elle importe plus de matières brutes ».

Au final, l'adaptation des institutions économiques par rapport aux phases de développement économique, passe par un dosage idoine des mesures protectionnistes : d'abord avec un dosage protectionniste faible pour assurer « l'éducation commerciale » du pays (notamment en phase agricole, avec niveau « inférieur d'intelligence ») ; puis avec un dosage protectionniste fort pour assurer « l'éducation industrielle » du pays lorsque celui-ci aura un niveau suffisant « d'éducation intellectuelle, politique et économique » pour que sa production commence à se substituer aux importations (ibid, p. 104).

Arrivé au stade agricole-industriel-manufacturier, le pays devrait adopter une politique de libre-échange (i.e. lever au maximum les barrières protectionnistes), afin d'importer un maximum de matières premières et brutes et exporter ses produits manufacturé (*ibid*, p.105). Cette politique de libre-échange est à moduler inversement par rapport au niveau de développement industriel des partenaires commerciaux.
A ce propos, List (ibid, p. 104) énonce le principe de gestion économique suivant : « Le droit protecteur en faveur d'une industrie ne doit pas descendre assez bas pour que l'existence de celle-ci puisse être compromise par la concurrence étrangère. ***On doit prendre pour règle invariable de conserver ce qui existe, de protéger l'industrie nationale dans son tronc et dans ses racines*** ».

Dans tous les cas, quel que soit le degré de protectionnisme appliqué, l'économie intérieure du pays, doit garder des institutions économiques garantissant libertés et droits de propriété afin de permettre l'expression des talents et la concurrence saine entre les acteurs économiques ; sans quoi elles faibliraient et ne seraient pas en mesure de soutenir à terme la concurrence internationale, ni obtenir une croissance de la productivité (ibid, p. 101).
Ces mêmes institutions de liberté/propriété internes, ont joué historiquement (ibid, p. 225), notamment en Angleterre, un rôle attractif pour les capitaux et talents étrangers.

Pour List il n'y a donc pas de bonnes institutions économiques invariantes quelles que soient la phase de développement économique et de la concurrence internationale subie, surtout, s'agissant des politiques de commerce extérieur ; à l'exception notable des institutions internes (politiques et économiques), qu'il recommande libres et garantissant le droit de propriété en tout temps.

La seule constante étant l'objectif économique à savoir « conserver, développer et perfectionner » les « forces productives » de son pays ou territoire (List, 1841, pp. 95-98).
Les institutions, ne sont alors bonnes que dans la mesure où elles s'adaptent et concourent à l'atteinte de cet objectif.
Les institutions sont dans cette perspective, un moyen stratégique, donc soumis au principe d'adaptation nécessaire selon les circonstances (phases de développement, niveau de puissance, et situation vis-à-vis des partenaires étrangers).

Aghion et Howitt (2006) : Les institutions adaptées
Aghion et Howitt (2006), considèrent les institutions à travers trois aspects politiques (concurrence et entrée sur les marchés, éducation et politique macroéconomique), puis montrent que la politique institutionnelle optimale varie selon la distance à la frontière technologique et le développement financier.
Ainsi, Aghion et Howitt (2006) montrent que « l'augmentation du nombre d'entrées et de sorties (rotation des entreprises) et l'accent mis sur l'enseignement supérieur sont plus propices à la croissance dans les pays plus proches de la frontière technologique », que de plus, « les politiques budgétaires contra-cycliques favorisent davantage la croissance dans les pays à développement financier faible », et qu'enfin, il existe « des effets d'interaction importants entre politiques et variables d'état, telles que la distance à la frontière technologique ou le développement financier ».

Les résultats de Aghion et Howitt (2006), corroborent ceux de List, notamment sur la nécessité d'adapter les institutions politiques et économiques, aux états (ou phases) de développement économique ; Ils concluent qu'il n'existe pas d'institutions optimales qui vaillent par tout temps et en toutes circonstances, quel que soit l'état d'avancement économique du pays. Ces institutions optimales sont contingentes à l'étape de développement d'un pays. Leur modèle d'inspire des idées de Gershenkron (1962) qui introduit la notion d'institutions « appropriées ».

Une synthèse de Smith, List, North, Acemoglu-Robinson et Rodrik
En combinant les recommandations institutionnelles de Smith (1776), List (1841), North (1994, 1996), Rodrik (2001), Rodrik et Subramanian (2003), nous pouvons conclure que les bonnes institutions économiques sont des instruments stratégiques (donc adaptatives et non rigides) au service du développement économique, de l'indépendance et de la puissance politique.

Ce développement économique inclut non seulement la hausse de la productivité et du niveau général du pouvoir d'achat, mais également le bien-être issu entre autres de « la qualité de l'environnement physique et social » (North, 1996) ; ce développement se fait par un renforcement progressif (à travers ses phases successives) de la base industrielle manufacturière identifiée comme la principale activité productive source de la hausse de productivité de tous les secteurs d'activité (List, 1841).
En effet pour List (1841, p.96), la base industrielle manufacturière, sa conservation, son développement et sa perfection, constituent le fondement du développement économique et institutionnel et la puissance du pays.
Il précise que : « L'industrie manufacturière (…) est favorable aux sciences, aux arts et aux progrès politiques ; elle augmente le bien-être général, la population, le revenu de l'État, et la puissance du pays ; elle fournit à celui-ci les moyens d'étendre ses relations dans toutes les parties du monde, et de fonder des colonies ; elle alimente les pêcheries, la navigation marchande et la marine militaire. Par elle seulement l'agriculture du pays atteint un haut point de perfection ».

Par ailleurs, North (1994) indique que les institutions formelles ne sont "légitimes" (nous dirions soutenables, efficaces et effectives) que dans la mesure où elles sont cohérentes avec la culture du pays (institutions informelles, normes informelles et valeurs).

Donc, les bonnes institutions économiques doivent s'adapter (sans idéologie rigide) à la culture du pays et sa phase de développement, mais aussi à sa situation concurrentielle vis-à-vis de ses partenaires commerciaux.
Les institutions économiques des échanges externes (dans la mesure où l'indépendance et la puissance politique le permettent) servent de levier d'adaptation, tandis que les institutions économiques **internes** doivent tendre dans une direction de renforcement inclusif et de respect des droits de propriété et libertés économiques individuelles, assurés par l'Etat de droit et la démocratie (Acemoglu et Robinson, 2012) ; le tout au service des activités productives, notamment celles qui procurent des gains de productivité importants, donc historiquement l'industrie manufacturière, sources de prospérité, d'indépendance et de puissance du pays.

Sur le rôle fondamental de l'industrie manufacturière, les constats historiques de List, ainsi que les hypothèses de Smith sont confirmés par des études empiriques, comme par exemple : Petri (1993), Fishlow et al. (1994), Rodrik (2001, 2002), Deubel (2008, pp.485-489), Colletis (2012), Cohen et Buigues (2014).

6.3. Mesures des institutions

6.3.1. Revue de la littérature, sur la mesure des institutions

Une large revue bibliographique nous a permis de recenser les mesures des institutions suivantes (dont la plupart sont déjà disponibles dans notre base de données) :

Tableau 1.12 Revue des mesures des institutions

Article	Mesure des institutions (intitulé dans l'article)	Mesure des institutions (traduction en français)	Cité dans …
Bosma et al. (2018)	Size of government	Taille de l'Etat (ou de l'administration publique)	Bosma N., Content J., Sanders M., Stam E. (2018)
Bosma et al. (2018)	Regulation of credit, labor, and business	Régulation du crédit, du travail et de l'activité économiques	Bosma N., Content J., Sanders M., Stam E. (2018)
Bosma et al. (2018)	Financial stability (also known as access to sound money).	Stabilité financière (également connu sous le nom d'accès à une monnaie saine).	Bosma N., Content J., Sanders M., Stam E. (2018)
Bosma et al. (2018)	Fear of failure	Peur de l'échec	Bosma N., Content J., Sanders M., Stam E. (2018)
Bosma et al. (2018)	Entrepreneurship as a good career choice.	Entreprenariat comme un bon choix de carrière.	Bosma N., Content J., Sanders M., Stam E. (2018)
Bosma et al. (2018)	Perceived knowledge and skills to start a business. It is proxied by the percentage of the working age population who indicate that they believe to have the required skills and knowledge to start a business, as measured via the GEM Adult Population Surveys	Auto-évaluation de la population active par rapport au niveau de connaissances et compétences nécessaires pour créer une entreprise. Cet indicateur est approximé par le pourcentage de la population en âge de travailler qui croit avoir les compétences et connaissances nécessaires pour créer une entreprise, tel que mesuré par The GEM Adult Population Surveys	Bosma N., Content J., Sanders M., Stam E. (2018)
Basnet (2017)	Control of Corruption	Lutte contre la corruption	Basnet (2017)
Basnet (2017)	Government Effectiveness	Efficacité de l'Etat	Basnet (2017)
Basnet (2017)	Political Stability and absence of violence	Stabilité politique et absence de violence	Basnet (2017)
Basnet (2017)	Regularity Quality	Qualité de la règlementation, Qualité de régulation	Basnet (2017)
Basnet (2017)	Rule of Law	Etat de droit	Basnet (2017)
Basnet (2017)	Voice and Accountability Estimates	Participation citoyenne et responsabilité publique	Basnet (2017)
Basnet (2017)	Political Rights	Droits politiques	Basnet (2017)
Basnet (2017)	Civil Liberty	Liberté civile	Basnet (2017)
Basnet (2017)	Property Rights	Droits de propriété	Basnet (2017)
Basnet (2017)	Freedom from Corruption	Liberté face à la corruption	Basnet (2017)
Basnet (2017)	Fiscal Freedom	Liberté fiscale	Basnet (2017)
Basnet (2017)	Government Spending	Dépenses publiques (ou de l'Etat)	Basnet (2017)
Basnet (2017)	Business Freedom	Liberté d'entreprise	Basnet (2017)
Basnet (2017)	Monetary Freedom	Liberté monétaire	Basnet (2017)
Basnet (2017)	Trade Freedom	Liberté du commerce	Basnet (2017)
Basnet (2017)	Investment Freedom	Liberté d'investissement	Basnet (2017)
Basnet (2017)	Financial Freedom	Liberté financière	Basnet (2017)
Das, Quirk (2016)	Capital Account Openness	Ouverture du compte de capital	Das et Quirk (2016)
Das, Quirk (2016)	Corruption	Corruption	Das et Quirk (2016)
Das, Quirk (2016)	Democracy	Démocratie	Das et Quirk (2016)
Das, Quirk (2016)	FDI (% GDP)	IDE (% PIB)	Das et Quirk (2016)

Article	Mesure des institutions (intitulé dans l'article)	Mesure des institutions (traduction en français)	Cité dans …
Das, Quirk (2016)	Government Consumption (% GDP)	Consommation de l'administration publique (ou de l'Etat) en % du PIB	Das et Quirk (2016)
Das, Quirk (2016)	Human Rights	Droits de l'homme	Das et Quirk (2016)
Das, Quirk (2016)	Income	Revenu, Revenu par habitant	Das et Quirk (2016)
Das, Quirk (2016)	Inflation	Inflation	Das et Quirk (2016)
Das, Quirk (2016)	Inflation Volatility	Volatilité de l'inflation	Das et Quirk (2016)
Das, Quirk (2016)	Investment (% GDP)	Investissement (% du PIB)	Das et Quirk (2016)
Das, Quirk (2016)	Investment Profile	Profil d'investissement	Das et Quirk (2016)
Das, Quirk (2016)	Law and Order	Ordre public	Das et Quirk (2016)
Das, Quirk (2016)	Liquid Liabilities	Passifs exigibles	Das et Quirk (2016)
Das, Quirk (2016)	Political Rights	Droits politiques	Das et Quirk (2016)
Das, Quirk (2016)	Population Growth	Croissance démographique	Das et Quirk (2016)
Das, Quirk (2016)	Private Credit (% GDP)	Crédits privés (% PIB)	Das et Quirk (2016)
Das, Quirk (2016)	Regulation	Régulation, règlementation	Das et Quirk (2016)
Das, Quirk (2016)	Risk Budget Balance	Risque de deficit budgétaire	Das et Quirk (2016)
Das, Quirk (2016)	Schooling	Scolarisation, nombre d'années de scolarisation	Das et Quirk (2016)
Das, Quirk (2016)	Trade (% GDP)	Commerce (% PIB)	Das et Quirk (2016)
Hartmann et al. (2015)	Corruption Control	Maîtrise de la corruption, Lutte contre la corruption	Hartmann, Guevara, Jara-Figueroa, Aristarán,Hidalgo (2015)
Hartmann et al. (2015)	Government Effectiveness	Efficacité de l'Etat	Hartmann, Guevara, Jara-Figueroa, Aristarán1,Hidalgo (2015)
Hartmann et al. (2015)	Political Stability	Stabilité politique	Hartmann, Guevara, Jara-Figueroa, Aristarán1,Hidalgo (2015)
Hartmann et al. (2015)	Regulatory Quality	Qualité de la règlementation	Hartmann, Guevara, Jara-Figueroa, Aristarán1,Hidalgo (2015)
Hartmann et al. (2015)	Voice and accountability	Participation citoyenne et responsabilité de l'administration publique	Hartmann, Guevara, Jara-Figueroa, Aristarán1,Hidalgo (2015)
Eicher et al. (2006)	Quality of institutions	Qualité des institutions	Efendic, Pugh (2007)
Eicher et al. (2006)	Rule of law and protection of property rights	Etat de droit et la protection des droits de propriété	Efendic, Pugh (2007)
Eicher et al. (2006)	Integration = Trade/nominal GDP	Intégration (ouverture) au commerce international	Efendic, Pugh (2007)
Chousa et al. (2005a)	Institutional efficiency	Efficience des institutions	Efendic, Pugh (2007)
Chousa et al. (2005a)	Years under Communism (to reflect the initial conditions of the reform process)	Années sous régime communiste (pour refléter les conditions initiales du processus de réforme)	Efendic, Pugh (2007)
Chousa et al. (2005a)	Democratization index	Indice de démocratisation	Efendic, Pugh (2007)

Article	Mesure des institutions (intitulé dans l'article)	Mesure des institutions (traduction en français)	Cité dans …
Chousa et al. (2005a)	FDI inflow	IDE flux entrant	Efendic, Pugh (2007)
Redek et al. (2005)	Index of Institutional quality (obtained from the Heritage Foundation database. This index measures overall institutional quality by evaluating ten aspects of institutions: trade policy; fiscal burden of government; government intervention in the economy; monetary policy; capital flows and foreign investment; banking and finance; wages and prices; property rights; regulation; and informal market activity).	Indice de qualité des institutions (obtenu à partir de la base de données Heritage Foundation. Cet indice mesure la qualité institutionnelle globale en évaluant dix aspects des institutions: la politique commerciale; prélèvements obligatoires ; intervention du gouvernement dans l'économie; politique monétaire; flux de capitaux et investissements étrangers; banque et finance; salaires et prix; droits de propriété; réglementation; et activité sur le marché informel).	Efendic, Pugh (2007)
Glaeser et al. (2004)	Executive constraints	Contraintes sur le pouvoir exécutif	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Expropriation risk	Risque d'expropriation	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Autocracy	Autocratie	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Government effectiveness	Efficacité de l'Etat	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Judicial independence	Indépendance de la justice	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Constitutional review	Réforme constitutionnelle	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Plurality	Pluralisme	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	Proportional representation	Représentation proportionnelle	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Glaeser et al. (2004)	French legal origin	Législation d'origine française	Glaeser E., LaPorta R., Lopes de-Silanes F., Shleifer A. (2004)
Assane et al. (2003)	Economic freedom (the socioeconomic liberty index).	Liberté économique (indice de liberté socio-économique).	Efendic, Pugh (2007)
Assane et al. (2003)	Institutional efficiency (unweighted average of nine indicators of government performance : the higher the value of institutional efficiency, the higher the "quality" of institutions)	Efficience des institutions (moyenne non pondérée de neuf indicateurs de la performance étatiques : plus la valeur de l'efficacité institutionnelle est élevée, plus la «qualité» des institutions est élevée).	Efendic, Pugh (2007)
Assane et al. (2003)	Institutional quality (based on political risk ratings published in the International Country Risk Guide. Political risk is an unweighted average of 13 indicators of government performance).	Qualité des institutions (basée sur les évaluations politiques publiée par International Country Risk Guide. Le risque politique est une moyenne non pondérée de 13 indicateurs de performance étatique)	Efendic, Pugh (2007)

Notons que les mesures des institutions peuvent être directes (exemple : indice de démocratie) ou indirectes (proxy) à travers des indicateurs corrélés avec les effets des institutions (indicateurs de résultats).

En ce qui concerne la typologie des institutions économiques, Rodrik (1999) distingue les catégories suivantes : Institutions de droits de propriété, institutions de régulation, institutions pour la stabilisation macroéconomique, institutions d'assurances sociales, et institutions de gestion des conflits.
Das et Quirk (2016), quant à eux, proposent la typologie ci-après : Institutions créatrices de marché (exemple : ordre public), institutions de légitimation de marché (exemple : démocratie), institutions de stabilisation du marché (indicateur de résultat : inflation), institutions de régulation du marché (exemple : règlementation bancaire).

6.3.2. Mesures originales de l'inclusivité économique

Dans la revue des mesures effectuée, nous constatons une carence de mesures de l'inclusivité économique. Aussi, proposons-nous quelques pistes.
Nous partons de Flaux (2019), qui introduit un indice visant à distinguer les « des économies les plus prédatrices », des « plus redistributrices », en classant 188 pays « selon le niveau d'inclusion de leur économie, en confrontant leur niveau de richesse par habitant avec leur indice développement humain ».

Nous observons que l'indice de Flaux (noté ci-dessous III.GAP), ne mesure pas l'inclusivité, mais plutôt la non-inclusivité.

$III.GAP = Country.Rank(GDP\ per\ Capita) - Country.Rank(HDI) = RG - RH$
Avec :
$HDI: Human\ Developement\ Index$
$RG = Country.Rank(GDP\ per\ Capita) = Rang\ des\ pays\ selon\ le\ GDP\ per\ capita$
$RH = Country.Rank\ (HDI) = Rank\ des\ pays\ selon\ le\ HDI$
$n = Nombre\ de\ pays$
Nous utilisons *III* comme sigle de *Institutional Inclusivity Index.*

En effet, l'auteur semble interpréter les valeurs positives de son indice comme un signe d'inclusivité.
Mais, si les valeurs négatives indiquent un classement humain faible par rapport au classement économique (problème d'inégalité et de prédation économique et de non-redistribution), les valeurs positives élevées du *III.GAP* indiquent aussi un problème tout aussi grave : un classement économique indigne du classement humain, donc clairement un problème de liberté et d'accès aux opportunités économiques, alors même qu'un Etat autoritaire assure une éducation et un système sanitaire d'une certaine efficacité.

Pour mesurer l'inclusion, nous proposons des mesures alternatives qui seraient plus appropriée pour mesurer l'inclusivité et la non-inclusivité en distinguant les 2 profils de non-inclusivité.
Nous observons que : $1 \leq RG \leq n\ \ et\ 1 \leq RH \leq n\ \Longrightarrow 1 - n \leq RG - RH \leq n - 1$
C'est-à-dire $1 - n \leq III.GAP \leq n - 1$
Les valeurs de cet indice dépendent de *n* : Cette propriété mathématique n'est pas adaptée à une situation où le nombre de pays de l'étude peut varier d'une année à l'autre.

On a : $0 \leq |RG - RH| \leq n - 1, n > 1$
D'où : $1 \leq n - |RG - RH| \leq n$

Nous définissons la mesure d'inclusivité relative suivante :
$III.1 = n - |III.GAP|, avec\ 1 \leq III.1 \leq n$

L'indice *III.1* n'est pas satisfaisant car sa borne supérieure dépend de *n*, qui peut varier selon les années.

Pour avoir un indice compris entre 0 et 1, nous définissons la mesure d'inclusivité relative suivante :
$$III.2 = 1 - \frac{|III.GAP|}{n} = \frac{III.1}{n}, avec\ \ \frac{1}{n} \leq III.2 \leq 1$$

Nous définissons la mesure de non-inclusivité relative associée suivante :
$$III.GAP.2 = III.GAP/n, avec\ \ \frac{1-n}{n} \leq III.GAP.2 \leq \frac{n-1}{n}$$

Ces 2 derniers indices peuvent être améliorés par rapport à leurs bornes atteignables.

On a :

$$1 \leq RG \leq n \implies 0 \leq RG - 1 \leq n - 1$$

$$\implies 0 \leq \frac{RG - 1}{n - 1} \leq 1$$

De même, on peut écrire :

$$0 \leq \frac{RH - 1}{n - 1} \leq 1 \implies -1 \leq \frac{-RH + 1}{n - 1} \leq 0$$

D'où :

$$-1 \leq \frac{RG - 1 - RH + 1}{n - 1} \leq 1$$

$$\implies -1 \leq \frac{RG - RH}{n - 1} \leq 1$$

Nous définissons la mesure de non-inclusivité relative suivante :

$$III.GAP.3 = \frac{RG - RH}{n - 1}, et \ -1 \leq III.GAP.3 \leq 1$$

On a $0 \leq |III.GAP.3| \leq 1$

$\implies -1 \leq -|III.GAP.3| \leq 0$

$\implies 0 \leq 1 - |III.GAP.3| \leq 1$

Nous définissons la mesure d'inclusivité relative suivante :

$$III.3 = 1 - |III.GAP.3|, \text{avec } 0 \leq III.3 \leq 1$$

En utilisant cette dernière formule, nos estimations sont représentées par la cartographie suivante.

Figure 1.37 Indice d'inclusivité des institutions économiques, 2015

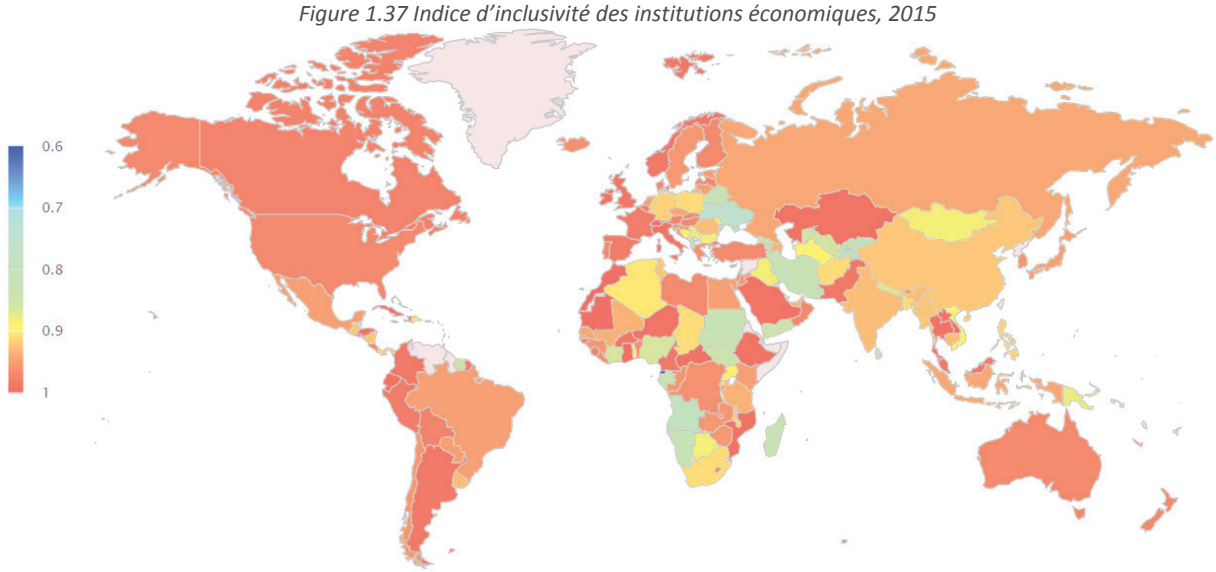
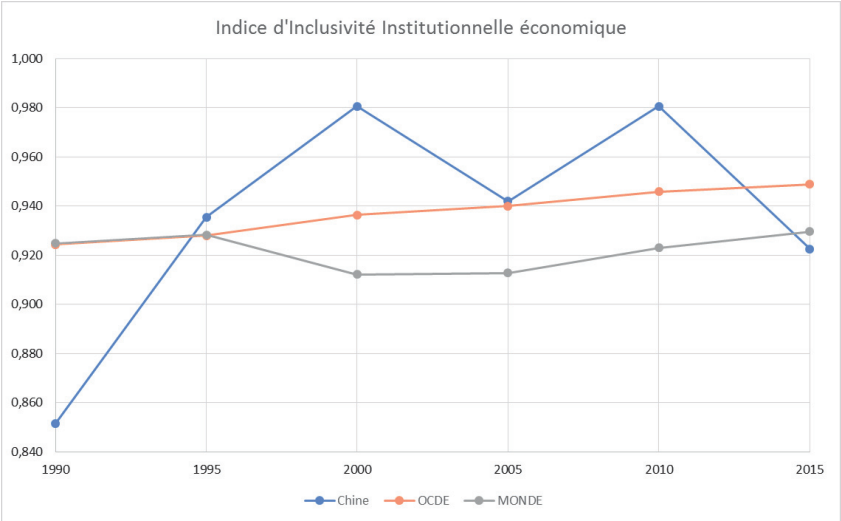


Tableau 1.13 Indice d'inclusivité des institutions économiques, Chine vs moyenne OCDE et moyenne Monde

Année	Chine	OCDE	MONDE
2015	0,923	0,949	0,930
2010	0,981	0,946	0,923
2005	0,942	0,940	0,913
2000	0,981	0,936	0,912
1995	0,935	0,928	0,928
1990	0,852	0,924	0,925

Figure 1.38 Indice d'inclusivité des institutions économiques, Chine vs moyenne OCDE et moyenne Monde



D'après nos calculs, nous constatons que contrairement à l'hypothèse avancée par Acemoglu et Robinson (2015, p.130) en se basant sur les institutions politiques, la Chine a été plus inclusive économiquement que la moyenne de l'OCDE et la moyenne mondiale de 1995 à 2010.
Toutefois l'inclusivité économique chinoise a chutée depuis lors, mais reste relativement proche de la moyenne mondiale.
Et donc que par conséquent, la croissance économique chinoise n'est pas éphémère et non-durable, comme anticipée par Acemoglu et Robinson (2015, p.130), qui s'inspiraient du précédent soviétique, basé sur des institutions politiques extractives et qui s'est soldé par un effondrement politique et économique.
Nos analyses nous indiquent plutôt une soutenabilité de la croissance économique Chinoise, qui ne saurait être réduit à un phénomène non-durable qui conduirait à un effondrement économique à la soviétique.

6.3.3. Originalité de notre approche systémique de la mesure des institutions

Matrice IEPE (Interne, Externe, Politique, Economique) de mesure des institutions

L'originalité de notre approche consiste à caractériser les institutions politiques par des indicateurs du double point de vue des deux souverainetés issues des institutions des Lumières (dès la Glorieuse Révolution de 1689) : la souveraineté du peuple et les droits « naturels » de l'individu.
L'affirmation de cette double souveraineté, renvoie aux deux libérations (externe et interne) de Rostow (1997) : L'effectivité de la souveraineté de peuple libère du joug des puissances étrangères, et la conquête des droits individuels commence par les avancées de la démocratie et de l'Etat de droit.

Nous-nous proposons donc de caractériser les institutions via : d'une part, des indicateurs « externes » liés à la souveraineté nationale (comme le solde commercial avec les puissances « dominantes », les ratios de dépenses militaires (par rapport au PIB), les déclarations d'indépendance politique), et d'autre part, des indicateurs « internes » de gouvernance et de respect des droits constitutionnels de l'individu.

Nous regroupons les indicateurs de gouvernance et de respect des droits constitutionnels de l'individu en trois catégories : d'abord, le respect de l'Etat de droit (hiérarchie des normes, égalité juridique, indépendance judiciaire), ensuite, le respect des règles démocratiques (indices de démocratie), et enfin, le respect des droits fondamentaux individuels (la liberté, la propriété, la sûreté, et la résistance à l'oppression en référence à la constitution française de 1789).

Ensuite, à chacun de ces niveaux (externe et interne à la Rostow), nous distinguerons deux sous-catégories d'institutions : Les institutions politiques, puis les institutions économiques.

Enfin, en ce qui concerne les institutions internes, nous veillerons à distinguer les institutions relevant de l'Etat de droit (comme le respect du droit de propriété), de celles relevant de la démocratie.

Tableau 1.14 Matrice IEPE (Interne, Externe, Politique, Economique) de mesure des institutions

	Politique	**Economique**
Interne	- Droits et libertés politiques internes - Régime politique - Droits civiques	- Inclusivité économique - Dépenses publiques (consommation, investissement, redistribution, autres) - Prélèvements obligatoires - Production manufacturière (comme indicateur de compétitivité des pays) - Indépendance de la banque centrale par rapport au gouvernement
Externe	- Régime politique (c'est-à-dire régime colonial ou non) - Poids économique (PIB du pays, en % du PIB mondial) - Poids militaire (Dépenses militaires du pays, en % des dépenses militaires mondiales)	- Ouverture commerciale - Commerce de biens matériels - Commerce de Services - Balance commerciale et tension commerciale - Taxes sur commerce international

Nous listons ci-après quelques exemples d'indicateurs statistiques utilisables pour mesurer les items de la matrice IEPE de mesure des institutions.

Institutions politiques internes

- Droits et libertés politiques internes :
  - Indice global de démocratie
- Régime politique
  - Régime politique
- Droits civiques
  - Niveau de libertés civiles

Institutions économiques internes

- Inclusivité économique
  - Indices d'inclusivité
  - Indices de non-inclusivité
  - Indice de Gini avant taxes et redistributions
  - Indice de Gini après taxes et redistributions
- Dépenses publiques : Redistribution
  - Indice de Gini après redistribution
  - Ratio : Indice de Gini après taxes et redistributions / Indice de Gini avant taxes et redistribution
  - Taux transferts et subsides dans le PIB
- Dépenses publiques : Autres
  - Dépenses publiques (% PIB)
  - Dépenses de consommation finale des administrations publiques (% PIB)
  - Part des investissements publics par rapport à l'ensemble des investissements
  - Rémunération des employés (% des dépenses)
  - Paiements d'intérêts (% des dépenses)
- Prélèvements obligatoires
  - Taux de prélèvements obligatoires [*total Tax Revenues, % of GDP*]
  - Degré de liberté fiscale
  - Taxes sur les biens et services (% des recettes)
  - Impôts sur le revenu, les bénéfices et les revenus du capital (% des recettes)
  - Taux d'imposition total (% des bénéfices commerciaux)
  - Impôt sur les bénéfices (% des bénéfices commerciaux), *[profit tax, % of commercial profits]*

    - Impôts et cotisations sur le travail (% des bénéfices commerciaux), *[labor tax and contributions (% of commercial profits)]*
- Production manufacturière (comme indicateur de compétitivité des pays)
    - Valeur ajoutée du secteur manufacturier (% PIB)
    - Emplois dans le secteur manufacturier (% Emplois)
    - Indice de complexité des produits exportés *[Exported Products Complexity index (Economic Complexity Index)]*
- Indépendance de la banque centrale par rapport au gouvernement
    - Degré de liberté monétaire

Institutions politiques externes

- Régime colonial (ou non)
- Poids économique (PIB du pays, en % du PIB mondial)
- Poids militaire (Dépenses militaires du pays, en % des dépenses militaires mondiales)
- Dépenses militaires (% du PIB)
- Dépenses militaires par habitant [Dépenses militaires (% du PIB)] x [GDP per capita (constant USD 2010)]
- Dépenses militaires (% des dépenses du gouvernement central), WDI
- Personnel militaire (% de la main-d'œuvre totale)
- Niveau d'exportation des armes conventionnelles pour 100 000 habitants (Global Peace Index)
- Niveau de capacité et de sophistication militaire (Global Peace Index)
- Niveau de paix extérieure (Global Peace Index)

Institutions économique externes

- Ouverture commerciale
    - Ouverture commerciale (Importations et Exportations, en % PIB))
    - Exportations de biens et services (% PIB)
    - Importations de biens et services (% PIB)

- Commerce de biens matériels
    - Indice de complexité des produits exportés *[Exported Products Complexity index (Economic Complexity Index)]*
    - Exportations de haute technologie (% des exportations de produits manufacturés). *[High-technology exports (% of manufactured exports)].*
    - Ouverture commerciale en biens (Importations et Exportations de biens, en % PIB)

- Commerce de services
    - Ouverture commerciale en services (Importations et Exportations de services, en % PIB). *[Trade in services (% of GDP)].*
    - Assurances et services financiers (% des exportations de services commerciaux). *[Insurance and financial services (% of commercial service exports)]*
    - Assurances et services financiers (% des importations de services commerciaux). *[Insurance and financial services (% of commercial service imports)]*
    - Tourisme international, recettes (% des exportations totales*). [International tourism, receipts (% of total exports)]*
    - Services de transport (% des exportations de services commerciaux). *[Transport services (% of commercial service exports)]*

- Balance commerciale et tension commerciale
    - Balance commerciale. *[Trade balance (% of GDP)]*
    - Tension commerciale sur 10 ans (cumul du solde commercial en % PIB). Variable originale.

- Tension commerciale sur 20 ans (cumul du solde commercial en % PIB). Variable originale.

- Taxes
  - Taxes sur le commerce international (% des recettes). *[Taxes on international trade (% of revenue)]*
  - Taxes sur les exportations (% des recettes fiscales). *[Taxes on exports (% of tax revenue)]*
  - Taxes sur les importations (% des recettes fiscales
  - Taux tarifaire appliqué, moyenne pondérée, tous produits (%). *[Tariff rate, applied, weighted mean, all products (%)]*
  - Taux tarifaire appliqué, moyenne pondérée, produits manufacturés (%). *[Tariff rate, applied, weighted mean, manufactured products (%)]*
  - Taux tarifaire, appliqué, moyenne pondérée, matières premières (%). *[Tariff rate, applied, weighted mean, primary products (%)]*

7. Cognition, éducation et capital humain

7.1. Définitions

Cognition
Selon le dictionnaire de l'*American Psychological Association* (2019), la cognition désigne « toutes les fonctions de connaissance et de conscience, telles que la perception, la conception, la mémorisation, le raisonnement, le jugement, l'imagination et la résolution des problèmes ».

Par extension, les sciences cognitives rassemblent l'ensemble des domaines scientifiques qui étudient la cognition, humaine ou non-humaine, telles que : la psychologie, les neurosciences, la linguistique, l'intelligence artificielle, les mathématiques appliquées.

La cognition est considérée classiquement (*ibid*.) comme l'un des trois composants psychologiques des processus mentaux, dont les deux autres sont l'affect (sentiments ou émotions) et la conation (motivations proactives).

Compétences cognitives
Les compétences cognitives désignent les capacités opérationnelles relavant de la cognition, et liées à un domaine (ou discipline) particulier, comme par exemple les mathématiques, les sciences physiques, la linguistique.

Education
L'éducation, notamment scolaire, permet aux élèves d'acquérir des connaissances et des compétences cognitives dans de nombreux domaines, telles que le langage, la lecture, la rédaction, les sciences, l'art.

Capital humain
Becker (2002) définit le capital humain comme un stock de « connaissances, informations, idées, compétences et santé des individus ». L'investissement en capital humain inclut « la scolarisation, la formation continue, les soins médicaux, la consommation de vitamines et l'acquisition d'informations sur le système économique » (Becker, 1962).

Goldin (2016), définit le capital humain comme « le stock de compétences que possède la main-d'œuvre », avec l'idée « qu'il existe des investissements dans les personnes (éducation, formation, santé, par exemple) et que ces investissements augmentent la productivité d'un individu ».

Ainsi défini, le concept de capital humain est large et inclut « la santé, les capacités cognitives, les connaissance, les habiletés physiques et même des traits de comportement » (Hearns, Baten et Crayen, 2009).

7.2. Compétences cognitives et performances économiques

7.2.1. Rendement micro-économique du capital humain

Au niveau micro-économique, le capital humain a été essentiellement mesuré par le nombre d'années de scolarisation.
Les rendements micro-économiques (i.e. rémunérations) de l'éducation sont abondamment étudiés à partir des analyses sur le capital humain de Mincer (1970, 1974) : ces études montrent une corrélation positive entre la rémunération et le nombre d'années d'études. La relation entre le niveau d'éducation et les inégalités de revenu est donc empiriquement établie, même si « les compétences cognitives en elles-mêmes ne déterminent évidemment pas la répartition complète, car d'autres facteurs tels que les institutions du marché du travail, les taxes» sont à prendre en compte (Hanushek et Woessmann, 2008, p.627)

Becker (1974) développe également un modèle théorique illustrant le *trade-off* entre durée d'éducation (investissement dans l'éducation) et salaire que l'individu pourrait tirer d'un emploi immédiat.

Si les modèles de rémunération de type « Mincer » mesurent le rendement individuel de la scolarisation, de nombreuses externalités positives pourraient intervenir pour l'estimation du rendement social de la scolarisation.
A cet effet, Hanushek et Woessmann (2008, p.615), évoquent des effets positifs « de l'éducation sur la criminalité, la santé, la fécondité, une participation accrue des citoyens » et « sur la croissance et la productivité de l'économie dans son ensemble ». Ils citent à cet effet de nombreuses études empiriques des années 2000, qui concluent dans ce sens.

Aghion et Cohen (2004, p.16), distinguent quant à eux, trois types d'externalités non-prises en compte dans les rendements micro-économiques de l'éducation : les externalités intra-générationnelles, les externalités intergénérationnelles (apport des parents), les externalités sur le progrès technique.

7.2.2. Approche macro-économique du capital humain

Du point de vue de la théorie économique l'intégration du capital humain dans les modèles macroéconomiques a véritablement démarré par la théorie de la croissance endogène développée par Romer (1986) et Lucas (1988) : Le capital humain y joue un rôle crucial en rendant endogène le progrès technique et l'innovation.
Aghion et Cohen (2004) mentionnent que l'hypothèse d'un lien positif entre la productivité (représenté par le PIB par tête) et le capital humain (mesuré par le taux de scolarisation des 12-17 ans) est empiriquement vérifiée, ce qui a été trouvé dans un premier temps par Mankiw, Romer et Weil (1992).
De Aghion et Cohen (2004), nous tirons les points théoriques essentiels suivants : l'innovation et le progrès technique sont les sources des gains de productivité et donc de la croissance à long terme. Le stock de capital humain détermine en partie la capacité à l'innovation et au progrès technique d'un pays ou d'un secteur économique. La R&D et l'innovation se font majoritairement dans les entreprises. Sans institutions « appropriées » le potentiel du capital humain n'est pas pleinement réalisé.

Sur un plan empirique, mesurer l'effet macro-économique du capital humain, a d'abord consisté à adapter le modèle « Mincer » en remplaçant la rémunération individuelle par le PIB par habitant ou la croissance du PIB par habitant. Mais encore une fois, les compétences humaines, sont mesurées essentiellement par le niveau d'éducation scolaire, du moins dans les premiers temps (Hanushek et Woessmann, 2008, p.614).

Toutefois, les compétences globales d'un individu sont multifactorielles et varieront au sein d'un pays, en fonction « de l'investissement familial, de la qualité de l'école, des différences de capacités et d'autres facteurs spécifiques au pays » (*ibid*.).

Il faut également rappeler, que dans la littérature psychologique, les compétences cognitives ne sont pas les seules compétences mentales prises en considération, puisque les domaines de l'affect (sentiments ou émotions) et de la conation (motivations) peuvent être également mobilisés (American Psychological Association [2019] et Hanushek et Woessmann [2008]).
Par ailleurs, les différences qualitatives entre systèmes scolaires au niveau international, entament la pertinence du seul nombre d'année d'études comme mesure des compétences cognitives.
Se sont donc développés depuis les années 60, de nombreux tests internationaux (dont certains s'étalent sur plusieurs années) pour tenter d'établir des mesures universelles des compétences cognitives des élèves adolescents (de 9 à 15 ans) : Hanushek et Woessmann (2008, p.661) en recensent 12, dont TIMSS (*Trends in International Mathematics and Science Study*) et PISA (*Programme for International Student Assessment*).
Ces tests tentent (avec leurs limites et biais) de mesurer une ou plusieurs compétences cognitives en ce qui concerne la lecture, les mathématiques et les sciences.
En 2018, l'étude PISA organisée par l'OCDE, s'étend désormais à 80 pays (*OECD Programme for International Student Assessment*, 2019).
Pour ce qui concerne l'évaluation des compétences cognitives des adultes, nous pouvons citer l'étude internationale PIAAC (*Programme for the International Assessment of Adult Competencies*) : Ce programme concerne 24 pays et a débuté en 2011. Les domaines cognitifs couverts sont les lettres (*literacy*), les mathématiques (*numeracy*) et la résolution de problèmes dans un environnement technique (*problem solving in technology-rich environments*).

La pertinence empirique de ces mesures internationales des compétences cognitives, est solidement établie par Hanushek et Woessmann (2008) qui en ont montré la corrélation positive et significative avec la croissance économique, après neutralisation des conditions économiques initiales (PIB).

Du point de vue de l'histoire longue, il faut également noter l'étude cliométrique de référence de Hearns, Baten et Crayen (2009) pour mesurer les compétences calculatoires (si ce n'est mathématique) dans les pays Européens depuis le 14ème siècle., ainsi qu'aux Etats-Unis depuis le 19ème siècle. Cette recherche utilise comme proxy des compétences mathématiques (*numeracy*), une mesure inverse du taux de déclaration d'âge sous forme de nombres arrondis se terminant par 0 ou 5 (phénomène qualifié de « *age-heaping* », signe souvent avéré d'une incapacité à donner un nombre exact, en cliométrie, comme en démographie).
Le phénomène de« *age-heaping* » étant illustratif des lacunes numériques, il faut prendre l'inverse de son taux pour mesurer le niveau de compétence mathématique.

Pour notre part, nous limitons l'analyse du capital humain, à sa seule dimension cognitive (la dimension sanitaire est incluse dans la partie consacrée à la transition démographique, via l'analyse du taux de mortalité) et discutons dans la section suivante des mesures de ces compétences cognitives.

7.3. Mesures des compétences cognitives et du niveau d'éducation

Une large revue bibliographique nous a permis de recenser les mesures des compétences cognitives (dont la plupart sont déjà disponibles dans notre base de données).

Tableau 1.15 Sigles liés aux mesures des compétences cognitives et du niveau d'éducation

Sigle	Définition
EIAA	Enquête Internationale sur l'Alphabétisation des Adultes
ISATs	International Student Achievement Tests
LLECE	Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación
NAEP	National Assessment of Educational Progress
PASEC	Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (Conférence des Ministres de l'Education des Etats et Gouvernements de la Francophonie)
PIAAC	Programme for the International Assessment of Adult Competencies. Programme pour l'évaluation internationale des compétences des adultes (Population en age de travailler, 15-64 ans)
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment. Programme international pour le suivi des acquis des élèves (de 15 ans)
RSATs	Regional Student Achievement Tests
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
TIMSS	Trends in International Mathematics and Science Study. Elèves de 9-10 ans (4th grade), et 13-14 ans (8th grade)

Tableau 1.16 Revue des mesures des compétences cognitives et du niveau d'éducation

Année	Article	Compétences cognitives	Mesure de cognition (intitulé dans l'article)	Mesure de cognition (Traduction en français)
2019	Doyle, Meier zu Selhausen, Weisdorf (2019)	Numeracy	Whipple's Index: "ratio of the observed frequency of ages ending in 0 or 5 to the frequency predicted by assuming a uniform distribution of terminal digits".	Whipple Index : « rapport entre la fréquence observée des âges se terminant par 0 ou 5 et la fréquence prédite en supposant une distribution uniforme des chiffres terminaux".
2019	Doyle, Meier zu Selhausen, Weisdorf (2019)	Numeracy	ABCC Index: "a linear transformation of the Whipple's Index, represents the share of the population able to report their own age exactly, ranging between 0 and 100"	ABCC Index : « une transformation linéaire de l'indice de Whipple représentant la proportion de la population capable de déclarer son propre âge avec exactitude, comprise entre 0 et 100 »
2017	Altinok et Aydemir (2017)	Cognitive skills (Literacy, Mathematics, Science)	ISATs tests: TIMSS, PIRLS and PISA tests.	Les tests ISATs : tests TIMSS, PIRLS et PISA.
2017	Altinok et Aydemir (2017)	Cognitive skills (Literacy, Mathematics, Science)	RSATs tests: conducted in Africa and Latin America, such as LLECE, SACMEQ or PASEC tests.	Les tests RSATs : réalisés en Afrique et en Amérique latine, tels que les tests LLECE, SACMEQ ou PASEC.
2017	Altinok et Aydemir (2017)	Education	Years of schooling	Années de scolarité
2017	Hanushek, Ruhose and Woessmann (2017)	Education	School attainment (years of schooling) for the working-age population of each state (population aged 20–65 not currently in school)	Niveau scolaire (années d'études) pour la population en âge de travailler de chaque État (population âgée de 20 à 65 ans non scolarisée)
2017	Hanushek, Ruhose and Woessmann (2017)	Cognitive skills (Literacy, Mathematics, Science)	Aggregate tests TIMSS, PISA, PIRLS scores.	Les scores agrégés aux tests TIMSS, PISA, PIRLS.
2017	Hanushek, Ruhose and Woessmann (2017)	Cognitive skills (Literacy, Mathematics, Science)	Aggregate NAEP tests scores (USA)	Scores agrégés aux tests NAEP (USA)
2017	Hanushek, Woessmann (2017)	Cognitive skills (Literacy, Mathematics, Science)	PISA test, testing math, science, and reading performance of 15-year-olds on a three-year cycle since 2000	Le test PISA évalue les performances en mathématiques, en sciences et en lecture des élèves de 15 ans selon un cycle de trois ans depuis 2000

Année	Article	Compétences cognitives	Mesure de cognition (intitulé dans l'article)	Mesure de cognition (Traduction en français)
2017	Hanushek, Woessmann (2017)	Cognitive skills (Mathematics, Science)	TIMSS test, testing math and science performance (mostly) of eighth-graders on a four-year cycle since 1995	Le test TIMSS évalue principalement les performances mathématiques et scientifiques des élèves de huitième année selon un cycle de quatre ans depuis 1995
2017	Hanushek, Woessmann (2017)	Cognitive skills (Literacy)	PIRLS test, testing primary-school reading performance on a five-year cycle since 2001	Le test PIRLS évalue les performances de lecture dans les écoles primaires tous les cinq ans depuis 2001.
2017	Hanushek, Woessmann (2017)	Education	Cumulative educational expenditure per student (poorly ou not at all correlated with tests scores)	Dépenses d'éducation cumulées par élève (peu ou pas corrélées avec les Scores aux tests)
2017	Hanushek, Woessmann (2017)	Education	Class size (correlated with tests scores)	Taille de la classe (en corrélation avec les Score aux tests)
2016	Hanushek (2016)	Cognitive skills (Mathematics, Science)	International assessments of mathematics and science, as direct measures of cognitive skills	Évaluations internationales des mathématiques et des sciences, en tant que mesures directes des compétences cognitives
2016	Hanushek (2016)	Education	Years of schooling	Années de scolarité
2016	Hanushek (2016)	Cognitive skills (Literacy, Mathematics, Science)	PISA test reading scores	Scores de lecture au test PISA
2016	Hanushek (2016)	Education	Higher (tertiary) education attainment percentages for the 25–34-year-old population	Pourcentages de niveaux d'études supérieures (tertiaires) pour la population âgée de 25 à 34 ans
2016	Teixeira et Queirós (2016)	Education	Number of years of schooling of the population aged 25 or more (in log)	Nombre d'années de scolarité de la population âgée de 25 ans ou plus (en log)
2015	Hafer, Jones (2015)	Cognitive skills (IQ)	Intelligence quotient IQ, used as the equivalent of TIMSS ans PISA tests scores	Le quotient intellectuel QI, utilisé comme l'équivalent des scores aux tests TIMSS et PISA
2015	Hafer, Jones (2015)	Cognitive skills (Mathematics, Science)	TIMSS test scores	Scores au test TIMSS
2015	Hafer, Jones (2015)	Cognitive skills (Literacy, Mathematics, Science)	PISA test scores	Scores au test PISA
2014	Artus P., García-Peñalosa C., Mohnen P., (2014), p. 8	Cognitive skills (Literacy, Mathematics, Science)	Scores au test PISA	Scores au test PISA (Elèves de 15 ans)
2014	Artus P., García-Peñalosa C., Mohnen P., (2014), p. 8	Cognitive skills (Literacy, Numeracy, Problem solving)	Scores au test PIAAC	Scores au test PIAAC (Population en âge de travailler, 15-64 ans)
2014	Cohen et Buigues (2014)	R&D	Dépenses en R&D (% PIB) : Total, public, privé	Dépenses en R&D (% PIB) : Total, public, privé
2014	Cohen et Buigues (2014)	R&D	Nombre de brevets par million d'habitants	Nombre de brevets par million d'habitants
2014	Cohen et Buigues (2014)	Innovation	Rang du pays en matière d'innovation parmi les pays de l'UE	Rang du pays en matière d'innovation parmi les pays de l'UE

Année	Article	Compétences cognitives	Mesure de cognition (intitulé dans l'article)	Mesure de cognition (Traduction en français)
2012	Acemoglu et Autor (2012)	Education	Primary+Secondary School, Post-Secondary School, high school dropout, some college, college graduate, greater than college	Etudes primaire et secondaire, études post-secondaires, abandon des études au secondaire, quelques années d'études universitaires, diplômé de l'université, études post-universitaires
2012	Acemoglu et Autor (2012)	Education	Years of schooling	Années de scolarité
2012	Castello-Climent et Domenech (2012)	Education	Average years of schooling for the population 15 years and above in primary, secondary and tertiary education	Nombre moyen d'années de scolarité de la population âgée de 15 ans et plus dans l'enseignement primaire, secondaire et tertiaire
2012	Castello-Climent et Domenech (2012)	Education	Relative supply of population with high education with respect to population with low secondary education	Offre relative de la population hautement scolarisée par rapport à la population peu scolarisée dans le secondaire
2009	Hearns, Baten et Crayen (2009), p. 787	Numeracy	The Whipple Index: "ratio of the observed frequency of ages ending in 0 or 5 to the frequency predicted by assuming a uniform distribution of terminal digits ”	Whipple Index : "rapport entre la fréquence observée des âges se terminant par 0 ou 5 et la fréquence prédite en supposant une distribution uniforme des chiffres terminaux"
2009	Hearns, Baten et Crayen (2009), p. 788	Numeracy		ABCC Index : "transformation de l'indice de Whipple pouvant être interprétée comme la proportion d'individus qui déclarent correctement leur âge"
2008	Goldin et Katz (2008)	Education	Common school	Ecole primaire
2008	Goldin et Katz (2008)	Education	Grammar school (i.e. "11-plus")	Lycée (c’est-à-dire 11 ans et plus)
2008	Goldin et Katz (2008)	Education	High school	Lycée
2008	Goldin et Katz (2008)	Education	College	Université
2008	Hanushek et Woessmann (2008)	Education	School attainment, average years of schooling (conditional on the initial level of output)	Niveau de scolarité, nombre moyen d'années de scolarité (en fonction du niveau de sortie initial)
2008	Hanushek et Woessmann (2008)	Work experience	Labor market experience	Expérience du marché de l'emploi
2004	Wilson et Briscoe (2004)	Education	Average years of schooling	Nombre moyen d'années de scolarité
2004	Wilson et Briscoe (2004)	Education	Percentage of population 15+ with no schooling	Pourcentage de la population de 15 ans et plus non scolarisée
2004	Wilson et Briscoe (2004)	Education	Years of upper school for males over 25	Années d'études supérieures pour les hommes de plus de 25 ans
2001	Barro (2001)	Education	Male attainment at the secondary and higher levels of education	Niveau de scolarisation des hommes aux niveaux secondaire et supérieur
2001	Barro (2001)	Education	Female attainment at the secondary and higher levels of education lacks significant explanatory power	Le niveau de scolarisation des filles aux niveaux secondaire et supérieur manque de pouvoir explicatif significatif
2001	Barro (2001)	Education	Primary schooling (male, female)	Enseignement primaire (masculin, féminin)
2001	Barro (2001)	Science	Student test scores in science	Scores des élèves aux tests en sciences

Année	Article	Compétences cognitives	Mesure de cognition (intitulé dans l'article)	Mesure de cognition (Traduction en français)
2001	Barro (2001)	Mathematics	Student test scores in mathematics	Scores des élèves aux tests en mathématiques
2001	Barro (2001)	Reading	Student test scores in reading	Scores des élèves aux tests en lecture
2001	Barro (2001)	Science, mathematics, and reading	Student overall test score: "The overall test score, (…) equals the science score, where available, and uses the reading score, adjusted for differences in average levels from the science scores, to fill in some additional observations"	Score global de l'élève : "Le score global au test, (…) est égal au score scientifique, si disponible, et utilise le score en lecture ajusté par rapport aux différences entre les niveaux moyens des scores scientifiques, pour compléter certaines observations"
2001	Barro (2001)	Education	Average years of schooling for the population 15 years and above in primary, secondary and tertiary education	Nombre moyen d'années de scolarité de la population âgée de 15 ans et plus dans l'enseignement primaire, secondaire et tertiaire
1998	Todd (1998)	Education	Proportion d'individus ayant obtenu une licence (Bac+3) ou Bachelor of arts (BA)	Proportion d'individus ayant obtenu une licence (Bac+3) ou Bachelor of arts (BA)
1998	Todd (1998)	Mathematics	Score au test TIMSS mathématique des élèves	Score au test TIMSS mathématique des élèves. Elèves de 9-10 ans (4th grade), et 13-14 ans (8th grade)
1998	Todd (1998)	Littératie, Literacy	Scores EIAA des adultes 16-65 ans : Compréhension de textes, de textes schématiques, de textes à contenu quantitatif	Scores EIAA des adultes 16-65 ans : Compréhension de textes, de textes schématiques, de textes à contenu quantitatif
1998	Todd (1998)	Science	Nombre de diplômés supérieurs en sciences (ingénieurs, biologistes, mathématiciens)	Nombre de diplômés supérieurs en sciences (ingénieurs, biologistes, mathématiciens)
1998	Todd (1998)	Education	Dépenses publiques pour l'éducation (% PIB)	Dépenses publiques pour l'éducation (% PIB)
1997	Bandyopadhyay (1997)	Education	Relative Proportion of Adults with at least Secondary Level Education	Proportion relative d'adultes ayant au moins un niveau d'enseignement secondaire
1997	Bandyopadhyay (1997)	Education	Percentage of working age population in secondary school, Percentage with no education, Percentage with some secondary level education and Percentage with higher than secondary level education of the adult people of age 25 and above	Pourcentage de la population en âge de travailler au secondaire, Pourcentage de la population en âge de travailler sans éducation, Pourcentage de la population en âge de travailler avec niveau d'instruction secondaire et Pourcentage des adultes de 25 ans et plus avec un niveau d'instruction supérieur au secondaire
1997	Mincer (1997)	Education	Schooling groups: less than 12 years, 12 years, 16 years, more than 16 years	Groupes scolaires : moins de 12 ans, 12 ans, 16 ans, plus de 16 ans
1996	Mincer (1996)	Education	Years of schooling	Années de scolarité
1996	Mincer (1996)	Education	Years of experience (experience levels are in 5-year intervals)	Années d'expérience (les niveaux d'expérience sont à intervalles de 5 ans)
1994	Becker (1994), p. 106	Education	Averge education	Education moyenne
1994	Becker (1994), p. 106	Education	Low education (1-7 years of school completed)	Faible niveau d'études (1 à 7 années d'études terminées)
1994	Becker (1994), p. 106	Education	Medium education (8-12 years of school completed)	Niveau d'éducation moyen (8-12 années d'études terminées)

Année	Article	Compétences cognitives	Mesure de cognition (intitulé dans l'article)	Mesure de cognition (Traduction en français)
1994	Becker (1994), p. 106	Education	High education (more than 12 years of school completed)	Études supérieures (plus de 12 années d'études terminées)
1994	Becker (1994), p. 172	Education	College dropout	Abandon des études au niveau universitaire
1994	Becker (1994), p. 176	Intelligence quotient	Intelligence quotient (under 120, over 120)	Quotient intellectuel (moins de 120, plus de 120)
1994	Becker (1994), p. 230	Education	More than 15 years of school completed	Plus de 15 années d'études terminées
1994	Becker (1994), p. 234	Education	Elementary school graduates	Diplômés du primaire
1994	Becker (1994), p. 234	Education	High school graduates	Diplômés du secondaire
1994	Becker (1994), p. 234	Education	College graduates	Diplômés universitaires
1991	Mincer (1991)	Education	Schooling groups: < 12 years, 12 years, 13-15 years, 16 years, 16+ years, <12 years to 16+ years ratio	Groupes scolaires : <12 ans, 12 ans, 13-15 ans, 16 ans, 16 ans et plus, ratio de <12 ans à 16 ans et plus
1991	Mincer (1991)	Education	Education: Grade School, High School, some college, College Degree, Graduate Work	Education : école primaire, secondaire et lycée, quelques années d'études universitaires, diplôme universitaire, études post-universitaires
1991	Mincer (1991)	Education	Years of schooling	Années de scolarité
1991	Mincer (1991)	Work experience	Experience (years)	Expérience (en années)
1991	Mincer (1991)	Job training	Job training: Training on the job, in years. Question: "On a job like yours, how long would it take the average new person to become fully trained and qualified?"	Formation professionnelle : formation sur le tas, en années. Question : "Dans un emploi comme le vôtre, combien de temps faudrait-il à une nouvelle personne pour acquérir une formation et une qualification complètes?"

Au vu la littérature et des données disponibles, nous baserons nos travaux développés dans la section suivante, sur les mesures de compétence cognitive ci-après :

Education comme mesure indirecte de compétences cognitives

- Niveau d'éducation atteint à l'age adulte (*Adult education attainment*) : Barro-Lee database (Lee et Lee [2016], Barro et Lee [2015]), World Bank Education Statistics database, World Bank Development Indicators database (utilisé aux sous-niveaux)
    - Analphabétisme (*No education*, *no schooling*)
    - Education primaire (*Primary education*)
    - Education secondaire (*Secondary education*)
        - Secondaire inférieure (*Lower secondary*)
        - Secondaire supérieure (*Upper secondary*)
    - Education tertiaire (*Tertiary education*)
        - Cycle court (*Short-cycle*)
        - Licence (*Bachelor*)
        - Master
        - Doctorat (*PhD*)

- Nombre moyen d'années de scolarisation (*Average years of schooling*) : Barro-Lee database (Lee et Lee [2016], Barro et Lee [2015]). Tranches d'âge 15-64 ans, 15-24 ans, 25-64 ans.

Mesures directes de compétences cognitives (Tests internationaux)

- Lecture et compréhension des textes (*literacy, reading*) : World Bank Education Statistics database (PISA, PIAAC)
- Compétences mathématiques (*numeracy, mathematics*) : World Bank Education Statistics database (PISA, TIMSS-4, TIMSS-8, PIAAC). Pour le temps long, Hearns, Baten et Crayen (2009, table 4, p. 801)
- Résolution de problème en environnement technique élaboré (*Problem solving in technology-rich environment*) : World Bank Education Statistics database (PIAAC)
- Connaissances scientifiques : World Bank Education Statistics database (PISA, TIMSS-4, TIMSS-8)

- Quotient intellectuel (Lynn et Vanhanen [2002, 2006, 2012]) : Moyenne par pays du score global de tests de QI (mesure de compétences cognitives des domaines tels : verbal, observations et mémorisation, logique, calcul, spatial) menés par plusieurs auteurs. Le score 100 des tests de QI correspond à la moyenne britannique (Lynn et Vanhanen, 2012, p. 391).
  Nous n'excluons pas ces données, malgré les propos controversés (hors articles scientifiques) de Lynn, ayant entrainé la révocation de son statut de professeur émérite par l'université d'Ulster, ainsi que des accusations de racisme et sexisme par le syndicat des étudiants de la même université (BBC, 2018). De même nous n'ignorons pas les critiques sur les biais culturels potentiels pouvant entacher les tests cognitifs internationaux s'ils ne sont pas correctement reformulés et recalibrés : cela concerne tout aussi bien QI que les autres tests cognitifs à portée internationale (PISA, TIMSS, PIACC par exemple).
  Lynn et Vanhanen (2012, p. 1) présentent une revue des mesures de corrélation entre leurs estimations de QI nationaux, et les performances scientifiques :

Tableau 1.17 Variables cognitives corrélées au QI national

	Variable	N Pays	r × QI	Référence
1	Publications académiques	139	.87	Morse (2008)
2	Index des brevets	112	.51	Gelade (2008)
3	Autonomie intellectuelle	63	.63	Gelade (2008)
4	STEM	90	.74	Rindermann, Sailer, and Thompson (2009)
5	Brevets : 1960-2007	76	.40	Rindermann et al. (2009)
6	Prix Nobel : Littérature	97	.13	Rindermann et al. (2009)
7	Prix Nobel : Paix	97	.21	Rindermann et al. (2009)
8	Prix Nobel : Science	97	.34	Rindermann et al. (2009)
9	Scientifiques, ingénieurs	51	.61	Rindermann et al. (2009)
10	Exportations technologiques	61	.38	Rindermann et al. (2009)
11	Capacité des politiciens	90	.36	Rindermann et al. (2009)

De même, Lynn et Vanhanen (2012, p. 1) effectuent une revue des mesures de corrélation entre leurs estimations de QI nationaux, et les performances aux tests internationaux de compétences cognitives :

Tableau 1.18 Niveau d'éducation corrélé au QI national

	Variable	N Pays	r × QI	Référence
1	Maths : TIMSS 1999	38	.88	Lynn et Vanhanen (2002)
2	Science : TIMSS 1999	38	.87	Lynn et Vanhanen (2002)
3	Maths/science : 1964/86	38	.81	Lynn et Vanhanen (2006)
4	Maths : age 10. 1994	27	.86	Lynn et Vanhanen (2006)
5	Science : age 10. 1994	26	.79	Lynn et Vanhanen (2006)
6	Maths : age 14. 1994	30	.89	Lynn et Vanhanen (2006)
7	Science : age 14. 1994	37	.81	Lynn et Vanhanen (2006)
8	Maths : PISA. 2000	40	.88	Lynn et Vanhanen (2006)
9	Science : PISA. 2000	40	.83	Lynn et Vanhanen (2006)
10	Maths : PISA. 2003	39	.87	Lynn et Vanhanen (2006)
11	Lecture : age 10	35	.81	Barber (2006)

	Variable	N Pays	r × QI	Référence
12	Maths : age 10. 2003	46	.87	Lynn et Mikk (2007)
13	Science : age 10. 2003	46	.85	Lynn et Mikk (2007)
14	Maths : age 14. 2003	46	.92	Lynn et Mikk (2007)
15	Science : age 14. 2003	46	.91	Lynn et Mikk (2007)
16	Maths, science	63	.89	Rindermann (2007)
17	Maths, science, alphabétisation	56	.84	Lynn et Mikk (2007)
18	Maths, science	73	.90	Meisenberg (2009)
19	Maths, science, alphabétisation	108	.91	Lynn et Meisenberg (2010)
20	Maths, science, alphabétisation	82	.92	Meisenberg et Lynn (2011)

Les corrélations très élevées entre QI et performances aux tests de compétences cognitives, dans le tableau ci-dessus, nous autorise à utiliser le score QI comme mesure synthétique de compétences cognitives dans les études internationales, étant donné que les tests cognitifs usuelles actuels se limitent pour l'essentiel aux seuls pays de l'OCDE.

Les variables cognitives seront également analysées selon le sexe.
La variable qui mesure le stock de capital humain (Lee et Lee [2016], Barro et Lee [2015]) est utilisée comme variable passive, c'est-à-dire descriptive des classes statistiques.

8. Temps transitionnel

Une originalité et apport de notre approche consiste à développer une notion de temps transitionnel différent du temps calendaire, qui s'écoule au rythme des phases et évènements transitionnels et est mesurée à partir d'une série temporelle spécifique.

Ce temps transitionnel nous permet de comparer la trajectoire d'entités (géographiques ou autre) différentes traversés par un même phénomène transitionnel, mais qui se déroule à des périodes historiques (donc calendaires) différentes.
Techniquement, nous définissons le temps transitionnel, comme une série temporelle qui a une dépendance monotone forte avec le temps. Nous mesurons cette dépendance monotone (concordante ou discordante) par le coefficient de corrélation des rangs de Kendall τ, permettant ainsi de tester l'hypothèse qu'une série temporelle donnée puisse être utilisée comme temps transitionnel.
Nous définissons l'âge transitionnel d'un pays, comme différence entre le temps transitionnel et sa valeur minimale internationale dans la période de transition.
Ces notions sont détaillées dans la suite de cette section.

8.1. Propriété de concordance du subordinateur du temps

Une série temporelle X est un subordinateur du temps si et seulement si X est un processus croissant (à variations bornées).
Si X est un subordinateur du temps, alors X est parfaitement *concordant* avec le temps, i.e. que le coefficient de corrélation des rangs Kendall (entre X et le temps) est égal à 1. Car les rangs (du temps et de son subordinateur X) suivent le même ordre que les valeurs du temps et de son subordinateur X.
Rappel : le coefficient de corrélation des rangs de Kendall, mesure le niveau de *concordance* entre deux classements (rangs).
Le coefficient de corrélation de rangs est également un coefficient de dépendance monotone car il est invariant pour toute transformation monotone croissante des variables.

8.2. Coefficient de corrélation des rangs de Kendall et tendance temporelle monotone

Lecoutre, Tassi [1987], pp. 151-154, ont montré que la tendance temporelle monotone d'une série temporelle X, peut se mesurer par le coefficient de corrélation des rangs de Kendall τ.

Lecoutre, Tassi (*ibid*) développent également un test de tendance monotone d'une série temporelle X, basée sur la loi asymptotique (normale centrée) du coefficient de corrélation des rangs de Kendall τ.

Pour savoir si deux séries temporelles X et Y définies sur un même intervalle de temps, varient dans le même sens ou en sens contraire, on peut étudier le signe du produit $(X_\mathrm{i} - X_\mathrm{j})(Y_\mathrm{i} - Y_\mathrm{j})$ où $(X_\mathrm{i}, Y_\mathrm{i})\ et\ (X_\mathrm{j}, Y_\mathrm{j})$ sont deux réalisations indépendantes du couple (X, Y). Avec *i* et *j* entiers naturels non-nuls distincts.

Si $\left((X_i - X_j)(Y_i - Y_j) > 0\right) = 1/2$, il y a autant de chances d'observer une variation de X et Y dans le même sens (concordance) que dans des sens contraires (discordance).

Saporta ([2011], p. 138-139) définit théoriquement le coefficient de corrélation des rangs de Kendall τ par la différence entre la probabilité de concordance (p_c) et la probabilité de discordance (p_d):

$$\tau = P\left((X_i - X_j)(Y_i - Y_j) > 0\right) - P\left((X_i - X_j)(Y_i - Y_j) < 0\right) = p_c - p_d$$

Donc : $\tau = p_c - (1 - p_c) = 2.p_c - 1$

Ce coefficient est donc compris entre -1 (discordance parfaite) et +1 (concordance parfaite) et s'annule lorsque X et Y sont indépendantes (mais pas seulement dans ce cas).

Essayons de mesurer à présent la monotonie (croissante ou décroissante) de X (continue, sans exæquos) dans le temps.
Cette dépendance monotone entre X et la variable déterministe de temps (que nous supposons discret) peut se mesurer par le coefficient de corrélation des rangs de Kendall τ (Lecoutre, Tassi [1987], pp. 151-154).
Lecoutre, Tassi (*ibid*) développent également un test de tendance monotone d'une série temporelle X, basée sur la loi asymptotique (normale centrée) du coefficient de corrélation des rangs de Kendall τ.

Le temps est une variable déterministe monotone croissante par définition.
Donc toute série temporelle concordante avec le temps, est monotone croissante.
Nous pouvons donc écrire :
Probabilité de concordance $p_c = P\left(\left(X_i - X_j\right) > 0\right)$, $\forall\, i > j$
Et probabilité de discordance $p_d = P\left(\left(X_i - X_j\right) < 0\right)$, $\forall\, i > j$

Or $\left\{\left(X_i - X_j\right) > 0, \forall\, i > j\right\} = \{(X_i - X_{i-1}) > 0, \forall\, i\}$
Et $\left\{\left(X_i - X_j\right) < 0, \forall\, i > j\right\} = \{(X_i - X_{i-1}) < 0, \forall\, i\}$

Donc
$p_c = P\big((X_i - X_{i-1}) > 0\big)$, $\forall\, i$
$p_d = P\big((X_i - X_{i-1}) < 0\big)$, $\forall\, i$

D'où :

$$\tau = p_c - p_d = P\big((X_i - X_{i-1}) > 0\big) - P\big((X_i - X_{i-1}) < 0\big) = 2.P\big((X_i - X_{i-1}) > 0\big) - 1$$

Empiriquement, pour un échantillon de taille n on a : $p_c = \frac{\sum_{i=1}^{n} 1_{R_+^*}(X_i - X_{i-1})}{n}$
Donc :

$$\tau = 2.\frac{\sum_{i=1}^{n} 1_{R_+^*}(X_i - X_{i-1})}{n} - 1$$

8.3. Temps transitionnel

L'objectif ici est de définir un *temps transitionnel* universel, qui permet de situer chaque pays dans son processus de transition (démographique, ou économique par exemple), à un moment (temps) donné de son histoire.

Nous définissons le *temps transitionnel*, comme une série temporelle qui est *fortement* concordante avec le temps. Autrement dit, un *temps transitionnel*, est une série temporelle qui a une dépendance monotone *forte* avec le temps.

Nous faisons l'hypothèse statistique d'une concordance *forte*, pour les cas où le coefficient de corrélation des rangs de Kendall τ entre la série temporelle testée (pour jouer le rôle de temps transitionnel) et le temps, est significatif (au seuil de 5%) et supérieur ou égal à 50% (en nous inspirant de l'échelle de Cohen, Cohen [1988], p. 227).

Remarque : Toute transformée symétrique (par rapport à l'axe horizontal du temps, et à une transformation monotone près) d'une série temporelle discordante au temps, est concordante au temps, donc peut jouer le rôle de *temps transitionnel* (si la discordance est forte).
Dans notre quête d'un *temps transitionnel*, nous utiliserons la propriété ci-dessus pour rechercher les séries temporelles ayant une concordance ou une discordance forte (donc une monotonie forte) avec le temps.

8.4. Age transitionnel

Nous définissons l'âge transitionnel d'un pays, comme différence entre le temps transitionnel et sa valeur minimale (quel que soit le pays) dans la période de transition. Cette valeur minimale [notée $Min(X)$]est un seuil de passage, i.e. un évènement définissant le début théorique de la période transitoire, quel que soit le pays.

$$Age_{transitionnel} = X - Min(X)$$

Où X est une série temporelle de temps transitionnel.

Remarque : Si $X_{symetrique}$ est une transformée symétrique (par rapport à l'axe horizontal du temps, et à une transformation monotone près) du temps transitionnel X, alors il est possible de donner une définition d'âge transitionnel par :

$$\begin{aligned} Age_{transitionnel} &= -X_{symetrique} - \left(-Max\left(X_{symetrique}\right)\right) \\ &= Max\left(X_{symetrique}\right) - X_{symetrique} \end{aligned}$$

$Max\left(X_{symetrique}\right)$ étant un seuil indépassable durant la période de transition.

9. Conclusion

Ce premier chapitre a pour finalité de définir le cadre et les outils conceptuels nous permettant de poursuivre nos recherches cliométrique de l'histoire globale.

Après les précisions sur les notions d'histoire globale et de cliométrie, nous avons pu identifier les domaines d'apport de la cliométrie à l'histoire globale : à savoir, son approche quantitative et ses modèles de mathématiques appliquées.

Nous avons également pu définir les domaines et les phénomènes transitionnels (lorsque cela est établi dans la littérature) de la grille multidisciplinaire ACIDE (anthropologie, cognition, institutions, démographie et économique), ainsi que les indicateurs statistiques permettant une étude quantitative de ces domaines.

Pour chacun des domaines anthropologique, cognitif, institutionnel et démographique, nous avons également montré les liens avec le domaine économique, légitimant ainsi leur usage pour des analyses multidisciplinaires en sciences économiques.

Enfin, nous avons complété la grille quantitative et multidisciplinaire ACIDE, avec une mesure alternative du temps, qui est adaptée au suivi comparatif des trajectoires entre pays, soumis à des phénomènes transitionnels se déroulant à des périodes historiques distinctes et à des intensités et rythmes différents. Ainsi nous avons introduit les notions de temps transitionnel commun à plusieurs pays et d'âge transitionnel propre à chaque pays.

Passons à présent en revue, les principaux résultats de ce chapitre.

L'histoire globale est une méthode d'étude historique qui se distingue (Testot *et al.*, 2015, p.6) par quatre caractéristiques essentielles : la transdisciplinarité (anthropologie, démographie, science politique, archéologie, sociologie, par exemple), l'approche en temps long, la vision globale (pays interconnectés à l'échelle mondiale), et les échelles multiples d'analyse (jeu de va-et-vient contrastés, entre le global et le local, le long terme et le court terme, le niveau macro avec le niveau micro).

La cliométrie se définit comme l'application de modèles économiques enrichis d'applications économétriques à l'histoire quantitative et institutionnelle. Elle permet une lecture des évènements historiques dans une perspective de temps long, avec une capacité de prospective à long terme dans les domaines étudiés.

La transition économique moderne (i.e. associé au phénomène de Révolution Industrielle) désigne le passage d'une économie agraire, vers une économie de service, avec une poussée transitoire industrielle, une hausse constante de la productivité des secteurs agricole et industriel, une baisse constante des prix réels des produits alimentaires et industriels et une consommation de masse des produits alimentaires et industriels.
Dans le cadre de la transition économique, nous utilisons les composantes structurelles suivantes de l'économie : le pourcentage de la population en âge de travailler (15-64 ans), le taux d'emploi de la population en âge de travailler, les variables transitionnelles classiques (i.e. les taux d'emploi sectoriels), la productivité sectorielle par emploi.

Le concept de transition démographique désigne le passage d'un régime relativement stable avec fécondité et mortalité « élevées », vers un régime à fécondité et mortalité nettement plus faibles. Cette

transition est initiée par une chute de la mortalité, suivie quelques décennies plus tard par celle de la fécondité.
Malgré l'intérêt certain du solde migratoire dans le parcours démographique d'un pays, les taux de mortalité et de natalité ont été suffisants pour l'analyse de la transition démographique.
Toutefois, nous avons observé que durant les transitions démographiques (notamment européennes), les périodes de plus fort accroissement naturel correspondent aux périodes de solde migratoire négatif, et se manifestant par des sorties massives nettes de population. Il existe donc historiquement une corrélation forte et positive entre le solde naturel et l'émigration.

L'anthropologie culturelle couvre les habitudes acquises par les êtres humains en société.
Dans le cadre des relations entre l'anthropologie et l'économie, nous avons testé l'hypothèse de Todd (sur la disparité des inégalités de revenu au sein de l'OCDE qui pourrait s'expliquer par la diversité des types familiaux), en mesurant au sein de l'OCDE, les associations entre variables anthropologiques (types familiaux, dimensions) d'une part et inégalités de revenu (moyenne mobile des 10 dernières années) d'autre part. Les inégalités de revenus *avant* taxes et redistribution, s'expliquent moyennement par les variables anthropologiques : 19% de variance expliquée. Par contre, les inégalités de revenus *après* taxes et redistribution, s'expliquent fortement par les variables anthropologiques : 63% de variance expliquée.
La conclusion est donc que les politiques publiques de taxation et de redistributions des revenus sont très fortement déterminées par les profils anthropologiques dans différents pays.
En ajoutant une dimension inégalitaire au modèle anthropologique Inglehart-Welzel, nous proposons les variables suivantes pour caractériser la dynamique anthropologique : valeurs laïques-rationnelles (vs valeurs traditionnelles), valeurs d'expression personnelle (vs valeurs de survie), valeurs inégalitaires (mesurées par l'indice de Gini après taxes et redistribution, à très forte corrélation avec les variables anthropologiques).

Les institutions économiques désignent les règles légales régissant l'activité économique.
En combinant les recommandations institutionnelles de List (1841) et North (1994, 1996), nous pouvons conclure que les bonnes institutions économiques sont des instruments stratégiques au service du développement économique, de l'indépendance et de la puissance politique.
Les bonnes institutions économiques doivent donc s'adapter à la culture du pays et sa phase de développement, mais aussi à sa situation concurrentielle vis-à-vis de ses partenaires commerciaux.
Les institutions économiques des échanges externes servent de levier d'adaptation aux opportunités et risques mondiaux, tandis que les institutions économiques internes doivent tendre dans une direction de renforcement inclusif et de respect des droits de propriété et libertés économiques individuelles ; le tout au service des activités productives, notamment celles qui procurent des gains de productivité importants, historiquement l'industrie manufacturière, source de prospérité, d'indépendance et de puissance du pays.
Dans la revue des mesures effectuée, nous constatons une carence de mesures de l'inclusivité économique (pluralité de participants et de bénéficiaires). Aussi, proposons-nous plusieurs indicateurs originaux.
Considérons la différence entre le rang en PIB par habitant, et le rang en Indice de développement humain d'un pays : si les valeurs négatives indiquent un classement humain faible par rapport au classement économique (problème d'inégalité et de prédation économique et de non-redistribution), les valeurs positives élevées indiquent un classement économique indigne du classement humain, donc clairement un problème de liberté et d'accès aux opportunités économiques, alors même qu'un Etat autoritaire assure une éducation et un système sanitaire d'une certaine efficacité.
En utilisant la différence entre les rangs économique et humain d'un pays, nous avons défini deux indices : l'un mesurant l'inclusivité et l'autre la non-inclusivité (en distinguant les deux profils de non-inclusivité).

En calculant notre indice d'inclusivité des institutions économiques en 2015, nous constatons que contrairement à l'hypothèse avancée par Acemoglu et Robinson (2015, p.130) en se basant sur les

institutions politiques, la Chine a été plus inclusive économiquement que la moyenne de l'OCDE et la moyenne mondiale de 1995 à 2010. Même si l'inclusivité économique chinoise a chutée depuis lors, elle reste relativement proche de la moyenne mondiale.

Une autre originalité de notre approche institutionnelle consiste à caractériser les institutions via, d'une part, des indicateurs « externes » liés à la souveraineté, et d'autre part, des indicateurs « internes » de gouvernance et de respect des droits constitutionnels de l'individu.
Ensuite, à chacun de ces niveaux (externe et interne), nous distinguerons deux sous-catégories d'institutions : Les institutions politiques, puis les institutions économiques.
Nous obtenons ainsi une matrice 2x2 que nous dénommons Matrice IEPE (Interne, Externe, Politique, Economique) de mesure des institutions.
Dans cette matrice, nous avons donc quatre catégories d'institutions : les institutions politiques internes, les institutions politiques externes, les institutions économiques internes, et enfin les institutions économiques externes.
Les institutions politiques internes se caractérisent par exemple par les indicateurs du régime politique ou des droits civiques. Les institutions politiques externes peuvent se mesurer par l'aspect colonial ou non du régime politique, le poids économique, le poids militaire du pays. Les institutions économiques internes peuvent se caractériser par l'inclusivité économique, les dépenses publiques, les prélèvements obligatoires, la compétitivité (mesuré par la production manufacturière) du pays. Enfin, les institutions économiques externes peuvent se mesurer par l'ouverture commerciale, la balance commerciale, les taxes sur commerce international.

La cognition désigne l'ensemble des fonctions de connaissance et de conscience, telles que la perception, la conception, la mémorisation, le raisonnement, le jugement, l'imagination et la résolution des problèmes (*American Psychological Association*, 2019).
Les compétences cognitives désignent les capacités opérationnelles relavant de la cognition, et liées à un domaine particulier, comme par exemple les mathématiques, les sciences physiques, la linguistique.
Le lien entre la cognition et l'économie, se fait à travers le capital humain (Becker, 2002) qui est défini comme un stock de « connaissances, informations, idées, compétences et santé des individus ».
Nous caractérisons la dynamique cognitive par les variables suivantes : Le taux d'adultes de 25-64 ans, non scolarisés ; le taux d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint ; compétences cognitives numériques-littéraires-logiques (mesurées par le QI moyen dont la très forte corrélation avec les scores de tests de compétences cognitives, a été démontré dans de nombreuses études).

L'objectif du temps transitionnel, ici est de définir un temps universel, qui permet de situer chaque pays dans son processus de transition (économique par exemple), à temps calendaire donné de son histoire.
Nous définissons le temps transitionnel, comme une série temporelle qui a une dépendance monotone forte avec le temps. Nous mesurons cette dépendance monotone (concordante ou discordante) par le coefficient de corrélation des rangs de Kendall τ entre la série temporelle testée et le temps calendaire.
Nous pouvons ainsi tester l'hypothèse qu'une série temporelle puisse être utilisée comme temps transitionnel.
Nous notons également que toute transformée symétrique (par rapport à l'axe horizontal du temps) d'une série temporelle discordante au temps, est concordante au temps, donc peut jouer le rôle de temps transitionnel.
Nous définissons l'âge transitionnel d'un pays, comme différence entre le temps transitionnel et sa valeur minimale (quel que soit le pays) dans la période de transition.

Chapitre 2.
Analyse séquentielle des trajectoires macro-économiques, dans l'histoire globale récente

1. Table des matières

Chapitre 2. Analyse séquentielle des trajectoires macro-économiques, dans l'histoire globale récente *100*

1. Table des matières **101**

2. Introduction **102**

3. Détection de phases **104**

3.1. Indicateurs structurels de l'économie 104
3.2. Sélection des variables à étudier 104
3.3. Transformation préalable des variables à étudier 105
3.4. Sélection des pays à étudier 105
3.5. Période historique étudiée 105
3.6. Classification des (pays, années) sur composantes principales 105
3.7. Indexation temporelle des classes de pays 106
3.8. Description statistique des classes de couples (pays, années) 106

4. Analyse des séquences d'états *(State sequences Analysis STS)* **108**

4.1. Notions préalables 108
4.2. Visualisation des données de séquences d'états 109
4.3. Statistiques descriptives des séquences d'états 112
4.4. Typologie des trajectoires 117
4.5. Analyse de la discrépance des séquences 121

5. Analyse des séquences d'évènements *(Time stamped Events analysis TSE)* **125**

5.1. Découverte des sous-séquences les plus fréquentes 125
5.2. Sous-séquences discriminantes des classes de trajectoires 126
5.3. Recherche de sous-séquences et calcul des occurrences 128
5.4. Découverte des règles d'associations séquentielles 128

6. Contributions des composantes structurelles macroéconomiques, à la variation du PIB par habitant **130**

7. Conclusion **142**

8. Bibliographie **145**

2. Introduction

L'analyse séquentielle des parcours biographiques (*life course analysis, event history analysis*) est un champ de recherche très actif dans les sciences sociales depuis le début des années 1980. Il s'agit notamment de réaliser des analyses séquentielles statistiques : visualisations, statistiques descriptives, typologies, modèles explicatifs des séquences d'états ou d'évènements.

La *Sequence Analysis Association* (2020) a recensé depuis 1970, 397 articles scientifiques dans le domaine de l'analyse des séquences ; dont 37% des publications sur les cinq dernières années seulement.
Une revue des thèmes traités dans les publications des cinq dernières années révèle que les trajectoires analysées concernent pour l'essentiel, les parcours types suivants : vie professionnelle, vie familiale, résidences, soins (pour des malades chroniques), santé (bien-être), finances personnelles.
Ces trajectoires sont analysées pour l'ensemble des individus, ou pour des populations spécifiques : retraités et personnes âgées (Azar *et al.*, 2019), adolescents (Kang, 2019), adultes sans enfants (Raab et Struffolino, 2019), parents prenant des congés parentaux (Zhelyazkova et Ritschard, 2018), malades chroniques (Le Meur *et al.*, 2019), personnes souffrant d'invalidités (Perhoniemi, Blomgren, et Laaksonen, 2019), jeunes en début de carrière (Salvà-Mut, Thomás-Vanrell et Quintana-Murci, 2016), immigrés (Fuller, 2015) ou enfants d'immigrés (Ferrari et Pailhé, 2017).
Sur la même période (cinq dernières années), 32 articles (22%) au moins, traitent essentiellement de méthodologie (Cornwell, 2018). ou de logiciels (Ritschard et Studer [2018], Helske S. et Helske J. [2019]).

En ce qui concerne la méthodologie de l'analyse séquentielle des parcours biographiques, une des démarches les plus complètes (Robette, 2014) est celle proposée par (Gabadinho *et al.* [2011], Müller *et al.* [2011]), avec un logiciel à la clé (Gabadinho *et al.*, [2011]).
Cette démarche permet notamment de : visualiser l'ensemble des données de séquences, explorer ces séquences avec des statistiques descriptives, construire une typologie des trajectoires, analyser les associations statistiques entre trajectoires types et covariables, analyser les séquences d'événements.

Si l'analyse des séquences d'états ou d'évènement est activement utilisé pour étudier les trajectoires biographiques dans les sciences sociales ou parfois en merchandising et marketing (Tufféry, 2012, p.365), donc à un niveau « micro », les usages au niveau des agrégats « macro » sont plus rares comme parfois en sciences politiques (Casper et Wilson, 2015) ou en sciences économiques (Emmenegger *et al.* [2013], sur la dualisation *insiders-outsiders* sur les marchés nationaux du travail en Europe).

L'originalité de notre approche consiste à appliquer, dans une perspective d'histoire globale, les méthodes et démarches d'analyse séquentielle biographique, à l'analyse longitudinale macroéconomique.

Pour cette étude, nous-nous sommes restreints à l'histoire récente, car malheureusement, les données historiques sur l'emploi sectoriel ne sont disponibles qu'à partir de 1991.
La période étudiée pour l'ensemble des pays dans le monde (hors « petits » Etats) va de 1991 à 2017. Ces 27 années de l'histoire économique récente, correspondent à peu près une génération.

Nos contributions se situent à quatre principaux niveaux.

D'abord, la détection des phases transitionnelles historiques, par des méthodes de classification statistiques sur composantes principales, appliquée aux données longitudinales économiques. Nous formulons alors, à partir de cet ensemble de phases économiques détectées, l'historique économique

de chaque pays, année après année. Le résultat est que chaque pays a désormais sa trajectoire économique décrite par une séquence de phases (ou états).

Deuxièmement, grâce aux méthodes d'analyse des séquences d'états (*State sequences Analysis STS*), nous utilisons les états (phases) économiques précédemment construites pour réaliser une analyse « biographique » des trajectoires macro-économiques (encore appelées séquences) des pays à l'échelle planétaire sur la période étudiée.
Cela se fera par des visualisations synthétiques de l'ensemble des trajectoires économiques nationales, les statistiques descriptives de ces trajectoires (le niveau des disparités mondiales, le mode des états de chaque année, la durée moyenne passée par les pays dans chaque état, les probabilités de transition entre états économiques), la mise en évidence et la caractérisation des typologies de trajectoires économiques nationales, et enfin, les mesures, tests d'association et modèle non-linéaire (par arbre de segmentation) entre les trajectoires et les caractéristiques individuelles des pays.

Le troisième niveau de contribution, utilise l'analyse des séquences d'évènements (*Time stamped Events analysis TSE*) pour étudier les séquences d'évènements (changement d'état) économiques majeures. Cela revient à identifier les sous-séquences d'évènements (donc de changements) les plus fréquentes dans la trajectoire des pays sur la période étudiée, puis caractériser les typologies de trajectoires établies précédemment par les sous-séquences d'évènements les plus discriminantes. Viennent ensuite les recherches et les calculs d'occurrence de sous-séquences spécifiques, puis surtout, la découverte de règles d'associations temporelles entre sous-séquences d'évènement qui peut jouer un rôle prédictif car basé sur des approches datamining.

Enfin, toujours dans l'analyse longitudinale, nous proposons une méthode d'estimation séquentielle (puis par agrégation) des contributions à la variation pluriannuelle du PIB par habitant, des principaux secteurs économiques (agriculture et pêcheries, industrie, services), et de la dynamique démographique de la population en âge de travailler.

3. Détection de phases

La détection des *phases* économiques consiste à effectuer une classification statistique des couples (pays, années) à partir des indicateurs structurelles des économies nationales. Les classes ainsi créées constituent des *états* ou *phases* dans la trajectoire macro-économique de chaque pays.

3.1. Indicateurs structurels de l'économie

Les indicateurs structurels de l'économie que nous utiliserons sont ceux que nous avons définis au chapitre 1 (section 3.2) : la population en âge de travailler, le taux d'emploi, la part d'emploi de chaque secteur, les productivités sectorielles. Les secteurs considérés sont l'agriculture et la pêche, l'industrie et les services, telles que définies par la Banque Mondiale (World Bank, 2017). Ce sont ces variables (dites « actives ») qui serviront à caractériser les trajectoires des pays.

3.2. Sélection des variables à étudier

Les variables actives qui sont utilisées pour effectuer la classification statistique des couples (pays, années) sont listées ci-après :

Tableau 2.1 Variables actives

Variables	Termes en anglais	Termes en français
Employment.in.agriculture.Pct.of.total.employment	Employment in agriculture (% of total employment) (modeled ILO estimate)	Emploi dans l'agriculture (% de l'emploi total) (estimation modélisée du BIT)
Employment.in.industry.Pct.of.total.employment	Employment in industry (% of total employment) (modeled ILO estimate)	Emploi dans l'industrie (% de l'emploi total) (estimation modélisée du BIT)
Employment.in.services.Pct.of.total.employment	Employment in services (% of total employment) (modeled ILO estimate)	Emploi dans les services (% de l'emploi total) (estimation modélisée du BIT)
Population.ages.15.64.Pct.of.total	Population ages 15-64 (% of total)	Population âgée de 15 à 64 ans (% du total)
Agr.VA.per.worker.agr.product.const.2010USD	Agriculture value added per worker, measure of agricultural productivity (constant 2010 US$)	Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010)
Industry.value.added.per.worker.constant.2010.USD	Industry, value added per worker (constant 2010 US$)	Industrie, valeur ajoutée par travailleur (USD constants de 2010)
Services.value.added.per.worker.constant.2010.USD	Services, value added per worker (constant 2010 US$)	Services, valeur ajoutée par travailleur (USD constants de 2010)

3.3. Transformation préalable des variables à étudier

Les variables actives utilisées font l'objet des transformations suivantes, avant la sélection de l'année d'étude : estimation des valeurs manquantes par interpolation spline, et moyenne mobile arrière d'ordre 5.

3.4. Sélection des pays à étudier

Les analyses sont effectuées avec exclusion des petits Etats (*small states*) au sens de la Banque Mondiale (The World Bank, 2019), ainsi que des pays ayant une superficie inférieure au troisième quartile des *small states* de la Banque Mondiale.
Ces *small states*, du fait de leur « faible population (à savoir, celles avec une population de 1,5 million d'habitants ou moins) et de leur base économique, ces pays sont particulièrement vulnérables aux chocs exogènes, tels que les catastrophes naturelles et le changement climatique. Avec des opportunités économiques limitées et une migration importante, ils sont souvent confrontés à des contraintes de capacité » (The World Bank, 2019).

3.5. Période historique étudiée

La période historique étudiée pour l'ensemble des pays est de 1991 à 2017. L'usage des moyennes mobiles d'ordre 5, fait que les résultats des traitements statistiques sont obtenus à partir de 1995.
Evidemment l'objectif initial était d'inclure l'historique le plus ancien possible, mais les variables structurelles utilisées sont rarement disponibles avant-guerre. Et malheureusement les données historiques sur l'emploi sectoriel ne sont disponibles qu'à partir de 1991.
Toutefois, l'historique étudié couvre 27 ans de l'histoire économique récente, soit à peu près une génération.

3.6. Classification des (pays, années) sur composantes principales

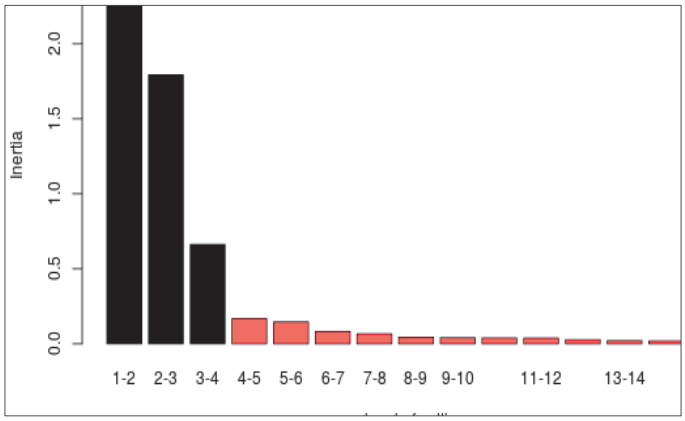
3.6.1. Sélection des composantes principales pour la classification

L'analyse en composantes principales (ACP) est réalisée à partir des 7 variables actives (Tableau 2.1).
Les 2 axes principaux expriment 84.1% de l'inertie totale du jeu de données ; cela signifie que 84.1% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan.

3.6.2. Choix du nombre optimal de classes de couples (pays, années)

Après classification statistique des pays, nous obtenons quatre classes optimales de couples (pays, années).

Figure 2.1 Gain d'inertie en fonction du nombre de classes



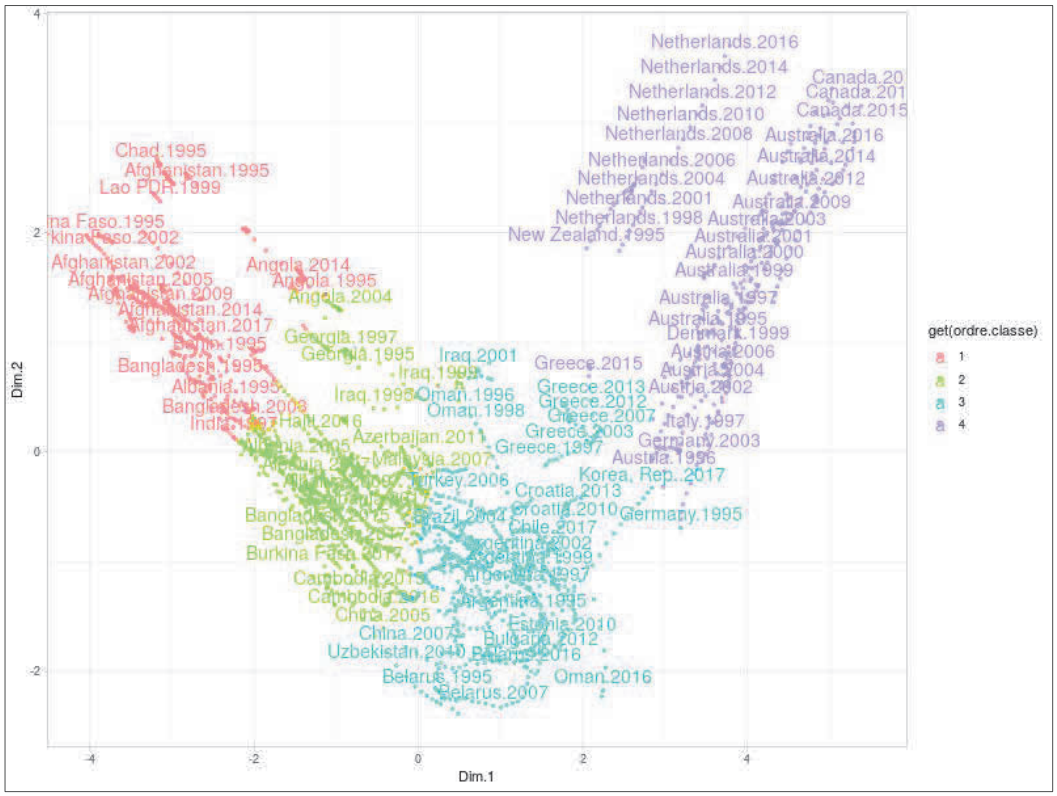
3.7. Indexation temporelle des classes de pays

Etant donné que le taux d'emplois agricoles est de tendance décroissante (en fonction du temps) durant le processus de transition économique, les classes de (pays, années) sont numérotées par ordre décroissant de leur moyenne des taux d'emplois agricoles : ainsi, la classe 4 qui a la moyenne de taux d'emplois agricoles lissé le plus faible, porte l'index temporel le plus élevé (c'est la classe la plus ancienne dans le processus de transition) ; et inversement pour la classe 1 qui est à l'autre bout du spectre.

3.8. Description statistique des classes de couples (pays, années)

Cette représentation des classes de (pays, années) suggère un parcours temporel des pays sous forme d'un V en allant de la gauche vers la droite.

Figure 2.2 Représentation des classes d'âge transitionnel de pays sur les deux premiers axes factoriels



Ces 4 classes illustrent également les 4 *phases* ou *états* observables sur la période historique étudiées, et sont caractérisés par les statistiques ci-après :

Tableau 2.2 Description des classes selon les variables actives

Classes	Emploi dans l'agriculture (% de l'emploi total)	Emploi dans les services (% de l'emploi total)	Emploi dans l'industrie (% de l'emploi total)	Population âgée de 15 à 64 ans (% de la population totale)	Valeur ajoutée agricole par travailleur (produit agricole, $US constant 2010)	Valeur ajoutée industrielle par travailleur ($US constant 2010)	Valeur ajoutée des services par travailleur ($US constant 2010)
1	70,2	21,7	8,1	52,6	764	5901	4223
2	40,3	41,0	18,7	59,0	2676	11391	6997
3	14,7	58,8	26,5	66,3	9013	NS	21217
4	4,1	70,9	25,0	66,5	45254	90890	83484
Monde	32,2	47,6	20,2	61,2	11032	27195	23377

NS : valeur non-significativement différente de la moyenne mondiale au seuil de 5%

Ces 4 *phases* ou *états* serviront de base pour les analyses séquentielles des trajectoires macro-économiques des pays.

4. Analyse des séquences d'états *(State sequences Analysis STS)*

L'analyse des séquences d'états (*State sequences Analysis STS*) consiste à réaliser une analyse « biographique » des trajectoires macro-économiques des pays à l'échelle planétaire sur les 23 dernières années. Les *états (states)* ou *phases* utilisées pour caractériser les séquences historiques nationales, ont été établis lors de la détection des phases précédemment effectuée.
Comme indiqué précédemment à la section 3.7, l'index temporel (1, 2, 3 et 4) des quatres phases détectées représente un ordre d'avancement des pays dans le processus de transition économique, selon le critère du taux d'emplois agricole (qui baisse durant la transition) ou encore de la productivité agricole (qui augmente durant la transition).

L'historique des états couvre une période de 23 ans (1995 à 2017).

4.1. Notions préalables

4.1.1. Séquences

Définitions :
Alphabet *A* : ensemble fini d'éléments.
Séquence de longueur *k* : liste ordonnée de *k* éléments successivement choisis dans *A*

Notations :
La séquence *x* de longueur *k* est notée : $x = (x_1, x_2, \ldots, x_k)$
S'il n'y a pas de risque de confusion, on peut noter également : $x = x_1 x_2 \cdots x_k$
L'insertion d'un séparateur entre éléments successifs peut être nécessaire si A inclut un ou plusieurs notations composées de plusieurs caractères. Ce séparateur peut être un tiret « - ».
Exemple : Si x_1=*C*, x_2=*M*, x_3=*ME*, alors *x=C-M-ME*

4.1.2. Types de séquences

L'alphabet *A* est composé d'un nombre fini d'éléments.
Les éléments sont tous d'une même nature qui peut être temporelle ou non-temporelle.

Un élément de nature temporelle est un *état* ou un *changement d'état* ou un *ensemble de changements d'états*.
Un *changement d'état* s'appelle *évènement*.
Exemple : la notation *état1>état2* désigne l'évènement de passage de l'état *état1* vers l'état *état2*

Une *transition* est un ensemble (non-ordonné) d'évènements simultanés. Ainsi, un évènement est transition n'ayant qu'un seul élément.
Exemples : La notation *(état1>état2)* désigne une transition composée du seul évènement *état1>état2.*
La notation *(état1)* désigne une transition composée de l'évènement qui consiste à être initialement dans l'état *état1* et à y rester.

Un élément de nature non-temporelle est un *objet* ou un *changement d'objet*.

En résumé, un élément de l'alphabet *A*, ne peut être que d'une des quatre natures : état, transition, objet, changement d'objet.

Par conséquent, une séquence ne peut être que d'une des quatre natures suivantes, selon la nature des éléments qui la composent : *séquence d'états*, *séquence de transitions*, *séquence d'objets*, *séquence de changements d'objet*.

Par convention la *séquence de transitions* est encore appelée (par abus de langage) *séquence d'évènements*.
Exemple : La notation *(état1>état2)-(état1)* désigne une séquence composée de 2 transitions *(état1>état2)* et *(état1)*

Une sous-séquence d'une séquence *S* d'évènements, est une séquence issue de la séquence *S* où l'ordre des évènements dans *S* est respecté.

4.1.3. Notation de séquences d'évènements avec délais entre les transitions

Chaque séquence d'évènements avec des délais est notée sous la forme :

(event1,event2,...)-délai-(event2,...)-délai

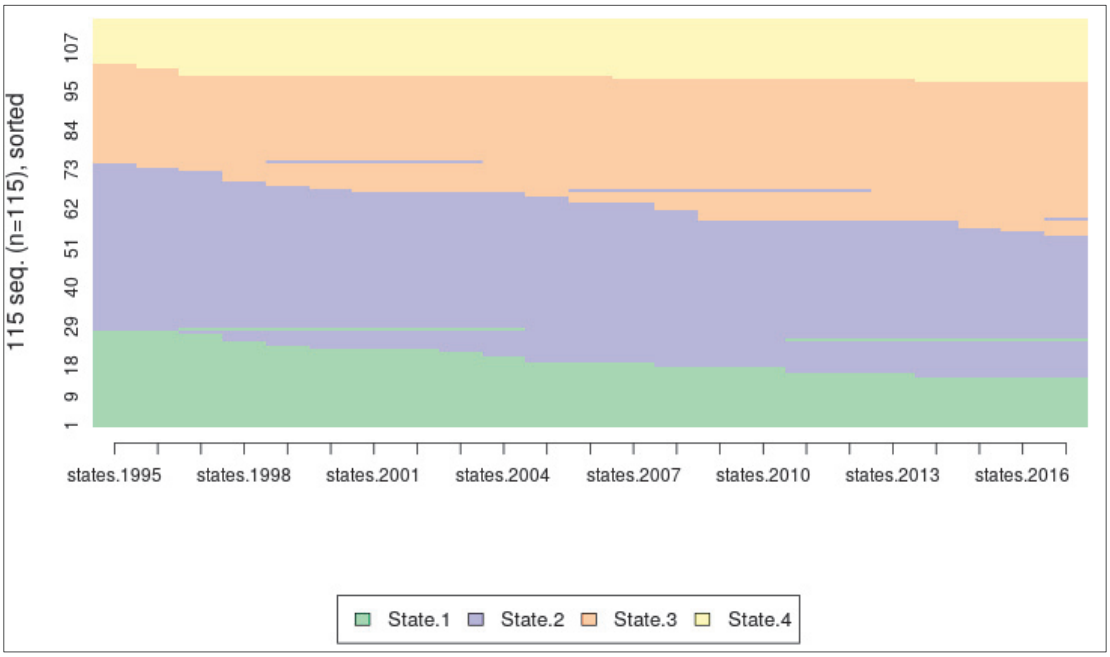
où *(event1,event2,...)* désigne une transition comportant un ou plusieurs évènements simultanés.
délai désigne le temps écoulé entre deux transitions ou jusqu'à la fin de la séquence.

4.2. Visualisation des données de séquences d'états

4.2.1. Visualisation des pays selon leur état initial sur la période étudiée

Dans ce graphique, chaque état (*state*) a une couleur distincte, et les pays ayant un même état initial (donc une même couleur initiale) sont superposés ensemble. Les lignes horizontales correspondent aux trajectoires des pays.

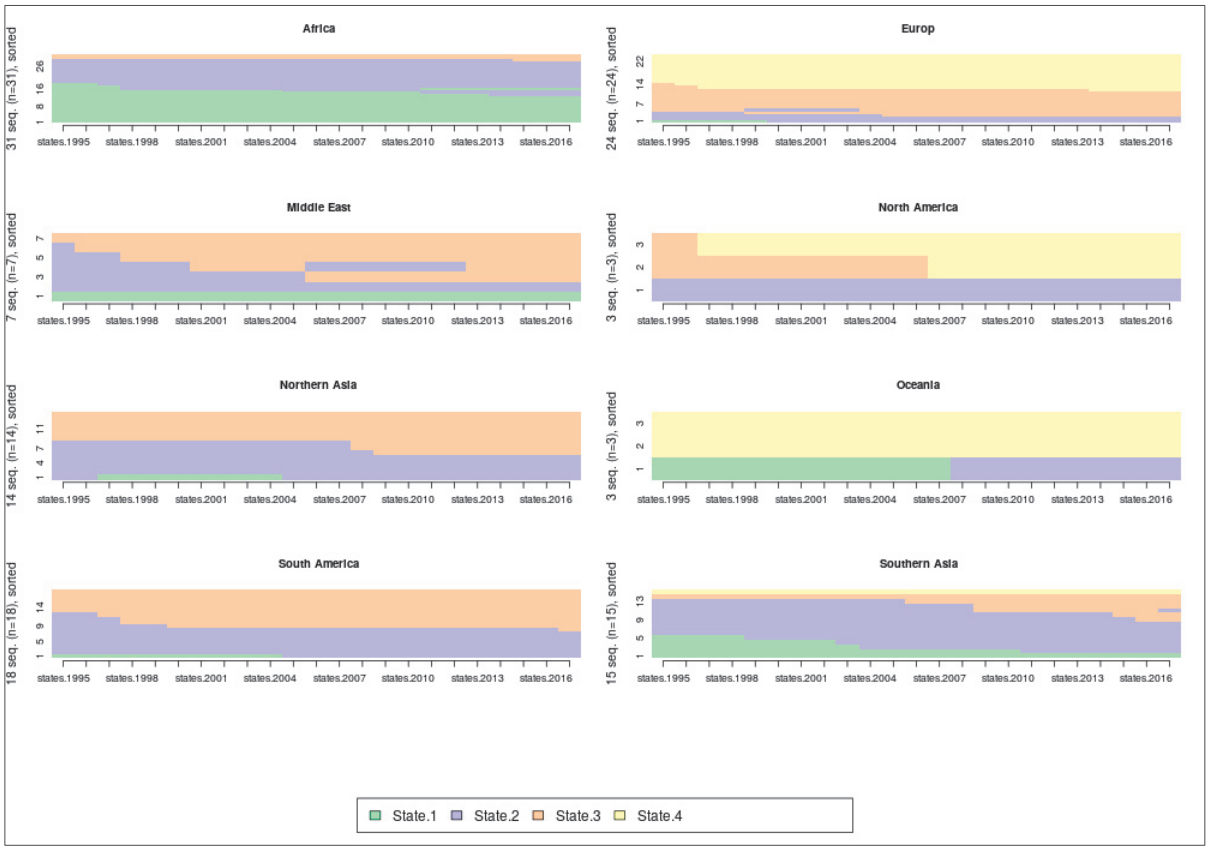
Figure 2.3. Visualisation de tous les pays selon leur état initial sur la période étudiée



Sur la période historique étudiée, nous observons l'état économique la moins avancée (en termes de transition économique) qui est State.1 est en décroissance continue. L'état économique suivant (State.2) décroit légèrement, tandis que nous constatons une importante hausse du nombre de pays en phase 3 (State.3). Le nombre de pays transitionnellement avancés (State.4) a également tendance à la hausse.

Le graphique ci-après décline le précédent descriptif par grande régions mondiales :

Figure 2.4. Visualisation par zone géographique (ou autre caractéristique de pays), des pays selon leur état initial sur la période étudiée



Ainsi, nous constatons des progressions transitionnelles sur tous les continents, caractérisés par des passages d'états économiques d'indexes plus faibles à des états économiques d'indexes plus élevés.

4.2.2. Visualisation des séquences d'états les plus fréquentes

Cette section descriptive offre une visualisation mondiale et par grandes régions des 10 séquences d'états les plus fréquentes. Les fréquences sont proportionnelles à l'épaisseur des barres horizontales.

Figure 2.5. Visualisation des 10 séquences d'états les plus fréquentes

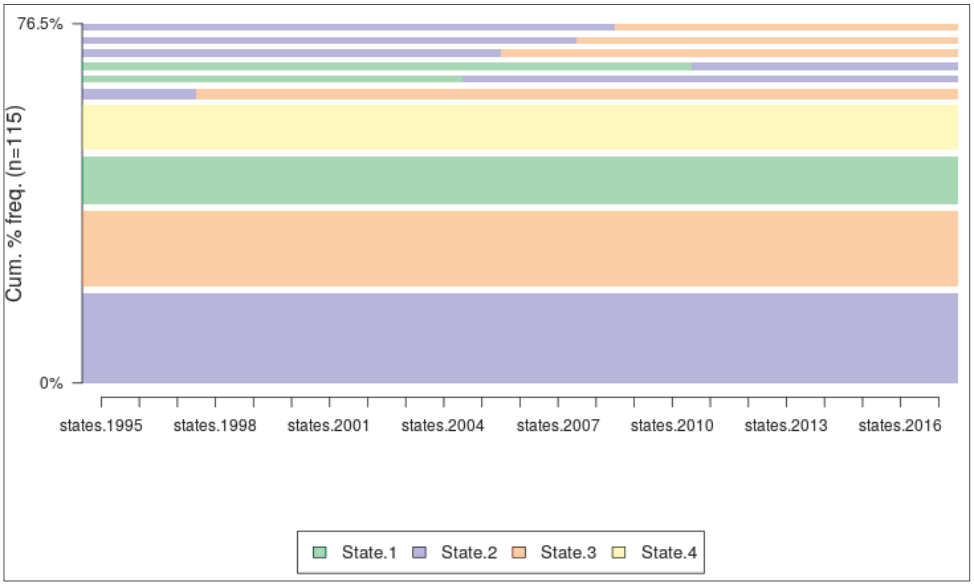
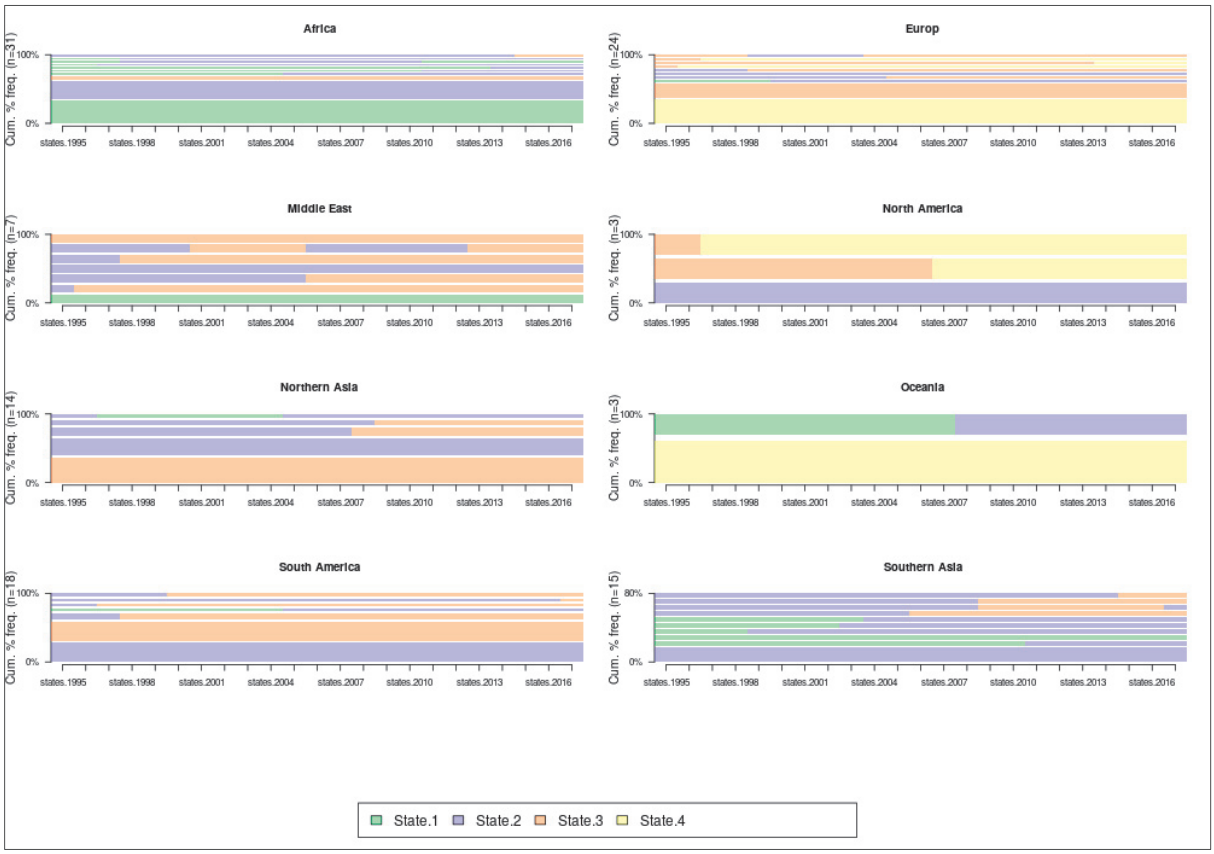


Figure 2.6. Visualisation des 10 séquences d'états les plus fréquentes, par zone géographique (ou autre caractéristique de pays)



Nous observons ainsi que quelle que soit la perspective géographique (mondiale ou par région), le fait de rester dans un même état sur la période étudiée est toujours la trajectoire la plus fréquente. Cependant les cas de changement d'état (vers plus de développement) font systématiquement partie des 10 séquences types les plus fréquentes.

4.2.3. Décompte des occurrences de toutes les séquences observées

Dans ce tableau, les séquences d'états observées sont triées par ordre décroissant de leur nombre (ou pourcentages) d'occurrences.

Le nombre qui suit (après le signe « / ») le nom de l'état désigne la durée (ou délai) de l'état concerné.

Comme nous l'avions précédemment noté, les séquences mono-état sont les plus fréquentes. Evidemment plus il y a d'états dans une séquence, moins elle est fréquente.

Nous constatons également l'absence des cas de régression durables sur les 23 années observées.

Tableau 2.3. Nombre et Pourcentage d'occurrences de toutes les séquences d'états

	Freq	Percent
State.2/23	26	22.61
State.3/23	22	19.13
State.1/23	14	12.17
State.4/23	13	11.30
State.2/3-State.3/20	3	2.61
State.1/10-State.2/13	2	1.74
State.1/16-State.2/7	2	1.74
State.2/11-State.3/12	2	1.74
State.2/13-State.3/10	2	1.74
State.2/14-State.3/9	2	1.74
State.2/20-State.3/3	2	1.74
State.3/2-State.4/21	2	1.74
State.1/13-State.2/10	1	0.87
State.1/19-State.2/4	1	0.87
State.1/2-State.2/21	1	0.87
State.1/3-State.2/13-State.1/7	1	0.87
State.1/3-State.2/20	1	0.87
State.1/4-State.2/19	1	0.87
State.1/5-State.2/18	1	0.87
State.1/8-State.2/15	1	0.87
State.1/9-State.2/14	1	0.87
State.2/1-State.3/22	1	0.87
State.2/10-State.3/13	1	0.87
State.2/14-State.3/8-State.2/1	1	0.87
State.2/2-State.1/8-State.2/13	1	0.87
State.2/2-State.3/21	1	0.87
State.2/21-State.3/2	1	0.87
State.2/22-State.3/1	1	0.87
State.2/4-State.3/19	1	0.87
State.2/5-State.3/18	1	0.87
State.2/6-State.3/5-State.2/7-State.3/5	1	0.87
State.3/1-State.4/22	1	0.87
State.3/12-State.4/11	1	0.87
State.3/19-State.4/4	1	0.87
State.3/4-State.2/5-State.3/14	1	0.87

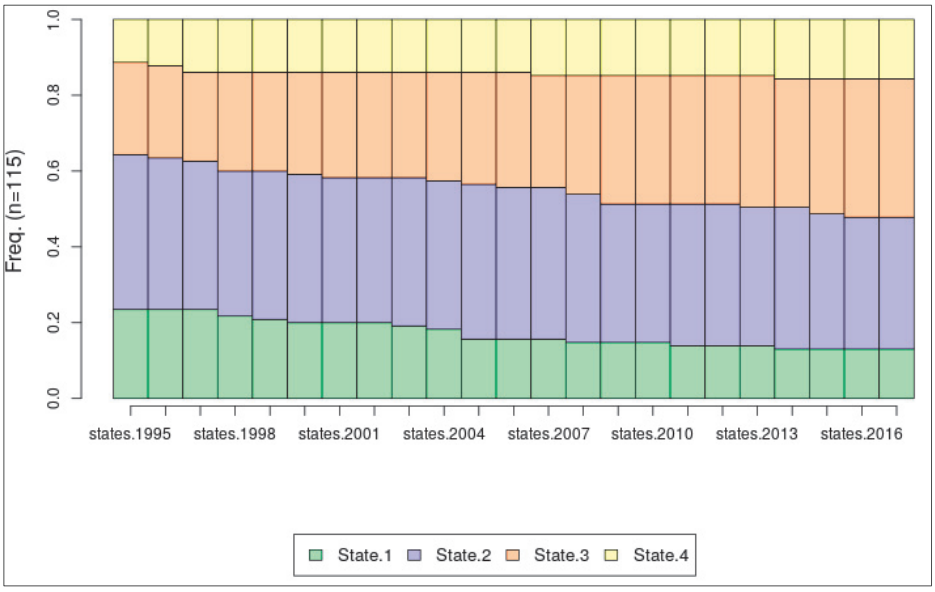
4.3. Statistiques descriptives des séquences d'états

4.3.1. Distributions transversales des états par année

Dans le graphique analysé ici, la visualisation se fait de manière verticale (transversale) i.e. par année et non de manière horizontale séquentielle.
On peut y observer 23 barres verticales correspondant aux 23 années de la période étudiée.

Pour chaque année, le graphique indique la proportion de chaque état parmi les pays étudiés.
Les conclusions sont naturellement les mêmes que celles de la section 4.2.1 : le nombre de pays dans l'état économique la moins avancée dans la transition économique (State.1) est en décroissance continue. Le nombre de pays dans l'état économique State.2 décroit légèrement, tandis qu'en phase 3 (State.3), il y a une importante hausse du nombre de pays. Le nombre de pays les plus développés (State.4) est également en tendance haussière.

Figure 2.7. Distributions transversales des états par année (par tranches verticales)

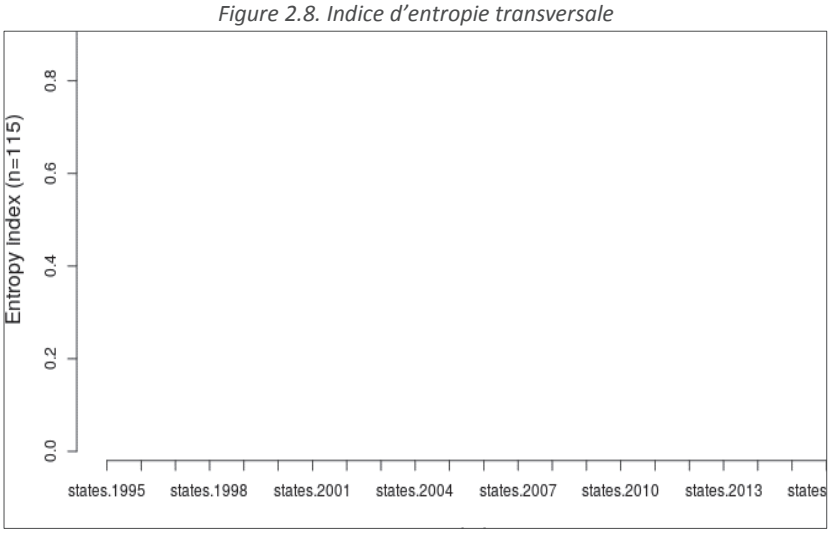


4.3.2. Séquence d'entropies des distributions transversales d'états

L'indice d'entropie transversale (*transversal entropy index*), mesure chaque année l'hétérogénéité de la distribution des états parmi les pays étudiés.
Chaque distribution transversale est notée $(p_1, \ldots, p_a)$, avec $a = |A|$ nombre d'éléments de l'alphabet A
L'entropie de Shannon est mesurée par :

$$h(p_1, \ldots, p_s) = -\sum_{\mathrm{i}=1}^{s} p_{\mathrm{i}} \log_2(p_{\mathrm{i}})$$

Figure 2.8. Indice d'entropie transversale



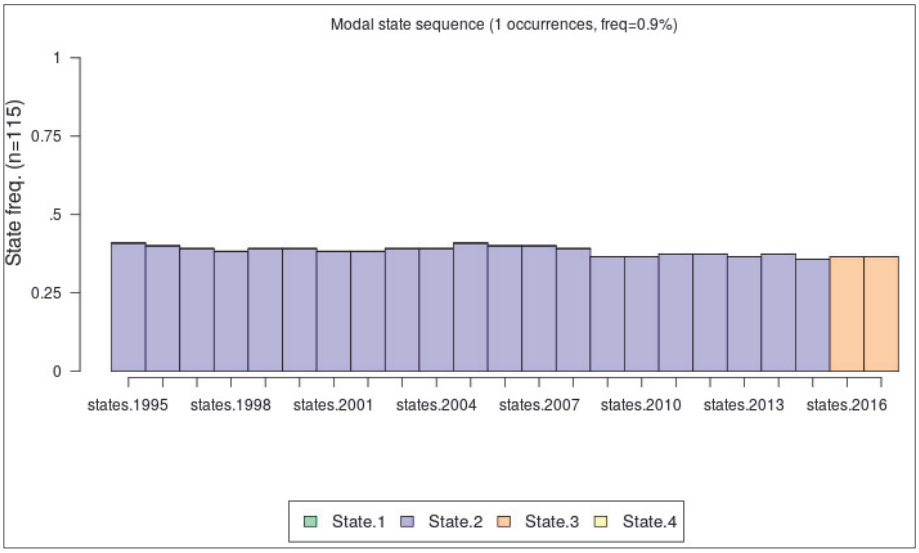
Nous pouvons noter que le niveau de variabilité des états dans lesquels sont les pays, baisse très légèrement. Donc réciproquement, cela signifie une légère hausse de l'homogénéité des pays.

4.3.3. Séquence du mode des états dans les distributions transversales

La figure ci-dessous indique pour chaque année le mode des états, dans des visions transversales.

Ainsi, de 1995 à 2015, l'état (phase) le plus fréquent par an, est State.2. Cependant une bascule vers un état plus avancé (State.3) est observable dès 2016.

Figure 2.9. Mode des états dans les distributions transversales (une colonne par année)



4.3.4. Temps passé dans les différents états, pour chaque pays

Pour chaque pays, nous calculons le temps passé dans chaque état de l'alphabet (ensemble des quatre états), durant la période étudiée.
Nous constatons comme anticipé, que moins un pays est développé (part d'emplois agricoles élevé), plus il a passé du temps dans les états d'indice faible. Et inversement.
Le tableau ci-après, malgré sa taille, permet de mettre un nom sur les pays étudiés et leur profil séquentiel, en termes de durée passée dans les différents états.

Tableau 2.4 Temps passé dans les différents états, pour chaque pays

	State.1	State.2	State.3	State.4
Afghanistan	23	0	0	0
Cameroon	23	0	0	0
Central African Republic	23	0	0	0
Chad	23	0	0	0
Congo, Dem. Rep.	23	0	0	0
Cote d'Ivoire	23	0	0	0
Guinea	23	0	0	0
Lao PDR	23	0	0	0
Sierra Leone	23	0	0	0
Somalia	23	0	0	0
Tanzania	23	0	0	0
Uganda	23	0	0	0
Zambia	23	0	0	0
Zimbabwe	23	0	0	0
Burkina Faso	19	4	0	0
Benin	16	7	0	0
Cambodia	16	7	0	0
Papua New Guinea	13	10	0	0
Angola	10	13	0	0
Haiti	10	13	0	0
Sudan	10	13	0	0
Bangladesh	9	14	0	0
Tajikistan	8	15	0	0
Vietnam	8	15	0	0
Albania	5	18	0	0
India	4	19	0	0
Kenya	3	20	0	0
Ghana	2	21	0	0
Armenia	0	23	0	0
Azerbaijan	0	23	0	0
Bolivia	0	23	0	0
Congo, Rep.	0	23	0	0
Ecuador	0	23	0	0
Georgia	0	23	0	0
Guatemala	0	23	0	0
Honduras	0	23	0	0
Indonesia	0	23	0	0
Kyrgyz Republic	0	23	0	0
Macedonia, FYR	0	23	0	0
Madagascar	0	23	0	0
Malawi	0	23	0	0
Mali	0	23	0	0
Mauritania	0	23	0	0
Mexico	0	23	0	0
Morocco	0	23	0	0
Mozambique	0	23	0	0
Nepal	0	23	0	0
Nicaragua	0	23	0	0
Niger	0	23	0	0
Nigeria	0	23	0	0
Pakistan	0	23	0	0
Peru	0	23	0	0
Philippines	0	23	0	0
Senegal	0	23	0	0
Paraguay	0	22	1	0
Thailand	0	21	2	0
Egypt, Arab Rep.	0	20	3	0
Sri Lanka	0	20	3	0

	State.1	State.2	State.3	State.4
Mongolia	0	15	8	0
Malaysia	0	14	9	0
Uzbekistan	0	14	9	0
Iraq	0	13	10	0
Kazakhstan	0	13	10	0
Moldova	0	13	10	0
China	0	11	12	0
Syrian Arab Republic	0	11	12	0
Turkey	0	10	13	0
Colombia	0	5	18	0
Romania	0	5	18	0
Bosnia and Herzegovina	0	4	19	0
Dominican Republic	0	3	20	0
Iran, Islamic Rep.	0	3	20	0
Panama	0	3	20	0
Brazil	0	2	21	0
Oman	0	1	22	0
Argentina	0	0	23	0
Belarus	0	0	23	0
Bulgaria	0	0	23	0
Chile	0	0	23	0
Costa Rica	0	0	23	0
Croatia	0	0	23	0
Cuba	0	0	23	0
Estonia	0	0	23	0
Hungary	0	0	23	0
Jordan	0	0	23	0
Korea, Rep.	0	0	23	0
Latvia	0	0	23	0
Libya	0	0	23	0
Lithuania	0	0	23	0
Poland	0	0	23	0
Portugal	0	0	23	0
Russian Federation	0	0	23	0
Serbia	0	0	23	0
South Africa	0	0	23	0
Ukraine	0	0	23	0
Uruguay	0	0	23	0
Venezuela, RB	0	0	23	0
Greece	0	0	19	4
Canada	0	0	12	11
Iceland	0	0	2	21
United States	0	0	2	21
Germany	0	0	1	22
Australia	0	0	0	23
Austria	0	0	0	23
Denmark	0	0	0	23
Finland	0	0	0	23
France	0	0	0	23
Italy	0	0	0	23
Japan	0	0	0	23
Netherlands	0	0	0	23
New Zealand	0	0	0	23
Spain	0	0	0	23
Sweden	0	0	0	23
Switzerland	0	0	0	23
United Kingdom	0	0	0	23

4.3.5. Durée moyenne passée dans chaque état de l'alphabet (ensemble des états)

Nous mesurons ici, la moyenne de la durée passée dans chacun des états, dans les séquences des pays.

Figure 2.10. Durée moyenne passée dans chaque état de l'alphabet

Tableau 2.5. Statistiques descriptives du Temps passé dans les différents états

	State1	State2	State3	State4
Min	0	0	0	0
1st Qu	0	0	0	0
Median	0	3	0	0
Mean	3,957	8,783	6,974	3,287
3rd Qu	0	20,5	18,5	0
Max	23	23	23	23

Dans le cas présent, la durée moyenne passée par les pays dans un état donnée, n'apporte pas une information vraiment pertinente, étant donné que les distributions sont très asymétriques : globalement chaque pays semble avoir passé l'essentiel de son temps dans un état spécifique, avec quelques fois des évolutions vers un autre état. Il y a donc une inertie importante, avec des évolutions relativement lentes peuvent tenir à la relative étroitesse de la période retenue.

4.3.6. Taux de transition

Les taux de transition sont des estimations des probabilités de passage d'un état à un autre (pouvant rester identique) d'une année à l'autre : $p\left(x_{it}|x_{j(t-1)}\right)$

Tableau 2.6. Taux de transition entre états de l'alphabet

	[-> State.1]	[-> State.2]	[-> State.3]	[-> State.4]
[State.1 ->]	0,9682	0,0318	0	0
[State.2 ->]	0,0021	0,9753	0,0227	0
[State.3 ->]	0	0,0039	0,9895	0,0066
[State.4 ->]	0	0	0	1

Les taux de transition calculés, ne constituent toutefois pas une matrice de transition de Markov valable pour tous les pays : Notamment les pays ayant une transition économique précoce ont été plus lents dans leurs évolutions, que les pays à transition économique « tardive ». Nous pouvons citer à titre d'exemple le cas du japon et surtout des « dragons » de l'est asiatique (McMillan et Rodrik, [2014], Petri [1993])

De ce point de vue, les matrices des probabilités de transition, seraient plus pertinentes par groupes de pays (homogènes quant à leur niveau de développement), qu'au niveau mondial.

Toutefois, sur la période historique étudiée, nous observons l'absence de certains sauts d'une année à l'autre : ce sont les sauts de deux états ou plus. Les seules évolutions observées sont entre des classes d'indice contigües avec des probabilités de progrès (dans le processus de transition économique) nettement plus fortes que les probabilités de recul.
Le phénomène d'inertie est également très visible par des probabilités de rester dans le même état d'une année à l'autre, qui sont toutes supérieures à 96%.

4.4. Typologie des trajectoires

L'objectif de cette section est de réaliser une typologie des trajectoires nationales, en utilisant la méthode de classification hiérarchique (méthode de Ward) des séquences basée sur la mesure d'« alignement optimal » (ou Appariement optimal, ou *Optimal Matching* OM) de proximité entre séquences (Ritschard *et al.* 2009).

4.4.1. Dendrogramme de la classification des individus selon leurs trajectoires

Le dendrogramme ci-après, suggère une classification en 4 ou 6 classes de trajectoires. Cependant, une analyse plus poussée et basée sur des statistiques adaptées, permettra d'effectuer un arbitrage.

Figure 2.11. Dendrogramme de la classification des pays selon leurs trajectoires

4.4.2. Mesure de la qualité des partitions et choix du nombre optimal de classes

Nous disposons de plusieurs mesures de la qualité des partitions selon le nombre de classes souhaité. Ces mesures de qualité des partitions, permettent de choisir le nombre optimal de classes en minimisant ou maximisant lesdites mesures (cf. colonne Min/Max dans le tableau ci-dessous).

Tableau 2.7. Mesures de la qualité d'une partition (clustering)

Nom	Abrv.	Etendue	Min/Max	Interprétation
Point Biserial Correlation	PBC	[0.61;0.87]	Max	Mesure de la capacité du clustering à reproduire les distances.
Hubert's Gamma	HG	[0.75;1]	Max	Mesure de la capacité du clustering à reproduire les distances (ordre de grandeur)
Hubert's Somers D	HGSD	[-1;1]	Max	Mesure de la capacité du clustering à reproduire les distances (ordre de grandeur) avec prise en compte des égalités sur les distances.
Hubert's C	HC	[0;0.1]	Min	Ecart entre la partition obtenue et la meilleure partition qu'il serait théoriquement possible d'obtenir avec ce nombre de groupes et ces distances.
Average Silhouette Width	ASW	[0.49;0.96]	Max	Cohérence des assignations. Une cohérence élevée indique des distances intergroupes élevées et une forte homogénéité intragroupe.
Calinski-Harabasz index	CH	[0;+∞[	Max	Pseudo F calculé à partir des distances.
Calinski-Harabasz index	CHsq	[0;+∞[	Max	Idem que précédent, mais en utilisant les distances *au carré.*
Pseudo R^2	R^2	[0.32;1]	Max	Part de la dispersion expliquée par la solution de clustering (uniquement pour comparer des partitions avec nombre de groupes identiques).
Pseudo R^2	R2sq	[0;1]	Max	Idem que précédant, mais en utilisant les distances *au carré.*

Source : Studer (2012), p. 85

Pour les formules et explications complémentaires sur les mesures de qualité de partition ci-dessus, voir Studer (2012), p. 102-104.

Les figures ci-après indiquent les courbes des principales mesures de qualité de partition des séquences en fonction du nombre de classes :

Figure 2.12. Courbes des mesures de qualité des partitions selon le nombre de classes

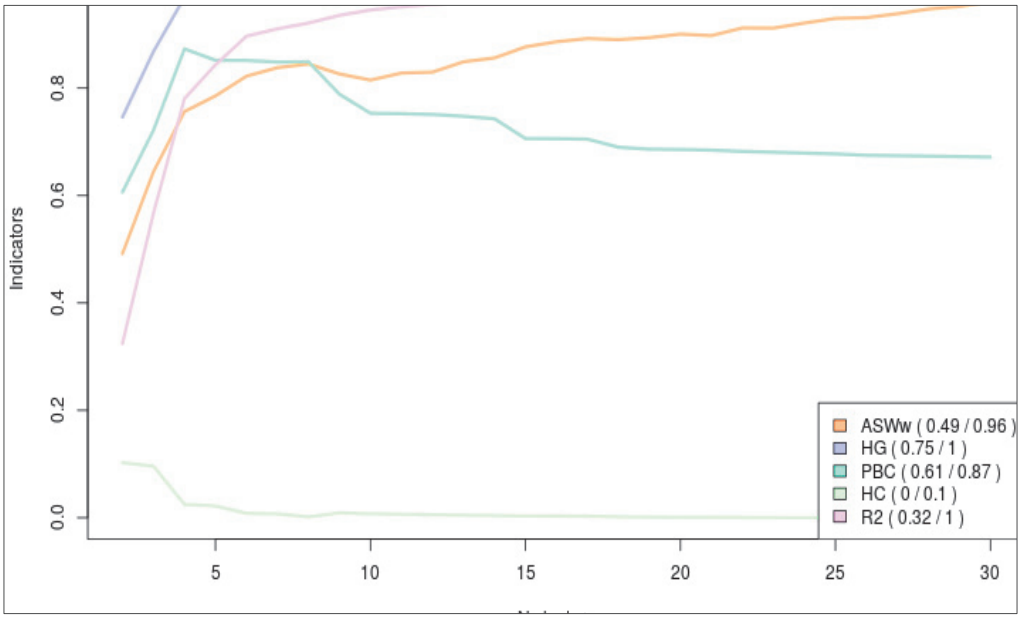
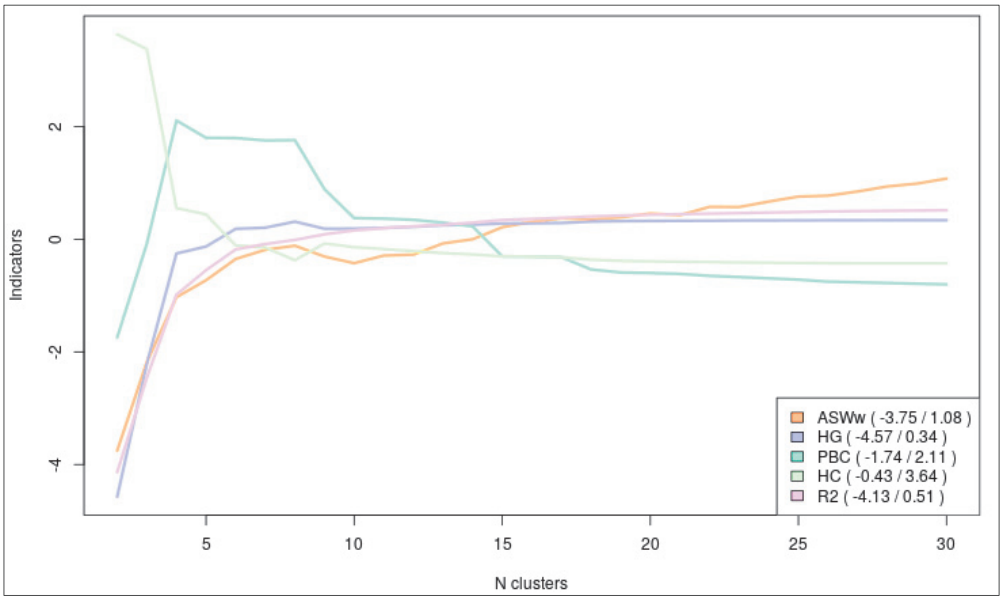


Figure 2.13. Courbes des mesures standardisées de qualité des partitions selon le nombre de classes

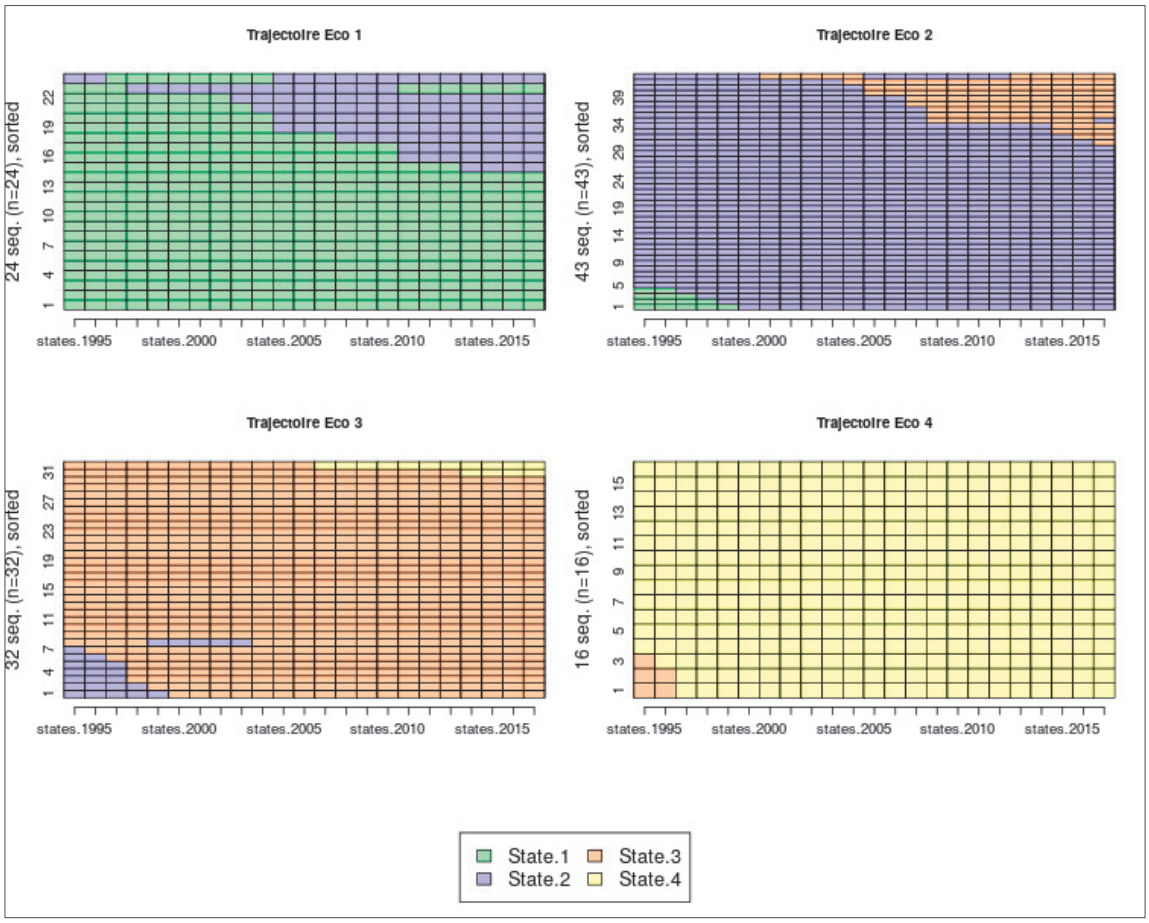


Les données ayant servi à construire les deux figures ci-dessus sont consignées dans des tableaux en annexe, respectivement aux sections 1.1 et 1.2 . Après analyse des courbes et données, et pour des raisons de robustesse, nous optons pour une classification en 4 types de trajectoires.

4.4.3. Visualisation des trajectoires au sein de chaque classe

Nous visualisons ici l'intégralité des séquences internes à chaque classe de trajectoires.

Figure 2.14. Séquences (ou trajectoires) de chaque classe, avec quadrillage selon le pays et l'année

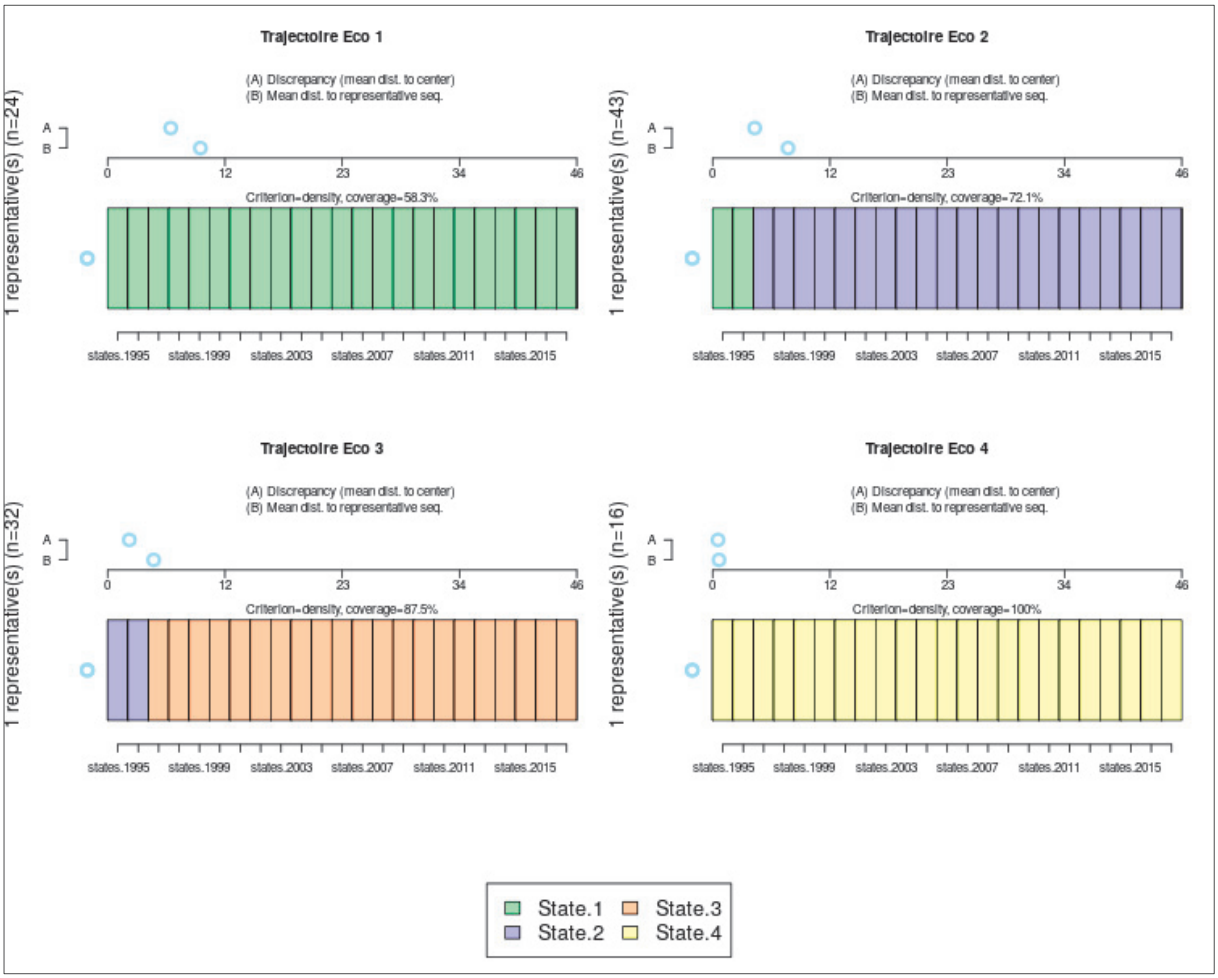


La première classe de trajectoires (*Trajectoire Eco 1*) est majoritairement caractérisée par une position permanente dans l'état 1 (State.1) et pour environ 35% des pays par un passage de State.1 à State.2.
La deuxième classe de trajectoires (*Trajectoire Eco 2*) est majoritairement caractérisée par une position stable dans l'état 2 (State.2), avec pour certains pays minoritaires, une évolution de State.2 à State.3. Et aussi pour une petite minorité, le passage de State.1 à State.2 en début d'historique.
La troisième classe de trajectoires (*Trajectoire Eco 3*) est essentiellement caractérisée par une position stable dans l'état 3 (State.3), avec deux petites minorités de pays caractérisés par une évolution de State.2 à State.3 en début d'historique, ou par le passage de State.3 à State.4.
La quatrième classe de trajectoires (*Trajectoire Eco 4*) est presque exclusivement caractérisée par une position permanente dans l'état 4 (State.4), avec une petite minorité de pays caractérisés par une évolution de State.3 à State.4 en début d'historique.
Fait notable, la classe de trajectoire qui a le plus faible effectif (24 pays) est la classe la moins avancée économiquement (*Trajectoire Eco 1)* ; le groupe le plus large (43 pays) est celui des pays ayant une trajectoire de type 2 (*Trajectoire Eco 2)* qui viennent donc de quitter la classe la moins avancée.

4.4.4. Visualisation de la trajectoire représentative de chaque classe

Le centrotype désigne la séquence représentative d'une classe : c'est la séquence dont la somme des distances aux autres séquences de la classe est minimale (Ritschard et al., 2009).
La figure ci-après illustre le centrotype de chaque classe. On y retrouve les grands traits de la description ci-dessus des classes de trajectoires.

Figure 2.15. Séquence représentative de chaque classe. Années représentées par des colonnes.



4.5. Analyse de la discrépance des séquences

L'objectif de cette section est de mesurer l'association entre les trajectoires et les caractéristiques individuelles des pays.
La discrépance mesure la variabilité des séquences. Il faudra donc mesurer la part de discrépance expliquée par une ou plusieurs caractéristiques (variables explicatives) des pays. La méthodologie utilisée est l'analyse de dispersion des séquences (Studer et al., 2011). La méthode est basée sur une généralisation de l'ANOVA (Studer et al., 2010), qui peut être étendue au cas multi-facteur et à la régression par arbre de segmentation.

4.5.1. Définition d'une mesure de la discrépance d'un ensemble de séquences en utilisant une mesure de dissimilarité

Soit le vecteur $\mathrm{y} = (\mathrm{y_i}), \mathrm{i} = 1, \dots, \mathrm{n}$.
Dans le cas euclidien, la somme des carrés SS (ou inertie) peut être exprimée en termes de distances euclidienne entre pairs :

$$SS = \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{y_i} - \bar{\mathrm{y}})^2 = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=\mathrm{i}+1}^{\mathrm{n}} \left(\mathrm{y_i} - \mathrm{y_j}\right)^2$$

$$= \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=\mathrm{i}+1}^{\mathrm{n}} \mathrm{d_{ij}}$$

En supposant que $\mathrm{d_{ij}}$ est une distance entre séquences (par ex. *Optimal Matching* OM), SS devient une mesure de dispersion des séquences.
La discrépance mesure la variabilité des trajectoires (séquences) individuelles de pays.

4.5.2. Estimation et test de la part de discrépance des séquences expliquée par une variable explicative

L'analyse ANOVA-généralisée est basée sur des dissimilarités par paires.
La somme des carrés SS définie précédemment se décompose en SS-intra (SS_W) et SS-inter (SS_B) :

$$SS_T = SS_B + SS_W$$

En reprenant la formule de la SS totale définie avec les distances $\mathrm{d_{ij}}$, on a :

$$SS_T = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=\mathrm{i}+1}^{\mathrm{n}} \mathrm{d_{ij}}$$

$$SS_W = \sum_{\mathrm{g}} \left(\frac{1}{\mathrm{n_g}} \sum_{\mathrm{i}=1}^{\mathrm{n_g}} \sum_{\mathrm{j}=\mathrm{i}+1}^{\mathrm{n_g}} \mathrm{d_{ij,g}} \right)$$

$$SS_B = SS_T - SS_W$$

g représente un groupe i.e. une modalité des individus. $\mathrm{n_g}$ est effectif du groupe g.

Tableau 2.8. ANOVA-généralisé pour m groupes (ou modalités)

	Discrepancy	*df*	Mean Discr.	*F*
Between	SS_B	$df_B = m - 1$	$\frac{SS_B}{df_B}$	$\frac{SS_B}{SS_W}\frac{df_W}{df_B}$
Within	SS_W	$df_W = \sum_g n_g - m$	$\frac{SS_W}{df_W}$	
Total	SS_T	$df_T = n - 1$		

Source : Ritschard et al. (2009)

Pseudo R^2

$$R^2 = \frac{SS_B}{SS_T}$$

Le pseudo-R^2 mesure la part de discrépance des séquences expliquée par la variable explicative ayant m modalités.

Pseudo F

$$F = \frac{SS_B/(m-1)}{SS_W/(n-m)}$$

La significativité de F est évaluée par un test de permutation, car on ne fait pas d'hypothèse a priori sur sa distribution.

Test de permutation

La procédure est la suivante (Ritschard et al., 2009) :

- Estimer la distribution empirique de F sous hypothèse d'indépendance :
  - Calculer F_{perm} la valeur F associée à une permutation aléatoire.
  - Permutation : réaffecter de manière aléatoire chaque modalité de covariable à l'une des séquences observées.
- Répéter cette étape *R* fois.
- La p-value associée au test est la probabilité $p = P(F_{obs} > F_{perm})$.
- L'intervalle de confiance de la p-value est $p \pm 1.96\sqrt{p(1-p)/R}$

Il est usuellement admis que 5000 permutations sont suffisantes pour évaluer un seuil de significativité de 1% et 1000 pour un seuil de 5%.

Analyse ANOVA-généralisée

Pour notre exemple, nous réalisons une analyse ANOVA-généralisée, en utilisant comme facteur explicatif, la zone géographique des pays.

Tableau 2.9. Estimation et test de la part de discrépance des séquences expliquée par la zone géographique du pays

Test values (p-values based on 1000 permutation)		
	t0	**p.value**
Pseudo F	4.6582108	0.001
Pseudo Fbf	5.1550218	0.001
Pseudo R2	0.2335654	0.001
Bartlett	0.2124377	0.905
Levene	0.3608388	0.924

La zone géographique du pays explique ici 23.4% (pseudo-R^2) de la variabilité des trajectoires et l'association est significative (p-value < %1).

La statistique de Levene sert à tester l'égalité des discrépances intra-groupe (Studer, 2012, p.116).

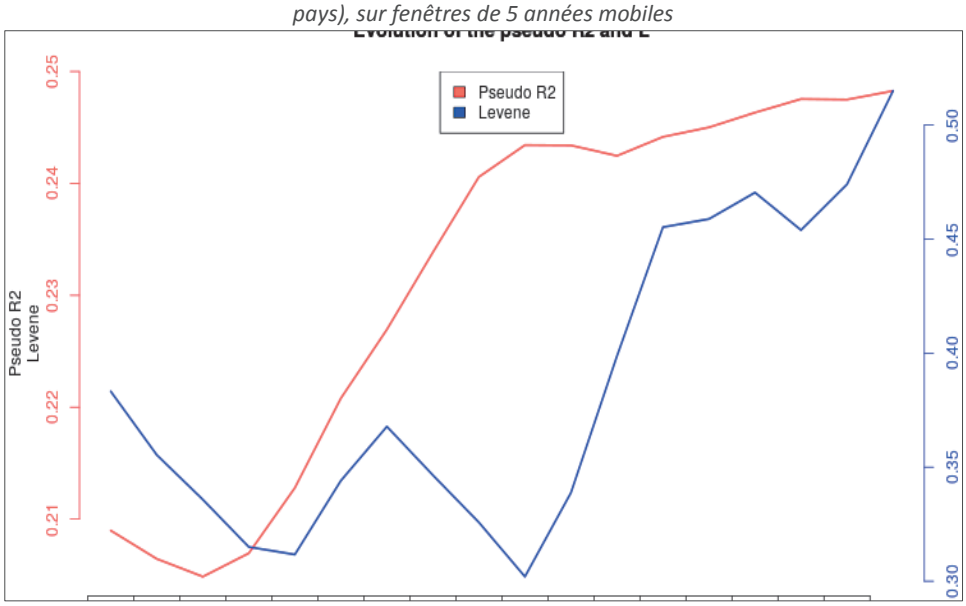
Tableau 2.10. Discrépance des séquences par modalité de la variable explicative

Discrepancy per level		
	n	**Discrepancy**
Africa	31	12.57895
Europ	24	13.36775
Middle East	7	12.20486
North America	3	12.39571
Northern Asia	14	11.17426
Oceania	3	10.22222
South America	18	11.51200
Southern Asia	15	11.79970
Total	115	15.93729

La discrépance des séquences par modalité de la variable explicative, mesure la variabilité des trajectoires au sein de chaque modalité (ici la zone géographique).
Ainsi la zone de plus forte variabilité des trajectoires des pays, est l'Europe. La zone ayant la plus grande homogénéité des trajectoires est l'Océanie.

4.5.3. Dynamique de la relation entre la discrépance des trajectoires et la variable explicative

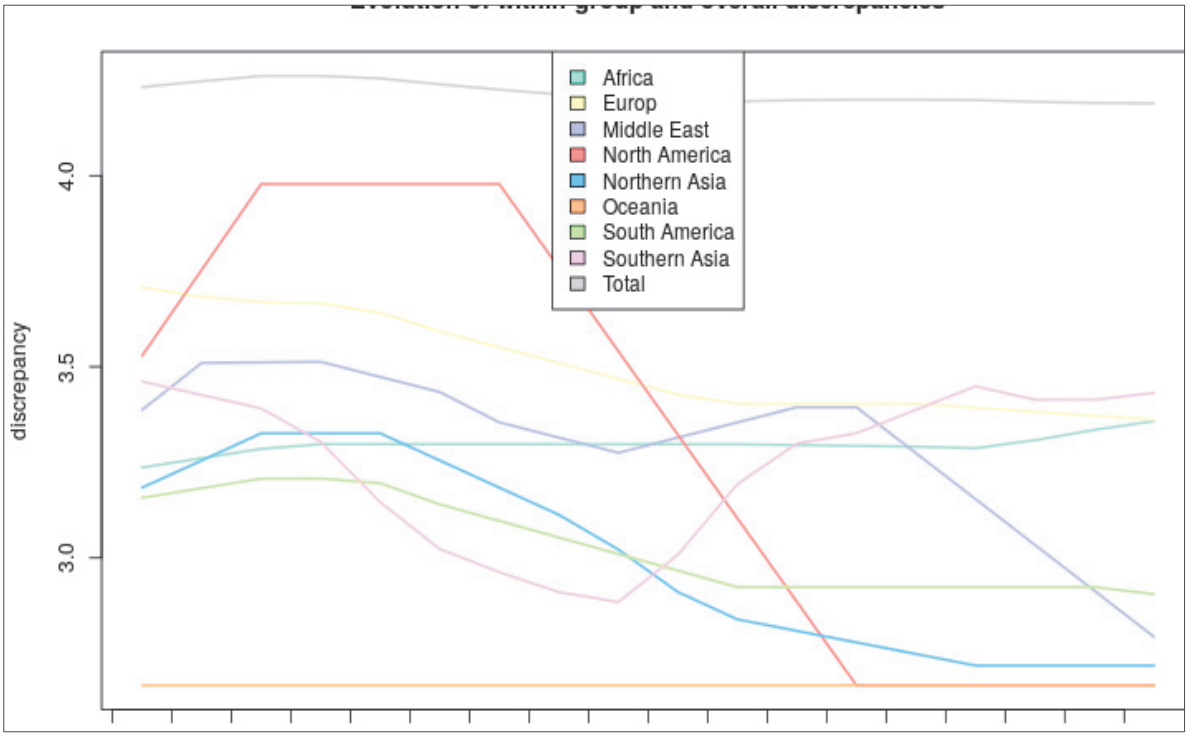
Figure 2.16. Pseudo R2 et statistique de Levene entre les trajectoires et la variable explicative (ici zone géographique des pays), sur fenêtres de 5 années mobiles



A travers la statistique de Levene, nous pouvons suivre la dynamique des dissimilarités de discrépance entre modalités (zones géographiques) de la variable explicative.

Le pseudo-R^2quant à lui, permet de suivre la dynamique de la part de discrépance (variabilité) des trajectoires (séquences d'états) expliquée par la variable explicative. Ici la hausse continue (puis la relative stabilité à partir de 2004) du pseudo-R^2 signifie une augmentation, puis une stabilité des différences de trajectoires entre régions.

Figure 2.17. Evolution des discrépances intra-groupes, sur fenêtres de 5 années mobiles



Une baisse de discrépance dans une zone indique une convergence des trajectoires au sein de la zone. Tandis qu'une hausse de la discrépance indique une divergence des trajectoires des pays de la zone.

Nous observons que la zone nord-américaine est composée des trajectoires les moins distinctes (discrépance la plus faible) et les plus stables dans le temps.
A contrario la zone sud-asiatique a la discrépance la plus forte, mais cependant relativement stable dans le temps.

4.5.4. Analyse multivariée de la discrépance des trajectoires

L'objectif de cette section est de modéliser la discrépance des trajectoires à partir de plusieurs variables explicatives. Cette approche permet également de détecter les variables explicatives les plus discriminantes, d'estimer la part de discrépance expliquée par l'ensemble des variables explicatives, et de produire des segments de pays ayant des trajectoires proches.
Le modèle utilisé est une généralisation de la régression par arbre de segmentation (Studer, 2012, p.123). À chaque étape de l'algorithme, chaque groupe (nœud) de pays est scindé en 2 par la covariable qui donne le R2 le plus élevé. La significativité de la scission est évaluée au moyen du test du pseudo-F. La croissance de l'arbre s'arrête lorsque la division sélectionnée n'est plus significative.

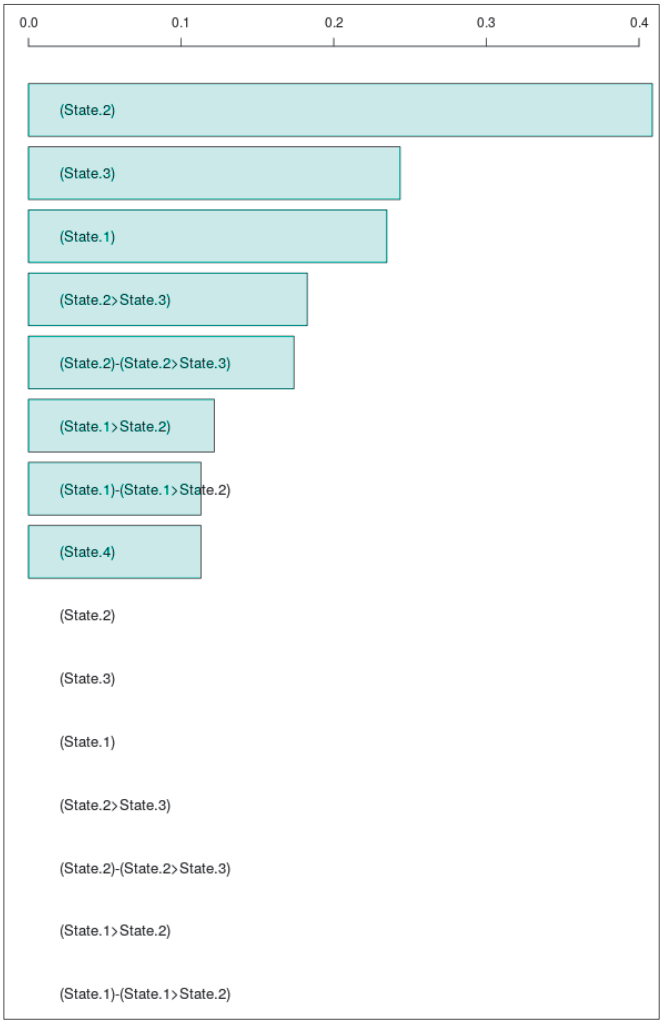
Ainsi, en utilisant les covariables, *Region, High income (yes, no), Middle income (yes, no) et Low income (yes, no)* nous obtenons une part de discrépance expliquée de 34% (pseudo R^2).
L'importance des variables pour la discrimination des trajectoires est dans l'ordre : *High income, Region, Low income.*
Les notions de *High*, *middle* et *low income* sont des catégories de la Banque Mondiale (World Bank, 2017).

5. Analyse des séquences d'évènements *(Time stamped Events analysis TSE)*

L'objectif à présent est d'étudier les séquences d'évènements, en complément des séquences d'états (STS analysis) analysées précédemment.
Pour rappel (cf. section 4.1), un *changement d'état* s'appelle *évènement*. Par exemple, la notation *état1>état2* désigne l'évènement de passage de l'état *état1* vers l'état *état2*
De même, une *transition* est un ensemble (non-ordonné) d'évènements simultanés. Un évènement est donc une transition n'ayant qu'un seul élément. La notation *(état1>état2)* désigne une transition composée du seul évènement *état1>état2*. La notation *(état1)* est équivalente à *(état1>état1).*
Une *séquence de transitions* est encore appelée *séquence d'évènements*. La notation *(état1>état2)-(état1)* désigne une séquence composée de 2 transitions ou évènements *(état1>état2)* et *(état1).*
Une sous-séquence d'une séquence *S* d'évènements, est une séquence issue de la séquence *S* où l'ordre des évènements dans *S* est respecté.
Dans notre cas, une séquence désigne une trajectoire entière, d'où sont issues des sous-séquences observées.

5.1. Découverte des sous-séquences les plus fréquentes

Figure 2.18. Les 15 sous-séquences les plus fréquentes



Nous constatons que les sous-séquences les plus fréquentes sont constituées d'un unique état (State.2, State.3, State.1). L'évènement State.2>State.3 constitue la quatrième sous-séquence la plus fréquente. La cinquième sous-séquence la plus fréquente est composée de l'état State.2, suivi dans un délai quelconque, par l'évènement State.2>State.3 : la sous-séquence est notée (State.2)-(State.2>State.3)

Huit sous-séquences sont particulièrement fréquentes parmi 115 séquences répertoriées.

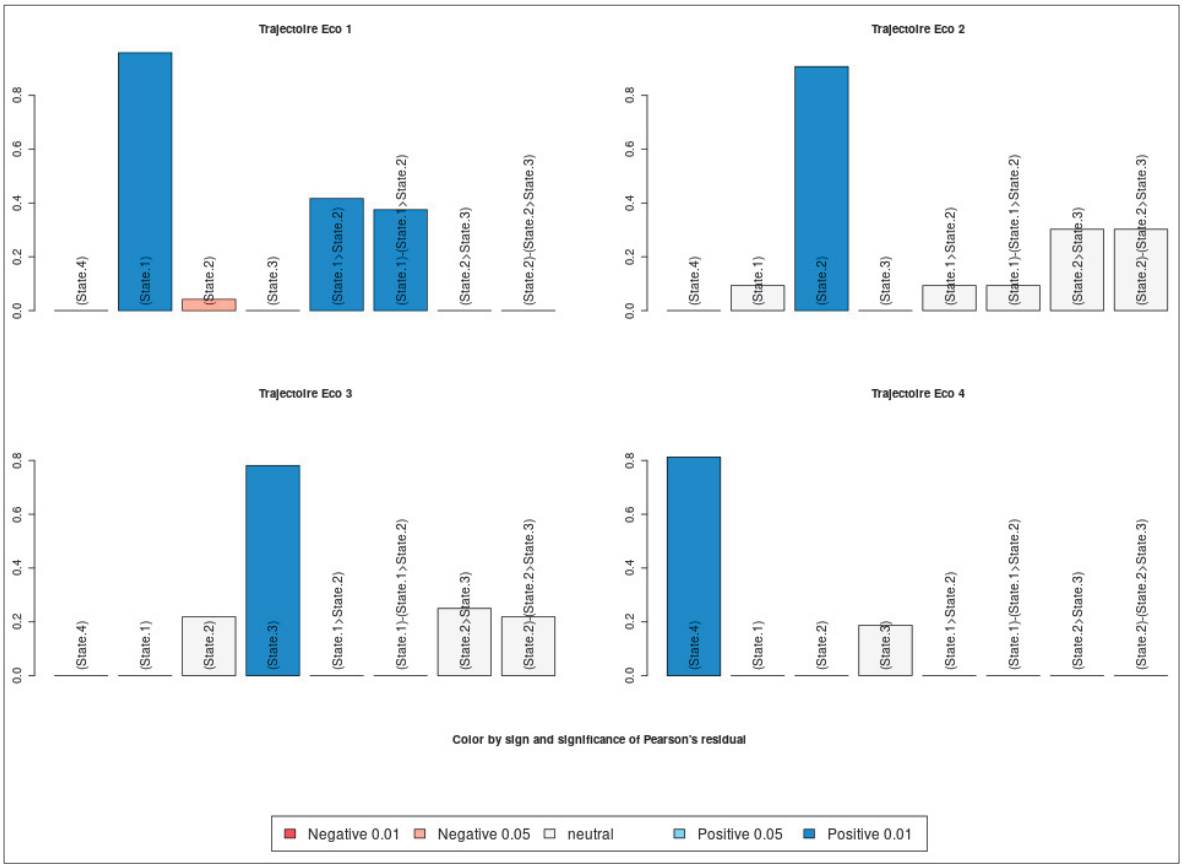
Tableau 2.11. Sous-séquences les plus fréquentes

	Subsequence	Support	Count
1	(State.2)	0.4086957	47
2	(State.3)	0.2434783	28
3	(State.1)	0.2347826	27
4	(State.2>State.3)	0.1826087	21
5	(State.2)-(State.2>State.3)	0.1739130	20
6	(State.1>State.2)	0.1217391	14
7	(State.1)-(State.1>State.2)	0.1130435	13
8	(State.4)	0.1130435	13

5.2. Sous-séquences discriminantes des classes de trajectoires

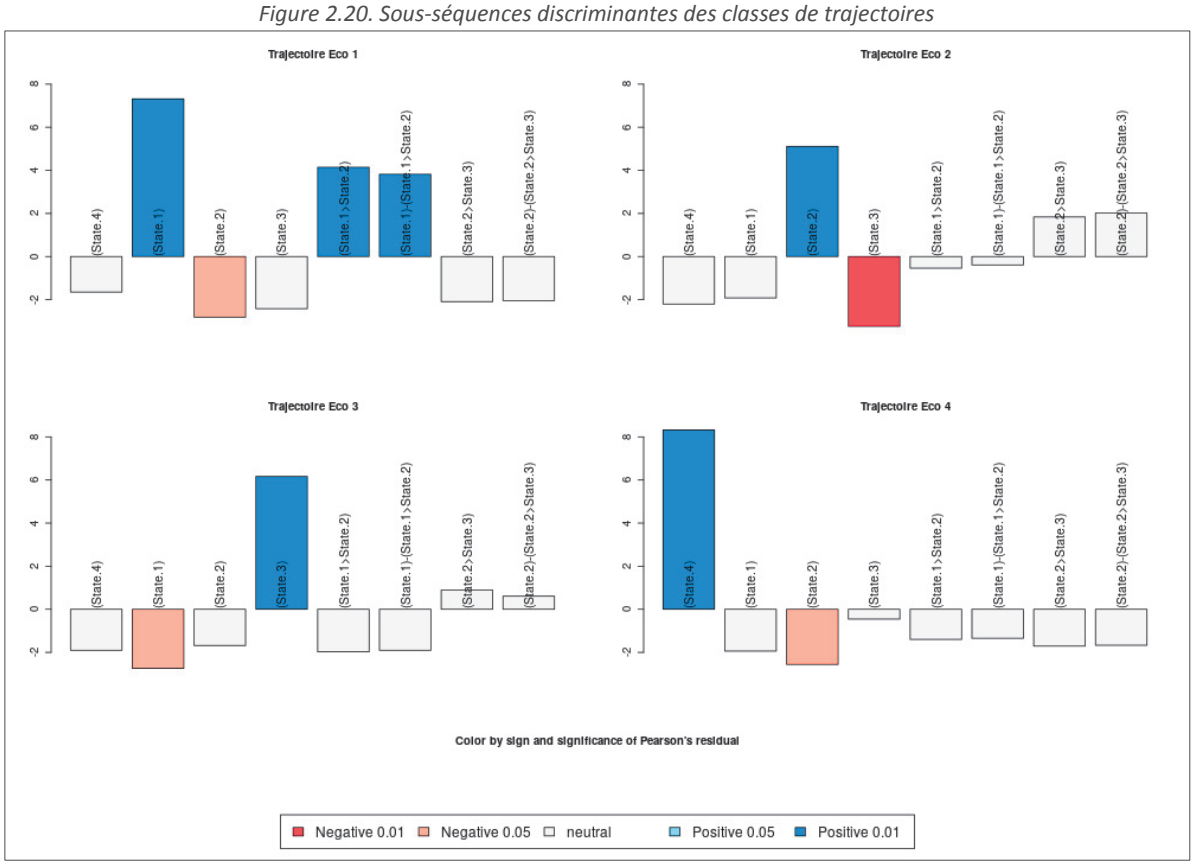
L'objectif de cette section est d'identifier les sous-séquences associées aux classes de trajectoires. Le pouvoir discriminant est déterminé par la p-value d'un test du Khi-2 (Ritschard, Bürgin et Studer, 2013). Dans les deux graphiques ci-dessous, la couleur de chaque barre est définie par le résidu de Pearson associé au test de Khi-2 : Pour les résidus inférieurs ou égaux -2 (rouge foncé), la sous-séquence est nettement moins fréquente que prévu sous hypothèse d'indépendance, tandis que pour les résidus supérieurs à 2 (bleu foncé), la sous-séquence est significativement plus fréquente.
Les fréquences relatives dans la figure ci-dessous correspondent au support relatif au sein de chaque groupe, c'est-à-dire qu'ils indiquent pour chaque groupe la proportion de séquences dans le groupe contenant la sous-séquence.

Figure 2.19. Sous-séquences discriminantes des classes de trajectoires. Fréquence relative.



Dans la figure ci-dessous, l'axe vertical représente le résidu de Pearson associé au test de Khi-2.

Figure 2.20. Sous-séquences discriminantes des classes de trajectoires



5.3. Recherche de sous-séquences et calcul des occurrences

L'analyse des séquences d'évènements (*Time stamped Events analysis TSE*) permet de rechercher et décompter le nombre de trajectoires (pays) dans lesquelles une sous-séquence est présente.
Les analyses TSE offrent également la possibilité de décompter le nombre de fois qu'une sous-séquence s'est réalisée sur la période d'histoire étudiée.
Il nous est même possible de personnaliser les sous-séquences recherchées par les paramétrages suivants :

- Le support minimum (en nombre de séquences ou trajectoires)
- Le support minimum (en pourcentage, le comptage correspondant sera arrondi) :
- Le nombre maximum d'événements autorisés dans une sous-séquence
- L'intervalle de temps maximum entre deux événements
- La durée maximale acceptée pour les sous-séquences
- La position de départ minimale autorisée pour les sous-séquences
- La position maximale de l'heure de début autorisée pour les sous-séquences
- La position de fin maximale autorisée pour les sous-séquences

Dans notre cas, il nous est parfaitement possible de connaitre le nombre de réalisation d'une sous-séquence particulière d'évènements, à l'échelle mondiale (globale) sur la période historique étudiée.

Ainsi, sur la période historique étudiée, nous avons pu établir que : l'évènement (State.1>State.2), s'est produit 14 fois, pour 14 (12%) trajectoires de pays (dont nous avons pu extraire la liste), tandis que l'évènement (State.3>State.4) ne s'est produit que dans 5 pays (4%). L'évènement intermédiaire (State.2>State.3), s'est déroulé 22 fois dans 21 (18%) de pays.
L'évènement (State.2>State.3), s'est produit 22 fois, pour 21 (18%) trajectoires de pays (dont nous avons pu établir la liste).
Comme constaté auparavant dans les autres analyses, aucun pays sur les 23 années étudiées, n'a pu faire un saut progressif de 2 états économiques successif.

5.4. Découverte des règles d'associations séquentielles

En s'inspirant des méthodes de détection de règles d'association sans prise en compte du temps (Tufféry, 2012, p.365), plusieurs chercheurs dont Müller *et al.* (2010) ont fait des adaptations pour développer une méthodologie pour la détection des règles d'associations temporelles entre sous-séquences.
Tout comme dans une règle d'association non-séquentielle, une règle d'association séquentielle, consiste à séparer une sous-séquence en une partie *prémisse* et une autre partie *conclusion*. Les deux parties se suivent temporellement dans l'ordre *prémisse*, *conclusion.*
En prenant l'exemple (Müller *et al.* 2010) d'une sous-séquence composée de trois événements, $(e_1) - (e_2) - (e_3)$, deux règles d'association séquentielles peuvent être extraites : $(e_1) \Longrightarrow (e_2) - (e_3)$ et $(e_1) - (e_2) \Longrightarrow (e_3)$.
Trois mesures, dérivées des mesures classiques de règles d'association non-séquentielles, sont utilisées pour caractériser les règles d'association séquentielles : il s'agit du support relatif, qui est égale à la probabilité qu'une séquence contienne la règle, la confiance (*confidence*), qui est la probabilité conditionnelle d'observer la conclusion sachant qu'on a observé auparavant le prémisse, et enfin le rappel (*lift*), qui désigne le ratio entre la probabilité d'observer la règle, et la probabilité d'observer la conclusion. Ce qui équivaut en continuant avec la première règle séquentielle de l'exemple ci-dessus, aux formulations suivantes :

$$Confidence\big((e_1) \Longrightarrow (e_2) - (e_3)\big) = \frac{P\big((e_1)-(e_2)-(e_3)\big)}{P\big((e_1)\big)}$$

$$Lift\big((e_1) \Longrightarrow (e_2) - (e_3)\big) = \frac{Confidence\big((e_1)\Longrightarrow(e_2)-(e_3)\big)}{P\big((e_2)-(e_3)\big)}$$

Des algorithmes spécifiques (Müller *et al.* 2010) permettent d'extraire les règles d'associations séquentielles, en se basant sur le seuil de support exigé et un niveau de confiance minimal.

Dans le tableau ci-dessous, nous présentons les règles d'associations séquentielles obtenues avec un support relatif supérieur ou égal à 1% (règles établies avec au moins 2 pays sur les 115) : l'extraction donne 8 règles séquentielles.
Nous avons choisis un support relatif particulièrement faible pour avoir un nombre plus élevé de règles. Dans la pratique, c'est le seuil de 5% qui est le plus souvent utilisé : dans notre cas, cela conduit à retenir uniquement les règles 1 et 2.

Tableau 2.12 règles d'associations séquentielles obtenues avec un support relatif supérieur ou égal à 0.01%

N°	Rules	Support	Confidence	Lift
1	(State.2) => (State.2>State.3)	20	0,426	2,3
2	(State.1) => (State.1>State.2)	13	0,481	4,0
3	(State.3) => (State.3>State.4)	5	0,179	4,1
4	(State.2) => (State.2>State.3)-(State.3>State.2)	2	0,043	2,4
5	(State.2)-(State.2>State.3) => (State.3>State.2)	2	0,100	3,8
6	(State.2) => (State.3>State.2)	2	0,043	1,6
7	(State.2>State.3) => (State.3>State.2)	2	0,095	3,7
8	(State.3>State.2) => (State.2>State.3)	2	0,667	3,7

Support = Nombre de trajectoires utilisées pour établir la règle

Ainsi, les trois premières règles séquentielles indiquent le passage d'un état économique vers l'état contigüe qui le suit dans l'ordre de la productivité agricole (ou dans l'ordre inverse du taux d'emploi agricole) : Les principales règles observées indiquent donc que les pays progressent en terme de développement économique.
Les quatre règles suivantes, bien que de support faible, évoquent toutes un faible risque d'instabilité (progression suivi de régression) entre les états State.2 et State.3. Cependant ce faible risque de régression localisé est contrebalancé par la règle N°8, qui a un niveau de confiance élevé (67%), et qui indique que la régression est provisoire.

6. Contributions des composantes structurelles macroéconomiques, à la variation du PIB par habitant

Nous restons ici, toujours dans l'analyse longitudinale, avec pour objectif d'estimer les contributions des composantes structurelles macroéconomiques, à la variation du PIB par habitant d'un pays donné.
Pour cela, nous utilisons les mêmes composantes structurelles de l'économie que nous avons utilisées tout au long de ce chapitre, d'abord pour la détection des phases (section 3.1) qui ont servi de base à l'analyse séquentielle par la suite : la population en âge de travailler, le taux d'emploi, la part d'emploi de chaque secteur, les productivités sectorielles (agriculture et pêche, industrie et services).

Les contributions des composantes structurelles sont calculées séquentiellement (par année) sur une période définie, puis aggrégées sur l'ensemble de la période.
Pour effectuer ces calculs, nous-nous basons naturellement sur l'équation comptable qui relie le PIB par habitant aux composantes structurelles dont nous souhaitons mesurer les contributions.
Nous avions déjà formulée cette équation au chapitre 1, section 3.2 :

$$\frac{PIB}{Pop} = \sum_{i=1}^{3} \left(\frac{Pop_{15-64}}{Pop}\right) . \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right) . \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right) . \left(\frac{PIB_i}{E_i}\right)$$

où :

Pop est la population totale
Pop_{15-64} est la population âgée de 15 à 64 ans
E_i est le nombre d'emplois dans le secteur i
$\sum_{u=1}^{3} E_u$ est le nombre total d'emplois dans les trois secteurs
PIB_i est le PIB du secteur i

Les composants (entre parenthèses) de l'équation sont : la proportion de la population en âge de travailler (15-64 ans), le taux d'emploi de la population en âge de travailler, les variables transitionnelles classiques (i.e. taux d'emploi dans les différents secteurs), la productivité par emploi, dans chaque secteur.

Notre méthode est donc basée sur une équation comptable (toujours vérifiée), et aboutit à des contributions calculées par une formule mathématique explicite.
Notre approche diffère d'un modèle économétrique (de forme Cobb-Douglas par exemple) avec quelques variables supposées expliquer la croissance économique, d'après une théorie que l'on souhaiterait tester.
A la différence avec les modèles économétriques (comme par exemple, les modèles de Solow ou Carré-Dubois-Malinvaud [1984]), notre approche est plutôt une démarche de comptabilité nationale inspirée de Kuznets.

Les développements mathématiques de cette section à partir de l'équation ci-dessus, permettent d'estimer l'élasticité (sensibilité) instantanée (annuelle) du PIB par habitant par rapport aux composants de l'équation, ainsi que les contributions de ces composants à la variation sur une période pluriannuelle, du PIB par habitant.

Nous étudions ici, la contribution des variables exogènes à l'évolution de la variable endogène lorsque le modèle est de la forme (ce qui est le cas de l'équation du PIB par habitant développée plus haut) :

$$Y_t = \sum_{i=1}^{p} \prod_{j=1}^{q} X_{i,j,t} \qquad (1)$$

où
Y_t est la variable endogène à la date t ;
$X_{i,j,t}$ des variables exogènes continues portant deux indices $i \in \{1, \cdots, \mathrm{p}\}$ et $j \in \{1, \cdots, \mathrm{q}\}$, de numérotation à la date t telle que $X_{i,j,t} > 0$

Considérons l'intervalle de temps $I_k = [t, t+k]$ où k est un entier naturel. Notons :

$$\Delta_k Y_t = Y_{t+k} - Y_t$$

<u>Définition 1</u>
Nous définissons la *variation relative* de la variable Y sur l'intervalle de temps I_k comme étant la quantité :

$$\frac{\Delta_{\mathrm{k}} Y_t}{Y_t} \times 100 = \frac{Y_{t+k} - Y_t}{Y_t} \times 100 \qquad (2)$$

<u>Définition 2</u>
La contribution en points de pourcentage de la variable $X_{i,j}$ $\forall(i,j)$ à la variation relative de Y $\left(soit\ \frac{\Delta_k Y}{Y} \times 100\right)$ sur l'intervalle de temps I_k, est le nombre de points de ce pourcentage, attribuable à la variable exogène $X_{i,j}$.
Cette contribution est notée $cp(X_{i,j})$.
L'un des objectifs du présent développement est de calculer $cp(X_{i,j})$.

<u>Définition 3</u>
Soit $j^* \in \{1, \cdots, q\}$ la contribution de la variable $X_{.j^*}$ à la variation relative de Y sur l'intervalle de temps I_k est la contribution en points de pourcentage de toutes les variables exogènes $\left(X_{i,j^*}\right)_i$ c'est-à-dire toutes variables exogènes portant les indices j^*. Nous notons $cp(X_{.j^*})$ cette contribution.

<u>Définition 4</u>
Soit $i^* \in \{1, \cdots, p\}$ la contribution de la variable $X_{i^*.}$ à la variation relative de Y sur l'intervalle de temps I_k est la contribution en points de pourcentage de toutes les variables exogènes $\left(X_{i^*,j}\right)_j$ i.e. toutes variables exogènes portant les indices i^*. Nous notons $cp(X_{i^*.})$ cette contribution.

<u>Définition 5</u>
La contribution relative de la variable $X_{i,j}$ $\forall(i,j)$ à la variation relative de Y $\left(soit\ \frac{\Delta_k Y}{Y} \times 100\right)$ sur l'intervalle de temps I_k, est le pourcentage de cette variation relative, attribuable à la variable exogène $X_{i,j}$.
Cette contribution est notée $CR(X_{i,j})$.
Sa formule est donnée par :

$$CR\left(X_{i,j}\right) = \frac{cp(X_{i,j})}{\Delta_k Y_t} \times \frac{Y_t}{100}$$

Définition 6

Soit $j^* \in \{1, \cdots, q\}$, la contribution relative de la variable $X_{.j^*}$ à la variation relative de Y sur l'intervalle de temps I_k, est le pourcentage de cette variation relative attribuable à toutes les variables exogènes $\left(X_{i,j^*}\right)_i$ i.e. toutes variables exogènes portant les indices j^*.

Nous notons $CR(X_{.j^*})$ cette contribution relative.

Sa formule est donnée par :

$$CR\left(X_{.j^*}\right) = \frac{cp(X_{.j^*})}{\Delta_k Y_t} \times \frac{Y_t}{100}$$

Définition 7

Soit $i^* \in \{1, \cdots, p\}$, la contribution relative de la variable $X_{i^*.}$ à la variation relative de Y sur l'intervalle de temps I_k, est le pourcentage de cette variation relative attribuable à toutes les variables exogènes $\left(X_{i^*,j}\right)_j$ i.e. toutes variables exogènes portant les indices i^*.

Nous notons $CR(X_{i^*.})$ cette contribution relative.

Sa formule est donnée par :

$$CR(X_{i^*.}) = \frac{cp(X_{i^*.})}{\Delta_k Y_t} \times \frac{Y_t}{100}$$

Définition 8

On appelle élasticité d'une variable endogène Y par rapport à une variable exogène $X_{i,j}$ $\forall(\mathrm{i},\mathrm{j})$ noté $e_{y/x}$ le rapport entre la variation relative de Y et la variation relative de X c'est-à-dire

$$e_{y/x} = \frac{\frac{\Delta_k Y}{Y}}{\frac{\Delta_k X_{i,j}}{X_{i,j}}} ou \quad \frac{\Delta_k Y}{Y} = e_{y/x} \times \frac{\Delta_k X_{i,j}}{X_{i,j}}$$

Hypothèses 1

$\forall\, t \in \mathbb{N}, \forall\, i \in \{1, \cdots, \mathrm{p}\}$ et $\forall\, j \in \{1, \cdots, \mathrm{q}\}$, nous supposons dans un premier temps, une valeur de k suffisamment petite, telle que :

a) le quotient $\frac{X_{i,j,t+k}}{X_{i,j,t}}$ proche de 1

b) le quotient $\frac{S_{i,t+k}}{S_{i,t}}$ est proche de 1, avec $S_{i,t} = \prod_{j=1}^{q} X_{i,j,t}$

Y_t devient :

$$Y_t = \sum_{i=1}^{p} S_{i,t} \qquad \text{et}$$

$$\frac{\Delta_k Y_t}{Y_t} = \frac{\sum_{i=1}^{p} S_{i,t+k} - \sum_{i=1}^{p} S_{i,t}}{Y_t} = \frac{\sum_{i=1}^{p}\left(S_{i,t+k} - S_{i,t}\right)}{Y_t}$$

$$= \sum_{i=1}^{p} \frac{S_{i,t+k} - S_{i,t}}{Y_t} \qquad \text{(Car } Y_t \text{ est constante à la période } t \text{ fixée)}$$

$$\frac{\Delta_k Y_t}{Y_t} = \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \times \frac{S_{i,t+k} - S_{i,t}}{S_{i,t}}$$

$$= \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \times \left(\frac{S_{i,t+k}}{S_{i,t}} - 1\right)$$

D'après le b) des hypothèses 1, on a :

$\ln\left(\frac{S_{i,t+k}}{S_{i,t}}\right) \approx \frac{S_{i,t+k}}{S_{i,t}} - 1$ (Car $\ln(x) \approx x - 1$ lorsque x tend vers 1)

On en déduit que :

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \times \ln\left(\frac{S_{i,t+k}}{S_{i,t}}\right)$$

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \times \left(\ln(S_{i,t+k}) - \ln(S_{i,t})\right)$$

$$\approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \left(\ln\left(\prod_{j=1}^{q} X_{i,j,t+k}\right) - \ln\left(\prod_{j=1}^{q} X_{i,j,t}\right)\right)$$

$$\approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \left(\sum_{j=1}^{q} \ln(X_{i,j,t+k}) - \sum_{j=1}^{q} \ln(X_{i,j,t})\right)$$

$$\approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \left[\sum_{j=1}^{q} \left(\ln(X_{i,j,t+k}) - \ln(X_{i,j,t})\right)\right]$$

$$\approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \times \left[\sum_{j=1}^{q} \ln\left(\frac{X_{i,j,t+k}}{X_{i,j,t}}\right)\right]$$

En utilisant l'hypothèse 1 – a), on obtient

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \times \sum_{j=1}^{q} \left(\frac{X_{i,j,t+k}}{X_{i,j,t}} - 1\right)$$

Et donc

$$\boxed{\frac{\Delta_k Y_t}{Y_t} \approx \sum_{i=1}^{p} \frac{S_{i,t}}{Y_t} \left[\sum_{j=1}^{q} \frac{\Delta_k X_{i,j,t}}{X_{i,j,t}}\right]} \qquad (3)$$

Les élasticités de la variable endogène Y, par rapport à l'évolution de la variable exogène $X_{i,j,t}$, $1 \leq j \leq q$ et $1 \leq j \leq q$ (i, j) fixés, sont données par les quantités $\frac{S_{i,t-k}}{Y_{t-k}}$, et à une date donnée.

Décomposition du PIB par habitant

Reprenons la décomposition du PIB par habitant :

$$\boxed{\left(\frac{PIB}{Pop}\right) = \sum_{i=1}^{3}\left(\frac{Pop_{15-64}}{Pop}\right)\times\left(\frac{\sum_{u=1}^{3}E_u}{Pop_{15-64}}\right)\times\left(\frac{E_i}{\sum_{u=1}^{3}E_u}\right)\times\left(\frac{PIB_i}{E_i}\right)} \qquad (4)$$

Comme nous l'avons souligné dans l'introduction, cette décomposition est sous la forme du modèle (1) :

$$Y_t = \sum_{i=1}^{p}\prod_{j=1}^{q}X_{i,j,t}$$

Où le facteur étudié ici est le secteur d'activité de chaque pays.

En considérant la tranche d'âge de la population en âge de travailler (15-64 ans), la décomposition (4) devient :

$$Y_{t,15-64} = \sum_{i=1}^{3}\prod_{j=1}^{4}X_{i,j,t}$$

où

$$Y_{t,15-64} = \left(\frac{PIB}{Pop}\right)_t, \qquad X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, \qquad X_{i,2,t} = \left(\frac{\sum_{u=1}^{3}E_u}{Pop_{15-64}}\right)_t,$$

$$X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3}E_u}\right)_t \ et \ X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

On a établi au (1) que :

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{i=1}^{3}\frac{S_{i,t}}{Y_t}\left[\sum_{j=1}^{4}\frac{\Delta_k X_{i,j,t}}{X_{i,j,t}}\right]$$

$$\frac{\Delta_k Y_t}{Y_t} \approx \frac{\Delta_k X_{i,1,t}}{X_{i,1,t}} + \frac{\Delta_k X_{i,2,t}}{X_{i,2,t}} + \sum_{i=1}^{3}\left(\frac{S_{i,t}}{Y_t}\right)\left[\sum_{j=3}^{4}\frac{\Delta_k X_{i,j,t}}{X_{i,j,t}}\right]$$

On a

$$\boxed{\frac{\Delta_k Y_{t,15-64}}{Y_{t,15-64}} \approx \frac{\Delta_k X_{i,1,t}}{X_{i,1,t}} + \frac{\Delta_k X_{i,2,t}}{X_{i,2,t}} + \sum_{i=1}^{3}\left(\frac{S_{i,t}}{Y_t}\right)\frac{\Delta_k X_{i,3,t}}{X_{i,3,t}} + \sum_{i=1}^{3}\left(\frac{S_{i,t}}{Y_t}\right)\frac{\Delta_k X_{i,4,t}}{X_{i,4,t}}} \qquad (5)$$

où

$$Y_{t,15-64} = \left(\frac{PIB}{Pop}\right)_t, \qquad X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, X_{i,2,t} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)_t,$$

$$X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right)_t, \qquad X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

et

$$\frac{S_{i,t}}{Y_t} = \frac{\prod_{j=1}^{4} X_{i,j,t}}{Y_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_t}{\left(\frac{PIB}{Pop}\right)_t} = \left(\frac{PIB_i}{PIB}\right)_t$$

on en déduit les élasticités instantanées pour tout $i, 1 \leq i \leq 3$,

$$\frac{\left(\frac{\Delta_k Y_{t,15-64}}{Y_{t,15-64}}\right)}{\left(\frac{\Delta_k X_{i,1,t}}{X_{i,1,t}}\right)} \approx 1, \frac{\left(\frac{\Delta_k Y_{t,15-64}}{Y_{t,15-64}}\right)}{\left(\frac{\Delta_k X_{i,2,t}}{X_{i,2,t}}\right)} \approx 1$$

Et pour $j = 3,4$, on a :

$$\frac{\left(\frac{\Delta_k Y_{t,15-64}}{Y_{t,15-64}}\right)}{\left(\frac{\Delta_k X_{i,j,t}}{X_{i,j,t}}\right)} \approx \left(\frac{PIB_i}{PIB}\right)_t$$

où

$$Y_{t,15-64} = \left(\frac{PIB}{Pop}\right)_t, \qquad X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, X_{i,2,t} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)_t,$$

$$X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right)_t, \qquad X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

et

$$\frac{S_{i,t}}{Y_t} = \frac{\prod_{j=1}^{4} X_{i,j,t}}{Y_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_t}{\left(\frac{PIB}{Pop}\right)_t} = \left(\frac{PIB_i}{PIB}\right)_t$$

<u>Remarque</u> : Sur une période de temps assez longue, la notion d'élasticité serait insuffisante pour expliquer la variation de la variable endogène par les variables exogènes car elle ne prend pas en compte l'évolution des variables sur plusieurs années (notamment à cause des *Hypothèses 1*, qui ne seront plus respectées).

Généralisation de la décomposition sur une longue période

Considérons un intervalle de temps, $[t, t+k] = I_k$ avec k un entier naturel.

<u>Hypothèses 2</u>

$\forall\, t \in \mathbb{N}, \forall\, i \in \{1, \cdots, \mathrm{p}\}$ et $\forall\, j \in \{1, \cdots, \mathrm{q}\}$, nous supposons dans ce deuxième temps, une valeur de k suffisamment grande telle que :

a) le quotient $\frac{X_{i,j,t+k}}{X_{i,j,t}}$ soit différent de 1

b) le quotient $\frac{S_{i,t+k}}{S_{i,t}}$ soit différent de 1, avec $S_{i,t} = \prod_{j=1}^{q} X_{i,j,t}$

Les hypothèses 2 traduisent le fait que l'écart entre $X_{i,j,t}$ et $X_{i,j,t+k}$ d'une part et l'écart entre $S_{i,j,t}$ et $S_{i,j,t+k}$ sont importants, et donc la transformation directe en somme de Y_t n'est plus possible. Dans ce cas nous subdivisons l'intervalle I_k tel que les hypothèses 1 soient vérifiées sur chaque subdivision.

Considérons une subdivision de l'intervalle I_k en n intervalles d'amplitude $\frac{k}{n}$ tel que $\frac{k}{n}$ soit un entier naturel.

On a

$$\Delta_k Y_t = Y_{t+k} - Y_t$$

$$= \left(Y_{t+k} - Y_{t+\frac{(n-1)k}{n}}\right) + \left(Y_{t+\frac{(n-1)k}{n}} - Y_{t+\frac{(n-2)k}{n}}\right) + \cdots + \left(Y_{t+\frac{k}{n}} - Y_t\right)$$

$$\Delta_k Y_t = \sum_{l=1}^{n} \left(Y_{t+\frac{(n-l+1)k}{n}} - Y_{t+\frac{(n-l)k}{n}}\right)$$

donc $$\frac{\Delta_k Y_t}{Y_t} = \frac{\sum_{l=1}^{n}\left(Y_{t+\frac{(n-l+1)k}{n}} - Y_{t+\frac{(n-l)k}{n}}\right)}{Y_t}$$

$$= \sum_{l=1}^{n} \frac{Y_{t+\frac{(n-l+1)k}{n}} - Y_{t+\frac{(n-l)k}{n}}}{Y_t}$$

$$= \sum_{l=1}^{n} \frac{Y_{t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{Y_{t+\frac{(n-l+1)k}{n}} - Y_{t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}}$$

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{l=1}^{n} \frac{Y_{t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k Y_{t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}}$$

D'après la relation (3) on a :

$$\frac{\Delta_k Y_{t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}} \approx \sum_{i=1}^{p} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}} \times \left[\sum_{j=1}^{q} \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}\right]$$

Donc

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{l=1}^{n} \frac{Y_{t+\frac{(n-l)k}{n}}}{Y_t} \times \left[\sum_{i=1}^{p} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}} \times \left(\sum_{j=1}^{q} \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}\right)\right]$$

$$\approx \sum_{l=1}^{n}\sum_{i=1}^{p}\sum_{j=1}^{q} \frac{Y_{t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}$$

$$\approx \sum_{i=1}^{p}\sum_{j=1}^{q}\left(\sum_{l=1}^{n} \frac{Y_{t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_{t+\frac{(n-l)k}{n}}} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}\right)$$

$$\frac{\Delta_k Y_t}{Y_t} \approx \sum_{j=1}^{q}\sum_{i=1}^{p}\left(\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}\right)$$

$$\boxed{\frac{\Delta_k Y_t}{Y_t} \approx \sum_{j=1}^{q}\sum_{i=1}^{p}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}} \qquad (6)$$

Soit $j \in \{1, \cdots, q\}$ tel que pour tout $i, i' \in \{1, \cdots, p\}, i \neq i'$, on a

$$X_{i,j,t} = X_{i',j,t} \qquad (7)$$

Nous notons G_1 l'ensemble des variables $X_{i,j}$ vérifiant la relation (7) et par G_2 le complémentaire de G_1 c'est-à-dire

$$G_1 = \{X_{i,j} | \, X_{i,j} \perp i \, , \forall \, i, j\}$$

et

$$G_2 = \left\{X_{i,j} \mid \neg\left(X_{i,j} \perp (i,j)\right), \forall \, i, j\right\}$$

<u>Remarque</u> : $G_1 \cup G_2$ est donc l'ensemble de toutes les variables dans le modèle (1).

De la relation (6), suivant que $X_{i,j}$ appartiennent à G_1 ou à G_2, on en déduit les contributions en points de pourcentage suivantes :

$$cp(X_{i,j}) \approx \sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \times 100 \qquad \text{avec } X_{i,j} \in G_2$$

$$cp(X_{.,j}) \approx \sum_{i=1}^{p}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \times 100 \qquad \text{avec } X_{i,j} \in G_1$$

$$cp(X_{i,.}) \approx \sum_{X_{i,j}\in G_2}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \times 100 \qquad \text{avec } X_{i,j} \in G_2$$

Puis la contribution relative de $X_{i,j}$ (respectivement de $X_{i,.}$ *e*t $X_{.,j}$) par :

$$CR(X_{i,j}) = \frac{cp(X_{i,j})}{\Delta_k Y_t} \times \frac{Y_t}{100} \approx \sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \qquad \text{avec } X_{i,j} \in G_2$$

$$CR(X_{.,j}) = \frac{cp(X_{.,j})}{\Delta_k Y_t} \times \frac{Y_t}{100} \approx \sum_{i=1}^{p}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \qquad \text{avec } X_{i,j} \in G_1$$

$$CR(X_{i,.}) = \frac{cp(X_{i,.})}{\Delta_k Y_t} \times \frac{Y_t}{100} \approx \sum_{X_{i,j}\in G_2}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \qquad \text{avec } X_{i,j} \in G_2$$

Application au modèle (4)
On a

$$G_1 = \{X_{i,1}, X_{i,2}\} \text{ avec } X_{i,1} = \left(\frac{Pop_{15-64}}{Pop}\right), X_{i,2} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)$$

et

$$G_2 = \{X_{i,3}, X_{i,4}\} \text{ avec } X_{i,3} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right), X_{i,4} = \left(\frac{PIB_i}{E_i}\right)$$

La contribution en points de pourcentage de $X_{i,j}$ (Secteur $\boldsymbol{i}$, variable $\boldsymbol{j}$) appartenant à G_2 est donnée par :

$$cp(X_{i,j}) \approx \sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \times 100$$

Avec

$$\frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} = \frac{\prod_{j=1}^{q} X_{i,j,t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_{t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t},$$

$$X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, X_{i,2,t} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)_t, X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right)_t \quad et$$

$$X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

La contribution **relative** de $X_{i,j}$ appartenant à G_2 est donnée par :

$$CR(X_{i,j}) = \frac{cp(X_{i,j})}{\Delta_k Y_t} \times \frac{Y_t}{100} \approx \sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}$$

Avec :

$$\frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} = \frac{\prod_{j=1}^{q} X_{i,j,t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_{t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t},$$

$$X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, X_{i,2,t} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)_t, X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right)_t \; et$$

$$X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

La contribution $cp(X_{i,j})$ est la contribution de la variable d'indice j du secteur i dans la variation relative de $Y_{15-64} = \left(\frac{PIB}{Pop}\right)$ et $CR(X_{i,j})$ sa contribution relative.

Respectivement on a la contribution en points de pourcentage et la contribution relative de $X_{i,\cdot}$ et $X_{\cdot,j}$:

$$cp(X_{i,\cdot}) \approx \sum_{X_{i,j} \in G_2} \sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \times 100 \qquad \text{avec } X_{i,j} \in G_2$$

$$CR(X_{i,\cdot}) = \frac{cp(X_{i,\cdot})}{\Delta_k Y_t} \times \frac{Y_t}{100} \approx \sum_{X_{i,j} \in G_2} \sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \qquad \text{avec } X_{i,j} \in G_2$$

Avec :

$$\frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} = \frac{\prod_{j=1}^{q} X_{i,j,t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_{t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t},$$

$$\frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} = \frac{\prod_{j=1}^{q} X_{i,j,t+\frac{(n-l)k}{n}}}{\Delta_k \left(\frac{PIB}{Pop}\right)_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_{t+\frac{(n-l)k}{n}}}{\Delta_k \left(\frac{PIB}{Pop}\right)_t},$$

$$X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, X_{i,2,t} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)_t, X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right)_t \quad et$$

$$X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

La contribution en points de pourcentage $cp(X_{i,\cdot})$ est donc la contribution du secteur i dans la variation relative de $Y_{15-64} = \left(\frac{PIB}{Pop}\right)$ et $CR(X_{i,\cdot})$ sa contribution relative.
De même on a :

$$cp(X_{\cdot,j}) \approx \sum_{i=1}^{p}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}} \times 100$$

$$CR(X_{\cdot,j}) = \frac{cp(X_{\cdot,j})}{\Delta_k Y_t} \times \frac{Y_t}{100} \approx \sum_{i=1}^{p}\sum_{l=1}^{n} \frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} \times \frac{\Delta_k X_{i,j,t+\frac{(n-l)k}{n}}}{X_{i,j,t+\frac{(n-l)k}{n}}}$$

Avec :

$$\frac{S_{i,t+\frac{(n-l)k}{n}}}{Y_t} = \frac{\prod_{j=1}^{q} X_{i,j,t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_{t+\frac{(n-l)k}{n}}}{\left(\frac{PIB}{Pop}\right)_t},$$

$$\frac{S_{i,t+\frac{(n-l)k}{n}}}{\Delta_k Y_t} = \frac{\prod_{j=1}^{q} X_{i,j,t+\frac{(n-l)k}{n}}}{\Delta_k\left(\frac{PIB}{Pop}\right)_t} = \frac{\left(\frac{PIB_i}{Pop}\right)_{t+\frac{(n-l)k}{n}}}{\Delta_k\left(\frac{PIB}{Pop}\right)_t},$$

$$X_{i,1,t} = \left(\frac{Pop_{15-64}}{Pop}\right)_t, X_{i,2,t} = \left(\frac{\sum_{u=1}^{3} E_u}{Pop_{15-64}}\right)_t, X_{i,3,t} = \left(\frac{E_i}{\sum_{u=1}^{3} E_u}\right)_t \quad et$$

$$X_{i,4,t} = \left(\frac{PIB_i}{E_i}\right)_t$$

La contribution en points de pourcentage $cp(X_{\cdot,j})$ est donc la contribution de toutes les j-ieme variables de tous les secteurs dans la variation relative de $Y_{15-64} = \left(\frac{PIB}{Pop}\right)$ et $CR(X_{\cdot,j})$ leur contribution relative.

Remarque :
De façon générale,

$$cp(X_{i,\cdot}) = \sum_{X_{i,j}\in G_2} cp(X_{i,\cdot j}) \;\; si \;\; X_{i,j} \in G_2 \;\; et \;\; cp(X_{\cdot,j}) = \sum_{i=1}^{p} cp(X_{i,\cdot j}) \;\; si \; X_{i,j} \in G_1$$

A titre d'exemple synthétique, nous avons appliqué les formules à la France sur une période de 10 ans de 2007 à 2016.
Le total des contributions en points de pourcentage est de -10% pour l'ensemble des composants. Or le PIB par habitant a varié de -11.2% sur la même période. Le taux de couverture des contributions estimées par notre méthode est donc de 89.7%.
Etant donné que les contributions sont calculées à partir d'une équation comptable fermée aux seules composantes indiquées dans la formule, nous pouvons redresser les contributions calculées de sorte à ce que leur somme corresponde à la variation effective du PIB par habitant.
Après redressement, nous obtenons les contributions pour les composantes suivantes : secteur agricole, secteur industriel, secteur des services, population en âge de travailler, taux d'emploi de la population en âge de travailler

	contribution en points de pourcentage
Service	-4,012
Population.ages.15.64.ratio.to.total.population.20	-3,906
Industry	-3,256
Agriculture	-0,486
Employment.to.population.ratio.ages.15.64.20	0,486
Total	**-11,174**
3 secteurs	**-7,754**

Il apparait que la dynamique démographique de la population en âge de travailler a contribué à une baisse de -3.9% du PIB par habitant sur la période étudiée, soit 31.4% de la baisse du PIB par habitant sur la période.
A contrario, la dynamique du taux d'emploi de la population en âge de travailler a eu un effet légèrement positif de +0.49% sur le PIB par habitant.
Les trois secteurs économiques ont donc contribué à une baisse de -7.8% du PIB par habitant (qui a lui-même baissé de -11.2%).
Or tous les secteurs n'ont pas le même poids dans l'économie, que ce soit en parts d'emplois ou en parts de valeur ajoutée (77% pour les services, 21% pour l'industrie et 2% pour l'agriculture en 2007).
En pondérant les contributions des secteurs économiques par leurs parts de valeur ajoutée dans l'économie en début de période (2007), et en redressant les valeurs pour avoir le même total de points de contribution, nous obtenons une inversion des contributions :

	contribution en points de pourcentage
Service	-0,851
Industry	-2,490
Agriculture	-4,412
Total	**-7,754**

Ce résultat signifie qu'à parts de valeur ajoutée identiques, l'agriculture contribue 5.2 fois plus à la baisse du PIB par habitant que les services. Et l'industrie a contribué plus de 2.9 fois à la baisse du PIB par habitant que les services.
Nous n'en sommes qu'à nos débuts pour l'exploration de cette méthode de calcul des contributions et de leurs implications. De même, il conviendrait de travailler en monnaie constante. De futures recherches et applications permettront d'en tirer de plus grands retours d'expérience.

7. Conclusion

L'originalité de notre approche cliométrique de l'histoire économique globale a consisté à appliquer, les méthodes et démarches d'analyse séquentielle biographique, à l'analyse longitudinale macroéconomique.
Pour cette étude, nous-nous sommes restreints à l'histoire économique récente des 27 années dernières (23 avec les moyennes mobiles), soit à peu près une génération. Les indicateurs structurels de l'économie que nous avons utilisés, sont basés sur la population en âge de travailler, le taux d'emploi, la part d'emploi de chaque secteur, les productivités sectorielles. Les secteurs considérés sont l'agriculture et la pêche, l'industrie et les services.

Détection des phases (ou états) structurelles de l'économie
La détection des phases transitionnelles historiques, par des méthodes de classification statistiques sur composantes principales, appliquée aux données longitudinales économiques, nous a permis de détecter quatre phases (ou états) structurelles possibles de l'économie. Chaque pays a donc sa trajectoire économique décrite par une séquence de phases (ou états).

Analyse des séquences d'états (*State sequences Analysis STS*)
Grâce aux méthodes d'analyse des séquences d'états, nous utilisons les états (phases) économiques précédemment construites pour réaliser une analyse « biographique » des trajectoires macro-économiques (encore appelées séquences) des pays à l'échelle planétaire sur la période étudiée.
Le recours à cette méthodologie permet d'obtenir une visualisation mondiale synthétique des trajectoires économiques nationales : sur la période étudiée, nous observons que le nombre de pays dans l'état économique la moins avancée transitionnellement (State.1) est en décroissance continue. L'état économique suivant (State.2) décroit légèrement en effectif, tandis que nous constatons une importante hausse du nombre de pays en phase 3 (State.3). Le nombre de pays les plus avancés transitionnellement (State.4) connait également une tendance à la hausse. Nous constatons des progrèssions transitionnelles sur tous les continents, caractérisées par des passages d'états moins avancés à des états plus avancés dans la transition économique, ainsi que l'absence des cas de régression transitionnelle durable.

Mode annuel des états : de 1995 à 2015, l'état (phase) le plus fréquent par an, est State.2. Cependant c'est l'état State.3, plus avancé, qui est devenu le mode depuis 2016.

Durée moyenne passée par les pays dans chaque état : globalement chaque pays semble avoir passé l'essentiel de son temps dans un état spécifique, avec quelques fois des évolutions vers un autre état. Il y a donc une inertie importante, avec des évolutions relativement lentes.

Les probabilités de transition entre états économiques : les pays ayant connu une transition économique précoce ont été plus lents dans leur évolution que les pays à transition économique plus tardive ; la matrice de probabilités de transition, serait donc plus pertinente par groupes homogènes de pays, qu'au niveau mondial. Toutefois, sur la période historique étudiée, nous observons l'absence de certains sauts d'une année à l'autre : ce sont les sauts de deux états ou plus. Les seules évolutions observées, sont celles entre des classes d'indice contigües avec des probabilités de « progrès » nettement plus fortes que les probabilités de recul.
Le phénomène d'inertie est très visible par des probabilités de rester dans le même état d'une année à l'autre, qui sont toutes supérieures à 96%.

La mise en évidence et la caractérisation des typologies de trajectoires économiques nationales : le nombre optimal de classes de trajectoires nationales est de quatre. La première classe de trajectoires

est majoritairement caractérisée par une position permanente dans l'état 1 (State.1) et pour environ 35% des pays par un passage de State.1 à State.2. La deuxième classe de trajectoires est majoritairement caractérisée par une position stable dans l'état 2 (State.2), avec pour certains pays minoritaires, une évolution de State.2 à State.3. Et aussi pour une petite minorité, le passage de State.1 à State.2 en début d'historique. La troisième classe de trajectoires est essentiellement caractérisée par une position stable dans l'état 3 (State.3), avec deux petites minorités de pays caractérisés par une évolution de State.2 à State.3 en début d'historique, ou par le passage de State.3 à State.4. La quatrième classe de trajectoires est presque exclusivement caractérisée par une position permanente dans l'état 4 (State.4), avec une petite minorité de pays caractérisés par une évolution de State.3 à State.4 en début d'historique.
Fait notable, la classe de trajectoire qui a le plus faible effectif (24 pays) est la classe la moins avancée économiquement (*Trajectoire Eco 1*), tandis que le groupe le plus large (43 pays) est celui des pays de la trajectoire de type 2 (*Trajectoire Eco 2*) composé de pays ayant quitté la classe la moins avancée.

Analyse ANOVA-généralisée : la zone géographique du pays explique ici 23.4% (pseudo-R^2) de la variabilité des trajectoires et l'association est significative (p-value < %1). La statistique de Levene sert à tester l'égalité des discrépances intra-groupe. La zone de plus forte variabilité des trajectoires des pays, est l'Europe. La zone ayant la plus grande homogénéité des trajectoires est l'Océanie.

Analyse multivariée de la discrépance des trajectoires : nous avons modélisé la discrépance des trajectoires à partir de plusieurs variables explicatives, par généralisation de la régression par arbre de segmentation.
En utilisant les covariables, *Region, High income (yes, no), Middle income (yes, no) et Low income (yes, no)* nous obtenons une part de discrépance expliquée de 34% (pseudo R^2). L'importance des variables pour la discrimination des trajectoires est dans l'ordre : *High income, Region, Low income.*

L'analyse des séquences d'évènements (*Time stamped Events analysis TSE*)
L'analyse des séquences d'évènements a pour objectif d'étudier et possiblement anticiper les séquences d'évènements (changement d'état) économiques majeurs.
Découverte des sous-séquences les plus fréquentes : les sous-séquences les plus fréquentes sont constituées d'un unique état (State.2, State.3, State.1). L'évènement State.2>State.3 constitue la quatrième sous-séquence la plus fréquente. La cinquième sous-séquence la plus fréquente est composée de l'état State.2, suivi dans un délai quelconque, par l'évènement State.2>State.3 : la sous-séquence est notée (State.2)-(State.2>State.3).

Sous-séquences discriminantes des classes de trajectoires : l'objectif de cette section est d'identifier les sous-séquences associées aux classes de trajectoires. Le pouvoir discriminant est déterminé par la p-value d'un test du Khi-2. Ce qui a permis d'identifier des sous-séquences bien spécifiques à chacune des quatre classes de trajectoires nationaux, en cohérence avec les descriptions qui avaient été faites de ces classes.

Recherche de sous-séquences et calcul des occurrences : il nous a été possible de calculer le nombre de réalisation d'une sous-séquence particulière d'évènements, à l'échelle mondiale sur la période étudiée. Ainsi, sur la période historique étudiée, nous avons pu établir que : l'évènement (State.1>State.2), s'est produit sur 14 (12%) trajectoires de pays, tandis que l'évènement (State.3>State.4) ne s'est produit que dans 5 pays (4%). L'évènement intermédiaire (State.2>State.3) s'est déroulé dans 21 pays (18% des cas). Autre fait notable, aucun pays, sur les 23 années étudiées, n'a fait un saut progressif de 2 états économiques successifs.

Découverte des règles d'associations séquentielles : une règle d'association séquentielle, consiste à subdiviser une sous-séquence en une partie *prémisse* et une autre partie *conclusion*, qui se suivent temporellement. Des algorithmes spécifiques permettent d'extraire ces règles en se basant sur le seuil de support exigé et un niveau de confiance minimal. Avec un support relatif supérieur ou égal à 0.01%

(règles établies avec au moins 2 pays sur les 115), l'extraction donne 8 règles séquentielles. Avec le seuil de 5% qui est le plus utilisé en pratique, seules deux règles existent. Les trois premières règles séquentielles indiquent le passage d'un état économique vers l'état contigu qui le suit dans l'ordre de la productivité agricole : les principales règles observées indiquent donc que les pays progressent sans rupture, en termes de développement économique. Les quatre règles suivantes, bien que de support faible, évoquent toutes un faible risque d'instabilité (progression suivi de régression) entre les états State.2 et State.3. Cependant ce faible risque de régression localisé est contrebalancé par la huitième règle, qui a un niveau de confiance élevé (67%), et qui indique que la régression est provisoire.

Contributions à la variation pluriannuelle du PIB par habitant, de ses composantes structurelles

Toujours dans l'analyse longitudinale, nous avons développé une méthode d'estimation séquentielle (puis par agrégation) des contributions à la variation pluriannuelle du PIB par habitant, des principaux secteurs économiques (agriculture et pêche, industrie, services), et de la dynamique démographique de la population en âge de travailler. Ces variables sont les composantes structurelles de l'économie que nous avons utilisées tout au long de ce chapitre, d'abord pour la détection des phases qui ont servi de base à l'analyse séquentielle par la suite.

8. Bibliographie

Azar, A., Staudinger, U. M., Slachevsky, A., Madero-Cabib, I., & Calvo, E. (2019). From Snapshots to Movies: The Association Between Retirement Sequences and Aging Trajectories in Limitations to Perform Activities of Daily Living. Journal of Aging and Health, 31(2), 293-321. https://doi.org/10.1177/0898264318782096

Bairoch P. (2005). « Mythes et paradoxes de l'histoire économique ». La découverte

Casper, G., & Wilson, M. (2015). Using Sequences to Model Crises. Political Science Research and Methods, 3(2), 381-397. https://doi.org/10.1017/psrm.2014.27

Cornwell, B. (2018). Network Analysis of Sequence Structures. In G. Ritschard & M. Studer (Éds.), Sequence Analysis and Related Approaches: Innovative Methods and Applications (p. 103-120). Springer International Publishing. https://doi.org/10.1007/978-3-319-95420-2_7

Emmenegger, P., Häusermann, S., Palier, B., & Seeleib-Kaiser, M. (2013). Structural change and the politics of dualization. Rassegna Italiana di Sociologia, 2/2013. https://doi.org/10.1423/74142

Ferrari, G., & Pailhé, A. (2017). Transition to adulthood in France: Do children of immigrants differ from natives? Advances in Life Course Research, 31, 34-56. https://doi.org/10.1016/j.alcr.2016.10.001

Fuller, S. (2015). Do Pathways Matter? Linking Early Immigrant Employment Sequences and Later Economic Outcomes: Evidence from Canada. International Migration Review, 49(2), 355-405. https://doi.org/10.1111/imre.12094

Gabadinho, A., & Ritschard, G. (2016). Analysing state sequences with probabilistic suffix trees: the PST R library. Journal of Statistical Software, 72(3), 1-39. https://doi.org/10.18637/jss.v072.i03

Gabadinho, A., Ritschard, G., Müller, N.S. & Studer, M. (2011), "Analyzing and visualizing state sequences in R with TraMineR", Journal of Statistical Software. Vol. 40(4), pp. 1-37.

Helske S., Helske J. (2019). Mixture Hidden Markov Models for Sequence Data: The seqHMM Package in R, Journal of statistical software, january 2019

Kang, T. (2019). The Transition to Adulthood of Contemporary Delinquent Adolescents. Journal of Developmental and Life-Course Criminology, 5(2), 176-202. https://doi.org/10.1007/s40865-019-00115-6

Le Meur, N., Vigneau, C., Lefort, M., Lebbah, S., Jais, J.-P., Daugas, E., & Bayat, S. (2019). Categorical state sequence analysis and regression tree to identify determinants of care trajectory in chronic disease: Example of end-stage renal disease. Statistical Methods in Medical Research, 28(6), 1731-1740. https://doi.org/10.1177/0962280218774811

McVicar D., Anyadike-Danes M. (2002). Predicting successful and unsuccessful transitions from school to work by using sequence methods. Journal of the Royal Statistical Society. Series A (Statistics in Society) 165(2), 317-334.

Müller N, Studer M, Ritschard G, Gabadinho A (2010). Extraction de règles d'association séquentielle à l'aide de modèles semi-paramétriques à risques proportionnels. Revue des nouvelles technologies de l'information (RNTI), E-19, 25-36.

Perhoniemi, R., Blomgren, J., & Laaksonen, M. (2019). Sources of income following a rejected disability pension application: a sequence analysis study. Disability and Rehabilitation, 0(0), 1-9. https://doi.org/10.1080/09638288.2018.1555619

Petri P.A. (1993), The lessons of East Asia : common foundations of East Asian success, The World Bank, WVashington, D.C.

Raab, M., & Struffolino, E. (2019). The Heterogeneity of Partnership Trajectories to Childlessness in Germany. European Journal of Population, 1–18.

Ritschard G., Gabadinho A., Muller N.S., Studer M. (2009), Données séquentielles: concepts et principe d'anaylse, Conference EGC 2009, Strasbourg, 27-30 janvier 2009

Ritschard, G., & Studer, M. (2018b). Sequence Analysis: Where Are We, Where Are We Going? (p. 1-11). https://doi.org/10.1007/978-3-319-95420-2_1

Ritschard, G., Bürgin, R. & Studer, M. (2013), "Exploratory Mining of Life Event Histories", In McArdle, J.J. & Ritschard, G. (eds) Contemporary Issues in Exploratory Data Mining in the Behavioral Sciences. Series: Quantitative Methodology, pp. 221-253. New York

Robette (2014). L'étude des parcours de vie, démarche descriptive ou causale ?, Idées économiques et sociales 2014/3 (N° 177), pages 8 à 15

Salvà-Mut, F., Thomás-Vanrell, C., & Quintana-Murci, E. (2016). School-to-work transitions in times of crisis: the case of Spanish youth without qualifications. Journal of Youth Studies, 19(5), 593-611. https://doi.org/10.1080/13676261.2015.1098768

Saporta G. (2011). « Probabilités, Analyse des données et Statistique ». Technip

Saporta, G. (1975). « Liaisons entre plusieurs ensembles de variables et codage de données Qualitatives ». Thèse de 3° cycle, Paris VI : Université Pierre et Marie Curie.

Sequence Analysis Association (2020). SAA Bibliography, Sequence analysis. Consulté en ligne le 22/11/2020,https://www.zotero.org/groups/2268769/saa_bibliography/collections/T33535YZ/items/CB7J48ZW/collection

Studer M. (2012), Étude des inégalités de genre en début de carrière académique à l'aide de méthodes innovatrices d'analyse de données séquentielles., Thèse de doctorat : Univ. Genève, 2012, no. SES 777

Studer M., Ritschard G., Gabadinho A., Müller N. S. (2011). Discrepancy analysis of state sequences. Sociological Methods and Research 40 (3), 471–510.

Studer, M., Ritschard, G., Gabadinho, A. & Müller, N.S. (2011), "Discrepancy Analysis of State Sequences", Sociological Methods and Research. Vol. 40(3), pp. 471-510.

Tufféry S. (2012). « Data Mining et Statistique Décisionnelle ». Technip.

World Bank (2017), World Development Indicators, [en ligne] (page consultée le 22/11/2017). https://datacatalog.worldbank.org/dataset/world-development-indicators

Zhelyazkova, N., & Ritschard, G. (2018). Parental Leave Take-Up of Fathers in Luxembourg. Population Research and Policy Review, 1-25. https://doi.org/10.1007/s11113-018-9470-8

Chapitre 3.
Analyse cliométrique transversale des transitions ACIDE

1. Table des matières

Chapitre 3. Analyse cliométrique transversale des transitions ACIDE 147

1. Table des matières 148

2. Introduction 149

3. Rappels 151

4. La transition économique 152

4.1. Analyse transversale de la transition économique 152

4.2. Comparaison du schéma transitionnel par coupe transversale avec le schéma théorique longitudinal de la transition économique 178

5. La transition démographique 181

5.1. Analyse transversale de la transition démographique 181

6. Anthropologie culturelle 206

6.1. Analyse transversale anthropologique 206

6.2. Peut-on faire l'hypothèse d'une théorie transitionnelle anthropo-culturelle ? 222

7. Institutions politiques et économiques 225

7.1. Analyse transversale institutionnelle 225

7.2. Peut-on faire l'hypothèse d'une théorie transitionnelle des institutions ? 245

8. Cognition, éducation et capital humain 248

8.1. Analyse transversale cognitive 248

8.2. Peut-on faire l'hypothèse d'une théorie transitionnelle de la cognition ? 276

9. Analyse multi-transitionnelle 279

9.1. Economie vs Autres transitions 279

9.2. Démographie vs Autres transitions 283

9.3. Anthropologie vs Autres transitions 286

9.4. Institutions vs Autres transitions 289

9.5. Cognition vs Autres transitions 292

9.6. Synthèse des associations entre transitions ACIDE 295

10. Conclusion 297

2. Introduction

L'absence de données économiques structurelles sur le temps long a été une limite à l'exploration historique du chapitre 2. Notamment, le manque de données nationales de long terme sur le secteur industriel, qui est une composante fondamentale de la structure et la transition économique d'un pays. Ayant défini la grille multidisciplinaire au chapitre 1, et ayant constaté les limites de l'approche longitudinale en termes de profondeur d'historique au chapitre 2, nous allons à présent tenter une inférence longitudinale à travers une coupe transversale et l'usage du temps transitionnel.

Par analyse transversale ou en coupe instantanée, il faut comprendre un « arrêt sur image » sur une année donnée. Ceci semble exclure toute dimension temporelle de l'analyse transversale contemporaine. L'idée qu'une vision en coupe instantanée puisse préfigurer l'évolution future d'un pays et renseigner sur la trajectoire passée peut, au premier abord, être assimilé à une vue de l'esprit. Comment analyser la dimension temporelle dès lors que le temps est absent, par définition, de l'analyse transversale contemporaine ? Si le temps est figé, comment obtenir une dynamique temporelle ?

Nous prenons en compte cette objection en introduisant les notions de temps transitionnel et d'âge transitionnel définies au chapitre 1, dans l'analyse transversale contemporaine. Ceci rend l'analyse originale dans la mesure où dans les travaux d'économie traditionnels, cette question n'est pas explicitement traitée. Nous avons systématiquement, pour mener à bien l'analyse des transitions ACIDE, fait le choix nécessaire d'un temps transitionnel qui soit de tendance monotone et de pente aussi proche que possible de 1 (sachant que la première condition est suffisante pour ordonner des groupes de pays selon une échelle de temps transitionnel).

L'originalité et la contribution de l'approche développée dans cette partie est de réintroduire cette dimension du temps dans l'analyse transversale contemporaine. En effet, ce qui est recherché ici est l'âge transitionnel moyen par classe statistique de pays : par exemple, pour l'analyse de la transition démographique, nous avons fait le choix d'ordonner les classes statistiques de pays selon la moyenne par classe du taux de mortalité infantile. Pour l'analyse de la transition économique, le choix s'est porté sur le pourcentage des emplois dans le secteur agricole. Ces deux exemples illustrent que les classes statistiques de pays sont ordonnées selon un critère qui reflète correctement la phase transitionnelle. Comme ces deux indicateurs sont tendanciellement monotones et décroissants, une valeur élevée traduit une entrée en transition (économique et démographique) récente et correspond (dans les deux cas) à une phase proche du début du processus de transition.

Afin de caractériser le processus de transition, nous recourrons aux méthodologies suivantes : analyses factorielles, classifications statistiques, détection de valeurs atypiques sur données non gaussiennes, mesure et tests d'associations linéaire ou non-linéaires entre variables.

L'originalité de notre méthode consiste à créer des classes de pays par rapport aux variables structurelles du domaine étudié et d'ordonner temporellement ces classes leur âge transitionnel moyen. Une fois datées temporellement, ces classes de pays appartiennent à une même phase de transition. On est ainsi à même de savoir quelle phase de transition précède l'autre. Le contenu informationnel de l'approche transversale est ainsi rendu comparable et partiellement substituable à l'approche longitudinale, en l'absence de données historiques et dans les limites des phases observées à partir des données transversales.

De façon imagée, on pourrait comparer le contenu informationnel délivré notre approche transversale à la datation au carbone 14 (où la concentration d'un isotope particulier est le temps transitionnel) ou au carottage (où le temps transitionnel est la profondeur), ou encore en astrophysique (où la distance

mesurée par les télescopes est le temps transitionnel) : en effet le présent contient des éléments d'information sur la trajectoire passée et prévisible. Nous remarquerons ici que ce procédé est courant et classique dans les sciences physiques, biologiques, géologiques et en archéologie et que nous cherchons ici à en adapter une forme en cliométrie.

Dans les deux domaines ACIDE où un processus de transition moderne est déjà théorisé (i.e. économie et démographie), la question qui se pose est celle de la reproductibilité, et donc de l'actualité du processus transitionnel dans les pays émergents ou à revenus les plus faibles.
A ce propos, McMillan et Rodrik (2014) exprimaient récemment des inquiétudes sur le risque d'une désindustrialisation précoce ou sur la capacité à s'industrialiser conformément à la théorie de la transition économique de certains pays pauvres mais riches en ressources naturelles comme par exemple le Nigéria et le Vénézuela.

Dans les trois autres domaines ACIDE (i.e. anthropologie, cognition, institutions) pour lesquels il n'existe pas de théories transitionnelles quantitatives établies, se posent les questions suivantes : Quels temps transitionnels adopter dans une approche transversale ? Peut-on proposer une théorie transitionnelle moderne dans ces domaines ?

Du point de vue multi-transitionnel, les questions qui se posent concernent le timing croisé des phases entre les cinq transitions ACIDE : Quelles phases sont concomitantes ou antagonistes ? Certaines phases dans une transition spécifique peuvent-elles être considérées comme limitantes (ou au contraire stimulantes) pour le déroulement du processus transitionnel dans d'autres domaines ?
Ces sujets évoquent, entre autres, les craintes économiques malthusiennes par suite d'une forte hausse démographique typique aux premières phases de transition démographique. Les ressources supposées croitre arithmétiquement peuvent-elles compenser les besoins nés d'une explosion démographique supposément géométrique ? Ou alors est-ce l'hypothèse boserupienne d'une adaptation (sous réserve d'une évolution institutionnelle) extensive et intensive des systèmes de production qui l'emporte ? (Boserup [1965], Le Bras [1994, p.254], Le Bras [2003]).

Les sections 4 et 5 sont consacrées aux transitions dont les théories sont déjà formulées et étudiées dans la littérature : Il s'agit de la transition économique et de la transition démographique. L'originalité de notre analyse ici consiste à utiliser l'approche transversale pour vérifier si ces transitions sont toujours à l'œuvre « actuellement » et à les enrichir par des caractéristiques historiques complémentaires à travers l'abondance des nouvelles statistiques actuellement disponibles dans les bases de données.

Dans les sections 6, 7, et 8, nous abordons les trois autres domaines ACIDE (i.e. anthropologie, cognition, institutions) pour lesquels il n'existe pas de théories transitionnelles quantitatives établies dans la littérature, et nous proposons et testons statistiquement un temps transitionnel, qui permet d'établir un âge transitionnel associé.
Enfin, nous recherchons et formulons domaine par domaine, une théorie transitionnelle dont nous testons la robustesse et la cohérence empirique avec les évolutions historiques connues sur certains indicateurs ACIDE.

Pour finir ce chapitre, la 9ième section est consacrée à l'analyse transversale multi-transitionnelle. Nous y étudions les relations entre les phases des cinq transitions ACIDE pour reconstituer la dynamique multi-transitionnelle.

3. Rappels

Nous allons dans les sections qui suivent, utiliser le temps et l'âge transitionnel définis au chapitre 1, pour réaliser une inférence de trajectoire longitudinale moyenne dans une coupe transversale dans chacun des cinq domaines ACIDE.
En effet, nous avions défini (chapitre 1) le temps transitionnel, comme une série temporelle qui a une dépendance monotone forte avec le temps.
Nous mesurons la dépendance monotone (concordante ou discordante) par le coefficient de corrélation des rangs de Kendall τ entre la série temporelle testée pour jouer le rôle de temps transitionnel et le temps (Lecoutre, Tassi [1987], pp. 151-154).
Il est ainsi possible d'utiliser les tests d'hypothèse sur le coefficient de Kendall pour tester l'hypothèse qu'une série temporelle puisse être utilisée comme temps transitionnel dans une analyse transversale.
Nous notons également que toute transformée symétrique d'une série temporelle discordante au temps, est concordante au temps, donc peut jouer le rôle de temps transitionnel (si la discordance est forte).
Nous avions également défini (chapitre 1) l'âge transitionnel d'un pays, comme différence entre le temps transitionnel et sa valeur minimale (quel que soit le pays) dans la période de transition. Cette valeur est un seuil de passage définissant le début théorique de la période transitoire, quel que soit le pays.

Avec des données en coupe transversale, nous testons et sélectionnons dans une première étape le temps et l'âge transitionnel adapté au domaine ACIDE étudié.
Nous effectuons dans une deuxième étape, une classification statistique des pays, basée sur les variables structurelles du domaine étudié.
Dans une troisième étape, les classes de pays sont indicées et ordonnées selon leur moyenne de l'âge transitionnel.
Enfin, dans la quatrième étape, nous produisons une courbe de trajectoire moyenne avec en abscisse l'indice d'âge transitionnel des classes de pays, et en ordonnée les moyennes des variables structurelles des classes des pays.

Cette même démarche en quatre grandes étapes sera répétée dans chacun des cinq domaines ACIDE.
Nous commençons par les deux domaines ACIDE où un processus de transition moderne est déjà théorisé (i.e. économie et démographie), puis nous poursuivons avec les trois autres domaines ACIDE (i.e. anthropologie, cognition, institutions) pour lesquels il n'existe pas de théories transitionnelles quantitatives établies.

Les analyses sont effectuées avec exclusion des petits Etats (*small states*) au sens de la Banque Mondiale (The World Bank, 2019), modifiée en y intégrant tout autre pays ayant une superficie inférieure au troisième quartile des *small states* de la Banque Mondiale.
Ces *small states*, du fait de leur « faible population (à savoir, celles avec une population de 1,5 million d'habitants ou moins) et de leur base économique, ces pays sont particulièrement vulnérables aux chocs exogènes, tels que les catastrophes naturelles et le changement climatique. Avec des opportunités économiques limitées et une migration importante, ils sont souvent confrontés à des contraintes de capacité » (The World Bank, 2019).
Sont également exclus de l'étude, les pays ayant des valeurs atypiques (détection par box-plot après normalisation par la méthode de Box-cox) dans les variables analysées.

4. La transition économique

4.1. Analyse transversale de la transition économique

4.1.1. Sélection des variables à étudier

Les variables structurelles utilisées se répartissent en deux catégories : les variables de transition et les variables actives.

Variables de transition : elles sont utilisées classiquement pour décrire le processus de transition économique.

Ces variables sont : l'emploi dans l'agriculture (% de l'emploi total), l'emploi dans l'industrie (% de l'emploi total), et l'emploi dans les services (% de l'emploi total).

Variables actives : Elles sont utilisées pour effectuer la classification statistique des pays. Elles incluent les variables de transition classiques auxquelles sont ajoutées des variables d'homogénéisation des classes. En effet une classification sur la base des seules variables de transition n'est pas suffisante car elle aboutit à des classes très hétérogènes en termes de niveau de développement et de PIB par habitant.

Tableau 3.1 Variables actives

Variables	Signification en anglais	Signification en français	Variable active	Variable de transition
Employment.in.agriculture.Pct.of.total.employment	Employment in agriculture (% of total employment) (modeled ILO estimate)	Emploi dans l'agriculture (% de l'emploi total) (estimation modélisée du BIT)	oui	oui
Employment.in.industry.Pct.of.total.employment	Employment in industry (% of total employment) (modeled ILO estimate)	Emploi dans l'industrie (% de l'emploi total) (estimation modélisée du BIT)	oui	oui
Employment.in.services.Pct.of.total.employment	Employment in services (% of total employment) (modeled ILO estimate)	Emploi dans les services (% de l'emploi total) (estimation modélisée du BIT)	oui	oui
Population.ages.15.64.Pct.of.total	Population ages 15-64 (% of total)	Population âgée de 15 à 64 ans (% du total)	oui	non
Agr.VA.per.worker.agr.product.const.2010USD	Agriculture value added per worker, measure of agricultural productivity (constant 2010 US$)	Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010)	oui	non
Industry.value.added.per.worker.constant.2010.USD	Industry, value added per worker (constant 2010 US$)	Industrie, valeur ajoutée par travailleur (USD constants de 2010)	oui	non
Services.value.added.per.worker.constant.2010.USD	Services, value added per worker (constant 2010 US$)	Services, valeur ajoutée par travailleur (USD constants de 2010)	oui	non

Source des données : World Bank (2017)

4.1.2. Transformation préalable des variables à étudier

Les variables actives utilisées font l'objet des transformations suivantes, avant la sélection de l'année d'étude : estimation des valeurs manquantes par interpolation spline, et moyenne mobile arrière d'ordre 5.

4.1.3. Sélection de l'année à étudier

Pour l'analyse transversale, l'année étudiée est 2014. Cela pour des raisons de complétude et de fiabilité des données, car les dernières (2017, source Banque Mondiale) sont souvent complétées et révisées après publication.

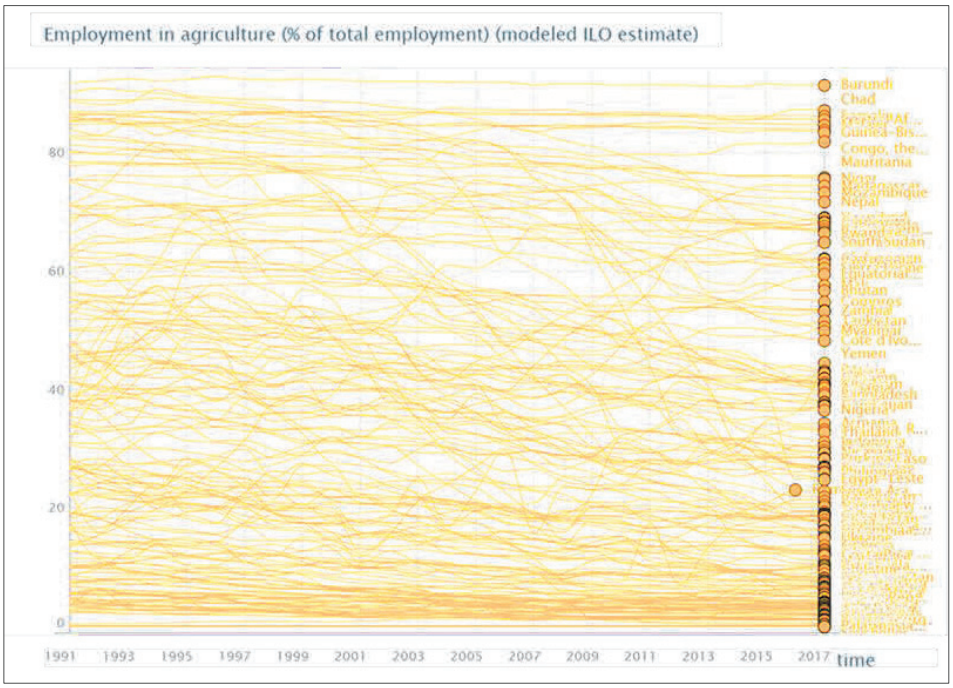
4.1.4. Choix du temps transitionnel

Variable à tester comme temps transitionnel

Le taux d'emploi dans l'agriculture est pressenti comme *symétrique* du temps transitionnel car il est supposé baisser dans le schéma théorique de la transition économique que nous avions déjà exposée.

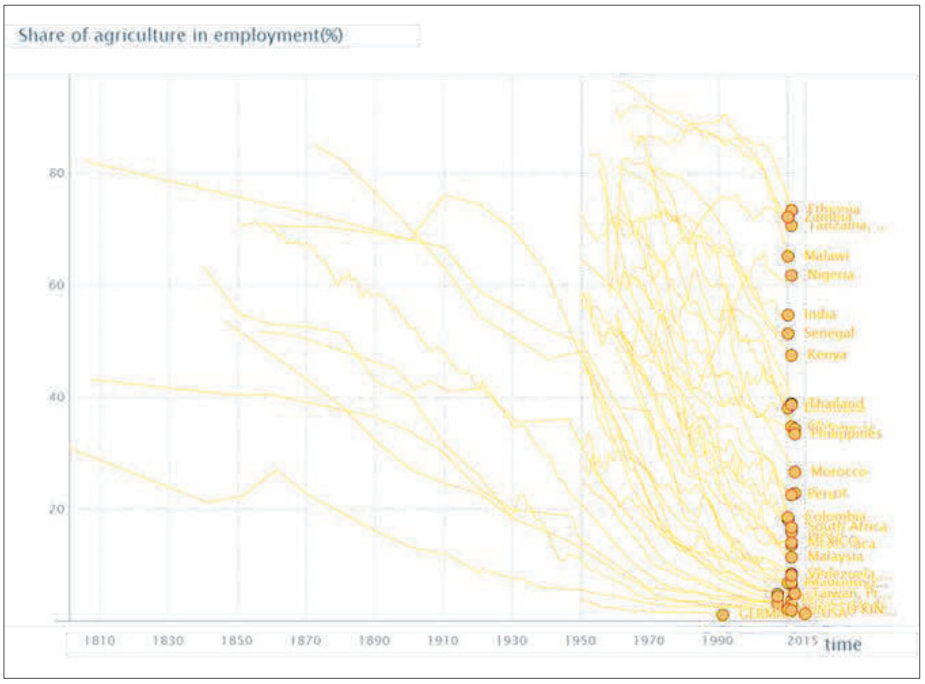
La variable à tester est une valeur transformée (moyenne mobile arrière d'ordre 5, pour lisser les fluctuations cycliques de court terme) de l'emploi dans l'agriculture (% de l'emploi total), estimation BIT.

Figure 3.1 Le taux d'emplois agricoles (% de l'emploi total), estimation BIT



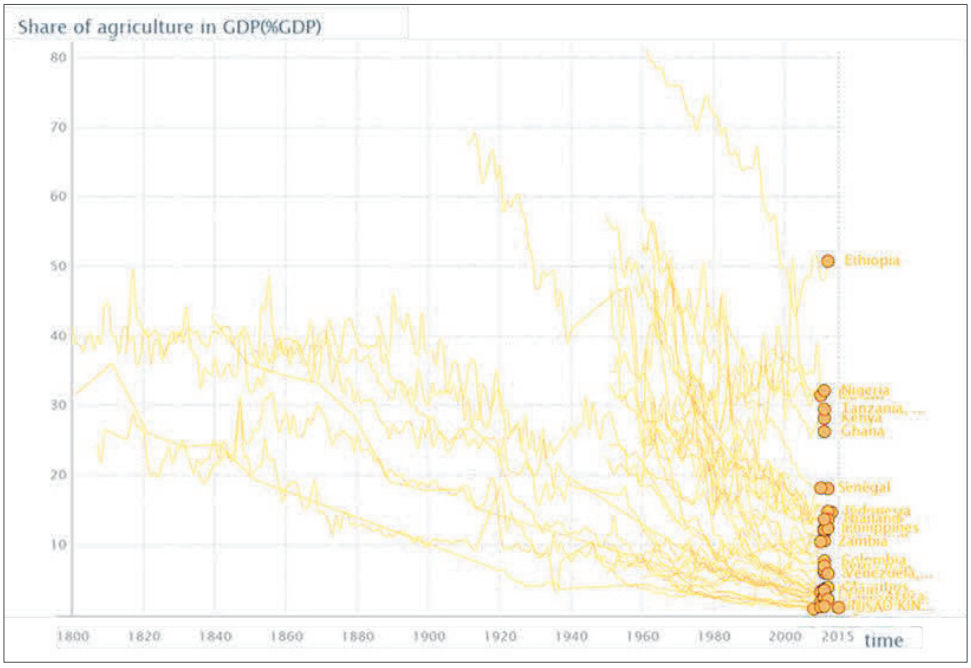
Source : World Bank (2018)

Figure 3.2 Le taux d'emplois agricoles (% de l'emploi total), historique long



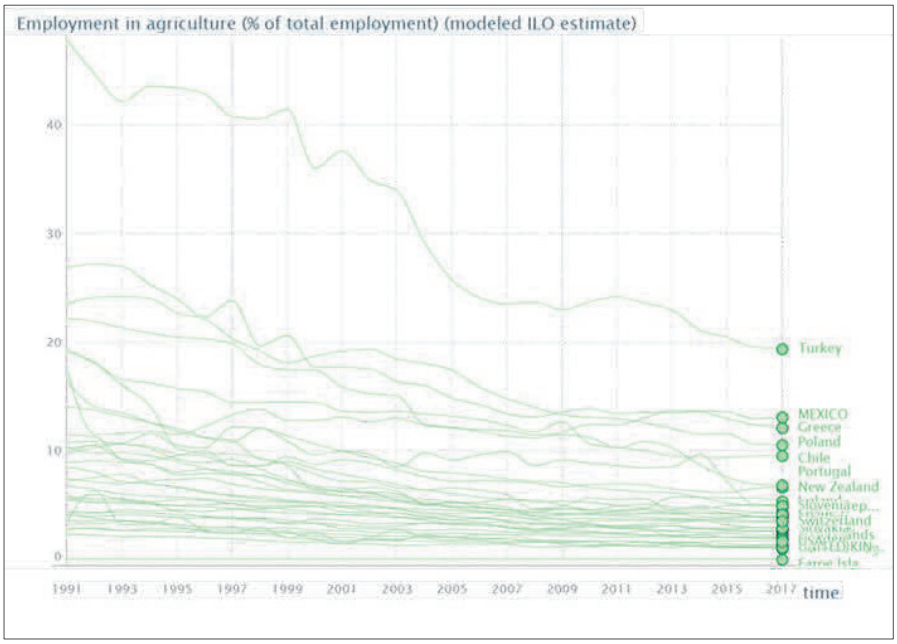
Source : Max Roser (2015)

Figure 3.3 Part de l'agriculture dans le PIB (% du PIB), historique long



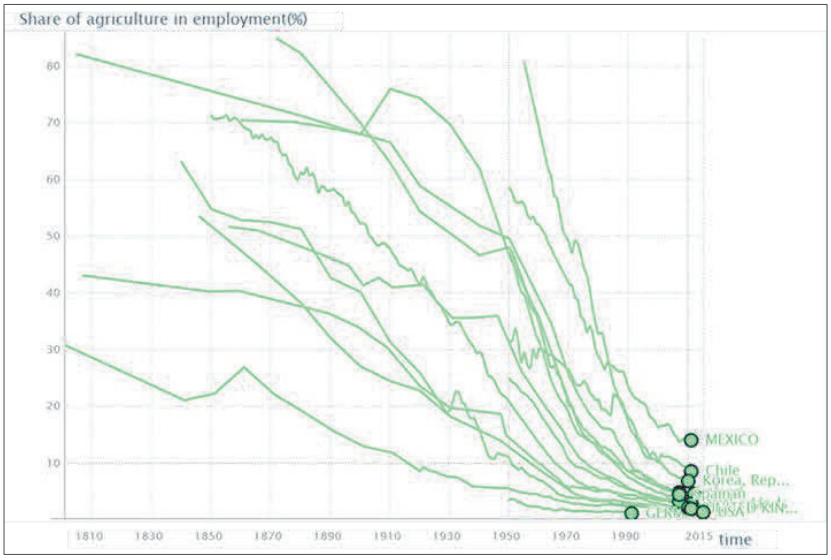
Source : Max Roser (2014)

Figure 3.4 Le taux d'emplois agricoles (non-lissé) des pays de l'OCDE



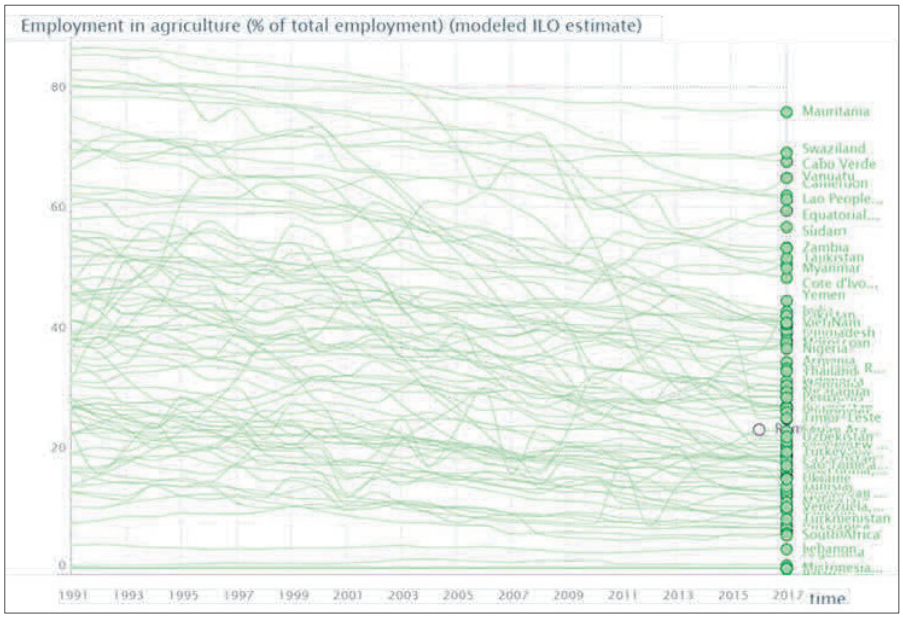
Source : World Bank (2018)

Figure 3.5 Le taux d'emplois agricoles (%) des pays de l'OCDE. Historique long.



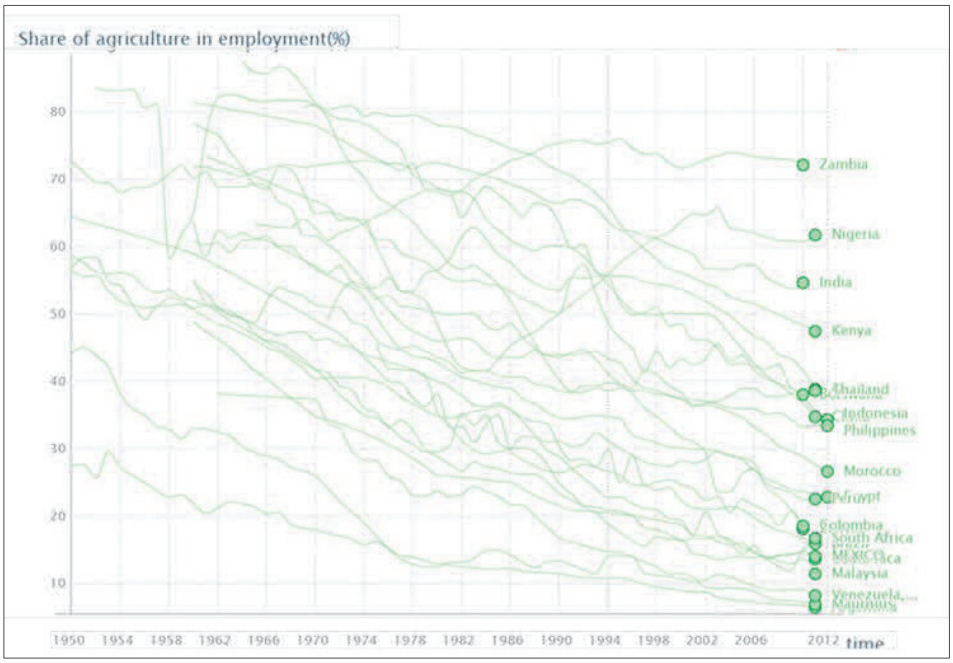
Source : Max Roser (2015)

Figure 3.6 Le taux d'emplois agricoles (non-lissé) des pays à moyen revenu par habitant (au sens de la Banque Mondiale)



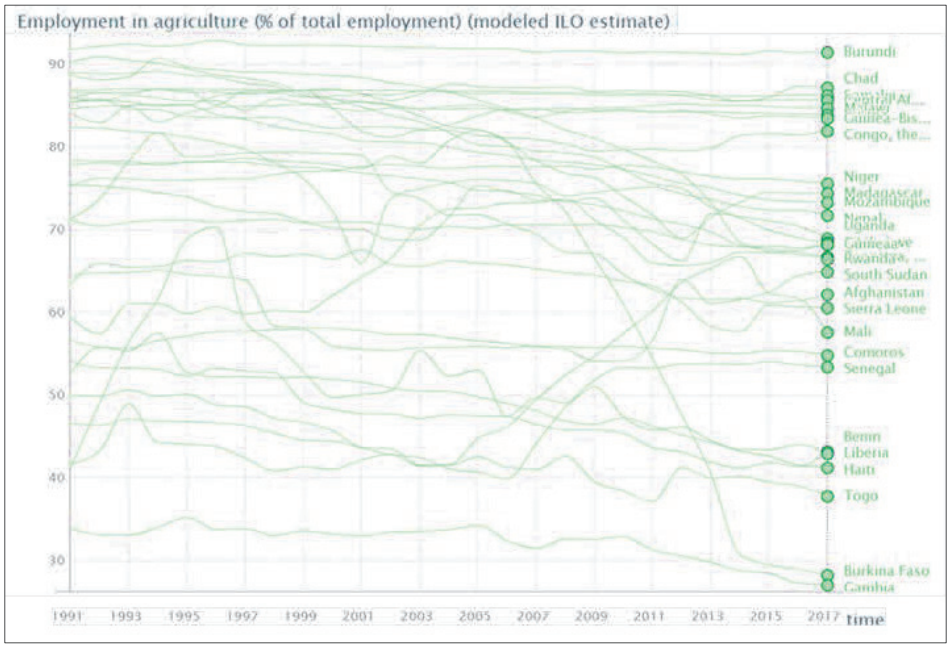
Source : World Bank (2018)

Figure 3.7 Le taux d'emplois agricoles (%) des pays à moyen revenu par habitant (au sens de la Banque Mondiale)



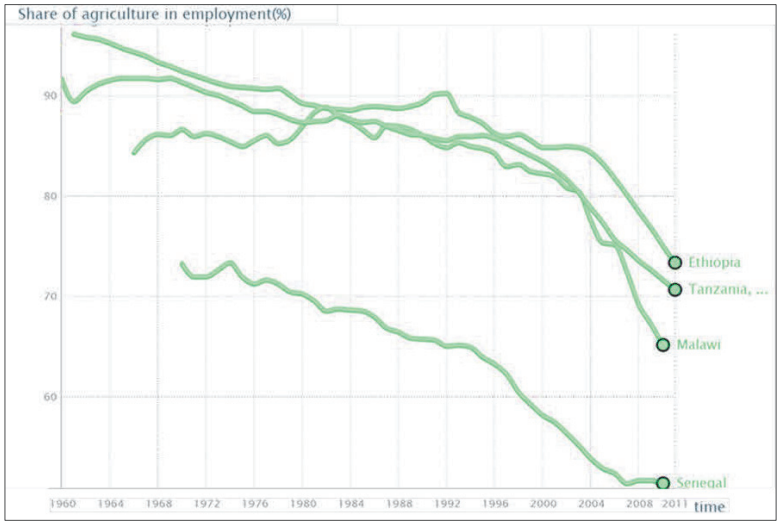
Source : Max Roser (2015)

Figure 3.8 Le taux d'emplois agricoles (non-lissé) des pays à faible revenu par habitant (au sens de la Banque Mondiale)



Source : World Bank (2018)

Figure 3.9 Le taux d'emplois agricoles (%) des pays à faible revenu par habitant, (au sens de la Banque Mondiale)



Source : Max Roser (2015)

Visuellement la tendance mondiale du taux d'emplois agricoles est à la baisse (et semble converger), même si on ne peut parler de monotonie mathématique : c'est la raison pour laquelle nous utilisons un lissage par moyenne mobile pour améliorer la monotonie de la variable pressentie pour jouer le rôle *symétrique* de temps transitionnel.

Vérification de la monotonie empirique de la variable à tester

La variable à tester Y est empiriquement monotone lorsque la série différenciée dY [dY(t)=Y(t) - Y(t-1)] est de signe constant sur la période étudiée, pour tout pays.

De plus pour assurer l'universalité de cette monotonie empirique, le signe de série différenciée doit être le même pour tous les pays étudiés

Pour chaque pays, nous calculons le pourcentage de valeurs négatives de la série différenciée, afin de vérifier l'unicité du signe sur la période étudiée.

Sur l'ensemble des 119 pays étudiés, le pourcentage moyen de valeurs négatives de la série différenciée est de 79.3%, avec un minimum à 38.5% (Zimbabwe), une médiane à 81%, et un maximum à 100% (Australie, Chili, Allemagne par exemple).
13 des 119 pays, soit 11%, ont 100% de valeurs négatives dans la série différenciée.
37 des 119 pays, soit 31%, ont plus de 90% de valeurs négatives dans la série différenciée.
66 des 119 pays, soit 55%, ont plus de 80% de valeurs négatives dans la série différenciée.

Par ailleurs, les valeurs centrales (moyenne=-0.8 et médiane=-0.9) du coefficient de Kendall indiquent une très forte relation monotone décroissante entre le taux d'emplois agricoles et le temps.
Sur les 119 pays, le coefficient de Kendall est systématiquement négatif (99% des pays), sauf pour le Zimbabwe.

Vérification empirique de l'unicité du signe de la pente de la variable testée, quel que soit le pays

A défaut d'une monotonie empirique parfaite, l'unicité pour tous les pays, du signe de la pente de la droite d'ajustement de la variable à tester, montre une universalité de la tendance (à la hausse ou à la baisse) sur l'ensemble des pays étudiés.
De plus, si pour un pays, pente de la droite d'ajustement de la variable à tester est proche en valeur absolue de 1, alors cette variable à tester se rapproche d'autant plus de la variable temps.

Nous régressions linéairement la variable à tester par rapport au temps. Le signe du coefficient du temps est positif, lorsque la variable à tester a une tendance linéaire décroissante dans le temps.
Ci-dessous les statistiques sommaires issues des 119 régressions (119 pays).

Summary(coef.Intercept)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -1.522   1.635   1.835   1.673   1.911   1.950
```

Summary(coef.year)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -2.125  -2.082  -1.999  -1.823  -1.781   1.658
```

Summary(coef.R2)

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
0.0000069 0.7008520 0.8832064 0.7890880 0.9577360 0.9974553
```

Pour 118 régressions (sur 119) soit 99%, le coefficient du temps (coef.year), est systématiquement négatif, confirmant ainsi l'hypothèse d'une baisse tendancielle du taux d'emplois agricoles (lissé), pour l'ensemble des pays.
L'exception est le Zimbabwe où la hausse du taux d'emplois agricole s'explique probablement par les politiques de retour à la terre.

Figure 3.10 Exemple de la France. Taux d'emplois agricoles (lissé)

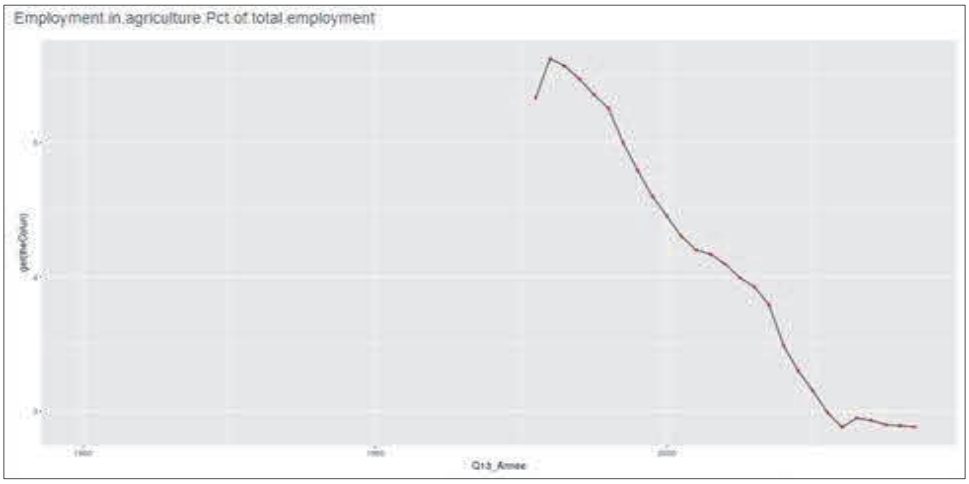
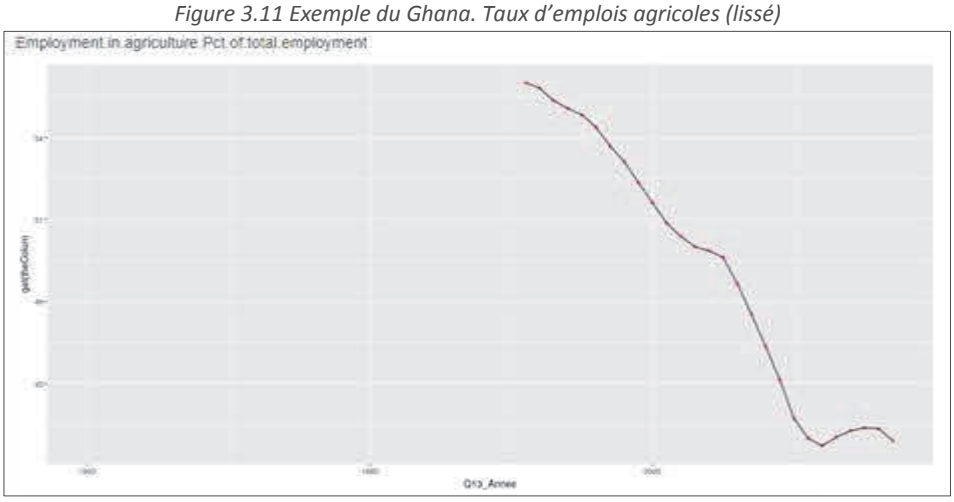


Figure 3.11 Exemple du Ghana. Taux d'emplois agricoles (lissé)



Conclusion

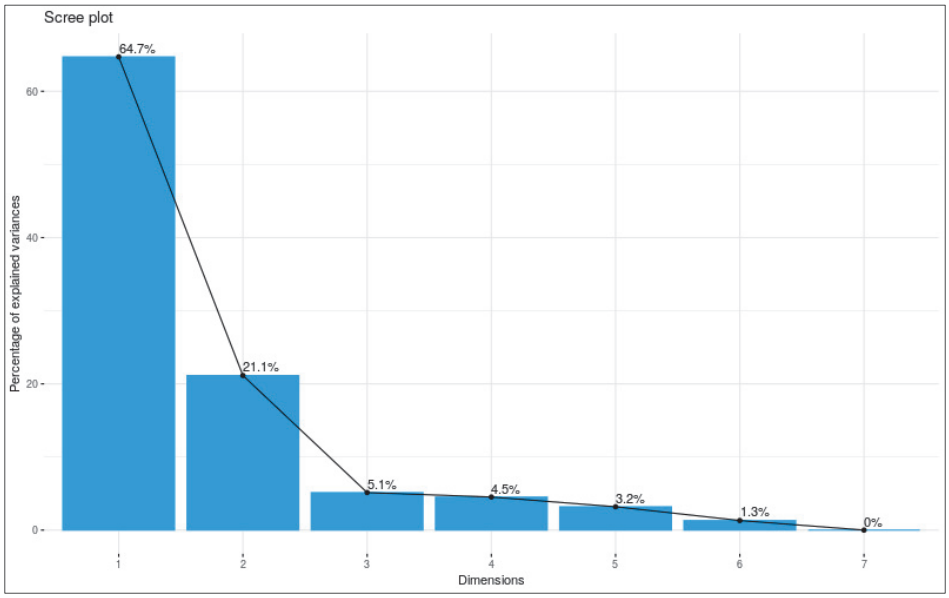
En conclusion, le taux d'emplois agricoles (lissé), est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale, du fait de sa tendance baissière généralisée (99% des pays) sur les 26 dernières années.

4.1.5. Classification des pays, sur composantes principales

Sélection des composantes principales pour la classification

L'analyse en composantes principales (ACP) est réalisée à partir des 7 variables actives (Tableau 2.1).

Figure 3.12 Pourcentage de variance expliquée par les axes principaux (ou dimensions)



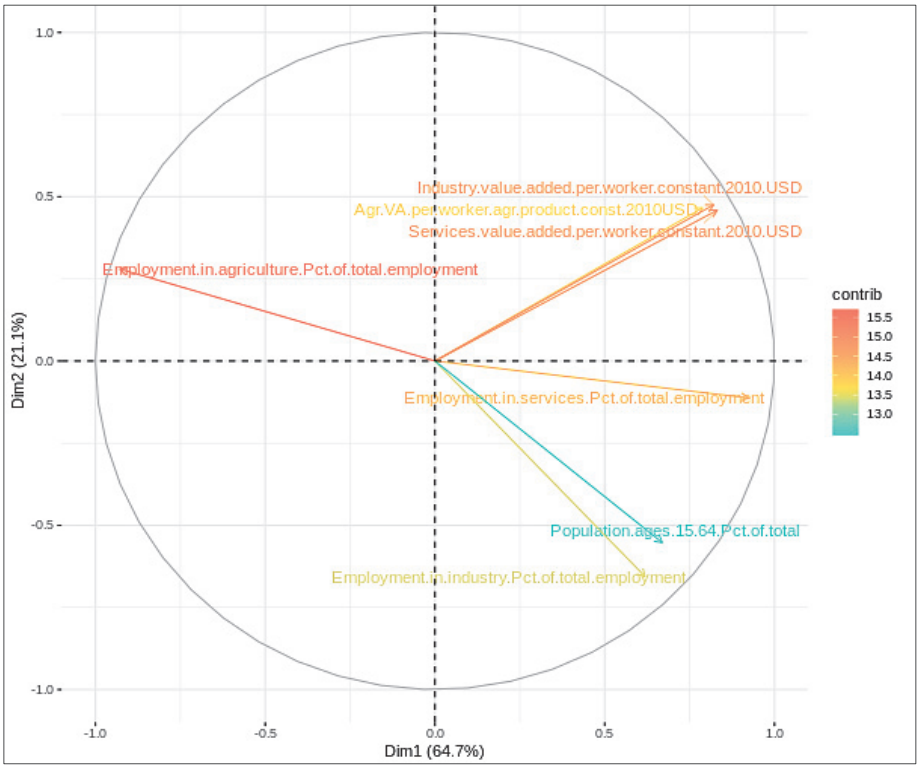
Les deux premiers axes principaux expriment 85.8% de l'inertie totale du jeu de données ; cela signifie que 85.8% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan.

Signification des composantes principales sélectionnées

Figure 3.13 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives

Employment.in.agriculture.Pct.of.total.employment	Emploi dans l'agriculture (% de l'emploi total) (estimation modélisée du BIT
Employment.in.industry.Pct.of.total.employment	Emploi dans l'industrie (% de l'emploi total) (estimation modélisée du BIT)
Employment.in.services.Pct.of.total.employment	Emploi dans les services (% de l'emploi total) (estimation modélisée du BIT)
Population.ages.15.64.Pct.of.total	Population âgée de 15 à 64 ans (% du total)
Agr.VA.per.worker.agr.product.const.2010USD	Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010)
Industry.value.added.per.worker.constant.2010.USD	Industrie, valeur ajoutée par travailleur (USD constants de 2010)
Services.value.added.per.worker.constant.2010.USD	Services, valeur ajoutée par travailleur (USD constants de 2010)

Figure 3.14 Cercle de corrélation entre axes principaux (Dim) et variables actives (avec leur contribution à la création des axes)



En analysant les Figure 3.13 et Figure 3.14, nous déduisons que la première composante (Dim 1) oppose les pays ayant un fort taux d'emplois dans l'agriculture aux autres pays : Les pays ayant les plus forts taux d'emplois dans l'agriculture, ont les valeurs les plus négatives sur la première composante.
A contrario, la première composante (Dim 1) a des valeurs d'autant plus positives que les productivités des emplois agricoles et industriels, le taux d'emplois dans l'industrie et les services, ainsi que le pourcentage de population en âge de travailler sont plus élevés, et les taux d'emplois agricoles sont plus faibles.

La deuxième composante (Dim 2) oppose les pays ayant des valeurs élevées pour le taux d'emplois industriels et le pourcentage de population en âge de travailler, aux pays ayant des valeurs élevées de productivités des emplois agricoles et industriels.
La deuxième composante (Dim 2) a des valeurs d'autant plus positives que les productivités des emplois agricoles et industriels sont plus élevées, et le taux d'emplois industriels plus faible.

Choix du nombre optimal de classes de pays

Après classification statistique des pays, nous obtenons quatre classes optimales de pays, qui seront décrites, après ordonnancement par le temps transitionnel qui est le taux d'emplois agricoles lissé.

Figure 3.15 Gain d'inertie en fonction du nombre de classes

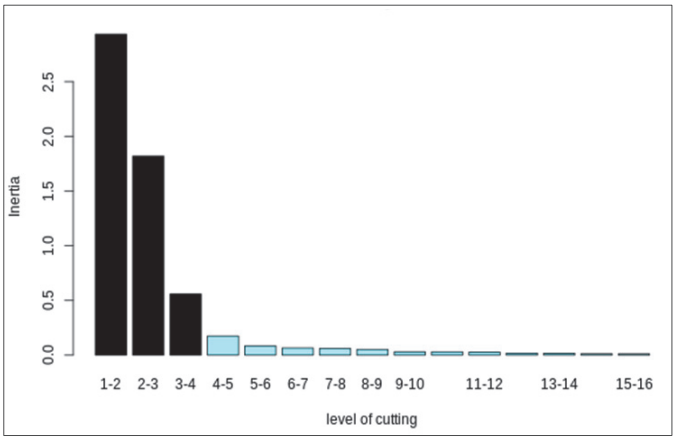
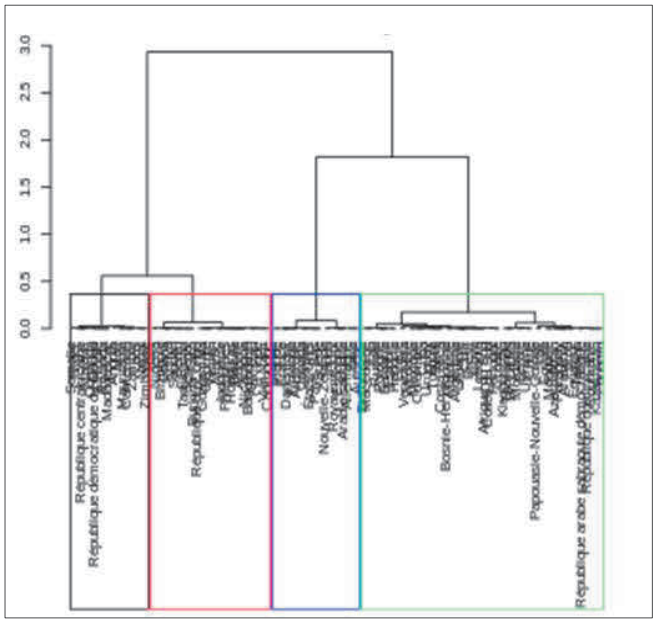


Figure 3.16 Dendrogramme de la classification ascendante hiérarchique



4.1.6. Indexation temporelle des classes de pays

Le temps transitionnel utilisé est l'opposé du taux d'emplois agricoles, lissé par une moyenne mobile d'ordre 5 : En effet si le taux d'emplois agricoles est décroissant dans le temps, son opposé est croissant dans le temps.

Etant donné que le taux d'emplois agricoles est décroissant (en fonction du temps), les classes de pays sont numérotées par ordre décroissant de leur moyenne : ainsi, la classe 4 qui a la moyenne de taux d'emplois agricoles lissé le plus faible, porte l'index temporel le plus élevé (c'est la classe la plus ancienne dans le processus de transition) ; et inversement pour la classe 1 qui est à l'autre bout du spectre.
Nous qualifions les classes ainsi indicées par l'âge transitionnel moyen de leurs membres, de *classes d'âge transitionnel*.

4.1.7. Description statistique des classes d'âge transitionnel de pays

Figure 3.17 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels

Cette représentation des classes d'âge transitionnel de pays, suggère un parcours temporel sous forme d'un V en allant de la gauche vers la droite.

Dans cette vue temporelle-transitionnelle, la classe 4, représente les pays plus anciens dans le processus de transition économique. A l'opposé, les pays de la classe 1 sont les pays les plus récents dans le processus de transition économique.

Ces 4 classes illustrent également les 4 phases observables « actuellement » dans le processus de transition économique. La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transition économique.

Figure 3.18 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes

La représentation elliptique ci-dessus suggère que les frontières des classes ne sont pas si tranchées que le laisse supposer la représentation en polygones.

Tableau 3.2 Description des classes selon les variables actives

Classes	Emploi dans l'agriculture (% de l'emploi total)	Emploi dans les services (% de l'emploi total)	Emploi dans l'industrie (% de l'emploi total)	Population âgée de 15 à 64 ans (% de la population totale)	Valeur ajoutée agricole par travailleur (produit agricole, $US constant 2010)	Valeur ajoutée industrielle par travailleur ($US constant 2010)	Valeur ajoutée des services par travailleur ($US constant 2010)
1	72,1	19,6	8,3	53,1	2 317	9 222	7 302
2	37,6	44,1	*NS*	*NS*	4 351	13 489	9 756
3	13,5	61,2	25,3	67,5	*NS*	*NS*	*NS*
4	3,4	75,1	*NS*	66,2	54 100	110 000	84 400
Monde	29,0	51,3	19,7	63,2	14 818	33 763	26 435

NS : valeur non-significativement différente de la moyenne mondiale au seuil de 5%

La **classe 1** est composée de pays tels que République centrafricaine, Malawi, Niger et Somalie. Ce groupe est caractérisé par :

- de fortes valeurs pour le taux d'emplois agricoles (71% dans la classe contre 30% au niveau mondial)
- de faibles valeurs pour le taux d'emplois dans l'industrie et les services, la population en âge de travailler, les productivités de l'industrie et de l'agriculture.

La **classe 2** est composée de pays (tels que Guatemala, Honduras, Philippines, R. Congo, Géorgie) partageant :

- de fortes valeurs pour le taux d'emplois agricoles (38% dans la classe contre 30% au niveau mondial)

- de faibles valeurs pour les productivités de l'industrie et de l'agriculture et le taux d'emplois dans les services.

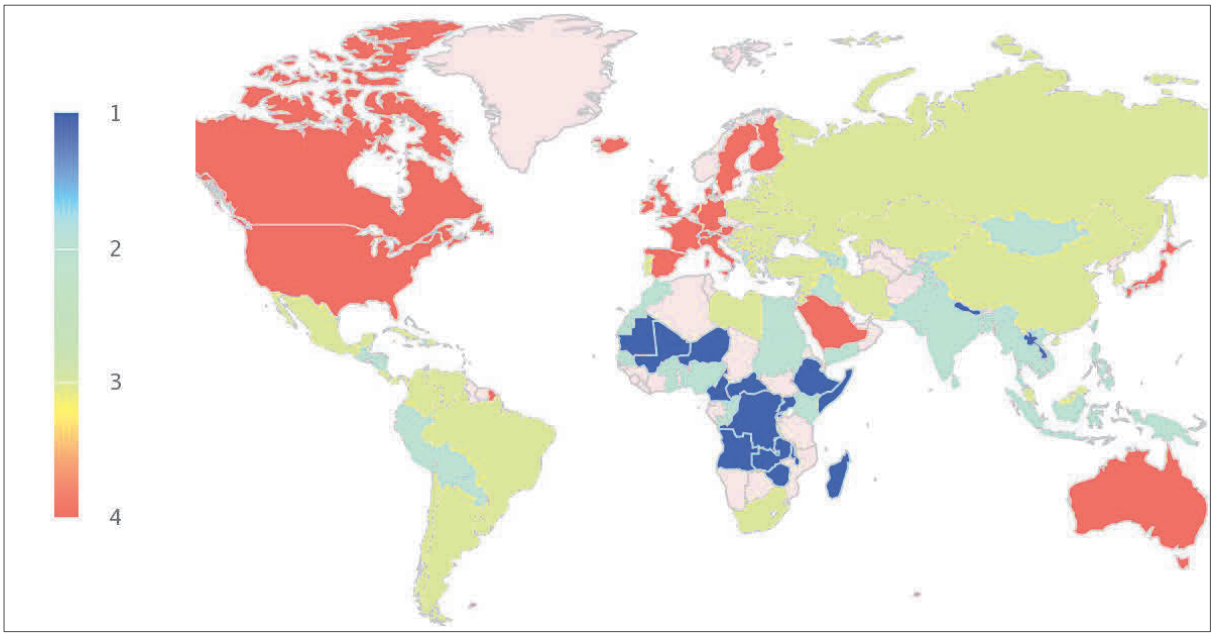
La **classe 3** est composée de pays tels que Brésil, Jordanie, Croatie, Malaisie, Afrique du Sud. Ce groupe est caractérisé par :

- de fortes valeurs pour les taux d'emplois industriels (25% dans la classe contre 20% au niveau mondial), le pourcentage de population en âge de travailler et les emplois dans les services.
- de faibles valeurs pour le taux d'emplois agricoles (14% dans la classe contre 30% au niveau mondial).

La **classe 4** est composée de pays tels que : Australie, Canada, Danemark, Islande, Irlande, Pays-Bas, Suède, Suisse et États-Unis. Ce groupe est caractérisé par :

- de fortes valeurs pour les productivités des emplois industriels et agricoles, le taux d'emplois dans les services, et le pourcentage de population en âge de travailler.
- de faibles valeurs pour l'emploi agricole (3.4% dans la classe contre 30% au niveau mondial).

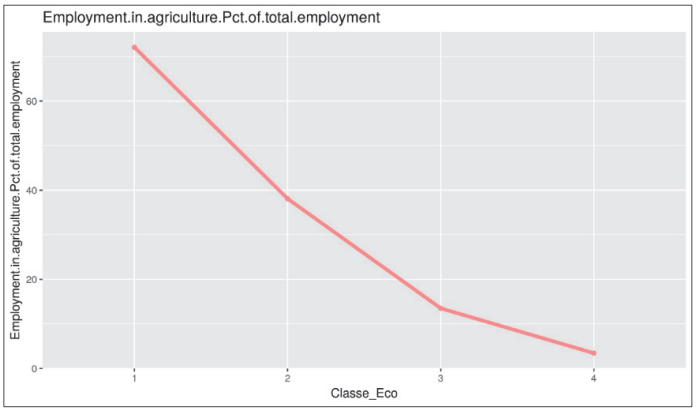
Figure 3.19 Cartographie des Classes temporelles de Pays



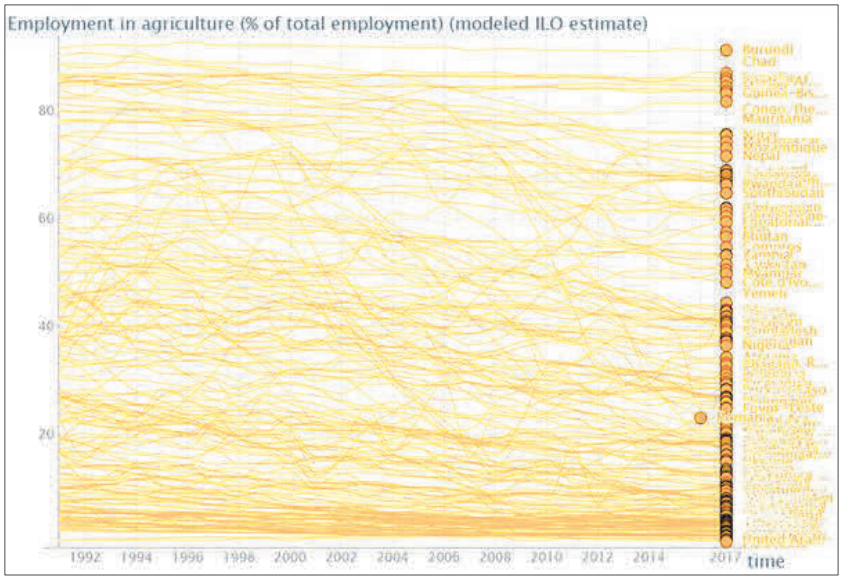
4.1.8. Variables de transition : courbes transitionnelles transversales selon l'indice d'âge transitionnel moyen

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales, sont nécessaires pour envisager la construction d'un schéma transitionnel à partir de l'analyse transversale.

Figure 3.20 Emploi dans l'agriculture (% de l'emploi total)

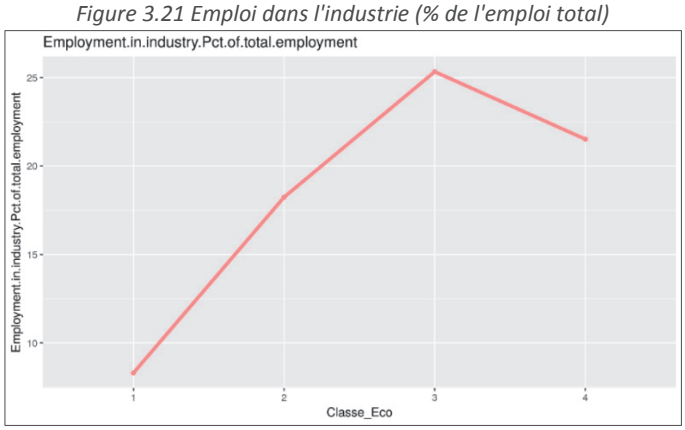


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures baissières similaires.

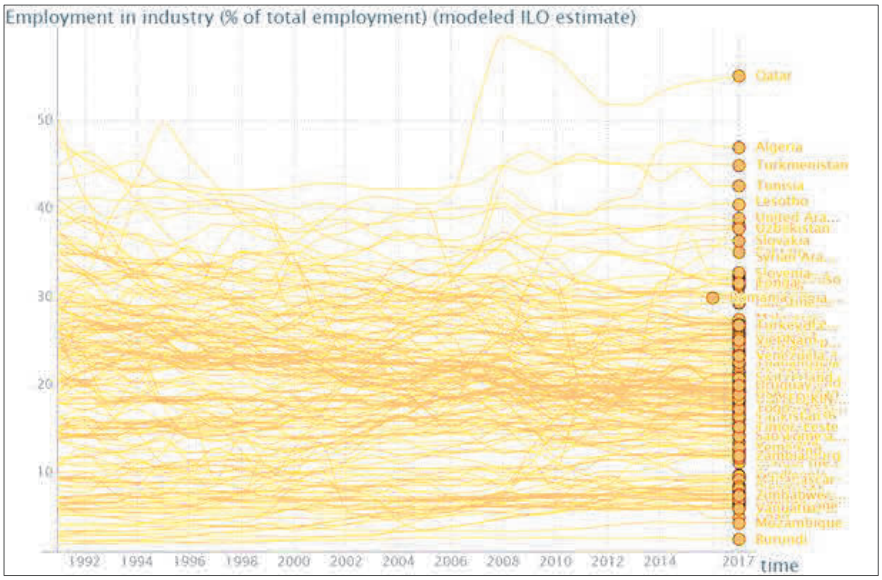


Source : Word Bank (2018)

Figure 3.21 Emploi dans l'industrie (% de l'emploi total)

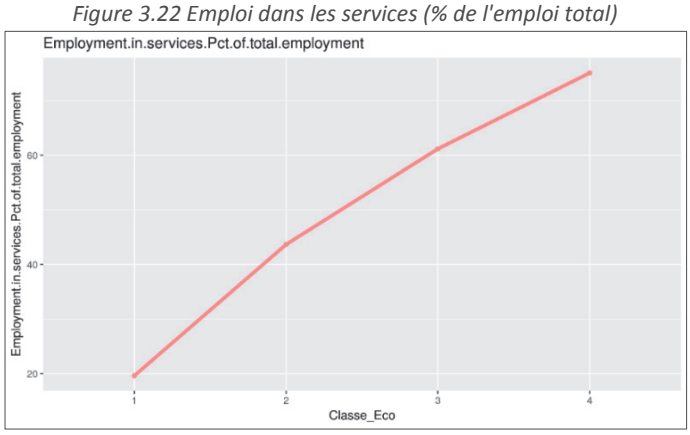


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.

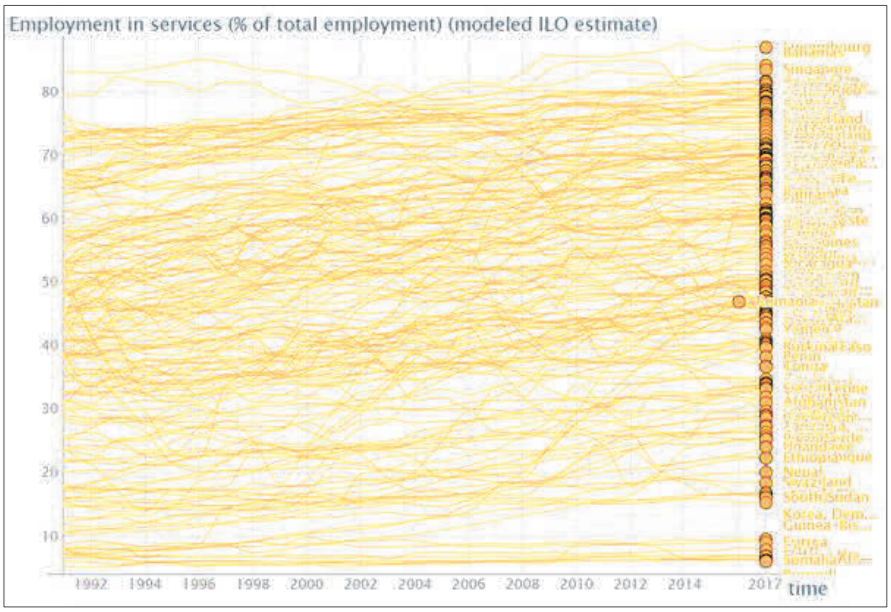


Source : Word Bank (2018)

Figure 3.22 Emploi dans les services (% de l'emploi total)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.

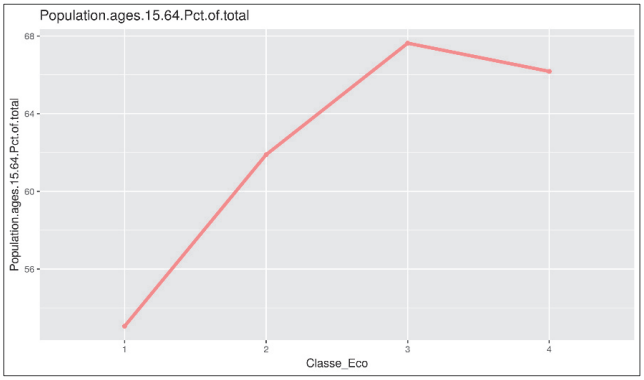


Source : Word Bank (2018)

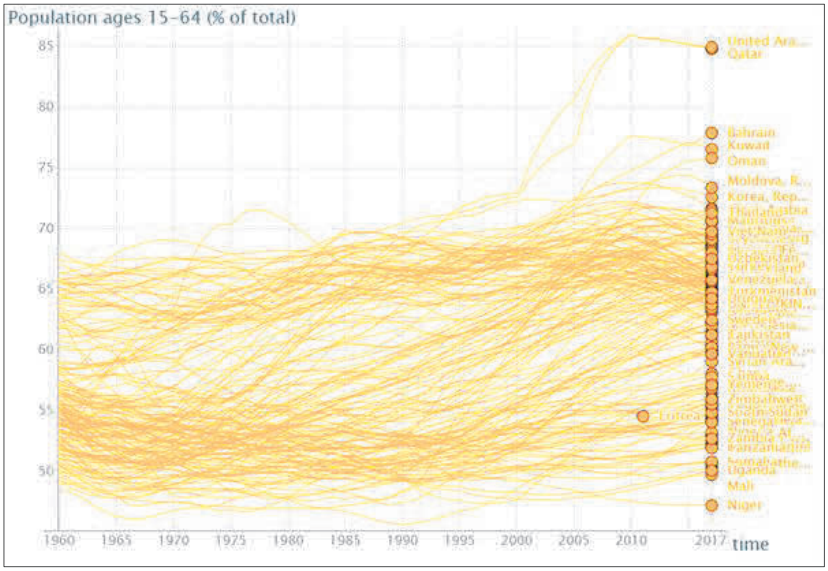
4.1.9. Autres variables actives : courbes transitionnelles transversales selon le temps transitionnel, avec la moyenne de la variable par classe d'âge transitionnel

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales, sont nécessaires pour envisager la construction d'un schéma transitionnel à partir de l'analyse transversale.

Figure 3.23 Population âgée de 15 à 64 ans (% de la population totale)



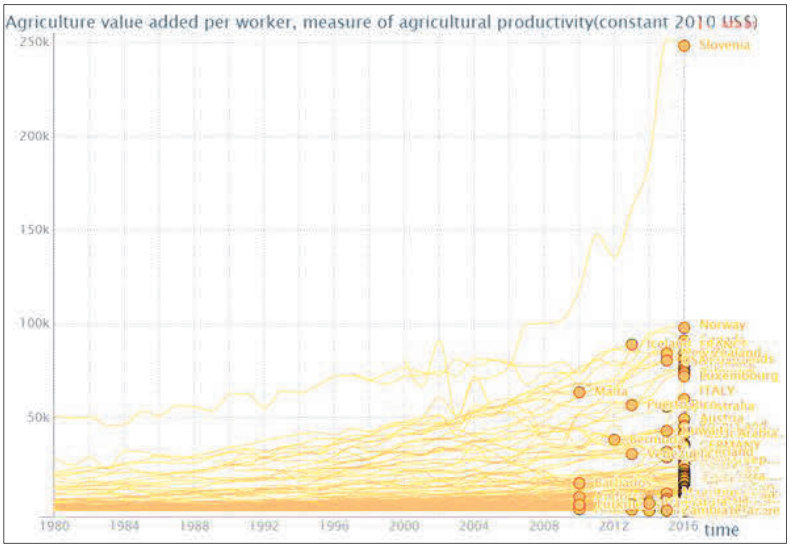
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.



Source : Word Bank (2018)

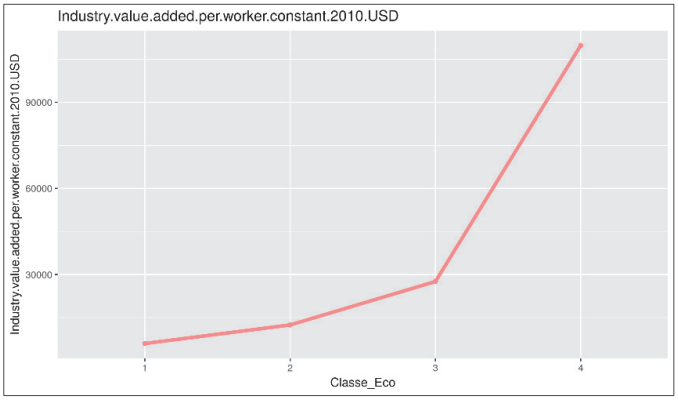
Figure 3.24 Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010)

La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.

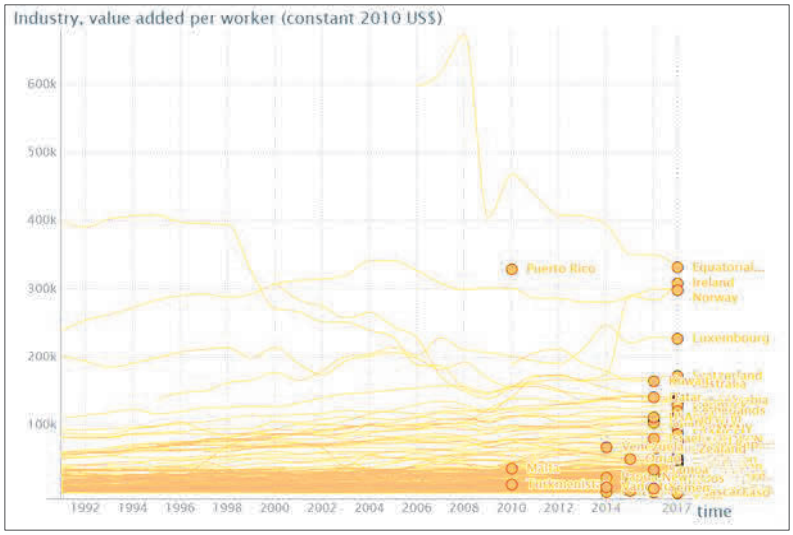


Source : Word Bank (2018)

Figure 3.25 Industrie, valeur ajoutée par travailleur (USD constants de 2010)

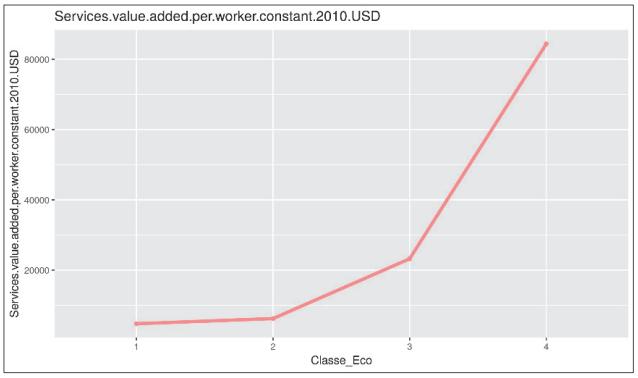


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.

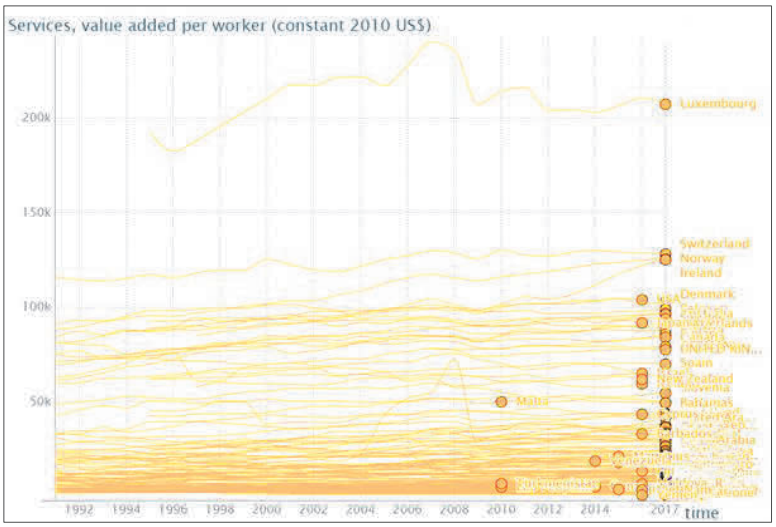


Source : Word Bank (2018)

Figure 3.26 Services, valeur ajoutée par travailleur (USD constants de 2010)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.



Source : Word Bank (2018)

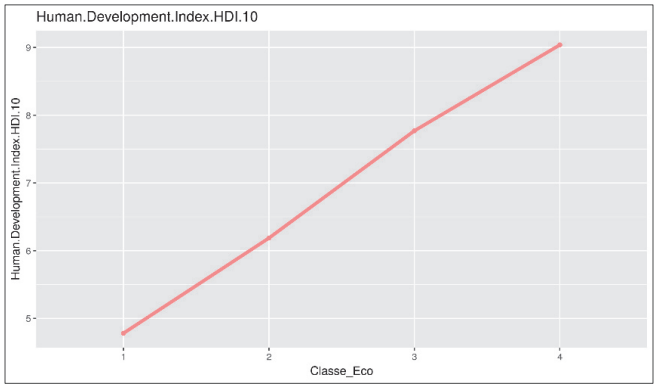
4.1.10. Variables passives : Comparaison des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales

Les variables passives sont des variables utilisées pour illustrer les classes, et qui ne sont pas des variables actives (qui elles sont utilisées pour créer les classes).

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales des variables passives, sont nécessaires pour vérifier la robustesse et la cohérence d'ordonnancement des classes d'âge transitionnelles construites.
La similitude globale des allures des courbes transversales et longitudinales que nous constatons dans les comparaisons ci-après dans plusieurs domaines, montre empiriquement cette robustesse.

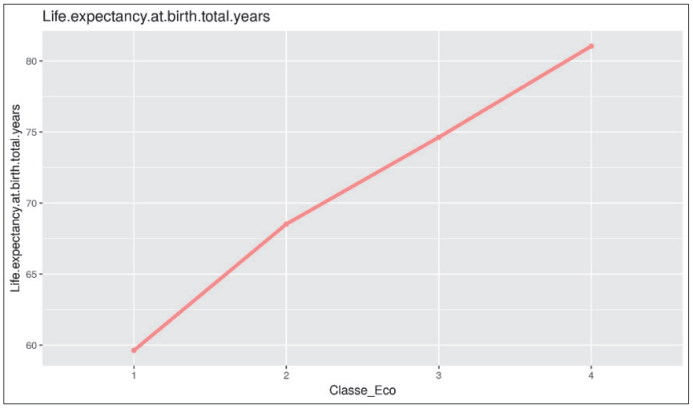
Moyennes par classe d'âge transitionnel

*Figure 3.27 Indice de développement humain (HDI*10)*



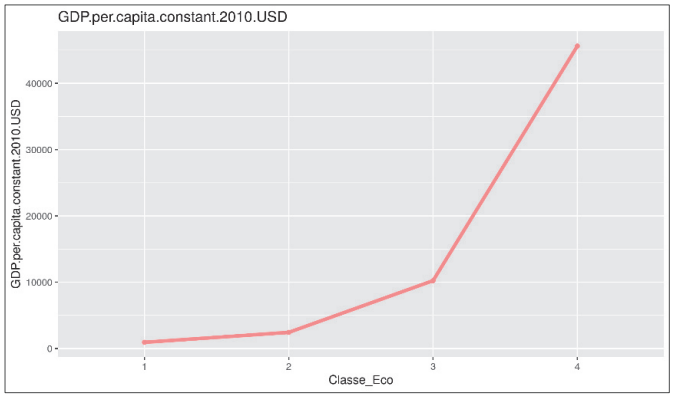
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.1*) nous permet de conclure à des allures haussières similaires.

Figure 3.28 Espérance de vie à la naissance, Total



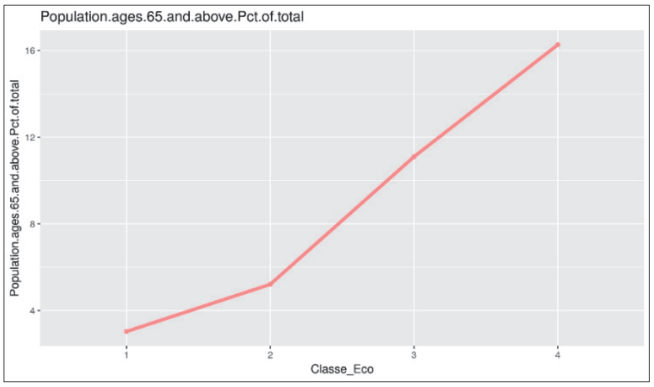
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.2*) nous permet de conclure à des allures haussières similaires.

Figure 3.29 PIB par habitant (USD constants de 2010)



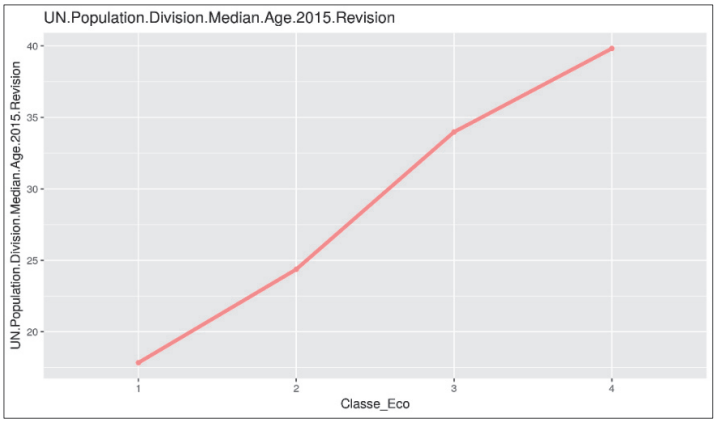
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.3*) nous permet de conclure à des allures haussières similaires.

Figure 3.30 Population âgée de 65 ans et plus (% de la population totale)



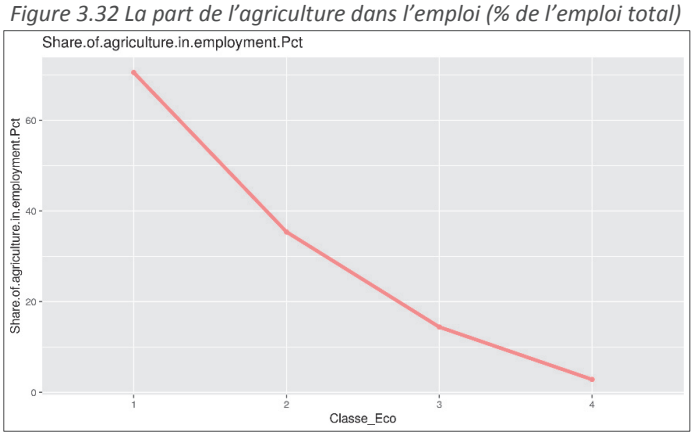
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.4*) nous permet de conclure à des allures haussières similaires.

Figure 3.31 Division de la population des Nations Unies (âge médian)



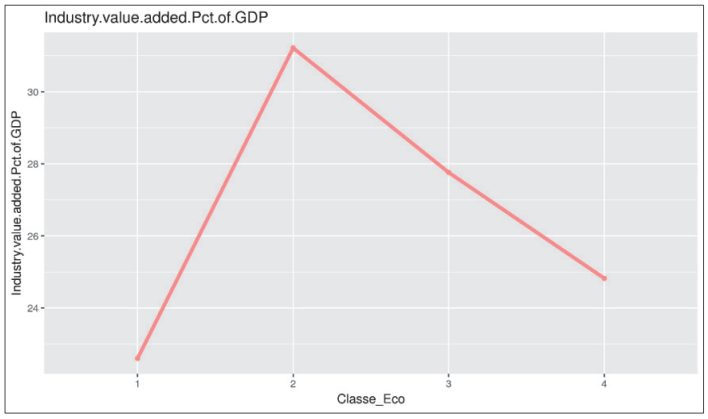
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.5)* nous permet de conclure à des allures haussières similaires.

Figure 3.32 La part de l'agriculture dans l'emploi (% de l'emploi total)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.6*) nous permet de conclure à des allures baissières similaires.

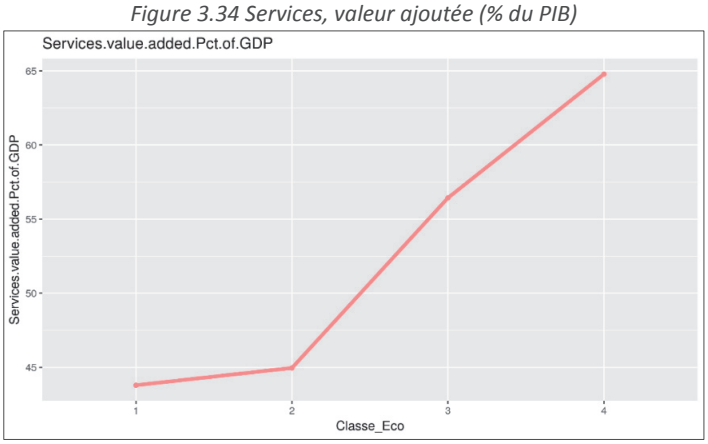
Figure 3.33 Industrie, valeur ajoutée (% du PIB)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (

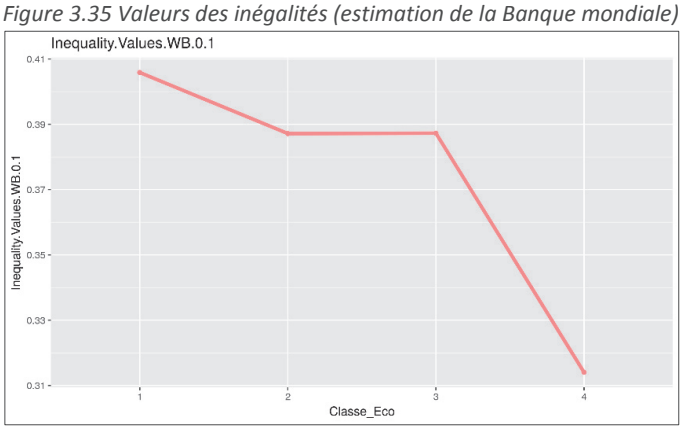
Figure 0.7) nous permet de conclure à des allures en cloche similaires.

Figure 3.34 Services, valeur ajoutée (% du PIB)



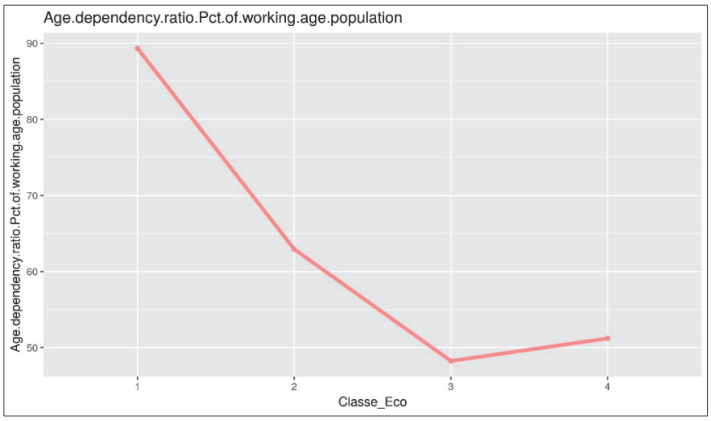
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.8*) nous permet de conclure à des allures haussières similaires.

Figure 3.35 Valeurs des inégalités (estimation de la Banque mondiale)



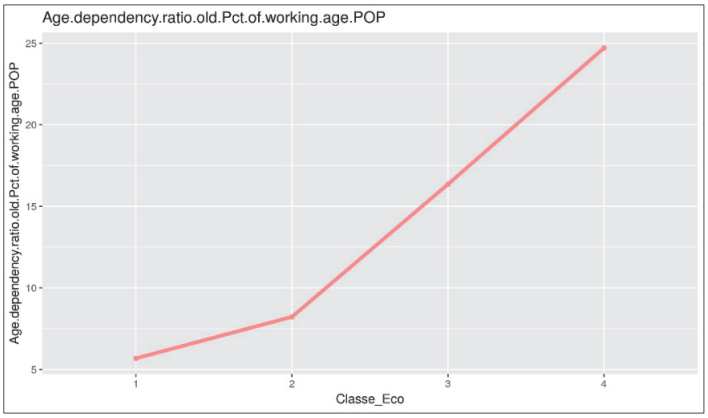
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.9)* nous permet de conclure à des allures baissières similaires.

Figure 3.36 Ratio de dépendance à l'âge (% de la population en âge de travailler)



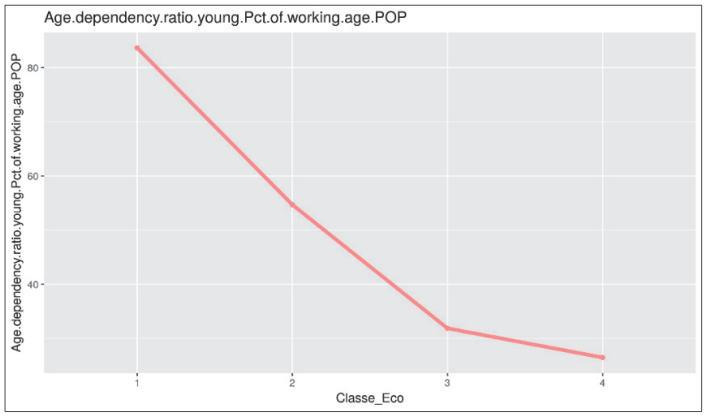
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.10*) nous permet de conclure à des allures baissières similaires.

Figure 3.37 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler)



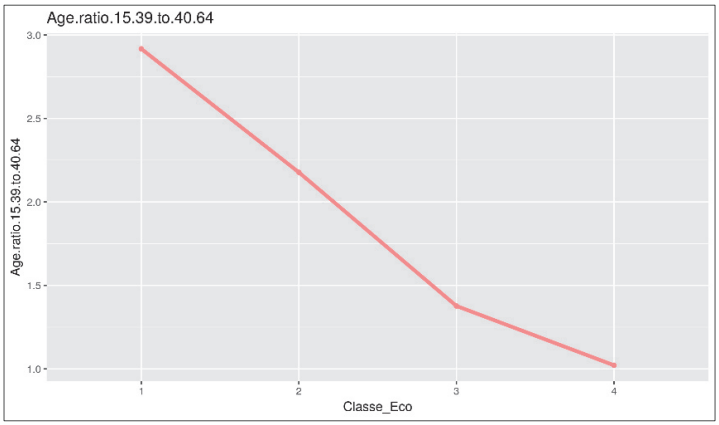
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.11*) nous permet de conclure à des allures haussières similaires.

Figure 3.38 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler)



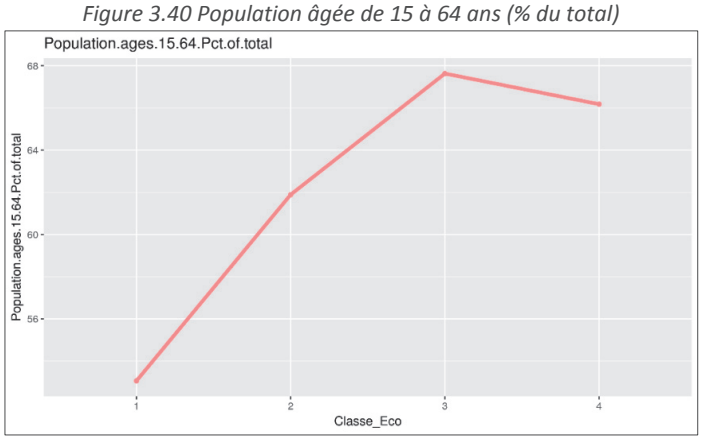
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.12*) nous permet de conclure à des allures baissières similaires.

Figure 3.39 Ratio d'âge 15-39 à 40-64



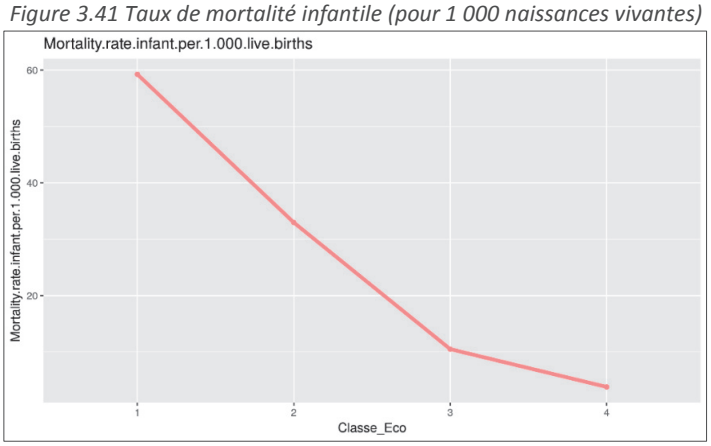
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.13*) nous permet de conclure à des allures baissières similaires.

Figure 3.40 Population âgée de 15 à 64 ans (% du total)



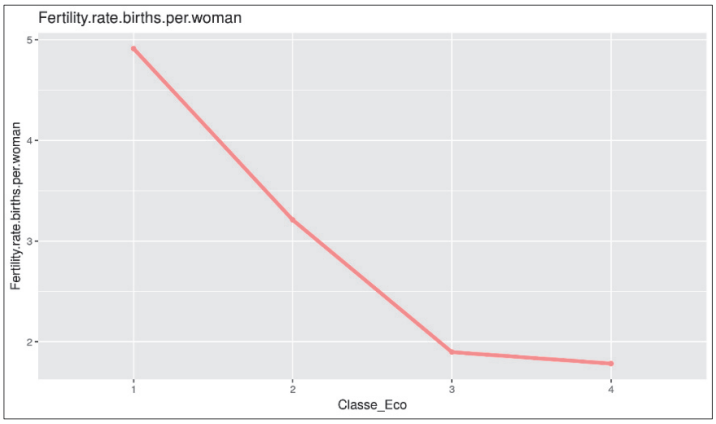
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.14)* nous permet de conclure à des allures haussières similaires.

Figure 3.41 Taux de mortalité infantile (pour 1 000 naissances vivantes)



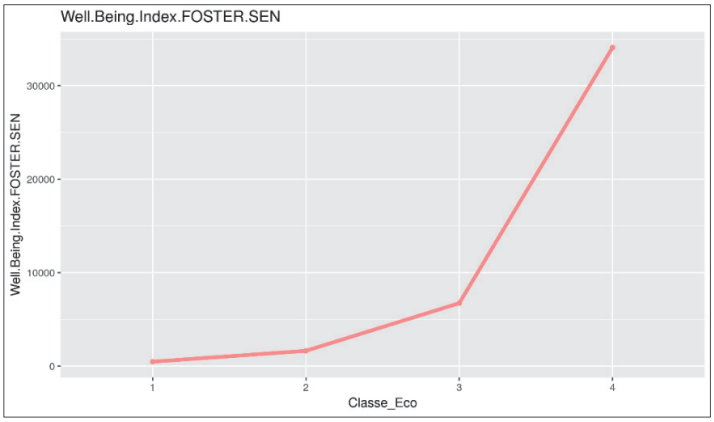
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.15*) nous permet de conclure à des allures baissières similaires.

Figure 3.42 Taux de fécondité, total (naissances par femme)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.16*) nous permet de conclure à des allures baissières similaires.

Figure 3.43 Indice de bien-être (FOSTER-SEN)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.17*) nous permet de conclure à des allures haussières similaires.

4.2. Comparaison du schéma transitionnel par coupe transversale avec le schéma théorique longitudinal de la transition économique

Variables de transition : schéma transitionnel transversal selon l'âge transitionnel

Figure 3.44 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne par classe d'âge transitionnel

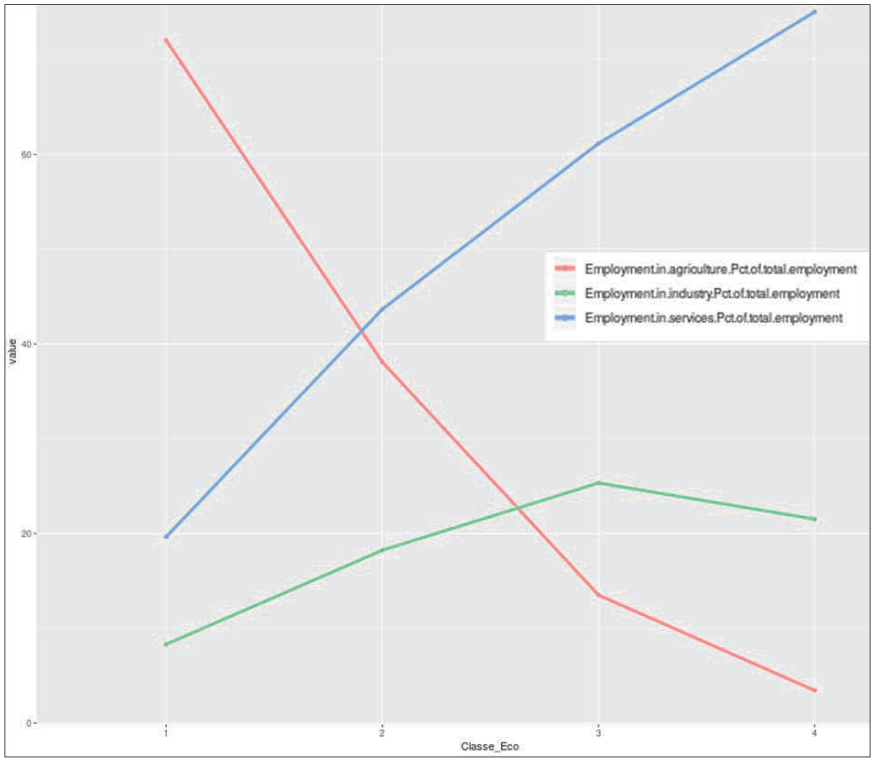
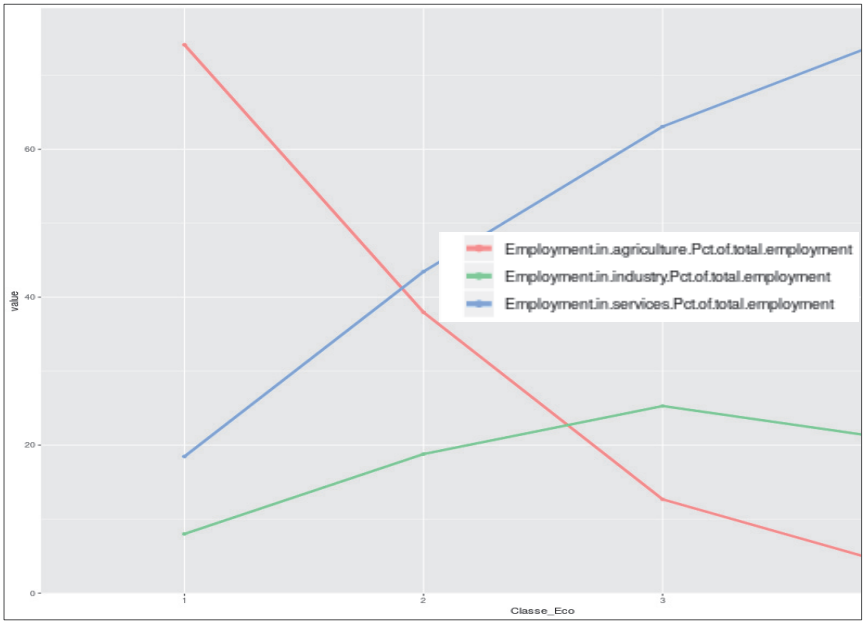


Figure 3.45 Schéma transitionnel transversal selon l'âge transitionnel. Médiane par classe d'âge transitionnel



L'allure des deux schémas ci-dessus est sensiblement la même, aussi bien en utilisant les moyennes que les médianes : C'est l'illustration d'une robustesse du schéma transitionnel obtenu en coupe transversale.

Figure 3.46 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse

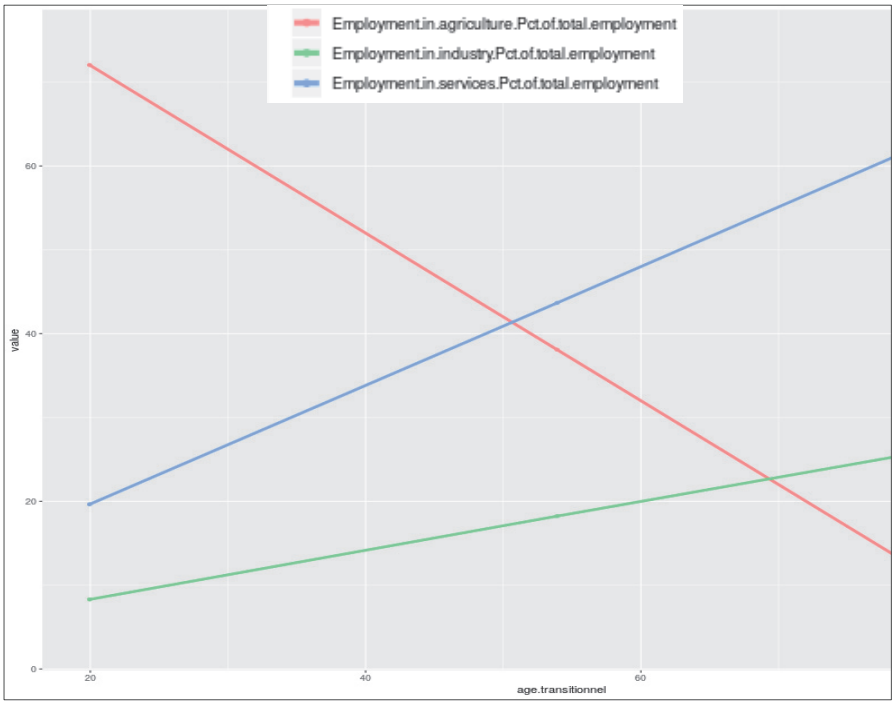
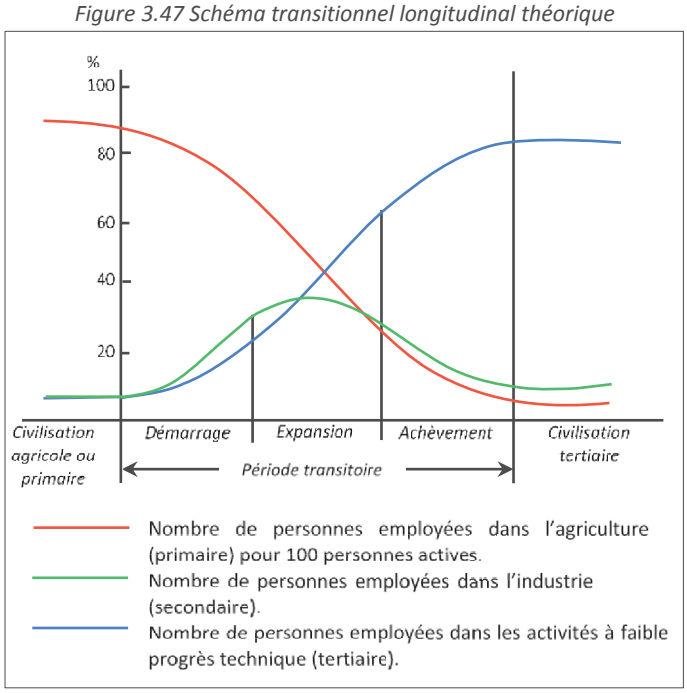


Figure 3.47 Schéma transitionnel longitudinal théorique



Source : Fourastié (1989), p. 148.

En comparant le schéma transitionnel transversal (*Figure 3.44*), avec le schéma transitionnel longitudinal théorique de Fourastié (1989) ci-dessus, nous observons la même dynamique des trois variables de transition.

En faisant le parallèle avec les phases de la transition économique (Fourastié, 1989), nous obtenons la situation mondiale actuelle suivante :

Figure 3.48 Correspondance entre Classes d'âge transitionnelles et Phases théoriques de la transition économique

Classe d'âge transitionnel	% de pays	Moyenne, Taux d'emplois agricoles	Moyenne, Taux d'emplois industriels	Moyenne, Taux d'emplois services	Phase Fourastié, Majoritaire	Phase Fourastié, Minoritaire
1	16%	71%	9%	21%	Civilisation agricole ou primaire	Démarrage
2	36%	38%	18%	44%	Expansion	Démarrage
3	32%	14%	25%	61%	Achèvement	
4	17%	3,4%	22%	75%	Achèvement	Civilisation tertiaire

L'étude réalisée confirme assez clairement la validité et la réalité actuelle du processus de transition économique présenté par Fourastié (1949).
Les phases transitionnelles économiques offrent une grille de lecture dynamique et pleine de sens sur la situation structurelle passée et future proche des pays de la planète.

5. La transition démographique

Nous allons ici utiliser le temps et l'âge transitionnel pour réaliser une inférence des trajectoires longitudinales dans une coupe transversale dans le domaine démographique.

5.1. Analyse transversale de la transition démographique

5.1.1. Sélection des variables à étudier

Les variables structurelles utilisées se répartissent en deux catégories : les variables de transition et les variables actives.

Variables de transition

Les « variables de transition » sont utilisées pour décrire le processus de transition démographique.

Variables actives : Elles sont utilisées pour effectuer la classification statistique des pays. Elles incluent les variables de transition auxquelles sont ajoutées des variables d'homogénéisation des classes. Dans le cas présent, il n'y a pas de variables d'homogénéisation : Les variables actives sont les mêmes que les variables de transition.

Tableau 3.3 Variables actives ou de transition

Variable	Signification en anglais	Signification en français	Variable active	Variable de transition
Death.rate.crude.per.1.000.people	Death rate, crude (per 1,000 people)	Taux de mortalité, brut (pour 1 000 personnes)	oui	oui
Birth.rate.crude.per.1.000.people	Birth rate, crude (per 1,000 people)	Taux de natalité, brut (pour 1 000 personnes)	oui	oui

5.1.2. Transformation préalable des variables à étudier

Les variables actives utilisées font l'objet des transformations suivantes, avant la sélection de l'année d'étude : estimation des valeurs manquantes par interpolation spline, puis moyenne mobile d'ordre 5.

5.1.3. Sélection de l'année à étudier

Pour l'analyse transversale, l'année étudiée est 2014. Cela pour des raisons de complétude et de fiabilité des données, car les dernières (2017, source Banque Mondiale) sont souvent complétées et révisées après publication.

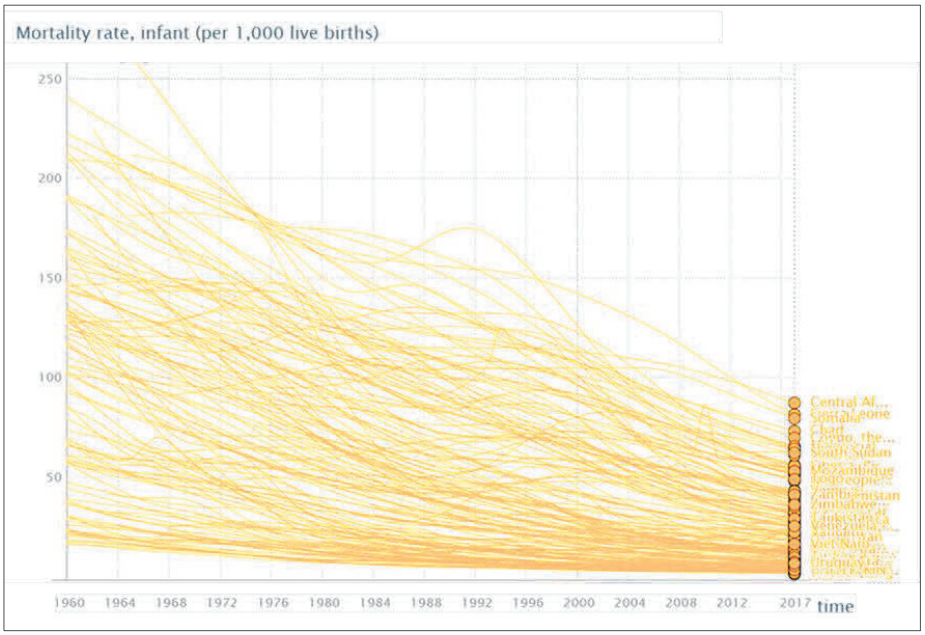
5.1.4. Choix du temps transitionnel

Variable à tester comme temps transitionnel

La variable à tester est une valeur transformée (moyenne mobile arrière d'ordre 5) du taux de mortalité infantile (pour 1000 naissances vivantes).

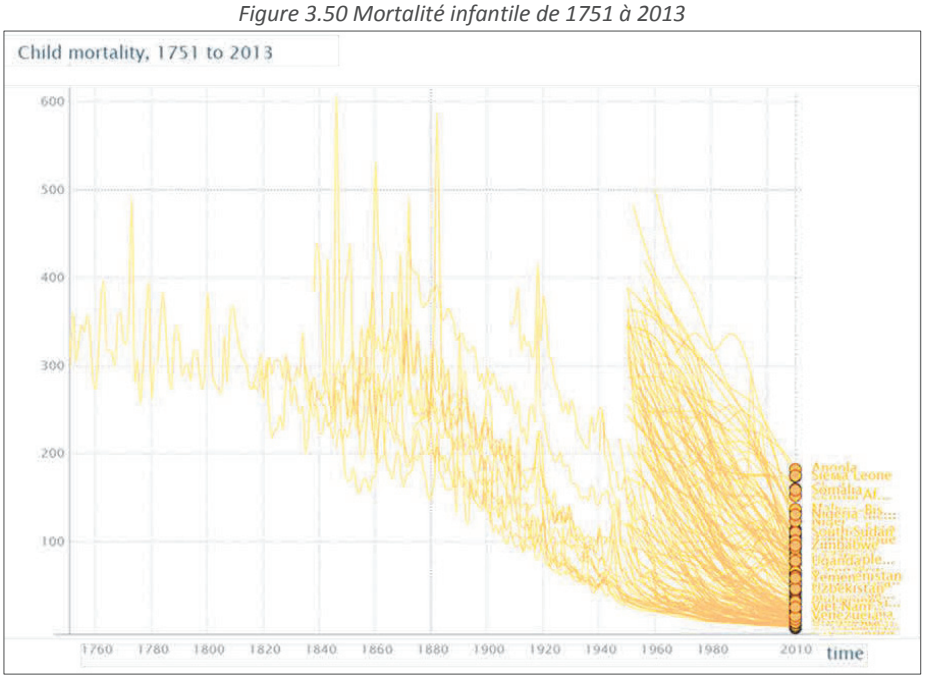
Le taux de mortalité infantile est choisi comme temps transitionnel car il baisse théoriquement durant la transition démographique.

Figure 3.49 Mortalité infantile (non-lissée), ensemble des pays du monde



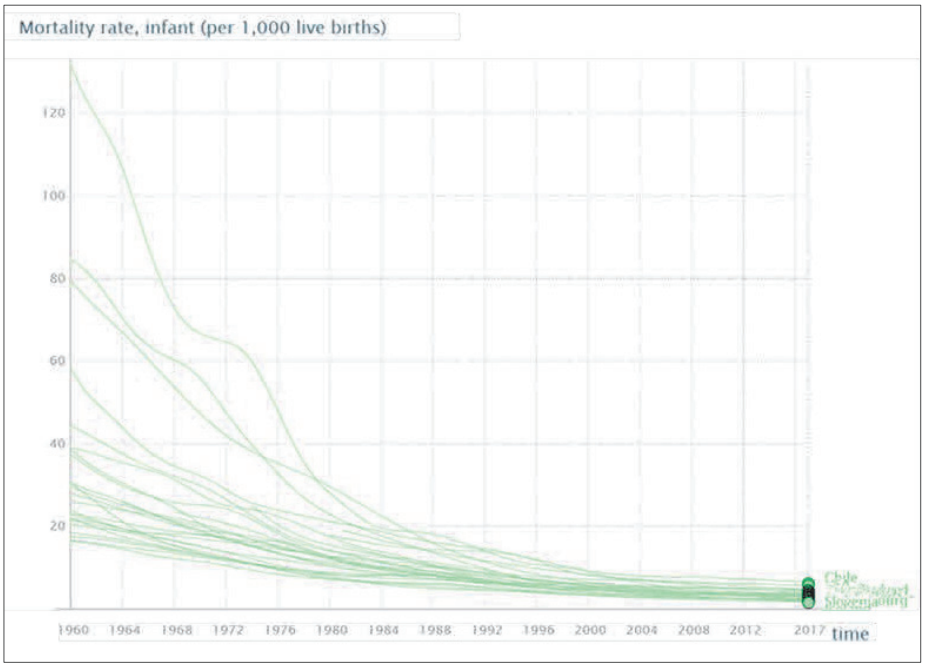
Source : World Bank (2018)

Figure 3.50 Mortalité infantile de 1751 à 2013



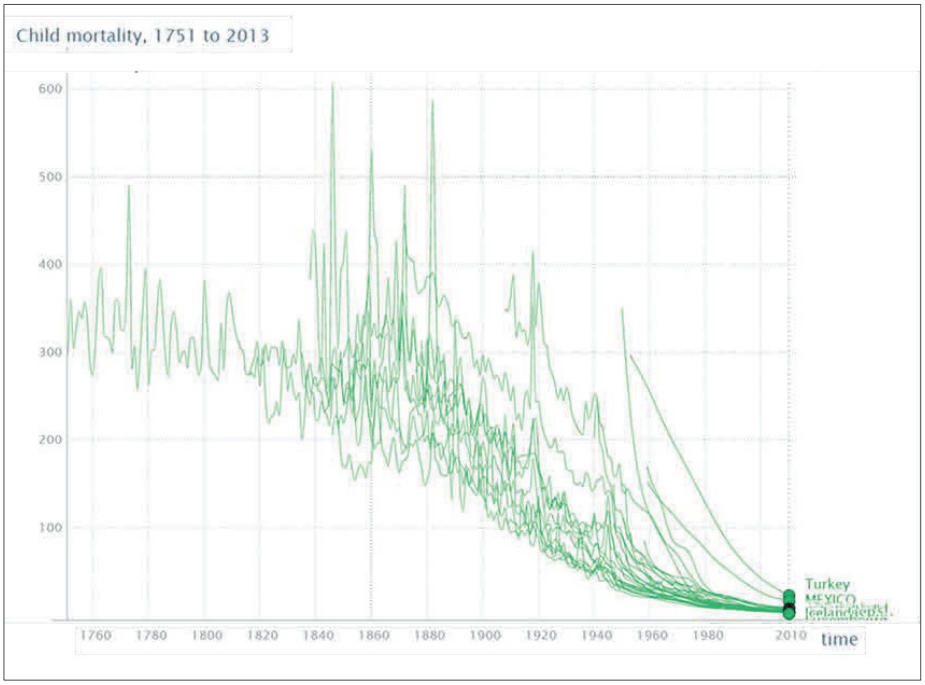
Source : Max Roser (2015)

Figure 3.51 Mortalité infantile des pays de l'OCDE



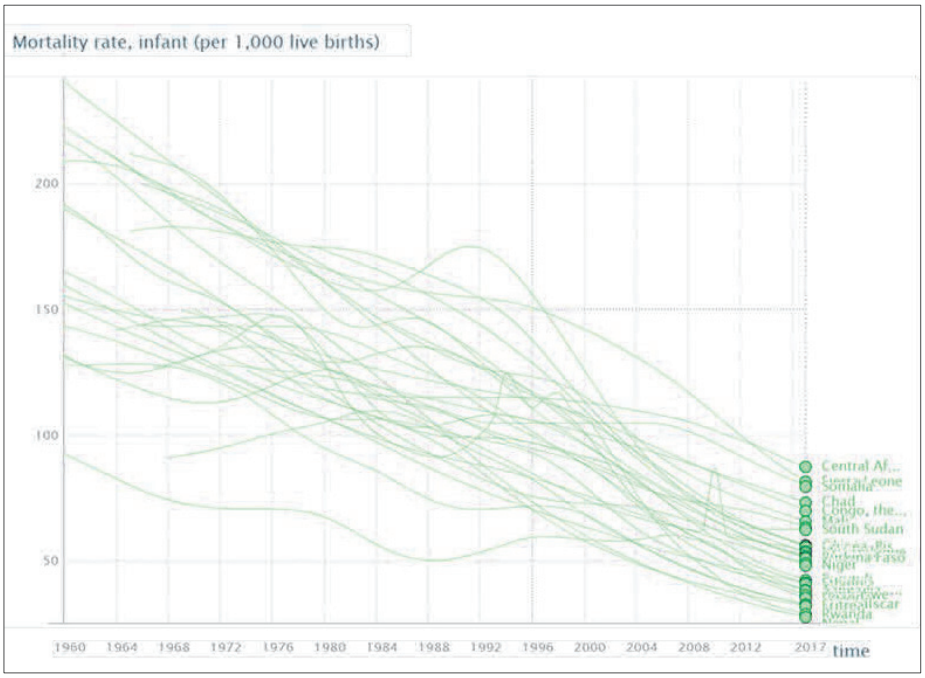
Source : World Bank (2018)

Figure 3.52 Mortalité infantile de 1751 à 2013



Source : Max Roser (2015)

Figure 3.53 Mortalité infantile (non-lissée) des pays à moyen revenu par habitant (au sens de la Banque Mondiale)



Source : World Bank (2018)

Figure 3.54 Mortalité infantile (non lissée) des pays à faible revenu par habitant (au sens de la Banque Mondiale)

Source : World Bank (2018)

Visuellement la tendance mondiale est à la baisse (et à une forme de convergence), même si on ne peut parler de monotonie mathématique : raison pour laquelle nous utilisons un lissage par moyenne mobile pour améliorer la monotonie de la variable pressentie pour jouer le rôle de temps transitionnel.

Vérification de la monotonie empirique de la variable à tester

La variable à tester Y est empiriquement monotone lorsque la série différenciée dY (dY=Y(t) – Y(t-1)) est de signe constant sur la période étudiée, pour tout pays.
De plus pour assurer l'universalité de cette monotonie empirique, le signe de série différenciée doit être le même pour tous les pays étudiés
Pour chaque pays, nous calculons le pourcentage de valeurs négatives de la série différenciée, afin de vérifier l'unicité du signe sur la période étudiée.
Sur l'ensemble des 130 pays étudiés, le pourcentage moyen de valeurs négatives de la série différenciée est de 96%, avec un minimum à 63% (Rwanda) et un maximum à 100% (USA).
Mieux, 91 des 130 pays, soit 70%, ont 100% de valeurs négatives dans la série différenciée.
Par ailleurs, 108 des 130 pays, soit 83%, ont plus de 90% de valeurs négatives dans la série différenciée.
Nous ne sommes pas en présence d'une monotonie parfaite, en ce qui concerne le taux de mortalité infantile lissé, même si cette monotonie est assurée empiriquement pour 70% des pays.

De plus, les valeurs centrales (moyenne=-0.96 et médiane=-1) du coefficient de Kendall indiquent une très forte relation monotone décroissante entre la mortalité infantile lissée et le temps.
Sur les 130 pays, le coefficient de Kendall est systématiquement négatif.

Vérification empirique de l'unicité du signe de la pente de la variable testée, quel que soit le pays

A défaut d'une monotonie empirique parfaite, l'unicité pour tous les pays, du signe de la pente de la droite d'ajustement de la variable à tester, montre une universalité de la tendance (à la hausse ou à la baisse) sur l'ensemble des pays étudiés.
De plus, si pour un pays, pente de la droite d'ajustement de la variable à tester est proche en valeur absolue de 1, alors cette variable à tester se rapproche d'autant plus de la variable temps.

Nous régressions linéairement la variable à tester par rapport au temps. Le signe du coefficient du temps est négatif, lorsque la variable à tester a une tendance linéaire décroissante dans le temps.
Ci-dessous les statistiques sommaires issues des 130 régressions (130 pays).

Summary(coef.Intercept)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 0.1022  0.1265  0.1306  0.3684  0.4817  1.6161       3
```

Summary(coef.year)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-1.8827 -1.2222 -1.0615 -1.1679 -1.0310 -0.8335       3
```

Summary(coef.R2)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 0.5999  0.8842  0.9330  0.9141  0.9701  0.9978       3
```

Sur les 130 régressions (130 pays) réalisées ici, le coefficient du temps (coef.year), est systématiquement négatif, confirmant ainsi l'hypothèse d'une baisse tendancielle de la mortalité infantile pour l'ensemble de pays.

Figure 3.55 Exemple de la France d'après-guerre (taux de mortalité infantile pour 1000 naissances vivantes lissé)

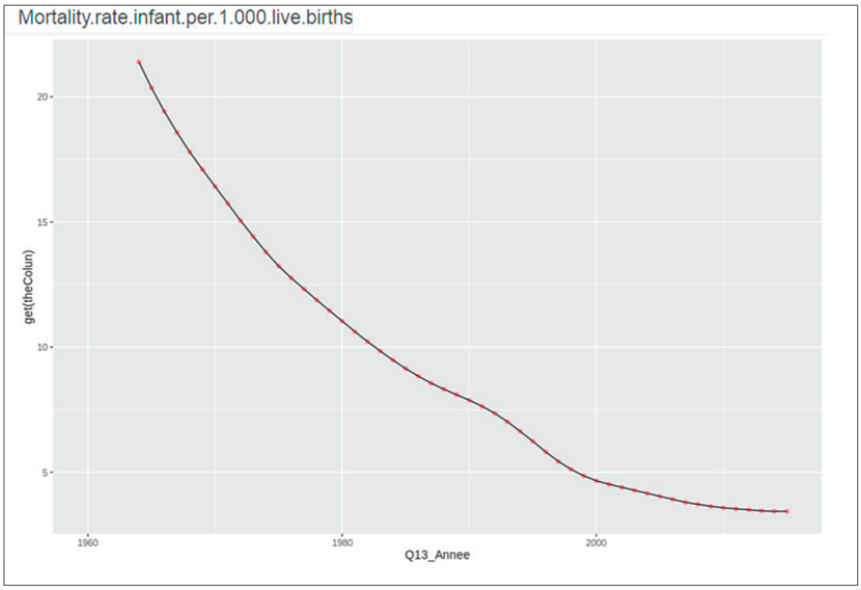
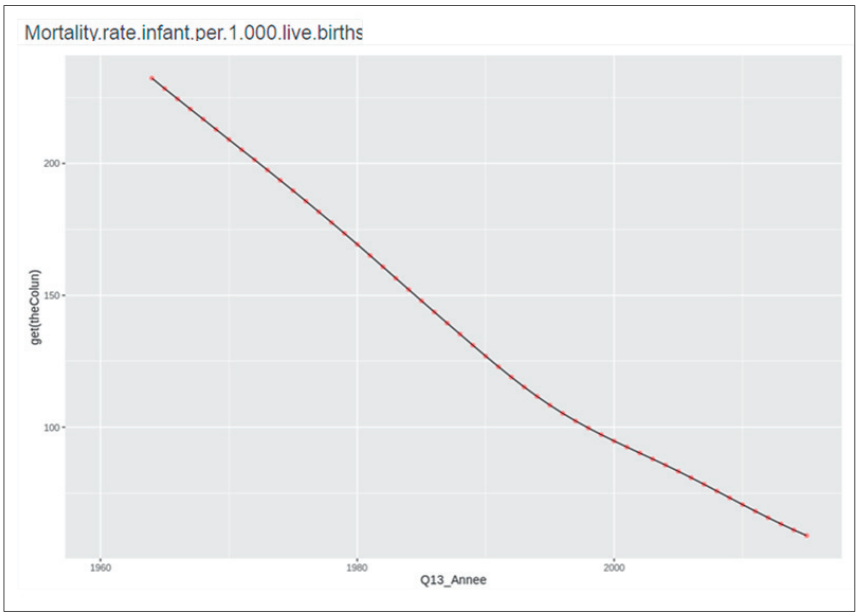


Figure 3.56 Exemple de l'Afghanistan ((taux de mortalité infantile pour 1000 naissances vivantes lissé)



Conclusion

En conclusion, le taux de mortalité infantile lissé est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale, du fait de sa tendance baissière généralisée depuis au moins 2 siècles (exception faite des crises sanitaires).

Notons également que la pente moyenne du taux de mortalité lissé est de -1.17 (médiane de -1.06), ce qui est assez proche en valeur absolue de 1 : ce qui renforce son rôle potentiel de temps transitionnel.

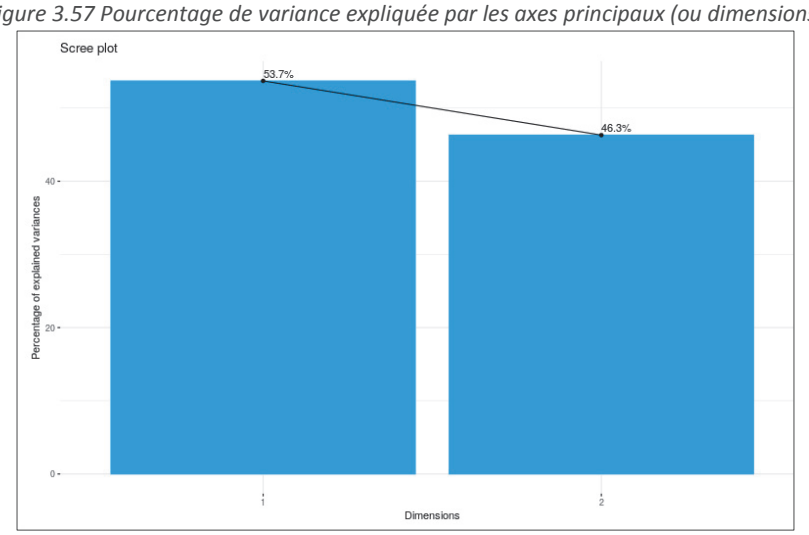
Précisons toutefois que le taux de mortalité est elle-même une statistique généralement lissée avant publication officielle, du moins dans période d'après-Guerre.

5.1.5. Classification des pays, sur composantes principales

Sélection des composantes principales pour la classification

L'analyse en composantes principales (ACP) est réalisée à partir de deux variables actives : Le taux de mortalité (*Death.rate.crude.per.1.000.people*) et le taux de natalité (*Birth.rate.crude.per.1.000.people*).

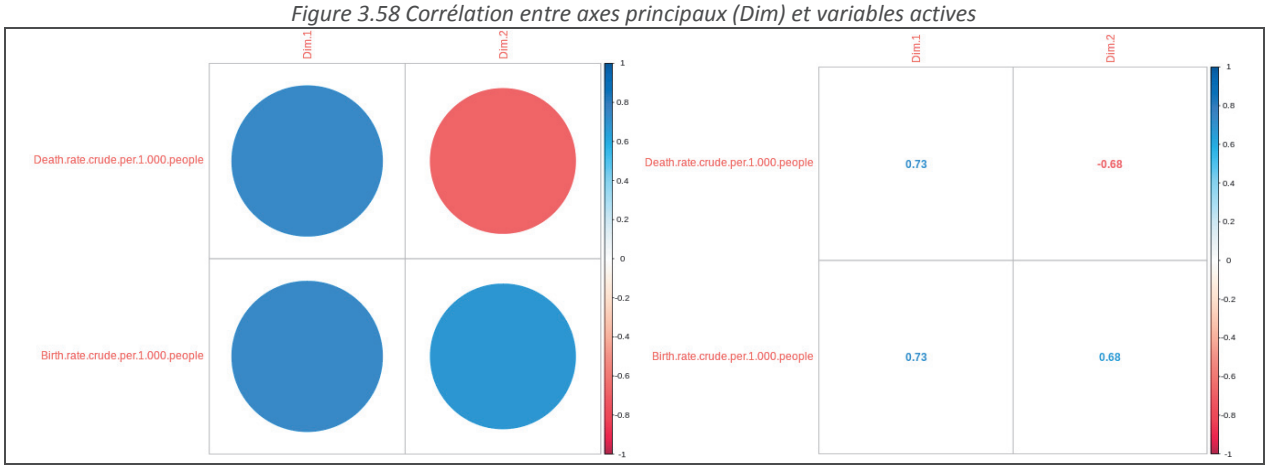
Figure 3.57 Pourcentage de variance expliquée par les axes principaux (ou dimensions)



Les 2 axes principaux expriment 100% de l'inertie totale du jeu de données ; cela signifie que 100% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan. Cela est normal puisque l'analyse ACP est réalisée avec seulement deux variables actives.

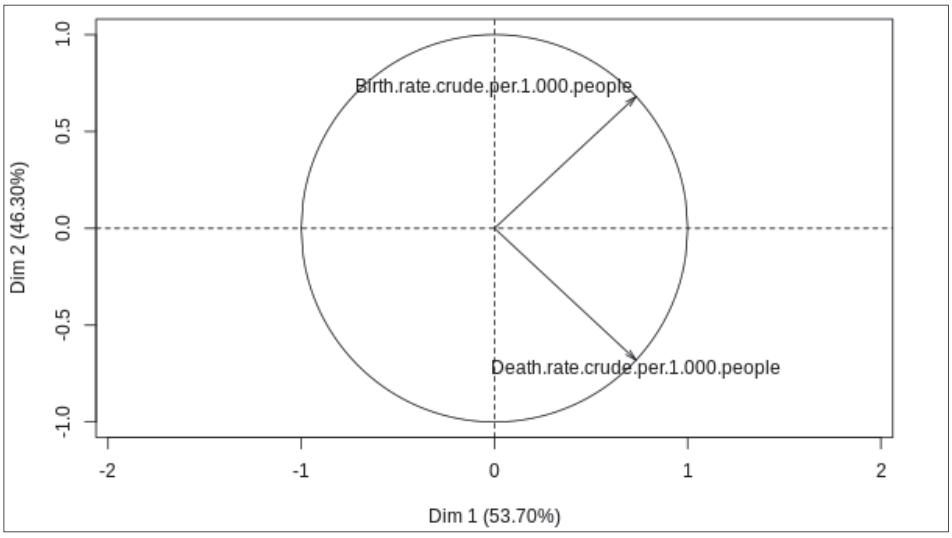
Signification des composantes principales sélectionnées

Figure 3.58 Corrélation entre axes principaux (Dim) et variables actives



Death.rate.crude.per.1.000.people	Taux de mortalité, brut (pour 1 000 personnes)
Birth.rate.crude.per.1.000.people	Taux de natalité, brut (pour 1 000 personnes)

Figure 3.59 Cercle de corrélation entre axes principaux (Dim) et variables actives



Le 1er axe factoriel (Dim 1) est positivement et identiquement corrélé avec les deux variables actives : il s'agit d'un effet taille associée simultanément au taux de mortalité et au taux de natalité.

Le 2ème axe factoriel (Dim 2) oppose le taux de moralité (corrélation négative), au taux de natalité (corrélation positive) : Ainsi les pays caractérisés par un taux de natalité supérieur au taux de mortalité se situeront dans la partie positive (haute) de Dim 2. Et inversement pour la partie négative (basse) de Dim 2.

Choix du nombre optimal de classes de pays

Après classification statistique des pays, nous obtenons quatre classes optimales de pays, qui seront décrites, après ordonnancement par le temps transitionnel qui est le taux de mortalité infantile lissé.

Figure 3.60 Gain d'inertie en fonction du nombre de classes

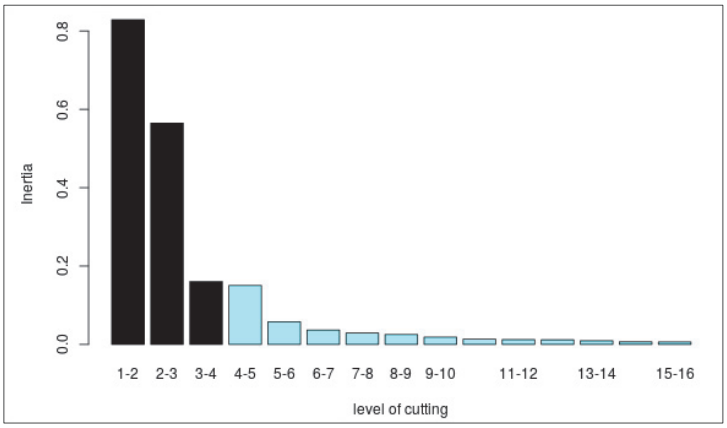
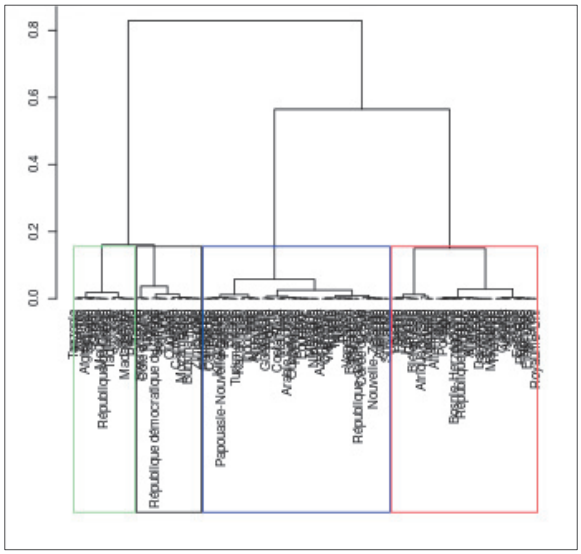


Figure 3.61 Dendrogramme de la classification ascendante hiérarchique



5.1.6. Indexation temporelle des classes de pays

Etant donné que le temps transitionnel (taux de mortalité infantile lissé) est décroissant en fonction du temps, les classes de pays sont numérotées par ordre décroissant de leur moyenne de la variable temps transitionnel.
Ainsi, la classe ayant la moyenne de proxy-temps la plus faible, porte l'index temporel le plus élevé (c'est la classe 4, qui est la plus « âgée » ou « ancienne » dans le processus de transition) ; et inversement pour la classe 1 qui est à l'autre bout du spectre.
Nous qualifions les classes ainsi indicées par le temps transitionnel, de *classes d'âge transitionnel*.

5.1.7. Description statistique des classes d'âge transitionnel de pays

Figure 3.62 Représentation en polygones des classes temporelles de pays sur les deux premiers axes factoriels

Cette représentation « temporelle » des classes de pays, suggère un parcours temporel sous forme d'un croissant en allant de la droite vers la gauche.
Dans cette vue « temporelle », la classe 4, représente les pays les plus anciens dans le processus de transition démographique. A l'opposé, les pays de la classe 1 sont les pays les plus récents dans le processus de transition démographique.
Ces 4 classes illustrent également les 4 phases observables « actuellement » dans le processus de transition démographique. La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transition démographique.

Figure 3.63 Représentation en ellipse des classes temporelles de pays sur les deux premiers axes

La représentation elliptique ci-dessus suggère que les frontières des classes ne sont pas si tranchées que le laisse supposer la représentation en polygones.
Nous marquons la position atypique de l'Afrique du sud qui se situe presque à mi-distance entre la chasse 1 majoritairement africaine et la classe 4 majoritairement européenne.

Tableau 3.4 Description des classes selon les variables actives

Classes	Birth rate, crude (per 1,000 people)	Death rate, crude (per 1,000 people)
1	19,7	6,0
2	11,1	10,6
3	34,7	*NS*
4	41,7	11,5
Monde	22,7	8,3

NS : valeur non-significativement différente de la moyenne mondiale au seuil de 5%

La **classe 1** est composée de pays tels que *Burundi*, *Tchad*, *Côte d'Ivoire*, *Mali*, *Niger*, *Nigeria*, *Sierra Leone* et *Somalie*. Ce groupe est caractérisé par de fortes valeurs pour le taux de natalité (41.7‰ dans la classe contre 22.7‰ au niveau mondial) et le taux de mortalité (11.5‰ dans la classe contre 8.3‰ au niveau mondial).

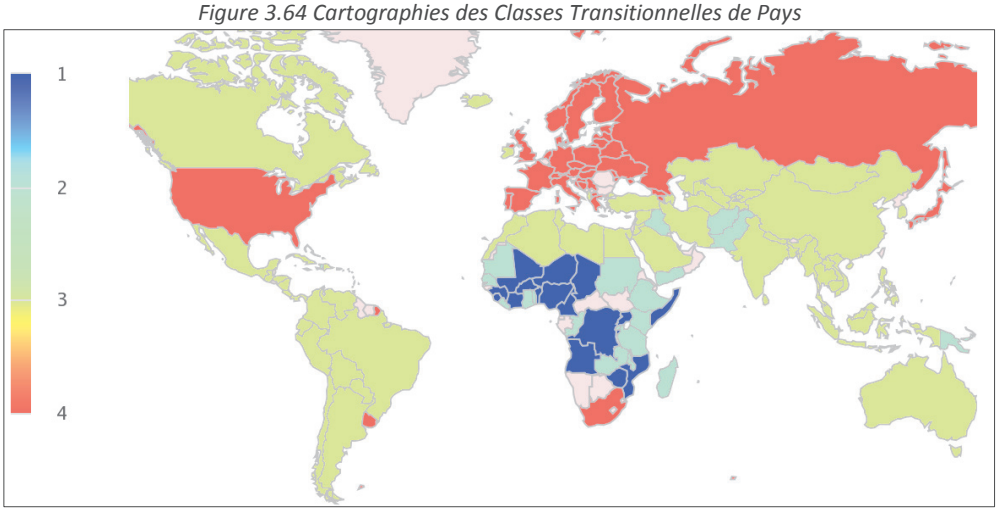
La **classe 2** est composée de pays (comme *Afghanistan, Ethiopie, Ghana, Madagascar, Rwanda, Sénégal*) partageant de forts taux de natalité (en moyenne 34.7‰ dans la classe contre 22.7‰ au niveau mondial)

La **classe 3** est composée de pays (tels que *Azerbaïdjan, Pérou, Bangladesh, Turquie, Vietnam, Tunisie, Népal*) ayant de faibles taux de mortalité (moyenne de 6‰ dans la classe contre 8.3‰ au niveau mondial) et de natalité (moyenne de 19.7‰ dans la classe contre 22.7‰ au niveau mondial).

La **classe 4** est constituée de pays tels que *: Allemagne, Bosnie, République Tchèque, Japon, Portugal, Italie, Pologne, Biélorussie*, *Estonie, Lettonie*, *Lituanie*, *Serbie* et *Ukraine*. Ce groupe est caractérisé par :

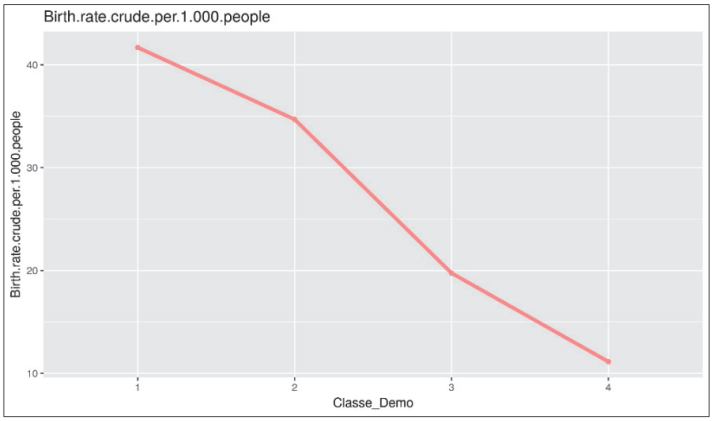
- de forts taux de mortalité (moyenne de 10.6‰ dans la classe contre 8.3‰ au niveau mondial)
- de faibles taux de natalité (moyenne de 11.1‰ dans la classe contre 22.7‰ au niveau mondial)

Figure 3.64 Cartographies des Classes Transitionnelles de Pays

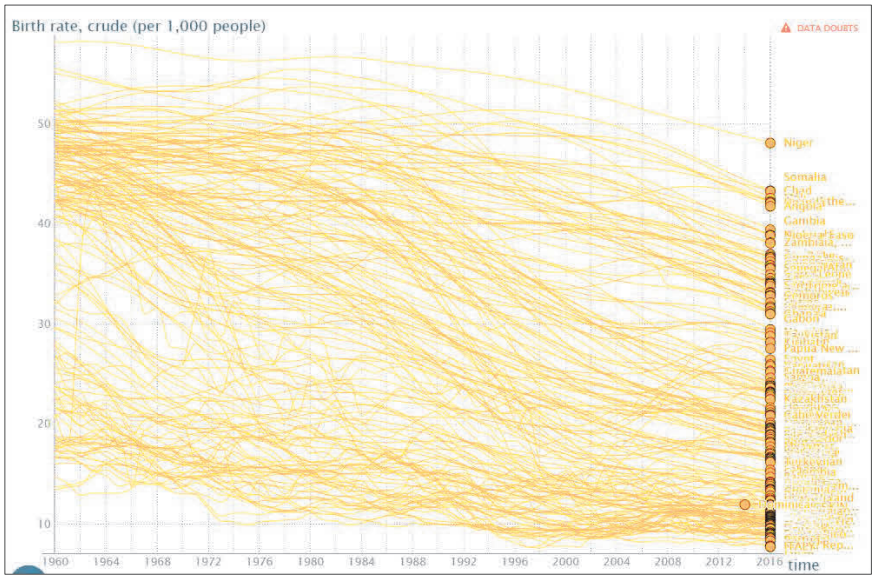


5.1.8. Variables actives : courbes transitionnelles transversales selon le temps transitionnel, avec la moyenne de la variable par classe d'âge transitionnel

Figure 3.65 Taux de natalité, brut (pour 1 000 personnes)

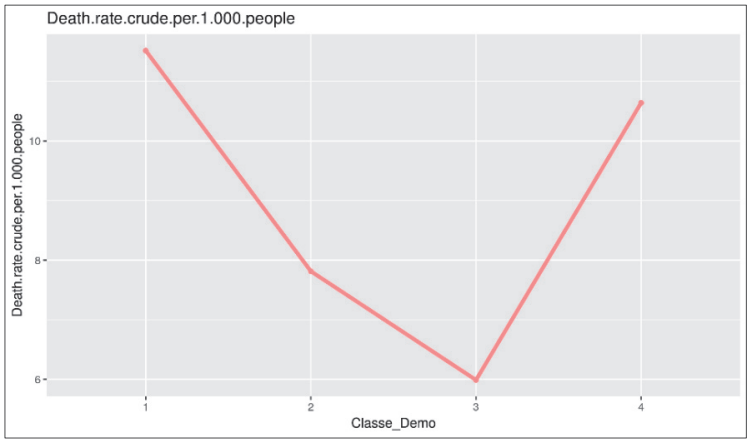


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures baissières similaires.

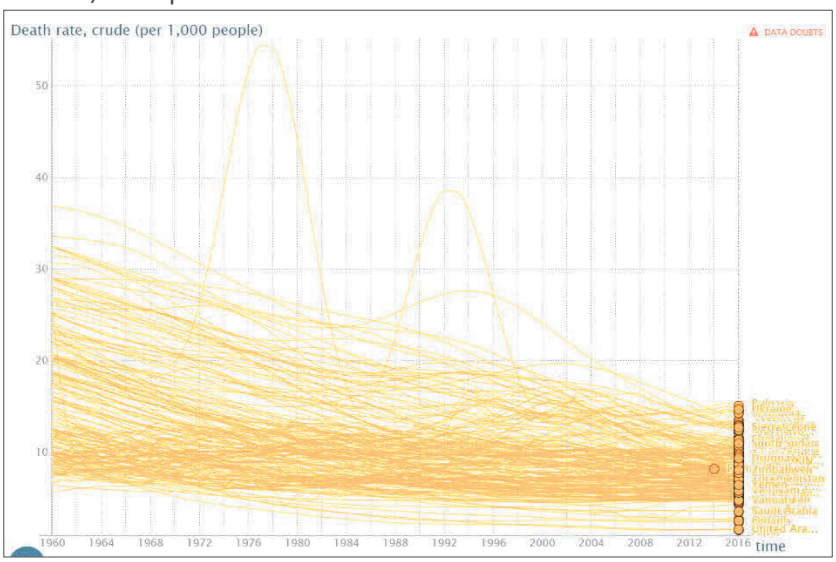


Source : World Bank (2018)

Figure 3.66 Taux de mortalité, brut (pour 1 000 personnes)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures en cloche similaires.



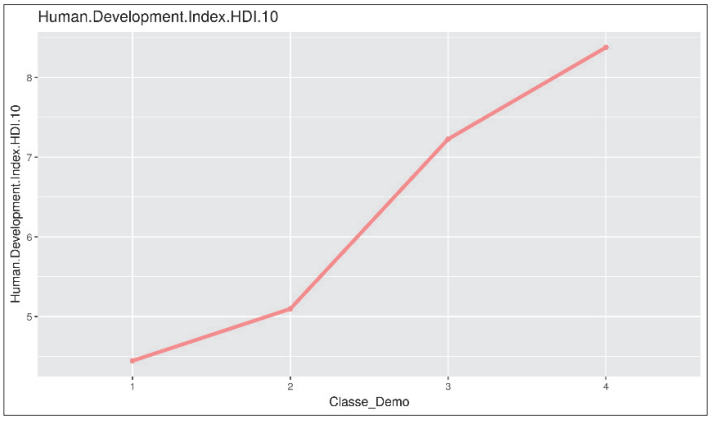
Source : World Bank (2018)

5.1.9. Variables passives : courbes transitionnelles transversales selon le temps transitionnel, avec la moyenne de la variable par classe d'âge transitionnel

Les variables passives sont des variables utilisées pour illustrer les classes, et qui ne sont pas des variables actives (qui elles sont utilisées pour créer les classes).

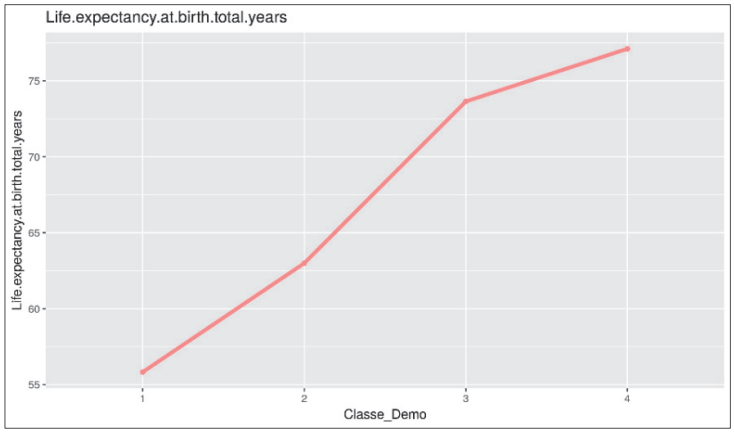
Moyennes par classes d'âge transitionnel : schéma transitionnel empirique selon le temps transitionnel

*Figure 3.67 Indice de développement humain (HDI*10)*



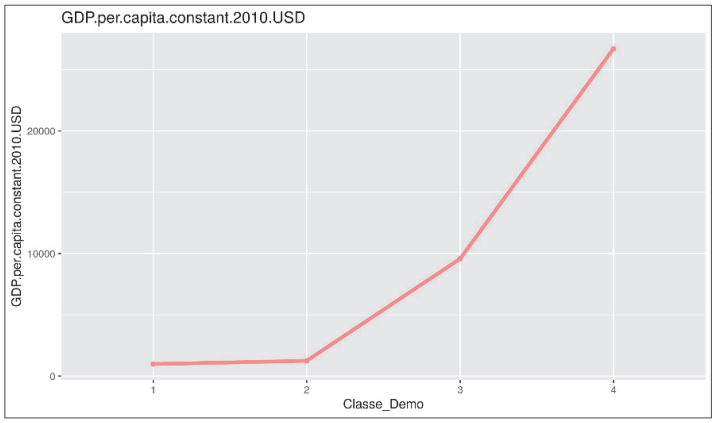
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.1*) nous permet de conclure à des allures haussières similaires.

Figure 3.68 Espérance de vie à la naissance



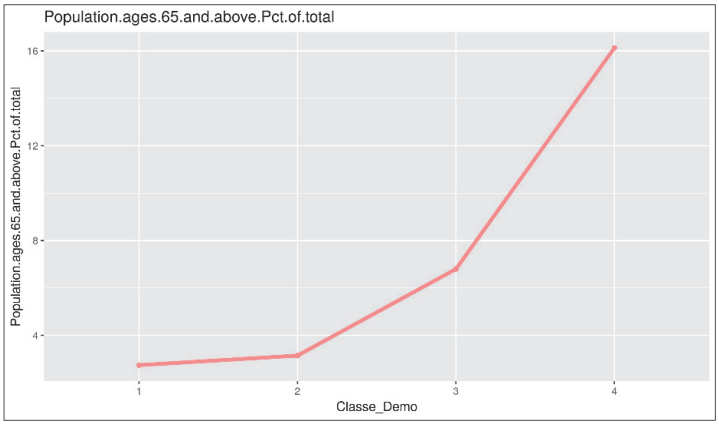
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.2*) nous permet de conclure à des allures haussières similaires.

Figure 3.69 PIB par habitant (USD constants de 2010)



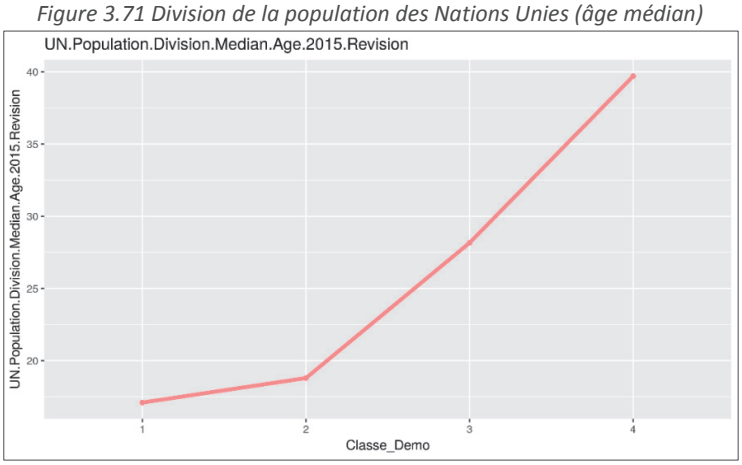
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.3*) nous permet de conclure à des allures haussières similaires.

Figure 3.70 Population âgée de 65 ans et plus (% population totale)



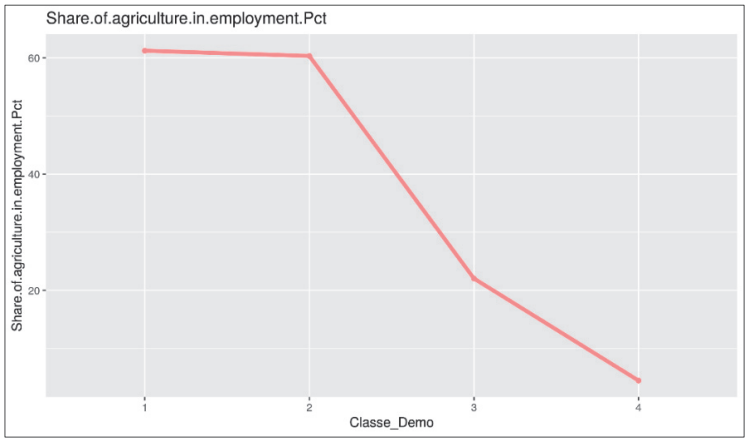
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.4*) nous permet de conclure à des allures haussières similaires.

Figure 3.71 Division de la population des Nations Unies (âge médian)



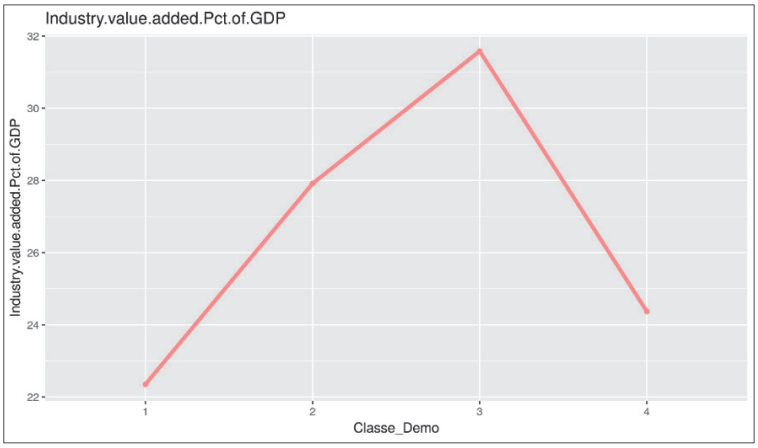
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.5*) nous permet de conclure à des allures haussières similaires.

Figure 3.72 La part de l'agriculture dans l'emploi (% de l'emploi total)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.6*) nous permet de conclure à des allures baissières similaires.

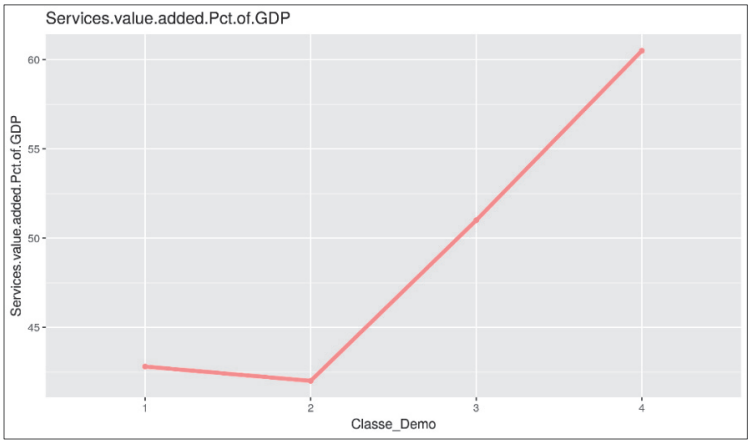
Figure 3.73 Industrie, valeur ajoutée



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (

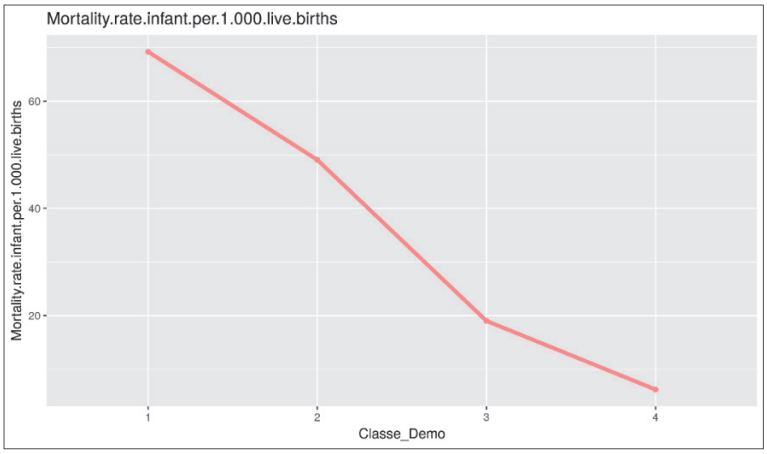
Figure 0.7) nous permet de conclure à des allures en cloche similaires.

Figure 3.74 Services, valeur ajoutée (% du PIB



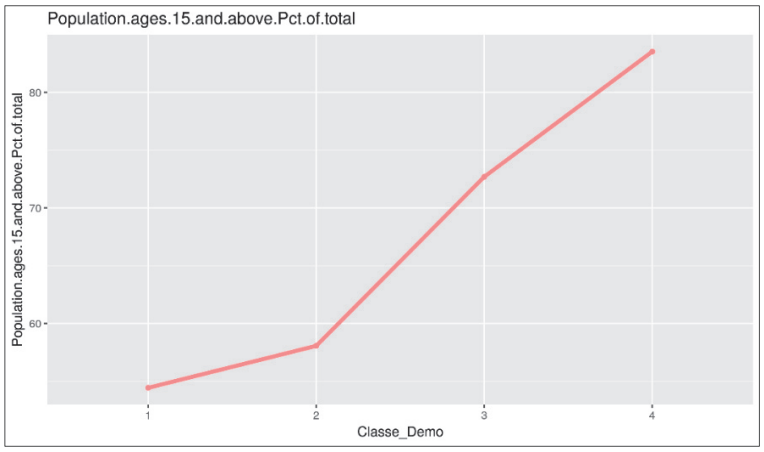
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.8*) nous permet de conclure à des allures haussières similaires.

Figure 3.75 Taux de mortalité infantile (pour 1 000 naissances vivantes)



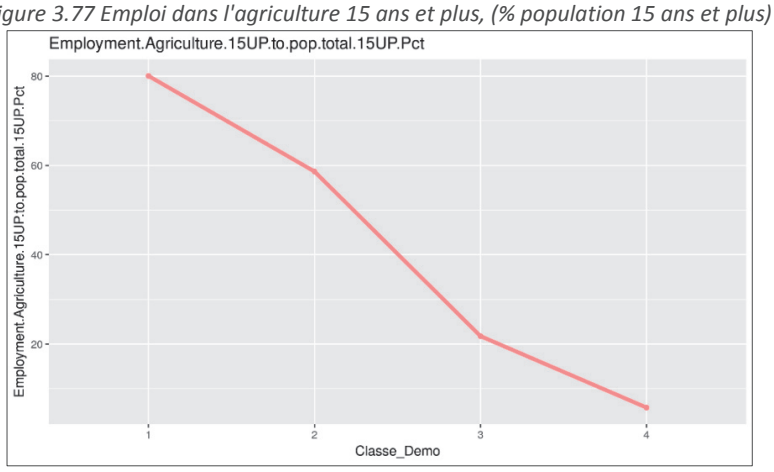
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.15)* nous permet de conclure à des allures baissières similaires.

Figure 3.76 Population de 15 ans et plus (% du total)



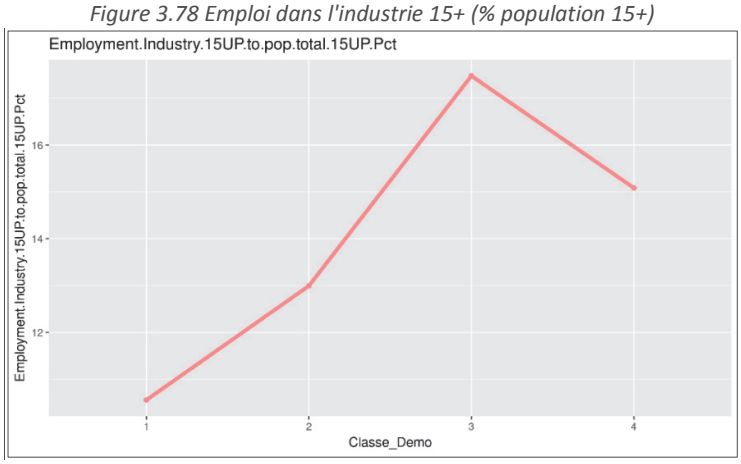
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.18*) nous permet de conclure à des allures haussières similaires.

Figure 3.77 Emploi dans l'agriculture 15 ans et plus, (% population 15 ans et plus)



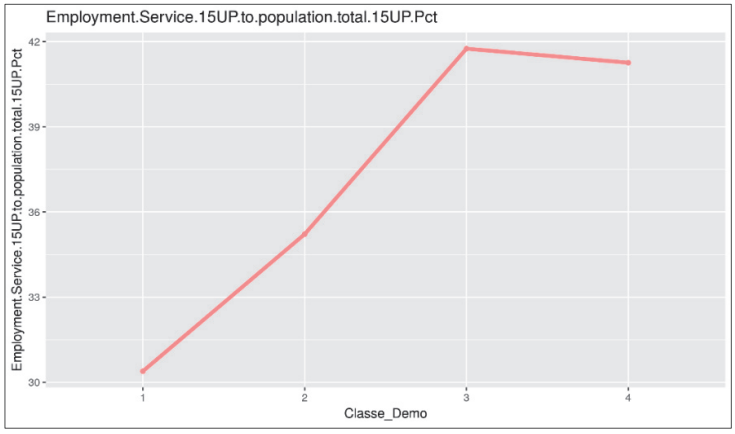
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.19*) nous permet de conclure à des allures baissières similaires.

Figure 3.78 Emploi dans l'industrie 15+ (% population 15+)



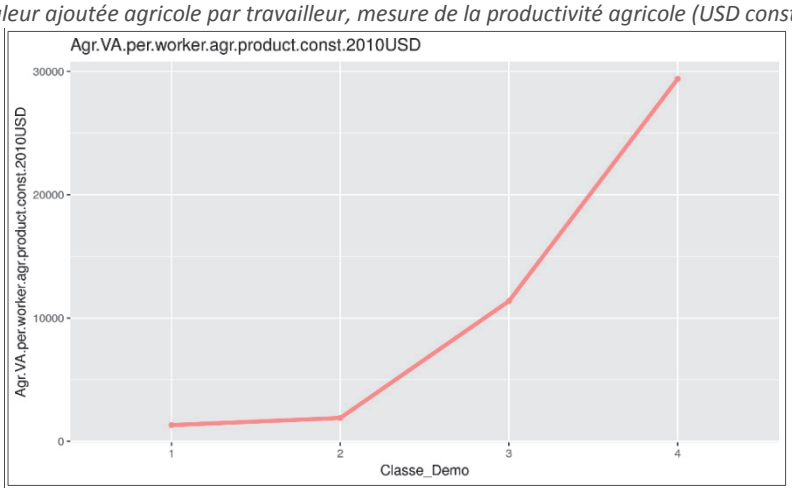
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.20*) nous permet de conclure à des allures en cloche similaires.

Figure 3.79 Emploi dans les services 15+ (% population 15+)



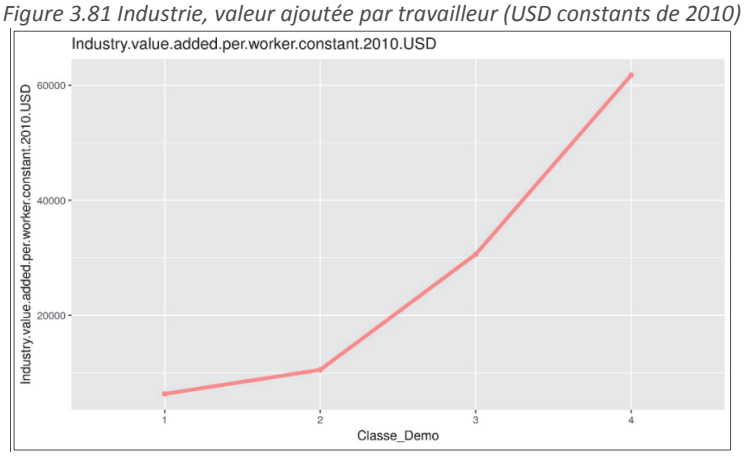
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.21*) nous permet de conclure à des allures haussières similaires.

Figure 3.80 Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010)



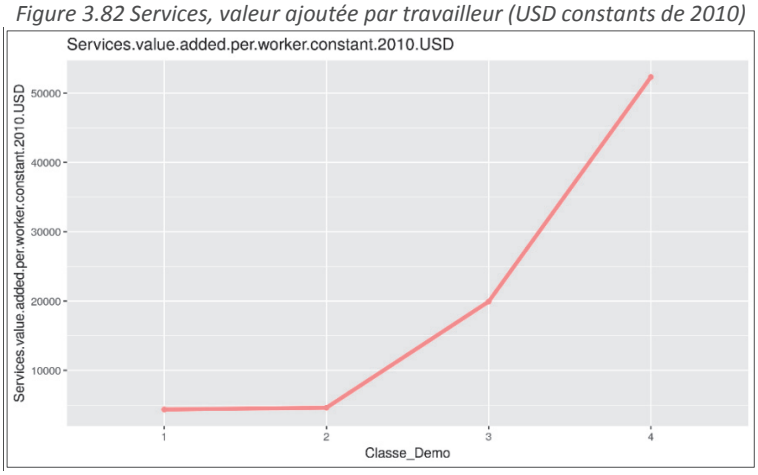
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.22*) nous permet de conclure à des allures haussières similaires.

Figure 3.81 Industrie, valeur ajoutée par travailleur (USD constants de 2010)



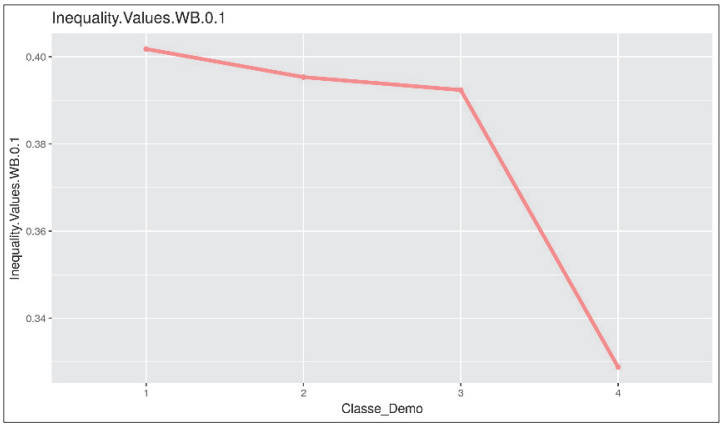
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.23*) nous permet de conclure à des allures haussières similaires.

Figure 3.82 Services, valeur ajoutée par travailleur (USD constants de 2010)



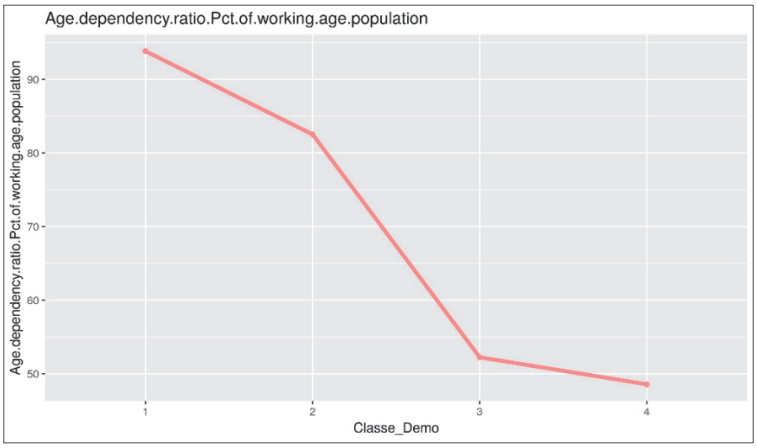
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.24*) nous permet de conclure à des allures haussières similaires.

Figure 3.83 Valeurs des inégalités (estimation de la Banque mondiale)



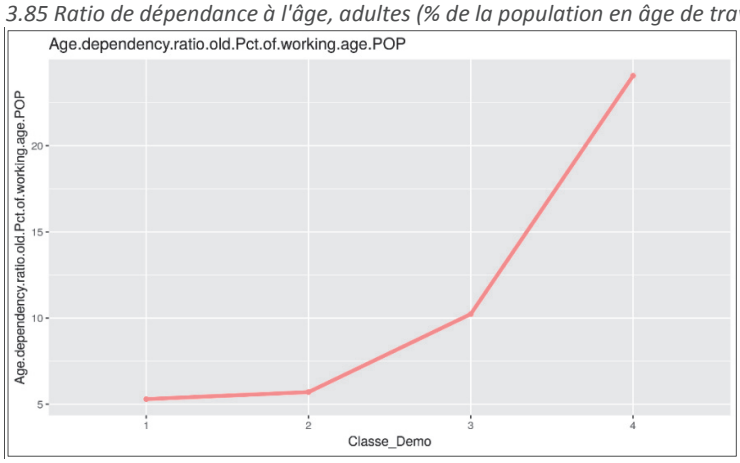
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.9)* nous permet de conclure à des allures baissières similaires.

Figure 3.84 Ratio de dépendance à l'âge (% de la population en âge de travailler)



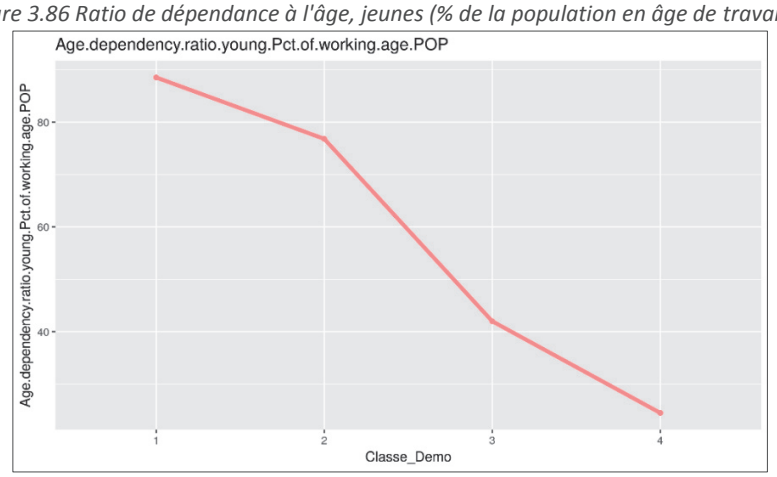
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.10*) nous permet de conclure à des allures baissières similaires.

Figure 3.85 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler)



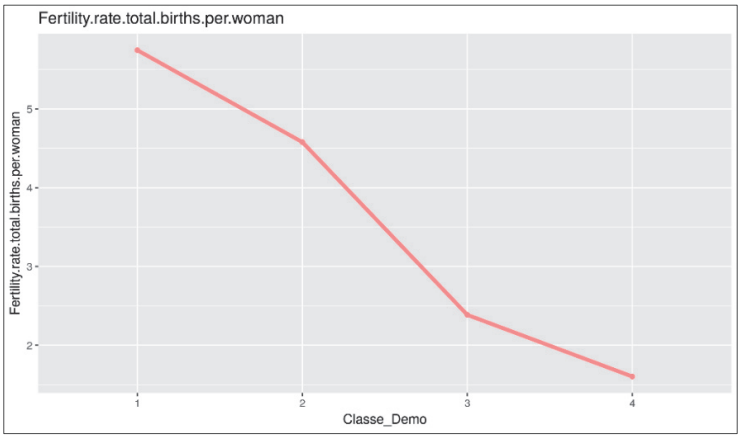
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.11*) nous permet de conclure à des allures haussières similaires.

Figure 3.86 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler)



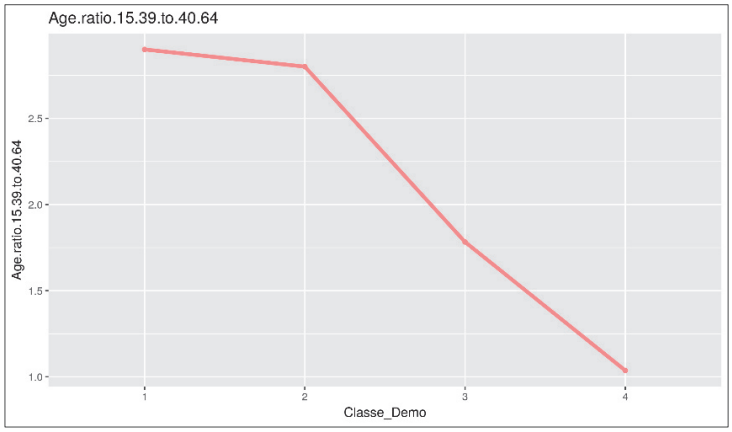
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.12*) nous permet de conclure à des allures baissières similaires.

Figure 3.87 Taux de fécondité, total (naissances par femme)



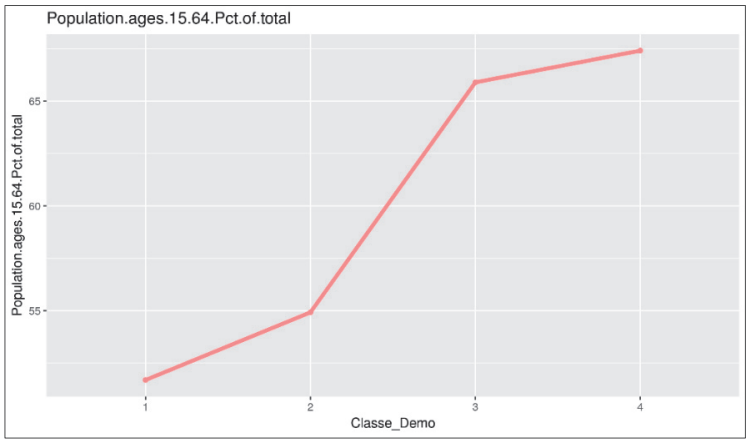
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.16*) nous permet de conclure à des allures baissières similaires.

Figure 3.88 Ratio d'âge 15-39 à 40-64



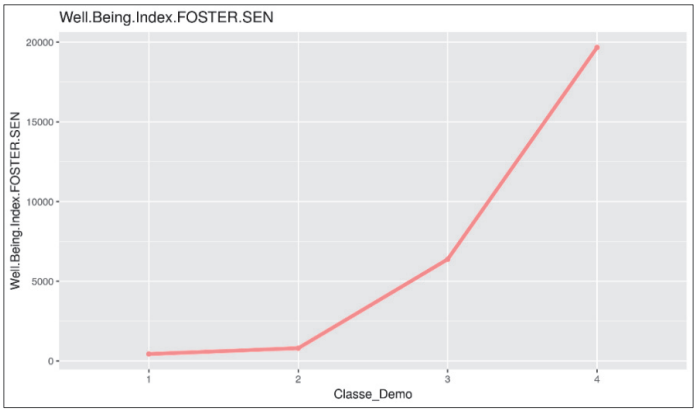
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.13*) nous permet de conclure à des allures baissières similaires.

Figure 3.89 Population âgée de 15 à 64 ans (% du total)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.14*) nous permet de conclure à des allures haussières similaires.

Figure 3.90 Indice de bien-être (FOSTER-SEN)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.17*) nous permet de conclure à des allures haussières similaires.

5.2. Comparaison du schéma transitionnel par coupe transversale avec le schéma théorique longitudinal de la transition démographique

Variables de transition démographique : schéma transitionnel transversal selon le temps transitionnel

Figure 3.91 Schéma transitionnel transversal selon le temps transitionnel. Moyennes par classes d'âge transitionnel

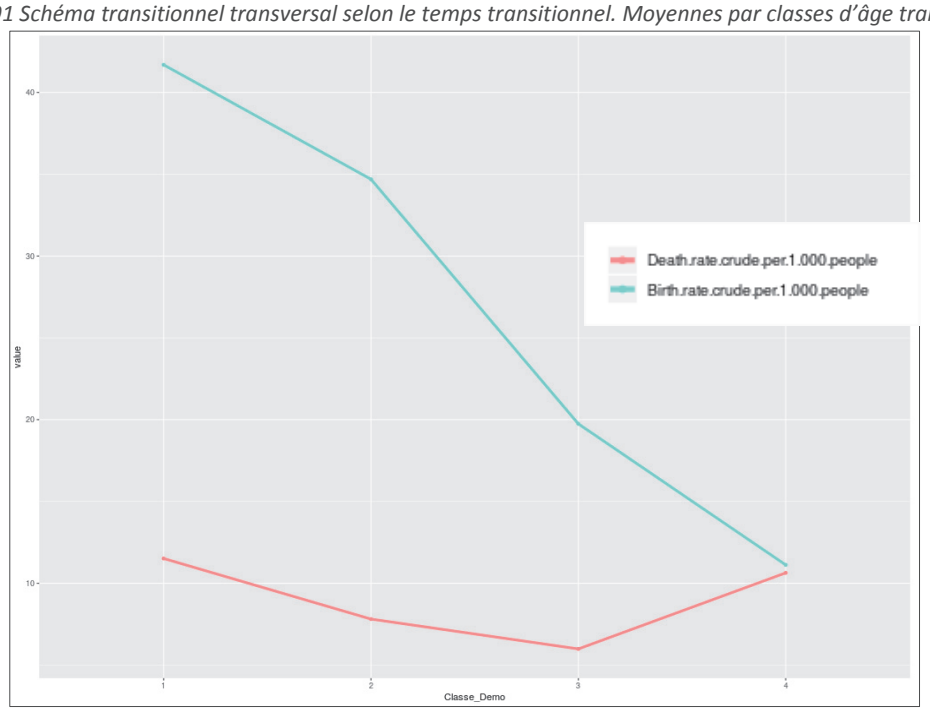


Figure 3.92 Schéma transitionnel transversal selon le temps transitionnel. Médianes par classes d'âge transitionnel

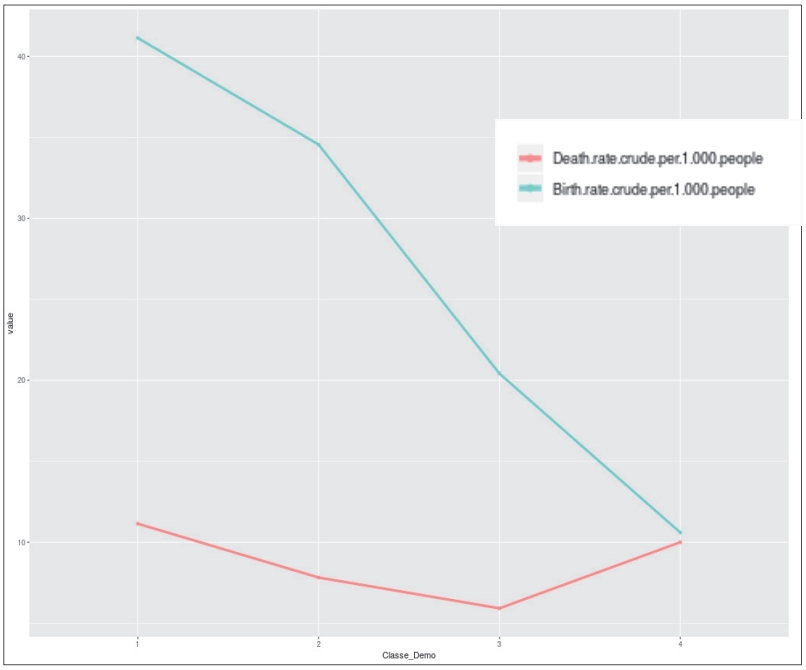
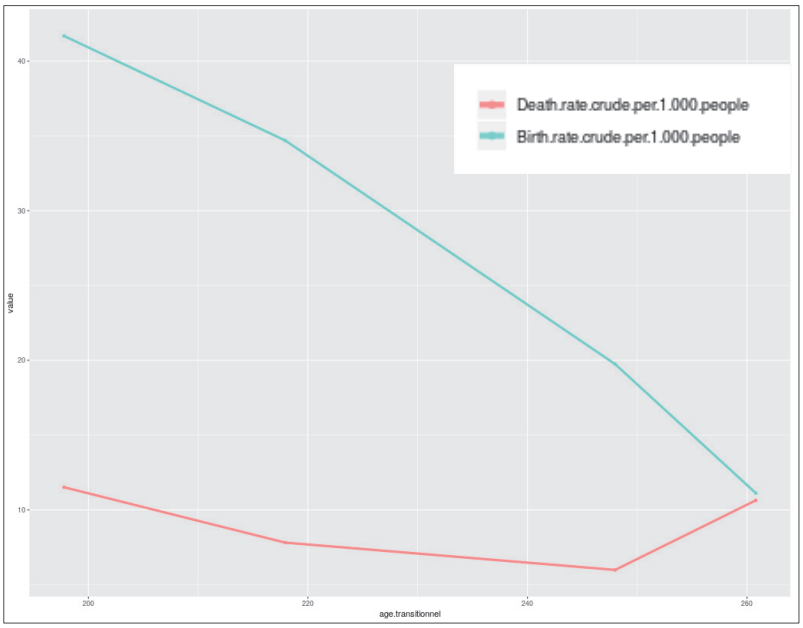


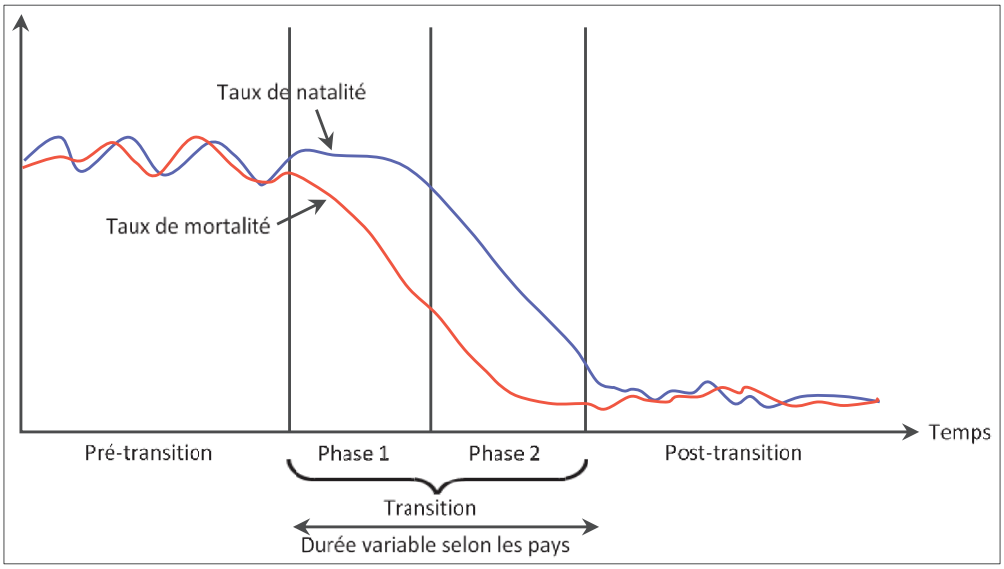
Figure 3.93 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse



L'allure des deux schémas ci-dessus est sensiblement la même, aussi bien en utilisant les moyennes que les médianes : C'est l'illustration d'une robustesse du schéma transitionnel démographique obtenu en coupe transversale par le temps transitionnel.

Schéma démographique transitionnel longitudinal théorique

Figure 3.94 Schéma démographique transitionnel longitudinal théorique

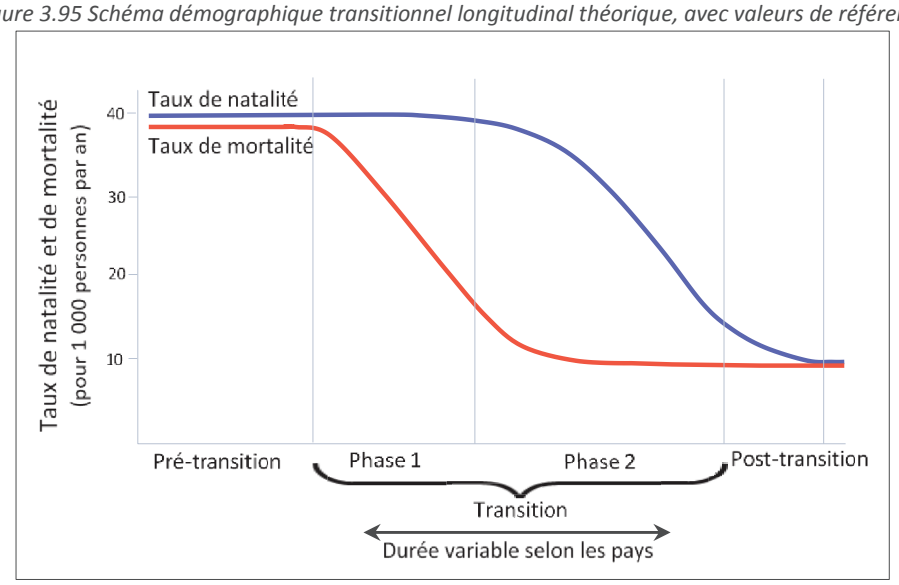


Source : Adapté de Durand et al. (2008).

En comparant le schéma transitionnel transversal selon le temps transitionnel (*Figure 3.91*), avec le schéma transitionnel longitudinal théorique (*Figure 3.94*) nous observons la même dynamique des deux variables de transition. Plus précisément, nous constatons une inclusion du schéma transversal contemporain, dans le schéma théorique complet, car le premier schéma s'identifie avec la phase 2 et le début de la post-transition du second schéma.
Ainsi, actuellement nous pouvons avancer qu'il n'y a plus de pays en période pré-transitoire, et seulement quelques rares pays en phase 1 de la transition démographique.

Une figure théorique de Roser et al. (2019) permet d'avoir des valeurs de référence théoriques des variables transitionnelles :

Figure 3.95 Schéma démographique transitionnel longitudinal théorique, avec valeurs de référence



Source : Adapté de Roser et al. (2019)

En faisant le parallèle avec les phases de la transition démographique (Durand *et al.*, 2008) et leurs valeurs de référence (Roser *et al.*, 2019), nous pouvons positionner les classes d'âge transitionnel contemporaines par rapport à ces phases théoriques :

Tableau 3.5 Correspondance entre Classes d'âge transitionnel et Phases théoriques de la transition démographique

		Moyenne	**Moyenne**	**Moyenne**	**Phase théorique**	**Phase théorique**	**Phase théorique**	**Phase théorique**
Classe d'âge transitionnel	**% de pays en 2014**	**Taux de Natalité**	**Taux de mortalité**	**Solde naturel**	**Natalité**	**Mortalité**	**Solde naturel**	**Synthèse**
1	13%	417%	115%	302%	Phase 1	Après Début Phase 2	Fin Phase 1 / Début Phase 2	Fin Phase 1 / Début Phase 2
2	16%	347%	78%	269%	Début Phase 2	Après Début Phase 2	Phase 2	Début Phase 2
3	43%	197%	60%	137%	Milieu Phase 2	Après Début Phase 2	Milieu Phase 2	Milieu Phase 2
4	28%	111%	106%	5%	Post-transition	Après Début Phase 2	Post-transition	Post-transition

L'étude réalisée confirme assez clairement la validité et la réalité actuelle du processus théorique de transition démographique présenté par Durand et al. (2008).
Les phases transitionnelles démographiques offrent une grille de lecture dynamique et pleine de sens sur la situation structurelle passée, présente et future proche des pays de la planète.

6. Anthropologie culturelle

Nous allons ici utiliser le temps et l'âge transitionnel pour réaliser une inférence des trajectoires longitudinales dans une coupe transversale dans le domaine d'anthropologie culturel.

6.1. Analyse transversale anthropologique

6.1.1. Sélection des variables à étudier

Les variables structurelles utilisées se répartissent en deux catégories : les variables de transition et les variables actives.
Variables de transition : Elles sont utilisées pour décrire le processus de changement anthropologique.
Variables actives : Elles sont utilisées pour effectuer la classification statistique des pays. Elles incluent les variables de transition auxquelles sont ajoutées des variables d'homogénéisation des classes. Dans le cas présent, il n'y a pas de variables d'homogénéisation : Les variables actives sont les mêmes que les variables de transition.

Tableau 3.6 Variables actives ou de transition

Variable	Libellé de la variable	Signification en anglais	Signification en français	Variable active	Variable de transition
Traditional.Secular.rational.values	Traditional / Secular-rational values	Secular-rational values (vs Traditional values)	Valeurs laïques-rationnelles (vs valeurs traditionnelles)	oui	oui
Survival.Self.expression.values	Survival / Self-expression values	Self-expression values (vs Survival values)	Valeurs d'expression personnelle (vs valeurs de survie)	oui	oui
Inequality values.World.Bank.0.1	Inequality values World Bank 0.1	Inequality values	Valeurs inégalitaires	oui	oui

6.1.2. Transformation préalable des variables à étudier

Les variables actives utilisées font l'objet des transformations suivantes, avant la sélection de l'année d'étude : estimation des valeurs manquantes par interpolation spline, puis moyenne mobile d'ordre 5

6.1.3. Sélection des pays à étudier

En plus des pays exclus pour l'ensemble des domaines (cf. ***Error! Reference source not found.***), nous excluons également les pays instables (conflits internes) suivants : Irak, Libye, Syrie, Yémen, Colombie, Mexique.

6.1.4. Sélection de l'année à étudier

Pour l'analyse transversale, l'année étudiée est 2014. Cela pour des raisons de complétude et de fiabilité des données, car les dernières (2017, source Banque Mondiale) sont souvent complétées et révisées après publication.

6.1.5. Choix du temps transitionnel

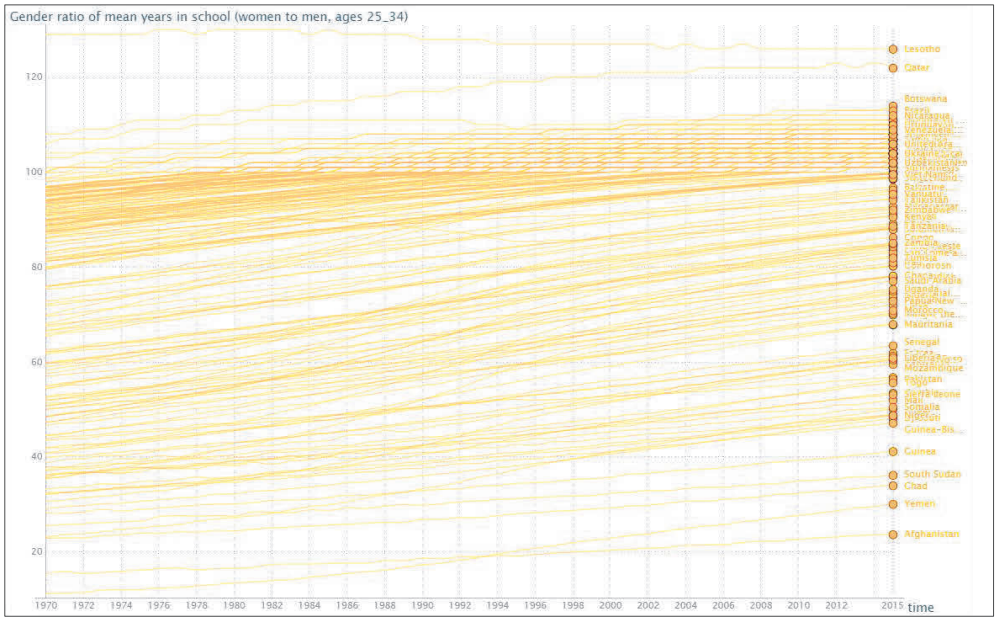
Variable à tester comme temps transitionnel

La variable à tester est une valeur transformée (moyenne mobile arrière d'ordre 5) du Ratio femmes vs hommes, du nombre moyen d'années d'études, de la population âgée de 25 à 34 ans *[Gender ratio of mean years in school (women to men, ages 25-34)]*.

Ce ratio, bien que basé sur des mesures éducatives, peut être légitiment utilisé dans le domaine anthropologique, car il illustre les choix scolaires familiaux et sociaux par rapport au sexe des enfants. Or les préférences familiales ou sociales par rapport au sexe relèvent du domaine anthropologique, notamment le *statut de la femme* (cf. page 50).

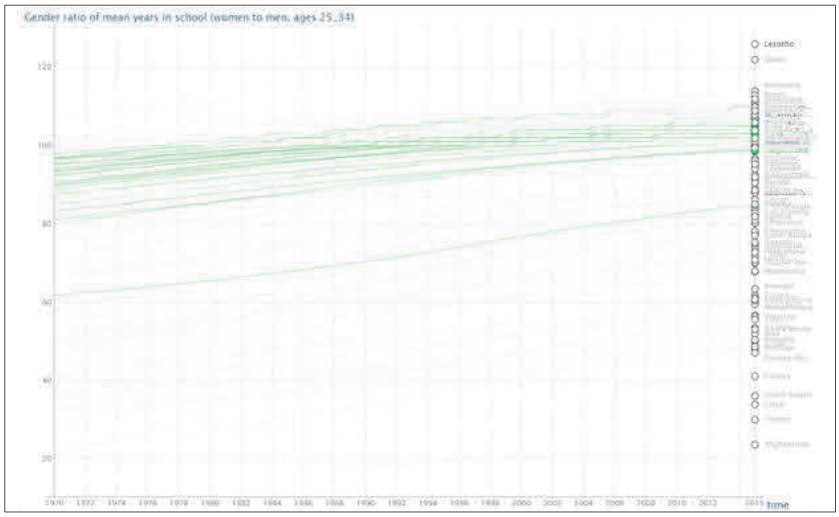
Une analyse faite visuellement (cf. *Figure 3.96* à *Figure 3.99* ci-dessous) de la tendance mondiale du ratio d'années d'études femmes vs hommes de 25-34 ans, permet d'avancer l'hypothèse haussière universelle, même si on ne peut parler de monotonie mathématique stricte. C'est la raison pour laquelle nous utilisons un lissage par moyenne mobile pour améliorer la monotonie de la variable pressentie pour jouer le rôle de temps transitionnel.

Figure 3.96 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (non-lissée), ensemble des pays du monde



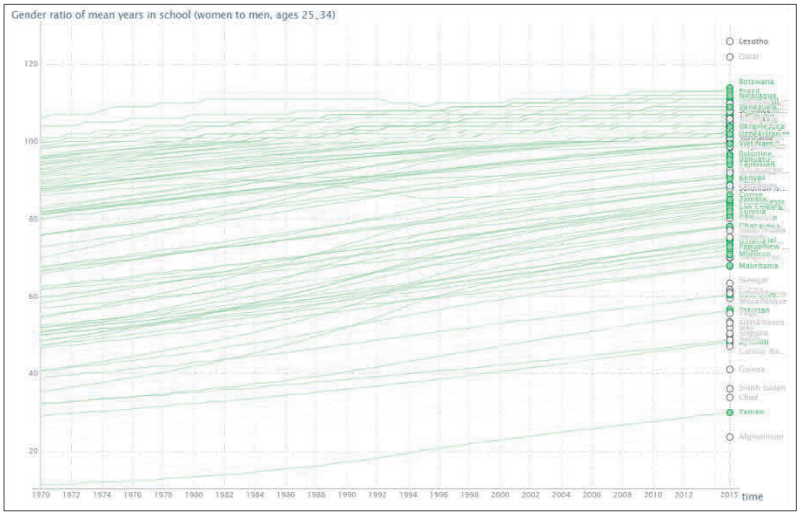
Source : World Bank (2020)

Figure 3.97 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans des pays de l'OCDE



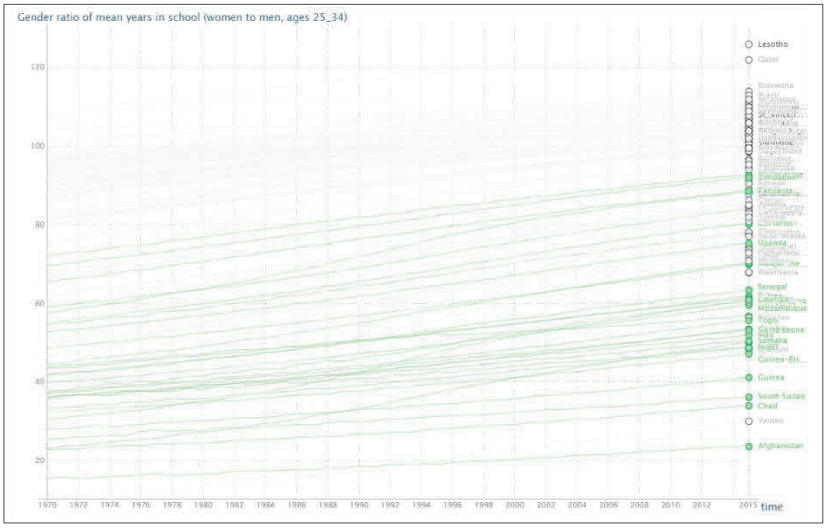
Source : World Bank *(2020)*

Figure 3.98 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (non-lissée) des pays à moyen revenu par habitant (au sens de la Banque Mondiale)



Source : World Bank (2020)

Figure 3.99 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (non lissée) des pays à faible revenu par habitant (au sens de la Banque Mondiale)



Source : World Bank (2020)

Vérification de la monotonie empirique de la variable à tester

La variable à tester Y est empiriquement monotone lorsque la série différenciée dY (dY(t)=Y(t) – Y(t-1)) est de signe constant sur la période étudiée, pour tout pays.
De plus pour assurer l'universalité de cette monotonie empirique, le signe de série différenciée doit être le même pour tous les pays étudiés
Pour chaque pays, nous calculons le pourcentage de valeurs positives de la série différenciée, afin de vérifier l'unicité du signe sur la période étudiée.
Sur l'ensemble des 124 pays étudiés, le pourcentage moyen de valeurs positives de la série différenciée est de 99.4%, avec un minimum à 53% (Algérie), une médiane à 100%, et un maximum à 100%.
Mieux, 119 des 124 pays, soit 96%, ont 100% de valeurs positives dans la série différenciée.

Par ailleurs, les valeurs centrales (moyenne=0.98 et médiane=1) du coefficient de Kendall indiquent une très forte relation monotone croissante entre le ratio d'années d'études femmes vs hommes de 25-34 ans lissé et le temps.
Sur les 124 pays, le coefficient de Kendall est systématiquement négatif (99% des pays), sauf pour 1 pays.

Vérification empirique de l'unicité du signe de la pente de la variable testée, quel que soit le pays

A défaut d'une monotonie empirique parfaite, l'unicité pour tous les pays, du signe de la pente de la droite d'ajustement de la variable à tester, montre une universalité de la tendance (à la hausse ou à la baisse) sur l'ensemble des pays étudiés.
De plus, si pour un pays, pente de la droite d'ajustement de la variable à tester est proche en valeur absolue de 1, alors cette variable à tester se rapproche d'autant plus de la variable temps.

Nous régressions linéairement la variable à tester par rapport au temps. Le signe du coefficient du temps est positif, lorsque la variable à tester a une tendance linéaire croissante dans le temps.
Ci-dessous les statistiques sommaires issues des 124 régressions (124 pays).

Summary(coef.Intercept)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-0.3723 -0.3715 -0.3706 -0.3643 -0.3693  0.2196       3
```

Summary(coef.year)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-0.7163  1.2047  1.2090  1.1884  1.2119  1.2145       3
```

Summary(coef.R2)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 0.3476  0.9830  0.9900  0.9771  0.9948  0.9991       3
```

Pour 123 régressions (sur 124), le coefficient du temps (coef.year), est systématiquement positif, confirmant ainsi l'hypothèse d'une hausse tendancielle du ratio d'années d'études femmes vs hommes de 25-34 ans (lissé), pour l'ensemble des pays.

Figure 3.100 Exemple de la France d'après-guerre (Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans lissé)

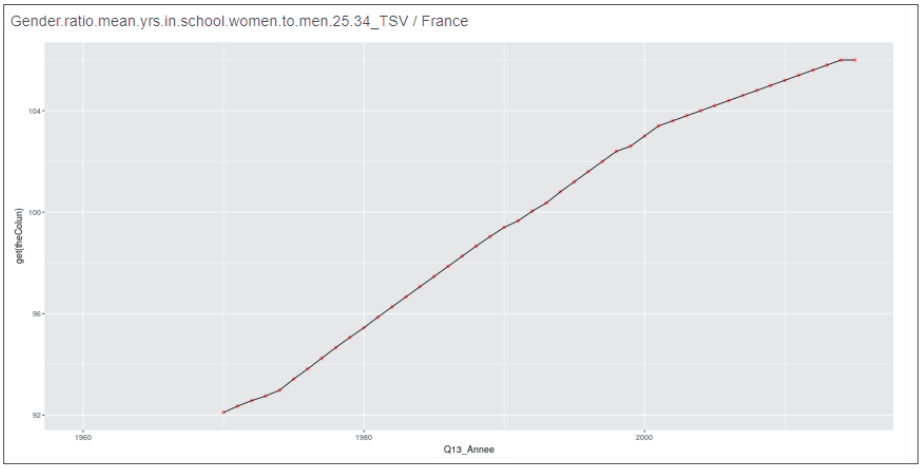
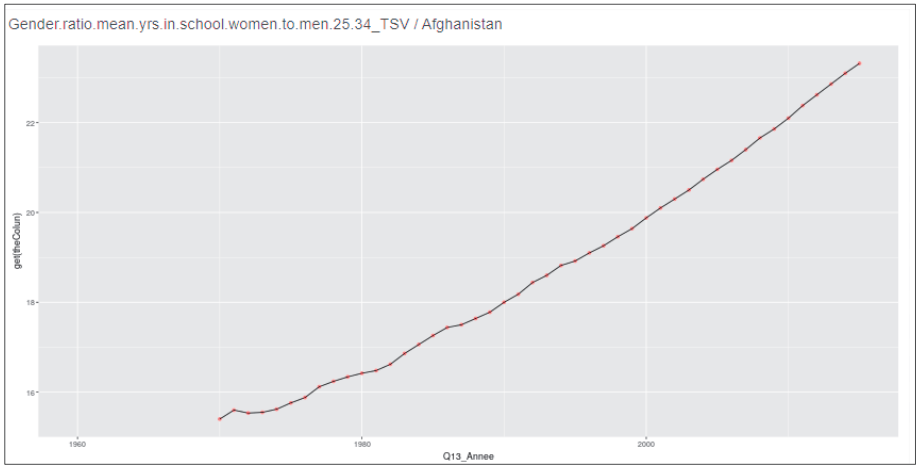
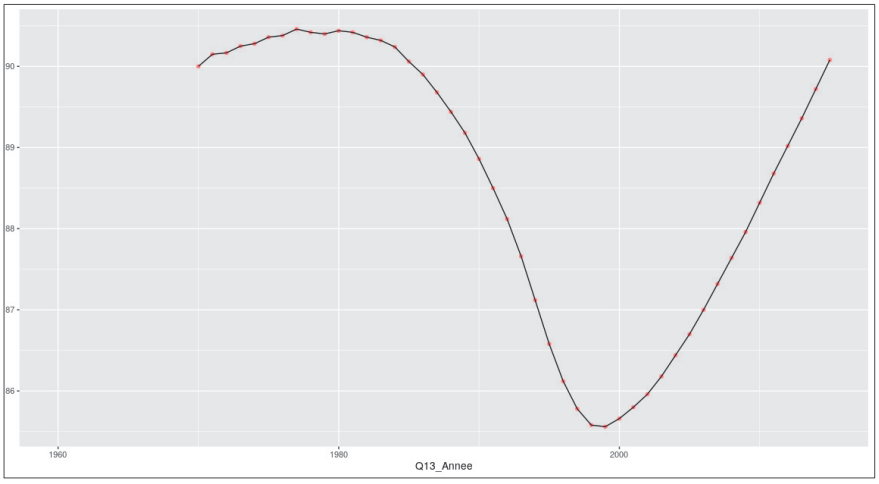


Figure 3.101 Exemple de l'Afghanistan (Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans lissé)



L'exception est l'Algérie qui a une pente historique négative, et même dans ce cas la pente est nettement positive, depuis les années 2000, comme le montre le graphique ci-après :

Figure 3.102 Algérie : Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (lissé)



Conclusion

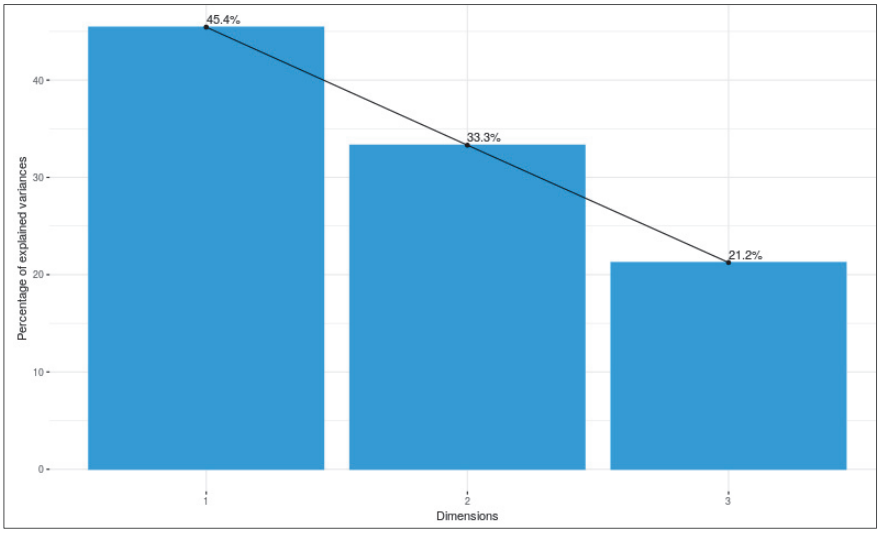
En conclusion, le ratio d'années d'études femmes vs hommes de 25-34 ans (lissé) est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale, du fait de sa tendance haussière généralisée, au moins depuis les années 1970.
Notons également que la pente moyenne du ratio d'années d'études femmes vs hommes de 25-34 ans (lissé) est de +1.18 (médiane de +1.2), ce qui est assez proche en valeur absolue de 1 : cela renforce le rôle de cette variable comme potentiel temps transitionnel.

6.1.6. Classification des pays, sur composantes principales

Sélection des composantes principales pour la classification

L'analyse en composantes principales (ACP) est réalisée à partir des 3 variables actives (cf. Tableau 3.6).

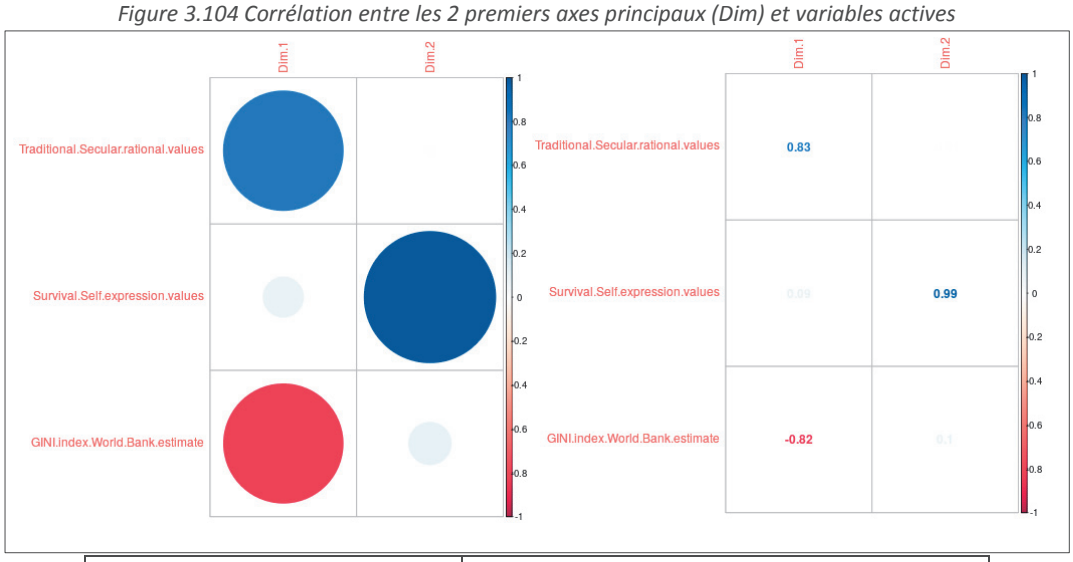
Figure 3.103 Pourcentage de variance expliquée par les axes principaux (ou dimensions)



Les 2 premiers axes principaux expriment 78.8% de l'inertie totale du jeu de données ; cela signifie que 78.8% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan.

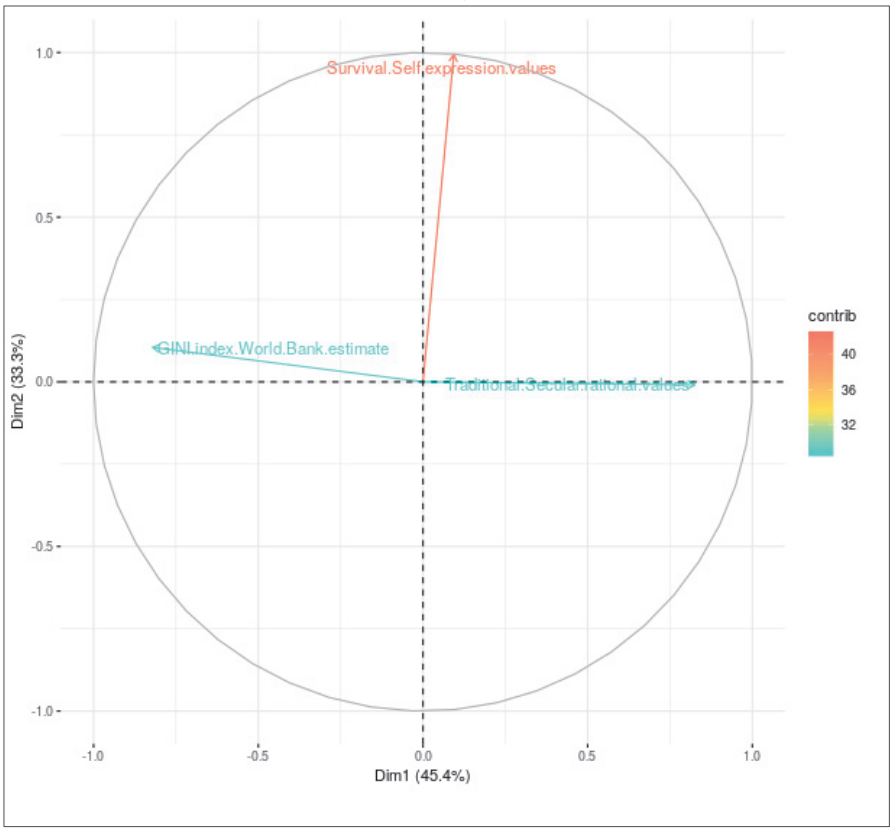
Signification des composantes principales sélectionnées

Figure 3.104 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives



Traditional/Secular-rational values	Valeurs laïques-rationnelles (vs valeurs traditionnelles)
Survival/Self-expression values	Valeurs d'expression personnelle (vs Valeurs survie)
Inequality values world bank 0.1	Valeurs inégalitaires

Figure 3.105 Cercle de corrélation entre axes principaux (Dim) et variables actives (avec leur contribution à la création des axes)



En analysant les Figure 3.104 et Figure 3.105, nous déduisons que la première composante (Dim 1) oppose les pays ayant des valeurs « *laïques-rationnelles* », et *égalitaires*, aux pays ayant des valeurs *inégalitaires*, *traditionnelles* et *religieuses*.

La première composante (Dim 1) a une corrélation positive avec l'indicateur des valeurs « *laïques-rationnelles »*, et négative avec l'indice de Gini après redistribution.

La deuxième composante (Dim 2) oppose les pays ayant des valeurs d'«*expression personnelle*», donc d'individualisme, aux pays ayant plutôt des valeurs de « *survie* ».
La deuxième composante (Dim 2) a une corrélation positive avec l'indicateur des valeurs d'«*expression personnelle*».

Pour les définitions précises des valeurs anthropologiques du modèle Inglehart-Welzel, voir la page 56.

Choix du nombre optimal de classes de pays

Après classification statistique des pays à partir des deux premiers axes principaux, nous obtenons quatre classes optimales de pays (cf. *Figure 3.106* et *Figure 3*.107 ci-dessous), qui seront décrites, après ordonnancement par le temps transitionnel qui est le ratio d'années d'études femmes vs hommes de 25-34 ans (lissé).

Figure 3.106 Gain d'inertie en fonction du nombre de classes

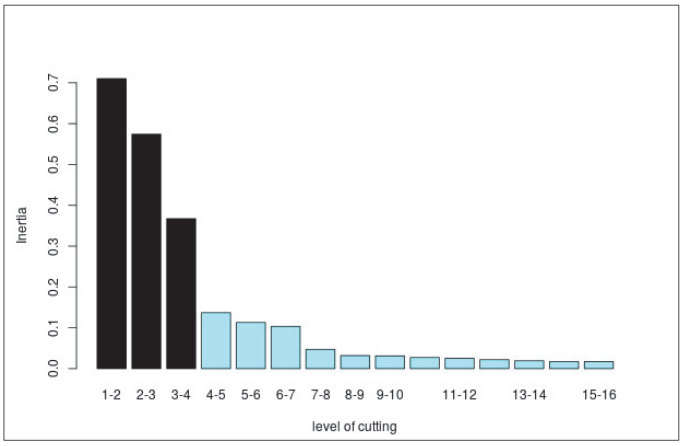
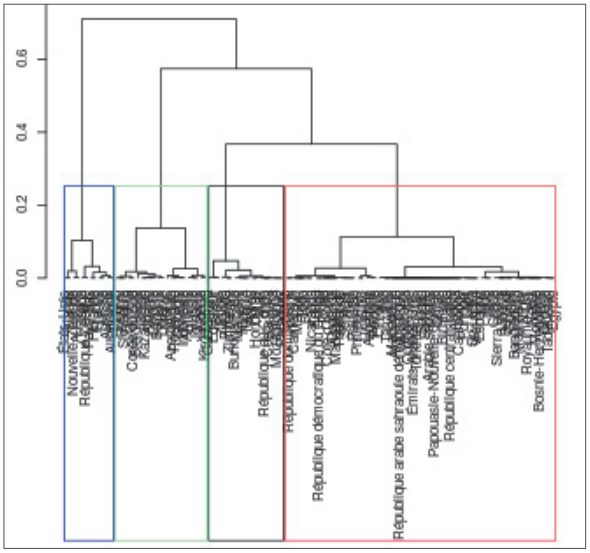


Figure 3.107 Dendrogramme de la classification ascendante hiérarchique

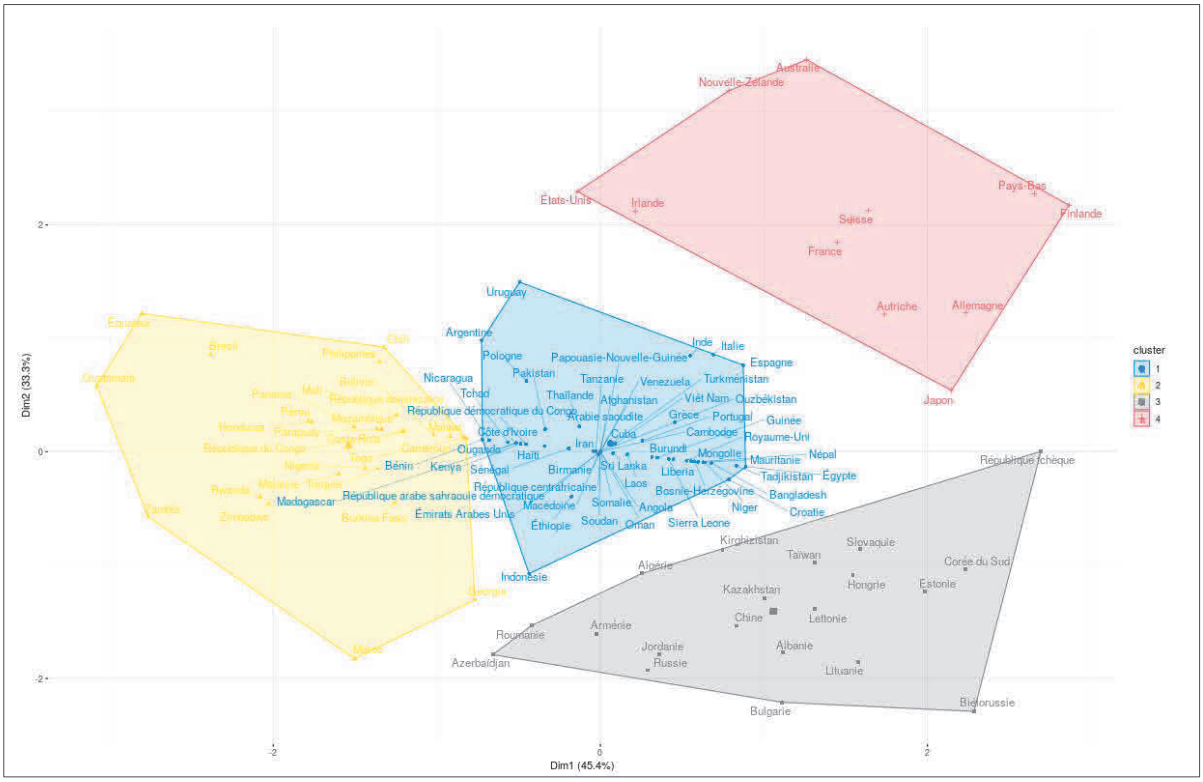


6.1.7. Indexation temporelle des classes de pays

Le temps transitionnel utilisé est le ratio d'années d'études femmes vs hommes de 25-34 ans (lissé) ; Etant donné que cette dernière variable est croissante dans le temps, les classes de pays sont numérotées par ordre croissant de leur moyenne. Ainsi, la classe 4 qui a la moyenne la plus élevée, porte l'index temporel le plus élevé (c'est la classe la plus ancienne dans le processus de transition) ; et inversement pour la classe 1 qui est à l'autre bout du spectre.
Nous qualifions les classes ainsi indicées par l'âge transitionnel moyen de leurs membres, de *classes d'âge transitionnel*.

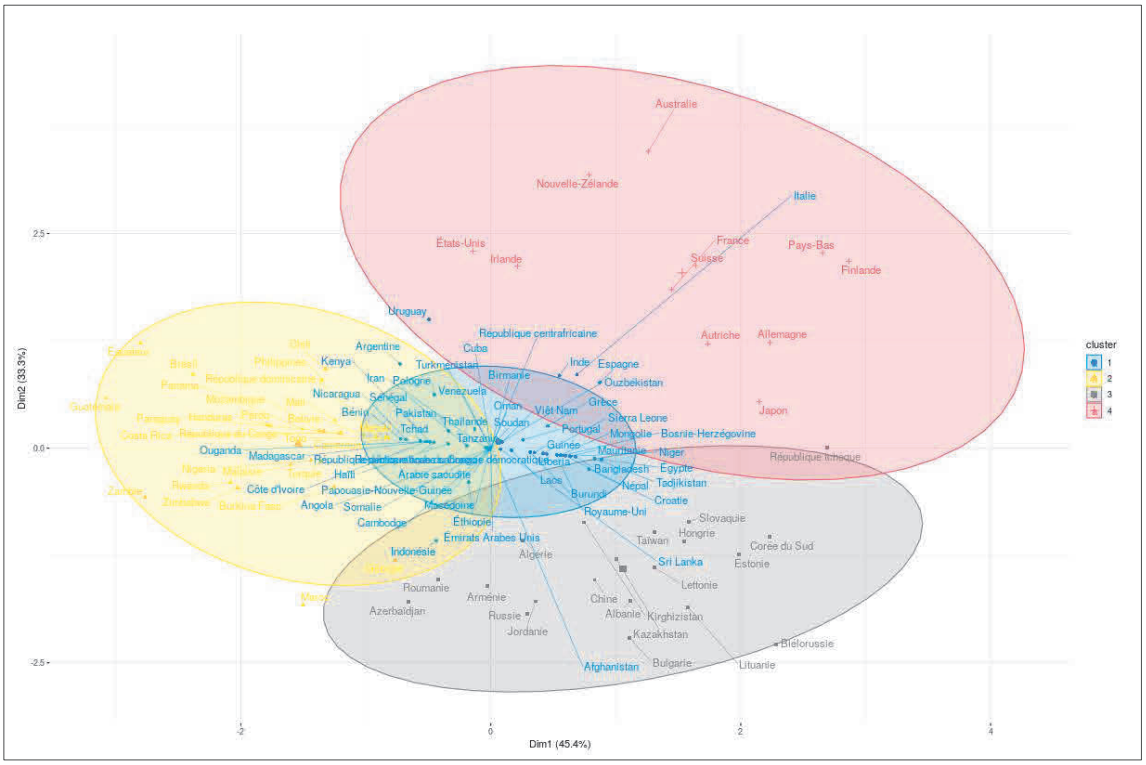
6.1.8. Description statistique des classes d'âge transitionnel de pays

Figure 3.108 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels



La représentation ci-dessus des classes d'âge transitionnel de pays, suggère un parcours temporel sous forme d'un spiral tournant à gauche en allant du centre vers l'extérieur.
Dans cette vue temporelle-transitionnelle, la classe 4, représente les pays plus anciens dans le processus de transformation cognitive. A l'opposé, les pays de la classe 1 sont les pays les plus récents dans le processus de transformation cognitive.
Ces 4 classes illustrent également les 4 phases observables « actuellement » dans le processus de transformation cognitive. La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transformation cognitive.

Figure 3.109 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes



La représentation elliptique ci-dessus suggère que les frontières des classes ne sont pas si tranchées aux frontières que le laisse supposer la représentation en polygones.

Tableau 3.7 Description des classes selon les moyennes des variables actives

Classes	Secular-rational values (vs Traditional values)	Self-expression values (vs Survival values)	Gin index after taxes and transfers (%)
1	*NS*	*NS*	*NS*
2	-0.689	*NS*	45.7
3	0.451	-0.870	32.8
4	0.683	1.111	33.3
Monde	-0.068	-0.116	38.1

NS : valeur non-significativement différente de la moyenne mondiale au seuil de 5%

La **classe 1** est composée de pays tels que : Viêt Nam, Sri Lanka, Angola, Cambodge, République Centrafricaine. Ce groupe est caractérisé par des valeurs anthropologiques qui ne diffèrent pas significativement de la moyenne mondiale.

La **classe 2** est composée de pays tels que : Brésil, Équateur, Guatemala et Zambie. Ce groupe est caractérisé par :

- de fortes valeurs inégalitaires (Indice Gini après redistribution de 45.7% dans la classe contre 38.1% au niveau mondial)
- de faibles valeurs « laïques-rationnelles », (-0.69 dans la classe contre -0.07 au niveau mondial). Donc, de fortes valeurs « traditionnelles ».

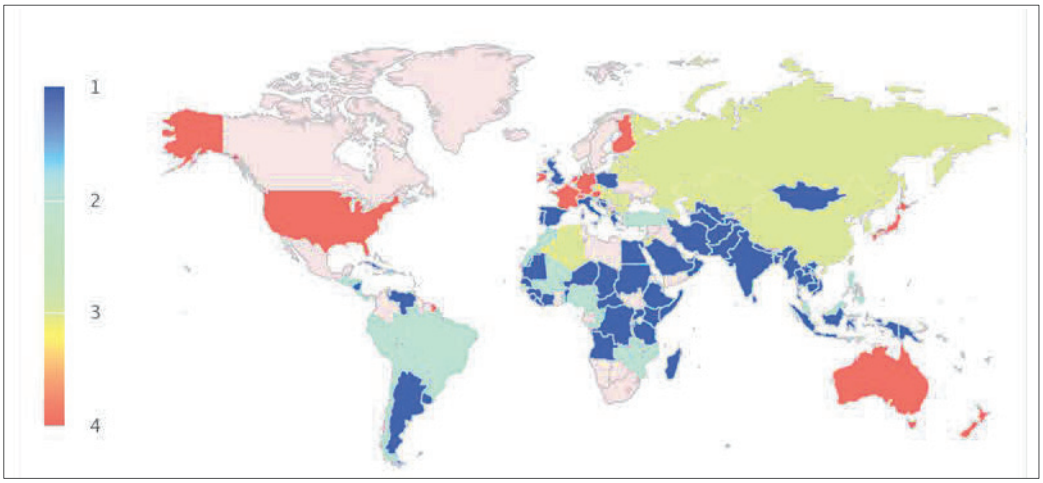
La **classe 3** est composée de pays tels que : Biélorussie, Bulgarie et République tchèque. Ces pays partagent :

- de fortes valeurs « laïques-rationnelles », (0.45 dans la classe contre -0.07 au niveau mondial). Donc, de faibles valeurs « traditionnelles ».
- de faibles valeurs inégalitaires (Indice Gini après redistribution de 32.8% dans la classe, contre 38.1% au niveau mondial)
- de faibles valeurs d'« expression personnelle » (Indicateur d'« expression personnelle » à -0.87 dans la classe, contre -0.12 au niveau mondial).

La **classe 4** est composée de pays tels que : Australie, Finlande, Irlande, Pays-Bas, Nouvelle-Zélande, Suisse et États-Unis. Ce groupe est caractérisé par :

- de fortes valeurs « laïques-rationnelles », (0.68 dans la classe contre -0.07 au niveau mondial). Donc, de faibles valeurs « traditionnelles ».
- de fortes valeurs d'« expression personnelle » (Indicateur d'« expression personnelle » à -1.11 dans la classe, contre -0.12 au niveau mondial).
- de faibles valeurs inégalitaires (Indice Gini après redistribution de 33.3% dans la classe, contre 38.1% au niveau mondial)

Figure 3.110 Cartographie des Classes temporelles de Pays



6.1.9. Variables de transition : Allures des courbes transversales selon l'indice d'âge transitionnel

Nous présentons ici, les courbes transversales des moyennes des variables de transition anthropologiques selon l'indice d'âge transitionnel.

Figure 3.111 Valeurs d'expression personnelle (vs Valeurs de survie)

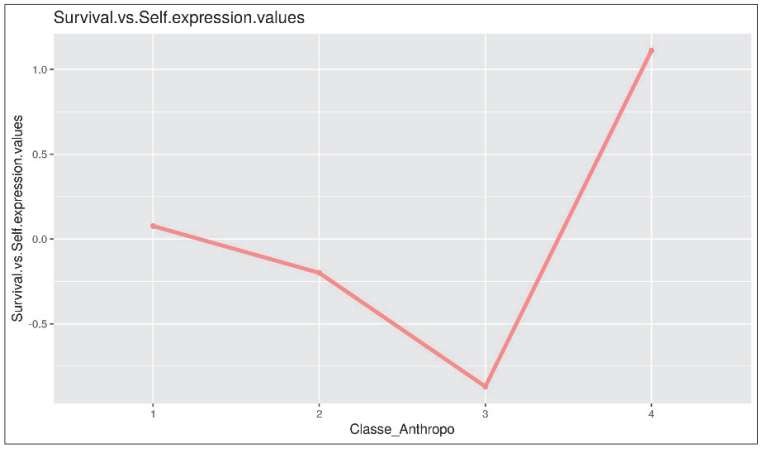


Figure 3.112 Valeurs laïques-rationnelles (vs Valeurs traditionnelles)

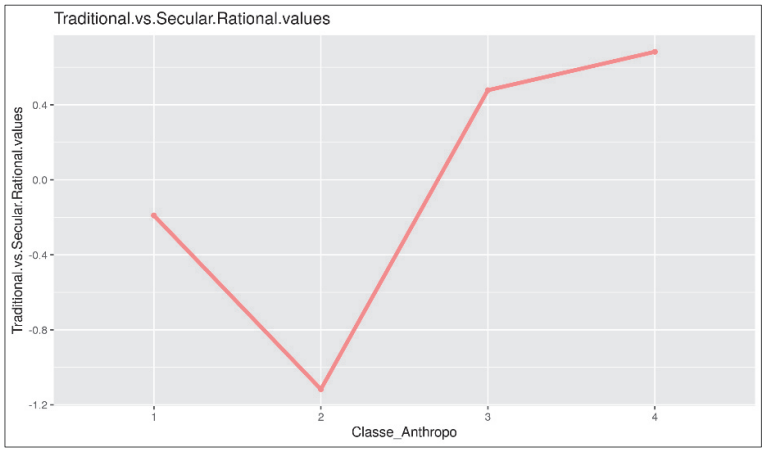
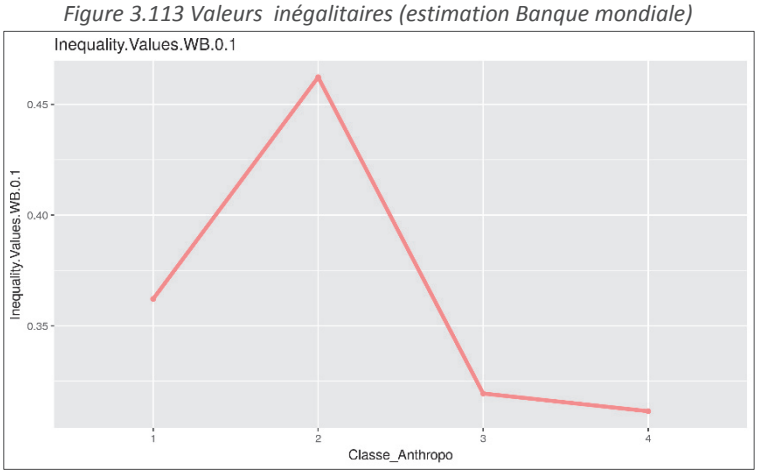


Figure 3.113 Valeurs inégalitaires (estimation Banque mondiale)



Idéalement, nous devrions effectuer des comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales afin d'évaluer la pertinence de la construction d'un schéma transitionnel à partir de l'analyse transversale. Mais le manque de profondeur historique des 3 variables de transition (pas de données avant 1980) ne permet pas ici de faire correctement de telles comparaisons.
Toutefois, ce travail comparatif entre les allures transversales et longitudinales sera effectué sur des variables passives qui ont la profondeur historique nécessaire.

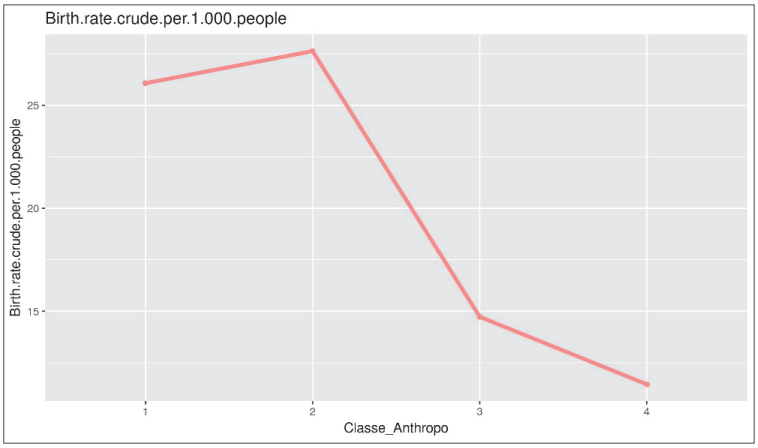
6.1.10. Variables passives : Comparaison des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales

Les variables passives sont des variables utilisées pour illustrer les classes, et qui ne sont pas des variables actives (qui elles sont utilisées pour créer les classes).

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales des variables passives, sont nécessaires pour vérifier la robustesse et la cohérence d'ordonnancement des classes d'âge transitionnelles construites.
La similitude globale des allures des courbes transversales et longitudinales que nous constatons dans les comparaisons ci-après dans plusieurs domaines, montre empiriquement cette robustesse.

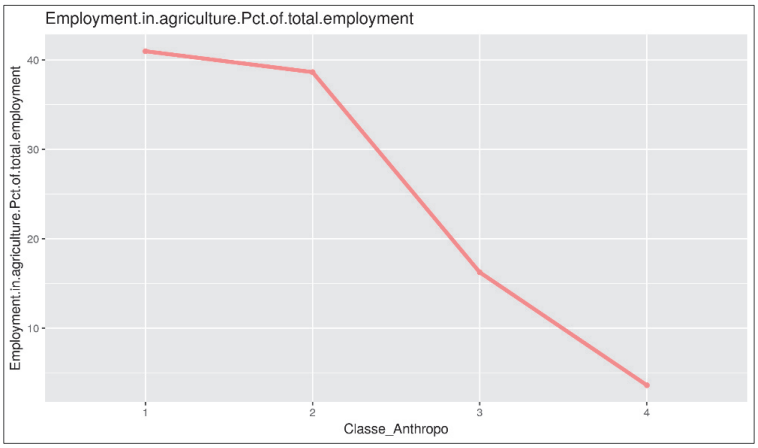
Moyennes par classe d'âge transitionnel

Figure 3.114 Taux de natalité, brut (pour 1 000 personnes)



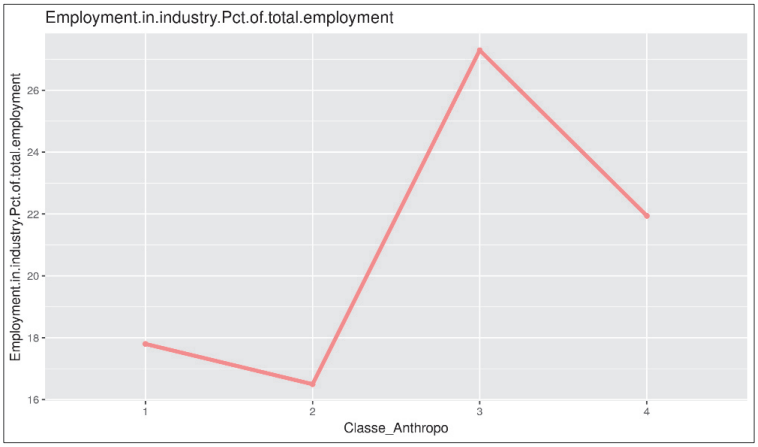
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 3.218*) nous permet de conclure à des allures baissières (après une légère hausse) similaires.

Figure 3.115 Emploi dans l'agriculture (% de l'emploi total)



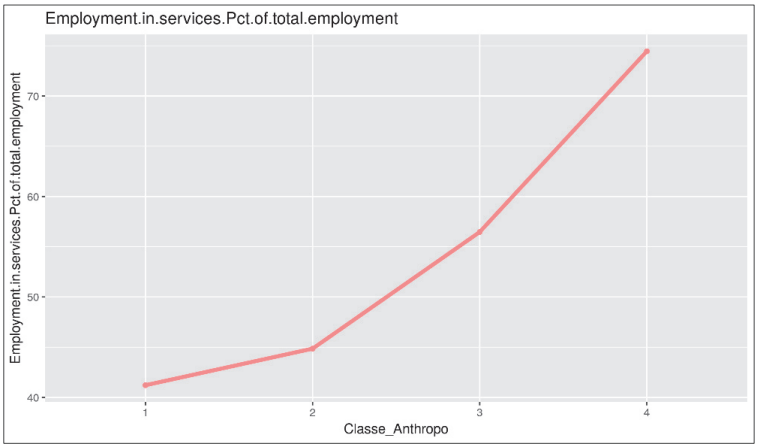
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.27*) nous permet de conclure à des allures baissières similaires.

Figure 3.116 Emploi dans l'industrie (% de l'emploi total)



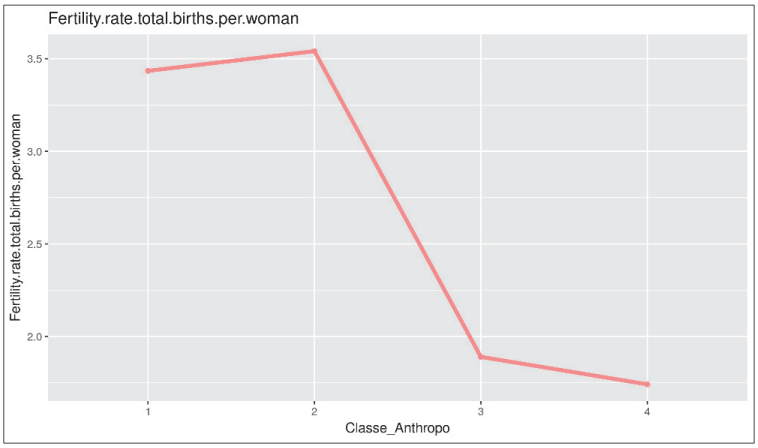
Nous ne disposons pas d'historique longue (au-delà de 1960) sur l'emploi industriel, mais l'allure en cloche de la courbe transversale rappelle celle de la courbe théorique de long terme évoquée par Fourastié (1989, p. 148) et présentée en *Figure 3.47* .

Figure 3.117 Emploi dans les services (% de l'emploi total)



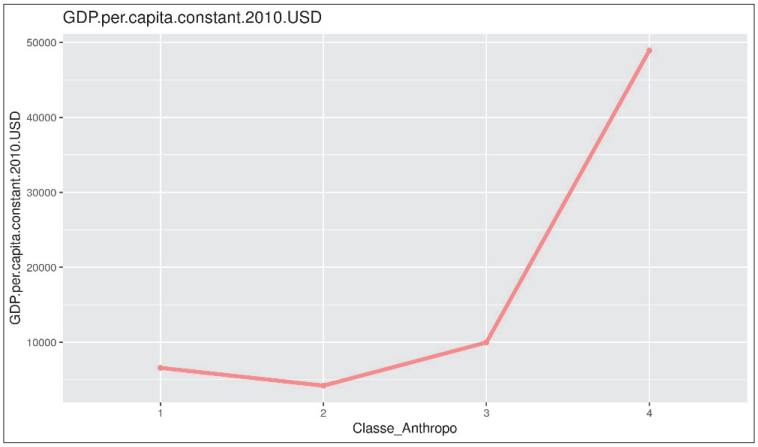
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.28*) nous permet de conclure à des allures haussières similaires.

Figure 3.118 Taux de fécondité, total (naissances par femme)



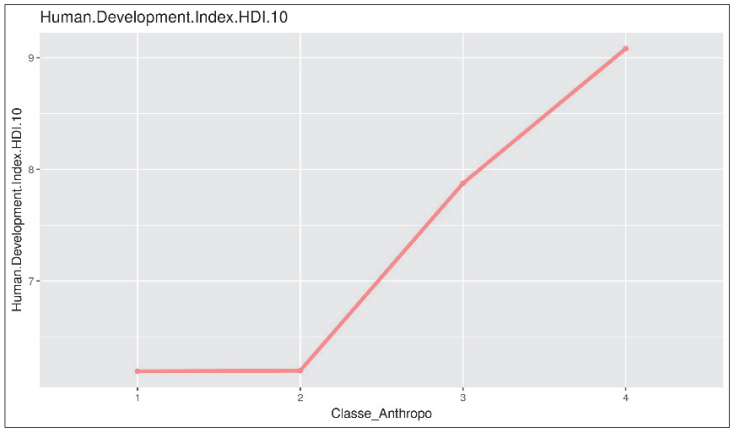
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.16*) nous permet de conclure à des allures baissières similaires. Evidemment, la courbe transversale montre les valeurs contemporaines.

Figure 3.119 PIB par habitant (USD constants de 2010)



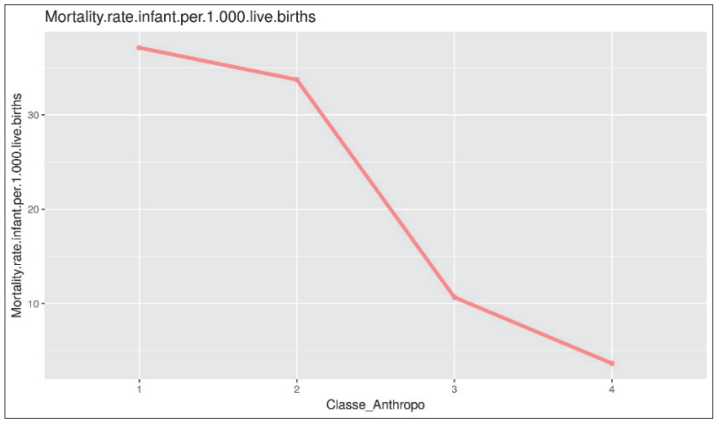
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.3*) nous permet de conclure à des allures haussières similaires (du moins sur les années récentes).

*Figure 3.120 Indice de développement humain (HDI*10)*



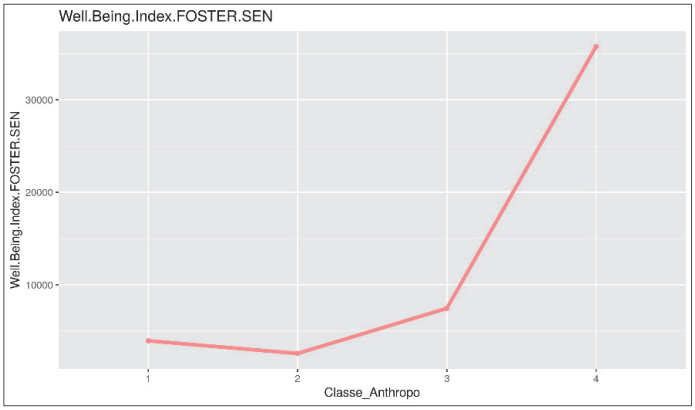
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.1*) nous permet de conclure à des allures haussières similaires (du moins sur les années récentes).

Figure 3.121 Taux de mortalité infantile (pour 1 000 naissances vivantes)



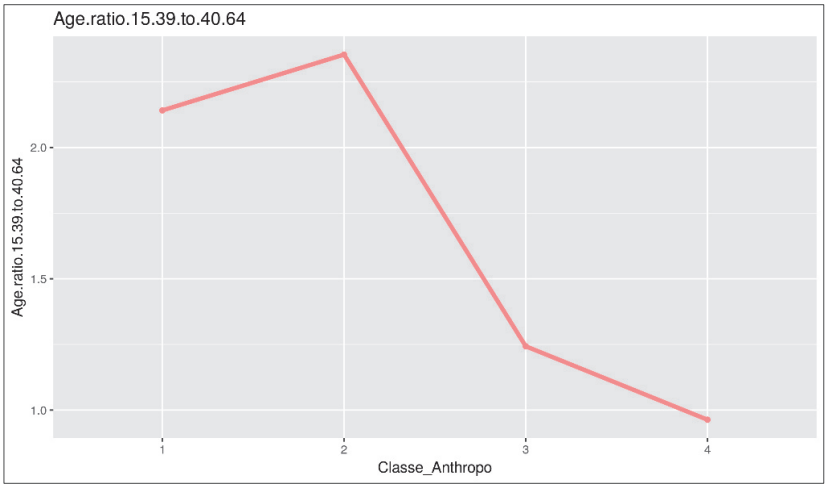
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.15)* nous permet de conclure à des allures baissières similaires (du moins sur les années récentes).

Figure 3.122 Indice de bien-être (FOSTER-SEN)



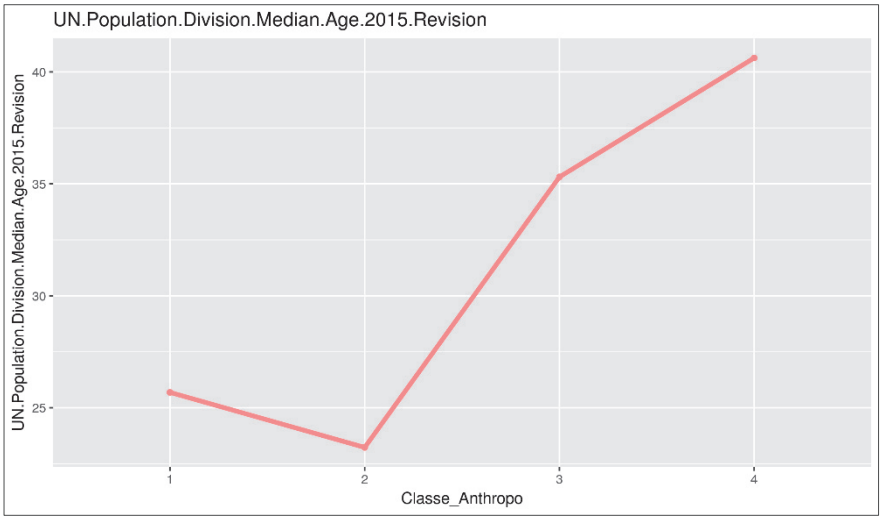
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.17*) nous permet de conclure à des allures haussières similaires (du moins sur les années récentes).

Figure 3.123 Ratio d'âge15-39 à 40-64



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.13*) nous permet de conclure à des allures haussières puis baissières similaires (du moins sur les années récentes).

Figure 3.124 Division de la population des Nations Unies (âge médian)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.5*) nous permet de conclure à des allures en S similaires (du moins sur les années récentes).

6.2. Peut-on faire l'hypothèse d'une théorie transitionnelle anthropo-culturelle ?

Nous désignons par transition anthropologique (ou anthropo-culturelle), une série de profondes et durables transformations anthropologiques de la population d'un pays, pouvant s'étaler sur quelques décennies à plusieurs générations.

Proposition de schéma transitionnel anthropologique selon l'indice d'âge transitionnel

A la section 6.1.9 nous n'avons pas pu comparer les courbes transversales des variables actives avec leurs versions longitudinales. Cependant, nous constatons que l'indice Gini des inégalités après taxes et transferts a la forme en cloche conformément à l'hypothèse de Kuznets et même en cohérence avec les *vagues de Kuznets* puisque les inégalités repartent sensiblement à la hausse dans la dernière classe d'âge transitionnelle (Milanović, 2019).
De même, la hausse dans la dernière classe d'âge transitionnelle, des valeurs d'« expression personnelle » est conforme aux observations d'évolutions historiques post-modernes (Inglehart et Welzel, 2005, .p. 5).
Enfin, Inglehart et Welzel (2005, p. 5) font également un lien historique entre la révolution industrielle, donc la modernité et le passage de valeurs « traditionnelles » aux valeurs « laïques-rationnelles ».
Par ailleurs, nous avions également montré à la section 6.1.10 que les courbes transversales des variables passives selon l'indice d'âge transitionnel, avaient sensiblement les mêmes allures que leurs courbes longitudinales.
En considérant les arguments ci-dessus énumérés, nous pouvons proposer le schéma théorique transitionnel suivant :

Figure 3.125 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne par classe d'âge transitionnel

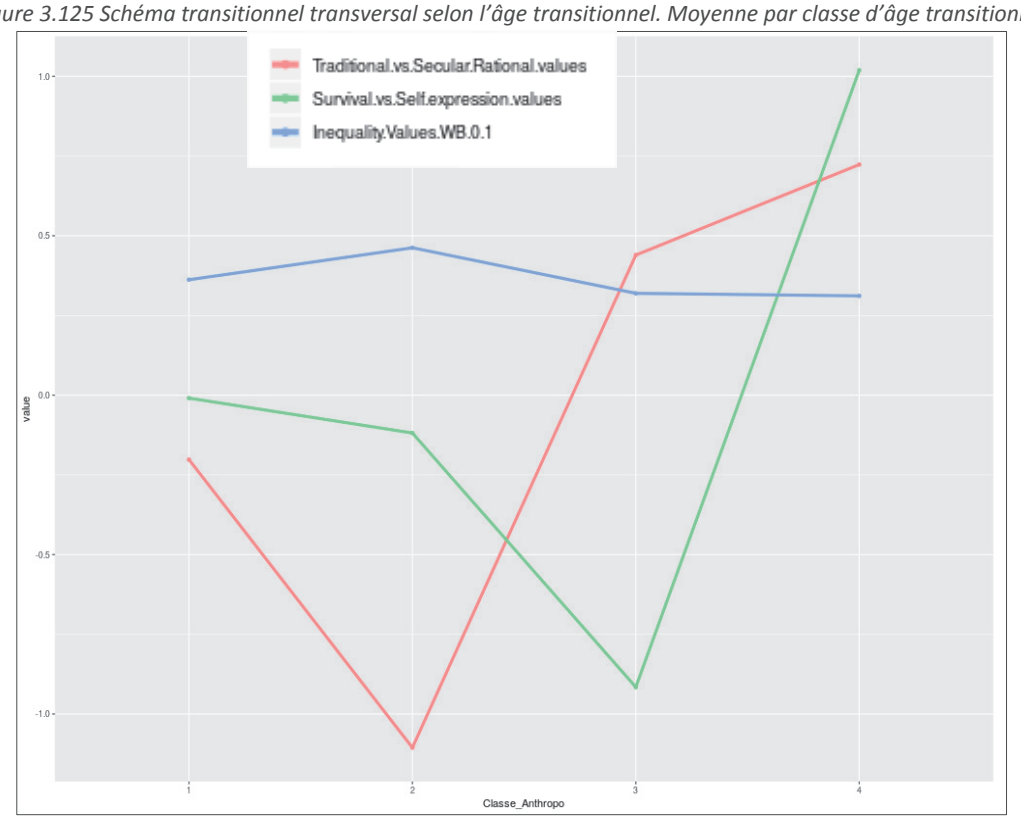
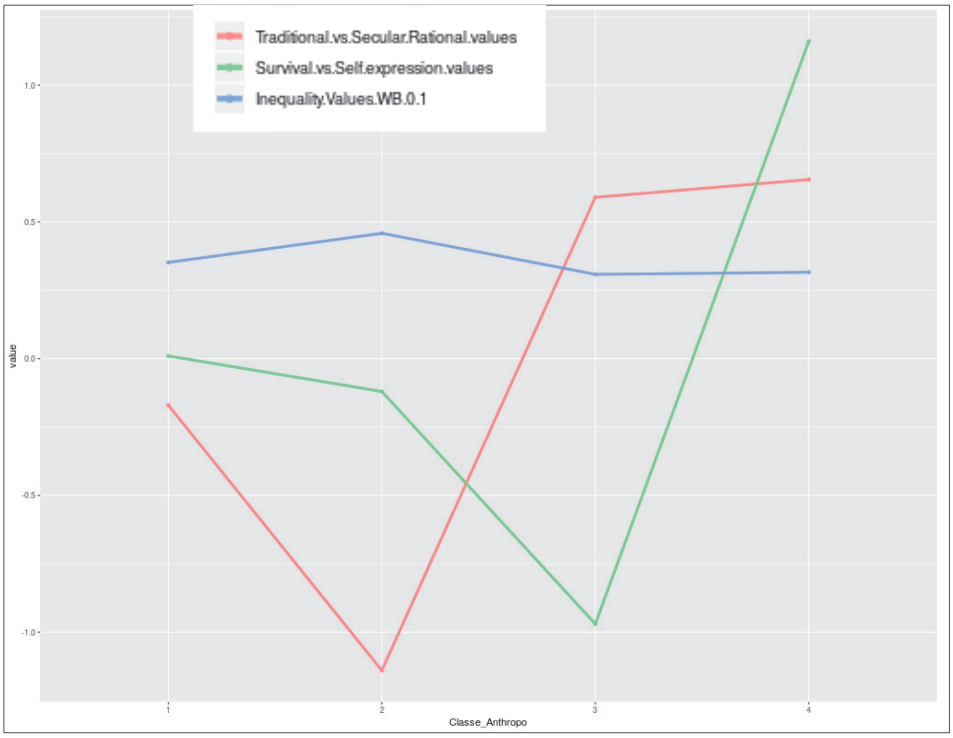
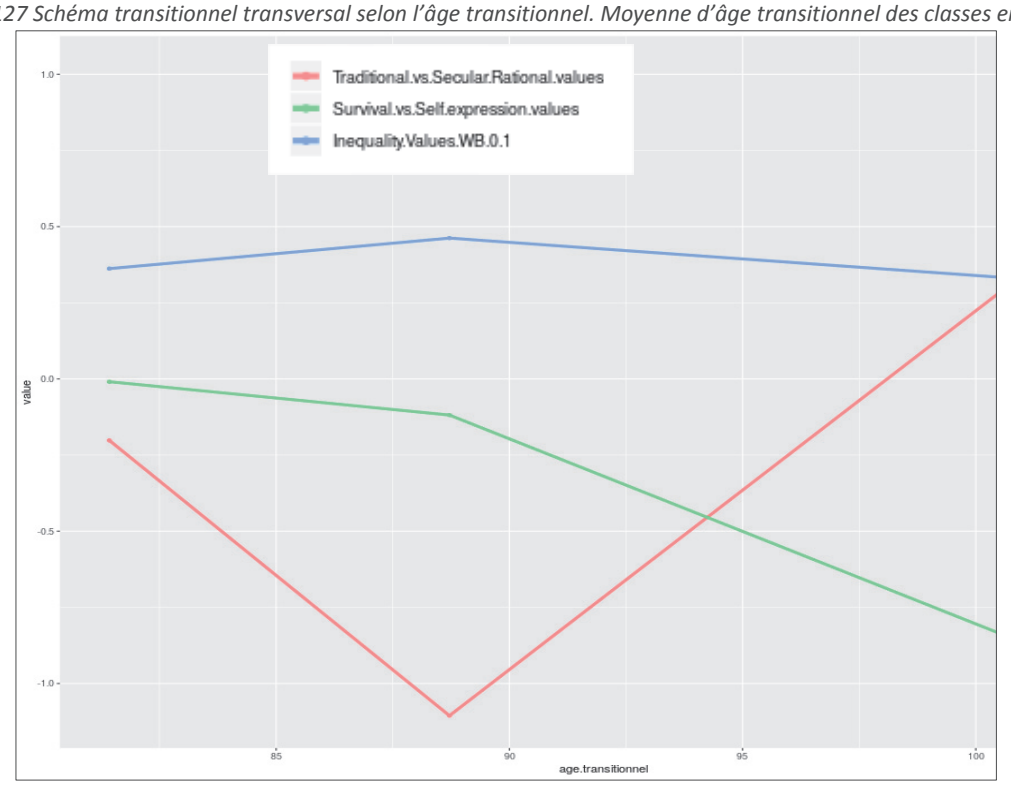


Figure 3.126 Schéma transitionnel transversal selon l'âge transitionnel. Médiane par classe d'âge transitionnel



L'allure des deux schémas ci-dessus est sensiblement la même, aussi bien en utilisant les moyennes que les médianes : C'est l'illustration d'une robustesse du schéma transitionnel obtenu en coupe transversale.

Figure 3.127 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse



Nous pouvons illustrer le schéma théorique de la transition anthropologique par le tableau des valeurs moyennes des variables transitionnelles par classes d'âge transitionnel.

Tableau 3.8 Valeurs moyennes des variables transitionnelles par classes d'âge transitionnel

Classe d'âge transitionnel	% de pays en 2014	Secular-rational values (vs Traditional values)	Self-expression values (vs Survival values)	Inequality values
1	50,0%	-0,19	0,077	0,370
2	23,3%	-0,689	-0,199	0,457
3	17,2%	0,451	-0,87	0,328
4	9,5%	0,683	1,111	0,333

Ces 4 classes d'âge transitionnel illustrent également les 4 phases observables « actuellement » dans le processus de transition anthropologique.
La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transition anthropologique.
Quant aux probables phases futures jamais observées, seules des études prospectivistes permettraient d'en spéculer les caractéristiques.

7. Institutions politiques et économiques

Nous allons ici utiliser le temps et l'âge transitionnel pour réaliser une inférence des trajectoires longitudinales dans une coupe transversale dans le domaine institutionnel.

7.1. Analyse transversale institutionnelle

7.1.1. Sélection des variables à étudier

Les variables structurelles utilisées se répartissent en deux catégories : les variables de transition et les variables actives.
Variables de transition : Elles sont utilisées pour décrire le processus de changement institutionnel. Dans le cas présent, nous-nous restreignons au sous-domaine institutionnel « politique interne » (cf. notre matrice institutionnelle IEPE, chapitre 1). Ce qui nous conduit au choix des variables de transition indiquées au Tableau 3.9
Variables actives : Elles sont utilisées pour effectuer la classification statistique des pays. Elles incluent les variables de transition auxquelles sont ajoutées des variables d'homogénéisation des classes. Dans le cas présent, il n'y a pas de variables d'homogénéisation : les variables actives sont les mêmes que les variables de transition.

Tableau 3.9 Variables actives ou transitionnelles

Variable	Signification en anglais	Signification en français	Variable active	Variable de transition
Overall.index.civil.liberties.1.free.7.repression	Overall index of civil liberties (1=free, 7=repression)	Indice global des libertés civiles (1=libre, 7=répression)	oui	oui
DGI.PR	Democracy Global Index (Polity IV-Roser)	Indice global de démocratie (Polity IV-Roser)	oui	oui
Human.Development.Index.(HDI*10)	Human Development Index (HDI*10)	Indice de développement Humain (IDH)	oui	oui

Les « libertés civiles » désignent ici, la liberté d'expression, d'assemblée, d'association, d'éducation et de religion ; ainsi que l'Etat de droit et une activité économique libre des citoyens.

L'indice global de démocratie est établi à partir de quatre facteurs : 1) le niveau d'ouverture par rapport à la provenance des candidats (y compris ceux qui sont externes à l'élite politique du pays); 2) le niveau de compétition entre les acteurs politiques (élection plutôt que sélection); 3) les limitations au pouvoir exécutif (tribunaux, pouvoir législatif, corps organisés de la société) et 4) le niveau de liberté dans l'espace public, des organisations qui défendent des opinions politiques. L'échelle va de 0 à 10, 0 étant le niveau le plus bas et 10 le niveau le plus élevé de démocratie.

Indice de développement humain (IDH), est un indicateur qui mesure la qualité de vie, en se basant sur l'économie (logarithme du PIB par habitant USD constants de 2010), la santé (espérance de vie à la naissance) et l'éducation (durée moyenne de scolarisation). Les 3 facteurs ont le même poids et l'indice varie de 0 à 1.

L'indice de développement humain est de fait, une mesure des résultats des dispositions institutionnelles dans les domaines politiques, économiques, sanitaires et éducatives. L'indice de développement humain est donc légitime pour être utilisé pour caractériser (au moins en partie) les institutions d'un pays à un moment donné.

7.1.2. Transformation préalable des variables à étudier

Les variables actives utilisées font l'objet des transformations suivantes, avant la sélection de l'année d'étude : estimation des valeurs manquantes par interpolation spline, et moyenne mobile d'ordre 5.

7.1.3. Sélection des pays à étudier

En plus des pays exclus pour l'ensemble des domaines (cf. ***Error! Reference source not found.***), nous excluons également les pays de l'OPEP, dont la trajectoire politique est susceptible de perturbations spécifiques à la présence de ressources pétrolières (Acemoglu et Robinson, 2015, p.426, 544).

7.1.4. Sélection de l'année à étudier

Pour l'analyse transversale, l'année étudiée est 2014. Cela pour des raisons de complétude et de fiabilité des données, car les dernières (2017, source Banque Mondiale) sont souvent complétées et révisées après publication.

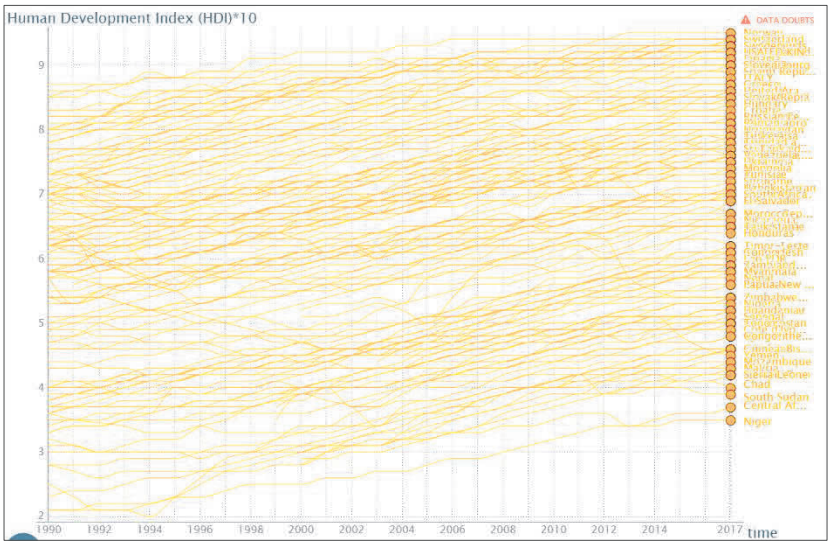
7.1.5. Choix du temps transitionnel

Variable à tester comme temps transitionnel

La variable à tester est une valeur lissée (moyenne mobile arrière d'ordre 5) de l'indice de développement humain (*Human Development Index*, HDI*10).
L'indice de développement humain est légitime pour être utilisé pour caractériser (au moins en partie) les institutions d'un pays à un moment donné (cf. arguments de la section 7.1.1).

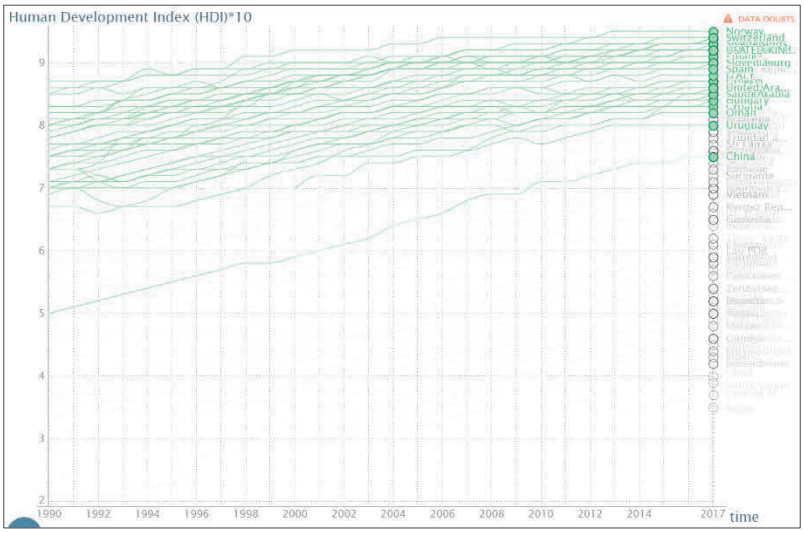
Une analyse visuellement (cf. *Figure 3.128* à *Figure 3.131* ci-dessous) de la tendance mondiale de l'indice de développement humain, permet d'avancer l'hypothèse haussière universelle (et peut-être à une forme de convergence), même si on ne peut parler de monotonie mathématique stricte. C'est la raison pour laquelle nous utilisons un lissage par moyenne mobile pour améliorer la monotonie de la variable pressentie pour jouer le rôle de temps transitionnel.

*Figure 3.128 Indice de développement humain (HDI*10)*



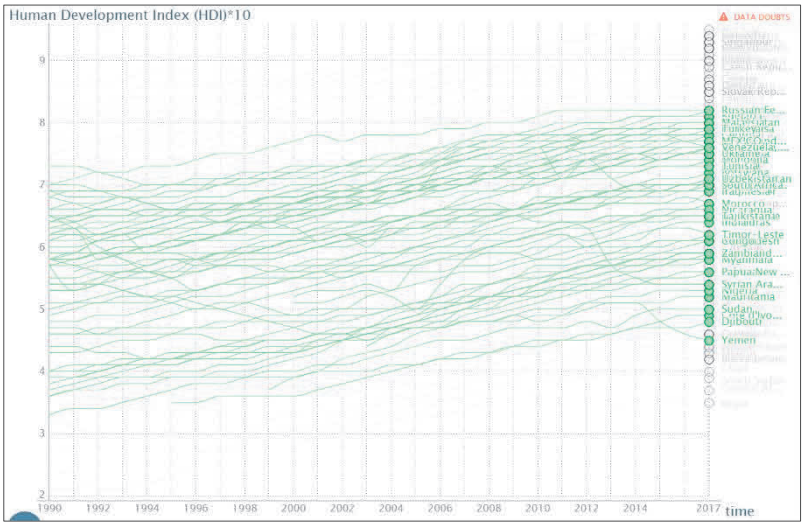
Source : Perspective Monde (2018)

*Figure 3.129 Indice de développement humain (HDI*10) des pays de l'OCDE*



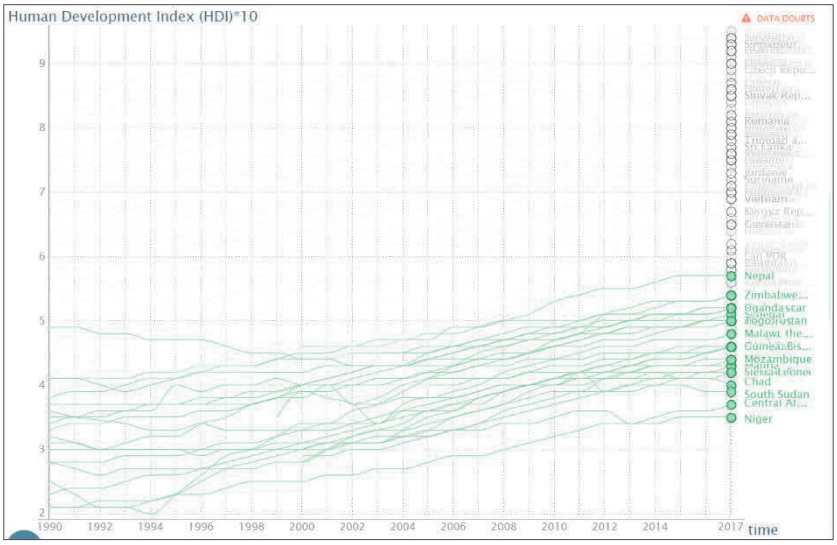
Source : Perspective Monde (2018)

*Figure 3.130 Indice de développement humain (HDI*10) des pays à moyen revenu par habitant (au sens de la Banque Mondiale)*



Source : Perspective Monde (2018)

*Figure 3.131 Indice de développement humain (HDI*10) des pays à faible revenu par habitant (au sens de la Banque Mondiale)*



Source : Perspective Monde (2018)

Vérification de la monotonie empirique de la variable à tester

La variable à tester Y est empiriquement monotone lorsque la série différenciée dY (dY(t)=Y(t) – Y(t-1)) est de signe constant sur la période étudiée, pour tout pays.
De plus pour assurer l'universalité de cette monotonie empirique, le signe de série différenciée doit être le même pour tous les pays étudiés

Pour chaque pays, nous calculons le pourcentage de valeurs positives de la série différenciée, afin de vérifier l'unicité du signe sur la période étudiée.
Sur l'ensemble des 118 pays étudiés, le pourcentage moyen de valeurs positives de la série différenciée est de 93.6%, avec un minimum à 42.3% (Zimbabwe), une médiane à 100%, et un maximum à 100%.
Mieux, 82 des 118 pays, soit 69.5%, ont 100% de valeurs positives dans la série différenciée, et 101 des 118 pays, soit 85.6%, ont plus de 80% de valeurs positives dans la série différenciée

Par ailleurs, les valeurs centrales (moyenne=0.9 et médiane=0.99) du coefficient de Kendall indiquent une très forte relation monotone croissante entre l'Indice de Développement Humain lissé et le temps.

Vérification empirique de l'unicité du signe de la pente de la variable testée, quel que soit le pays

A défaut d'une monotonie empirique parfaite, l'unicité pour tous les pays, du signe de la pente de la droite d'ajustement de la variable à tester, montre une universalité de la tendance (à la hausse ou à la baisse) sur l'ensemble des pays étudiés.
De plus, si pour un pays, pente de la droite d'ajustement de la variable à tester est proche en valeur absolue de 1, alors cette variable à tester se rapproche d'autant plus de la variable temps.

Nous régressions linéairement la variable à tester par rapport au temps. Le signe du coefficient du temps est positif, lorsque la variable à tester a une tendance linéaire croissante dans le temps.
Ci-dessous les statistiques sommaires issues des 118 régressions (118 pays).

Summary(coef.Intercept)

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
-10.1795  -1.8175  -1.8062  -1.9738  -1.7518  -0.2092        6
```

Summary(coef.year)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 0.2355  1.9721  2.0334  2.0859  2.0461  6.8760       6
```

Summary(coef.R2)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
0.01316 0.92005 0.97728 0.91107 0.99008 0.99725       6
```

La pente moyenne de l'indice de développement humain lissé est de +2.1 (médiane de +2).
Pour 100% des régressions, le coefficient du temps (coef.year), est systématiquement positif, confirmant ainsi l'hypothèse d'une hausse tendancielle de l'indice de développement humain, pour l'ensemble des pays.

Figure 3.132 Exemple de la France d'après-guerre, indice de développement humain (IDH) lissé

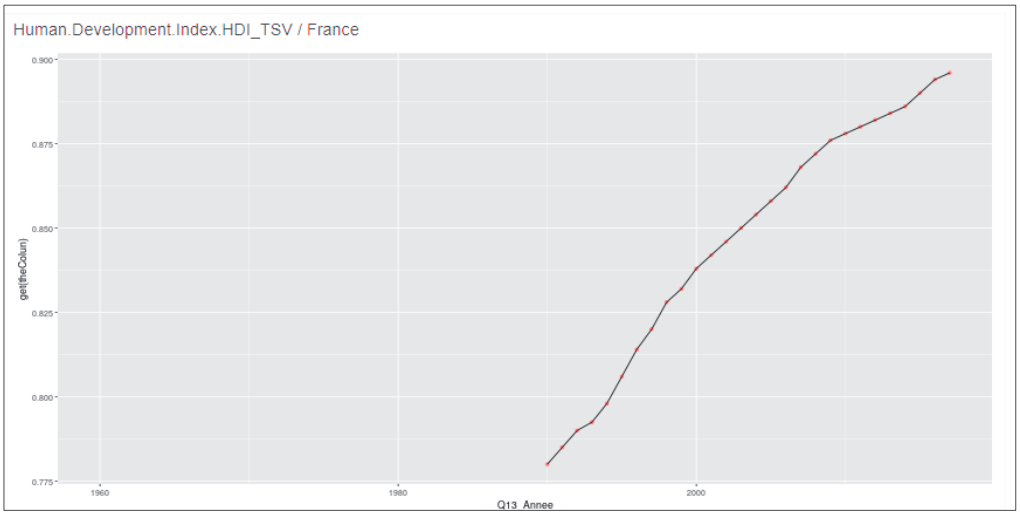
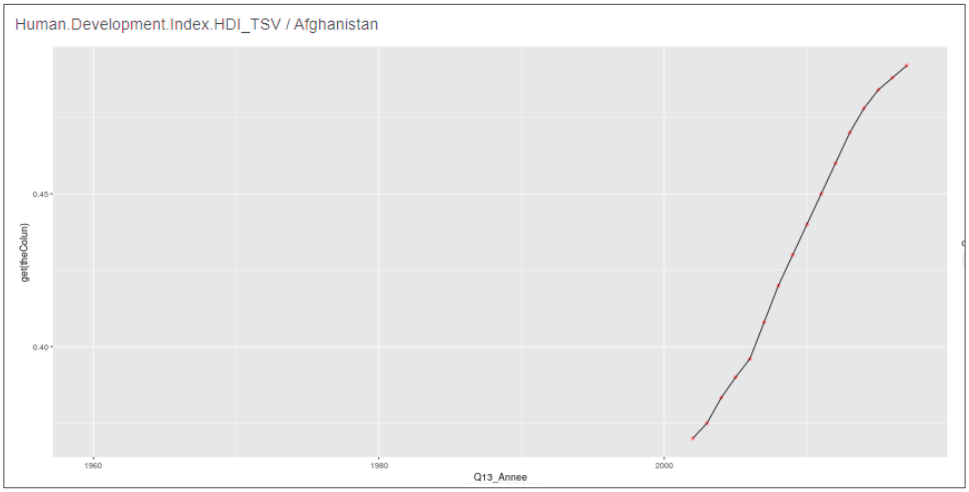


Figure 3.133 Exemple de l'Afghanistan, indice de développement humain (IDH) lissé



Conclusion

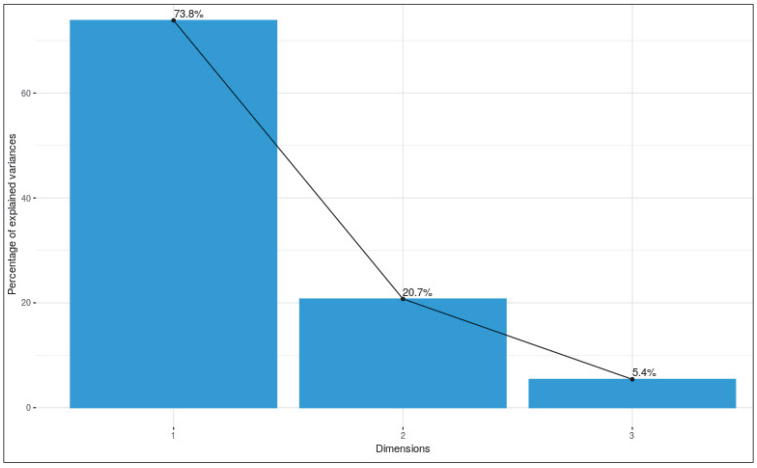
En conclusion, l'indice de développement humain (lissé) est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale, du fait de sa tendance haussière généralisée, au moins depuis les années 1990.

7.1.6. Classification des pays, sur composantes principales

Sélection des composantes principales pour la classification

L'analyse en composantes principales (ACP) est réalisée à partir des 3 variables actives (cf. Tableau 3.9).

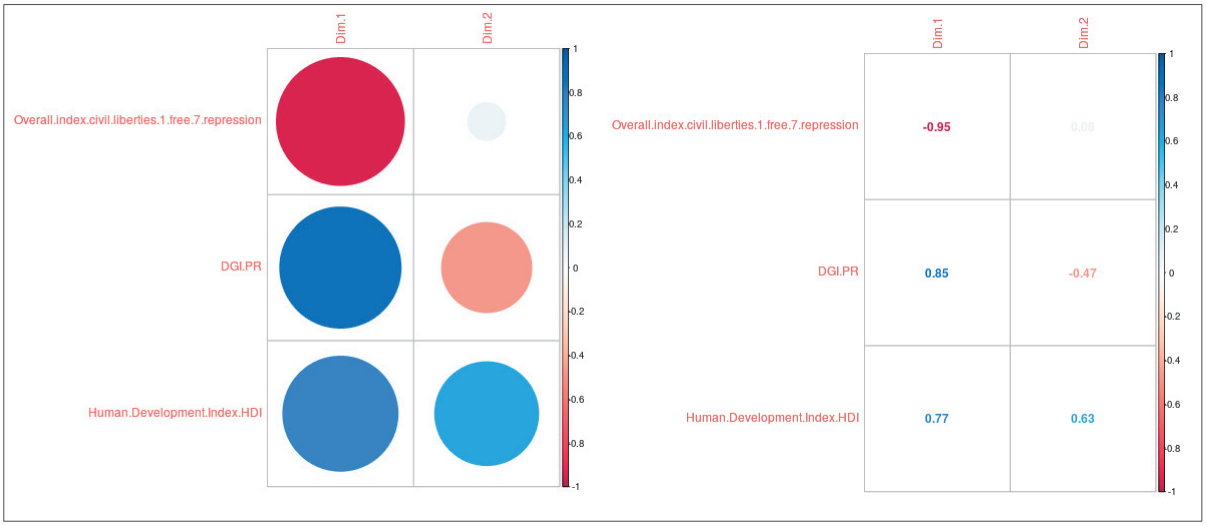
Figure 3.134 Pourcentage de variance expliquée par les axes principaux (ou dimensions)



Les 2 premiers axes principaux expriment 94.6% de l'inertie totale du jeu de données ; cela signifie que 94.6% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan.

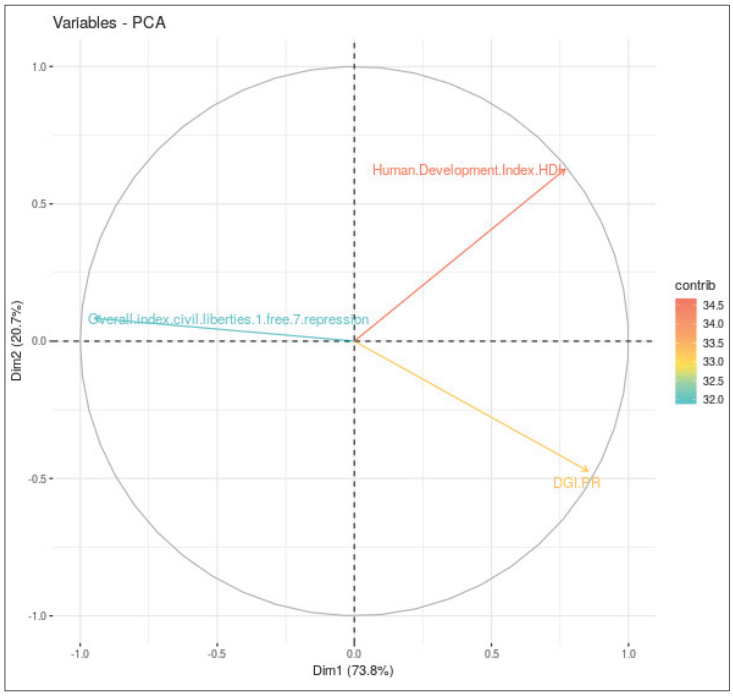
Signification des composantes principales sélectionnées

Figure 3.135 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives



Variable	Signification
Overall.index.civil.liberties.1.free.7.repression	Indice global des libertés civiles (1=libre, 7=répression)
DGI.PR	Indice global de démocratie (Polity IV-Roser)
Human.Development.Index.(HDI*10)	Indice de développement humain (IDH)

Figure 3.136 Cercle de corrélation entre axes principaux et variables actives (avec leur contribution à la création des axes)



En analysant *Figure 3.135* et *Figure 3.136*, nous déduisons que la première composante (Dim 1) oppose les pays ayant un fort indice de développement humain (HDI), et un fort Indice global de démocratie (DGI), aux pays ayant un fort niveau de répression des libertés civiles.

La deuxième composante (Dim 2) oppose les pays ayant un fort indice de développement humain (HDI), aux pays ayant un fort Indice global de démocratie (DGI).

Pour les définitions précises des variables utilisées ci-dessus, voir la page 225 .

Choix du nombre optimal de classes de pays

Après classification statistique des pays à partir des deux premiers axes principaux, nous obtenons quatre classes optimales de pays (cf. Figure 3.137 et Figure 3.138 ci-dessous), qui seront décrites, après ordonnancement par le temps transitionnel qui est l'indice de développement humain (lissé).

Figure 3.137 Gain d'inertie en fonction du nombre de classes

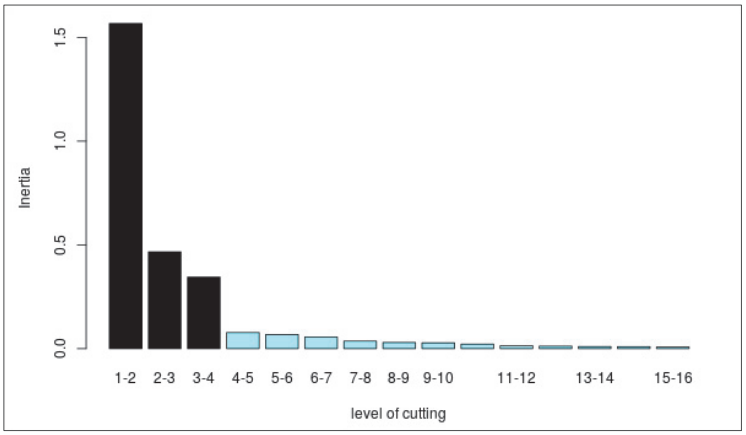
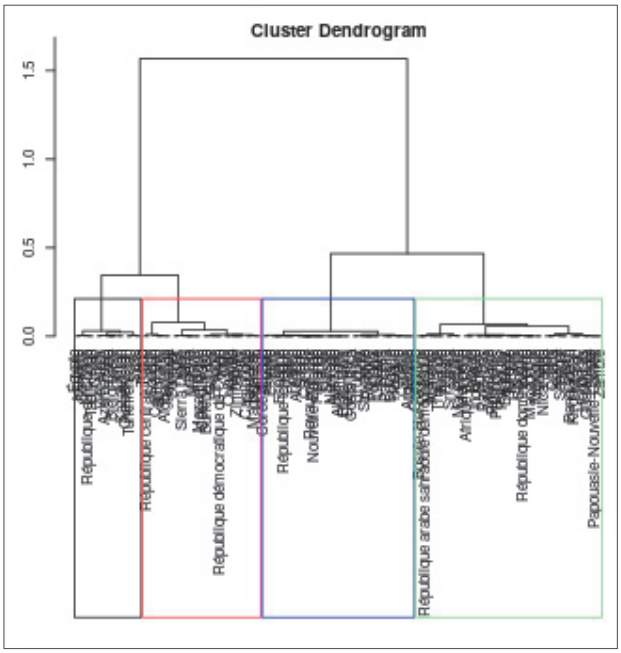


Figure 3.138 Dendrogramme de la classification ascendante hiérarchique



7.1.7. Indexation temporelle des classes de pays

Le temps transitionnel utilisé est l'indice de développement humain (lissé); Etant donné que cette dernière variable est croissante dans le temps, les classes de pays sont numérotées par ordre croissant de leur moyenne. Ainsi, la classe 4 qui a la moyenne la plus élevée, porte l'index temporel le plus élevé (c'est la classe la plus ancienne dans le processus de transition) ; et inversement pour la classe 1 qui est à l'autre bout du spectre.
Nous qualifions les classes ainsi indicées par l'âge transitionnel moyen de leurs membres, de *classes d'âge transitionnel*.

7.1.8. Description statistique des classes d'âge transitionnel de pays

La représentation ci-dessous des classes d'âge transitionnel de pays, suggère un parcours temporel sous forme d'un V inversé en allant de la gauche vers la droite.
Dans cette vue temporelle-transitionnelle, la classe 4, représente les pays plus anciens dans le processus de transformation cognitive. A l'opposé, les pays de la classe 1 sont les pays les plus récents dans le processus de transformation cognitive.
Ces 4 classes illustrent également les 4 phases observables « actuellement » dans le processus de transformation cognitive. La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transformation cognitive.

Figure 3.139 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels

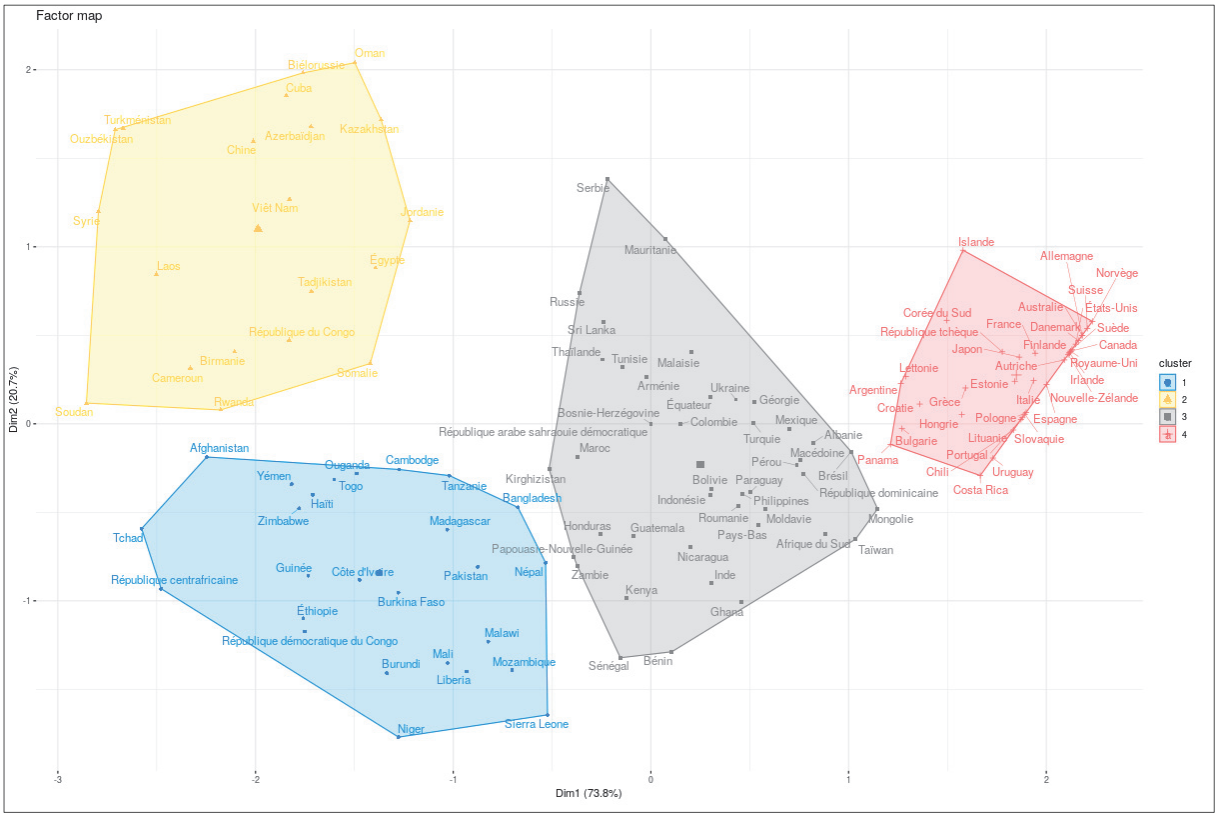
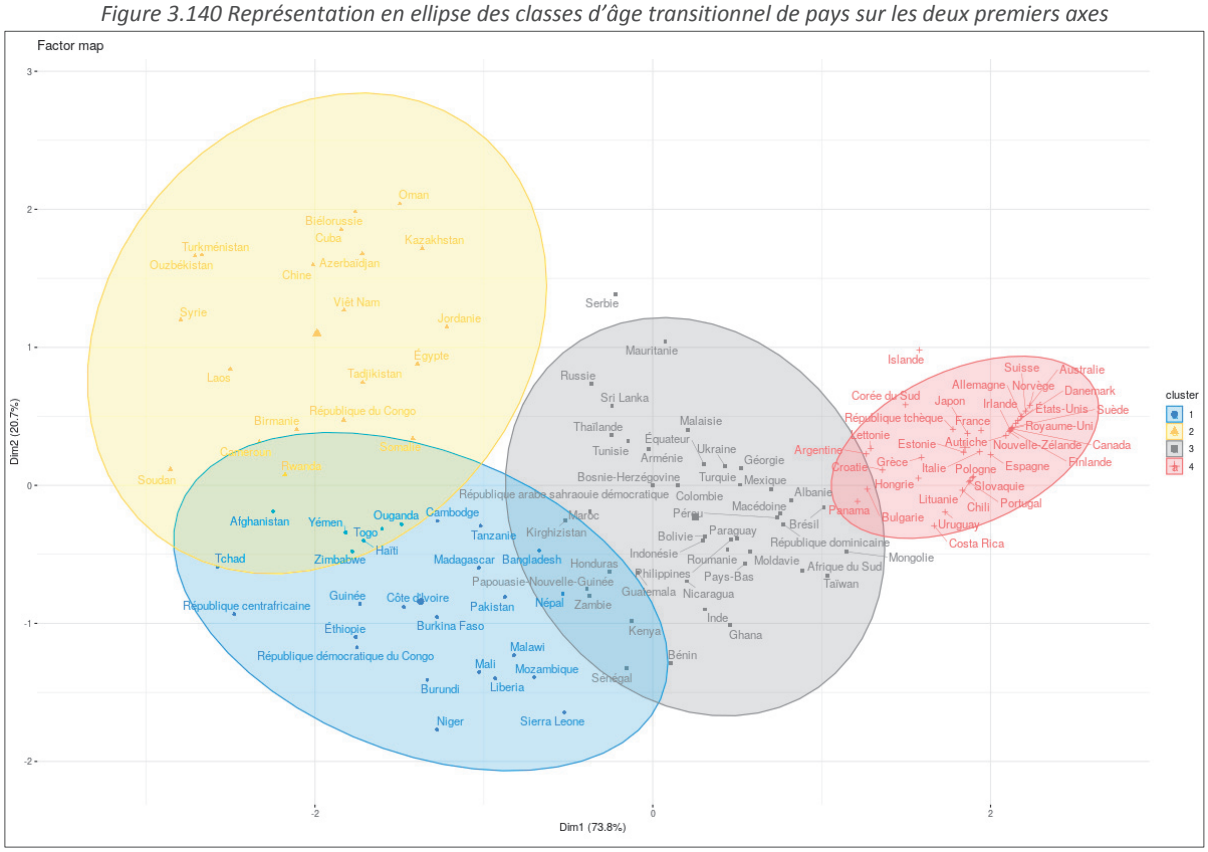


Figure 3.140 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes



La représentation elliptique ci-dessus suggère que les frontières des classes ne sont pas si tranchées que le laisse supposer la représentation en polygones.

Tableau 3.10 Description des classes selon les moyennes des variables actives

Classes	Moyenne Indice global des libertés civiles (1 : libre, 7 : répression)	Moyenne Indice global de démocratie (Polity IV-Roser)	Moyenne Indice de développement humain
1	4,6	2,7098836	0,4560769
2	5,835	-5,365006	*NS*
3	*NS*	6,623812	*NS*
4	1,2457143	9,4764183	0,8685714
Monde	3,370833	4,67464	0,6834915

NS : valeur non-significativement différente de la moyenne mondiale au seuil de 5%

La **classe 1** est composée de pays tels que : Burundi, Niger et Sierra Leone. Ce groupe est caractérisé par :

- de forts niveaux de répression des libertés civiles (4.6 dans la classe contre 3.4 au niveau mondial)
- de faibles valeurs de l'indice de développement humain (0.5 dans la classe contre 0.7 au niveau mondial) et de l'indice global de démocratie (2.7 dans la classe contre 4.7 au niveau mondial)

La **classe 2** est composée de pays tels que : Azerbaïdjan, Biélorussie, Chine, Cuba, Kazakhstan, Oman, Soudan, Syrie, Turkménistan et Ouzbékistan. Ce groupe est caractérisé par :

- de forts niveaux de répression des libertés civiles (5.8 dans la classe contre 3.4 au niveau mondial)
- de faibles valeurs de l'indice global de démocratie (-5.4 dans la classe contre 4.7 au niveau mondial)

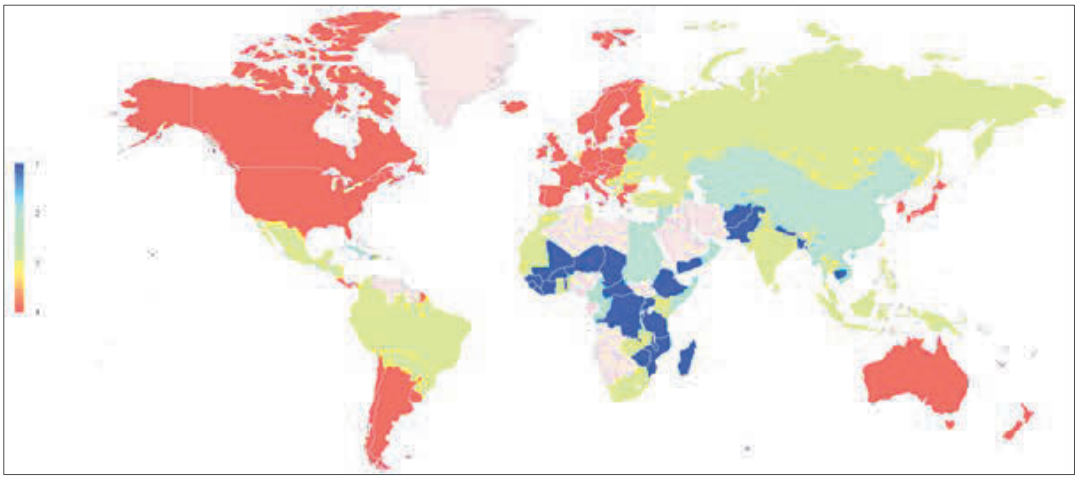
La **classe 3** est composée de pays tels que : Bolivie, Indonésie, Colombie, Philippines, Paraguay. Ce groupe est caractérisé par :

- de fortes valeurs de l'indice global de démocratie (6.6 dans la classe contre 4.7 au niveau mondial)

La **classe 4** est composée de pays tels que : Estonie, Italie, Japon, République tchèque. Ce groupe est caractérisé par :

- de fortes valeurs de l'indice de développement humain (0.9 dans la classe contre 0.7 au niveau mondial) et de l'indice global de démocratie (9.5 dans la classe contre 4.7 au niveau mondial)
- de faibles niveaux de répression des libertés civiles (1.2 dans la classe contre 3.4 au niveau mondial)

Figure 3.141 Cartographie des Classes temporelles de Pays



7.1.9. Variables de transition : Allures des courbes transversales selon l'indice d'âge transitionnel

Nous présentons ici, les courbes transversales des moyennes des variables de transition institutionnelles selon l'indice d'âge transitionnel.

La comparaison empirique entre chaque courbe transversale transitionnelle, et l'ensemble des courbes longitudinales associées, n'est pas possible à cause de l'allure à étages de ces dernières.
Ces étages en effets, induisent des superpositions des courbes de différents pays, ce qui empêche une visualisation claire de l'allure des différents pays.

Figure 3.142 Indice global des libertés civiles (1=libre, 7=répression). Courbe transversale.

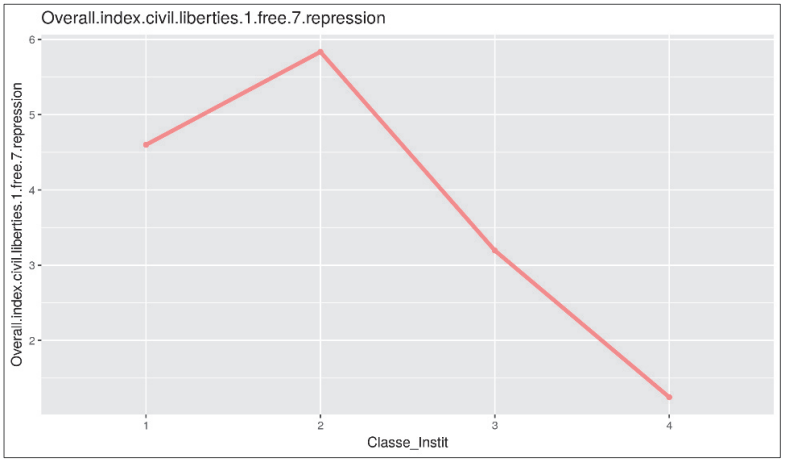
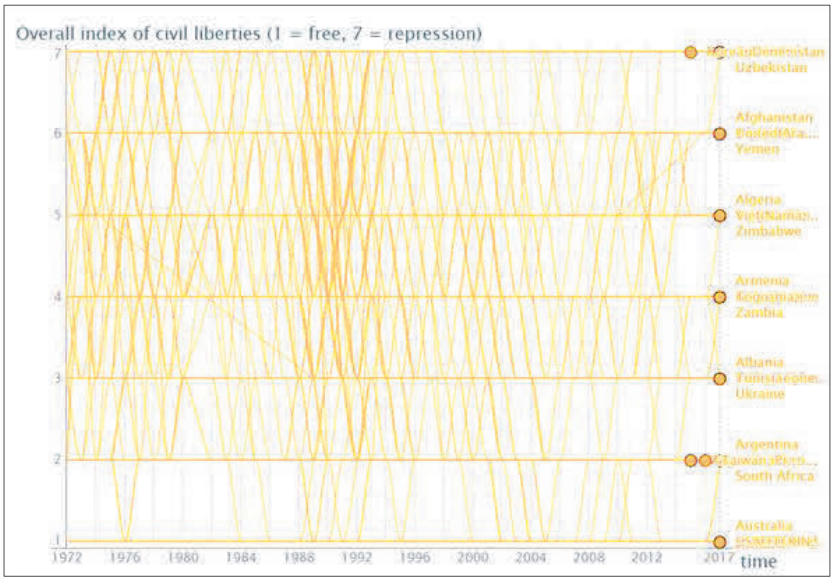


Figure 3.143 Indice global des libertés civiles (1=libre, 7=répression)



Source : Perspective Monde (2018)

Figure 3.144 Indice global de démocratie (Polity IV-Roser). Courbe transversale.

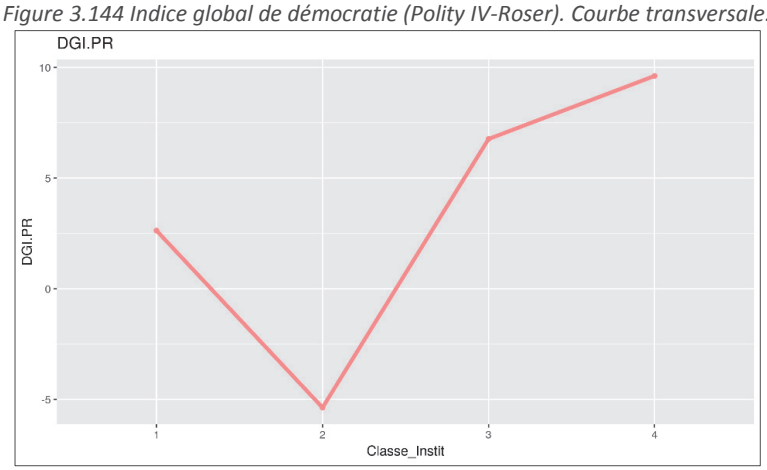
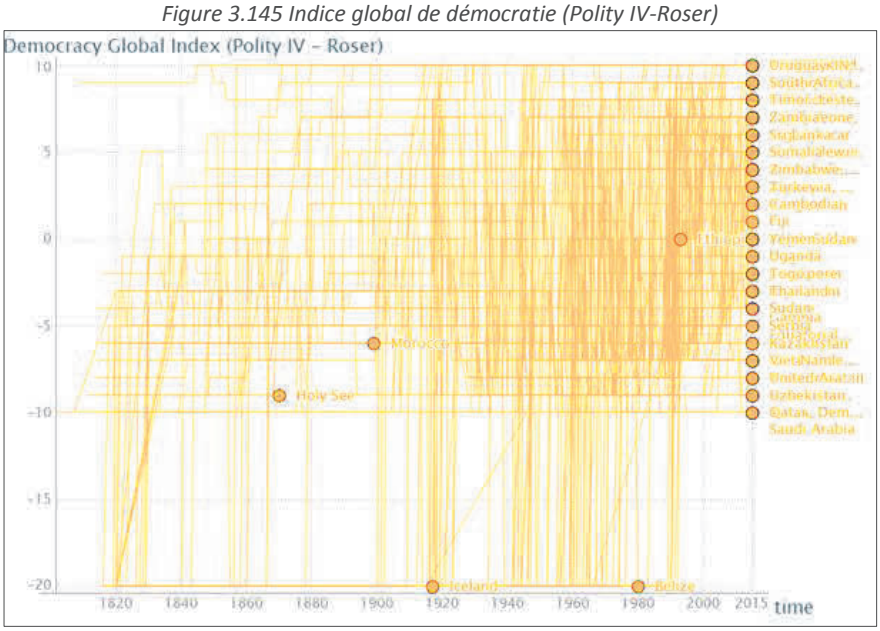
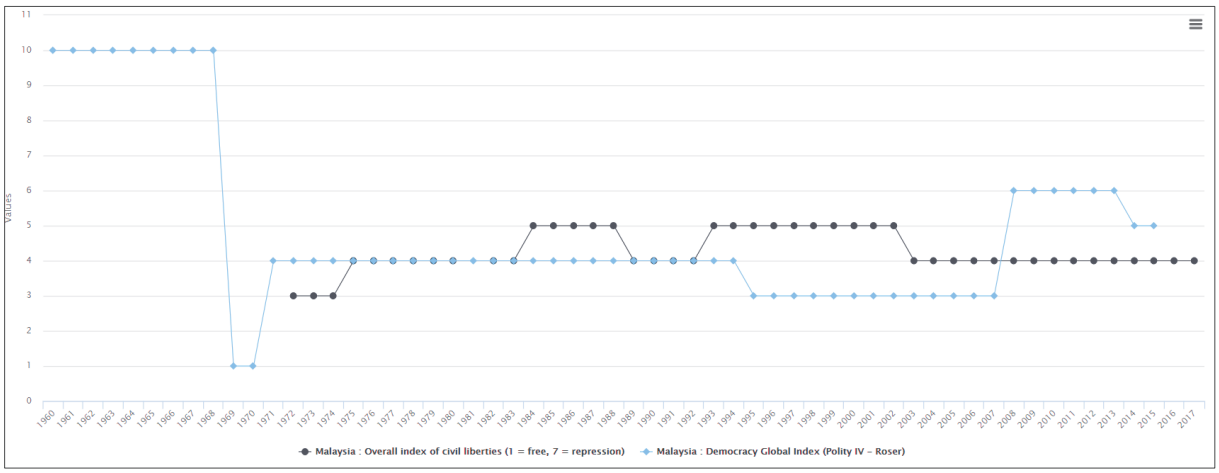


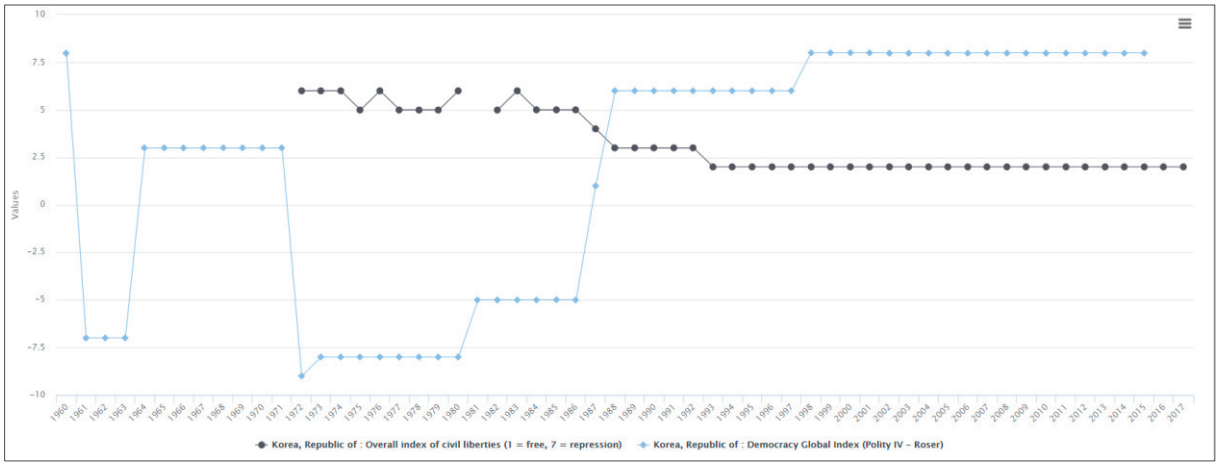
Figure 3.145 Indice global de démocratie (Polity IV-Roser)

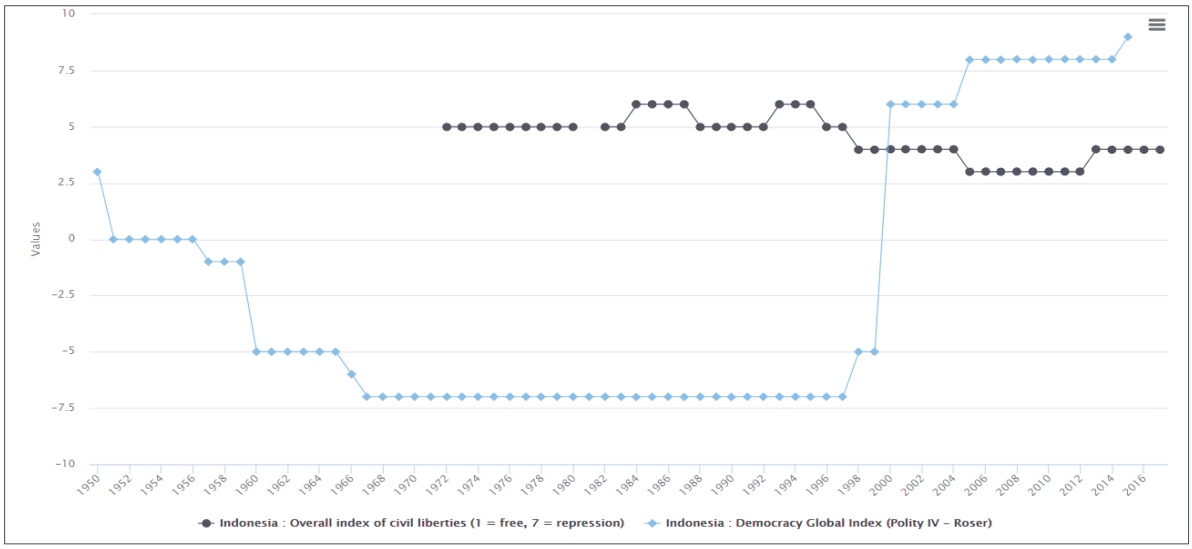


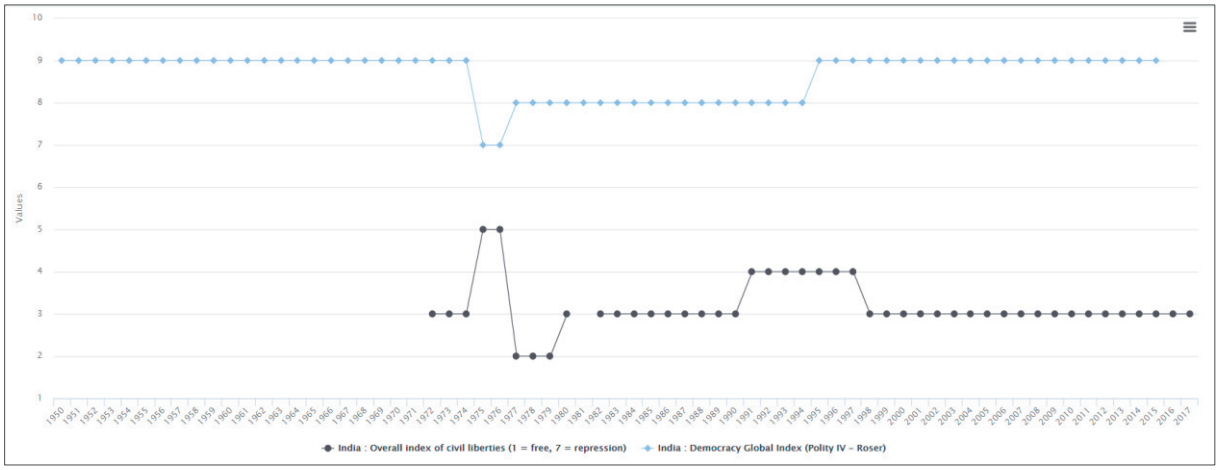
Source : Polity IV & Roser Max (2015)

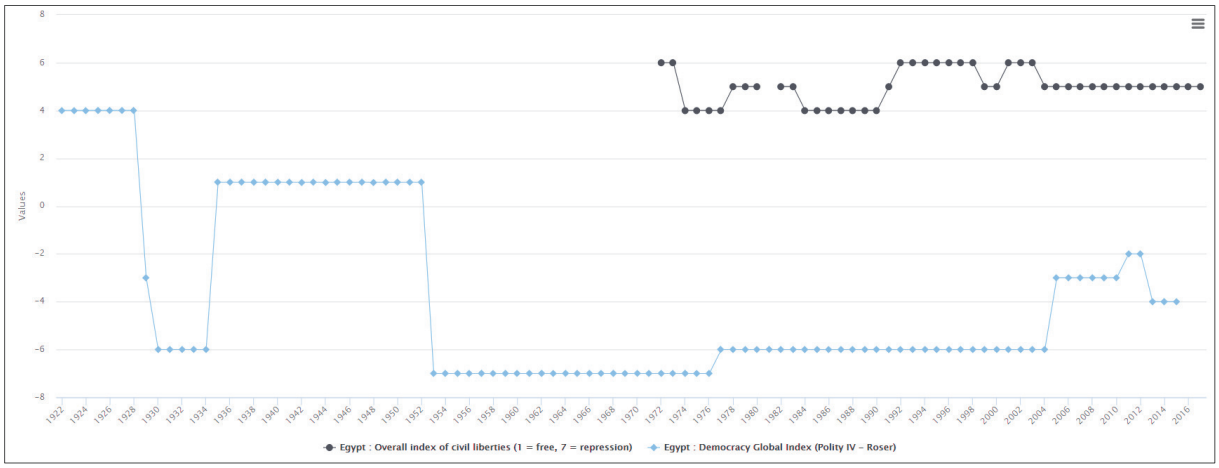
Cependant, en observant l'allure des pays issus des ex-empires coloniaux européens, nous observons des allures de courbes longitudinales similaires aux courbes transversales associées, sur la période postcoloniale (où commence réellement leur histoire institutionnelle moderne de pays indépendant). Notamment une période de hausse des répressions, suivie par une longue période de baisse des répressions. Nous y observons, une dégradation démocratique juste après l'indépendance du pays, et une remontée continue après un certain temps de stagnation.
Voici ci-après quelques exemples d'ex-colonies sélectionnées sur diverses zones géographiques :

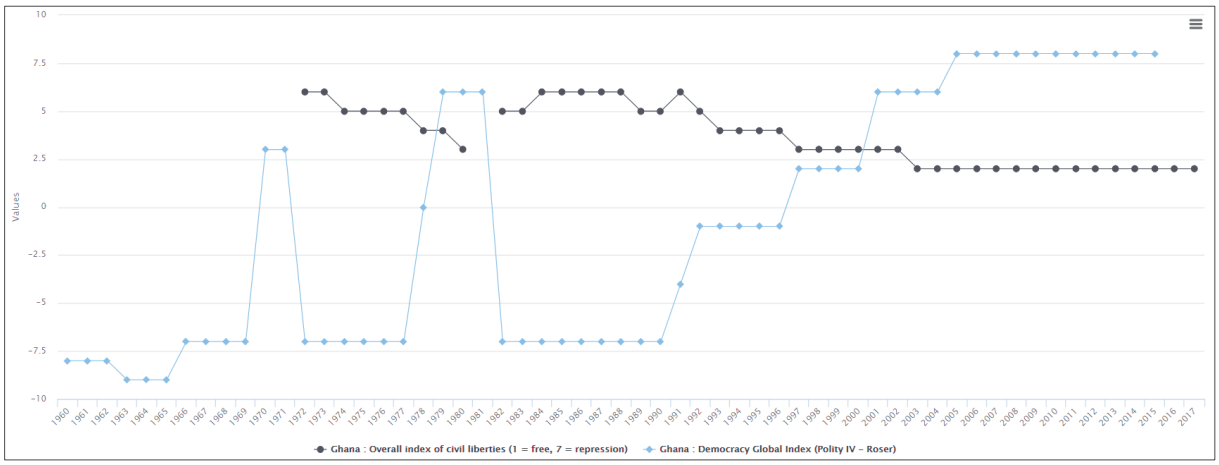


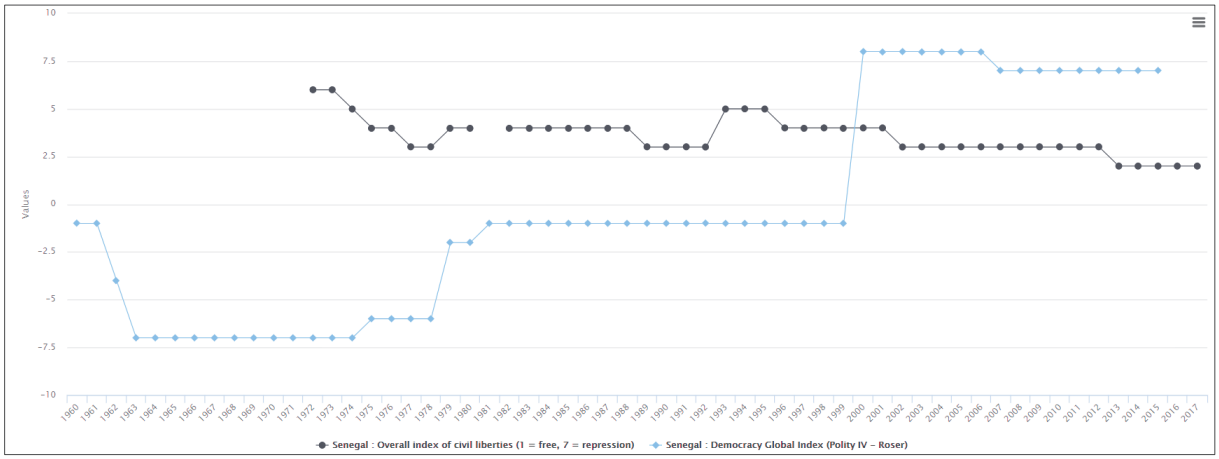





Values
India : Overall index of civil liberties (1 = free, 7 = repression)
India : Democracy Global Index (Polity IV – Roser)

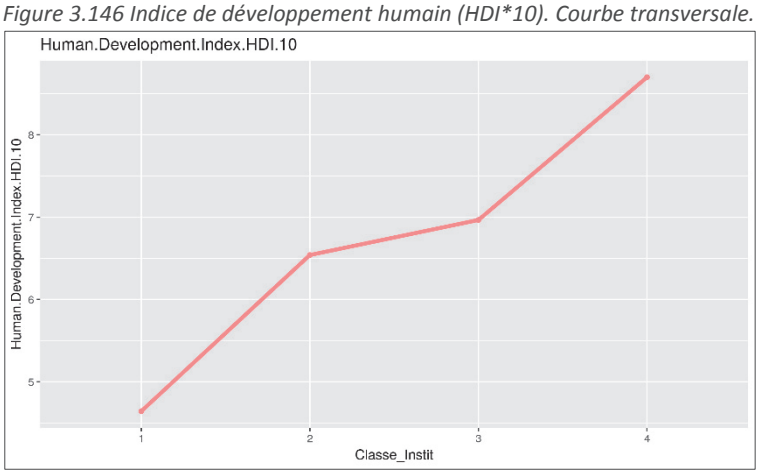

Values
Egypt : Overall index of civil liberties (1 = free, 7 = repression)
Egypt : Democracy Global Index (Polity IV – Roser)


Values
Ghana : Overall index of civil liberties (1 = free, 7 = repression)
Ghana : Democracy Global Index (Polity IV – Roser)



Quant à l'indice de développement humain, la comparaison empirique entre la courbe transversale transitionnelle, et les courbe longitudinales (cf. *Figure 3.128* à *Figure 3.131*), nous permet de conclure à des allures haussières similaires.

*Figure 3.146 Indice de développement humain (HDI*10). Courbe transversale.*



Source : Perspective Monde (2018)

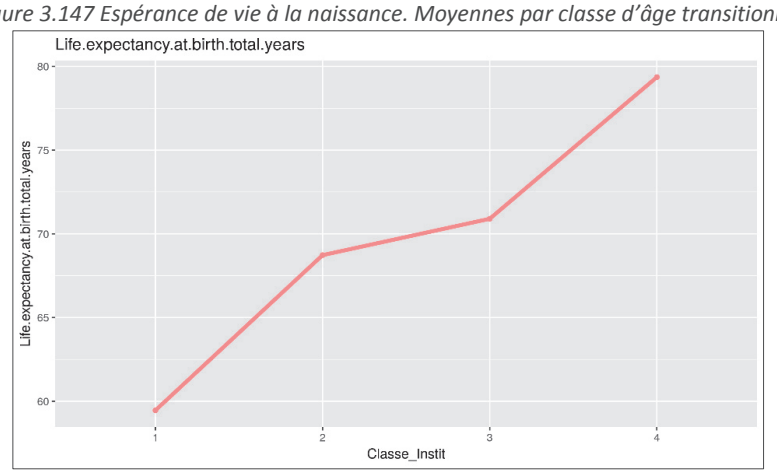
7.1.10. Variables passives : Comparaison des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales

Les variables passives sont des variables utilisées pour illustrer les classes, et qui ne sont pas des variables actives (qui elles sont utilisées pour créer les classes).

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales des variables passives, sont nécessaires pour vérifier la robustesse et la cohérence d'ordonnancement des classes d'âge transitionnelles construites.
La similitude globale des allures des courbes transversales et longitudinales que nous constatons dans les comparaisons ci-après dans plusieurs domaines, montre empiriquement cette robustesse.

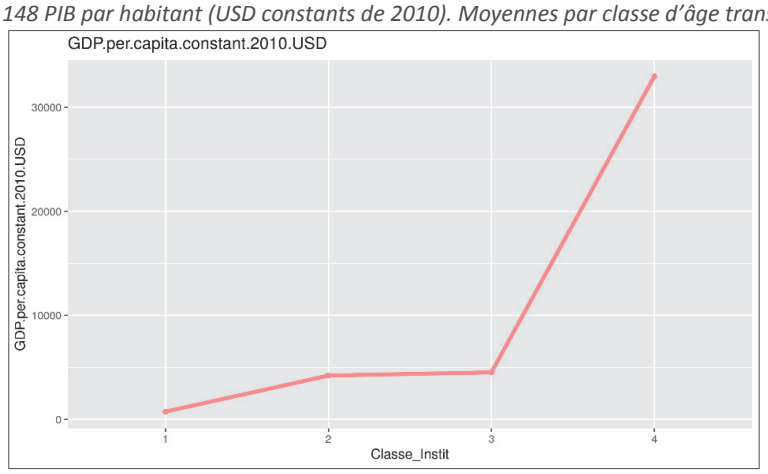
Moyennes par classe d'âge transitionnel

Figure 3.147 Espérance de vie à la naissance. Moyennes par classe d'âge transitionnel



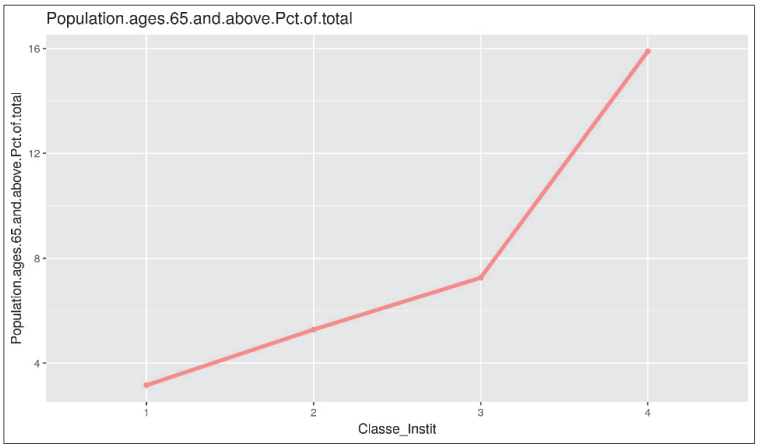
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.2*) nous permet de conclure à des allures haussières similaires.

Figure 3.148 PIB par habitant (USD constants de 2010). Moyennes par classe d'âge transitionnel



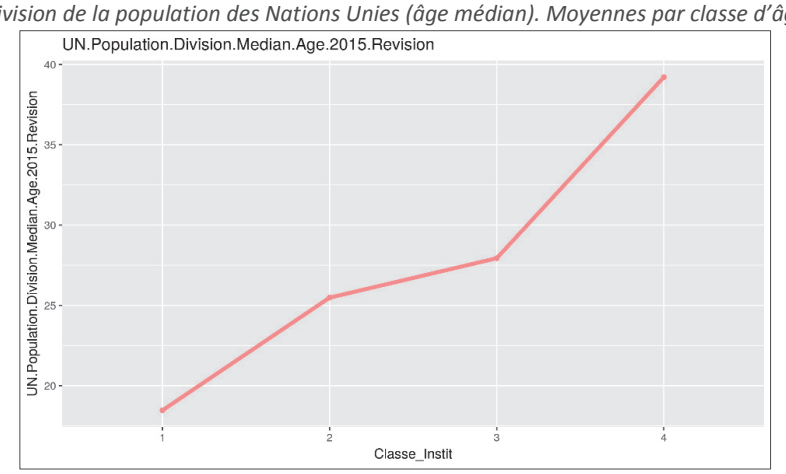
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.3*) nous permet de conclure à des allures haussières similaires.

Figure 3.149 Population âgée de 65 ans et plus. Moyennes par classe d'âge transitionnel



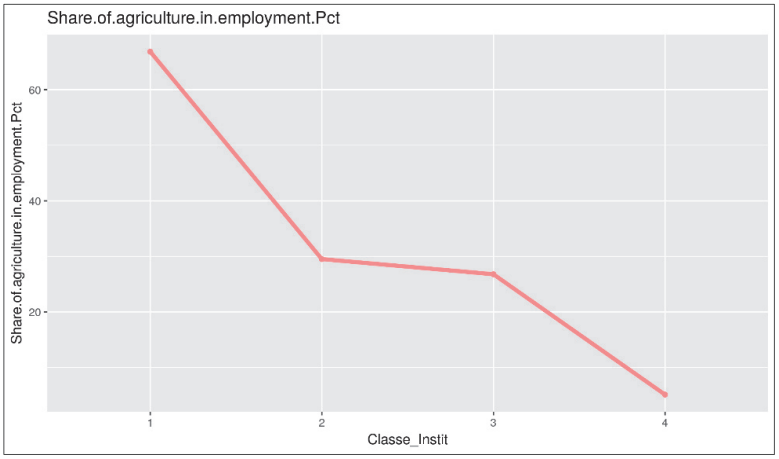
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.4*) nous permet de conclure à des allures haussières similaires.

Figure 3.150 Division de la population des Nations Unies (âge médian). Moyennes par classe d'âge transitionnel



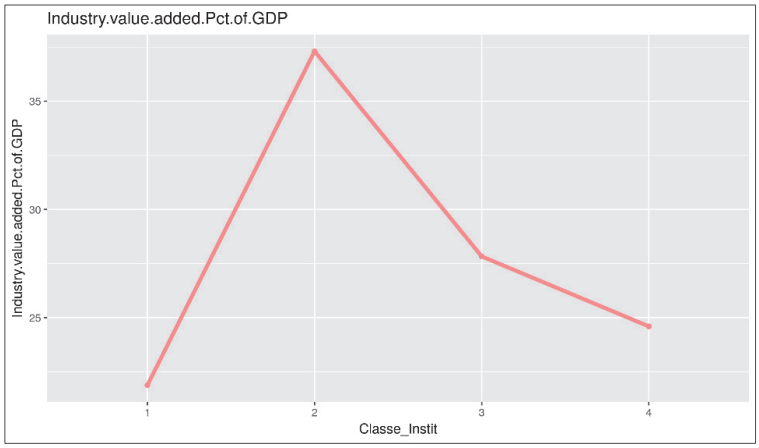
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.5*) nous permet de conclure à des allures haussières similaires.

Figure 3.151 La part de l'agriculture dans l'emploi (% de l'emploi total), historique long). Moyenne par classe d'âge transitionnel



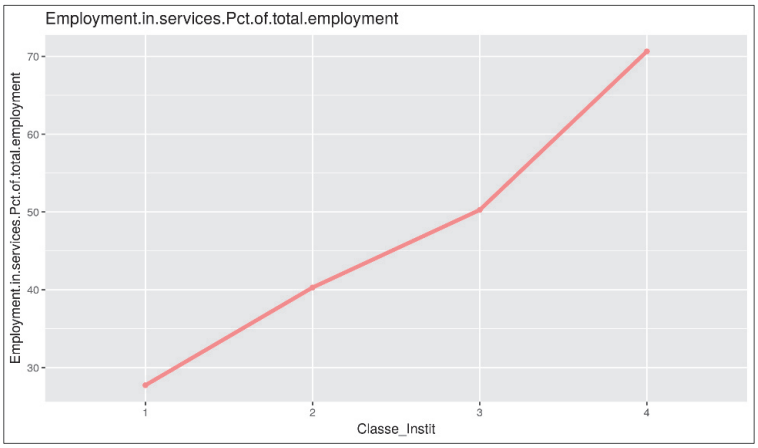
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.6)* nous permet de conclure à des allures baissières similaires.

Figure 3.152 Industrie, valeur ajoutée (% du PIB)



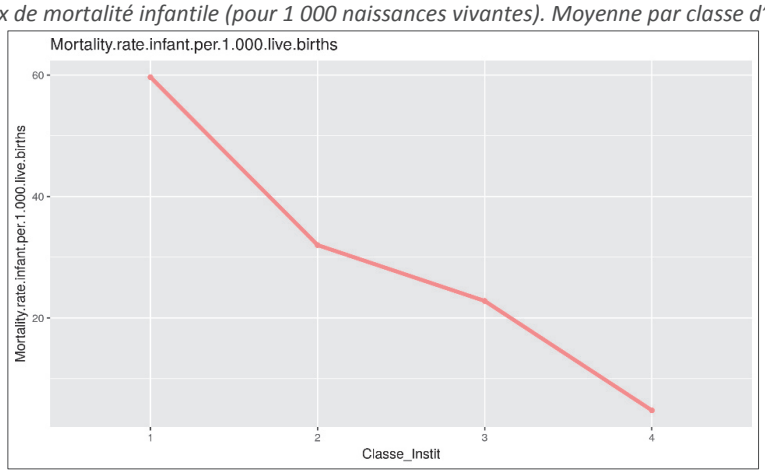
Nous ne disposons pas d'historique longue (au-delà de 1960) sur l'emploi industriel, mais l'allure en cloche de la courbe transversale rappelle celle de la courbe théorique de long terme évoquée par Fourastié (1989, p. 148) et présentée en *Figure 3.47*.

Figure 3.153 Emploi dans les services (% de l'emploi total). Moyenne par classe d'âge transitionnel



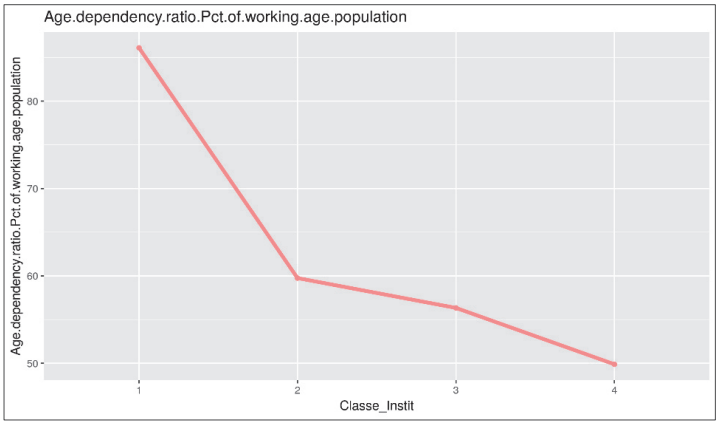
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.28*) nous permet de conclure à des allures haussières similaires.

Figure 3.154 Taux de mortalité infantile (pour 1 000 naissances vivantes). Moyenne par classe d'âge transitionnel



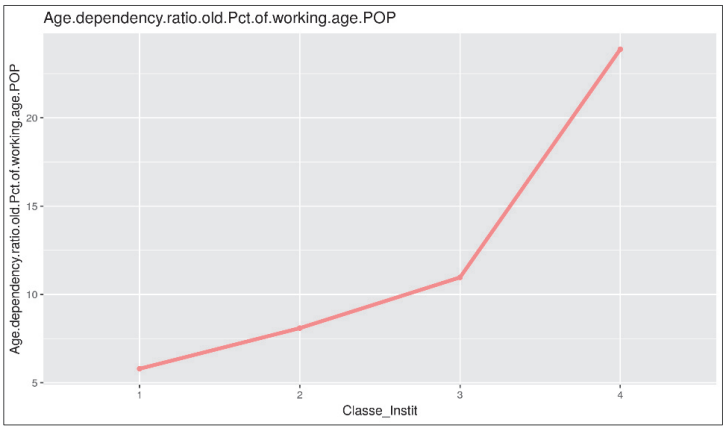
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.15)* nous permet de conclure à des allures baissières similaires.

Figure 3.155 Ratio de dépendance à l'âge (% de la population en âge de travailler). Moyenne par classe d'âge transitionnel



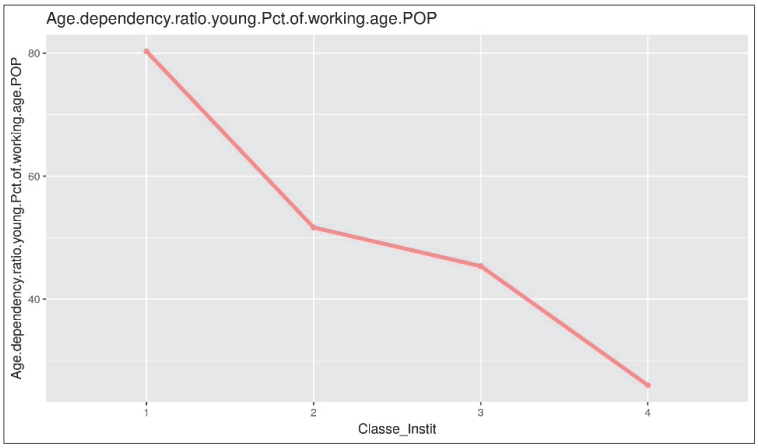
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.10*) nous permet de conclure à des allures baissières similaires.

Figure 3.156 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler). Moyenne par classe d'âge transitionnel



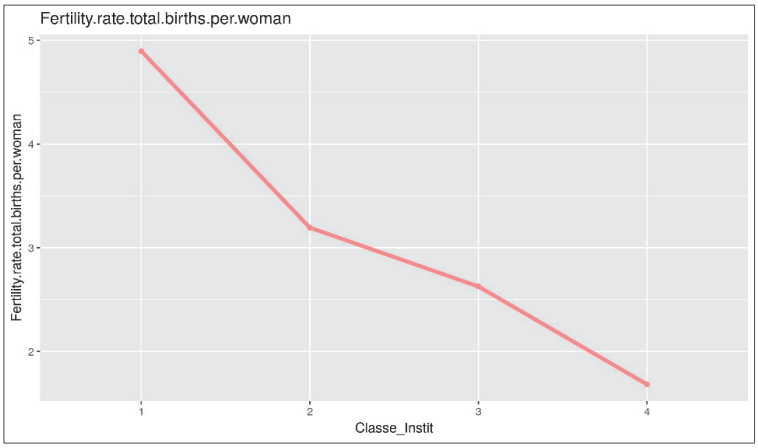
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.11*) nous permet de conclure à des allures haussières similaires.

Figure 3.157 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler). Moyenne par classe d'âge transitionnel



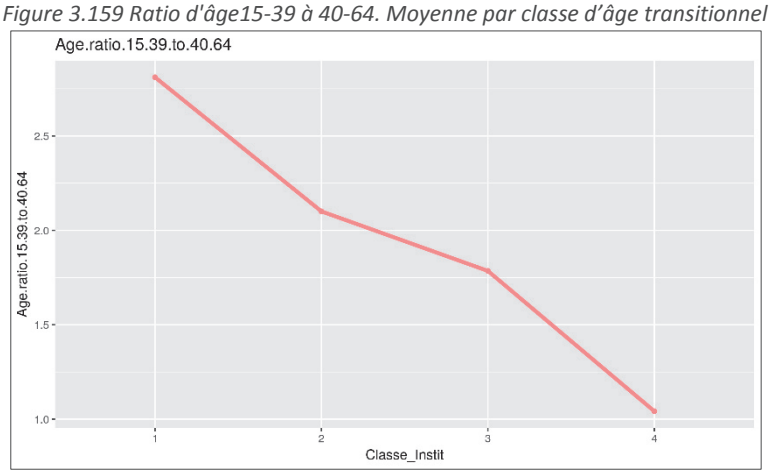
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.12)* nous permet de conclure à des allures baissières similaires.

Figure 3.158 Taux de fécondité, total (naissances par femme). Moyenne par classe d'âge transitionnel



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.16*) nous permet de conclure à des allures baissières similaires.

Figure 3.159 Ratio d'âge15-39 à 40-64. Moyenne par classe d'âge transitionnel



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.13*) nous permet de conclure à des allures baissières similaires.

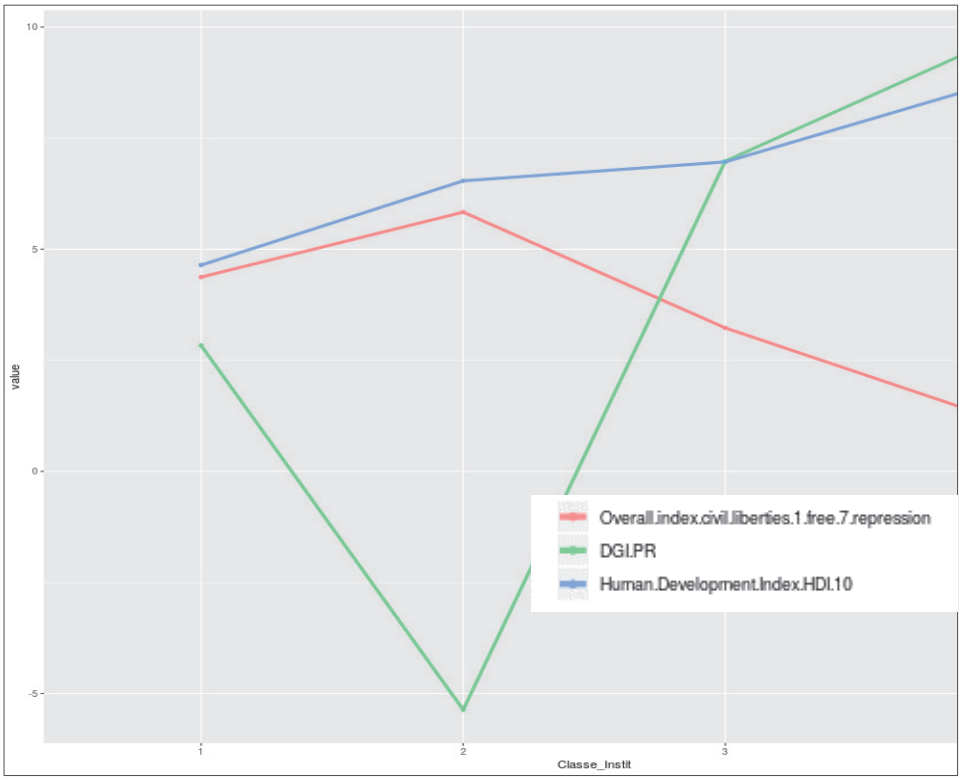
7.2. Peut-on faire l'hypothèse d'une théorie transitionnelle des institutions ?

Nous désignons par transition institutionnelle, une série de profondes et durables transformations institutionnelles dans un pays, pouvant s'étaler sur quelques décennies à plusieurs générations.

Proposition de schéma transitionnel institutionnel selon l'indice d'âge transitionnel

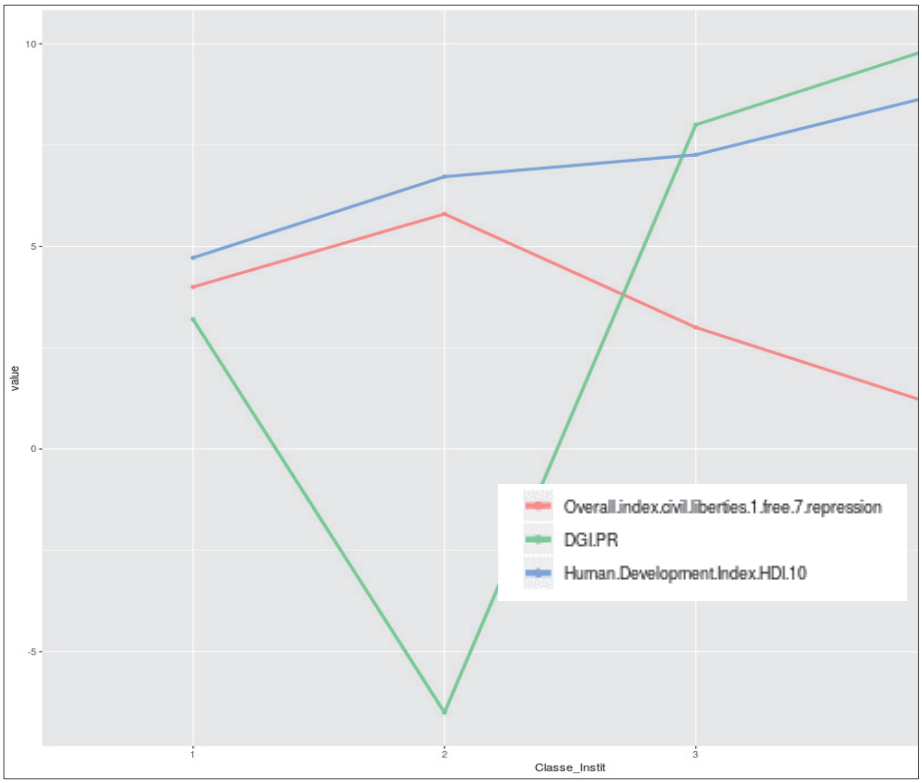
Ayant montré à la section 7.1.9 que les courbes transversales des variables de transition selon l'indice d'âge transitionnel, avaient sensiblement les mêmes allures que leurs courbes longitudinales, nous pouvons proposer un schéma théorique transitionnel.
Par ailleurs, nous avions également montré à la section 7.1.10 que les courbes transversales des variables passives selon l'indice d'âge transitionnel, avaient sensiblement les mêmes allures que leurs courbes longitudinales.
En considérant les arguments ci-dessus énumérés, nous pouvons proposer le schéma théorique transitionnel suivant :

Figure 3.160 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne par classe d'âge transitionnel



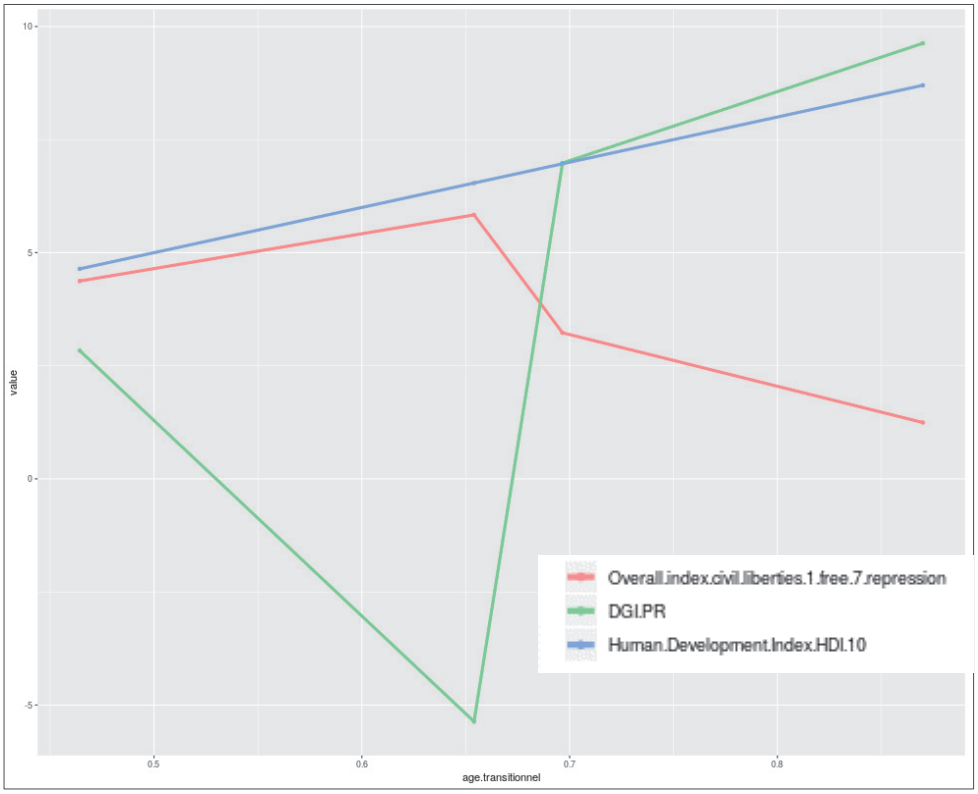
Variable	Signification
Overall.index.civil.liberties.1.free.7.repression	Indice global des libertés civiles (1=libre, 7=répression)
DGI.PR	Indice global de démocratie (Polity IV-Roser)
Human.Development.Index.(HDI*10)	Indice de développement humain (IDH*10)

Figure 3.161 Schéma transitionnel transversal selon l'âge transitionnel. Médiane par classe d'âge transitionnel



L'allure des deux schémas ci-dessus est sensiblement la même, aussi bien en utilisant les moyennes que les médianes : C'est l'illustration d'une robustesse du schéma transitionnel obtenu en coupe transversale.

Figure 3.162 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse



Nous pouvons illustrer le schéma théorique de la transition institutionnelle par le tableau des valeurs moyennes des variables transitionnelles par classes d'âge transitionnel.

Tableau 3.11 Valeurs moyennes des variables transitionnelles par classes d'âge transitionnel

Indice d'âge transitionnel	% de pays en 2014	Indice global des libertés civiles (1 : libre, 7 : répression). Moyenne	Indice global de démocratie (Polity IV-Roser). Moyenne	Indice de développement humain. Moyenne
1	21%	4,60	2,71	0,46
2	16%	5,84	-5,37	0,66
3	35%	3,19	6,62	0,68
4	28%	1,25	9,48	0,87

Ces 4 classes d'âge transitionnel illustrent également les 4 phases observables « actuellement » dans le processus de transition institutionnelle.
La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transition institutionnelle.
Quant aux probables phases futures jamais observées, seules des études prospectivistes permettraient d'en spéculer les caractéristiques.

8. Cognition, éducation et capital humain

Nous allons ici utiliser le temps et l'âge transitionnel pour réaliser une inférence des trajectoires longitudinales dans une coupe transversale dans le domaine cognitif.

8.1. Analyse transversale cognitive

8.1.1. Sélection des variables à étudier

Les variables cognitives peuvent être classées en deux catégories : Les variables mesurant le niveau d'éducation et les variables mesurant les compétences cognitives (souvent à travers des tests internationaux).
Les variables sélectionnées doivent couvrir des pays diversifiés en termes de développement humain (économique, éducatif et sanitaire), afin de couvrir plusieurs états historiques lors d'une coupe transversale.
Les variables sélectionnées répondent donc à cette double contrainte : diversité des catégories cognitives et diversité de niveau de développement humain des pays.

Les variables structurelles utilisées se répartissent en deux catégories : les variables de transition et les variables actives.
Variables de transition : Les variables *transitionnelles* sont utilisées pour décrire le processus de transition cognitive.
Variables actives : Elles sont utilisées pour effectuer la classification statistique des pays. Elles incluent les variables de transition auxquelles peuvent s'ajouter des variables d'homogénéisation des classes. Dans le cas présent, il n'y a pas de variables d'homogénéisation : Les variables actives sont les mêmes que les variables de transition.

Tableau 3.12 Variables actives ou transitionnelles

Variable	Signification en anglais	Signification en français	Variable active	Variable de transition
No.Schooling.Total.Pop.25.64	No Schooling (% of population aged 25-64)	Taux (%) d'adultes de 25-64 ans, non scolarisés	oui	oui
HLA.3rd.Cted.Pct.Pop25.64.Tot.25.64	Highest level attained, Tertiary completed (% of population aged 25-64)	Taux (%) d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint	oui	oui
HLA.1st.Cted.Pct.Pop25.64.Tot.25.64	Highest level attained, Primary completed (% of population aged 25-64)	Taux (%) d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint	oui	oui
HLA.2nd.Cted.Pct.Pop25.64.Tot.25.64	Highest level attained, Secondary completed (% of population aged 25-64)	Taux (%) d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint	oui	oui
Cognitive skills	Cognitive skills	Compétences cognitives	oui	oui

Compétences cognitives numériques-littéraires-logiques sont mesurées par le QI moyen.

8.1.2. Transformation préalable des variables à étudier

Les variables actives utilisées font l'objet des transformations suivantes, avant la sélection de l'année d'étude : estimation des valeurs manquantes par interpolation spline, puis Moyenne mobile d'ordre 5.

8.1.3. Sélection de l'année à étudier

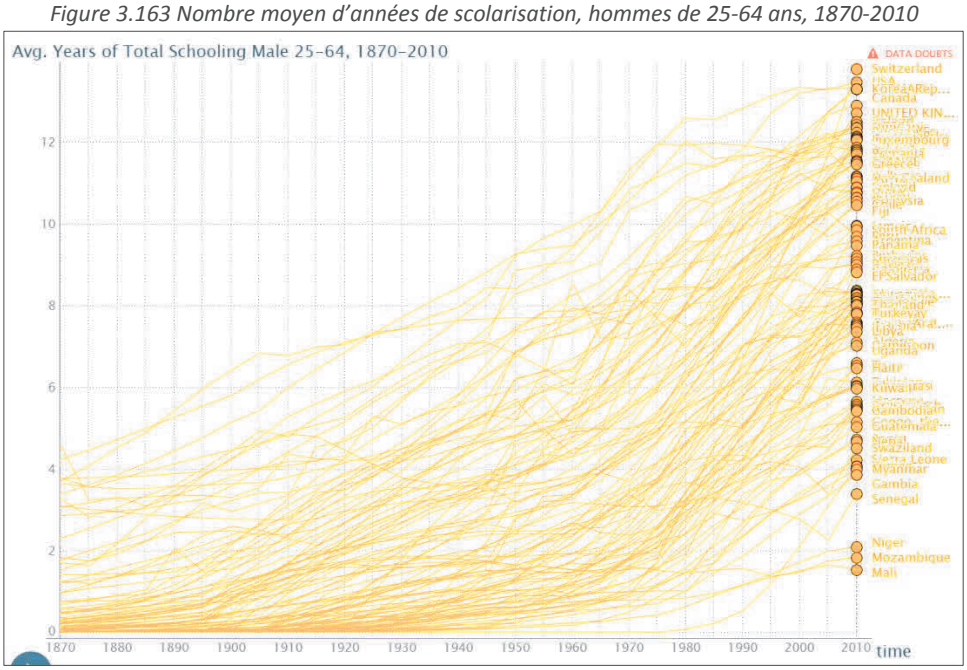
Pour l'analyse transversale, l'année étudiée est 2012. Cela pour des raisons de complétude et de fiabilité des données, car les dernières (2017, source Banque Mondiale) sont souvent complétées et révisées après publication.

8.1.4. Choix du temps transitionnel

Variable à tester comme temps transitionnel

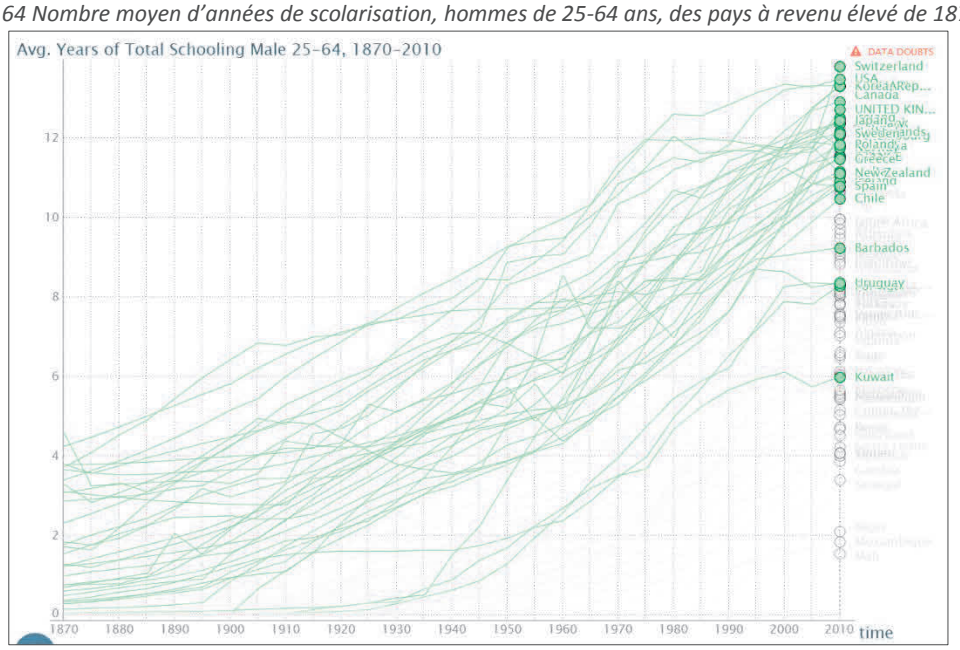
La variable à tester est une valeur lissée (moyenne mobile arrière d'ordre 5) du nombre moyen d'années de scolarisation (*average years of total schooling*) des hommes de 25-64 ans.

Figure 3.163 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, 1870-2010



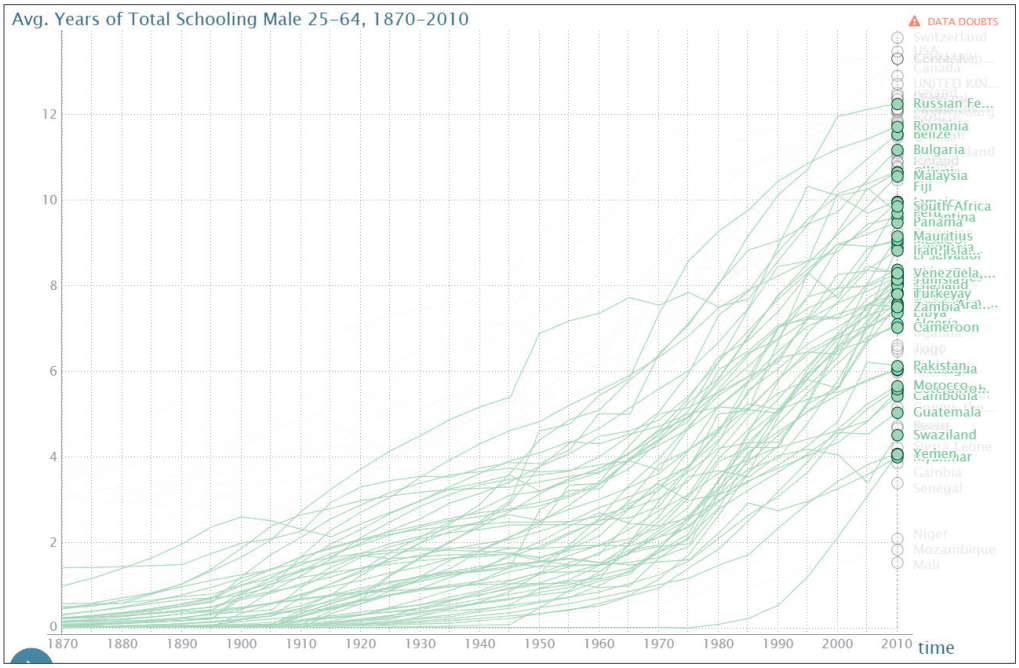
Source : Barro-Lee (2015)

Figure 3.164 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, des pays à revenu élevé de 1870 à 2010



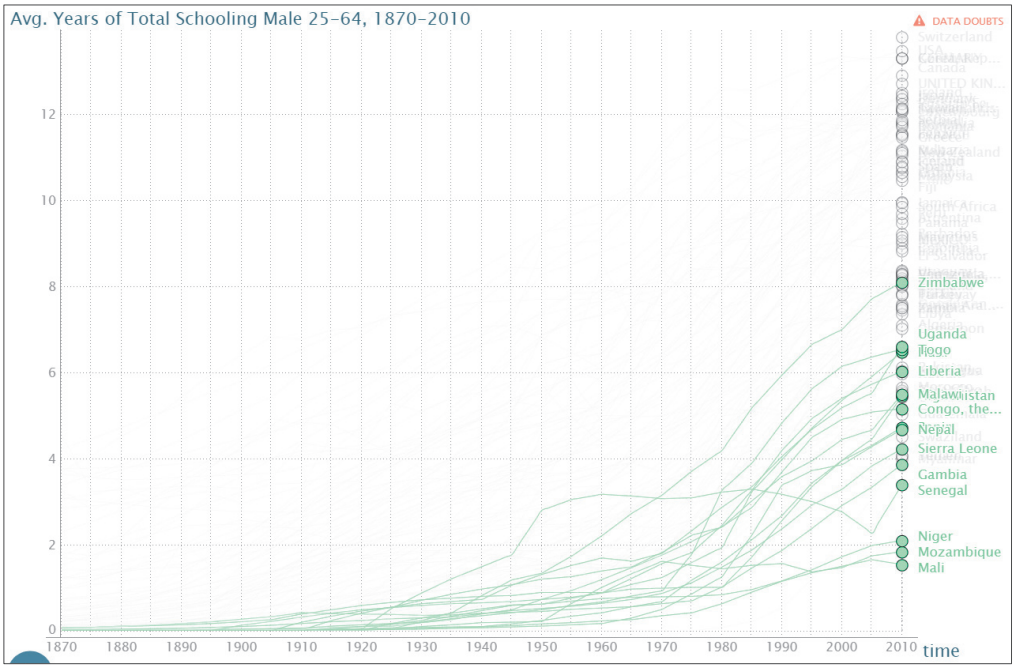
Source : Barro-Lee (2015)

Figure 3.165 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, (non-lissé) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) de 1870 à 2010



Source : Barro-Lee (2015)

Figure 3.166 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, (non lissé) des pays à faible revenu par habitant (au sens de la Banque Mondiale) de 1870 à 2010



Source : Barro-Lee (2015)

Visuellement la tendance mondiale est à la hausse, même si on ne peut parler de monotonie mathématique stricte : raison pour laquelle nous utilisons un lissage par moyenne mobile pour améliorer la monotonie de la variable pressentie pour jouer le rôle de temps transitionnel.

Vérification de la monotonie empirique de la variable à tester

La variable à tester Y est empiriquement monotone lorsque la série différenciée dY (dY(t)=Y(t) – Y(t-1)) est de signe constant sur la période étudiée, pour tout pays.
De plus pour assurer l'universalité de cette monotonie empirique, le signe de série différenciée doit être le même pour tous les pays étudiés
Pour chaque pays, nous calculons le pourcentage de valeurs positives de la série différenciée, afin de vérifier l'unicité du signe sur la période étudiée.
Sur l'ensemble des 91 pays étudiés, le pourcentage moyen de valeurs positives de la série différenciée est de 93.4%, avec un minimum à 38% (Sénégal), une médiane à 98%, et un maximum à 100% (Ghana et France par exemple).
Mieux, 70 des 91 pays, soit 48%, ont 100% de valeurs positives dans la série différenciée.
Par ailleurs, 108 des 91 pays, soit 77%, ont plus de 90% de valeurs positives dans la série différenciée.

De plus, les valeurs centrales (moyenne=0.94 et médiane=0.99) du coefficient de Kendall indiquent une très forte relation monotone décroissante entre le nombre moyen d'années de scolarisation, hommes de 25-64 ans (lissé) et le temps.
Sur les 91 pays, le coefficient de Kendall est systématiquement positif, sauf pour 1 pays.

Vérification empirique de l'unicité du signe de la pente de la variable testée, quel que soit le pays

A défaut d'une monotonie empirique parfaite, l'unicité pour tous les pays, du signe de la pente de la droite d'ajustement de la variable à tester, montre une universalité de la tendance (à la hausse ou à la baisse) sur l'ensemble des pays étudiés.
De plus, si pour un pays, pente de la droite d'ajustement de la variable à tester est proche en valeur absolue de 1, alors cette variable à tester se rapproche d'autant plus de la variable temps.

Nous régressions linéairement la variable à tester par rapport au temps. Le signe du coefficient du temps est positif, lorsque la variable à tester a une tendance linéaire croissante dans le temps.
Ci-dessous les statistiques sommaires issues des 91 régressions (91 pays).

Summary(coef.Intercept)

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
-0.13449 -0.13346 -0.13272 -0.12799 -0.13036  0.07286       42
```

Summary(coef.year)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-0.5836  1.0442  1.0631  1.0252  1.0690  1.0774      42
```

Summary(coef.R2)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 0.1831  0.9388  0.9732  0.9348  0.9840  0.9994      42
```

Pour 90 régressions (sur 91), le coefficient du temps (coef.year), est systématiquement positif, confirmant ainsi l'hypothèse d'une hausse tendancielle du nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé), pour l'ensemble des pays.
L'exception est le Sénégal qui a un coefficient négatif, et même dans ce cas la pente est nettement positive dans les 7 dernières années de l'étude, comme le montre le graphique ci-après :

Figure 3.167 Exemple du Sénégal. Nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé)

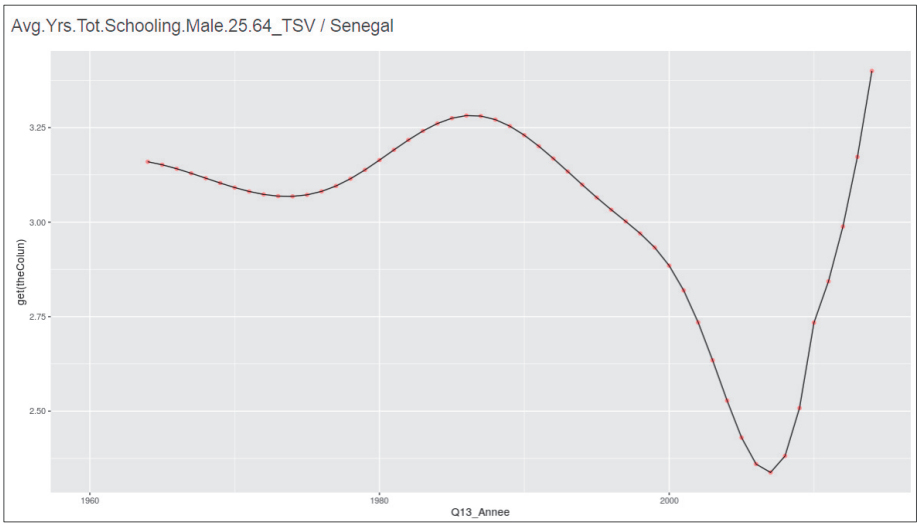


Figure 3.168 Exemple de la France. Nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé)

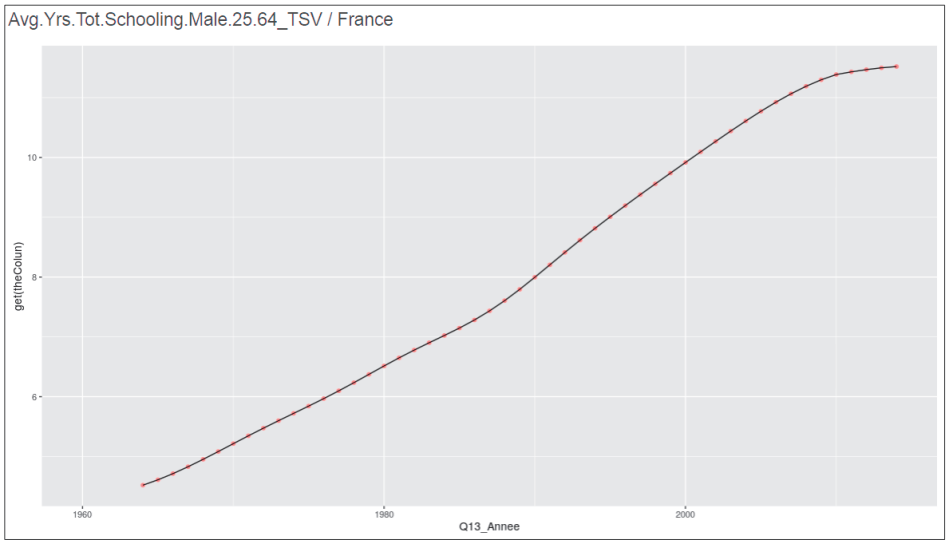
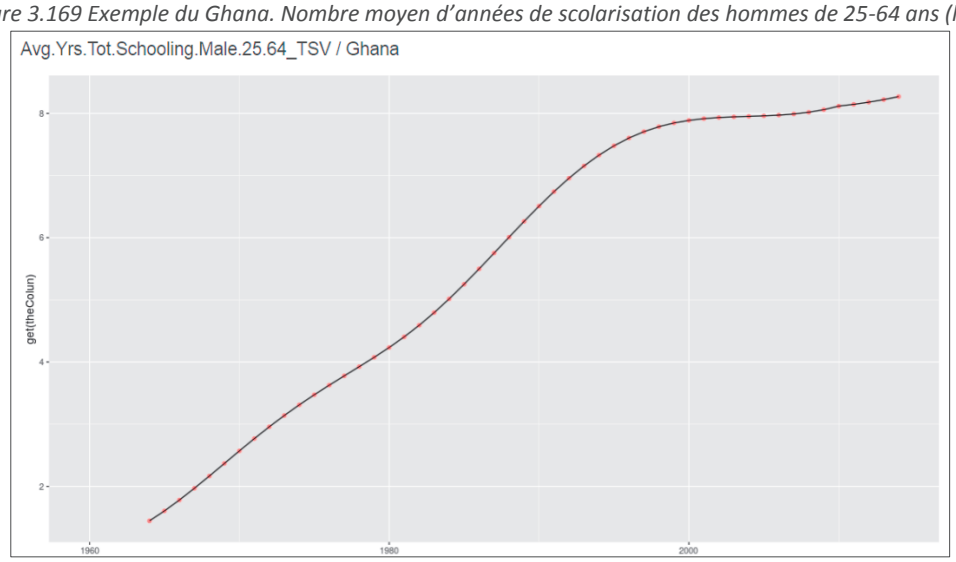


Figure 3.169 Exemple du Ghana. Nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé)



Conclusion

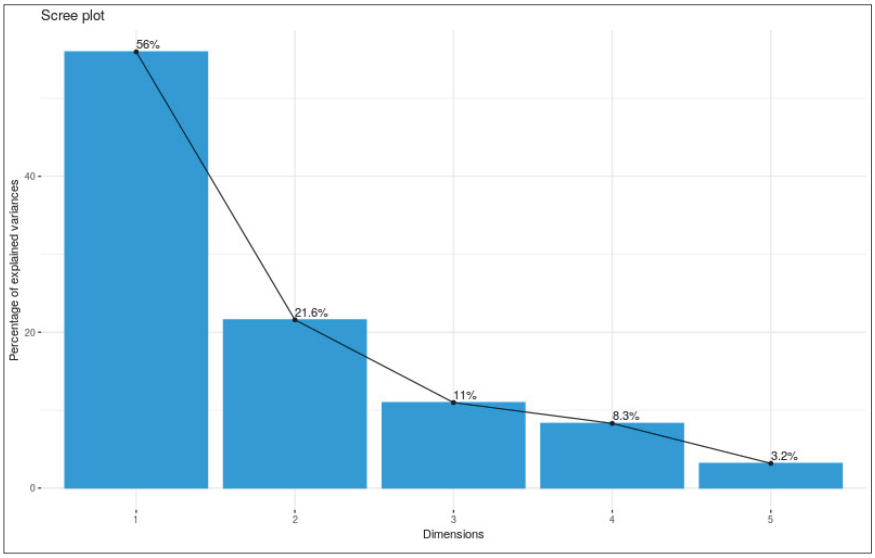
En conclusion, le *nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé)* est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale, du fait de sa tendance haussière généralisée depuis au moins 2 siècles.
Notons également que la pente moyenne du taux de mortalité lissé est de +1.03 (médiane de +1.06), ce qui est assez proche en valeur absolue de 1 : ce qui renforce son rôle potentiel de temps transitionnel.

8.1.5. Classification des pays, sur composantes principales

Sélection des composantes principales pour la classification

L'analyse en composantes principales (ACP) est réalisée à partir des 5 variables actives (cf. Tableau 3.12).

Figure 3.170 Pourcentage de variance expliquée par les axes principaux (ou dimensions)



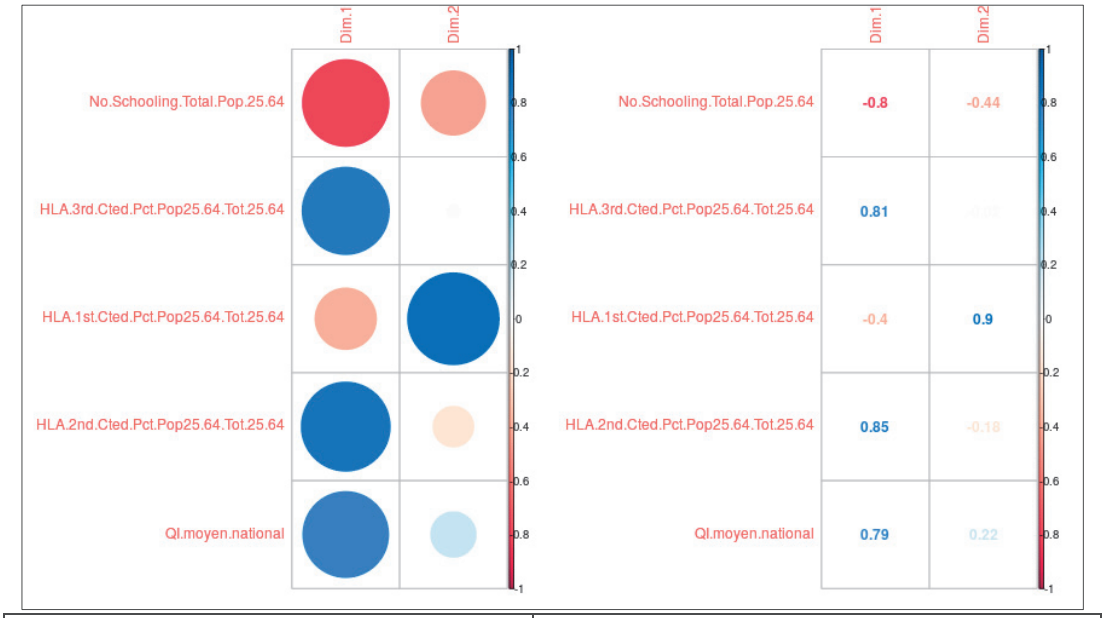
Les 2 axes principaux expriment 77.6% de l'inertie totale du jeu de données ; cela signifie que 77.6% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan.

Signification des composantes principales sélectionnées

En analysant les Figure 3.171 et Figure 3.172, nous déduisons que la première composante (Dim 1) oppose les pays ayant un fort taux d'analphabétisme des adultes de 25-64 ans, et un fort taux d'adultes (25-64 ans) dont l'éducation primaire constitue le niveau le plus élevé atteint, aux pays ayant un fort taux d'adultes (25-64 ans) dont l'éducation secondaire ou tertiaire constitue le niveau le plus élevé atteint, et présentant un QI élevé.
La première composante (Dim 1) a une corrélation positive avec les taux d'adultes (25-64 ans) dont l'éducation secondaire ou tertiaire constitue le niveau le plus élevé atteint, et le QI.

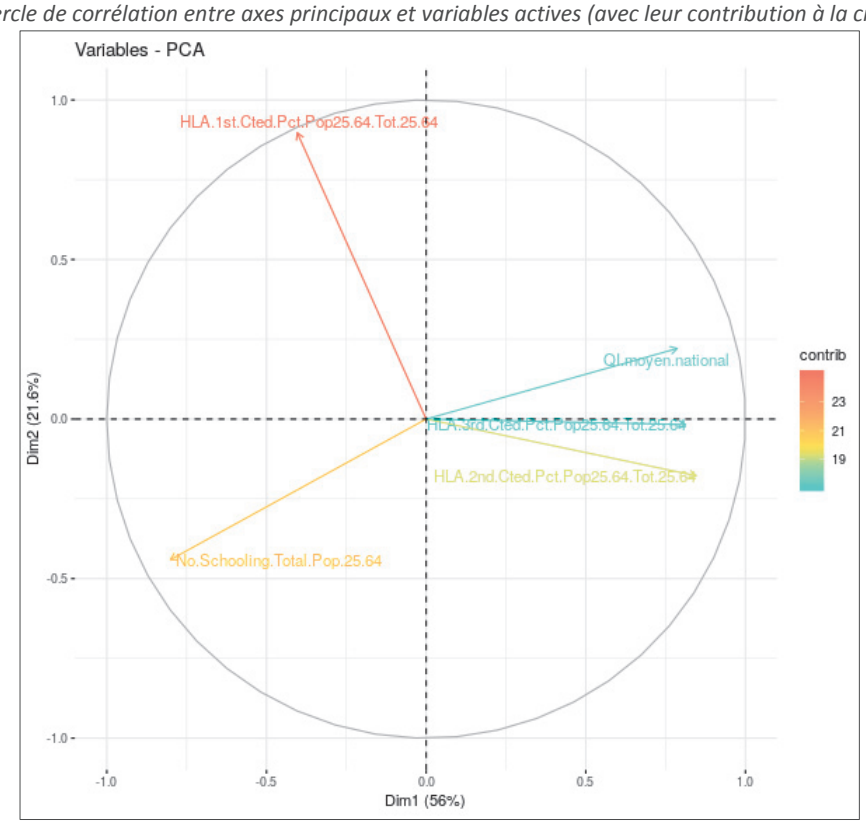
La deuxième composante (Dim 2) oppose les pays ayant un fort taux d'analphabétisme des adultes de 25-64 ans, aux pays ayant un fort taux d'adultes (25-64 ans) dont l'éducation primaire constitue le niveau le plus élevé atteint.
La deuxième composante (Dim 2) a une corrélation positive avec les taux d'adultes (25-64 ans) dont l'éducation primaire constitue le niveau le plus élevé atteint.

Figure 3.171 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives



No.Schooling.Total.Pop25.64	Taux (%) d'adultes de 25-64 ans, non scolarisés
HLA.3rd.Cted.Pct.Pop25.64.Tot25.64	Taux (%) d'adultes de 25-64 ans, dont l'éducation **tertiaire** achevée, constitue le niveau le plus élevé atteint
HLA.1rd.Cted.Pct.Pop25.64.Tot25.64	Taux (%) d'adultes de 25-64 ans, dont l'éducation **primaire** achevée, constitue le niveau le plus élevé atteint
HLA.2rd.Cted.Pct.Pop25.64.Tot25.64	Taux (%) d'adultes de 25-64 ans, dont l'éducation **secondaire** achevée, constitue le niveau le plus élevé atteint
Cognitive skills	Compétences cognitives

Figure 3.172 Cercle de corrélation entre axes principaux et variables actives (avec leur contribution à la création des axes)



Choix du nombre optimal de classes de pays

Après classification statistique des pays à partir des deux premiers axes principaux, nous obtenons quatre classes optimales de pays, qui seront décrites, après ordonnancement par le temps transitionnel qui est le nombre moyen d'années de scolarisation des hommes de 25-64 ans (après lissage sur 5 ans). Cela correspond à un compromis entre la classification en 3 classes, trop grossière, et la classification en 5 classes, trop détaillée, notamment avec deux classes ayant moins de 15 pays.

Figure 3.173 Gain d'inertie en fonction du nombre de classes

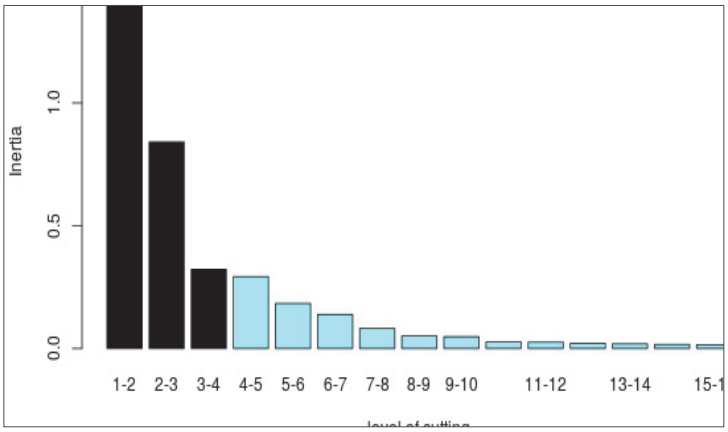
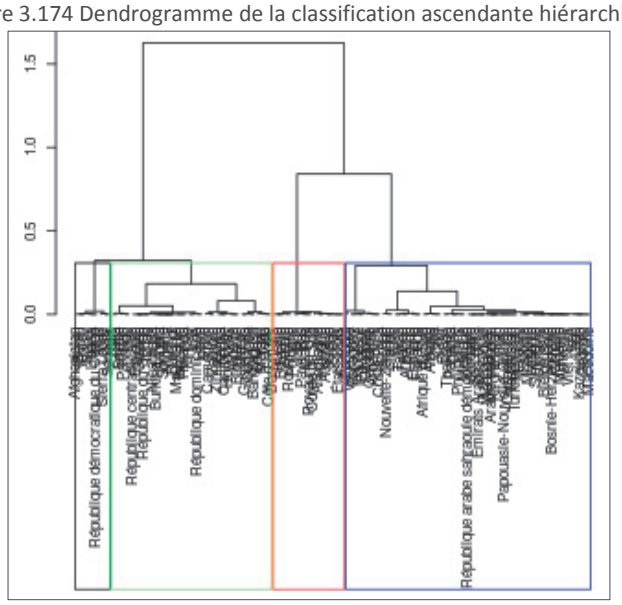


Figure 3.174 Dendrogramme de la classification ascendante hiérarchique



8.1.6. Indexation temporelle des classes de pays

Le temps transitionnel utilisé est le nombre moyen d'années de scolarisation des hommes de 25-64 ans (après lissage sur 5 ans).

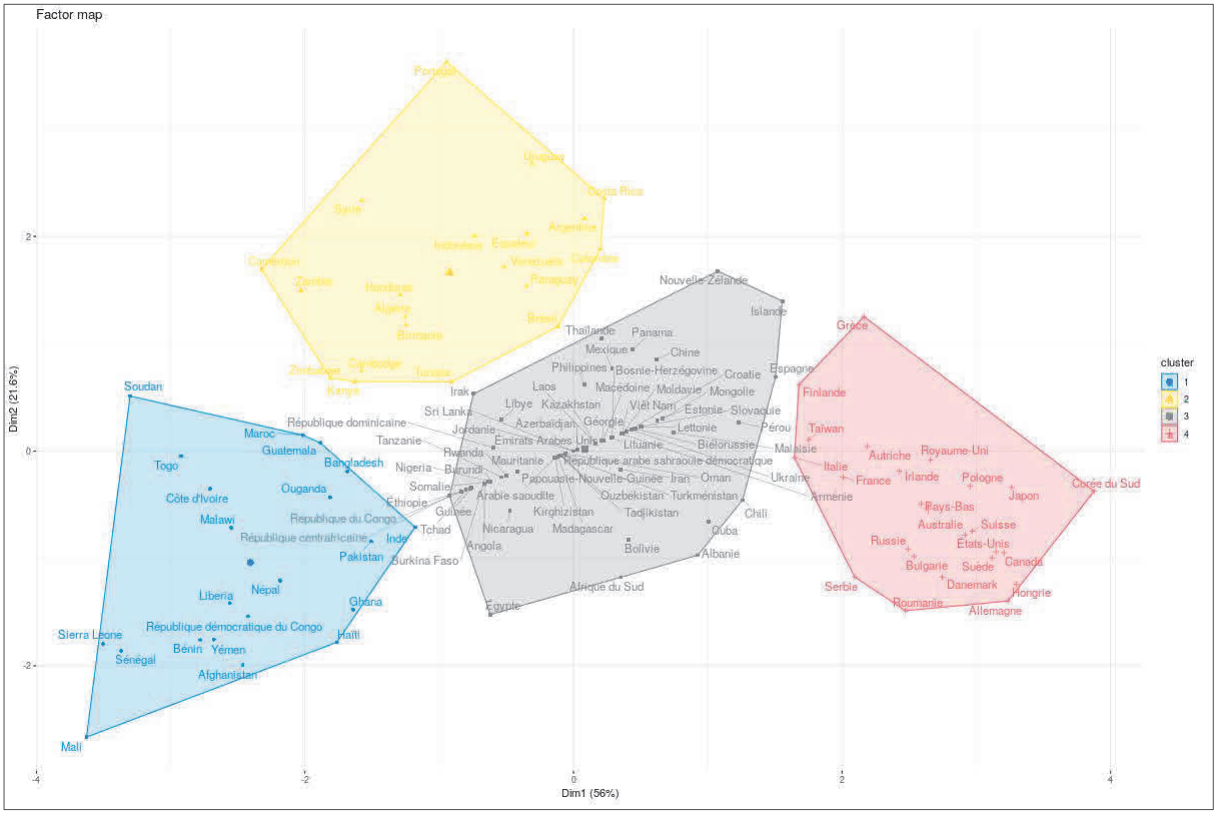
Etant donné que le nombre moyen d'années de scolarisation des hommes de 25-64 ans est croissant dans le temps, les classes de pays sont numérotées par ordre croissant de leur moyenne : ainsi, la classe 4 qui a la moyenne la plus élevée, porte l'index temporel le plus élevé (c'est la classe la plus ancienne dans le processus de transition) ; et inversement pour la classe 1 qui est à l'autre bout du spectre.

Nous qualifions les classes ainsi indicées par l'âge transitionnel moyen de leurs membres, de *classes d'âge transitionnel*.

8.1.7. Description statistique des classes d'âge transitionnel de pays

Cette représentation des classes d'âge transitionnel de pays, suggère un parcours temporel sous forme d'un V inversé en allant de la gauche vers la droite.
Dans cette vue temporelle-transitionnelle, la classe 4, représente les pays plus anciens dans le processus de transformation cognitive. A l'opposé, les pays de la classe 1 sont les pays les plus récents dans le processus de transformation cognitive.
Ces 4 classes illustrent également les 4 phases observables « actuellement » dans le processus de transformation cognitive. La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transformation cognitive.

Figure 3.175 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels



La représentation elliptique ci-dessus suggère que les frontières des classes ne sont pas si tranchées aux frontières que le laisse supposer la représentation en polygones.

Figure 3.176 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes

Tableau 3.13 Description des classes selon les variables actives

Classes	Taux (%) d'adultes de 25-64 ans, non scolarisés	Taux (%) d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint	Taux (%) d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint	Taux (%) d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint	Compétences cognitives
1	48,342536	*NS*	11,073881	2,592075	73,580952
2	*NS*	29,643244	18,095817	5,810229	*NS*
3	*NS*	*NS*	*NS*	*NS*	*NS*
4	1,024819	6,705193	43,720689	20,596263	97,93683
Monde	16,87665	15,53623	25,42666	10,43335	85,18392

NS : valeur non-significativement différente de la moyenne mondiale au seuil de 5%

La **classe 1** est composée de pays tels que : Afghanistan, Mali, Sénégal et Sierra Leone. Ce groupe est caractérisé par :

- de fortes valeurs pour le taux d'analphabétisme des adultes de 25-64 ans (48.3% dans la classe contre 16.9% au niveau mondial)
- de faibles valeurs pour le taux d'adultes de 25-64 ans dont l'éducation secondaire (11.1% dans la classe contre 25.4% au niveau mondial) ou tertiaire (2.6% dans la classe contre 10.4% au niveau mondial) constitue le niveau le plus élevé atteint, et le QI moyen (73.6% dans la classe contre 85.2% au niveau mondial).

La **classe 2** est composée de pays tels que : Argentine, Costa Rica, Portugal, Syrie et Uruguay. Ces pays partagent :

- de fortes valeurs pour le taux d'adultes de 25-64 ans dont l'éducation primaire constitue le niveau le plus élevé atteint (29.6% dans la classe contre 15.5% au niveau mondial)

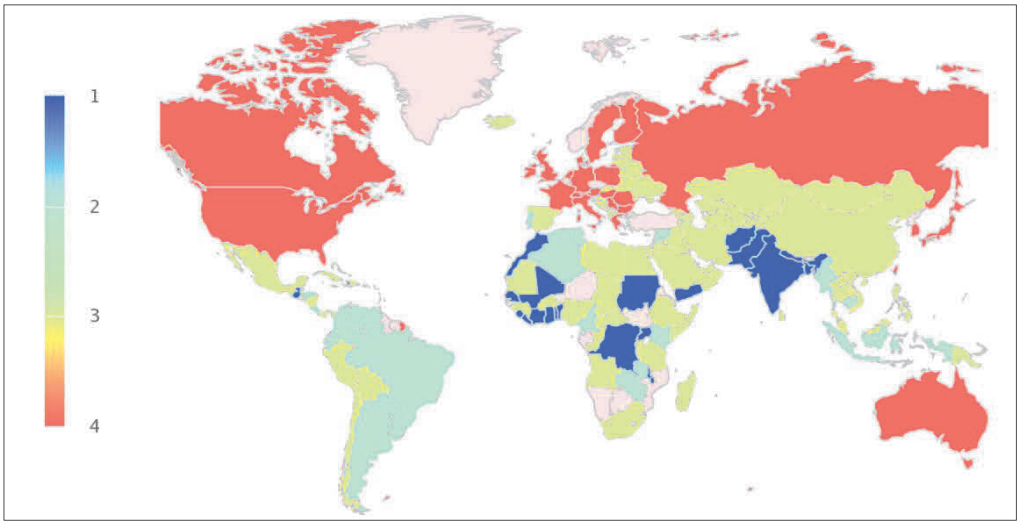
- de faibles valeurs pour le taux d'adultes de 25-64 ans dont l'éducation secondaire (18.1% dans la classe contre 25.4% au niveau mondial) ou tertiaire (5.8% dans la classe contre 10.4% au niveau mondial) constitue le niveau le plus élevé atteint.

La **classe 3** est composée de pays tels que : Burkina Faso, Croatie, Iran, Mexique, Arménie, Azerbaïdjan, Bolivie, Chine. Ce groupe est caractérisé par des variables cognitives dont les valeurs ne diffèrent pas significativement de la moyenne

La **classe 4** est composée de pays tels que : Allemagne, Hongrie et Corée du Sud. Ce groupe est caractérisé par :

- de fortes valeurs pour le taux d'adultes de 25-64 ans dont l'éducation secondaire (43.7% dans la classe contre 25.4% au niveau mondial) ou tertiaire (20.6% dans la classe contre 10.4% au niveau mondial) constitue le niveau le plus élevé atteint, et le QI moyen (97,9% dans la classe contre 85.2% au niveau mondial).
- de faibles valeurs pour le taux d'analphabétisme des adultes de 25-64 ans (1% dans la classe contre 16.9% au niveau mondial) et le taux d'adultes de 25-64 ans dont l'éducation primaire constitue le niveau le plus élevé atteint (6.7% dans la classe contre 15.5% au niveau mondial)

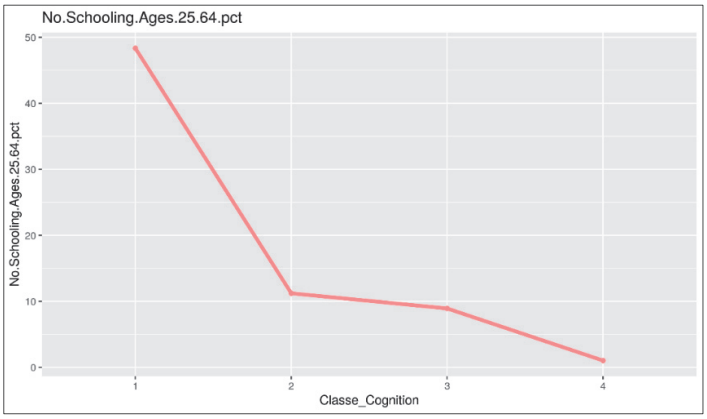
Figure 3.177 Cartographie des Classes temporelles de Pays



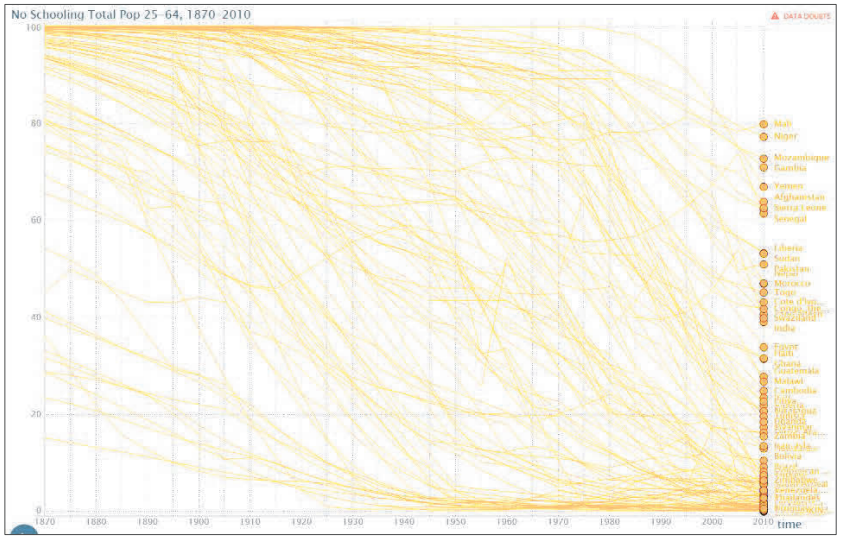
8.1.8. Variables de transition : Comparaison des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales, sont nécessaires pour envisager la construction d'un schéma transitionnel à partir de l'analyse transversale.

Figure 3.178 Population totale non scolarisée de 25 à 64 ans

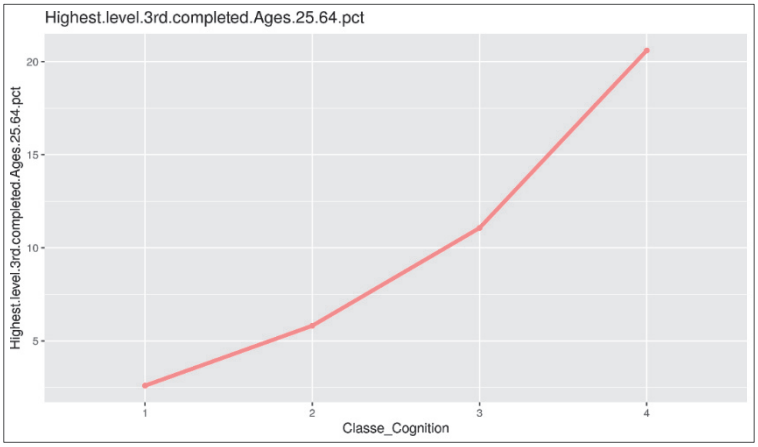


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures baissières similaires.

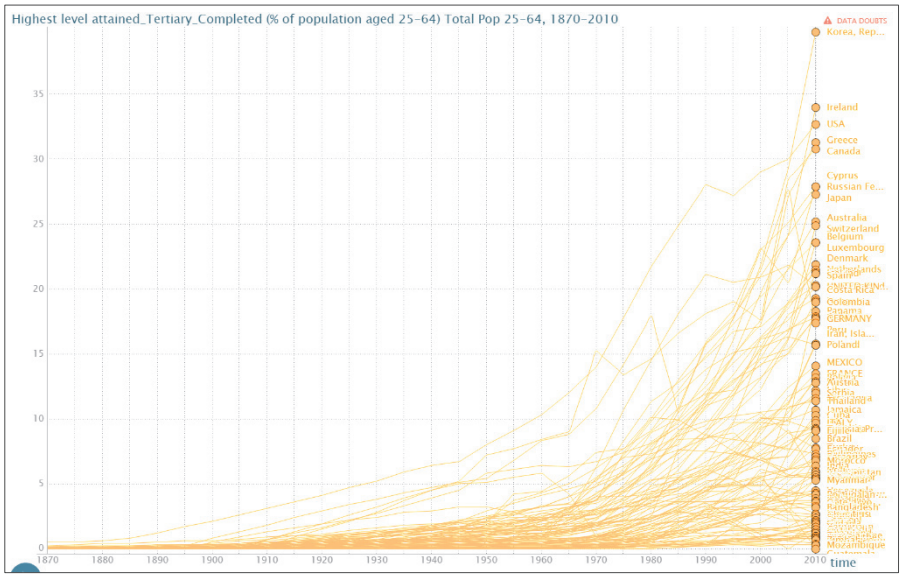


Source : Barro-Lee (2015)

Figure 3.179 Taux (%) d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint

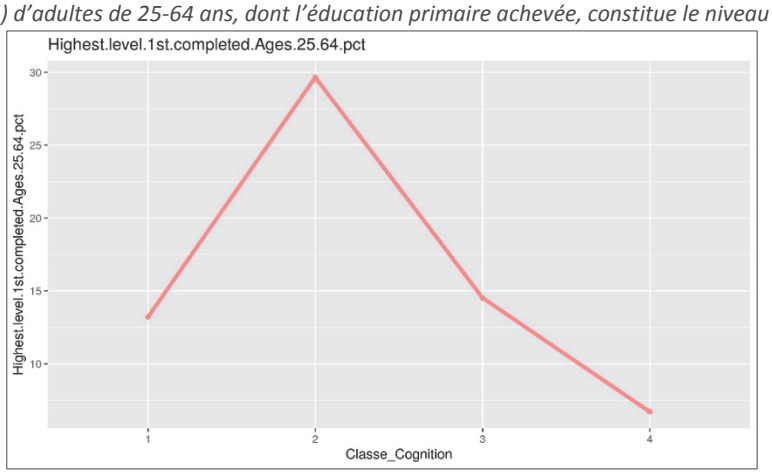


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.

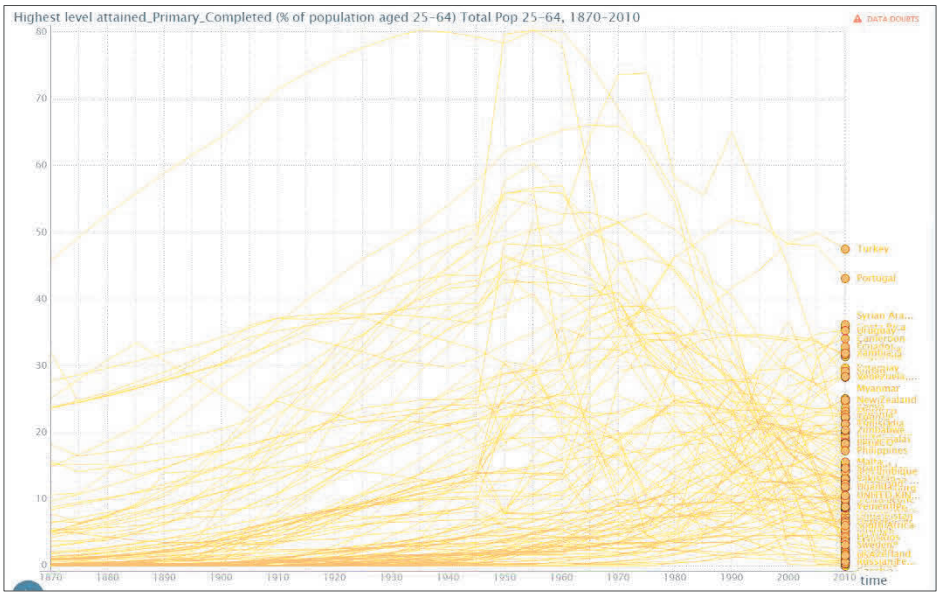


Source : Barro-Lee (2015)

Figure 3.180 Taux (%) d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint

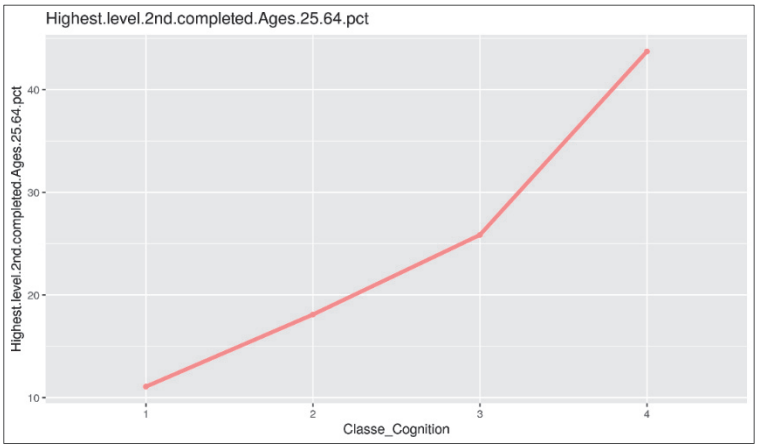


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures en V inversée (ou cloche) similaires.

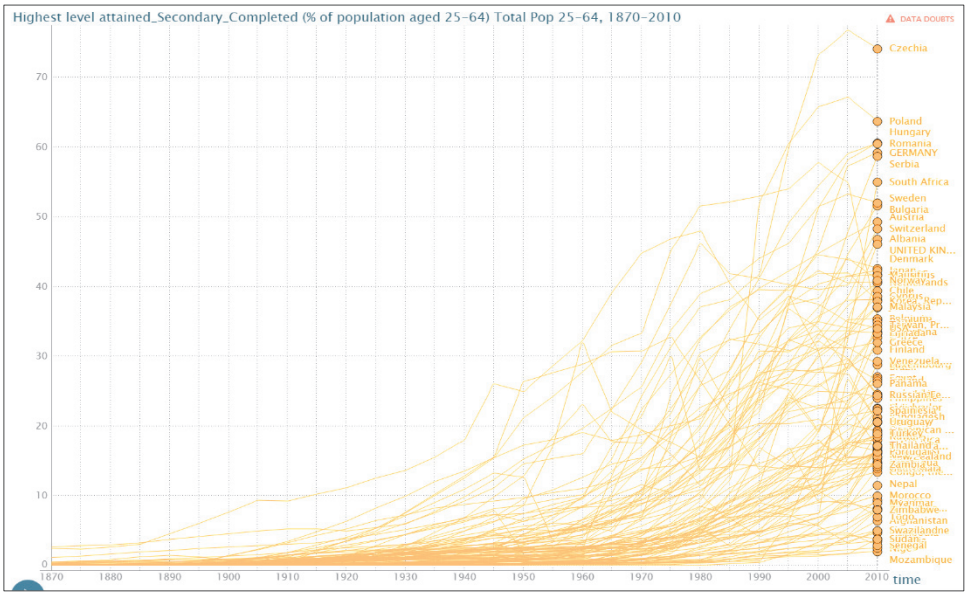


Source : Barro-Lee (2015)

Figure 3.181 Taux (%) d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint

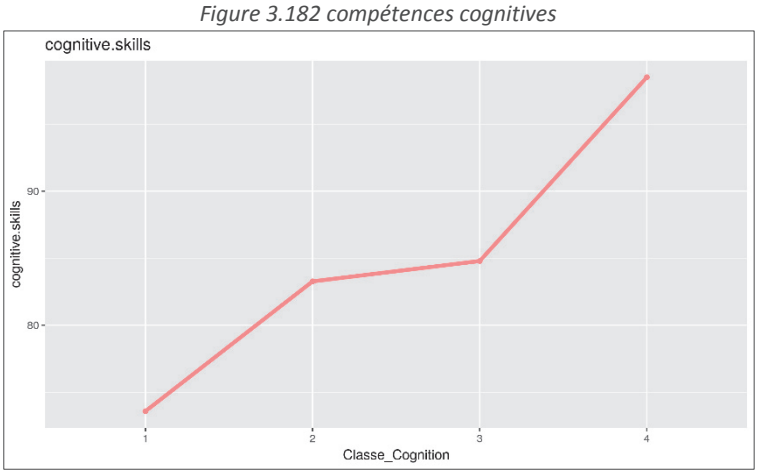


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, nous permet de conclure à des allures haussières similaires.

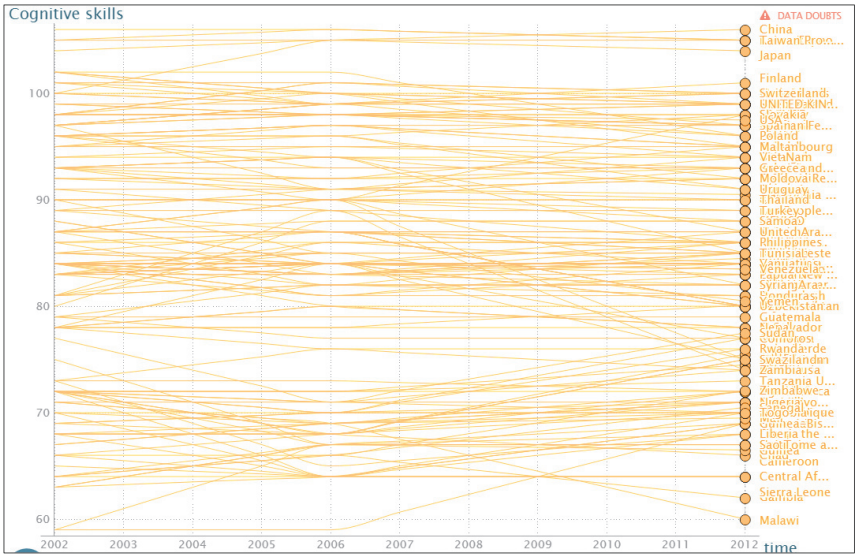


Source : Barro-Lee (2015)

Figure 3.182 compétences cognitives



Comparaison impossible entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale ci-dessous, car cette dernière n'a pas de profondeur historique suffisante.



Source : Richard Lynn & Tatu Vanhanen (2012)

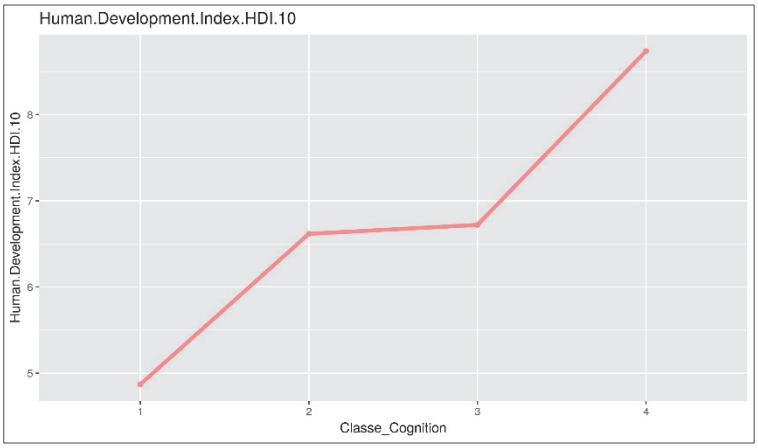
8.1.9. Variables passives : Comparaison des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales

Les variables passives sont des variables utilisées pour illustrer les classes, et qui ne sont pas des variables actives (qui elles sont utilisées pour créer les classes).

Les comparaisons des allures des courbes transversales selon l'indice d'âge transitionnel, avec celles de leurs courbes longitudinales des variables passives, sont nécessaires pour vérifier la robustesse et la cohérence d'ordonnancement des classes d'âge transitionnelles construites.
La similitude globale des allures des courbes transversales et longitudinales que nous constatons dans les comparaisons ci-après dans plusieurs domaines, montre empiriquement cette robustesse.

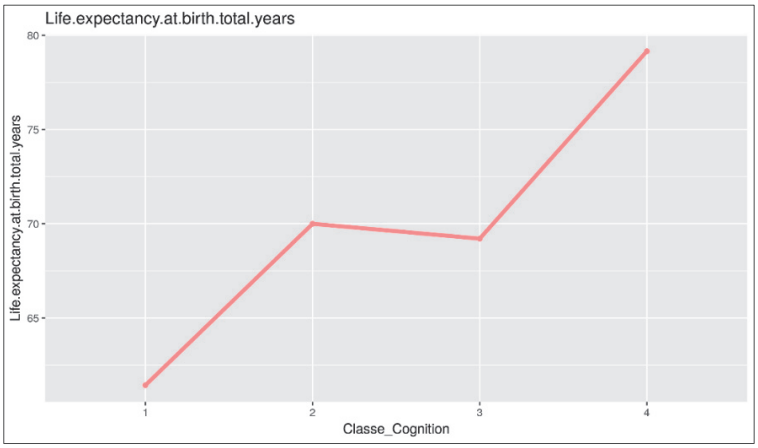
Moyennes par classe d'âge transitionnel

*Figure 3.183 Indice de développement humain (HDI*10)*



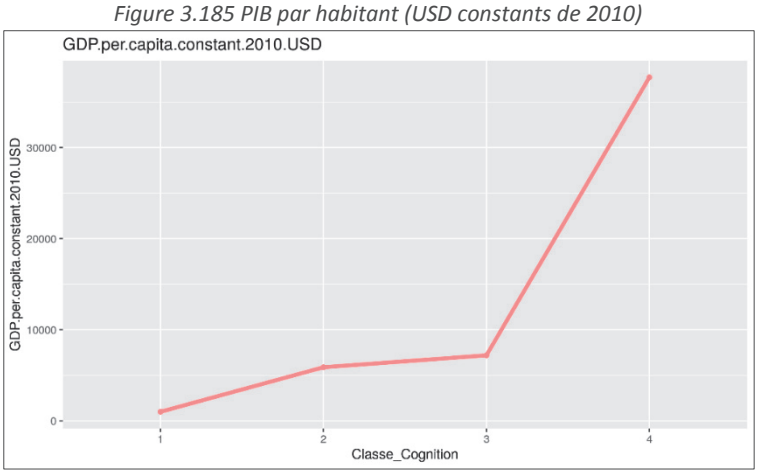
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.1*) nous permet de conclure à des allures haussières similaires.

Figure 3.184 Espérance de vie à la naissance, Total



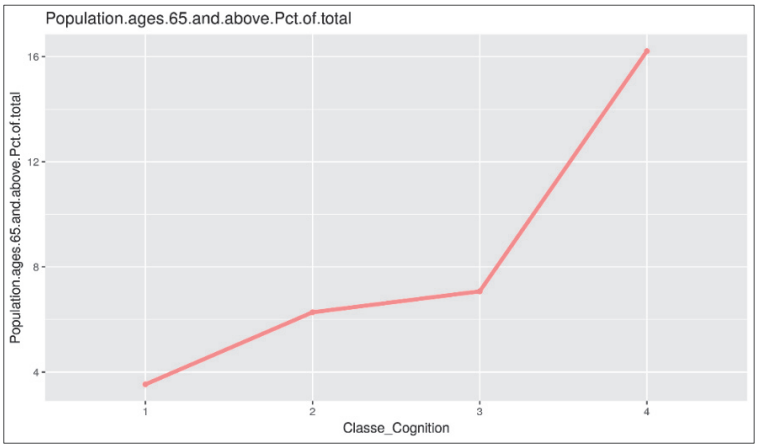
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.2*) nous permet de conclure à des allures haussières similaires.

Figure 3.185 PIB par habitant (USD constants de 2010)



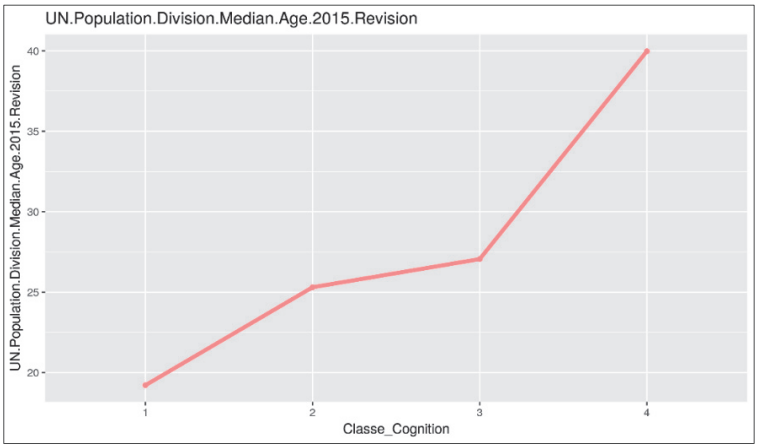
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.3*) nous permet de conclure à des allures haussières similaires.

Figure 3.186 Population âgée de 65 ans et plus, total



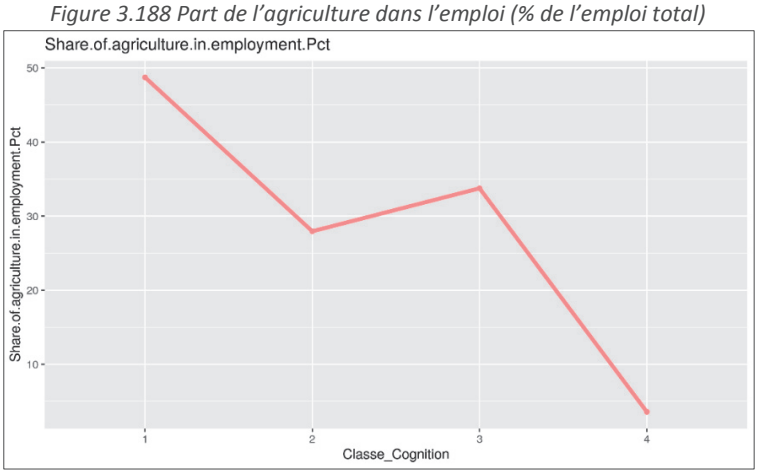
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.4*) nous permet de conclure à des allures haussières similaires.

Figure 3.187 Division de la population des Nations Unies (âge médian)



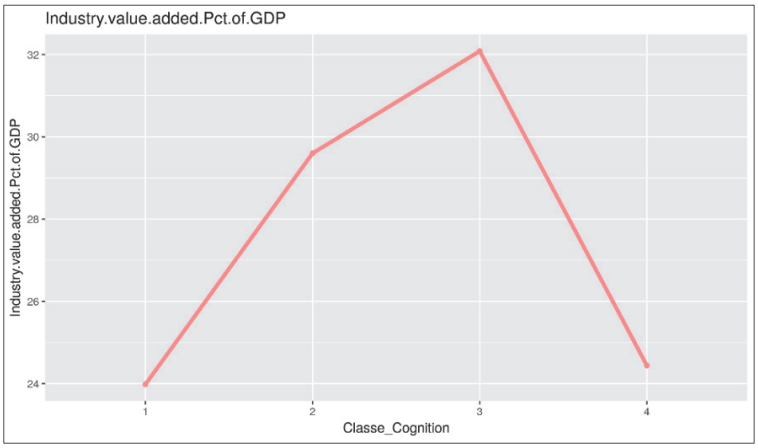
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.5*) nous permet de conclure à des allures haussières similaires.

Figure 3.188 Part de l'agriculture dans l'emploi (% de l'emploi total)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.6*) nous permet de conclure à des allures baissières similaires.

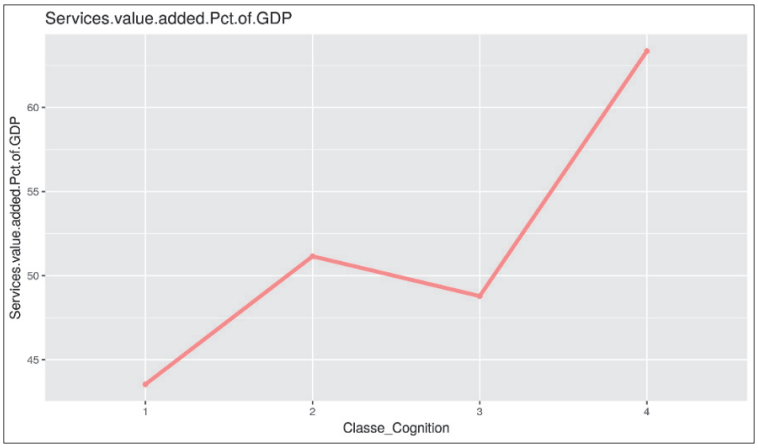
Figure 3.189 Industrie, valeur ajoutée (% du PIB)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (

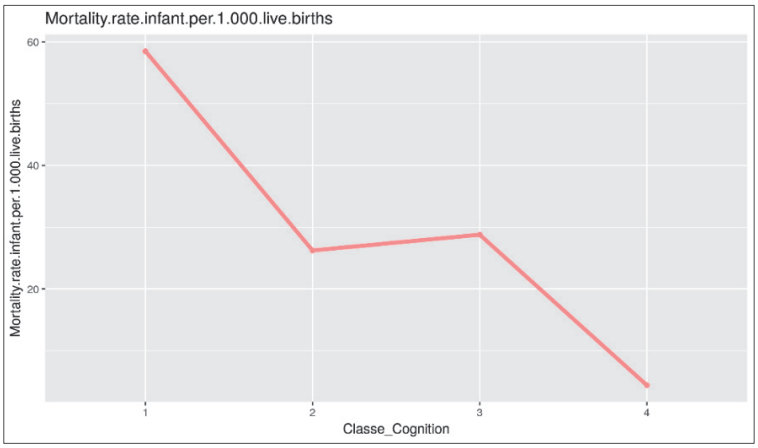
Figure 0.7) nous permet de conclure à des allures en cloche similaires.

Figure 3.190 Services, valeur ajoutée (% du PIB)



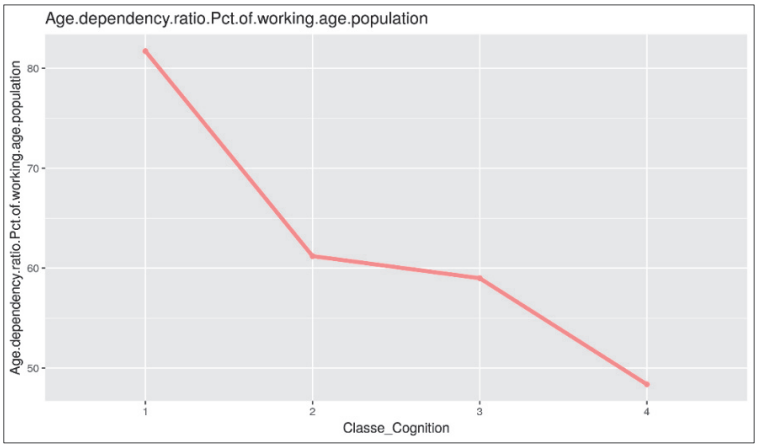
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.8*) nous permet de conclure à des allures haussières similaires.

Figure 3.191 Taux de mortalité infantile (pour 1 000 naissances vivantes)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.15*) nous permet de conclure à des allures baissières similaires.

Figure 3.192 Ratio de dépendance à l'âge (% de la population en âge de travailler)

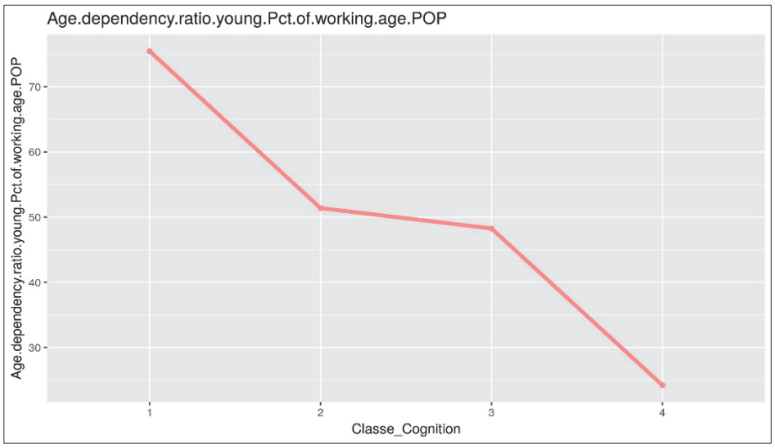


La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.10*) nous permet de conclure à des allures baissières similaires.

Figure 3.193 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler)

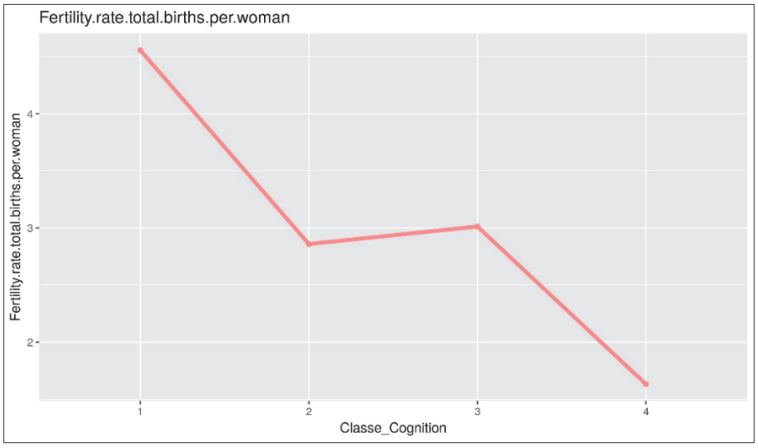
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.11*) nous permet de conclure à des allures haussières similaires.

Figure 3.194 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler)



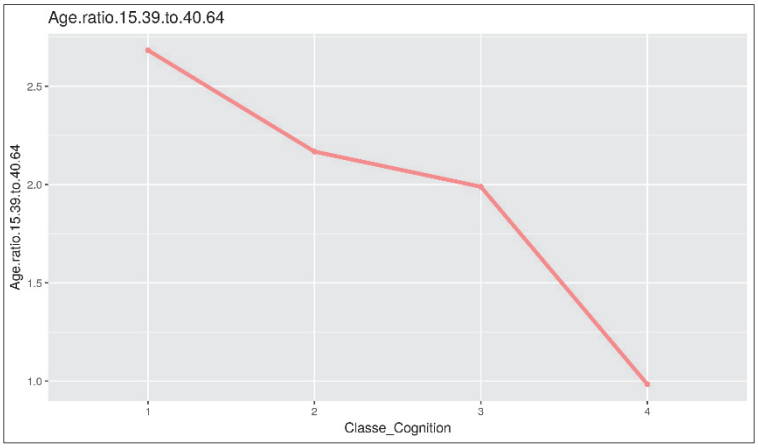
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.12*) nous permet de conclure à des allures baissières similaires.

Figure 3.195 Taux de fécondité, total (naissances par femme)



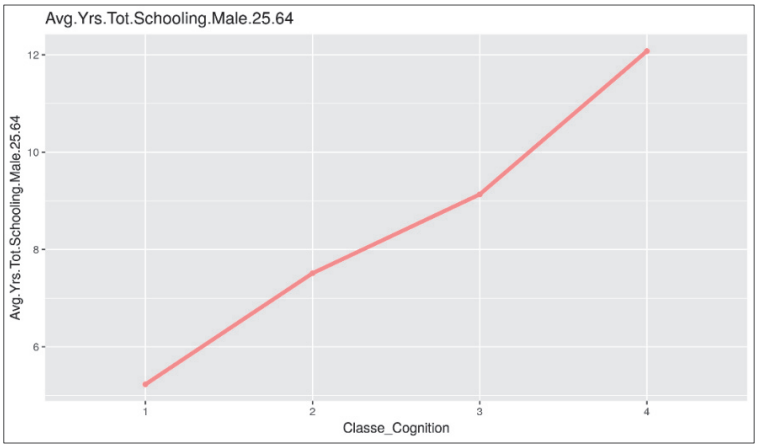
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.16*) nous permet de conclure à des allures baissières similaires.

Figure 3.196 Ratio d'âge15-39 à 40-64



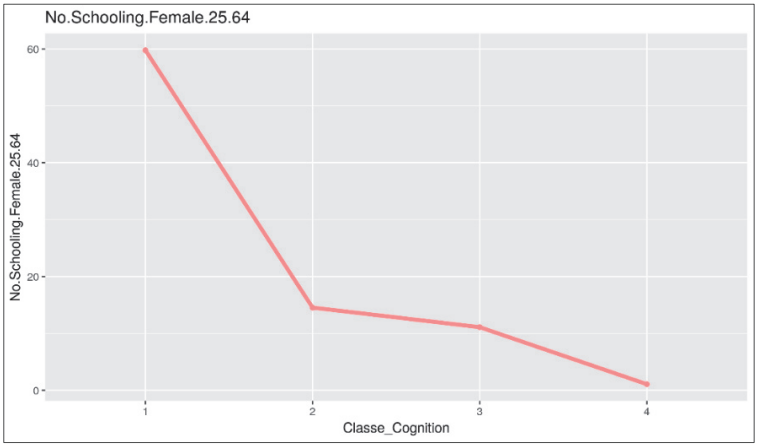
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.13*) nous permet de conclure à des allures baissières similaires.

Figure 3.197 Nombre moyen d'années de scolarisation totale des hommes de 25 à 64 ans



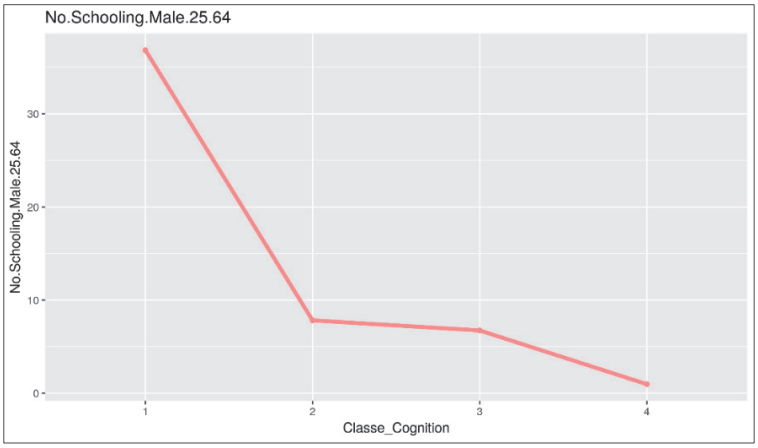
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.29*) nous permet de conclure à des allures haussières similaires.

Figure 3.198 Taux (%) des Femmes non scolarisées de 25 à 64 ans



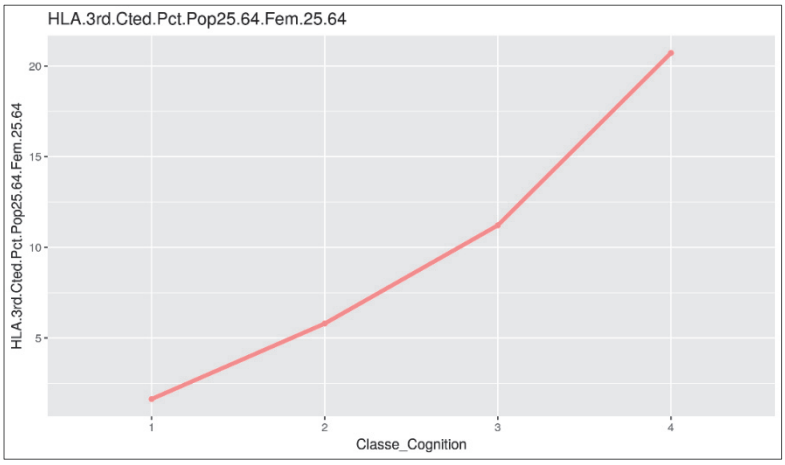
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.30*) nous permet de conclure à des allures baissières similaires.

Figure 3.199 Taux (%) des Hommes non scolarisés de 25 à 64 ans



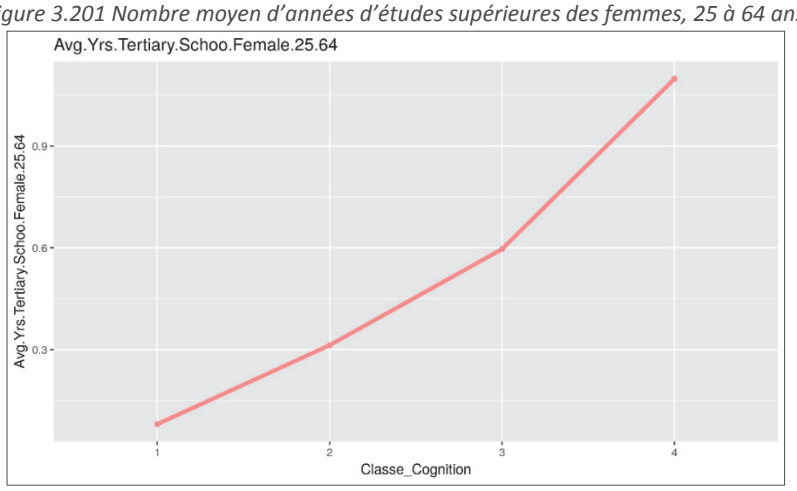
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.31*) nous permet de conclure à des allures baissières similaires.

Figure 3.200 Taux (%) des femmes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint



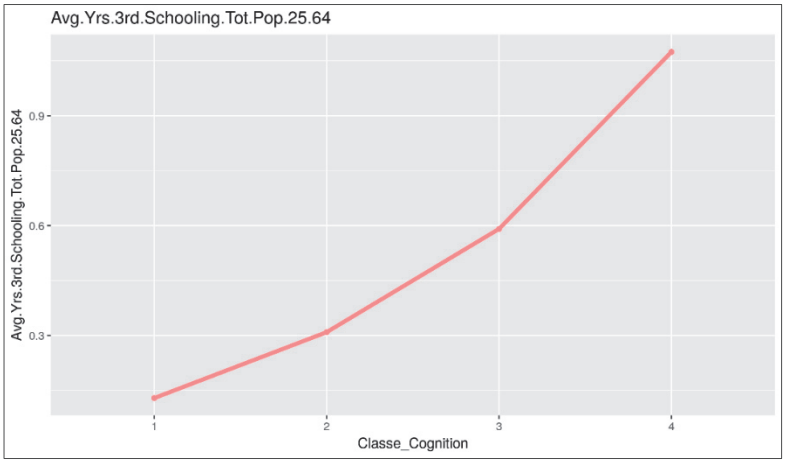
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.32*) nous permet de conclure à des allures haussières similaires.

Figure 3.201 Nombre moyen d'années d'études supérieures des femmes, 25 à 64 ans



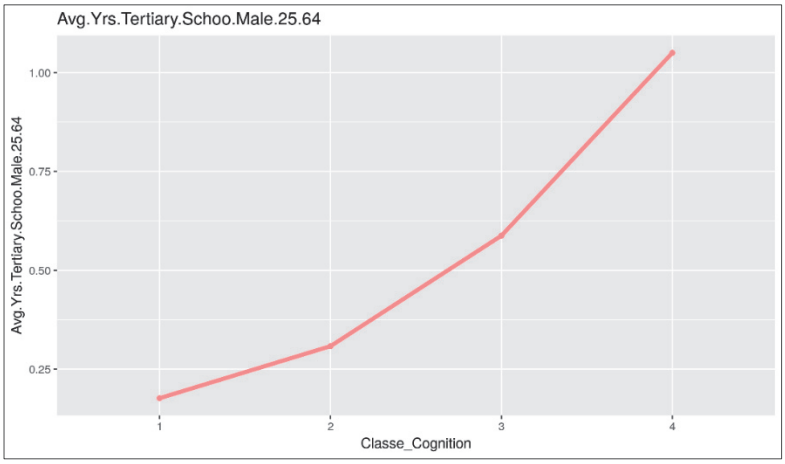
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.33*) nous permet de conclure à des allures haussières similaires.

Figure 3.202 Nombre moyen d'années d'études supérieures, population totale, 25 à 64 ans



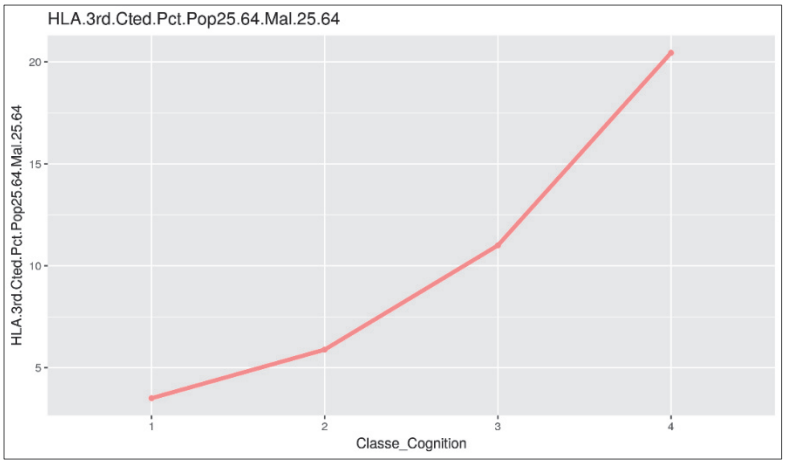
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.34*) nous permet de conclure à des allures haussières similaires.

Figure 3.203 Nombre moyen d'années d'études supérieures, des hommes, 25 à 64 ans



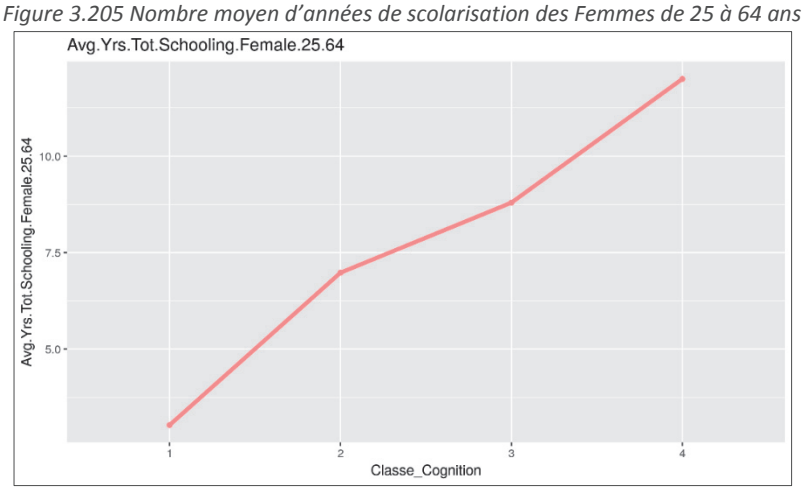
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.35*) nous permet de conclure à des allures haussières similaires.

Figure 3.204 Taux des hommes de 25-64 ans dont l'éducation tertiaire achevée constitue le niveau le plus élevé atteint



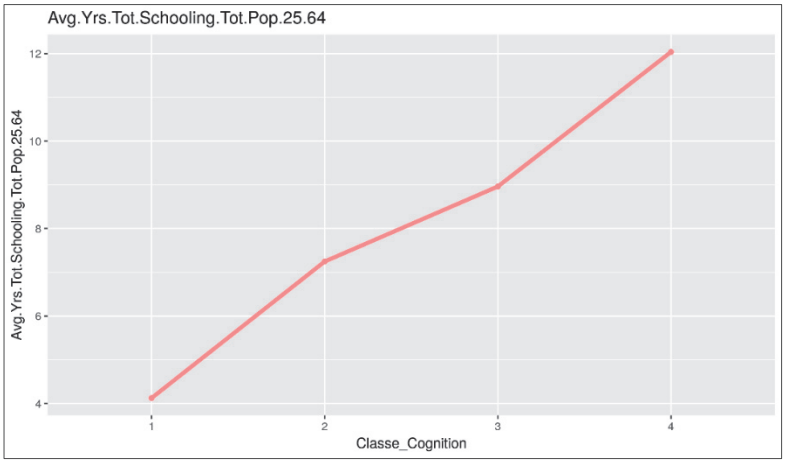
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.36*) nous permet de conclure à des allures haussières similaires.

Figure 3.205 Nombre moyen d'années de scolarisation des Femmes de 25 à 64 ans



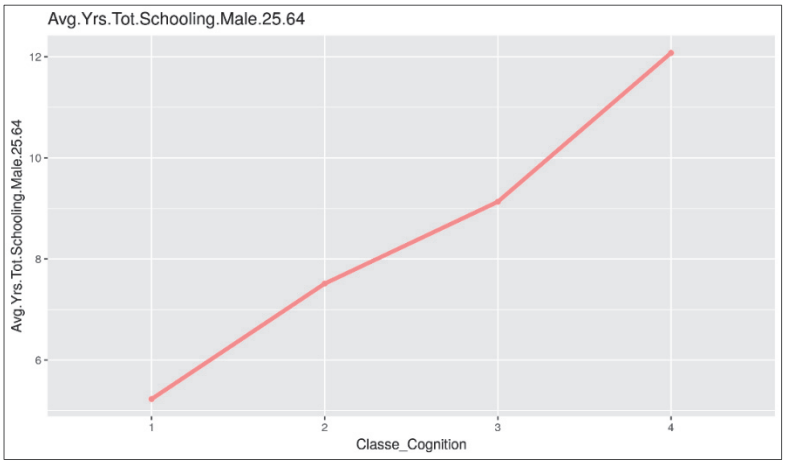
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.37)* nous permet de conclure à des allures haussières similaires.

Figure 3.206 Nombre moyen d'années de scolarisation (Population totale) 25 à 64 ans



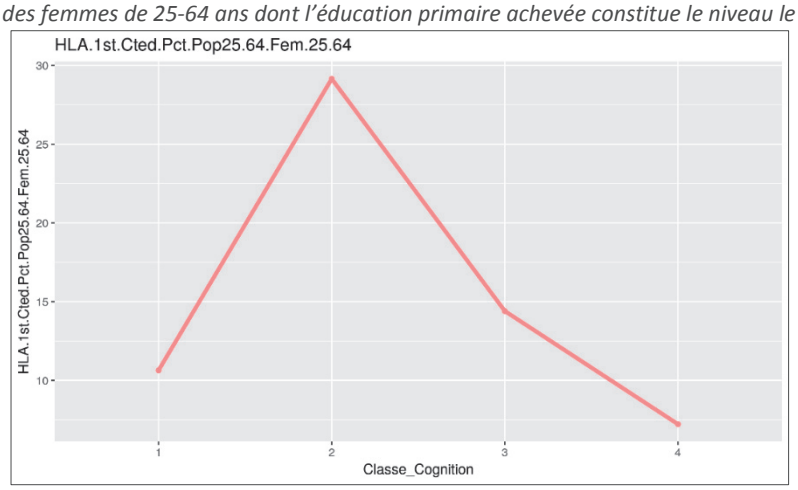
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.38*) nous permet de conclure à des allures haussières similaires.

Figure 3.207 Nombre moyen d'années de scolarisation des hommes, 25 à 64 ans



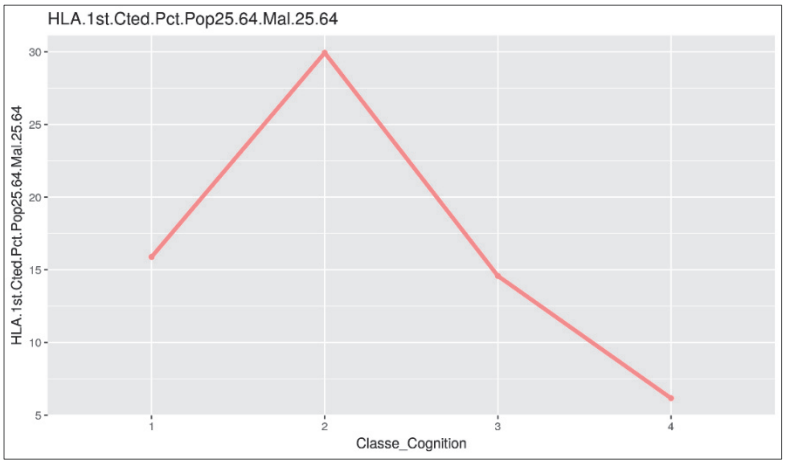
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.39*) nous permet de conclure à des allures haussières similaires.

Figure 3.208 Taux des femmes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint,



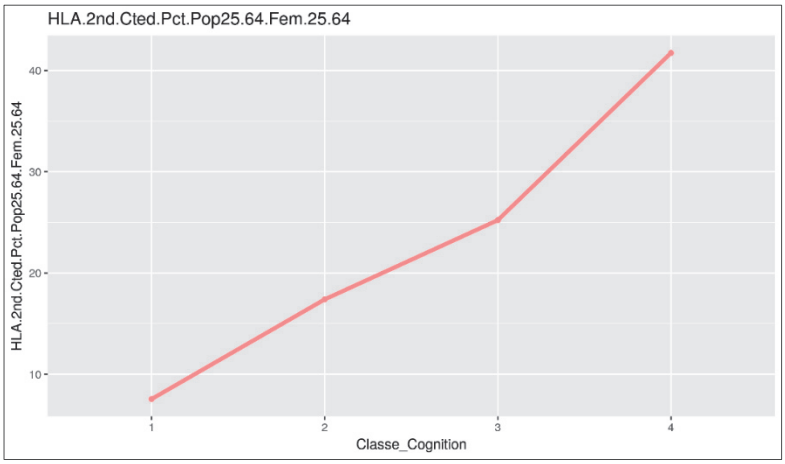
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.40*) nous permet de conclure à des allures en cloche similaires.

Figure 3.209 Taux des hommes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint



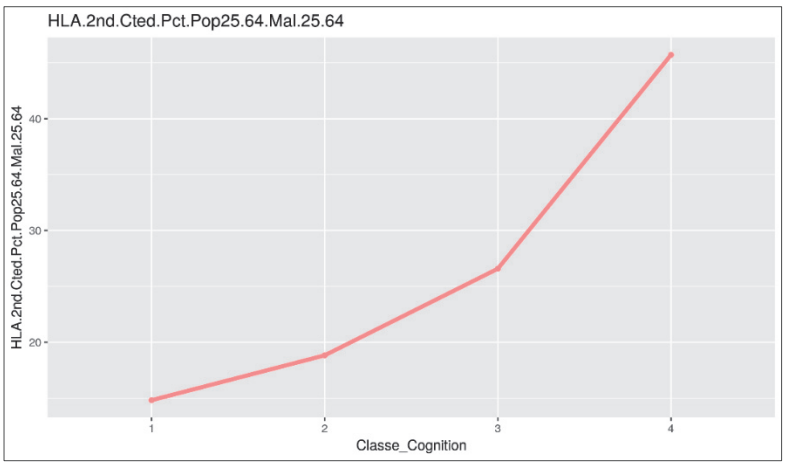
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.41)* nous permet de conclure à des allures en cloche similaires.

Figure 3.210 Taux des femmes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint



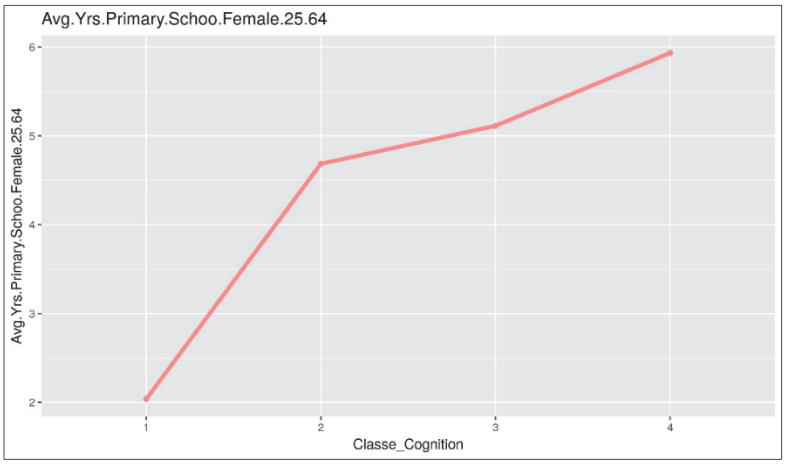
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.42*) nous permet de conclure à des allures haussières similaires.

Figure 3.211 Taux des hommes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint



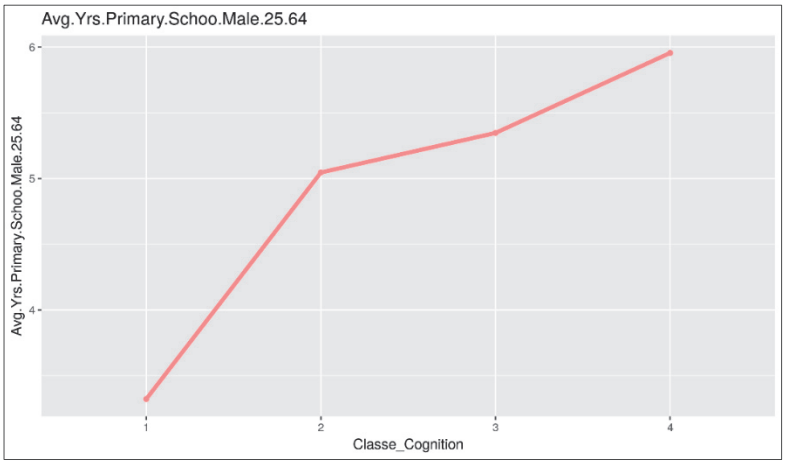
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.43*) nous permet de conclure à des allures haussières similaires.

Figure 3.212 Nombre moyen d'années d'études primaires des Femmes



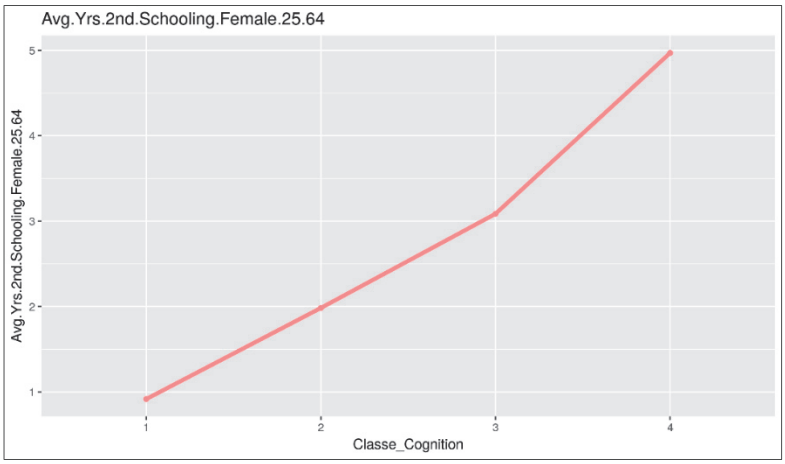
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.44*) nous permet de conclure à des allures haussières similaires.

Figure 3.213 Nombre moyen d'années d'études primaires des Hommes



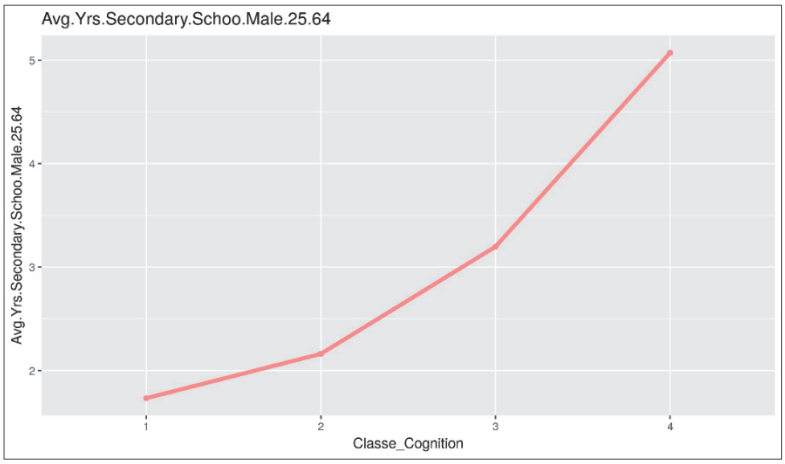
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.45*) nous permet de conclure à des allures haussières similaires.

Figure 3.214 Nombre moyen d'années d'études secondaires des femmes, 25 à 64 ans



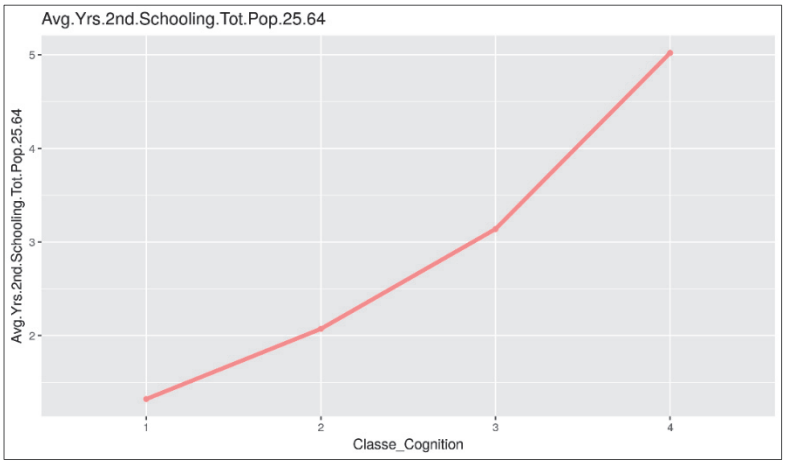
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.46*) nous permet de conclure à des allures haussières similaires.

Figure 3.215 Nombre moyen d'années d'études secondaires des hommes, 25 à 64 ans



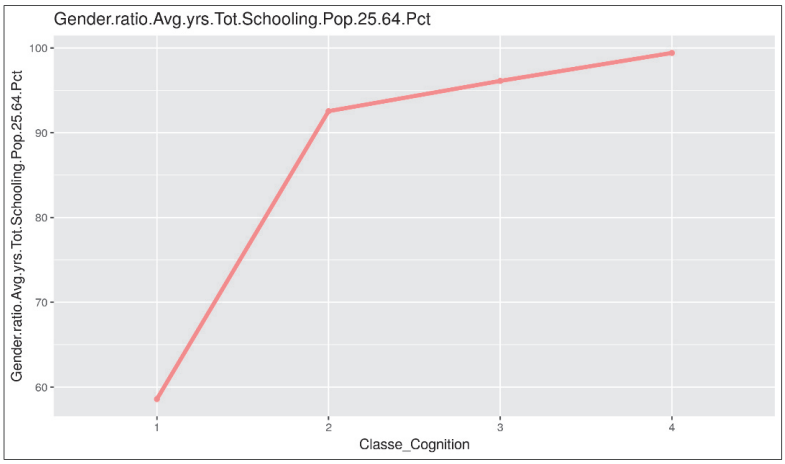
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante *(Figure 0.47*) nous permet de conclure à des allures haussières similaires.

Figure 3.216 Nombre moyen d'années d'études secondaires, population totale, 25 à 64 ans



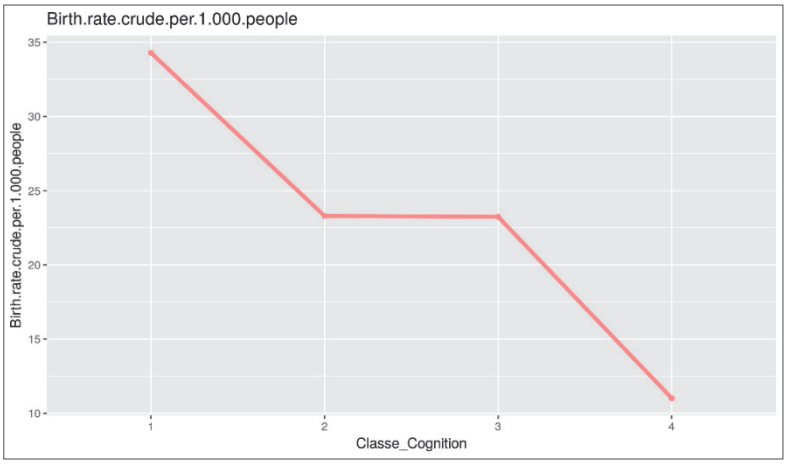
La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.48*) nous permet de conclure à des allures haussières similaires.

Figure 3.217 Ratio femmes vs hommes, nombre moyen d'années d'études (% population totale 25-64 ans)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.49*) nous permet de conclure à des allures haussières similaires.

Figure 3.218 Taux de natalité, brut (pour 1 000 personnes)



La comparaison empirique entre la courbe transversale transitionnelle ci-dessus, et la courbe longitudinale correspondante (*Figure 0.26*) nous permet de conclure à des allures baissières similaires.

8.2. Peut-on faire l'hypothèse d'une théorie transitionnelle de la cognition ?

Nous désignons par transition cognitive, une série de profondes et durables transformations cognitives de la population d'un pays, pouvant s'étaler sur quelques décennies à plusieurs générations.

Proposition de schéma transitionnel cognitif selon l'indice d'âge transitionnel

Ayant montré à la section 8.1.8 que les courbes transversales des variables de transition selon l'indice d'âge transitionnel, avaient sensiblement les mêmes allures que leurs courbes longitudinales, nous pouvons donc proposer le schéma théorique transitionnel suivant :

Figure 3.219 Schéma transitionnel selon l'indice d'âge transitionnel. Moyenne par classe d'âge transitionnel

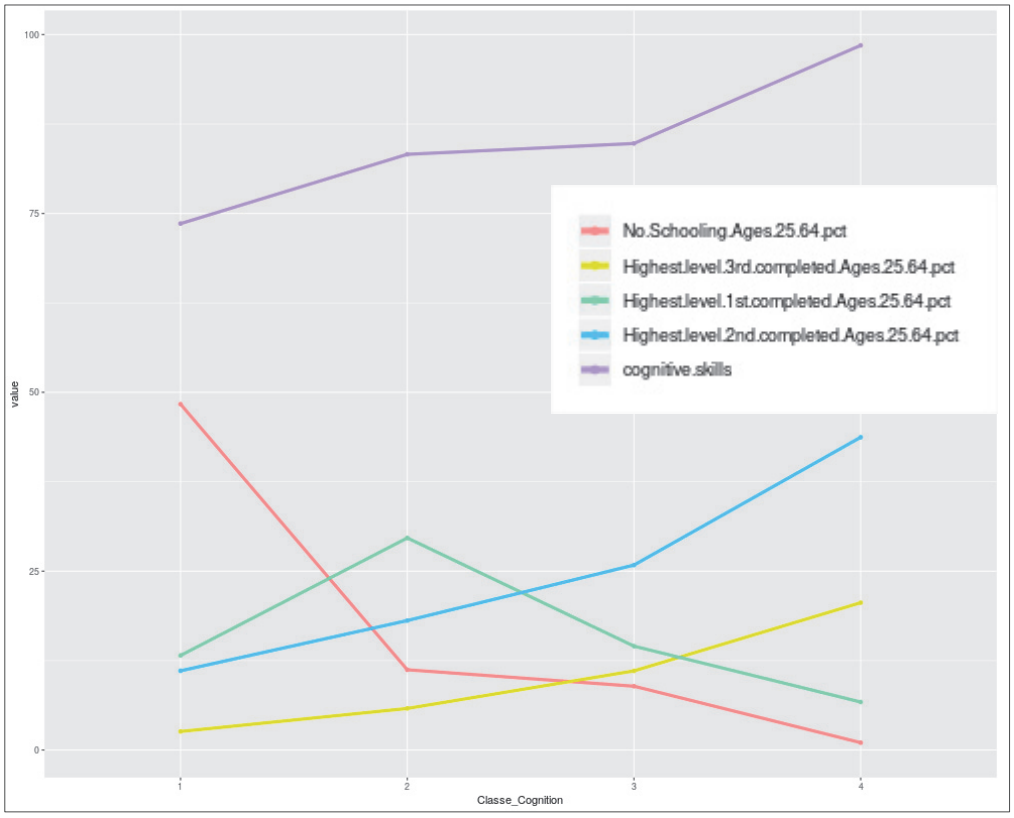
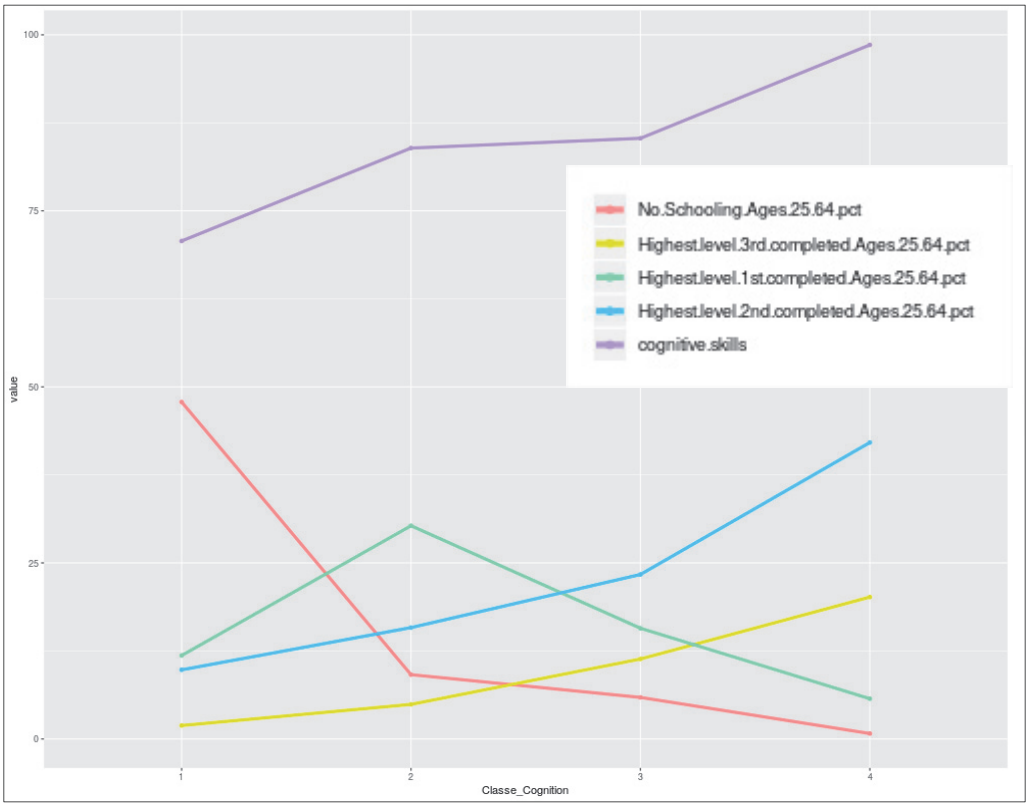
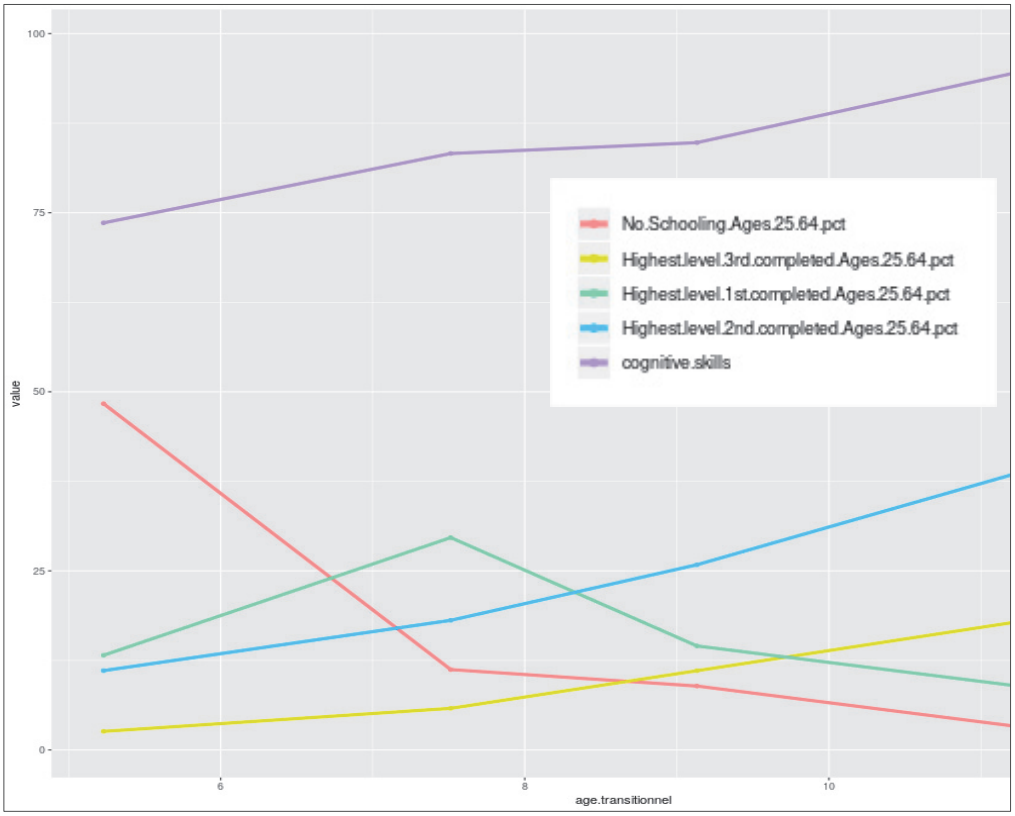


Figure 3.220 Schéma transitionnel selon l'indice d'âge transitionnel. Médiane par classe d'âge transitionnel



L'allure des deux schémas ci-dessus est sensiblement la même, aussi bien en utilisant les moyennes que les médianes : C'est l'illustration d'une robustesse du schéma transitionnel obtenu en coupe transversale.

Figure 3.221 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse



Nous pouvons illustrer le schéma théorique de la transition cognitive par le tableau des valeurs moyennes des variables transitionnelles par classes d'âge transitionnel.

Tableau 3.14 Valeurs moyennes des variables transitionnelles par classes d'âge transitionnel

Indice d'âge transitionnel	% de pays en 2012	Moyenne, Taux (%) d'adultes de 25-64 ans, non scolarisés	Moyenne, Taux (%) d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint	Moyenne, Taux (%) d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint	Moyenne, Taux (%) d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint	Compétences cognitives
1	16,4	48,3	13,2	11,1	2,6	73,6
2	15,6	11,2	29,6	18,1	5,8	83,3
3	49,2	8,9	14,5	25,9	11,1	84,8
4	18,8	1,0	6,7	43,7	20,6	97,9

Ainsi, selon le schéma théorique de transition cognitive proposé, les **compétences cognitives** (synthétisées par le QI moyen) d'un pays augmentent de manière continue en moyenne de 74 à 98.
Parallèlement, les adultes (25-64 ans) uniquement diplômés du **primaire**, représentent la majorité des adultes alphabétisés diplômés durant les deux premières classes d'âge transitionnel, avec un pic (à 30%) dans la deuxième classe.
A partir de la deuxième classe d'âge transitionnel, les adultes (25-64 ans), dont l'éducation **secondaire** achevée, constitue le niveau le plus élevé atteint, représentent la majorité des adultes alphabétisés diplômés. Ces diplômés du **secondaire** passent de 11% en classe transitionnelle 1, à 44% en classe transitionnelle 4. Et cela, tandis que les diplômés exclusifs du **primaire** décroissent en pourcentage (dans les classes transitionnelles 3 et 4), pour finir à 7% des diplômés adultes en classe transitionnelle 4.
Quant aux diplômés du **supérieur** (tertiaire), leur proportion parmi les adultes augmente continument de la classe transitionnelle 1 (à 3%) à la classe transitionnelle 4 (à 21%).

Ces 4 classes d'âge transitionnel illustrent également les 4 phases observables « actuellement » dans le processus de transition cognitive.
La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, se fera dans l'étude longitudinale de la transition cognitive.
Quant aux probables phases futures jamais observées, seules des études prospectivistes permettraient d'en spéculer les caractéristiques.

9. Analyse multi-transitionnelle

A partir de la découverte des phases « actuellement » observables parmi les pays du monde, pour chacune des cinq transitions ACIDE (anthropologique, cognitive, institutionnelle, démographique, économique), nous allons à présent rechercher les associations significatives et croisées entre plusieurs transitions.
Nous pourrons ainsi établir des relations de corrélation et de timing entre les phases issues de plusieurs transitions.
Nous allons d'abord, produire des graphiques descriptifs des croisements de phases entre les transitions ACIDE prises 2 à 2, puis, illustrer graphiquement les associations croisées (synchrones ou antagonistes) significatives des phases multi-transitionnelles, et enfin présenter des tableaux de synthèse de ces associations significatives.
Le mot « phase » est utilisé ici pour désigner une classe d'âge transitionnelle.

9.1. Economie vs Autres transitions

Nous produisons ci-après, les graphiques descriptifs des croisements de phases entre les phases actuelles de la transition économique et phases observables des autres transitions. Les comparaisons se font 2 à 2 entre l'économie et les 4 autres domaines (anthropologie, cognition, institutions et démographie).
Nous ne commenterons pas ici les liens observés car tous ne sont pas significatifs.

Figure 3.222 Economie vs démographie. Appariement descriptif des classes d'âge transitionnel des pays

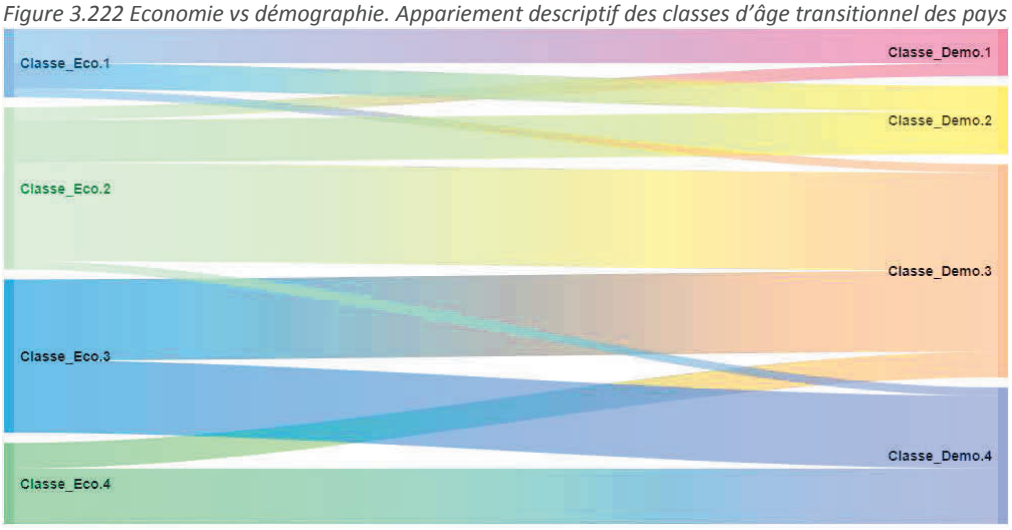


Figure 3.223 Economie vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays

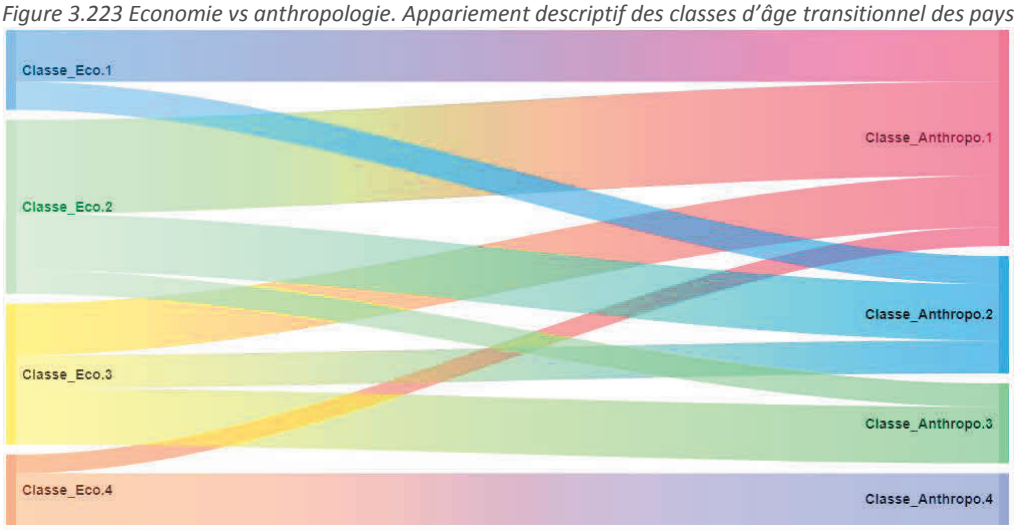


Figure 3.224 Economie vs cognition. Appariement descriptif des classes d'âge transitionnel des pays

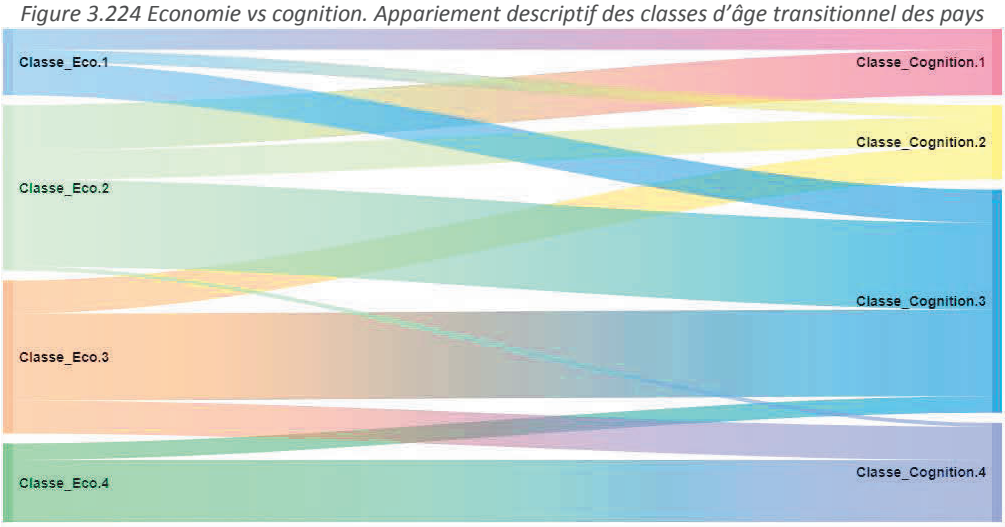
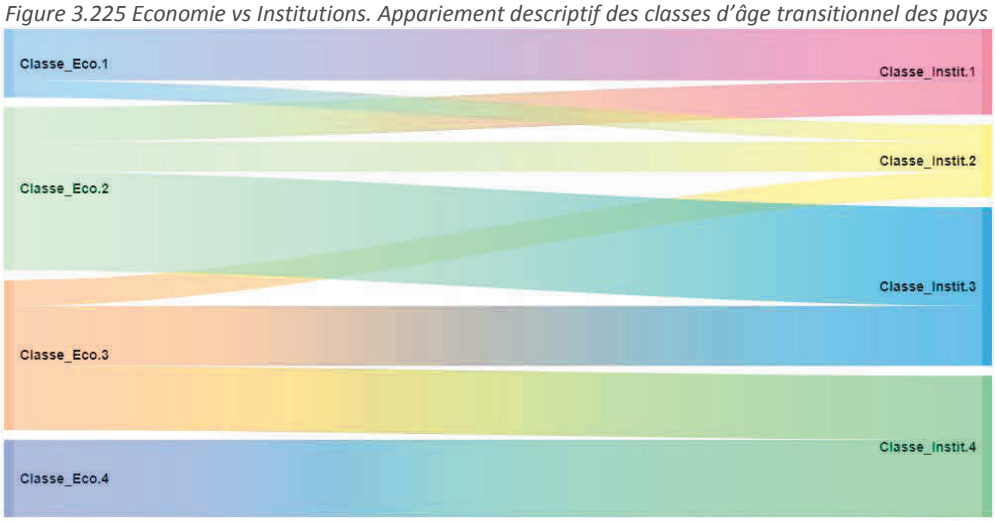
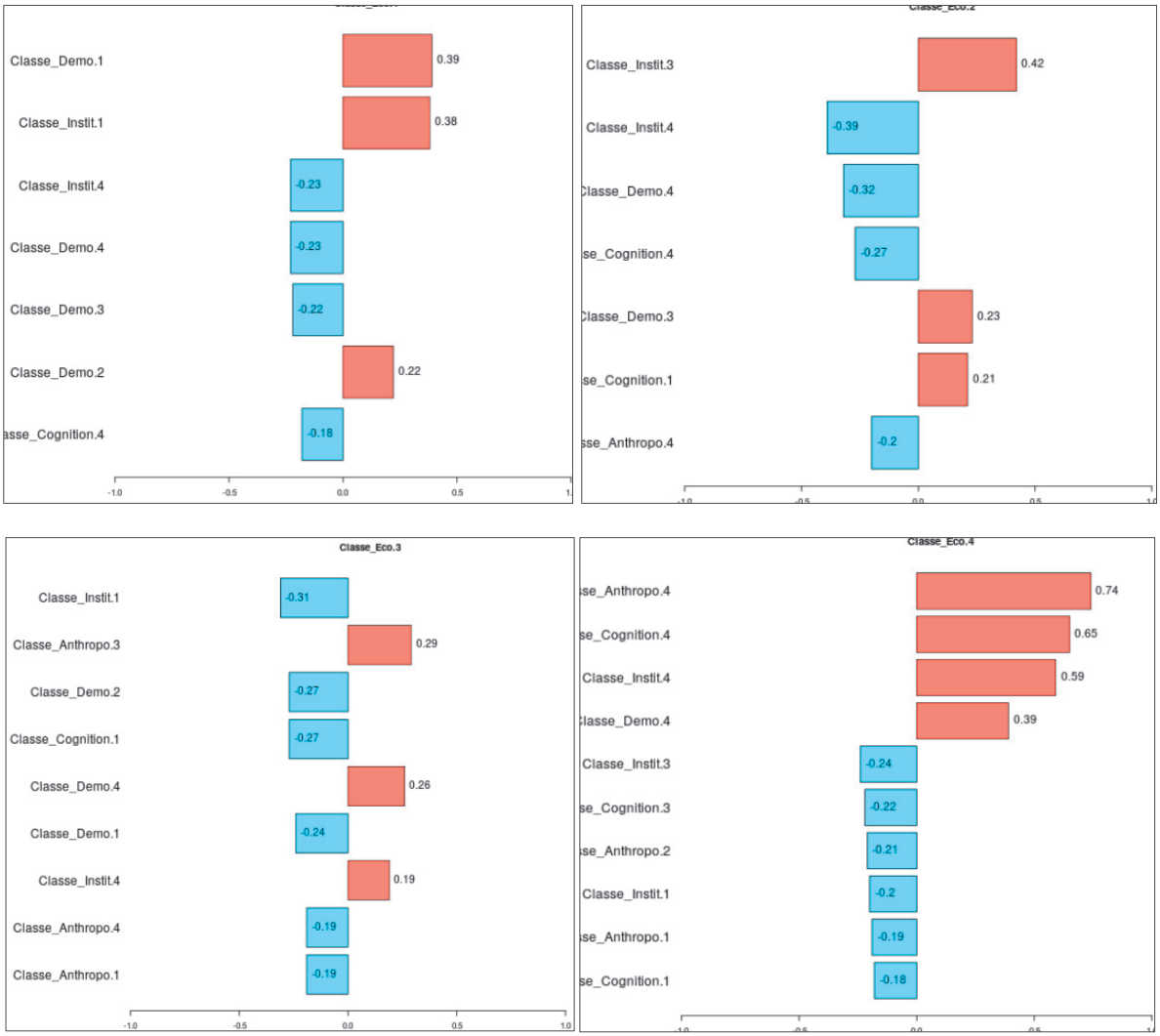


Figure 3.225 Economie vs Institutions. Appariement descriptif des classes d'âge transitionnel des pays



Dans le graphique ci-après, nous montrons les associations significatives, entre chaque phase transitionnelle économique et les phases observables dans les autres transitions (anthropologique, cognitive, institutionnelle et démographique).

Figure 3.226 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel économiques et classes des autres transitions



La phase économique 4 (classe d'âge transitionnelle économique 4) est significativement, fortement et exclusivement associée positivement avec la phase la plus avancée de chacune des 4 autres transitions (cf. graphique ci-dessous). Ce qui établit un lien clair et fort entre la phase la plus avancée de toutes transitions ACIDE.
Accessoirement, nous observons des associations significatives négatives, mais faibles, entre la phase économique 4 et les premières phases de 3 autres transitions (phases 1 et 2 anthropologiques, phases 1 et 3 cognitives, phase 1 et 3 institutionnelles).

A l'autre extrémité du spectre transitionnel, la phase économique 1 est significativement et positivement associée aux phases 1 et 2 démographiques, ainsi qu'à la phase 1 institutionnelle.
Marginalement, la phase économique 1 est significativement et négativement (mais faiblement) associée aux phases plus avancées 4 ou 3 des transitions institutionnelle, démographique et cognitive. Notons qu'aucun lien significatif n'existe entre la phase 1 économique et les phases anthropologiques.

La phase économique 2 se révèle incompatible (association négative) avec la phase la plus avancée (phase 4) de toutes les autres transitions (institutionnelle, démographique, cognitive et anthropologique, par ordre d'importance).

Cependant, la phase 2 économique est positivement associée aux phases 3 démographiques et institutionnelles, ainsi qu'à la phase 1 cognitive.

Quant à la phase économique 3, elle est significativement incompatible (association négative) avec la phase la moins avancée (phase 1) de toutes les autres transitions (institutionnelle, cognitive, démographique et anthropologique, par ordre d'importance).
Cependant, la phase 3 économique est positivement associée à la phase 3 anthropologique, ainsi qu'aux phases 4 démographiques et institutionnelles.

9.2. Démographie vs Autres transitions

Nous produisons ci-après, les graphiques descriptifs des croisements de phases entre les phases actuelles de la transition démographique et phases observables des autres transitions. Les comparaisons se font 2 à 2 entre l'économie et les 4 autres domaines (anthropologie, cognition, institutions et économique).
Nous ne commenterons pas ici les liens observés car tous ne sont pas significatifs.

Figure 3.227 Démographie vs économie. Appariement descriptif des classes d'âge transitionnel des pays

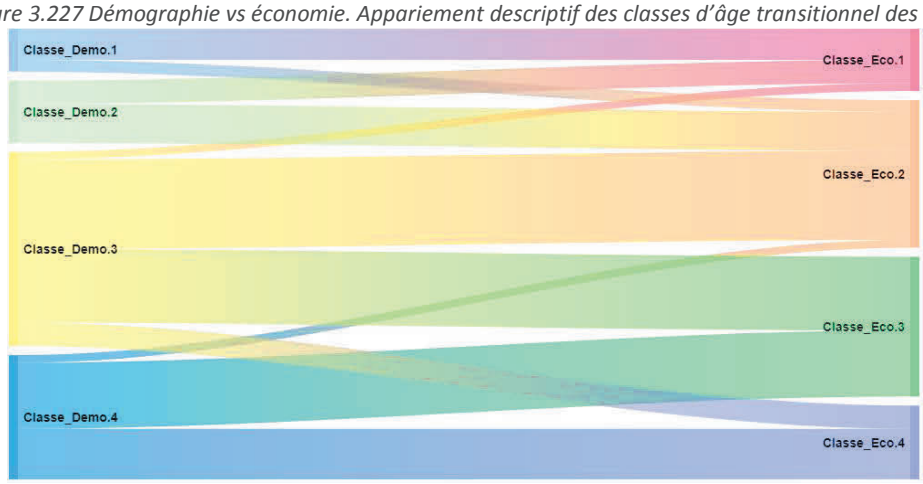


Figure 3.228 Démographie vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays

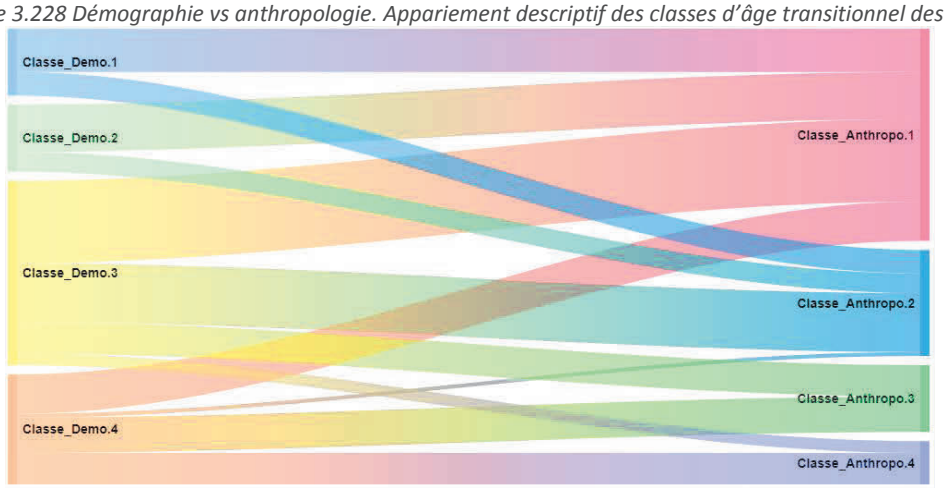


Figure 3.229 Démographie vs cognition. Appariement descriptif des classes d'âge transitionnel des pays

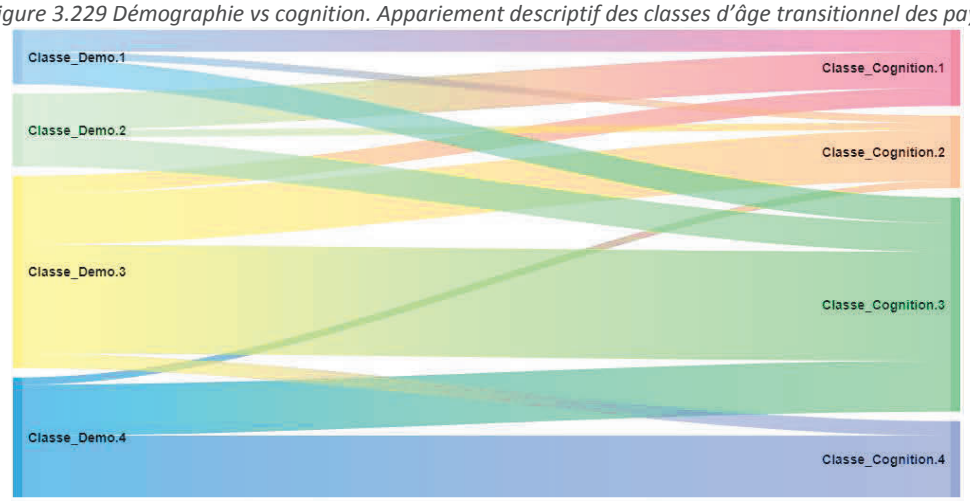
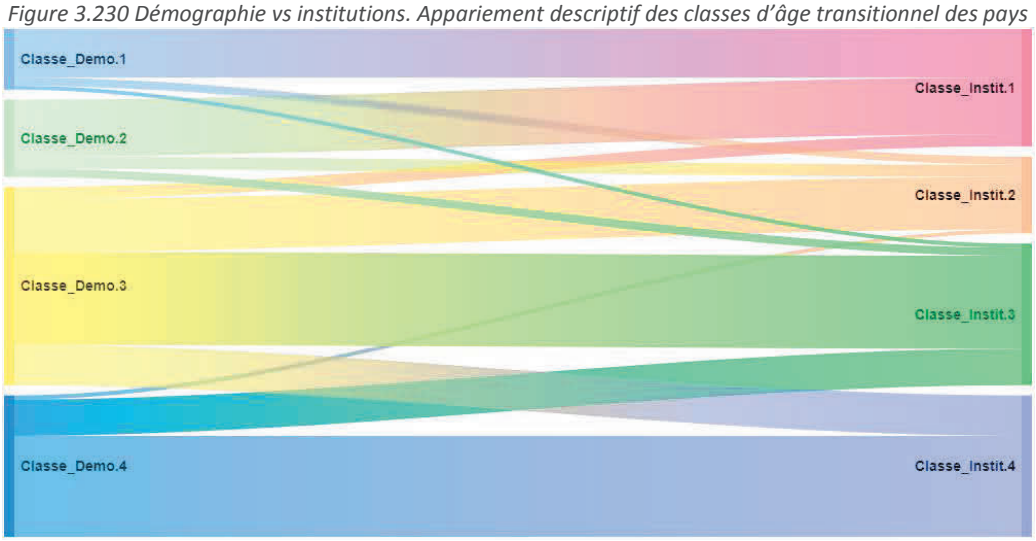
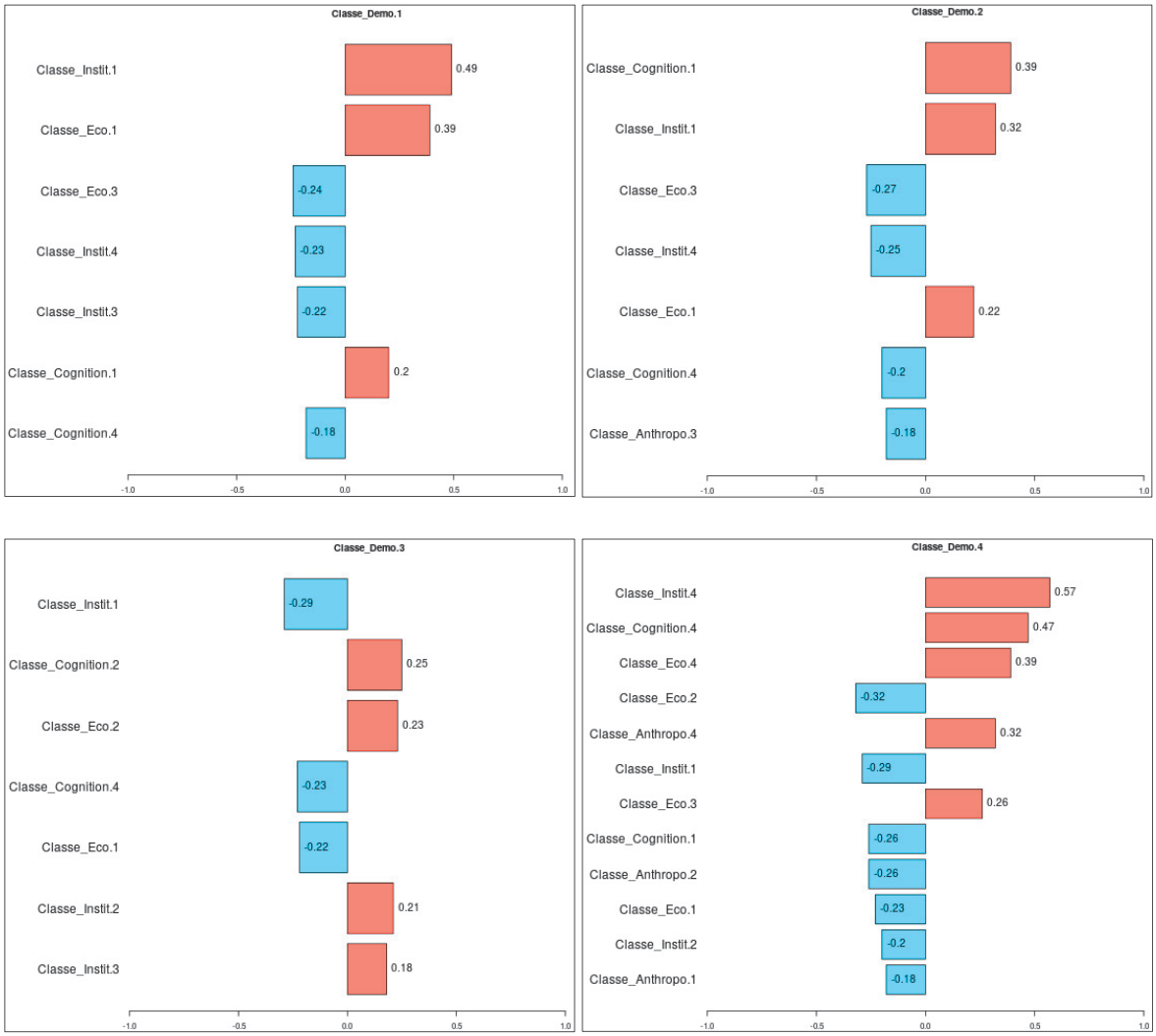


Figure 3.230 Démographie vs institutions. Appariement descriptif des classes d'âge transitionnel des pays



Dans le graphique ci-après, nous montrons les associations significatives, entre chaque phase transitionnelle démographique et les phases observables dans les autres transitions (anthropologique, cognitive, institutionnelle et économique).

Figure 3.231 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel démographiques et classes des autres transitions



La phase démographique 4 (classe d'âge transitionnelle démographique 4) est significativement et positivement associée avec la phase la plus avancée (phase 4) de chacune des 4 autres transitions (cf. graphique ci-dessous), ainsi qu'avec la phase 3 démographique (association faible).
Accessoirement, nous observons des associations significatives négatives, mais faibles, entre la phase démographique 4 et les premières phases de 3 autres transitions (phases 1 et 2 économiques, institutionnelles et anthropologiques, et phase 1 cognitive).

A l'autre extrémité du spectre transitionnel, la phase démographique 1 est significativement et positivement associée à la phase la moins avancée (phase 1) des transitions institutionnelle, économique et cognitive.
Marginalement, la phase démographique 1 est significativement et négativement (mais faiblement) associée aux phases plus avancées (phases 4 ou 3) des transitions institutionnelle, économique et cognitive. Notons qu'aucun lien significatif n'existe entre la phase 1 démographique et les phases anthropologiques.

La phase démographique 2 se révèle incompatible (association négative faible) avec la phase la plus avancée de toutes les autres transitions (phase 4 pour les transitions institutionnelle et cognitive, phase 3 pour les transitions économique et anthropologique).
A contrario, la phase 2 démographique est clairement et positivement associée à la phase la moins avancée (phase 1) des transitions cognitive, institutionnelle et économique.

Quant à la phase démographique 3, elle est significativement incompatible (association négative faible) avec la phase la moins avancée (phase 1) des transitions institutionnelle et économique, ainsi qu'à la phase la plus avancée (phase 4) cognitive.
Cependant, la phase 3 démographique est positivement associée à la phase 2 cognitive, économique et institutionnelle, ainsi qu'à la phase 3 institutionnelle.

9.3. Anthropologie vs Autres transitions

Nous produisons ci-après, les graphiques descriptifs des croisements de phases entre les phases actuelles de la transition anthropologique et phases observables des autres transitions. Les comparaisons se font 2 à 2 entre l'économie et les 4 autres domaines (cognition, institutions, démographie et économie).
Nous ne commenterons pas ici les liens observés car tous ne sont pas significatifs.

Figure 3.232 Anthropologie vs économie. Appariement descriptif des classes d'âge transitionnel des pays

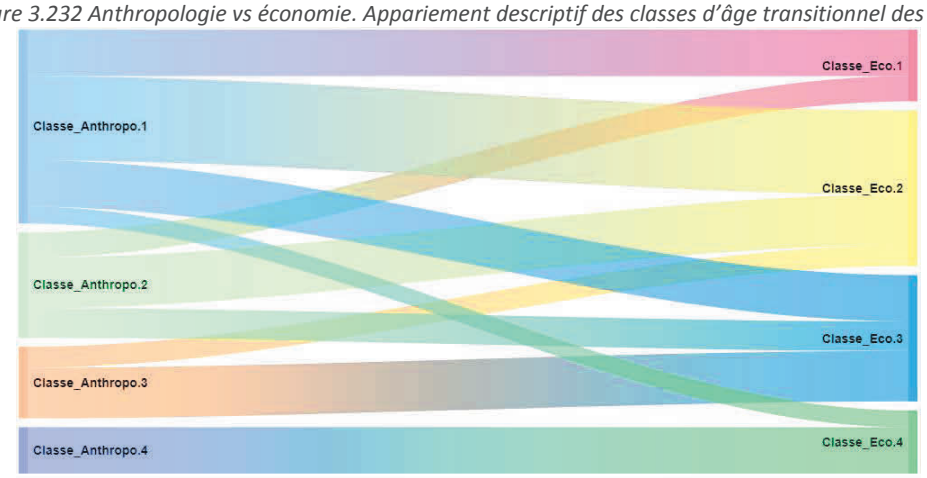


Figure 3.233 Anthropologie vs démographie. Appariement descriptif des classes d'âge transitionnel des pays

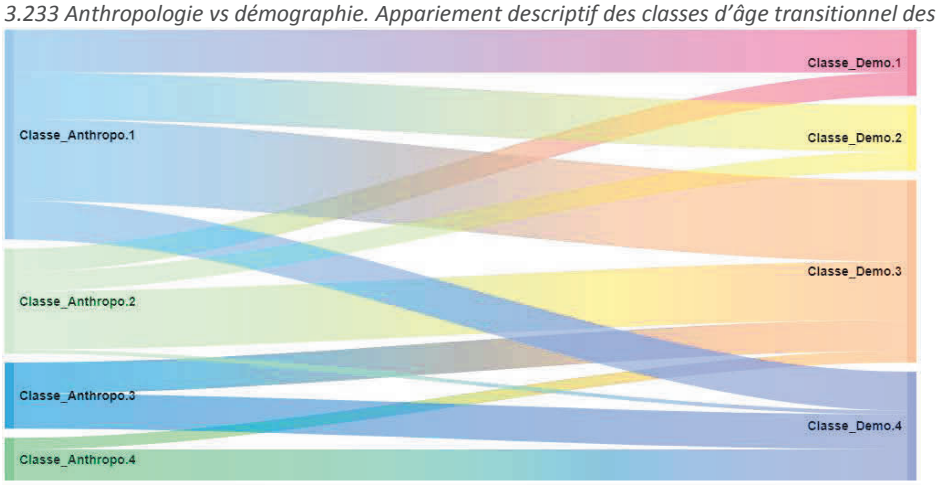


Figure 3.234 Anthropologie vs cognition. Appariement descriptif des classes d'âge transitionnel des pays

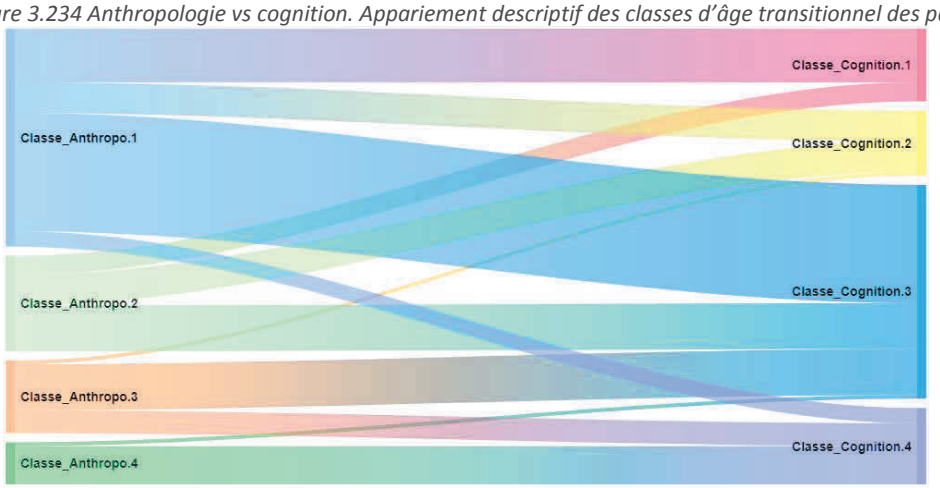
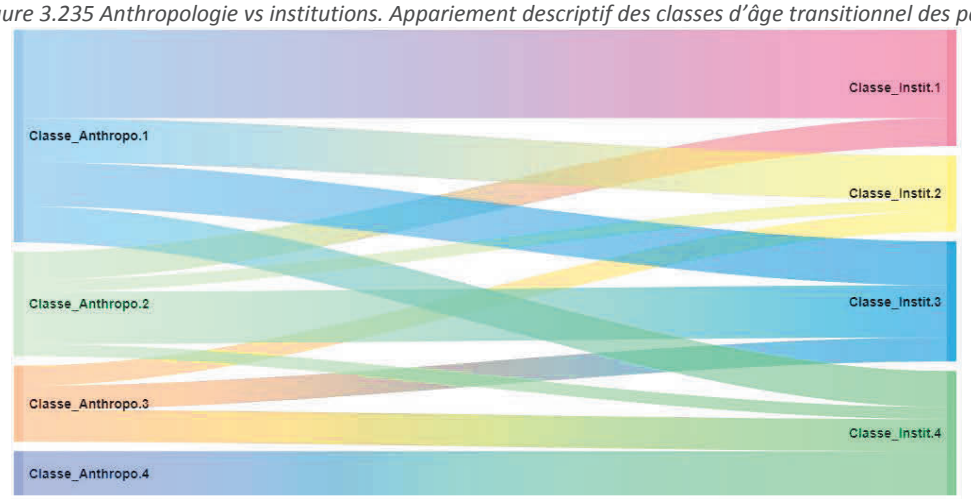
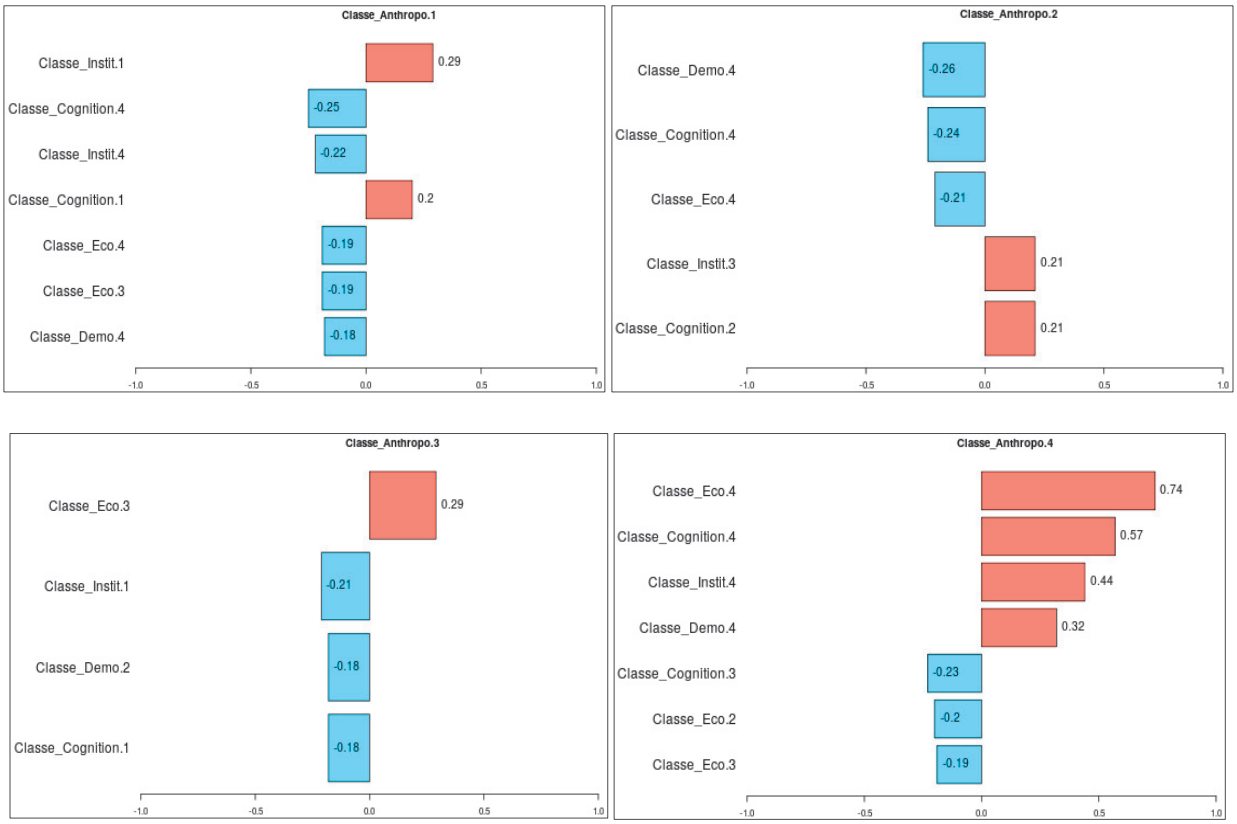


Figure 3.235 Anthropologie vs institutions. Appariement descriptif des classes d'âge transitionnel des pays



Dans le graphique ci-après, nous montrons les associations significatives, entre chaque phase transitionnelle anthropologique et les phases observables dans les autres transitions (cognitive, institutionnelle, démographique, économique).

Figure 3.236 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel anthropologiques et classes des autres transitions



La phase anthropologique 4 (classe d'âge transitionnelle anthropologique 4) est significativement et positivement associée uniquement avec la phase la plus avancée (phase 4) de chacune des 4 autres transitions (cf. graphique ci-dessous). Le lien est très fort avec l'économie, fort avec la cognition et moyennement fort avec les institutions et la démographie.

Accessoirement, nous observons des associations significatives négatives, mais faibles, entre la phase anthropologique 4 et les phases intermédiaires de 2 autres transitions (phases 2 et 3 économiques, institutionnelles et phase 3 cognitive).

A l'autre extrémité du spectre transitionnel, la phase anthropologique 1 est significativement et positivement (mais faiblement) associée à la phase la moins avancée (phase 1) des transitions institutionnelle et cognitive.
Marginalement, la phase anthropologique 1 est significativement et négativement (mais faiblement) associée aux phases plus avancées des transitions cognitive (phase 4), institutionnelle (phase 4), économique (phases 4 et 3) et démographique (phase 4).

La phase anthropologique 2 se révèle incompatible (association négative faible) avec la phase la plus avancée (phase 4) de 3 autres transitions (démographique, cognitive et économique).
A contrario, la phase 2 anthropologique est positivement (mais faiblement) associée aux phases intermédiaires des transitions institutionnelle (phase 3) et cognitive (phase 2).

Quant à la phase 3 anthropologique n'est positivement (et faiblement) associée qu'avec la phase 3 économique.
Cependant, la phase anthropologique 3, elle est significativement incompatible (association négative faible) avec la phase la moins avancée (phase 1) des transitions institutionnelle et démographiques, ainsi qu'avec la phase 2 cognitive.

9.4. Institutions vs Autres transitions

Nous produisons ci-après, les graphiques descriptifs des croisements de phases entre les phases actuelles de la transition institutionnelle et phases observables des autres transitions. Les comparaisons se font 2 à 2 entre l'économie et les 4 autres domaines (anthropologie, cognition, démographie, économie).
Nous ne commenterons pas ici les liens observés car tous ne sont pas significatifs.

Figure 3.237 Institutions vs économie. Appariement descriptif des classes d'âge transitionnel des pays

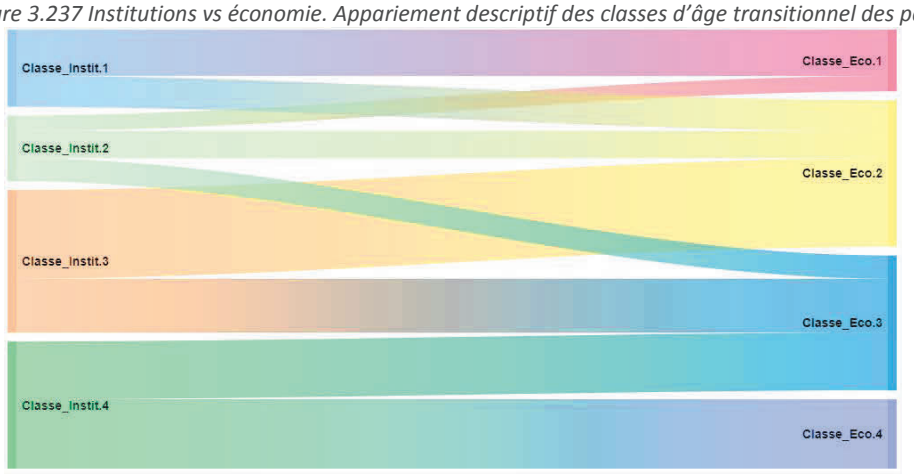


Figure 3.238 Institutions vs démographie. Appariement descriptif des classes d'âge transitionnel des pays

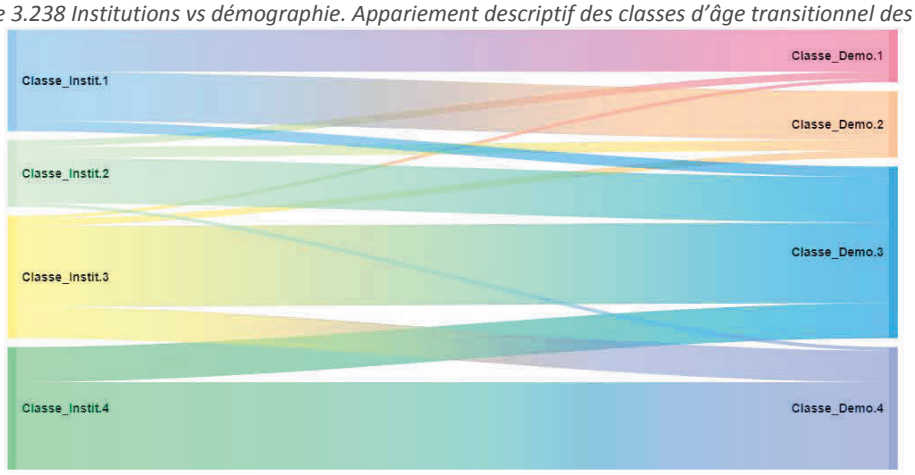


Figure 3.239 Institutions vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays

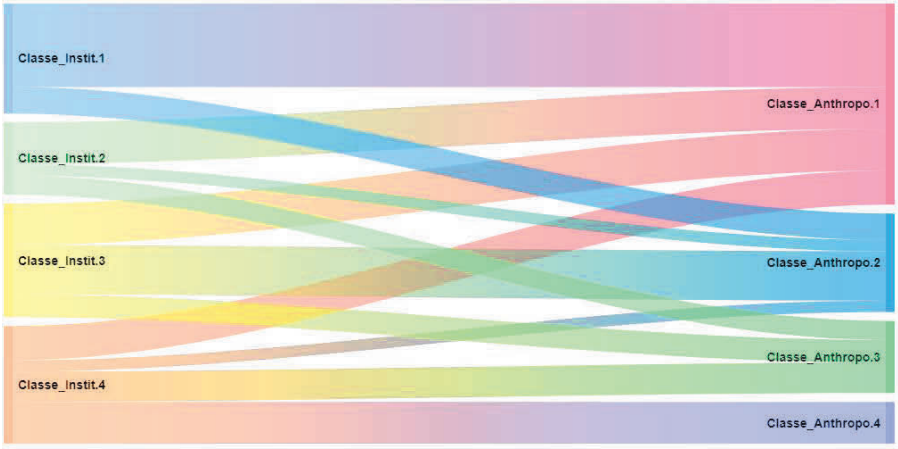
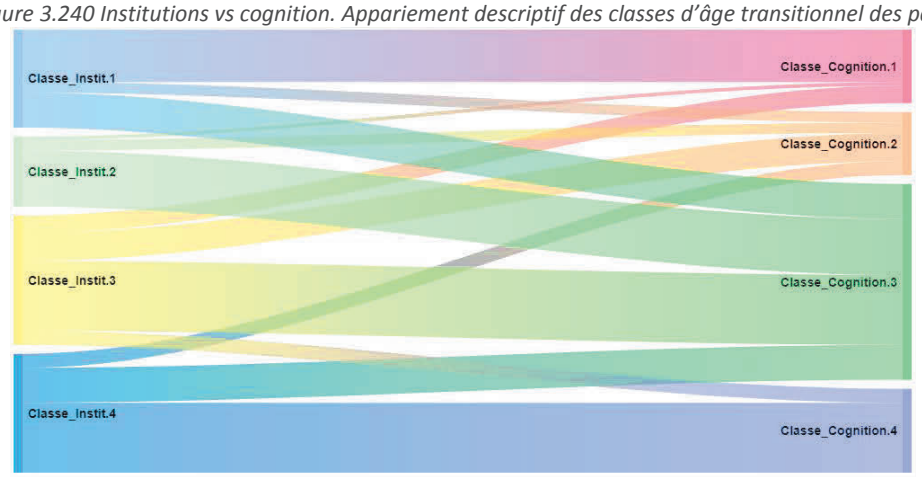
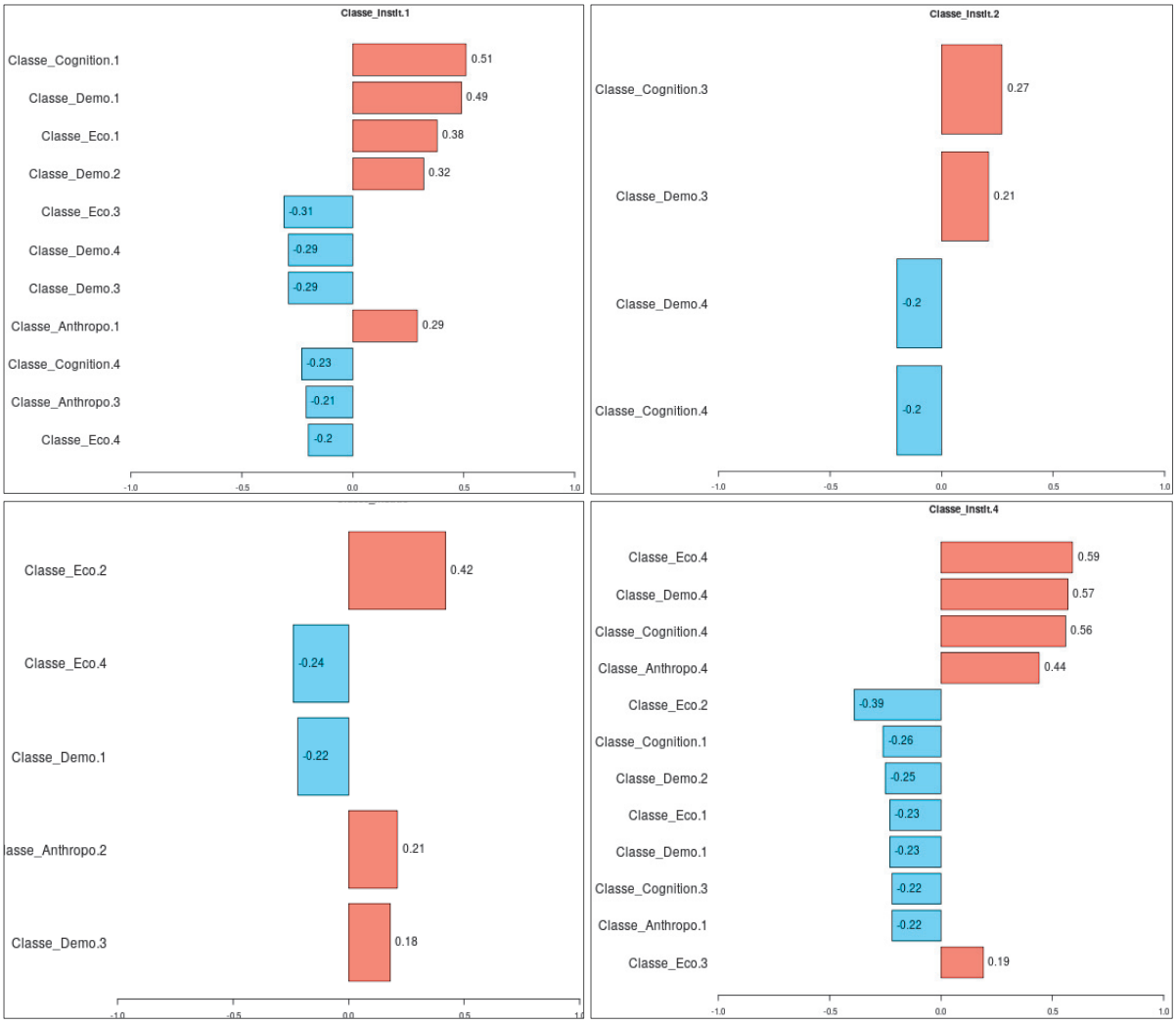


Figure 3.240 Institutions vs cognition. Appariement descriptif des classes d'âge transitionnel des pays



Dans le graphique ci-après, nous montrons les associations significatives, entre chaque phase transitionnelle institutionnelle et les phases observables dans les autres transitions (anthropologique, cognitive, démographique, économique).

Figure 3.241 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel institutionnelles et classes des autres transitions



La phase institutionnelle 4 (classe d'âge transitionnelle institutionnelle 4) est significativement et positivement associée uniquement avec la phase la plus avancée (phase 4) de chacune des 4 autres transitions (cf. graphique ci-dessous). Le lien est fort avec l'économie, la démographie et la cognition et moyennement fort avec l'anthropologie.
Par ailleurs, nous observons des associations significatives négatives (mais faibles sauf pour l'économie) entre la phase institutionnelle 4 et les phases « initiales » des 4 autres transitions (phases 1 et 2 économiques et démographiques, phases 1 et 3 cognitive, phase 1 anthropologique).

A l'autre extrémité du spectre transitionnel, la phase institutionnelle 1 est significativement et positivement associée à la phase la moins avancée (phase 1) des transitions cognitive (lien fort), démographique, économique (phases 1 et 2) et anthropologique. La force du lien étant moyenne avec les 3 dernières transitions énumérées.
A contrario, la phase institutionnelle 1 est significativement et négativement associée aux phases plus avancées des transitions économique (phases 4 et 3), démographique (phase 4 et 3), cognitive (phase 4), anthropologique (phase 4).

La phase 2 institutionnelle est positivement (mais faiblement) associée aux phases intermédiaires des transitions cognitive (phase 3) et démographique (phase 3).
Cependant, la phase institutionnelle 2 se révèle incompatible (association négative faible) avec la phase la plus avancée (phase 4) de 2 autres transitions (démographique et cognitive).

Quant à la phase 3 institutionnelle, elle est positivement associée avec la phase 2 économique (lien moyen), et faiblement associée avec la phase 2 anthropologique et la phase 3 démographique.
Par ailleurs, la phase institutionnelle 3, est significativement incompatible (association négative faible) avec la phase la moins avancée (phase 1) économique et la phase la plus avancée (phase 4) démographique.

9.5. Cognition vs Autres transitions

Nous produisons ci-après, les graphiques descriptifs des croisements de phases entre les phases actuelles de la transition cognitive et phases observables des autres transitions. Les comparaisons se font 2 à 2 entre l'économie et les 4 autres domaines (anthropologie, institutions, démographie, économie).
Nous ne commenterons pas ici les liens observés car tous ne sont pas significatifs.

Figure 3.242 Cognition vs économie. Appariement descriptif des classes d'âge transitionnel des pays

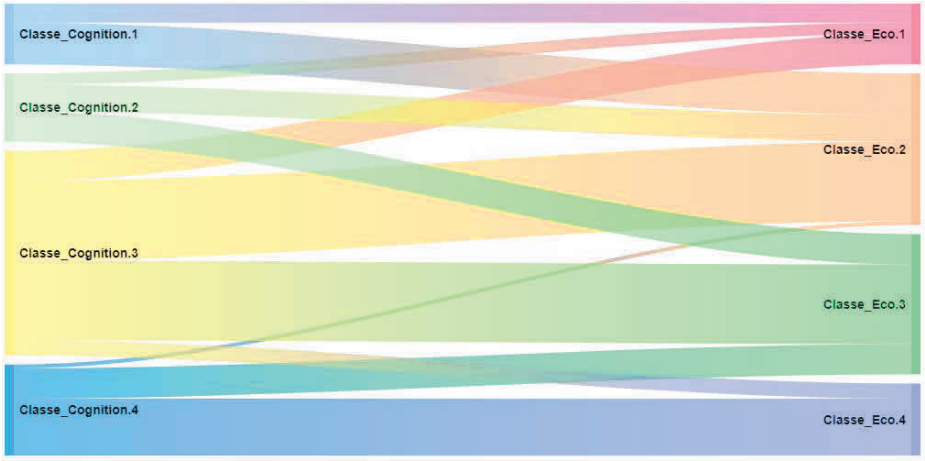


Figure 3.243 Cognition vs démographie. Appariement descriptif des classes d'âge transitionnel des pays

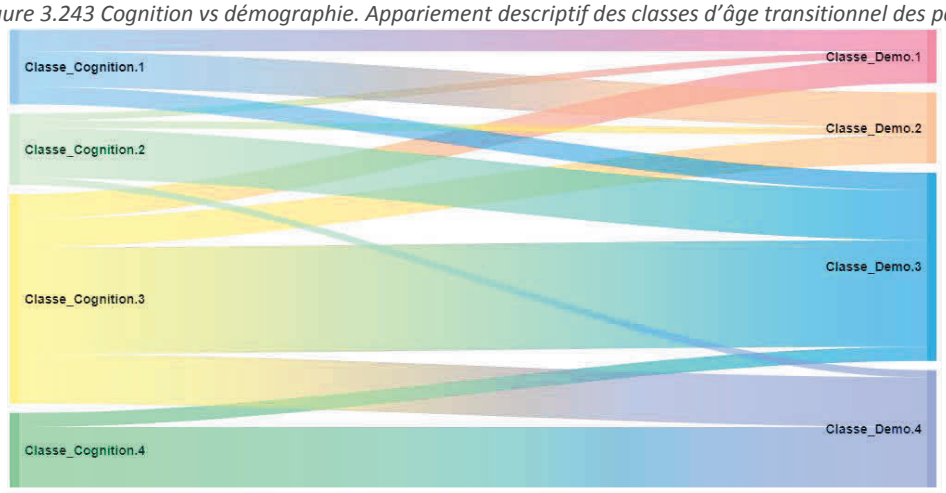


Figure 3.244 Cognition vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays

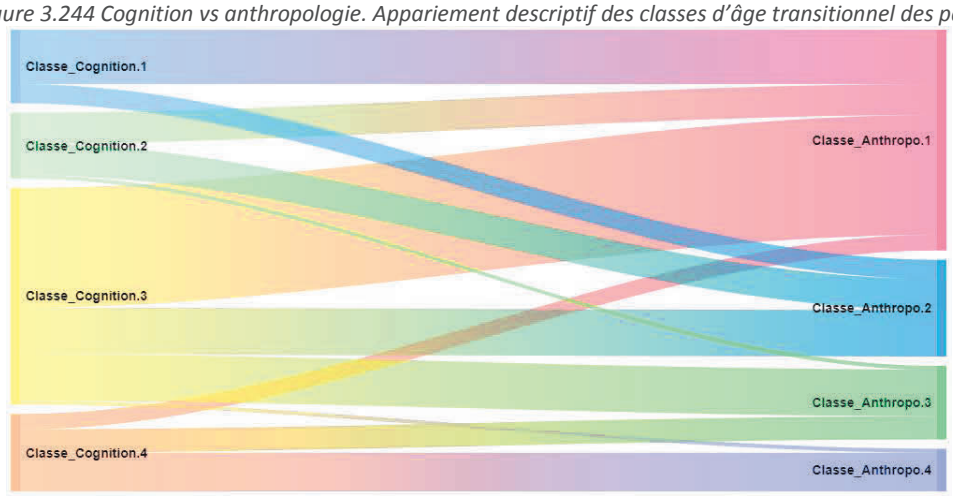
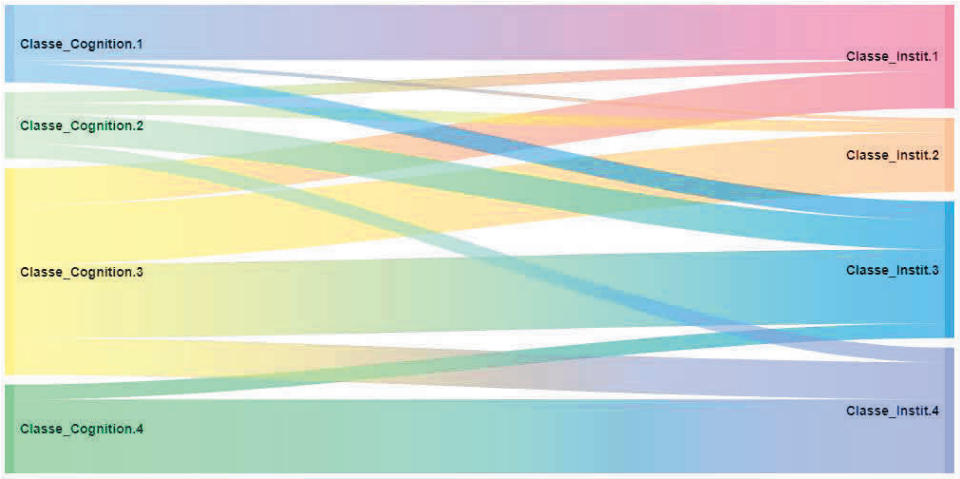
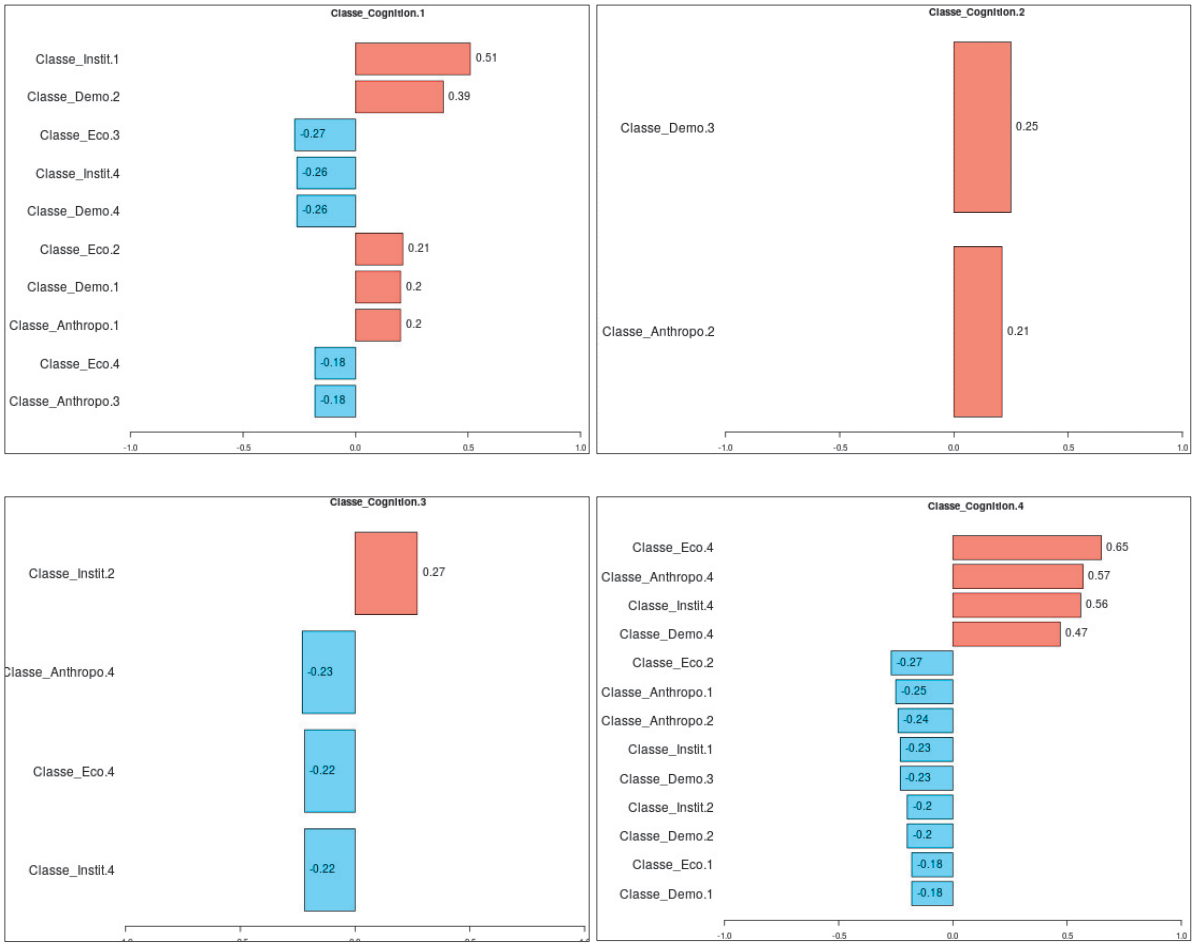


Figure 3.245 Cognition vs institutions. Appariement descriptif des classes d'âge transitionnel des pays



Dans le graphique ci-après, nous montrons les associations significatives, entre chaque phase transitionnelle cognitive et les phases observables dans les autres transitions (anthropologique, institutionnelle, démographique, économique).

Figure 3.246 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel cognitives et classes des autres transitions



La phase cognitive 4 (classe d'âge transitionnelle cognitive 4) est significativement, fortement et positivement associée uniquement avec la phase la plus avancée (phase 4) de chacune des 4 autres

transitions (économique, anthropologique, institutionnelle et démographique, par ordre d'importance du lien).
Par ailleurs, nous observons des associations significatives négatives (mais faibles) entre la phase cognitive 4 et les phases « initiales » des 4 autres transitions (phases 1 et 2 économique, anthropologique et institutionnelle ; et phases 1, 2 et 3 démographique).

A l'autre extrémité du spectre transitionnel, la phase cognitive 1 est significativement et positivement associée aux phases les moins avancées des transitions institutionnelle (lien fort avec phase 1), démographique (lien faible avec phase 1, moyen avec phase 2), économique (lien faible avec phase 2) et anthropologique (lien faible avec phase 1).
A contrario, la phase cognitive 1 est significativement et négativement (mais faiblement) associée aux phases plus avancées des transitions économique (phases 3 et 4), institutionnelle (phase 4), démographique (phase 4) et anthropologique (phase 3).

La phase 2 cognitive est positivement (mais faiblement) associée aux phases intermédiaires des transitions démographique (phase 3) et anthropologique (phase 2).

Quant à la phase 3 cognitive, elle est positivement associée avec la phase 2 institutionnelle (lien faible). Par ailleurs, la phase cognitive 3 est significativement incompatible (association négative faible) avec la phase la plus avancée (phase 4) de transitions anthropologique, économique et institutionnelle.

9.6. Synthèse des associations entre transitions ACIDE

Tableau 3.15 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel des transitions ACIDE

	A	C	I	D	E
A		43,10%	36,10%	32,40%	53,60%
C	43,10%		45,40%	41,60%	43,30%
I	36,10%	45,40%		50,30%	54,20%
D	32,40%	41,60%	50,30%		49,80%
E	53,60%	43,30%	54,20%	49,80%	

En tenant compte du degré de liberté et en utilisant l'échelle de Cohen, nous obtenons l'interprétation suivante sur l'intensité des liens entre les 5 transitions ACIDE :

Tableau 3.16 Intensité des liens entre les 5 transitions ACIDE, selon l'échelle de Cohen

	A	C	I	D	E
A		Très forte	Forte	Forte	Très forte
C	Très forte		Très forte	Très forte	Très forte
I	Forte	Très forte		Très forte	Très forte
D	Forte	Très forte	Très forte		Très forte
E	Très forte	Très forte	Très forte	Très forte	

Tableau 3.17 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel des transitions ACIDE. Modalités croisées.

		A	A	A	A	C	C	C	C	I	I	I	I	D	D	D	D	E	E	E	E
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
A	1					0,2			-0,26	0,29			-0,23				-0,18			-0,2	-0,19
A	2						0,2		-0,24			0,21					-0,26				-0,21
A	3					-0,18				-0,21					-0,18					0,29	
A	4							-0,23	0,57				0,44				0,32		-0,2	-0,19	0,74
C	1	0,2		-0,18						0,51			-0,26	0,2	0,39		-0,26		0,21	-0,28	-0,18
C	2		0,2													0,25					-0,18
C	3				-0,23						0,27		-0,22								-0,23
C	4	-0,26	-0,24		0,57					-0,23	-0,2		0,56	-0,18	-0,2	-0,23	0,47	-0,18	-0,27		0,64
I	1	0,29		-0,21		0,51			-0,23					0,49	0,32	-0,3	-0,3	0,38		-0,32	-0,2
I	2							0,27	-0,2							0,21	-0,21				
I	3		0,21											-0,22					0,42		-0,24
I	4	-0,23			0,44	-0,26		-0,22	0,56					-0,23	-0,25		0,57	-0,23	-0,4	0,18	0,58
D	1					0,2			-0,18	0,49		-0,22	-0,23					0,39		-0,25	
D	2			-0,18		0,39			-0,2	0,32			-0,25					0,22		-0,27	
D	3						0,25		-0,23	-0,3	0,21							-0,23	0,22		
D	4	-0,18	-0,26		0,32	-0,26			0,47	-0,3	-0,21		0,57					-0,23	-0,33	0,26	0,39
E	1								-0,18	0,38			-0,23	0,39	0,22	-0,23	-0,23				
E	2				-0,2	0,21			-0,27			0,42	-0,4			0,22	-0,33				
E	3	-0,2		0,29	-0,19	-0,28				-0,32			0,18	-0,25	-0,27		0,26				
E	4	-0,19	-0,21		0,74	-0,18	-0,18	-0,23	0,64	-0,2		-0,24	0,58				0,39				

Nous observons donc des associations fortes à très fortes entre les 5 processus de transition ACIDE (cf. *Tableau 3.16*).

De plus, nous constatons que les phases avancées des transitions sont associées positivement entre elles, tandis qu'il en est de même pour les phases les moins avancées.
Nous en déduisons donc que l'ensemble des 5 transitions ACIDE se déroulent plus ou moins simultanément dans le même sens.

Etant donné que les 5 processus transitionnels évoluent plus ou moins concomitamment, Il n'y a donc pas de phénomène d'antagonisme entre les transitions : l'évolution des uns n'empêche pas celle des autres. C'est même le contraire qui est observé.

Par-contre il y a des phénomènes limitatifs : pour atteindre les phases avancées dans une transition, il semble être nécessaire d'avoir franchi les phases « initiales » dans d'autres transitions. C'est la conclusion que nous tirons des associations croisées négatives entre les phases avancées de chaque transition et les phases « initiales » ou « intermédiaires » de certaines transitions. Ainsi la phase 4 économique (phase la plus avancée actuellement), a un certain degré d'incompatibilité avec les phases 1 et 2 anthropologiques, les phases 1, 2 et 3 cognitives, et les phases 1 et 3 institutionnelles.

10.Conclusion

L'approche quantitative et multidisciplinaire de l'histoire permet une lecture quantitative des évènements historiques par rapport aux tendances de temps long, et de faire de la prospective à long terme. Nous utilisons ici notre grille multidisciplinaire anthropologique, cognitive, institutionnelle, démographique et économique (ACIDE), associée aux notions de temps et d'âge transitionnels, pour inférer sur les trajectoires de temps long dans une approche en coupe transversale.

Dans ce chapitre, nous avons, grâce à notre démarche d'inférence longitudinale en coupe transversale, pu vérifier et confirmer la validité contemporaine de la transition économique à un niveau global (mondial). Cela peut répondre aux inquiétudes récentes sur le phénomène de désindustrialisation précoce (McMillan et Rodrik, 2014).
De même nous avons pu valider l'actualité de la transition démographique à l'échelle globale. Cela répond également aux inquiétudes d'une supposée surpopulation mondiale qui pourrait avoir une incidence négative sur la mortalité, l'espérance de vie et l'économie.

Les analyses concernant les trois autres domaines ACIDE (i.e. anthropologie, cognition, institutions) pour lesquels il n'existe pas de théories transitionnelles quantitatives établies, nous ont conduits à proposer trois nouvelles théories quantitatives transitionnelles respectivement en anthropologie, en institutions politiques internes et en cognition.
Ces contributions pourront être utilisées à des fins de reconstitution historiques, mais aussi prospectivistes pour les pays émergents ou en développement.

Enfin, notre approche transversale multi-transitionnelle nous a permis de montrer la très forte dépendance entre les dynamiques et transitions ACIDE au travers des concommitances et incompatibilités détectées entres les phases des différentes transitions.
Non seulement cela éclaire l'histoire globale, mais cela donne des pistes en termes de priorisation et d'ordonnancement des différentes composantes des politiques économiques, sociales et éducatives.

Les sections techniques qui suivent synthétisent les principaux résultats

Domaines ACIDE avec théories transitionnelles préexistantes.

Economie
La transition économique moderne est typiquement le passage d'une économie agraire, vers une économie de service. Avec au passage, une poussée transitoire industrielle (croissante, puis décroissante de la population active).
Choix du temps transitionnel
Le taux d'emplois agricoles lissé sur 5 ans peut empiriquement être utilisé comme symétrique du temps transitionnel.
Un âge transitionnel, se compte en unités de temps transitionnel et donc peut se définir par : 100 - (taux d'emploi agricole).
Comparaison du schéma transitionnel par coupe transversale avec le schéma théorique longitudinal de la transition économique
L'analyse transversale transitionnelle économique, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014. Les phases 1, 2, 3 et 4 contiennent respectivement 16%, 36%, 32% et 17% des pays. Par identification, ces 4 phases correspondent respectivement aux phases théoriques « agricole » ou parfois « démarrage », « démarrage » ou parfois « expansion », « achèvement », « achèvement » ou parfois « tertiaire ».

En comparant le schéma transitionnel transversal obtenu avec le schéma transitionnel longitudinal théorique de Fourastié, car nous observons la même dynamique des trois variables de transition.
L'étude transversale confirme assez clairement la validité et la réalité contemporaine du processus de transition économique théorique de Fourastié.

Démographie
Le concept de transition démographique désigne le passage d'un régime relativement stable avec fécondité et mortalité « élevées », vers un régime à fécondité et mortalité nettement plus faibles. Cette transition est initiée par une chute de la mortalité, suivie quelques décennies plus tard par celle de la fécondité.
Choix du temps transitionnel
Le taux de mortalité infantile lissé sur 5 ans peut empiriquement être utilisé comme temps transitionnel, dans une étude en coupe transversale.
Comparaison du schéma transitionnel par coupe transversale avec le schéma théorique longitudinal de la transition démographique
L'analyse transversale transitionnelle démographique, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014. Les phases 1, 2, 3 et 4 contiennent respectivement 13%, 16%, 43% et 28% des pays.
En comparant le schéma transitionnel transversal selon le temps transitionnel avec le schéma transitionnel longitudinal théorique, nous constatons une inclusion du schéma transversal contemporain, dans le schéma théorique complet, car le premier schéma s'identifie avec la phase 2 et le début de la post-transition du second schéma.
Ainsi, actuellement nous pouvons avancer qu'il n'y a plus de pays en période pré-transitoire, et seulement quelques rares pays en phase 1 théorique de la transition démographique.
L'étude réalisée confirme la validité et la réalité contemporaine du processus théorique de transition démographique.

Domaines ACIDE sans théories transitionnelles établies.

Anthropologie
L'anthropologie culturelle couvre les habitudes acquises par les êtres humains en société.
Variables de transition
En ajoutant une dimension inégalitaire au modèle anthropologique Inglehart-Welzel, nous proposons les variables suivantes pour caractériser la dynamique anthropologique : valeurs laïques-rationnelles (vs valeurs traditionnelles), valeurs d'expression personnelle (vs valeurs de survie), valeurs inégalitaires (mesurées par l'indice de Gini après taxes et redistribution, à très forte corrélation avec les variables anthropologiques).
Choix du temps transitionnel
Le ratio d'années d'études femmes vs hommes de 25-34 ans (lissé sur 5 ans) est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale.
Proposition de schéma transitionnel anthropologique selon l'indice d'âge transitionnel
Nous désignons par transition anthropologique (ou anthropo-culturelle), une série de profondes et durables transformations anthropologiques de la population d'un pays, pouvant s'étaler sur quelques décennies à plusieurs générations.
L'analyse transversale transitionnelle anthropologique, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014. Les phases 1, 2, 3 et 4 contiennent respectivement 50%, 23%, 17% et 10% des pays.
Le schéma transitionnel théorique obtenu révèle (en moyenne) que : Avec des niveaux en baisse dans les 3 premières phases, les valeurs d'« expression personnelle » (vs valeurs de « survie »), font un fort rebond vers un niveau sans précédent en phase 4.

Après une brève baisse « initiale », les valeurs « laïques-rationnelles » (vs valeurs « traditionnelles ») sont en progrès continu de la phase 2 à la phase 4. Nous notons toutefois une nette inflexion de la croissance de la phase 3 à la phase 4.
Quant aux valeurs « inégalitaires » (mesurées par l'indice de Gini après taxes et redistribution), elles suivent dans une première séquence une évolution en cloche suivi d'un léger rebond en phase 4.
L'allure des courbes transitionnelles est sensiblement la même, aussi bien en utilisant les moyennes que les médianes des variables transitionnelles : Cela montre la robustesse du schéma transitionnel obtenu en coupe transversale.

Institutions
Les institutions économiques désignent les règles de l'activité économique.
Variables transitionnelles
En nous restreignant au sous-domaine « politique interne » de notre matrice institutionnelle IEPE, nous proposons de suivre la dynamique institutionnelle à travers les 2 variables suivantes : les « libertés civiles », et l'indice global de démocratie.
Choix du temps transitionnel
L'indice de développement humain (lissé) est une variable qui peut empiriquement être utilisée comme temps transitionnel, dans une étude en coupe transversale.
L'indice de développement humain mesure les résultats des dispositions institutionnelles dans les domaines politiques, économiques, sanitaires et éducatives. Cet indice est donc légitime pour caractériser les institutions d'un pays à un moment donné.
Proposition de schéma transitionnel institutionnel selon l'indice d'âge transitionnel
Nous désignons par transition institutionnelle, une série de profondes et durables transformations institutionnelles dans un pays, pouvant s'étaler sur quelques décennies à plusieurs générations.
L'analyse transversale transitionnelle institutionnelle, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2014. Les phases 1, 2, 3 et 4 contiennent respectivement 21%, 16%, 35% et 28% des pays.
Le schéma institutionnel transitionnel théorique obtenu révèle (en moyenne) que : L'indice global de démocratie présente une forte baisse « initiale » phase 1, suivi d'un fort rebond en phase 2, qui se termine par une inflexion « finale » de la hausse entre les phases 3 et 4.
Quant à la répression des libertés civiles, son parcours est sans surprise, l'inverse de celui de l'indice global de démocratie : forte hausse de la phase 1 à la phase 2, suivi d'une forte baisse dès la phase 2, qui se termine par une baisse « finale » légèrement tassée entre les phases 3 et 4
Enfin, l'indice de développement humain, qui synthétise le fonctionnement opérationnel des institutions sanitaires, économiques et éducatives, est en hausse continue sur les 4 phases.
L'allure des courbes transitionnelles est sensiblement la même, aussi bien en utilisant les moyennes que les médianes des variables transitionnelles : Cela montre la robustesse du schéma transitionnel obtenu en coupe transversale.

Cognition
Variables de transition
nous caractérisons la dynamique cognitive par les variables suivantes : Le taux d'adultes de 25-64 ans, non scolarisés ; le taux d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint ; le taux d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint ; Compétences cognitives numériques-littéraires-logiques (mesurées par le QI moyen).
Choix du temps transitionnel
Nous utilisons empiriquement, le nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé) comme temps transitionnel, dans l'étude en coupe transversale.

Proposition de schéma transitionnel cognitif selon l'indice d'âge transitionnel

Nous désignons par transition cognitive, une série de profondes et durables transformations cognitives de la population d'un pays, pouvant s'étaler sur quelques décennies à plusieurs générations.

L'analyse transversale transitionnelle cognitive, nous a permis de regrouper les pays en 4 classes d'âge transitionnel (ou phases) en 2009-2012. Les phases 1, 2, 3 et 4 contiennent respectivement 16%, 16%, 49% et 19% des pays.

Ainsi, selon le schéma théorique de transition cognitive proposé, les compétences cognitives d'un pays augmentent de manière continue en moyenne de 74 à 98.

Parallèlement, les adultes uniquement diplômés du primaire, représentent la majorité des adultes alphabétisés diplômés durant les deux premières classes d'âge transitionnel, avec un pic (à 30%) dans la deuxième classe.

A partir de la deuxième classe d'âge transitionnel, les adultes dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint, représentent la majorité des adultes alphabétisés diplômés. Ces diplômés du secondaire passent de 11% en classe transitionnelle 1, à 44% en classe transitionnelle 4. Et cela, tandis que les diplômés exclusifs du primaire décroissent en pourcentage (dans les classes transitionnelles 3 et 4), pour finir à 7% des diplômés adultes en classe transitionnelle 4.

Quant aux diplômés du supérieur (tertiaire), leur proportion parmi les adultes augmente continument de la classe transitionnelle 1 (à 3%) à la classe transitionnelle 4 (à 21%).

L'allure des courbes transitionnelles est sensiblement la même, aussi bien en utilisant les moyennes que les médianes des variables transitionnelles : Cela montre la robustesse du schéma transitionnel obtenu en coupe transversale.

Synthèse des associations entre transitions ACIDE

Les associations sont d'intensité fortes à très fortes entre les 5 transitions ACIDE.

De plus, nous avons constaté que les phases avancées des transitions sont associées positivement entre elles, tandis qu'il en est de même pour les phases les moins avancées.

Nous en déduisons donc que l'ensemble des 5 transitions ACIDE se déroulent plus ou moins simultanément dans le même sens.

Ainsi, l'évolution des uns n'empêche pas celle des autres. C'est même le contraire qui est observé.

Par-contre il y a des phénomènes limitatifs : pour atteindre les phases avancées dans une transition, il peut être nécessaire d'avoir franchi les phases « initiales » ou « intermédiaires » dans d'autres transitions.

Ainsi la phase 4 économique (phase la plus avancée actuellement), a un certain degré d'incompatibilité avec les phases 1 et 2 anthropologiques, les phases 1, 2 et 3 cognitives, et les phases 1 et 3 institutionnelles.

Perspectives

Une suite logique serait de tenter d'établir une séquence de phases multi-transitionnelles ACIDE, ce qui permettrait de détecter d'éventuels décalages dans les évolutions transitionnelles.

Pour mener à bien cette étude, nous avons dû exclure les pays démographiquement ou géographiquement très « petits » (au sens de la Banque Mondiale), car ils sont particulièrement vulnérables aux chocs exogènes (telles les catastrophes naturelles), avec des opportunités économiques limitées, une migration importante, et des contraintes de capacité.

Nous avons également exclu pour l'analyse institutionnelle, les pays dotés de fortes ressources naturelles énergétiques (OPEP) à cause de particularités institutionnelles souvent induites par ces rentes.

Il sera donc particulièrement intéressant que des recherches complémentaires puissent investiguer les processus transitionnels ACIDE spécifiques aux « petits » pays et aux pays particulièrement riches en ressources naturelles. Ce qui permettrait d'enrichir les théories transitionnelles ACIDE.

Les classes d'âge transitionnel détectées illustrent également les phases observables « actuellement » dans les processus de transitionnels ACIDE.
La recherche d'éventuelles classes ayant pu exister historiquement et n'étant plus observable aujourd'hui, pourra se faire dans une approche longitudinale des transitions ACIDE.

Quant aux probables phases futures jamais observées, seules des études prospectivistes permettraient d'en spéculer les caractéristiques.

Chapitre 4.
Prospective actuarielle et cliométrique des taux de mortalité

1. Table des matières

Chapitre 4. Prospective actuarielle et cliométrique des taux de mortalité 302

1. Table des matières 303

2. Introduction 304

3. Relation entre le modèle Lee-Carter et la régression PCR 307

3.1. Présentation du modèle de Lee-Carter 307

3.2. Estimation des paramètres du modèle 307

3.3. Modèle Lee-Carter, comme régression sur la 1ère composante principale de Z 310

4. Modèle CBD 311

4.1. Présentation du modèle CBD 311

4.2. Estimation des paramètres $\mathrm{ki(t)}$ 311

5. Régression sur composantes principales 312

5.1. Définition 312

5.2. Principe 312

6. Modèle Logit-PCR 315

7. Amélioration du modèle Logit-PCR 316

8. Régression PLS 318

9. Améliorations du modèle Logit-PCR en utilisant la régression PLS 320

9.1. Amélioration du modèle Logit-PCR : Modèle Logit-PLS 320

10. Variable et modèle PCR cliométrique 321

10.1. Concept de variable cliométrique 321

10.2. Méthodologie de création d'une variable cliométrique 321

10.3. Amélioration du modèle Logit-PCR, en y intégrant la variable cliométrique 328

11. Résultats empiriques 329

11.1. Modèle non-cliométrique gaussien : logit PCR optimal 329

11.2. Modèles cliométriques 340

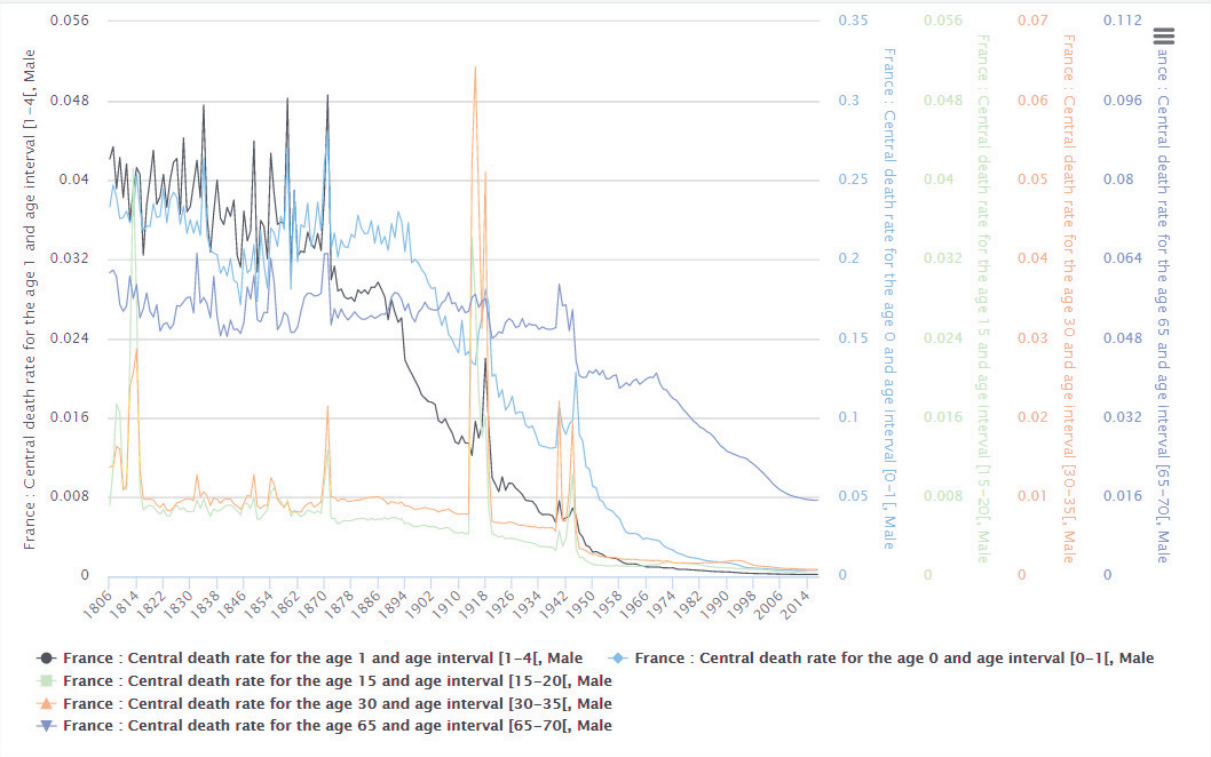
14 Conclusion 355

15 Bibliographie 359

2. Introduction

La transition démographique moderne est le processus historique par lequel une population passe d'un régime démographique caractérisé par un taux de mortalité global et un taux de natalité élevés, à un nouveau régime caractérisé par des valeurs nettement plus faibles (division par 3, voire 4) de ces deux taux.
Nous remarquons empiriquement des décalages temporels et des différences d'allure dans les courbes des taux de mortalité selon les âges (cf. *Figure 3.44*). Si une composante temporelle unique reflète la transition démographique du point de vue de la mortalité globale, la prise en compte de la transition sanitaire qui affecte les âges de manière différenciée, nécessite des composantes temporelles différenciées elles aussi par tranche d'âge.

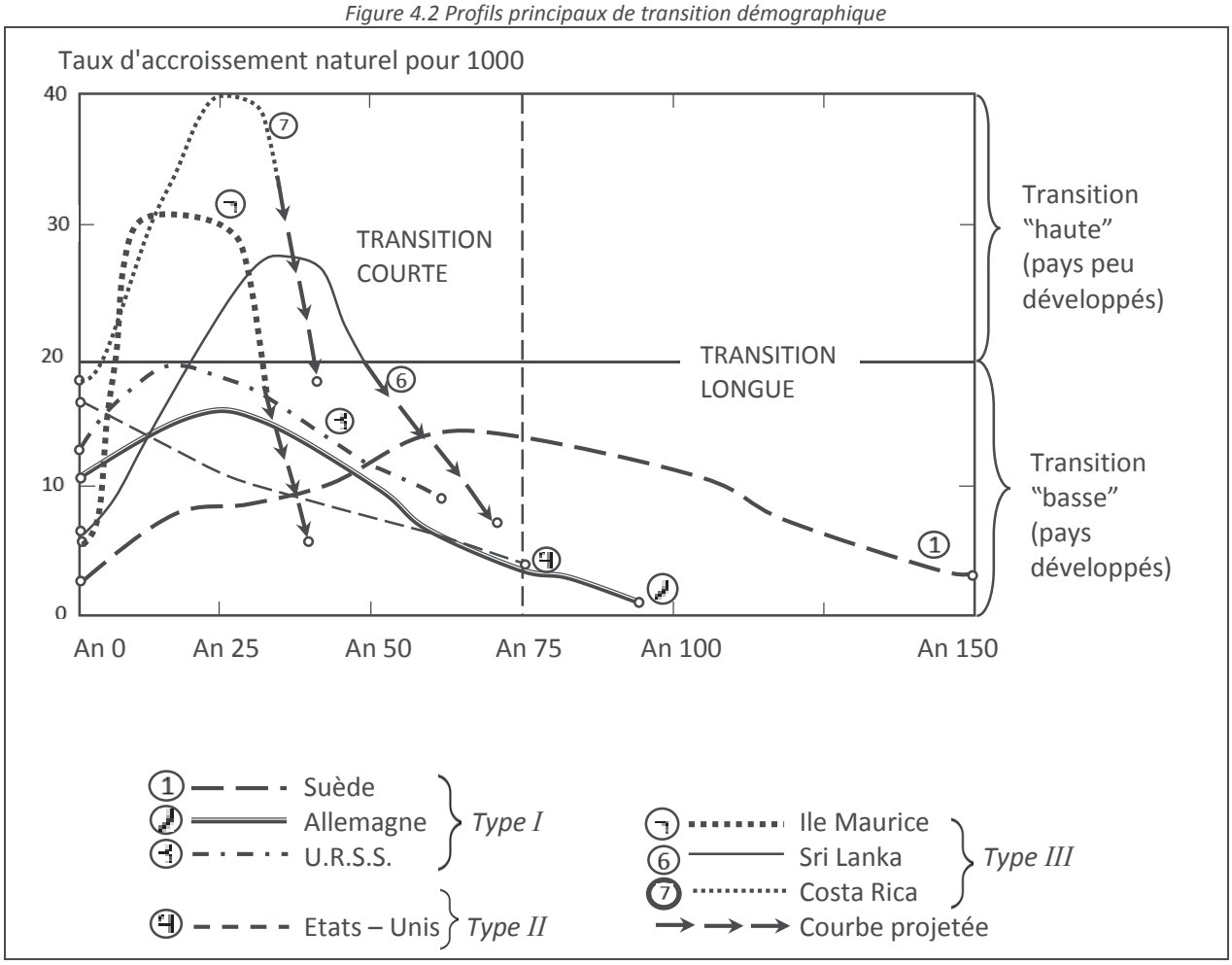
Figure 4.1 Décalages temporels et différences d'allure des taux mortalité selon les tranches d'âge. France



La première originalité de notre approche consiste à utiliser le constat ci-dessus pour généraliser les modèles de mortalité classiques (notamment les modèles de Lee-Carter et CBD), en sélectionnant de manière optimale (minimisation du *MSE* de validation croisée), une ou plusieurs composantes principales (régression PCR) ou PLS (régression PLS), pour chaque âge modélisé. Nous appellerons ces extensions, des modèles respectivement, de type PCR-optimal et PLS-optimal.

Le second constat historique est celui d'une transition démographique universelle, se déroulant par vagues (ou cohortes) de pays, avec des rythmes des changements transitionnels qui sont assez variables d'une cohorte de pays à l'autre. En effet, les pays récemment entrés en transition ont généralement des rythmes de changement plus rapides (cf. *Figure 4.2*).

Figure 4.2 Profils principaux de transition démographique



Source: Chesnais (1986), p. 270

Certains pays sont donc plus « *précoces* » dans la transition démographique que d'autres qui sont alors considérés comme étant en situation de « *rattrapage* » relatif.

Une originalité de ce chapitre consiste à utiliser l'histoire quantitative des mortalités des pays transitionnellement « plus âgés », pour tenter d'améliorer les prévisions à long terme de mortalité dans les pays transitionnellement « moins âgés » (cf. notion d'âge transitionnel défini au chapitre 1).
En pratique, ces améliorations pourraient s'appliquer aux études prospectives de la mortalité dans les pays émergents ou en développement.

Nous avions défini le *temps transitionnel* (cf. chapitre 1) comme étant une série temporelle qui est *fortement* concordante avec le temps. Autrement dit, un *temps transitionnel*, est une série temporelle qui a une dépendance monotone *forte* avec le temps. Nous notons également que toute transformée symétrique d'une série temporelle discordante au temps, est concordante au temps, donc peut jouer le rôle de temps transitionnel.
De même, nous avons défini l'âge transitionnel d'un pays, comme étant différence entre le temps transitionnel et sa valeur minimale (quel que soit le pays) dans la période de transition. Cette valeur minimale est un seuil de passage, quel que soit le pays.

Dans ce contexte de transition démographique par cohortes de pays, avec des rythmes différents de changement, notre démarche consiste à utiliser comme temps transitionnel, la symétrique (par rapport à l'axe horizontal du temps) du taux de mortalité infantile (éventuellement lissé pour accentuer la monotonie et la linéarité).

Les différences des rythmes du processus de transition démographique entre les pays de cohortes distinctes, ne permet pas de faire des prévisions en utilisant des modèles classiques de séries temporelles avec variable exogène retardé, car le retard n'est pas stable dans le temps.
La variable de temps transitionnel, permet de faire un appariement temporel entre le pays à modéliser et le pays transitionnellement « plus âgé».

La variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est une variable créée à partir des données historiques du pays A (transitionnellement « plus âgé » que le pays R), pour jouer un rôle de variable exogène dans les modèles de prévisions de mortalité du pays R.
Cette variable cliométrique exogène de mortalité sera construite de sorte à avoir la même allure temporelle que dans les pays « transitionnellement plus âgé», mais adaptée au rythme du pays à modéliser.
La variable cliométrique sera ainsi utilisée comme variable explicative supplémentaire dans les modèles prospectifs de mortalité du pays R. Nous nommons les modèles utilisant la variable cliométrique, modèles cliométriques : nous en proposons deux types qui sont les modèles PCR-optimal-cliometric et PLS-optimal-cliometric.

Pour les travaux empiriques, nous utilisons dans ce chapitre, comme modèles classiques de référence, les modèles suivants : Lee-Carter, Lee-Carter Log-Poisson, CBD-logit-binomial.

Notre modèle généralisé et optimisé de type PCR-optimal sera testé en contre les modèles classiques de référence, en France et en Italie. L'horizon de prévision sera de 50 ans.

Les modèles cliométriques de prévisions de long terme destiné *a priori* aux pays émergents ou en développement, sera testé pour Inde. L'horizon de prévision testé sera de 40 ans

3. Relation entre le modèle Lee-Carter et la régression PCR

Le modèle de Lee et Carter (1992) est un des modèles de classiques que nous allons modèle de référence dont nous allons comparer les performances avec celles des modèles que nous proposons.
Afin d'en proposer une généralisation, nous étudions et formalisons ici les relations entre le modèle classique de Lee-Carter et la régression sur composantes principales *(principal component regression, PCR)*.

3.1. Présentation du modèle de Lee-Carter

Le modèle de Lee et Carter (1992) s'intéresse à la variable $m_{x,t}$ qui représente le taux de mortalité central entre les âges x et $x+1$ mesuré au cours de l'année t. Le modèle de Lee-Carter, peut s'écrire comme suit (Boyer et al. [2015], p. 536):

$ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$
Où :

- a_x correspond à la moyenne du logarithme des taux de mortalité par âge calculée sur l'ensemble de la période considérée,
- k_t représente la tendance générale du niveau de mortalité au cours du temps,
- b_x est la sensibilité au facteur commun k_t spécifique à chaque âge $\left(\frac{d\, ln(m_{x,t})}{dt} = b_x \frac{dk_t}{dt}\right)$,
- $\varepsilon_{x,t}$ est le résidu du modèle, supposé *i.i.d.*, d'espérance nulle et de variance σ_ε^2

3.2. Estimation des paramètres du modèle

L'estimation des paramètres s'effectue par le critère des moindres carrés ordinaires et par la décomposition en valeurs singulières.

- Estimation de a_x par un critère des moindres carrés ordinaires (Planchet, Lelieur [2007]; Safitri *et al* [2018]) :

$$(\hat{a}_x, \hat{b}_x, \hat{k}_t) = \underset{a_x, b_x, k_t}{\arg\ min} \sum_{x,t} \left(ln(m_{x,t}) - a_x - b_x k_t\right)^2$$

Sous contrainte que $\sum_{x=x_m}^{x_M} b_x^2 = 1$ ou $\sum_{x=x_m}^{x_M} b_x = 1$ et $\sum_{t=t_m}^{t_M} k_t = 0$,
avec :
$x_m = min(x)$
$x_M = max(x)$
$t_m = min(t)$
$t_M = max(t)$

Il convient donc de résoudre le problème d'optimisation :

$$\frac{\partial \sum_{x,t}\left(ln(m_{x,t}) - a_x - b_x k_t\right)^2}{\partial a_x} = 2(t_M - t_m + 1)a_x - 2\sum_{t=t_m}^{t_M}\left(ln(m_{x,t}) - b_x k_t\right) = 0$$

$$= 2(t_M - t_m + 1)a_x - 2\sum_{t=t_m}^{t_M} ln(m_{x,t}) + 2b_x \sum_{t=t_m}^{t_M} k_t = 0$$

Or $\sum_{t=t_m}^{t_M} k_t = 0 \Rightarrow (t_M - t_m + 1)a_x = \sum_{t=t_m}^{t_M} ln(m_{x,t})$

$$\hat{a}_x = \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} ln(m_{x,t})$$

- Estimation de b_x et k_t

L'estimation de b_x et k_t se fait via le premier terme de la décomposition en valeurs singulières de la matrice $Z = ln(m_{x,t}) - \hat{a}_x$

Pour obtenir une meilleure visualisation de la composante temporelle k_t , nous modifions légèrement la présentation classique en mettant les informations temporelles en lignes et celle liées à l'âge en colonnes :

$$m_{x,t} = \begin{bmatrix} m_{x_m,t_m} & \cdots & m_{x_M,t_m} \\ \vdots & \ddots & \vdots \\ m_{x_m,t_M} & \cdots & m_{x_M,t_M} \end{bmatrix}$$

$$Z = \begin{bmatrix} Z_{x_m,t_m} = ln(m_{x_m,t_m}) - \hat{a}_{x_m} & \cdots & Z_{x_M,t_m} = ln(m_{x_M,t_m}) - \hat{a}_{x_M} \\ \vdots & \ddots & \vdots \\ Z_{x_m,t_M} = ln(m_{x_m,t_M}) - \hat{a}_{x_m} & \cdots & Z_{x_M,t_M} = ln(m_{x_M,t_M}) - \hat{a}_{x_M} \end{bmatrix}$$

Z est d'ordre $(t_M - t_m + 1, x_M - x_m + 1)$.

L'approximation de Z donnant $Z \approx \widehat{k_t}\hat{b}_x^T$ (ou encore $\hat{Z} = \widehat{k_t}\hat{b}_x^T$) découle de la formule du modèle de Lee-Carter :

$Z = ln(m_{x,t}) - \hat{a}_x = \widehat{k_t}\hat{b}_x^T + \hat{\varepsilon}_{x,t}$,

avec :

$\hat{b}_x^T = (\hat{b}_{x_m}, \dots, \hat{b}_{x_M})$ et $\widehat{k_t}^T = (\hat{k}_{t_m}, \dots, \hat{k}_{t_M})$.

Soit u_i le i^{ème} vecteur propre normé de ZZ^T de dimension $(t_M - t_m + 1) \times (t_M - t_m + 1)$ correspondant à la valeur propre λ_i on a :

$ZZ^T u_i = \lambda_i u_i$, avec $u_i^T u_i = 1$

En multipliant les deux membres de la relation précédente par Z^T, on a :

$Z^T Z Z^T u_i = Z^T \lambda_i u_i$

$(Z^T Z) Z^T u_i = Z^T \lambda_i u_i$

$(Z^T Z) Z^T u_i = \lambda_i Z^T u_i \Rightarrow$ $Z^T u_i$ est un vecteur propre de $Z^T Z$ associé à la même valeur propre λ_i, mais n'est pas normé car :

$(Z^T u_i)^T (Z^T u_i) = {u_i}^T Z Z^T u_i = {u_i}^T \lambda_i u_i = \lambda_i$, car ${u_i}^T u_i = 1$

Ou encore : $\|Z^T u_i\| = \sqrt{\lambda_i}$

Soit v_i le i^{ème} vecteur propre normé ($v_i^T v_i = 1$) de $Z^T Z$ d'ordre $(x_M - x_m + 1) \times (x_M - x_m + 1)$

Donc : $v_i = \frac{1}{\sqrt{\lambda_i}} Z^T u_i$ (formule de transfert)

On a :

$(Z^T Z) v_i = \lambda_i v_i$

En multipliant à gauche par Z , on a :
$(ZZ^T)Zv_i = \lambda_i Zv_i$

Zv_i est un vecteur propre non-normé de ZZ^T associé à la valeur propre λ_i (comme u_i sauf que u_i est normé) car :
$(Zv_i)^T(Zv_i) = {v_i}^T(Z^TZv_i) = {v_i}^T\lambda_i v_i = \lambda_i$, car ${v_i}^T v_i = 1$
Ou encore : $\|Zv_i\| = \sqrt{\lambda_i}$
Donc $u_i = \frac{1}{\sqrt{\lambda_i}} Zv_i$ (2ième formule de transfert)

La décomposition en valeurs singulières (SVD) de Z est donnée par :
$Z = U\Sigma V^T$
Avec
$U = [u_1 \dots u_{t_M - t_m + 1}]$

$$V^T = \begin{pmatrix} v_1^T \\ \dots \\ v_{x_M - x_m + 1}^T \end{pmatrix}$$

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$
D est d'ordre $(t_M - t_m + 1, x_M - x_m + 1)$ et est la matrice des valeurs singulières non-nulles de Z

Donc : $Z = \sum_{i \geq 1} \sqrt{\lambda_i}\, u_i v_i^T$

Puisque la première valeur propre λ_1 est bien supérieure aux autres, la meilleure approximation de Z est alors donnée par :
$Z \approx \sqrt{\lambda_1} u_1 v_1^T$

En confrontant la relation $Z \approx \widehat{k_t}\widehat{b_x}^T$ et $Z \approx \sqrt{\lambda_1} u_1 v_1^T$, on a : $\widehat{k_t}\widehat{b_x}^T = \sqrt{\lambda_1} u_1 v_1^T$.
En tenant compte des ordres de $\widehat{k_t}$ et $\widehat{b_x}$ et de la contrainte ${b_x}^T b_x = 1$, nous obtenons par identification :
$\widehat{b_x} = v_1$ et $\hat{k}_t = \sqrt{\lambda_1} u_1$

Si nous avions utilisé la contrainte alternative $\sum_{x=x_m}^{x_M} b_x = 1$, nous aurions obtenu le résultat ci-après :

En multipliant et divisant le second membre de la relation $\widehat{k_t}\widehat{b_x}^T = \sqrt{\lambda_1} u_1 v_1^T$ par $\sum_j v_{1j}$, on a :

$$\widehat{k_t}\widehat{b_x}^T = \sqrt{\lambda_1} u_1 v_1^T \times \frac{\sum_j v_{1j}}{\sum_j v_{1j}}$$

$$\widehat{k_t}\widehat{b_x}^T = \frac{v_1^T}{\sum_j v_{1j}} \times \sqrt{\lambda_1} \sum_j v_{1j} u_1$$

De ce qui précède, on a par identification :
$\widehat{b_x} = \frac{v_1}{\sum_j v_{1j}}$ et $\hat{k}_t = \sqrt{\lambda_1} \sum_j v_{1j} u_1$
avec $\sum_j v_{1j} \neq 0$.

3.3. Modèle Lee-Carter, comme régression sur la 1ère composante principale de Z

Si z_1^* est la première composante principale de Z , alors $z_1^* = Zv_1$, (Saporta, 2011, p. 168)
Avec v_1 premier vecteur propre normé de $Z^T Z$
Nous avions établi à la section précédente que sous la contrainte ${b_x}^T b_x = 1$, $\hat{k}_t = \sqrt{\lambda_1} u_1$

Or d'après la 2ème formule de transfert établi à la section précédente : $Zv_1 = \sqrt{\lambda_1} u_1$

Donc :
$\hat{k}_t = \sqrt{\lambda_1} u_1 = Zv_1 = z_1^*$

Donc $\hat{k}_t$ est égal à la 1ère composante principale de Z (z_1^*) et donc par conséquent le modèle de Lee-Carter est une régression (avec un terme constant) sur z_1^* :
$ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} = a_x + b_x z_1^* + \varepsilon_{x,t}$

Si nous avions utilisé la contrainte alternative $\sum_{x=x_m}^{x_M} b_x = 1$, nous aurions obtenu le résultat ci-après :
$\hat{k}_t = \sqrt{\lambda_1} \sum_j v_{1j} u_1 = Zv_1 \sum_j v_{1j} = z_1^* \sum_j v_{1j}$

Donc dans ce cas alternatif, $\hat{k}_t$ est égal à la 1ère composante principale de Z (notée z_1^*) à facteur multiplicatif ($\sum_j v_{1j}$) près et donc par conséquent le modèle de Lee-Carter reste une régression (avec un terme constant) sur z_1^* :
$ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} = a_x + b_x (z_1^* \sum_j v_{1j}) + \varepsilon_{x,t}$

4. Modèle CBD

Le modèle CBD est un des modèles de classiques que nous allons modèle de référence dont nous allons comparer les performances avec celles des modèles que nous proposons. Nous en faisons ici un exposé succinct.

4.1. Présentation du modèle CBD

Le modèle CBD (inspiré ici de Boyer et al. [2015], p. 538) est basé sur la transformation logistique du taux de mortalité. Il considère les réalisations des probabilités de décès selon la structure suivante :

$\text{logit}\, q_x\,(t) = \left(log\, \frac{q_x(t)}{1-q_x(t)}\right) = k_1(t) + k_2(t)(x-\bar{x}) + \varepsilon_x(t)$,

avec :

- $q_x(t) = 1 - \exp(-m_{x,t})$, estimation de la probabilité qu'un individu d'âge x décède à l'an t (avant d'atteindre l'âge $x+1$)
- $k_1(t)$ est un paramètre lié au temps. Il correspond à l'évolution temporelle moyenne du logit de la probabilité de mortalité pour l'âge moyen $\bar{x}$.

$$E\big(\text{logit}\, q_x\,(t)\big) = E\big(k_1(t) + k_2(t)(x-\bar{x}) + \varepsilon_x(t)\big)$$

$$= k_1(t) + k_2(t)E(x-\bar{x}) + E\big(\varepsilon_x(t)\big)$$

$E\big(\text{logit}\, q_x\,(\text{t})\big) = k_1(t)$, car $E(x-\bar{x}) = 0$ et $\big(\varepsilon_x(t)\big) = 0$.

- $k_2(t)$ est le second paramètre lié au temps. Il représente la sensibilité de la mortalité par rapport à l'âge x.
- $\bar{x}$ représente la moyenne des âges.
- $\varepsilon_x(t)$ sont les résidus du modèle de moyenne 0 et de variance σ^2.

4.2. Estimation des paramètres $k_i(t)$

Le modèle CBD est un modèle paramétrique avec $2(t_M - t_m + 1)$ coefficients, puisque t_m est l'année de début et t_M l'année de fin.

Pour estimer les $k_i(t)$ suffit de régresser linéairement pour chaque année t, le $\text{logit}\, q_x\,(t)$ sur l'âge centré $(x-\bar{x})$: la constante et la pente fournissent des estimateurs de $k_1(t)$ et $k_2(t)$ au temps t.

Pour pouvoir simuler l'évolution de la mortalité à des dates futures, le modèle CBD suggère de modéliser $k_i(t)$ par une marche aléatoire multivariée avec dérive μ :

$k(t+1) = \mu + k(t) + C\eta(t+1)$, où

- μ est un vecteur constant 2×1
- C est une matrice triangulaire supérieure constante 2×2
- $\eta(t+1)$ est une variable aléatoire normale centrée réduite.

$$\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \qquad C = \begin{pmatrix}c_{11} & c_{12}\\ 0 & c_{22}\end{pmatrix} \qquad \eta = \begin{pmatrix}\eta_{1,t+1}\\ \eta_{2,t+1}\end{pmatrix} \qquad k(t+1) = \begin{pmatrix}k_1(t+1)\\ k_2(t+1)\end{pmatrix}$$

5. Régression sur composantes principales

La régression sur composantes principales (*Principal Component Regression, PCR*) a été introduite principalement pour gérer les problèmes de multicolinéarité entre variables explicatives. Nous en faisons ici un exposé car il nous sera utile pour proposer une généralisation à d'autres composantes principales du modèle de Lee-Carter.

5.1. Définition

La régression sur composantes principales (*Principal Component Regression, PCR*) est une analyse en régression sur les composantes d'une analyse en composantes principales ACP (Cornillon et al., 2011)
Soit :

$$Y = x_1\beta_1 + \cdots + x_p\beta_p + \varepsilon\,,$$

où

- $\beta_1, \dots, \beta_p$ sont des paramètres à estimer
- Y : variable expliquée, centrée réduite
- $x_1, \dots, x_p$ sont des variables explicatives, centrées réduites

Ce modèle peut être réécrit sous la forme :

$$Y = X\beta + \varepsilon \quad \text{avec} \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad \text{et} \quad X = \left[x_1, \dots, x_p\right].$$

La PCR est utilisée le plus souvent lorsque la matrice X n'est pas de plein rang ou lorsque $p > n$ avec n= nombre d'observations.

5.2. Principe

La PCR introduit des composantes issues d'une ACP à la place des variables explicatives (X) et trouve un modèle équivalent :
$Y = X^*\beta^* + \varepsilon$
X^* étant la matrice des composantes principales.

La méthode se déroule en trois (3) phases :

- Analyse en composantes principales sur X

La matrice $(X'X)$ est une matrice symétrique donc diagonalisable, et elle peut s'écrire sous la forme

$X'X = U \wedge U^T$ où U = Matrice des vecteurs propres normalisés de $(X'X)$ c'est-à-dire $(U^TU = UU^T = I)$.
$\wedge$ = diag$\left(\lambda_1, \lambda_2 \dots, \lambda_p\right)$ est une matrice diagonale des valeurs propres $\left(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p\right)$

Dans l'ACP, U est la matrice des axes principaux normés à l'unité et $\left\{\lambda_j\right\}$ j: 1, … , p sont les valeurs propres de $X'X$.

En écrivant : $X = XUU^T$ (car $UU^T = I$).
Alors, l'équation $Y = X\beta + \varepsilon$ devient : $Y = XUU^T\beta + \varepsilon$

$Y = X^*\beta^* + \varepsilon$, où $\beta^* = U^T\beta$ et $X^* = XU$

Les colonnes de X^* sont les p composantes principales : $x_j^* = Xu_j$.
Le modèle $Y = X^*\beta^* + \varepsilon$ est appelé modèle de régression sur les composantes principales.
${X^*}^T X^* = U^T X^T XU = U^T U \wedge U^T U = \wedge \implies$ les colonnes de X^* sont orthogonales entre elles et de variance λ_j

- Régression linéaire de la variable indépendante sur les composantes
  - $x_j^* = Xu_j$ étant la j^ième^ composante principale.
  - Régression sur la 1ère composante

Soit $x_1^* = Xu_1$ où u_1 est un vecteur unitaire engendré par la plus grande valeur propre λ_1
$Y = X_1^*\beta_{(1)}^* + \varepsilon$ est le modèle de régression sur la 1ère composante principale x_1^*
Utilisons la méthode des moindres carrés pour estimer $\beta_{(1)}^*$
Soit à minimiser $S(\beta_{(1)}^*)$ avec

$$S(\beta_{(1)}^*) = \left\|Y - x_1^*\beta_{(1)}^*\right\|^2$$

$$= Y^TY - Y^Tx_1^*\beta_{(1)}^* - \beta_{(1)}^{T^*}{x_1^*}^TY + \beta_{(1)}^{T^*}{x_1^*}^Tx_1^*\beta_{(1)}^*$$

$$\frac{\partial S(\beta_{(1)}^*)}{\partial \beta_{(1)}^*} = 0 \iff -Y^Tx_1^* - {x_1^*}^TY + 2{x_1^*}^Tx_1^*\beta_{(1)}^* = 0$$

$$\iff -2{x_1^*}^TY + 2{x_1^*}^Tx_1^*\beta_{(1)}^* = 0$$

$$\hat{\beta}_{(1)}^* = \left({x_1^*}^Tx_1^*\right)^{-1}{x_1^*}^TY$$

$\frac{\partial^2 S(\beta_{(1)}^*)}{\partial^2 \beta_{(1)}^*} = 2{x_1^*}^Tx_1^* > 0$ car ${x_1^*}^Tx_1^* = \|x_1^*\|^2$ donc $\hat{\beta}_{(1)}^*$ est bien un minimum strict.

- Régression sur k composantes ($k < p$).

$$Y = x_1^*\beta_1^* + \cdots + x_k^*\beta_k^* + \varepsilon$$

$\hat{\beta}^*(k) = \left({X_{[k,1]}^*}^TX_{[1,k]}^*\right)^{(-1)}{X_{[k,1]}^*}^TY$, de façon analogue à la régression sur une composante

Etudions la dépendance des paramètres estimés :

Soient $\hat{\beta}_i^*(k) = \left({x_i^*}^Tx_i^*\right)^{-1}{x_i^*}^TY$ et $\hat{\beta}_j^*(k) = \left({x_j^*}^Tx_j^*\right)^{-1}{x_j^*}^TY$

$$Cov(\hat{\beta}_i^*(k), \hat{\beta}_j^*(k)) = Cov\left[\left({x_i^*}^Tx_i^*\right)^{-1}{x_i^*}^TY, \left({x_j^*}^Tx_j^*\right)^{-1}{x_j^*}^TY\right]$$
$$= E\left[\left({x_i^*}^Tx_i^*\right)^{-1}{x_i^*}^TYY^Tx_j^*\left({x_j^*}^Tx_j^*\right)^{-1}\right]$$
$$= \left({x_i^*}^Tx_i^*\right)^{-1}{x_i^*}^TE(Y^TY)x_j^*\left({x_j^*}^Tx_j^*\right)^{-1}, \text{ car } Y^TY = YY^T$$

$$Cov(\hat{\beta}_i^*(k), \hat{\beta}_j^*(k)) = \left({x_i^*}^Tx_i^*\right)^{-1}{x_i^*}^TV(Y)x_j^*\left({x_j^*}^Tx_j^*\right)^{-1}, \text{ où } V(Y) = E(Y^TY) = Variance\ de\ Y$$

Si $i \neq j \Rightarrow {x_i^*}^Tx_j^* = 0$ car x_i^* et x_j^* sont orthogonaux d'où $Cov(\hat{\beta}_i^*(k), \hat{\beta}_j^*(k)) = 0$.

Si $i = j$,

$$Cov\left(\hat{\beta}_i^*(k), \hat{\beta}_j^*(k)\right) = \left({x_i^*}^T x_i^*\right)^{-1} {x_i^*}^T x_i^* V(Y) \left({x_i^*}^T x_i^*\right)^{-1}$$
$$= V(Y) \left({x_i^*}^T x_i^*\right)^{-1} \left({x_i^*}^T x_i^*\right) \left({x_i^*}^T x_i^*\right)^{-1}$$
$$= V(Y) \left({x_i^*}^T x_i^*\right)^{-1} \text{ , car } \left({x_i^*}^T x_i^*\right) \left({x_i^*}^T x_i^*\right)^{-1} = 1$$
$$= \sigma^2 \left({x_i^*}^T x_i^*\right)^{-1}$$

$$Cov\left(\hat{\beta}_i^*(k), \hat{\beta}_j^*(k)\right) = \sigma^2 \frac{1}{\lambda_i}, \text{ car } {x_i^*}^T x_i^* = \lambda_i$$

D'où :

$$cov\left(\hat{\beta}_{i(k)}^*, \hat{\beta}_{j(k)}^*\right) = \sigma^2 \begin{cases} \frac{1}{\lambda_i} & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

- Calcul des paramètres de la régression en fonction des variables d'origine

Pour tout k fixé, le paramètre $\hat{\beta}_{(k)}^*$ estimé en PCR est :

$$\hat{\beta}_{PCR}^*(k) = \left({X_{[k,1]}^*}^T X_{[1,k]}^*\right)^{(-1)} {X_{[k,1]}^*}^T Y^T$$

En utilisant la relation $\beta^* = U^T \beta$, on peut passer de $\hat{\beta}_{PCR}^*(k)$ à $\hat{\beta}_{PCR}(k)$:
Soit

$$\beta^* = U^T \beta$$

En multipliant chaque membre de $\beta^* = U^T \beta$ par U on a :

$$U\beta^* = UU^T \beta$$

$$U\beta^* = \beta \quad car\ (UU^T = I)$$

On a alors $\hat{\beta}_{PCR}(k) = U_{[1:k]} \hat{\beta}_{PCR}^*(k)$

6. Modèle Logit-PCR

Nous avions démonté (cf. 3.3) que le modèle de Lee-Carter [$ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$]
est une régression sur la 1ère composante principale (à facteur multiplicatif près) de la matrice $Z = ln(m_{x,t}) - \hat{a}_x$

Le modèle CBD est basé sur les logit $q_x\,(t)$, aussi bien pour la variable à modéliser, que pour l'estimation des variables explicatives.

En combinant les approches Lee-Carter (et sa relation avec la régression PCR) et CBD, nous pouvons formuler un modèle de régression PCR emboité avec les k premières composantes principales et basée sur les logit $q_x\,(t)$:

$$\text{logit}\, q_x\,(t) = a_x + b_{x,1}x_1^*(t) + b_{x,2}x_2^*(t) + b_{x,3}x_3^*(t) + \cdots + b_{x,k}x_k^*(t) + \varepsilon_{x,t}$$

Où :

$q_x(t)$ est la probabilité de décès à l'âge x, au temps t
a_x : moyenne des logit $q_x\,(t)$ pour l'âge x calculée sur l'ensemble de la période considérée
x_i^* est la i ème composante principale de la matrice $\mathrm{X} = \text{logit}\, q_x\,(t) - \hat{a}_x$
$p = x_M - x_m + 1$
k $(k \leq p)$: Nombre de premières composantes principales sélectionnées (par méthode *Foreward*).
$\varepsilon_{x,t}$ est le résidu du modèle, supposé *i.i.d.*, d'espérance nulle et de variance σ_ε^2

Nous désignons ce modèle par : *Logit-PCR.*

7. Amélioration du modèle Logit-PCR

Boyer et al [2015] semblent associer le pouvoir explicatif (par rapport à la variable à expliquer) d'une composante principale, au pourcentage de variance qu'elle explique (du côté des variables explicatives). Cette méthode de sélection des composantes principales à conserver dans le modèle, n'est pas optimale du point de vue explicatif ou prédictif, car la variable à expliquer n'intervient pas dans le choix de ces composantes.
Les améliorations envisagées ci-dessus se justifient d'autant plus que nous remarquons empiriquement des décalages temporels, des différences d'allure dans la dynamique des taux de mortalité selon les tranches d'âge (cf., Figure xxx).
Cela questionne l'utilisation d'une composante temporelle unique, quel que soit l'âge modélisé.

Si une composante temporelle unique reflète la transition démographique du point de vue de la mortalité globale, la prise en compte de la transition épidémiologique et sanitaire qui affecte les âges de manière différentiée (en débutant par les plus jeunes âges) nécessite des composantes temporelles différentiées par tranche d'âge.

L'objectif ici est d'améliorer le modèle Logit-PCR en utilisant une (ou "plusieurs" selon le nombre d'observations i.e. années d'historique disponibles) composante principale optimale, pour chaque tranche d'âge (ou âge) modélisée. On appellera ce modèle Logit-PCR-***optimal***.

Nous ne souhaitons pas conserver toutes les composantes : c'est une réduction de dimension qui se justifie souvent par le grand nombre de variables explicatives initiales (et donc de composantes principales calculées), leurs colinéarités potentielles, et aussi par le nombre limité d'observations utilisées pour l'estimation (apprentissage du modèle).
Il est donc nécessaire de sélectionner les k_x meilleures composantes principales à conserver dans le modèle de chaque âge x.
Nous pouvons proposer quelques méthodes classiques de sélection de variables explicatives :

a. Critères de sélection du modèle : R^2 ajusté, AIC, BIC ou MSEP. Nous utiliserons préférentiellement MSEP.
b. Procédure de sélection du modèle : *stepwise (*utilisé avec R^2 ajusté, AIC, BIC), ou *recherche exhaustive*. Nous utiliserons préférentiellement la recherche exhaustive.

Le nouveau modèle est désormais formulé séparément pour chaque âge :

$$\text{logit}\, q_x\,(t) = a_x + \beta^*_{c_1} x^*_{c_1} + \beta^*_{c_2} x^*_{c_2} + \cdots + \beta^*_{c_{k_x}} x^*_{c_{k_x}} + \varepsilon_{x,t}$$

Où :

$q_x(t)$ est la probabilité de décès à l'âge x, au temps t
a_x : moyenne des $\text{logit}\, q_x\,(t)$ pour l'âge x calculée sur l'ensemble de la période considérée
$x^*_{c_i}$ est une composante principale de la matrice $\mathrm{X} = \text{logit}\, q_x\,(t) - \hat{a}_x$ (il s'agit pas de la $i^{\text{ème}}$ composante principale)
$p = x_M - x_m + 1$
$k\ \ (k \leq p)$: Nombre de premières composantes principales sélectionnées (par méthode *Foreward*).
$\varepsilon_{x,t}$ est le résidu du modèle, supposé *i.i.d.*, d'espérance nulle et de variance σ^2_ε

Nous désignons ce modèle par : ***Logit-PCR-optimal***.

Tableau 4.1 Comparaison théorique des modèles Logit-PCR et Logit-PCR-Optimal

Etapes	**Modèle Logit-PCR**	**Modèle Logit-PCR-Optimal**
Critère du choix du modèle	MAPE, MSE obtenus par validation croisée	MAPE, MSE obtenus par validation croisée
Procédure de sélection du modèle	Méthode emboîtée (*foreward selon la variance expliquée par composante*)	Recherche exhaustive ou *stepwise*
Modèles à estimer	$\text{logit}\, q_x\,(t) = a_x + b_{x,1}x_1^*(t) + \cdots + b_{x,k}x_k^*(t) + \varepsilon_{x,t}$	$\text{logit}\, q_x\,(t) = a_x + \beta_{c_1}^* x_{c_1}^* + \beta_{c_2}^* x_{c_2}^* + \cdots + \beta_{c_{k_x}}^* x_{c_{k_x}}^* + \varepsilon_{x,t}\,, x = x_m, \dots, x_M$

Tests de validation du modèle

Idéalement, chaque modèle, avant son utilisation, devrait être validé par les tests suivants :

a. Test d'homoscédasticité des erreurs (test de Harrison-McCabe)
Hypothèse (H_0) :$var(\varepsilon_t) = \sigma^2$ (homoscédasticité) contre (H_1): $var(\varepsilon_t) \neq \sigma^2$(hétéroscédasticité)

b. Test d'autocorrélation des erreurs (test de Breusch-Godfrey)
Hypothèse (H_0):$Cov(\varepsilon_t, \varepsilon_s) = 0$ (non-autocorrélation) contre (H_1): $Cov(\varepsilon_t, \varepsilon_s) \neq 0$ (autocorrélation) $\forall\ t \neq s$

c. Test de normalité des erreurs (test de Jarque-Bera)
Hypothèse (H_0) :ε_t suivent la loi normale contre (H_1): ε_t ne suivent pas la loi normale $\forall\ t$
L'hypothèse de normalité n'est pas indispensable pour valider le modèle.

Si les résidus sont auto-corrélés, il sera nécessaire de les modéliser par un processus ARMA de sorte à obtenir un bruit blanc.

8. Régression PLS

Comme la régression sur composantes principales (*Principal Component Regression, PCR*), le regression PLS (*Partial Least Square*) a été introduite principalement pour gérer les problèmes de multicolinéarité entre variables explicatives, à la nuance que ses composantes sont calculées en tenant compte de leur covariance avec la variable à expliquer. Nous en faisons ici un exposé car il nous sera utile pour proposer une généralisation du modèle de Lee-Carter, basée sur les composantes PLS.

Soient les nouvelles variables explicatives $l^{(1)}, l^{(2)}, \dots, l^{(k)}$, combinaisons linéaires des variables de départ $l^{(j)} = Xc_j$, qui sont orthogonales entre elles et classées par ordre décroissant. Nous remarquons que les composantes principales X_j^* obéissent à ces mêmes critères.
Lorsque Y est univarié, la régression PLS est appelée PLS1 et elle se définit itérativement. Pour cela, une procédure itérative va être utilisée (Cornillon et al. [2011]).

Etape 1 : La matrice X est notée $X^{(1)}$ et Y la variable à expliquer est notée $Y^{(1)}$.
La 1ère composante (ou variable latente) $l^{(1)} = X^{(1)}w_1 \in \mathbb{R}^n$ est choisie telle que :

$l^{(1)} = \underset{l=X^{(1)}w, w\in\mathbb{R}^P, \|w\|^2=1}{\operatorname{argmax}} \operatorname{cov}\left(l, Y^{(1)}\right)$, où w_1 désigne le poids
C'est-à-dire maximiser $\operatorname{cov}\left(l, Y^{(1)}\right)$ sous contrainte $\|w_1\|^2 = 1$. Pour le faire, utilisons la fonction Lagrangienne qui est définie par :
$L = \left(Y^{(1)}\right)^T X^{(1)} w_1 + \lambda(1 - \|w_1\|^2)$
On a alors :
$\frac{\partial L}{\partial w_1} = \left(Y^{(1)}\right)^T X^{(1)} - 2\lambda w_1 = 0$ et $\frac{\partial L}{\partial \lambda} = (1 - \|w_1\|^2) = 0$

L'équation $\left(Y^{(1)}\right)^T X^{(1)} - 2\lambda w_1 = 0$ montre que w_1 est colinéaire au vecteur $\left(Y^{(1)}\right)^T X^{(1)}$ et $\|w_1\|^2 = 1$ montre qu'il est normé. Si l'on veut un maximum, il suffit de prendre le vecteur
$w_1 = \frac{\left(Y^{(1)}\right)^T X^{(1)}}{\left\|\left(Y^{(1)}\right)^T X^{(1)}\right\|}$, avec $\lambda = \frac{1}{2}$.
Ensuite, nous effectuons la régression simple de $Y^{(1)}$ sur $l^{(1)}$, $Y^{(1)} = r_1 l^{(1)} + \hat{\varepsilon}_1$ où $r_1 \in \mathbb{R}$ est le coefficient de la régression estimée par moindres carrés et $\hat{\varepsilon}_1 = P_{l^{(1)}} \perp X^{(1)}$ est le vecteur des résidus de la régression simple sans constante, où $P_{t^{(1)}\perp}$ est l'opérateur de projection orthogonale sur la droite engendrée par $t^{(1)}$ et donné par $P_{t^{(1)}} \perp = (I - P_{l^{(1)}})$ et $P_{l^{(1)}} = l^{(1)}(l^{(1)^T} l^{(1)})^{-1} l^{(1)^T}$

Etape 2 : On note $Y^{(2)} = \hat{\varepsilon}_1 = P_{l^{(1)}} \perp X^{(1)}$, le résidu de $Y^{(1)}$ sur $l^{(1)}$ ou soit la partie encore non expliquée de Y.
Soit $X^{(2)} = P_{l^{(1)}} \perp X^{(1)}$ la partie de $X^{(1)}$ n'ayant pas encore servi à expliquer, ou soit le vecteur des résidus des régressions des variables de X sur $l^{(1)}$.
La seconde composante PLS $l^{(2)} = X^{(2)}w_2 \in \mathbb{R}^n$ est choisie telle que :
$l^{(2)} = \underset{l=X^{(2)}w, w\in\mathbb{R}^P, \|w\|^2=1}{\operatorname{argmax}} \operatorname{cov}\left(l, Y^{(2)}\right)$

En faisant le même raisonnement que dans l'étape 1, on a :
$w_2 = \left(Y^{(2)}\right)^T X^{(2)} / \left\|\left(Y^{(2)}\right)^T X^{(2)}\right\|$
Nous effectuons ensuite la régression simple de $Y^{(2)}$ sur $l^{(1)}$; $Y^{(2)} = r_2 l^{(2)} + \hat{\varepsilon}_2$ où $r_2 \in \mathbb{R}$ est le coefficient de la régression estimée par les moindres carrés et $\hat{\varepsilon}_2 = P_{l^{(2)}} \perp Y^{(2)}$ est le vecteur des résidus de la régression simple sans constante.

Nous poursuivons le même calcul, jusqu'à la $k^{ième} - 1$ étape.
Etape k : On pose $Y^{(k)} = \hat{\varepsilon}_{k-1} = P_{l^{(k-1)}} \perp Y^{(k-1)}$ la partie non encore expliquée de Y et $X^{(k)} = P_{l^{(k-1)}} \perp X^{(k-1)}$ la partie de $X^{(k-1)}$ n'ayant pas servi à expliquer la composante PLS $l^{(k)}$.
La composante PLS k; $l^k = X^{(k)} w_k \in \mathbb{R}^n$ est choisie telle que :

$$l^{(k)} = \underset{l=X^{(k)}w, w\in\mathbb{R}^P, \|w\|^2=1}{\operatorname{argmax}} \operatorname{cov}\left(l, Y^{(k)}\right)$$

Soit par le même raisonnement que dans les étapes précédentes, on a :

$$w_k = \left(Y^{(k)}\right)^T X^{(k)} / \left\| \left(Y^{(k)}\right)^T X^{(k)} \right\|$$

Ensuite, nous effectuons la régression simple de $Y^{(k)}$ sur $l^{(k)}$: $Y^{(k)} = r_k l^{(k)} + \hat{\varepsilon}_k$ où $r_k \in \mathbb{R}$ est le coefficient de la régression estimée par les moindres carrés et $\hat{\varepsilon}_k = P_{l^{(k)}} \perp Y^{(k)}$.
En utilisant le postulat que les composantes PLS sont orthogonales entre elles, le modèle PLS s'écrit sous la forme suivante :

$$Y = P_{l^{(1)}} Y^{(1)} + \cdots + P_{l^{(k)}} Y^{(k)} + \hat{\varepsilon}_k$$
$$= r_1 l^{(1)} + \cdots + r_k l^{(k)} + \hat{\varepsilon}_k \text{ , avec } \hat{\varepsilon}_k = P_{l^{(k)}} \perp Y^{(k)} = P_{B\left(l^{(1)},\ldots,l^{(k)}\right)} \perp Y.$$

9. Améliorations du modèle Logit-PCR en utilisant la régression PLS

La finalité ici est d'améliorer le modèle Logit-PCR en utilisant la régression PLS et en choisissant de manière optimale (MSEP par validation croisée) les composantes principales PLS pour chaque tranche d'âge (ou âge) modélisée. On appellera ce modèle ***logit-PLS***.

9.1. Amélioration du modèle Logit-PCR : Modèle Logit-PLS

A l'image de la régression sur composantes principales, l'objectif ici est de trouver pour chaque âge x le meilleur modèle à k_x composantes PLS.
Nous rappelons que ces composantes PLS sont notées $l^{(1)}, l^{(2)}, \dots, l^{(p)}$, combinaisons linéaires des variables de départ $l^{(j)} = Xc_j$, et qui soient orthogonales entre elles et classées par ordre d'importance, c'est-à-dire leurs liens avec la variable à expliquer. Cependant, le calcul de ces composantes $l^{(j)}$ est se fait par leur lien avec la variable à expliquer et non par la variabilité qu'elles représentent parmi les variables explicatives originales (comme en régression sur composantes principales).
Et le modèle sera donc défini par :

$\text{logit}\, q_x\,(t) = r_1 l^{(1)} + \cdots + r_{k_x} l^{(k_x)} + \varepsilon_{k_x}$, qui sera appelé **Logit-PLS,**

où $l^{(j)} = \underset{l=X^{(j)}w, w\in\mathbb{R}^P, \|w\|^2=1}{\text{argmax cov}} \left(l, Y^{(j)}\right)$ et $r_j \in \mathbb{R} \;\; \forall\, 1 \le j \le k_x$ est le coefficient de la régression estimé par les moindres carrés et $\hat{\varepsilon}_{k_x} = P_{l^{(k_x)}} \perp Y^{(k_x)}$. Pour plus de détails sur la procédure itérative des composantes PLS (cf.§ 8).

Tableau 4.2 Comparaison théorique des modèles Logit-PCR-Optimal et Logit-PLS

Etapes	**Modèle Logit-PCR-Optimal**	**Modèle Logit-PLS**
Critère du choix du modèle	MAPE, MSE par validation croisée	MAPE, MSE par validation croisée
Procédure de sélection du modèle	Recherche exhaustive ou *stepwise*	Méthode emboîtée (intégration progressive des composantes selon la variance expliquée)
Modèles à estimer	$\text{logit}\, q_x\,(t) = a_x + \beta^*_{c_1} x^*_{c_1} + \beta^*_{c_2} x^*_{c_2} + \cdots + \beta^*_{c_{k_x}} x^*_{c_{k_x}} + \varepsilon_{x,t}$, $x = x_m, \dots, x_M$	$\text{logit}\, q_x\,(t) = r_1 l^{(1)} + \cdots + r_{k_x} l^{(k_x)} + \varepsilon_{k_x}$, $x = x_m, \dots, x_M$

10.Variable et modèle PCR cliométrique

La variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est une variable créée à partir des données historiques du pays A (transitionnellement « plus âgé » que le pays R), pour jouer un rôle de variable exogène dans les modèles de prévisions de mortalité du pays R.
Cette variable cliométrique exogène de mortalité sera construite de sorte à avoir la même allure temporelle que dans les pays « transitionnellement plus âgé», mais adaptée au rythme du pays à modéliser.
La variable cliométrique sera ainsi utilisée comme variable explicative supplémentaire dans les modèles prospectifs de mortalité du pays R. Nous nommons les modèles utilisant la variable cliométrique, modèles cliométriques : nous en proposons deux types qui sont les modèles PCR-optimal-cliometric et PLS-optimal-cliometric.

10.1. Concept de variable cliométrique

On notera le pays transitionnellement « plus âgé », pays « A», et le pays transitionnellement « moins âgé » à modéliser, pays « R ».

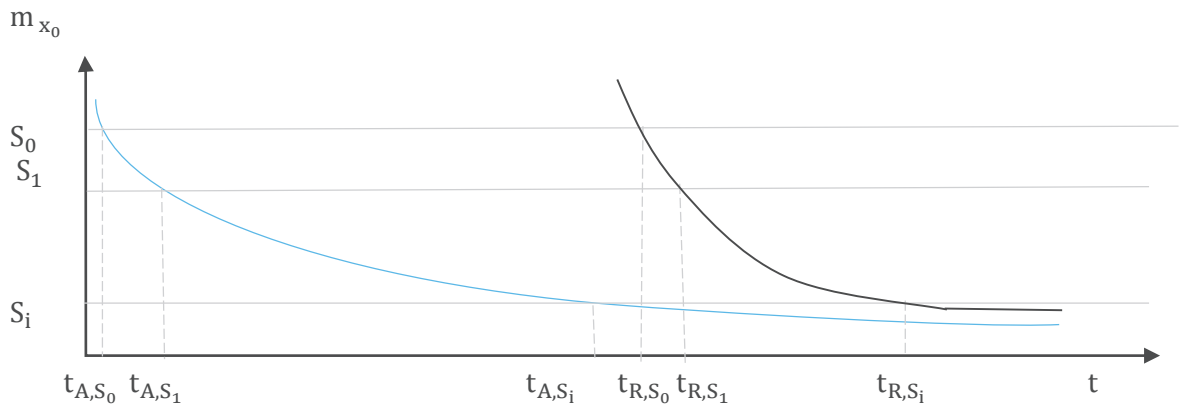
La variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est une variable crée à partir des données historiques du pays A (transitionnellement « plus âgé » que le pays R), pour jouer un rôle de variable exogène dans les modèle de prévisions de mortalité du pays R.
L'idée est d'utiliser l'histoire quantitative des mortalités (et éventuellement d'autres variables démographiques) du pays A pour tenter d'améliorer les prévisions de mortalité dans le pays R.
Une telle variable peut être utilisée comme variable active explicative supplémentaire dans les modèles prospectifs de mortalité du pays R.

Nous présentons ici la méthodologie d'estimation de la variable *cliométrique*.

10.2. Méthodologie de création d'une variable cliométrique

Notre objectif est de créer une variable cliométrique pour un pays R, pour un âge x donné.
Pour créer cette variable cliométrique, nous allons chercher à estimer une fonction qui associe chaque point temporel (t_A) du processus de transition du pays A et chaque point temporel (t_R) du processus de transition du pays R.

Figure 4.3 Courbes de mortalité des pays A (bleue) et R (noire)



10.2.1. Variable de temps transitionnelle

Notations :

$m_{A,\,x_0}$: taux de mortalité pour le pays A, à l'âge x_0,

$m_{R,\,x_0}$: taux de mortalité pour le pays R, à l'âge x_0,

$m_{R,x}$: taux de mortalité pour le pays R, pour un âge donné (ou une tranche d'âge donnée) x. C'est le taux de mortalité que nous cherchons à modéliser pour le pays R.

$m_{A,x}$: taux de mortalité pour le pays A, pour un âge donné (ou une tranche d'âge donnée) x,

Soit une variable m_{x_0} (taux de mortalité à l'âge x_0) dont la symétrique (par rapport à toute axe parallèle à l'axe horizontal du temps) est un temps transitionnelle (cf. chap. 1, §5.5, §7.2.5) mesurable dans les deux pays A ($m_{A,\,x_0}$) et R ($m_{R,\,x_0}$).

Cette variable m_{x_0} a les propriétés suivantes :

- La monotonie en tendance (nécessaire au rôle de temps transitionnel tel que préalablement défini)
- Des valeurs dans un même intervalle dans les deux pays (pour permettre les appariements entre les points temporels des 2 processus liés aux 2 pays).

Cette variable m_{x_0}, du fait de sa monotonie statistique par rapport au temps (cf. chap. 1, §5.4, §5.5, §7.2.5), nous permet de faire un appariement entre les points du processus transitionnel du pays A ($m_{A,x}(t_A)$)et celui du pays R ($m_{R,x}(t_R)$) .

Pour créer la variable *cliométrique de mortalité à l'âge x* du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$), nous proposons deux méthodes distinctes, dont nous comparerons les performances par la suite.

10.2.2. Première méthode d'estimation de la variable cliométrique

Description de la méthode

Dans cette méthode, nous cherchons à déterminer 2 fonctions déterministes monotones qui modélisent au mieux les courbes des variables m_{A,x_0} et m_{R,x_0}. A partir de ces fonctions déterministes, il sera alors possible d'établir une relation entre les points temporels t_A et t_R, car m_{A,x_0} et m_{R,x_0} sont des fonctions bijectives respectivement de t_A et t_R.

Soit s_0 un seuil arbitraire (des variables m_{A,x_0} et m_{R,x_0}) dont l'atteinte ou le franchissement (à la baisse) indique une présence du pays (A ou R) dans son processus de transition démographique.

En vertu de l'hypothèse d'universalité de l'allure des courbes de mortalité à l'âge x_0 durant la transition démographique, les courbes du taux de mortalité à l'âge x_0 des pays A et R (m_{A,x_0} et m_{R,x_0}) ont même allure après le passage du seuil s_0.

Soit t_{A,s_0} le premier temps d'atteinte ou de franchissement (à la baisse) de la valeur s_0 par la variable $m_{A,\,x_0}$:

$$t_{A,s_0} = \inf\{t \in \mathbb{N}, m_{A,x_0}(t) \leq s_0\}$$

De la même manière t_{R,x_0} est le premier temps d'atteinte ou de franchissement (à la baisse) de la valeur s_0 par la variable m_{R,s_0} :

$$t_{R,s_0} = \inf\{t \in \mathbb{N}, m_{R,x_0}(t) \leq s_0\}$$

Notons :

T le temps (année) présent.

H l'horizon maximal de prévision envisagé.

Soit f_{A,x_0} la fonction déterministe monotone modélisant la variable $m_{A,\,x_0}$
Soit f_{R,x_0} la fonction déterministe monotone modélisant la variable $m_{R,\,x_0}$
On peut donc écrire :
$m_{A,x_0}(t_{A,s_i}) = f_{A,x_0}(t_{A,s_i}) + \mathcal{E}_{t_{A,s_i}} \Rightarrow s_i = f_{A,x_0}(t_{A,s_i}) + \mathcal{E}_{t_{A,s_i}}$ **(a),**
avec f_A une fonction strictement monotone, c'est-à-dire f_A est bijective.

De même, $m_{R,x_0}(t_{R,s_i}) = f_{R,x_0}(t_{R,s_i}) + \mathcal{E}_{t_{R,s_i}} \Rightarrow s_i = f_{R,x_0}(t_{R,s_i}) + \mathcal{E}_{t_{R,s_i}}$ **(b)**,
avec f_R une fonction strictement monotone c'est-à-dire f_R est bijective.

Puisque ces deux fonctions prennent des valeurs dans le même intervalle, nous avons la relation suivante déduite de **(a)** et **(b)** : $f_{A,x_0}(t_{A,s_i}) = f_{R,x_0}(t_{R,s_i}) \Rightarrow t_{R,s_i} = f_{R,x_0}^{-1}\left(f_{A,x_0}(t_{A,s_i})\right)$

$\mathcal{E}_{t_{A,s_i}}$ et $\mathcal{E}_{t_{R,s_i}}$ sont des erreurs *i.i.d.* qui doivent respecter les conditions suivantes : espérance nulle, homoscédasticité, non-autocorrélation : Autrement dit $\left(\varepsilon_{t_{A,s_i}}\right)_i$ et $\left(\varepsilon_{t_{R,s_i}}\right)_i$ sont des bruits blancs.

Si les résidus suivent les conditions usuelles du modèle linéaire, alors nous utilisons le modèle fourni par la régression pour réaliser les prévisions. Sinon, les modèles ne sont pas valides et il sera nécessaire de compléter le modèle en modélisant les résidus par un ARMA.

Nous avons le tableau initial suivant pour le pays A :

Tableau 4.3 Tableau du pays A

$m_{A,x}$	$m_{A,\,x_0}$	t_A	$f_{A,\,x_0}$
	s_0	t_{A,s_0}	$f_{A,\,x_0}(\ t_{A,s_0})$
	s_1	t_{A,s_1}	$f_{A,\,x_0}(\ t_{A,s_1})$
	$\vdots$	$\vdots$	$\vdots$
	s_i	t_{A,s_i}	$f_{A,\,x_0}(\ t_{A,s_i})$
	$\vdots$	$\vdots$	$\vdots$
	s_d	t_{A,s_d}	$f_{A,\,x_0}(\ t_{A,s_d})$
	s_{d+1}	$t_{A,s_{d+1}}$	$f_{A,\,x_0}(\ t_{A,s_{d+1}})$
	$\vdots$	$\vdots$	$\vdots$
	s_r	t_{A,s_r}	$f_{A,\,x_0}(\ t_{A,s_r})$

d : Nombre d'observations disponibles de la variable $m_{R,\,x_0}$ à partir du temps t_{R,s_0}
$m_{A,\,x_0} = (S_i\ ,0 \leq i \leq r)^T$, avec $r > d$ et $S_r < S_d$

Nous souhaitons apparier chaque observation du pays A avec une observation du pays R.
Nous ferons l'appariement entre les valeurs de $f_{A,\,x_0}$ et de $f_{R,\,x_0}$
Pour chaque observation du pays A, nous cherchons à associer une observation du pays R, telle que :
$f_{A,\,x_0}(t_{A,s_i}) = f_{R,\,x_0}(t_{R,s_i}) = s_i$
D'où :

$t_{R,s_i} = f_{R,x_0}^{-1}\left(f_{A,x_0}(t_{A,s_i})\right)$, si la réciproque existe.

Puisque t_R doit être entier (année), nous prendrons la valeur arrondie à l'entier le plus proche.
Après association des observations du pays R à celles du pays A, nous obtenons le tableau suivant :

Tableau 4.4 Tableau d'appariement des tables du pays A et du pays R

$m_{A,x}$	m_{A,x_0}	t_A	f_{A,x_0}	$t_R = f_{R,x_0}^{-1}\left(f_{A,x_0}(t_A)\right)$	$m_{R,x}(t_R)$
	s_0	t_{A,s_0}	$f_{A,\,x_0}(t_{A,S_0})$	t_{R,s_0}	
	s_1	t_{A,s_1}	$f_{A,\,x_0}(t_{A,S_1})$	t_{R,s_1}	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	s_i	t_{A,s_i}	$f_{A,\,x_0}(t_{A,S_i})$	t_{R,S_i}	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	s_d	t_{A,s_d}	$f_{A,\,x_0}(\ t_{A,s_d})$	t_{R,S_d}	
	s_{d+1}	$t_{A,s_{d+1}}$	$f_{A,\,x_0}(\ t_{A,s_{d+1}})$	$t_{R,S_{d+1}}$	NA
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	s_r	t_{A,s_r}	$f_{A,\,x_0}(\ t_{A,s_r})$	t_{R,S_r}	NA

La mention NA désigne les valeurs manquantes.

La variable *cliométrique de mortalité à l'âge x* du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est la variable définie par les points du couple des variables (t_R , $m_{A,x}$) :

$$\mathcal{C}_{R,A,x}:\ \ t_{R,S_i} \mapsto m_{A,x}(t_{A,s_i})$$

Si les arrondis de t_R donnent des valeurs identiques pour des valeurs différentes de $m_{A,x}$, la moyenne des valeurs de $m_{A,x}$ est associée aux valeurs identiques de t_R , lors de la création des points de la variable cliométrique.

Pour éviter les valeurs manquantes dans la variable cliométrique, due aux valeurs manquantes dans la variable t_R (qui devrait contenir des valeurs consécutives des années), une interpolation spline (ou linéaire) pourra être utilisée.

Si l'horizon $T + H > t_{A,s_r}$ alors nous aurons recours aux prévisions par séries temporelles pour estimer les valeurs de la variable cliométrique sur la période $[\,t_{A,s_r} + 1; T + H]$

Application : Choix de la fonction mathématique à utiliser

a. Modélisation du taux de mortalité à l'âge x_0 par une fonction exponentielle

Nous faisons l'hypothèse que les courbes de taux de mortalité à l'âge x_0 peuvent être estimées par une fonction exponentielle, c'est-à-dire la courbe de taux de mortalité x_0 pour le pays A ($m_{A,\,x_0}$) en fonction de t_{A,s_i} peut être estimée par :

$$m_{A,\,x_0}(t_{A,s_i}) = ae^{-bt_{A,s_i}} + \mathcal{E}_{t_{A,s_i}},\ a, b > 0$$

et la courbe de taux de mortalité x_0 pour le pays en R ($m_{R,\,x_0}$) en fonction de t_{R,s_i} est donnée par :

$$m_{R,\,x_0}(t_{R,s_i}) = ce^{-dt_{R,s_i}} + \mathcal{E}_{t_{R,s_i}},\ \ c, d > 0.$$

Nous effectuons les régressions Log-linéaires des taux de mortalité à l'âge x_0 sur le temps (année) de chaque pays (pays A et pays en R) en nous basant sur leurs courbes de taux de mortalité. En appliquant Log à ces taux de mortalité, nous obtenons des régressions Log-linéaires données par :

$$\begin{cases} m_{A,\,x_0}(t_{A,s_i}) = ae^{-bt_{A,s_i}} + \mathcal{E}_{t_{A,s_i}} \\ m_{R,\,x_0}(t_{R,s_i}) = ce^{-dt_{R,s_i}} + \mathcal{E}_{t_{R,s_i}} \end{cases} \Longrightarrow \begin{cases} log\left(m_{A,\,x_0}(t_{A,s_i})\right) = log\,(a) - bt_{A,s_i} + \mathcal{E}_{t_{A,s_i}} \\ log\left(m_{R,\,x_0}(t_{R,s_i})\right) = log\,(c) - dt_{R,s_i} + \mathcal{E}_{t_{R,s_i}} \end{cases}$$

On a : $s_i = m_{A,\,x_0}(t_{A,s_i})$ et $s_i = m_{R,\,x_0}(t_{R,s_i})$

$$\Longrightarrow \; m_{A,\,x_0}\left(t_{A,s_i}\right) = m_{R,\,x_0}\left(t_{R,s_i}\right)$$
$$\Longrightarrow log(a) - bt_{A,s_i} = log(c) - dt_{R,s_i}$$
$$\Longrightarrow log(a) - bt_{A,s_i} = log(c) - dt_{R,s_i}$$
$$\Longrightarrow \; t_{R,s_i} = \frac{1}{d} log\left(\frac{c}{a}\right) + \frac{b}{d} t_{A,s_i}$$

Finalement t_{R,s_i}est sous la forme $t_{R,s_i} = D + Bt_{A,s_i}$ où $D, B \in \mathbb{R}_+$, avec $D = \frac{1}{d} log\left(\frac{c}{a}\right), B = \frac{b}{d}$.

b. Modélisation du taux de mortalité à l'âge x_0 par une fonction logistique avec pente décroissante

En faisant l'hypothèse que les courbes de taux de mortalité à l'âge x_0 peuvent être estimées par une fonction logistique, la courbe de taux de mortalité à l'âge x_0.pour le pays A ($m_{A,\,x_0}$) en fonction de t_{A,s_i} est définie par :
$m_{A,\,x_0}\left(t_{A,s_i}\right) = \frac{S_0}{1+e^{-(a_0+a_1 t_{A,s_i})}} + \mathcal{E}_{t_{A,s_i}}$, a_0 et a_1 sont des paramètres à estimer avec $a_1 \leq 0$.

Celle de taux de mortalité à l'âge x_0 pour le pays R ($m_{R,\,x_0}$) en fonction de t_{R,s_i} est définie par :

$m_{R,\,x_0}\left(t_{R,s_i}\right) = \frac{S_0}{1+e^{-(b_0+b_1 t_{R,s_i})}} + \mathcal{E}_{t_{R,s_i}}$, b_0 et b_1 sont des paramètres à estimer avec $b_1 \leq 0$.

Pour linéariser ces fonctions, nous procédons à leurs transformation, puis nous effectuons les régressions Log-linéaires des taux de mortalité à l'âge x_0 sur le temps (année) de chaque pays (pays A et pays R).

$$\frac{m_{A,x_0}\left(t_{A,s_i}\right)}{S_0} = \frac{1}{1+e^{-\left(a_0+a_1 t_{A,s_i}\right)}} \Rightarrow \frac{S_0}{m_{0,A}\left(t_{A,s_i}\right)} = 1 + e^{-\left(a_0+a_1 t_{A,s_i}\right)}$$
$$\Rightarrow \frac{S_0}{m_{A,x_0}\left(t_{A,s_i}\right)} - 1 = e^{-\left(b_0+b_1 t_{A,s_i}\right)}$$
$$\Rightarrow \frac{1}{\frac{S_0}{m_{A,x_0}\left(t_{A,s_i}\right)} - 1} = e^{\left(a_0+a_1 t_{A,s_i}\right)}$$
$$\Rightarrow log\left(\frac{1}{\frac{S_0}{m_{A,x_0}\left(t_{A,s_i}\right)}-1}\right) = a_0 + a_1 t_{A,s_i}$$

De même pour le pays R nous avons :

$$\frac{m_{R,x_0}\left(t_{R,s_i}\right)}{S_0} = \frac{1}{1+e^{-\left(b_0+b_1 t_{R,s_i}\right)}} \Rightarrow \frac{S_0}{m_{R,x_0}\left(t_{R,s_i}\right)} = 1 + e^{-\left(b_0+b_1 t_{R,s_i}\right)}$$
$$\Rightarrow \frac{S_0}{m_{R,x_0}\left(t_{R,s_i}\right)} - 1 = e^{-\left(b_0+b_1 t_{R,s_i}\right)}$$
$$\Rightarrow \frac{1}{\frac{S_0}{m_{R,x_0}\left(t_{R,s_i}\right)} - 1} = e^{\left(b_0+b_1 t_{R,s_i}\right)}$$
$$\Rightarrow log\left(\frac{1}{\frac{S_0}{m_{R,x_0}\left(t_{R,s_i}\right)}-1}\right) = b_0 + b_1 t_{R,s_i}$$

Soit s_i une valeur de m_{A,x_0} interne au processus de transition quel que soit le pays.

$\forall\ i \in [0\ ;\ T]$ on a :

$s_i = m_{A,x_0}(t_{A,s_i}) = m_{R,x_0}(t_{R,s_i})$
Donc :

$$log\left(\frac{\frac{1}{s_0}}{\frac{}{m_{A,x_0}(t_{A,s_i})}-1}\right) = log\left(\frac{\frac{1}{s_0}}{\frac{}{m_{R,x_0}(t_{R,s_i})}-1}\right).$$

Par conséquent :

$$a_0 + a_1 t_{A,s_i} = b_0 + b_1 t_{R,s_i} \Rightarrow t_{R,s_i} = \frac{a_0 - b_0}{b_1} + \frac{a_1}{b_1} t_{A,s_i}$$

D'où :

$$t_{R,s_i} = \alpha + \beta t_{A,s_i}, \text{ avec } \alpha = \frac{a_0 - b_0}{b_1}\ \ et\ \ \beta = \frac{a_1}{b_1}$$

A partir des coefficients $\widehat{b_0}$ et $\widehat{b_1}$ estimés, la fonction logistique donne :

$$\widehat{m}_{R,x_0}(t_R) = \frac{S_0}{1 + e^{-(\widehat{b_0} + \widehat{b_1} t_R)}}$$

La fonction réciproque f_{R,x_0}^{-1} utilisée pour déterminer les t_R est calculée après estimation des coefficients $\widehat{b_0}$ et $\widehat{b_1}$ dans la fonction $f_{R,\,x_0}$.

$$\frac{S_0}{1 + e^{-(\widehat{b_0} + \widehat{b_1} t_R)}} = \widehat{m}_{R,x_0}$$
$$\Longrightarrow \frac{1}{1 + e^{-(\widehat{b_0} + \widehat{b_1} t_R)}} = \frac{\widehat{m}_{R,x_0}}{S_0}$$
$$\Longrightarrow 1 + e^{-(\widehat{b_0} + \widehat{b_1} t_R)} = \frac{S_0}{\widehat{m}_{R,x_0}}$$
$$\Longrightarrow \widehat{b_0} + \widehat{b_1} t_R = log\left(\frac{1}{\frac{S_0}{\widehat{m}_{R,x_0}} - 1}\right)$$

La fonction réciproque est donnée par :

$$t_R(m_{R,x_0}) = \frac{1}{\widehat{b_1}} log\left(\frac{1}{\frac{S_0}{m_{R,x_0}} - 1}\right) - \frac{\widehat{b_0}}{\widehat{b_1}}$$

D'où :

$$\widehat{m}_{R,x_0}^{-1}(t_R) = \frac{1}{\widehat{b_1}} log\left(\frac{1}{\frac{S_0}{t_R} - 1}\right) - \frac{\widehat{b_0}}{\widehat{b_1}}$$

Cette fonction réciproque est utilisée pour déterminer les valeurs t_{R,s_d} à t_{R,s_r}
La variable cliométrique est créée en affectant les taux de mortalité à l'âge x du pays A aux t_R obtenus après appariement.

10.2.3. Deuxième méthode d'estimation de la variable cliométrique

Pour estimer la variable cliométrique, nous utilisons les données sans faire l'hypothèse d'une fonction explicite du taux de mortalité à l'âge x_0.

Nous avons le tableau initial suivant pour le pays A :

Tableau 4.5 Tableau du pays A. Méthode 2

$\mathrm{m_{A,x}}$	$\mathrm{m_{A,\,x_0}}$	$\mathrm{t_A}$
	$\mathrm{s_0}$	$\mathrm{t_{A,s_0}}$
	$\mathrm{s_1}$	$\mathrm{t_{A,s_1}}$
	$\vdots$	$\vdots$
	$\mathrm{s_i}$	$\mathrm{t_{A,s_i}}$
	$\vdots$	$\vdots$
	$\mathrm{s_d}$	$\mathrm{t_{A,s_d}}$
	$\mathrm{s_{d+1}}$	$\mathrm{t_{A,s_{d+1}}}$
	$\vdots$	$\vdots$
	$\mathrm{s_r}$	$\mathrm{t_{A,s_r}}$

d : Nombre d'observations disponibles de la variable $\mathrm{m_{R,\,x_0}}$ à partir du temps $\mathrm{t_{R,s_0}}$
$\mathrm{m_{A,\,x_0}} = (S_i\,, 0 \leq i \leq r)^T$, avec $r > d$ et $S_r < S_d$

Dans le cas présent $\mathrm{t_{R,s_i}}$ est le premier temps d'atteinte ou de franchissement (à la baisse) de la valeur $\mathrm{s_i}$ par la variable $\mathrm{m_{R,x_0}}$:

$$\mathrm{t_{R,s_i}} = \inf\{\mathrm{t} \in \mathbb{N}, \mathrm{m_{R,x_0}(t)} \leq \mathrm{s_i}\}, \mathrm{i} = 0, \dots, \mathrm{d}$$

Pour estimer $\mathrm{t_{R,s_i}}$ sur l'intervalle $[\mathrm{t_{R,s_{d+1}}}; \mathrm{t_{R,s_r}}]$ nous estimons un modèle linéaire entre $\mathrm{t_R}$ et $\mathrm{t_A}$ en utilisant les valeurs connues de ses 2 variables.
L'écart entre les 2 variables n'étant pas constant, on peut écrire :
$t_{R,s_i} - t_{A,s_i} = \alpha_0 + \alpha_1 t_{A,s_i} + \mathcal{E}_{s_i}$, avec α_0>0 et $\alpha_1 < 0$.

Le modèle final est donné par :
$t_{R,s_i} = \alpha_0 + (\alpha_1 + 1) t_{A,s_i} + \mathcal{E}_{s_i}$, avec α_0>0 et $\alpha_1 < 0$.

$\mathcal{E}_{s_i}$ sont des erreurs *i.i.d.* qui doivent respecter les conditions suivantes : espérance nulle, homoscédasticité, non-autocorrélation : Autrement ce sont des bruits blancs.
Si les résidus suivent les conditions usuelles du modèle linéaire, alors nous utilisons le modèle fourni par la régression pour réaliser les prévisions. Sinon, les modèles ne sont pas valides et il sera nécessaire de compléter le modèle en modélisant les résidus par un ARMA.

Puisque $\mathrm{t_R}$ doit être entier (année), nous prendrons la valeur arrondie à l'entier le plus proche.

Contrainte de calcul dans le cas où $t_{A,s_i} < t_{R,s_i}$: Si $t_{R,s_i} - t_{A,s_i} < 0$, alors on fixe $t_{R,s_i} = t_A, s_i$. Dans ce cas on n'autorise pas le pays R à atteindre la valeur s_i avant le pays A. Cette contrainte pourra être relâchée dans une approche alternative.

Nous souhaitons à présent apparier chaque observation du pays A avec une observation du pays R.
A chaque observation du pays A avec $\mathrm{t_A} = \mathrm{t_{A,s_i}}$, nous associons une observation du pays R avec $\mathrm{t_R} = \mathrm{t_{R,s_i}}$

Après association des observations du pays R à celles du pays A, nous obtenons le tableau suivant :

Tableau 4.6 Tableau d'appariement des tables du pays A et du pays R

$m_{A,x}$	m_{A,x_0}	t_A	t_R	$m_{R,x}(t_R)$
	s_0	t_{A,s_0}	t_{R,s_0}	
	s_1	t_{A,s_1}	t_{R,s_1}	
	$\vdots$	$\vdots$	$\vdots$	
	s_i	t_{A,s_i}	t_{R,S_i}	
	$\vdots$	$\vdots$	$\vdots$	
	s_d	t_{A,s_d}	t_{R,S_d}	
	s_{d+1}	$t_{A,s_{d+1}}$	$t_{R,S_{d+1}}$	NA
	$\vdots$	$\vdots$	$\vdots$	
	s_r	t_{A,s_r}	t_{R,S_r}	NA

La mention NA désigne les valeurs manquantes.

La variable *cliométrique de mortalité à l'âge x* du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est la variable définie par les points du couple des variables (t_R , $m_{A,x}$) :

$$\mathcal{C}_{R,A,x} : \ t_{R,S_i} \mapsto m_{A,x}(t_{A,s_i})$$

Si les arrondis de t_R donnent des valeurs identiques pour des valeurs différentes de $m_{A,x}$, la moyenne des valeurs de $m_{A,x}$ est associée aux valeurs identiques de t_R , lors de la création des points de la variable cliométrique.

Pour éviter les valeurs manquantes dans la variable cliométrique, due aux valeurs manquantes dans la variable t_R (qui devrait contenir des valeurs consécutives des années), une interpolation spline (ou linéaire, ou moyenne mobile) pourra être utilisée.

Si l'horizon $T + H > t_{A,s_r}$ alors nous aurons recours aux prévisions par séries temporelles pour estimer les valeurs de la variable cliométrique sur la période $[\, t_{A,s_r} + 1; T + H]$

10.3. Amélioration du modèle Logit-PCR, en y intégrant la variable cliométrique

Nous souhaitons améliorer le modèle Logit-PCR, en y intégrant la variable cliométrique correspondante à chaque tranche d'âge (ou âge) modélisée. On appellera ce modèle **Logit-PCR-cliométrique**.
Cette intégration de la variable cliométrique peut se faire selon 2 modalités :

- **Avant** le calcul et la sélection des composantes principales (donc prise en compte dans le calcul des composantes principales) : cette option présuppose une prévision des mortalités par âge. Or ces prévisions sont précisément l'objectif à atteindre. Cette option n'est donc pas pertinente.
- **Après** le calcul et la sélection des composantes principales (donc non-prise en compte dans le calcul des composantes principales). La variable cliométrique est visible dans le modèle final, mais il y a un risque de colinéarité avec les composantes principales. Ce risque de colinéarité obligerait à recourir à une méthode d'estimation adaptée (régressions de Ridge, Lasso, elastic-net, sélection *stewise* ou *recherche exhaustive* des variables, PCR, PLS), le cas échéant. Nous adopterons cette approche.

11.Résultats empiriques

11.1. Modèle non-cliométrique gaussien : logit PCR optimal

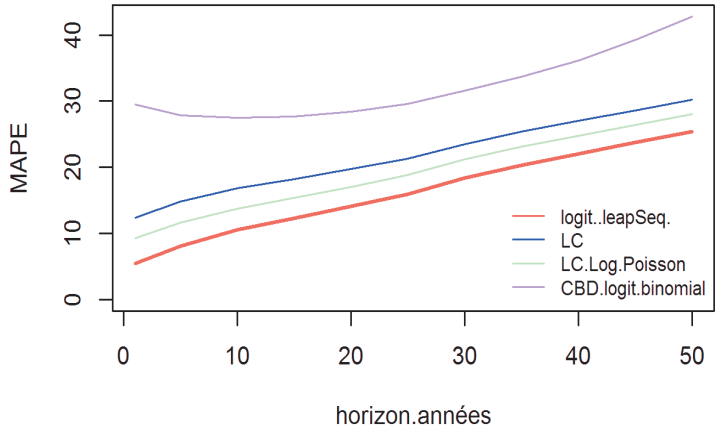
Les données par âge de la France et l'Italie sont issues de Human Mortality Database (2020).

11.1.1. Etude de cas : France

Variable étudiée : Probabilité de décès (hommes et femmes)
Ages : Tous les âges de 0 à 100 ans
Début de l'historique étudié : 1880
Année actuelle supposée : 1964
Horizons de prévision : de 1965 à 2014 (50 ans)
Correction de valeurs atypiques : Correction des taux de mortalités des années atypiques (1914 à 1918 et 1939 à 1945) des deux Guerres Mondiale par imputation avec interpolation linéaire.
Modèles comparés : Lee-carter (LC), LC Log-Poisson, CBD logit-binomial, logit PCR optimal (sélection *stepwise*)

MAPE tous âges confondus, selon les horizons testés

Figure 4.4 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision



Nous constatons un avantage systématique pour le modèle logit PCR optimal (sélection stepwise), quel que soit l'horizon de prévision étudié.

	MAPE (%)										
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	29,5	27,8	27,5	27,7	28,4	29,6	31,6	33,7	36,2	39,3	42,8
LC	12,4	14,8	16,8	18,2	19,7	21,3	23,5	25,4	27,0	28,6	30,2
LC.Log.Poisson	9,3	11,7	13,8	15,4	17,0	18,8	21,2	23,1	24,8	26,4	28,0
logit PCR optimal	5,5	8,1	10,5	12,3	14,1	15,9	18,4	20,3	22,0	23,7	25,4

Les baisses de MAPE obtenues par le modèle *logit PCR optimal* sont synthétisées ci-après :

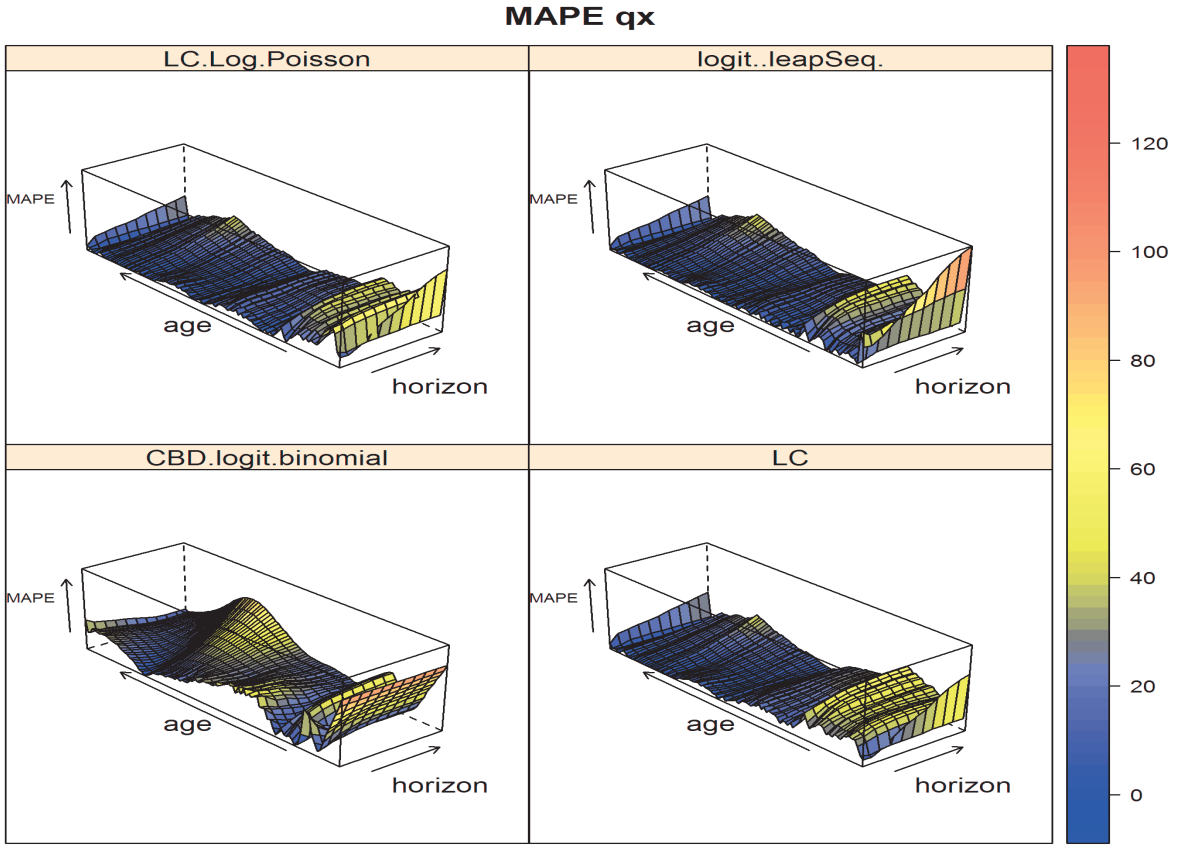
	Baisse MAPE : logit PCR optimal										
par rapport à :	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	-82%	-71%	-62%	-56%	-50%	-46%	-42%	-40%	-39%	-40%	-41%
LC	-56%	-45%	-37%	-32%	-29%	-25%	-22%	-20%	-18%	-17%	-16%
LC.Log.Poisson	-41%	-31%	-24%	-20%	-17%	-15%	-13%	-12%	-11%	-10%	-9%

Pour chaque horizon, les prévisions de toutes les années précédentes sont incluses dans le calcul du MAPE (et cela en considérant tous les âges). Chaque MAPE est donc calculé sur une fenêtre dont l'étendue augmente au fur et à mesure que l'horizon augmente : cela peut expliquer l'allure lissée des courbes MAPE.
(*logit..leapSeq* désigne le modèle *logit PCR optimal*)

MAPE des probabilités de décès selon les âges et les horizons

L'avantage tous âges confondus du modèle logit PCR optimal (sélection *stepwise*) quel que soit l'horizon envisagé est, toutefois, à nuancer pour les très jeunes âges comme l'illustrent les figures ci-après.

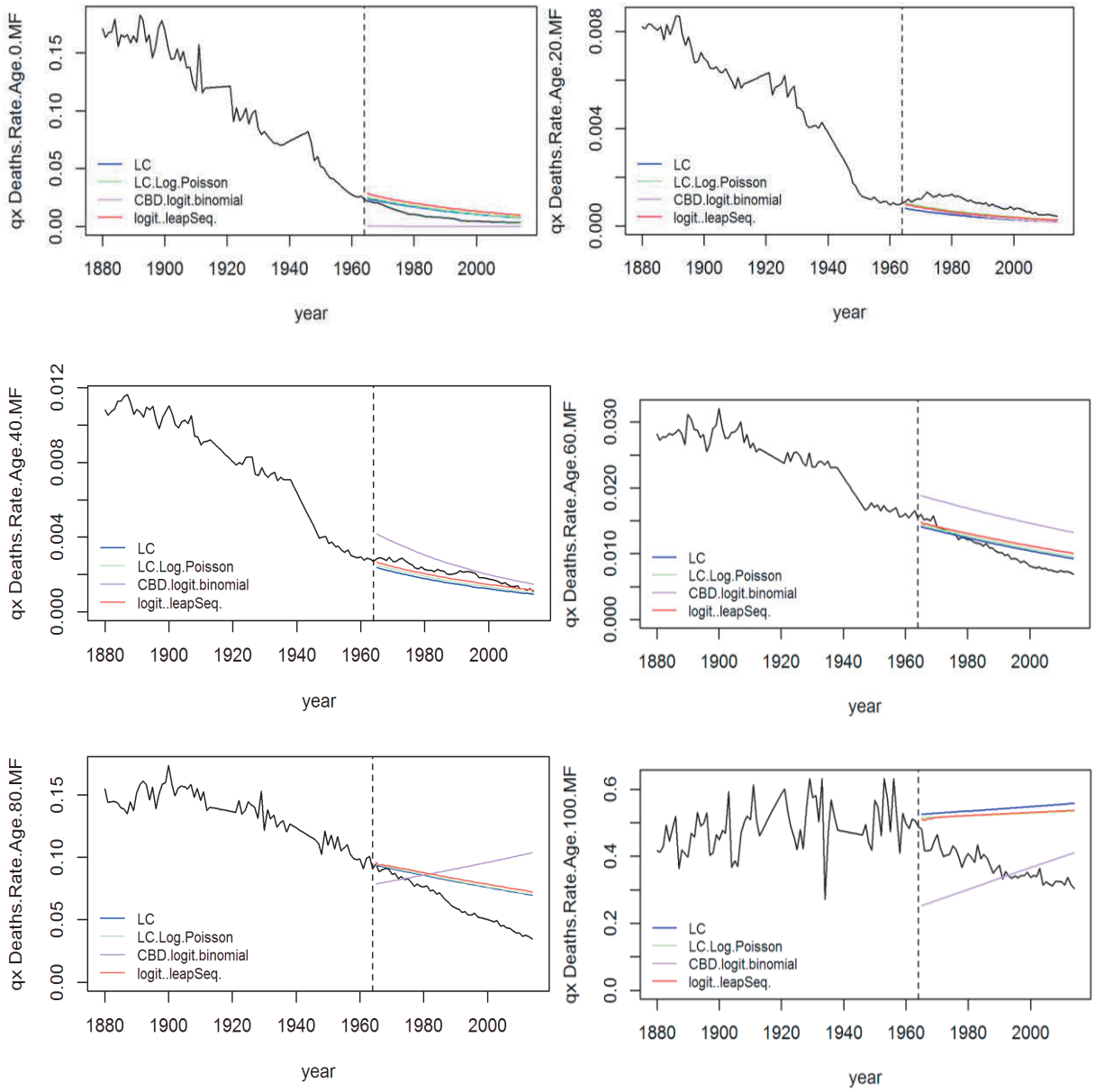
Figure 4.5 MAPE (%) des probabilités de décès selon les âges et les horizons



Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques âges : 0, 20, 40, 60, 80 et 100 ans.

Figure 4.6 Prévisions comparées des modèles pour quelques âges



Robustesse

Après réintroduction des taux de mortalités des années atypiques (1914 à 1918 et 1939 à 1945) des deux Guerres Mondiale, nous obtenons les performances suivantes :

Si les performances de tous les modèles ont baissé (à l'exception du CDB qui n'a sensiblement pas évolué), la dégradation la plus forte s'observe pour le modèle LC Log-Poisson dont les performances globales sont assez proches du LC (voire pire à partir de l'horizon de 45 ans).

Figure 4.7 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Présence de valeurs atypiques

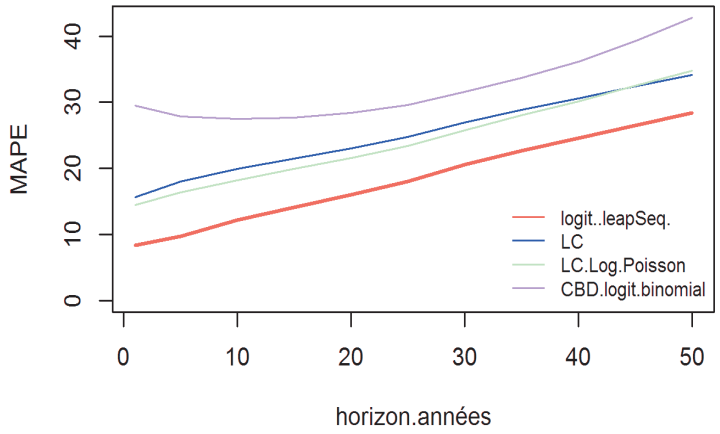
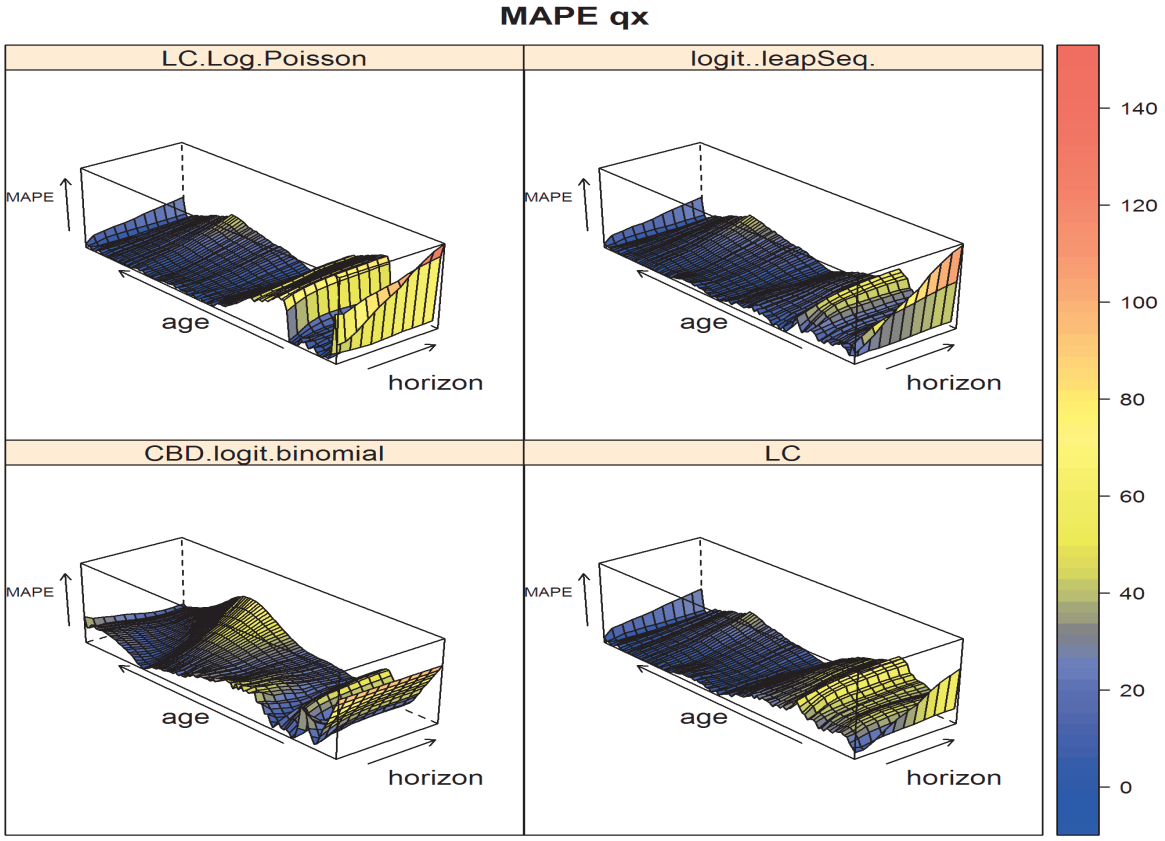


Figure 4.8 MAPE (%) des probabilités de décès selon les âges et les horizons. Présence de valeurs atypiques



	MAPE (%)										
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	29,5	27,8	27,5	27,7	28,4	29,6	31,6	33,7	36,2	39,3	42,8
LC	15,6	18,0	19,9	21,5	23,1	24,7	26,9	28,8	30,6	32,4	34,2
LC.Log.Poisson	14,5	16,3	18,2	19,9	21,6	23,4	25,8	28,0	30,2	32,5	34,8
logit PCR optimal	8,3	9,7	12,2	14,1	16,0	18,0	20,5	22,7	24,6	26,5	28,4

Les baisses de MAPE obtenues par le modèle *logit PCR optimal* sont synthétisées ci-après :

	Baisse MAPE : logit PCR optimal										
par rapport à :	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	-72%	-65%	-56%	-49%	-44%	-39%	-35%	-33%	-32%	-33%	-34%
LC	-47%	-46%	-39%	-34%	-30%	-27%	-24%	-21%	-20%	-18%	-17%
LC.Log.Poisson	-42%	-41%	-33%	-29%	-26%	-23%	-20%	-19%	-19%	-18%	-18%

Ainsi, en présence de valeurs atypiques, le modèle logit-PCR-optimal conserve globalement son avantage quel que soit l'horizon de prévision envisagé. Il accroit même cet avantage par rapport au LC et surtout le LC Log-Poisson.

En utilisant un historique plus court (20 ans, au lieu des 85 ans précédemment), i.e. commençant en 1945 au lieu de 1880, nous obtenons les évolutions suivantes des performances des modèles :

	Evolution MAPE : Début historique 1945 vs Début historique 1880										
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	-5%	-9%	-12%	-10%	-5%	1%	7%	12%	15%	17%	18%
LC	25%	34%	40%	44%	45%	45%	42%	41%	40%	39%	38%
LC.Log.Poisson	2%	18%	27%	32%	33%	34%	31%	30%	30%	29%	29%
logit PCR optimal	-23%	22%	31%	34%	33%	32%	29%	28%	27%	26%	25%

Figure 4.9 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Historique d'apprentissage de 20 ans au lieu de 85 ans

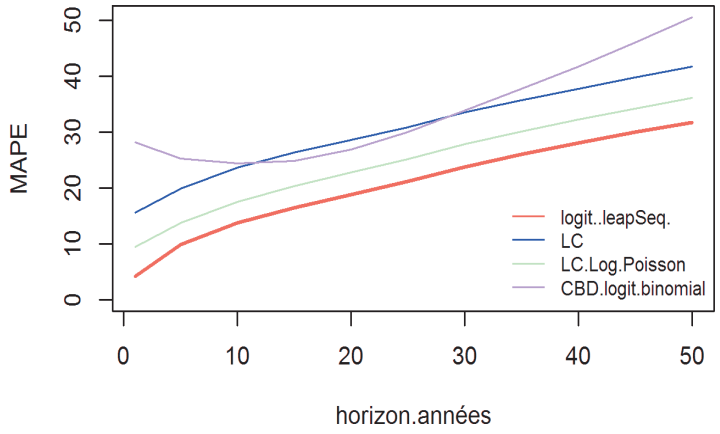
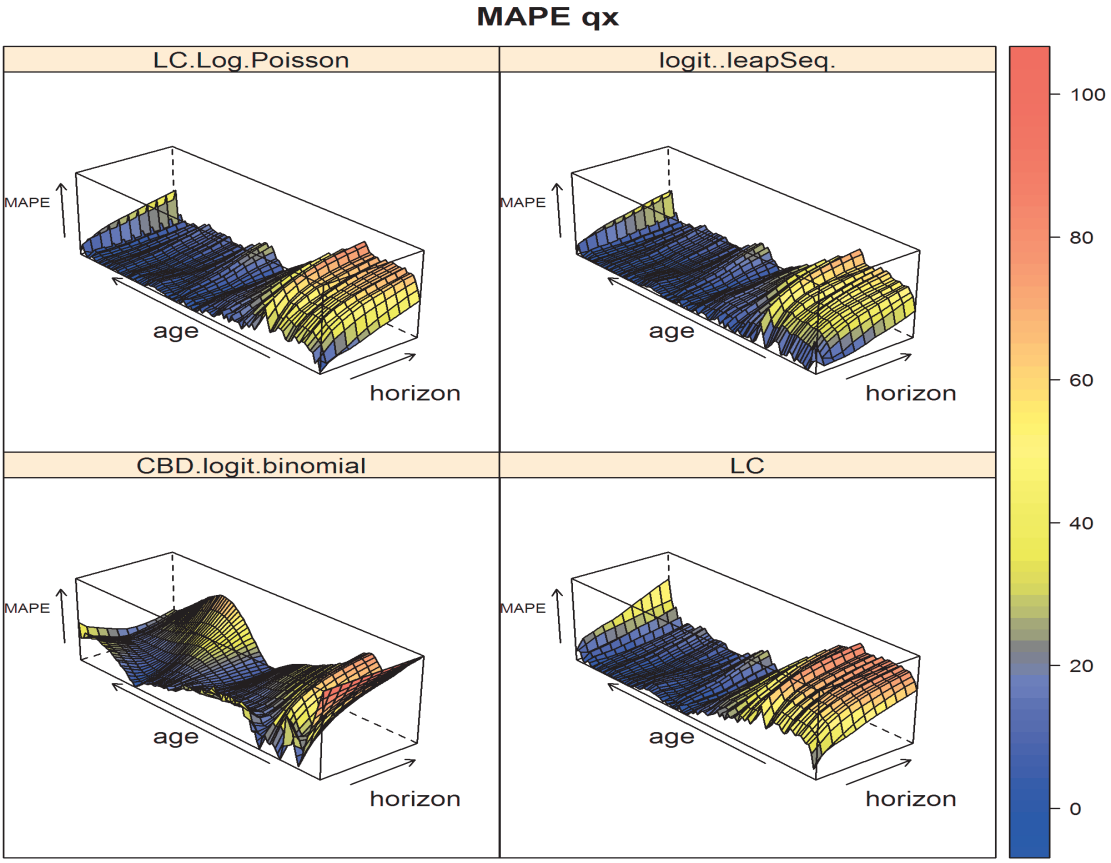


Figure 4.10 MAPE (%) des probabilités de décès selon les âges et les horizons. Historique d'apprentissage de 20 ans au lieu de 85 ans



A l'exception des modèles CBD et logit PCR optimal qui ont une amélioration aux faibles horizons, l'historique d'apprentissage de 20 ans se révèle moins prédictif que celui de 85 ans précédemment testé. Toutefois, l'ordre de performance des modèles reste globalement identique, et le modèle logit PCR optimal reste le meilleur.

11.1.2. Etude de cas : Italie

Variable étudiée : Taux de mortalité mixte (hommes et femmes)
Ages : Tous les âges de 0 à 100 ans
Début de l'historique étudié : 1872
Année actuelle supposée : 1964
Horizons de prévision : de 1965 à 2014 (50 ans)
Correction de valeurs atypiques : Correction des taux de mortalités des années atypiques (1914 à 1920 et 1939 à 1945) des deux Guerres Mondiale par imputation avec interpolation linéaire.
Modèles comparés : Lee-carter et logit PCR optimal (sélection stepwise)

MAPE tous âges confondus, selon les horizons testés

Comme pour la France, le modèle CBD se révèle être le moins performant des modèles testés.

Figure 4.11 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision

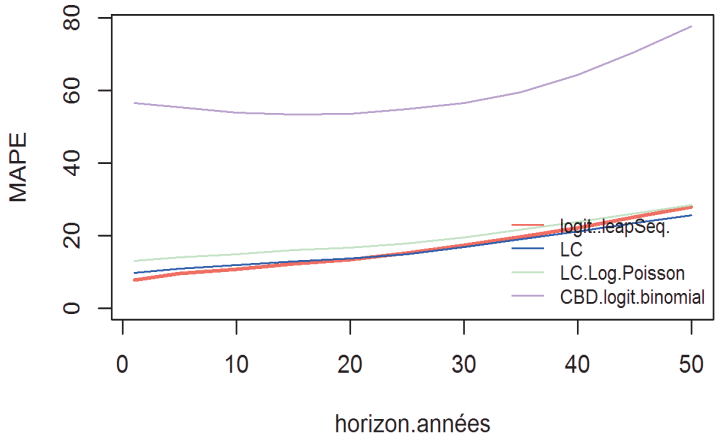
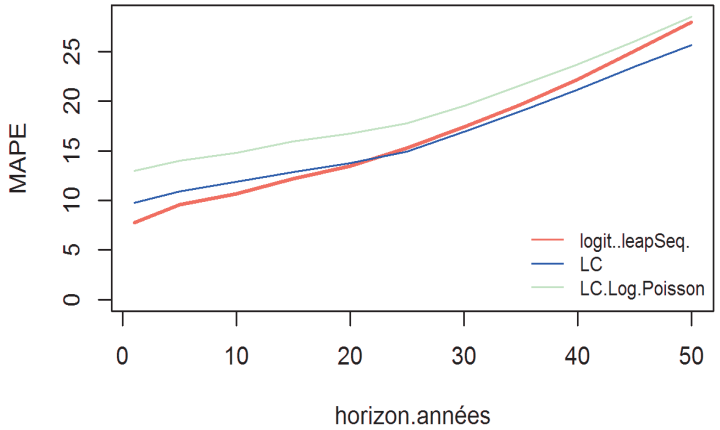


Figure 4.12 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Sans le modèle CBD



(*logit..leapSeq* désigne le modèle *logit PCR optimal*)

Nous constatons donc que, contrairement au cas Français où le modèle logit PCR optimal (sélection stepwise) est le plus performant jusqu'à 50 ans d'horizon de prévision, cet avantage n'excède pas l'horizon prédictif de 25 ans face au modèle LC.
Toutefois, l'avantage systématique de performance modèle logit PCR optimal est conservé vis-à-vis des deux autres modèles testés, i.e. LC.Log.Poisson et CBD.logit.binomial.
Autre différence avec le cas Français, le modèle LC performe systématiquement mieux que LC.Log.Poisson, quel que soit l'horizon prédictif testé (maximum 50 ans).

Les baisses de MAPE obtenues par le modèle *logit PCR optimal* sont synthétisées ci-après :

Performances comparées des modèles testés

	MAPE (%)										
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	56,6	55,4	54,0	53,4	53,6	54,9	56,5	59,5	64,3	70,6	77,7
LC	9,7	10,9	11,9	12,9	13,8	14,9	16,9	19,0	21,2	23,5	25,7
LC.Log.Poisson	13,0	14,0	14,8	15,9	16,8	17,8	19,5	21,6	23,7	26,0	28,5
logit PCR optimal	7,8	9,6	10,7	12,2	13,4	15,3	17,4	19,7	22,2	25,1	28,0

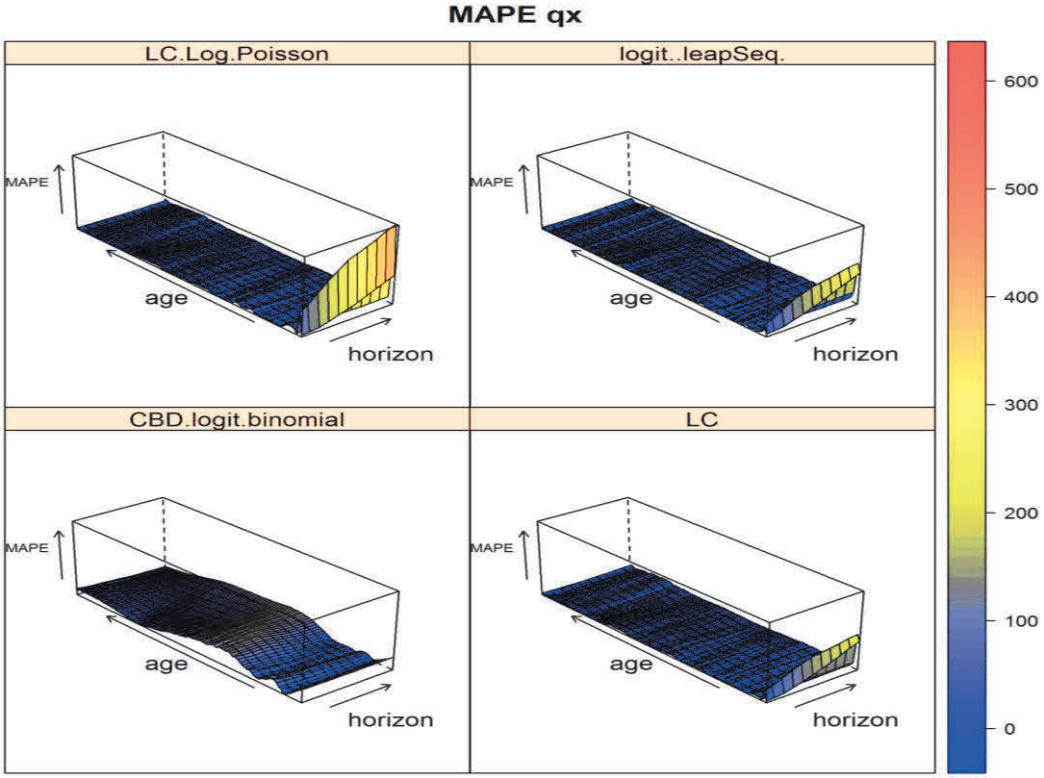
	Evolution du MAPE : logit PCR optimal										
par rapport à :	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	-86%	-83%	-80%	-77%	-75%	-72%	-69%	-67%	-65%	-65%	-64%
LC	-20%	-12%	-10%	-5%	-2%	2%	3%	4%	5%	7%	9%
LC.Log.Poisson	-40%	-32%	-28%	-23%	-20%	-14%	-11%	-9%	-6%	-4%	-2%

Rappel : Pour chaque horizon, les prévisions de toutes les années précédentes sont incluses dans le calcul du MAPE (et cela en considérant tous les âges). Chaque MAPE est donc calculé sur une fenêtre dont l'étendue augmente au fur et à mesure que l'horizon augmente : cela peut expliquer l'allure lissée des courbes MAPE.

MAPE des probabilités de décès selon les âges et les horizons

L'avantage tous âges confondus du modèle logit PCR optimal (sélection stepwise) quel que soit l'horizon envisagé est, toutefois, à nuancer pour les très jeunes âges comme l'illustrent les figures ci-après.

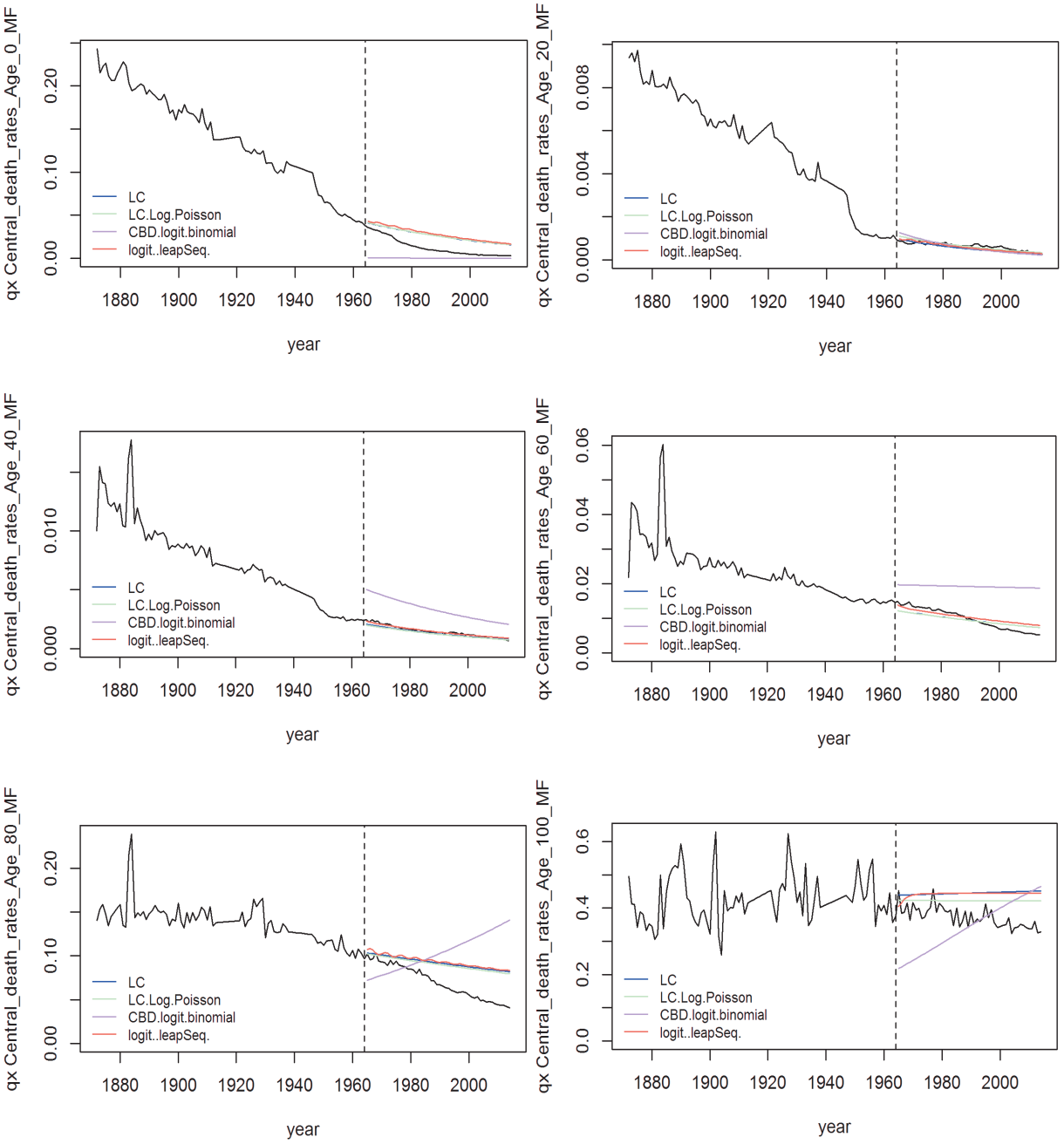
Figure 4.13 MAPE (%) des probabilités de décès selon les âges et les horizons



Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques âges : 0, 20, 40, 60, 80 et 100 ans.

Figure 4.14 Prévisions comparées des modèles pour quelques âges



Robustesse

Après réintroduction des taux de mortalités des années atypiques (1914 à 1918 et 1939 à 1945) des deux Guerres Mondiale, nous obtenons les performances ci-après.

Si les performances de tous les modèles ont baissé (à l'exception du CDB qui n'a sensiblement pas évolué), la dégradation la plus forte s'observe pour le modèle logit PCR optimal.

Figure 4.15 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Présence de valeurs atypiques

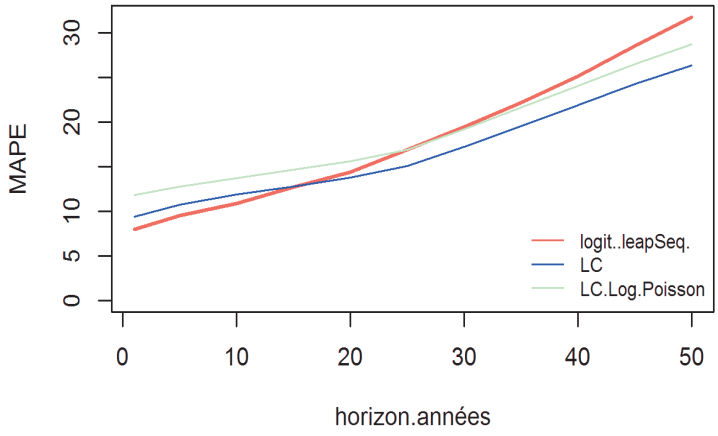
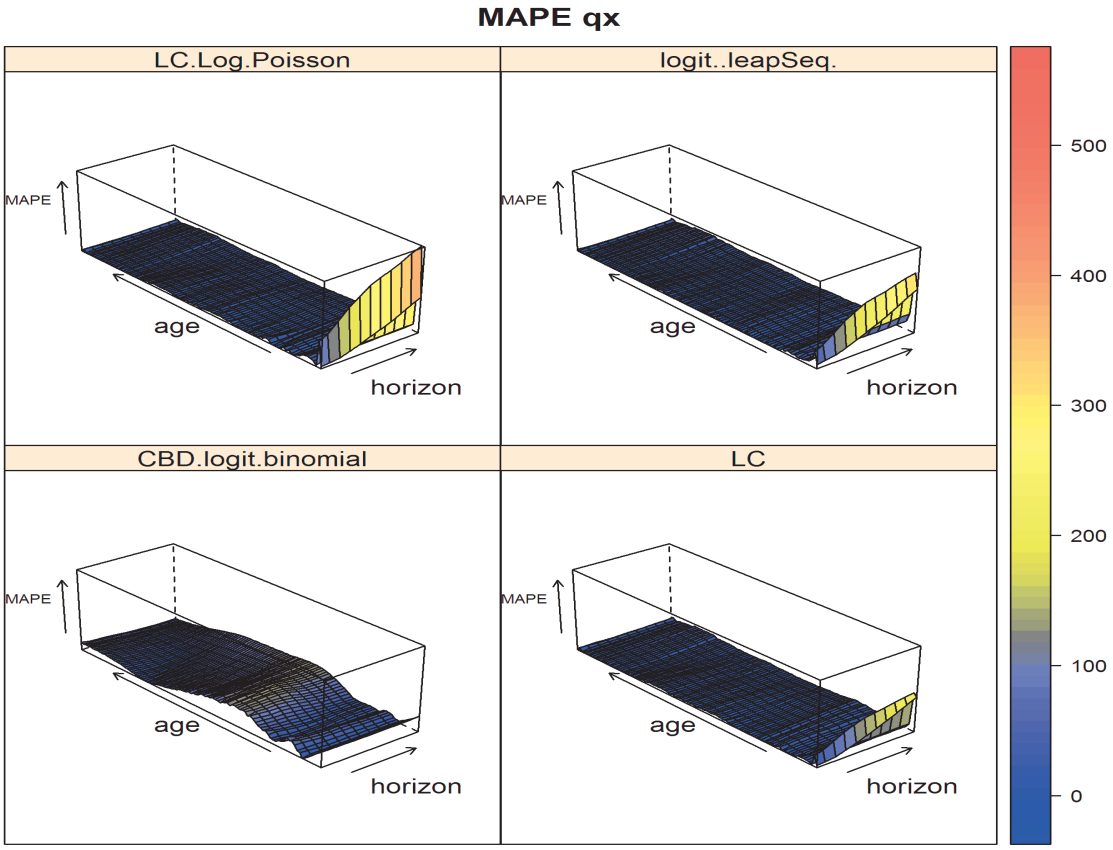


Figure 4.16 MAPE (%) des probabilités de décès selon les âges et les horizons. Présence de valeurs atypiques



	MAPE (%)										
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	56,6	55,4	54,0	53,4	53,6	54,9	56,5	59,5	64,3	70,6	77,7
LC	9,3	10,7	11,9	12,8	13,7	15,1	17,2	19,5	21,8	24,2	26,3
LC.Log.Poisson	11,8	12,8	13,7	14,7	15,6	16,9	19,2	21,6	24,0	26,5	28,7
logit PCR optimal	8,0	9,5	10,8	12,7	14,4	16,9	19,4	22,1	25,1	28,5	31,7

Les baisses de MAPE obtenues par le modèle *logit PCR optimal* sont synthétisées ci-après :

	Ecart MAPE : logit PCR optimal										
par rapport à :	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	-86%	-83%	-80%	-76%	-73%	-69%	-66%	-63%	-61%	-60%	-59%
LC	-15%	-11%	-9%	-1%	5%	12%	13%	14%	15%	18%	20%
LC.Log.Poisson	-33%	-26%	-21%	-13%	-8%	0%	2%	3%	5%	8%	11%

Ainsi, en présence de valeurs atypiques, le modèle logit-PCR-optimal perd son cet avantage par rapport au LC (après l'horizon prédictif de 15 ans) et le LC Log-Poisson (dès l'horizon prédictif de 25 ans).

En utilisant un historique plus court (20 ans, au lieu des 93 ans précédemment), i.e. commençant en 1945 au lieu de 1872, nous obtenons les évolutions suivantes des performances des modèles :

	Evolution MAPE : Début historique 1945 vs Début historique 1872										
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**	**45 ans**	**50 ans**
CBD.logit.binomial	81%	67%	50%	38%	31%	30%	32%	37%	41%	45%	47%
LC	-7%	-3%	5%	12%	15%	18%	19%	22%	24%	27%	29%
LC.Log.Poisson	-18%	-11%	0%	7%	11%	15%	17%	20%	23%	26%	29%
logit PCR optimal	23%	12%	15%	20%	21%	21%	21%	23%	26%	28%	29%

Figure 4.17 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Historique d'apprentissage de 20 ans au lieu de 93 ans

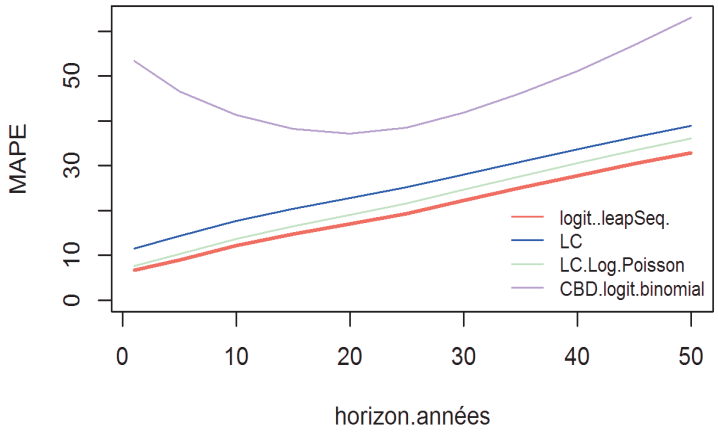
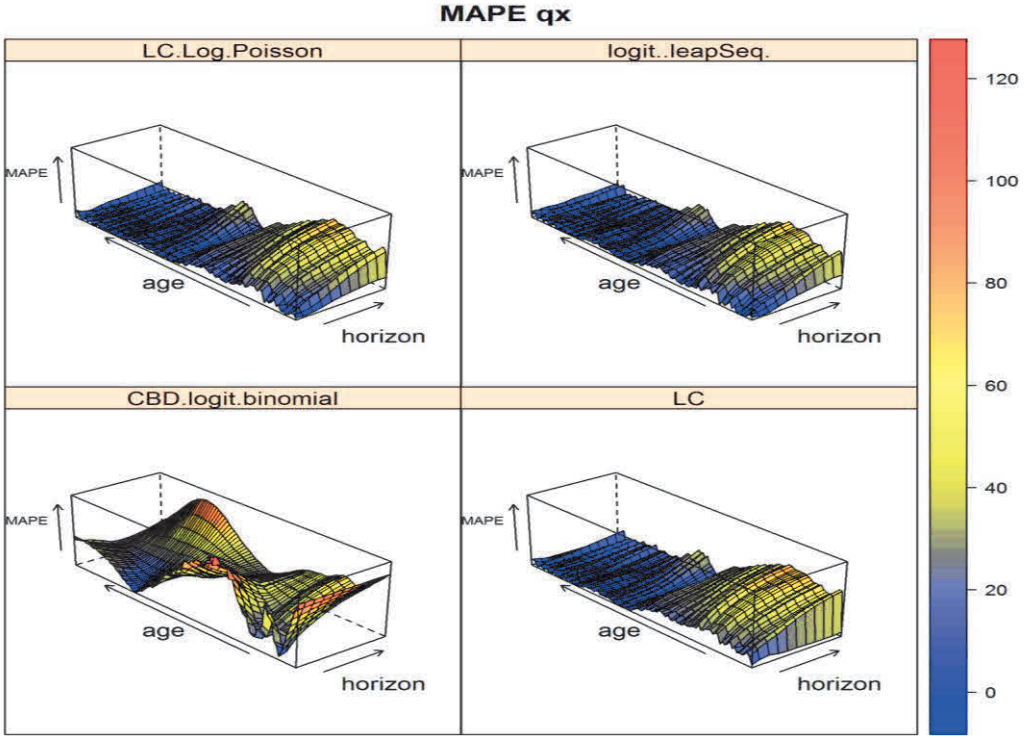


Figure 4.18 MAPE (%) des probabilités de décès selon les âges et les horizons. Historique d'apprentissage de 20 ans au lieu de 93 ans



A l'exception des modèles LC et LC Log-Poisson qui ont une amélioration aux faibles horizons, l'historique d'apprentissage de 20 ans se révèle moins prédictif que celui de 93 ans précédemment testé. Toutefois contrairement aux modélisations basées sur l'historique d'apprentissage de 93 ans, le modèle logit PCR optimal devient à présent (20 ans d'historique) le meilleur quel que soit l'horizon de prévision.

11.1.3. Conclusion sur le modèle non-cliométrique gaussien logit PCR optimal

La supériorité empirique de performance du modèle logit-PCR-optimal par rapport aux modèles de référence LC et LC-Log-Poisson, est toute relative selon le pays, l'horizon de prévision et l'âge (on peut conjecturer sur la période étudiée également).
Ainsi, il revient à l'utilisateur de sélectionner les conditions d'utilisation du modèle logit-PCR-optimal, selon les résultats des prévisions tests.

11.2. Modèles cliométriques

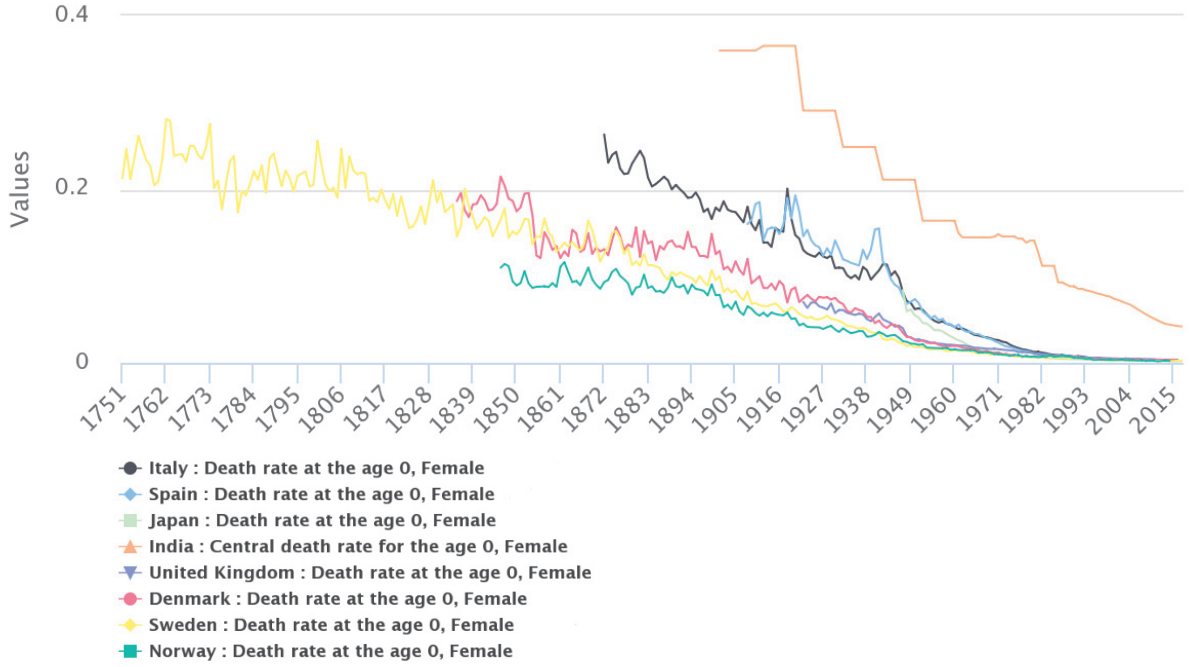
Les données par tranche d'âge de l'Inde et l'Italie sont issues de Max Planck Institute for Demographic Research (2020).

Variable étudiée : Probabilité de décès (femmes)
Ages étudiées : Age 0, et tranches d'âge de 1 à 80 ans

Pays transitionnellement « moins âgé » à modéliser, pays « R » : Inde
- Début de l'historique étudié : 1901
- Année actuelle supposée : 1973
- Horizons de prévision : de 1974 à 2013 (40 ans)

Pays transitionnellement « plus âgé », pays « A» : Italie. L'Italie étant un des derniers pays européens (donc transitionnellement « agés ») à entrer en transition démographique, elle fait donc partie des pays européens les plus « proches » transitionnellement de l'Inde, comme l'illustre le graphique ci-après :

Figure 4.19 Taux de mortalité à l'âge 0 des femmes. Quelques pays européens vs Inde

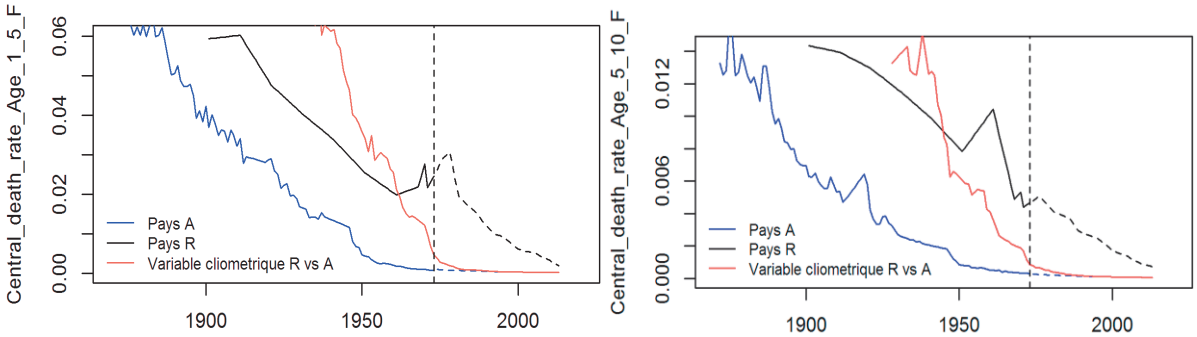


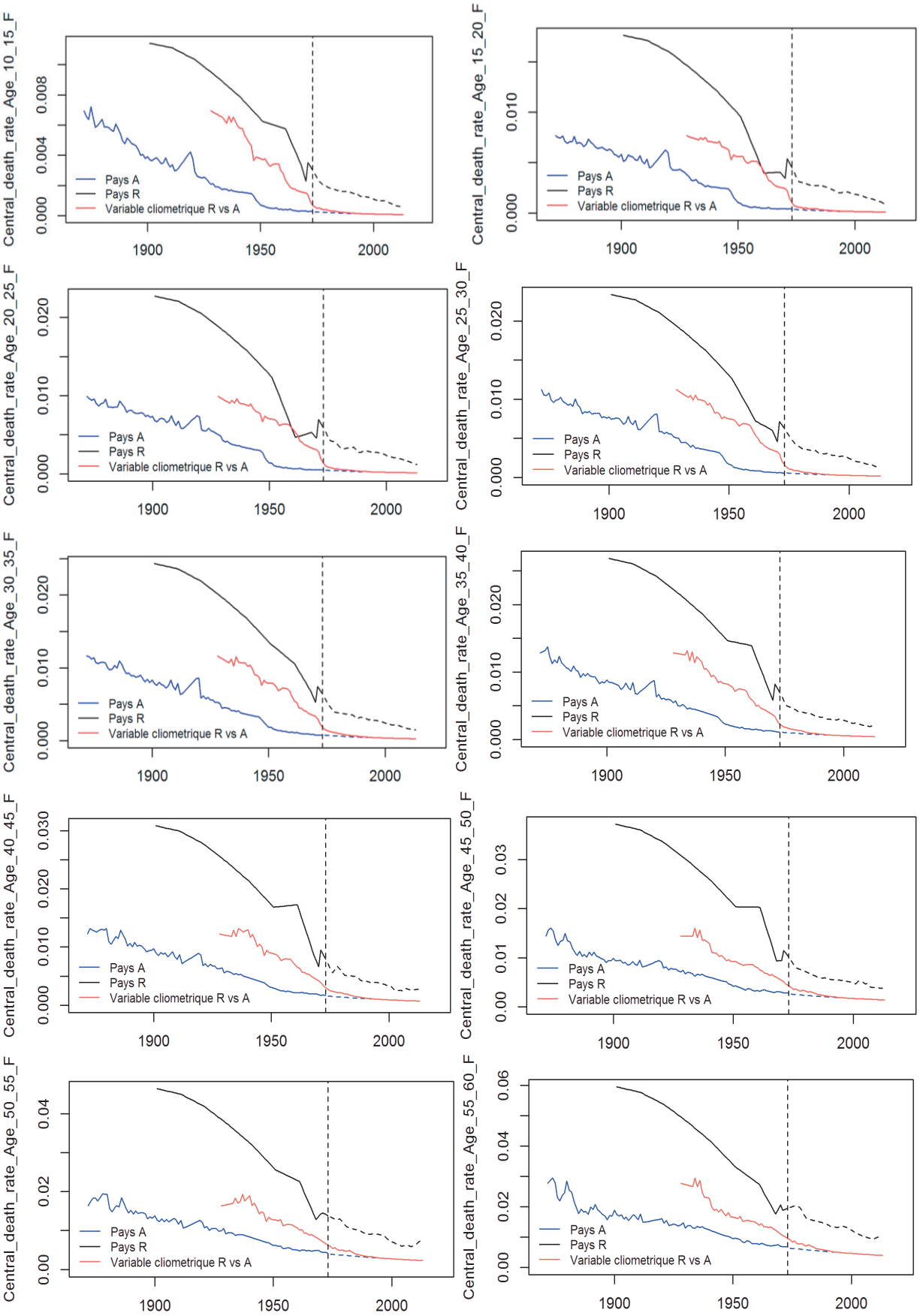
Nous utilisons comme symétrique de la variable de temps transitionnel, le taux de mortalité à l'âge 0 ($x_0 = 0$).

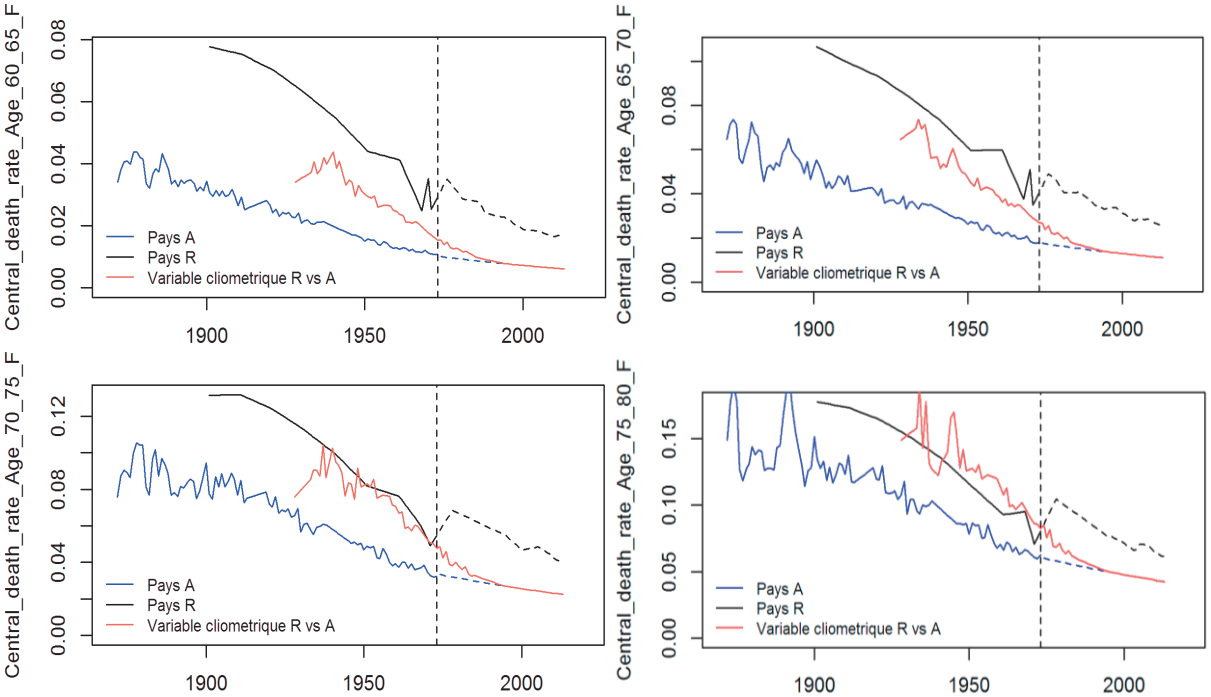
11.2.1. Variables cliométriques des mortalités de l'Inde par rapport à l'Italie (Méthode 2)

Les figures ci-dessous permettent de comparer selon les tranches d'âge, les allures des taux de mortalité italiens (bleu), indiens (noir), ainsi que la variable cliométrique (rouge) déduite pour l'inde.
Les variables cliométriques sont calculées avec la méthode d'estimation 2, exposée à la section 10.2.3 .
Ces variables cliométriques seront utilisées comme candidates explicatives dans les modèles de mortalité de l'Inde, que nous testerons dans la section suivante.

Figure 4.20 Courbes de variables cliométriques selon l'âge Inde




Central_death_rate_Age_10_15_F
Central_death_rate_Age_15_20_F
Central_death_rate_Age_20_25_F
Central_death_rate_Age_25_30_F
Central_death_rate_Age_30_35_F
Central_death_rate_Age_35_40_F
Central_death_rate_Age_40_45_F
Central_death_rate_Age_45_50_F
Central_death_rate_Age_50_55_F
Central_death_rate_Age_55_60_F
Pays A
Pays R
Variable cliometrique R vs A
1900
1950
2000



11.2.2. Comparaison des modèles pour l'Inde

Modèles de référence : Lee-Carter, Lee-Carter Log-Poisson, CBD logit binomial

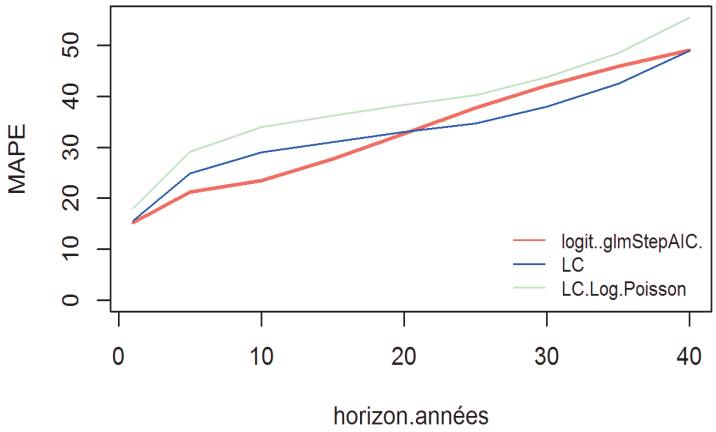
Modèles alternatifs (5) testés :

- Modèle logit-binomial-PCR optimal (sélection par stepwise et AIC), non-cliométrique : logit-binomial-PCR-optimal *(logit..glmStepAIC)*
- Modèle cliométrique logit-binomial-PCR-optimal, avec la variable cliométrique ($\mathcal{C}_{R,A,x}$) : logit-binomial-PCR-optimal Cliometric *(logit..glmStepAIC.Cliometric)*
- Modèle cliométrique PCR optimale (sélection par stepwise et MSE crossvalidée), avec la variable cliométrique ($\mathcal{C}_{R,A,x}$) : logit-PCR-optimal-Cliometric *(logit..leapSeq.Cliometric)*
- Modèle cliométrique PCR emboité, avec la variable cliométrique ($\mathcal{C}_{R,A,x}$) : logit-PCR-Cliometric *(logit..pcr.Cliometric)*
- Modèle cliométrique PLS, avec la variable cliométrique ($\mathcal{C}_{R,A,x}$) candidate explicative : logit-PLS-Cliometric *(logit..pls.Cliometric)*

Modèle logit-Binomial-PCR-optimal (non-cliométrique)

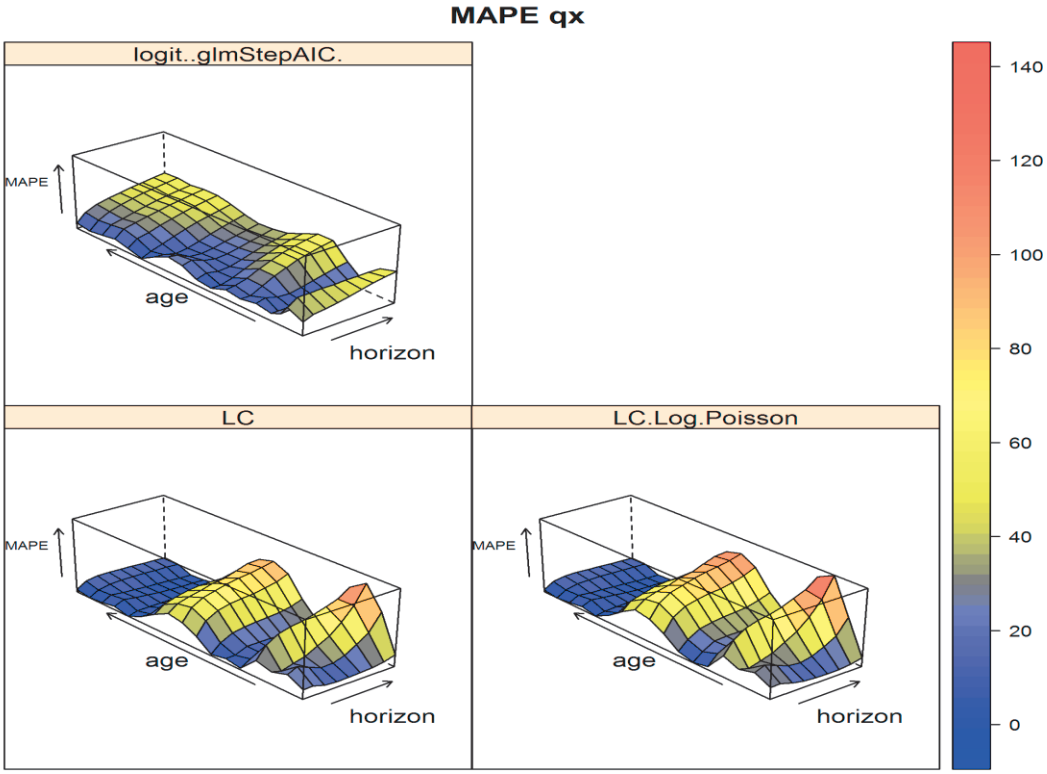
MAPE tous âges confondus, selon quelques horizons

Figure 4.21 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision



MAPE selon les âges et les horizons

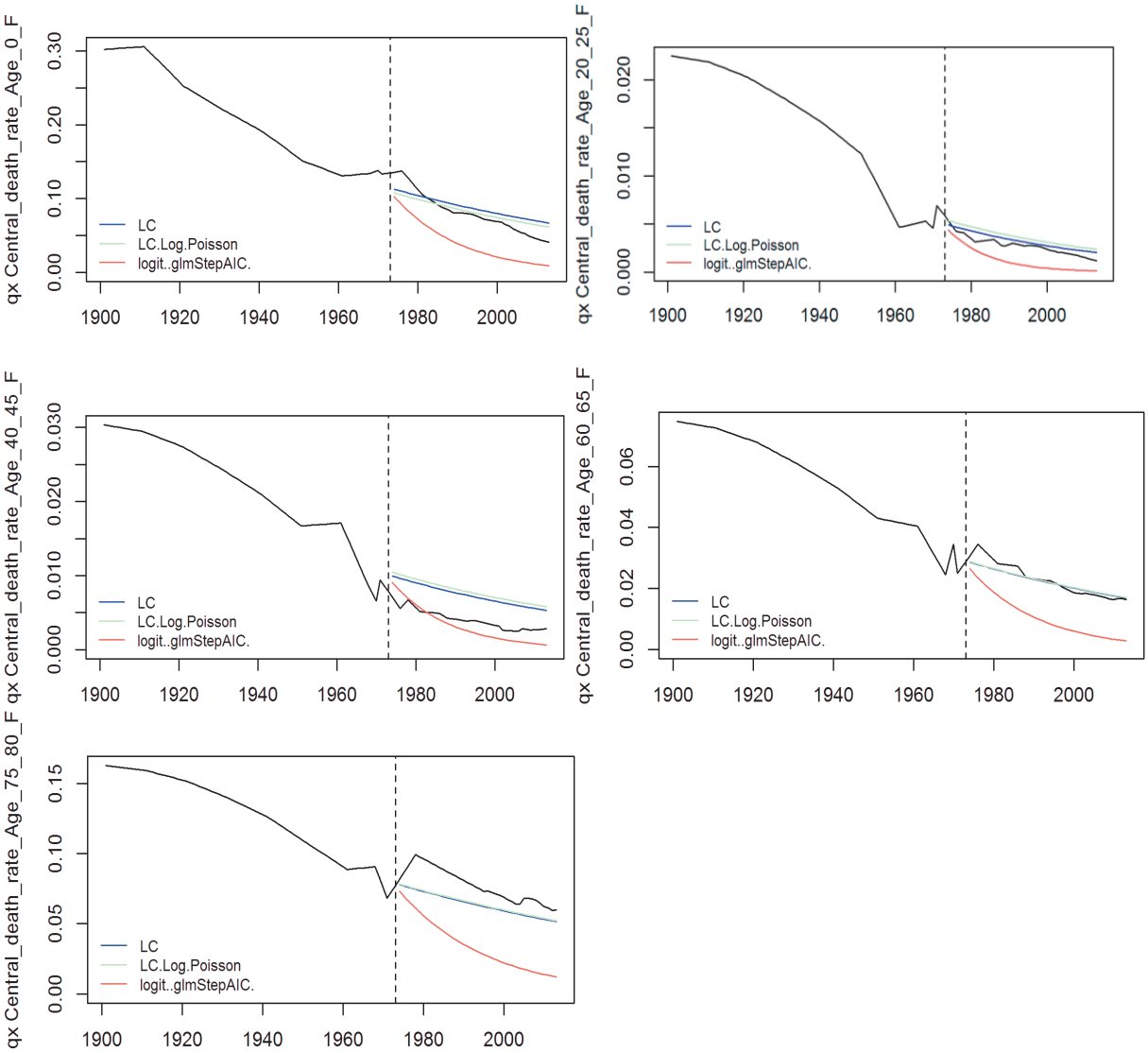
Figure 4.22 MAPE des modèles selon les âges et les horizons



Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques tranches d'âge : 0, [20 ; 25[, [40 ; 45[, [60 ; 65[, [75 ; 80[

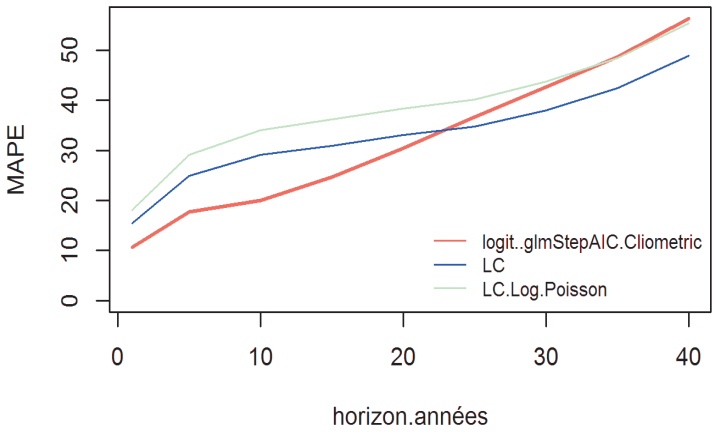
Figure 4.23 Prévisions comparées des modèles, pour quelques âges



Modèle logit-binomial-PCR-optimal Cliometric

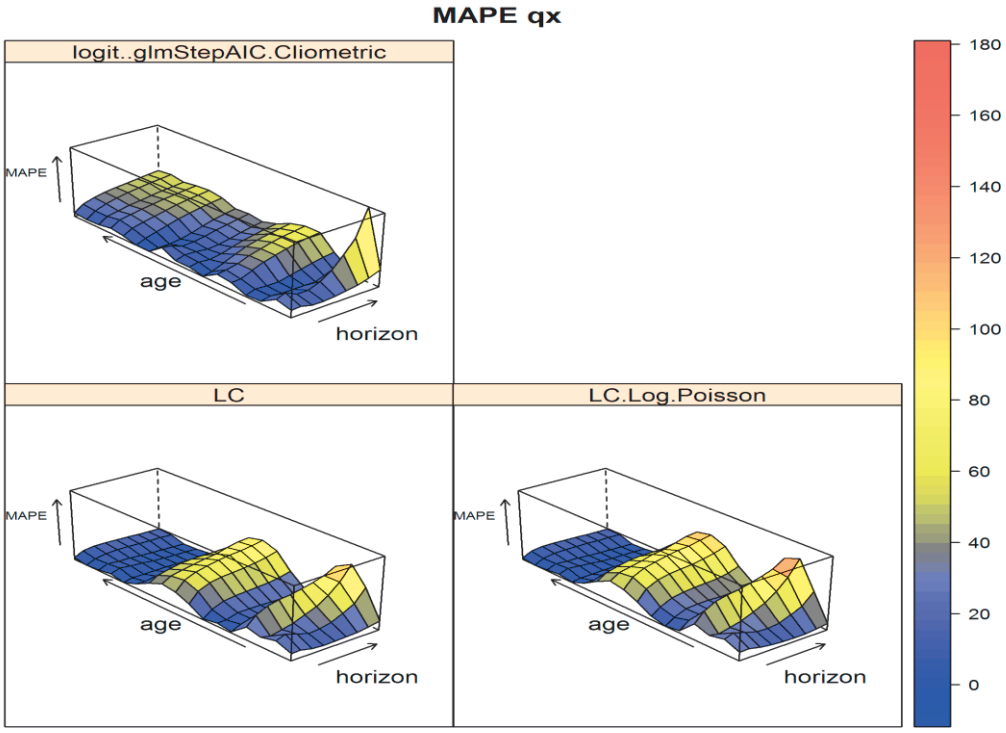
MAPE tous âges confondus, selon quelques horizons

Figure 4.24 MAPE (%) tous âges confondus, selon quelques horizons



MAPE selon les âges et les horizons

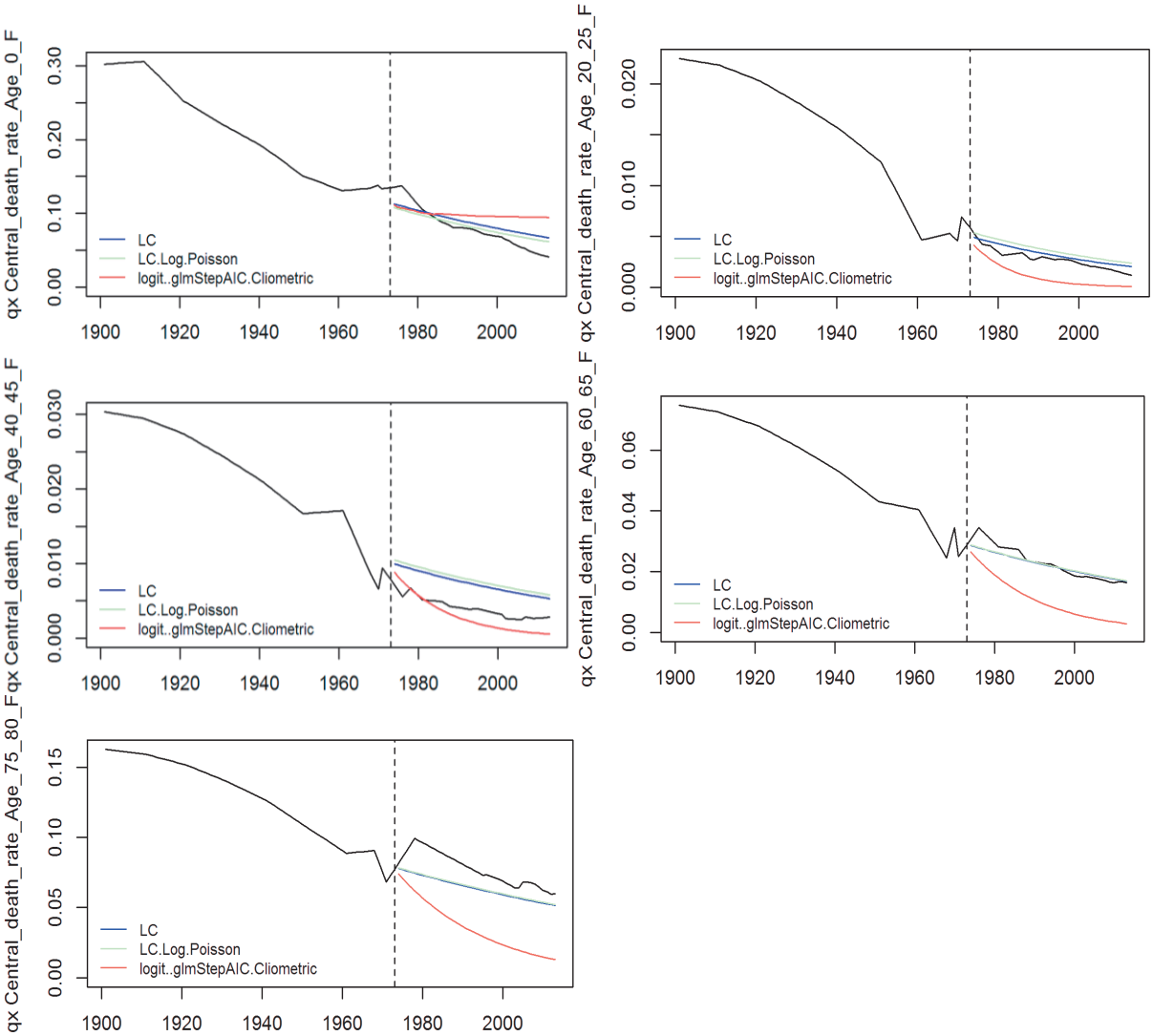
Figure 4.25 MAPE (%) selon les âges et les horizons



Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques tranches d'âge : 0, [20 ; 25[, [40 ; 45[, [60 ; 65[, [75 ; 80[

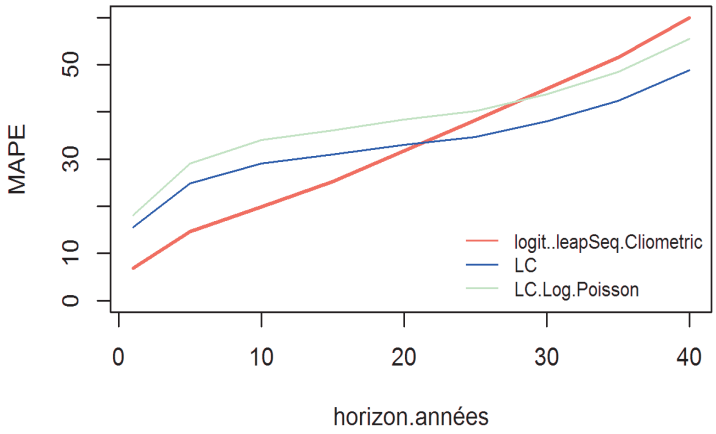
Figure 4.26 Prévisions comparées des modèles, pour quelques âges



Modèle logit-PCR-optimal-Cliometric

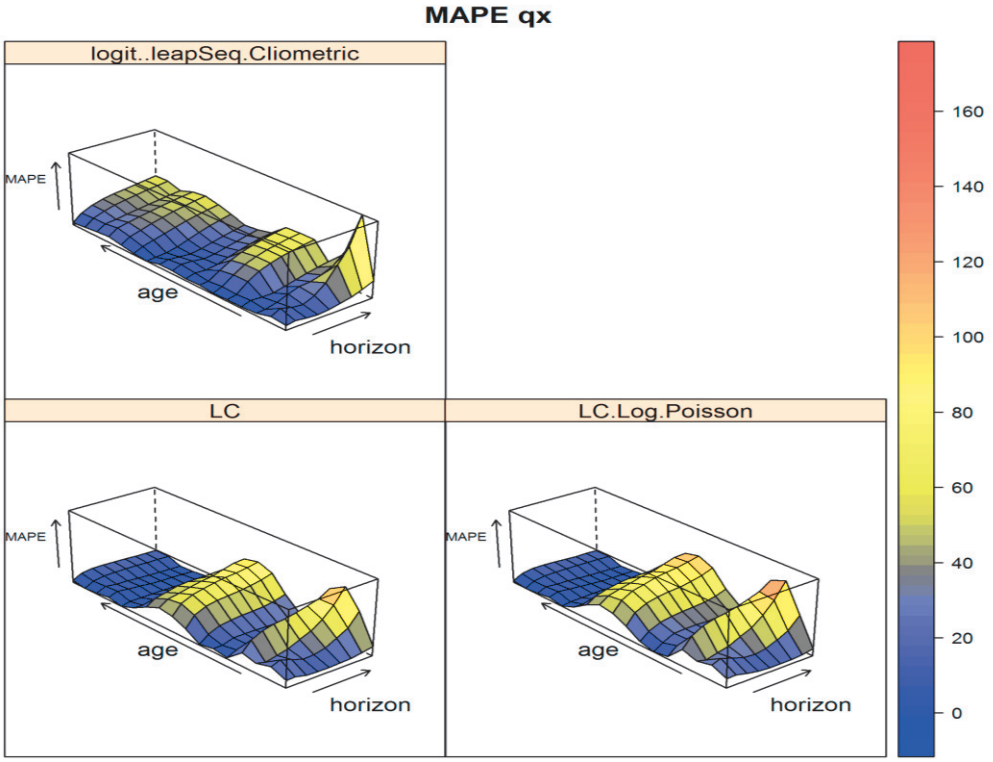
MAPE tous âges confondus, selon quelques horizons

Figure 4.27 MAPE (%) tous âges confondus, selon quelques horizons



MAPE selon les âges et les horizons

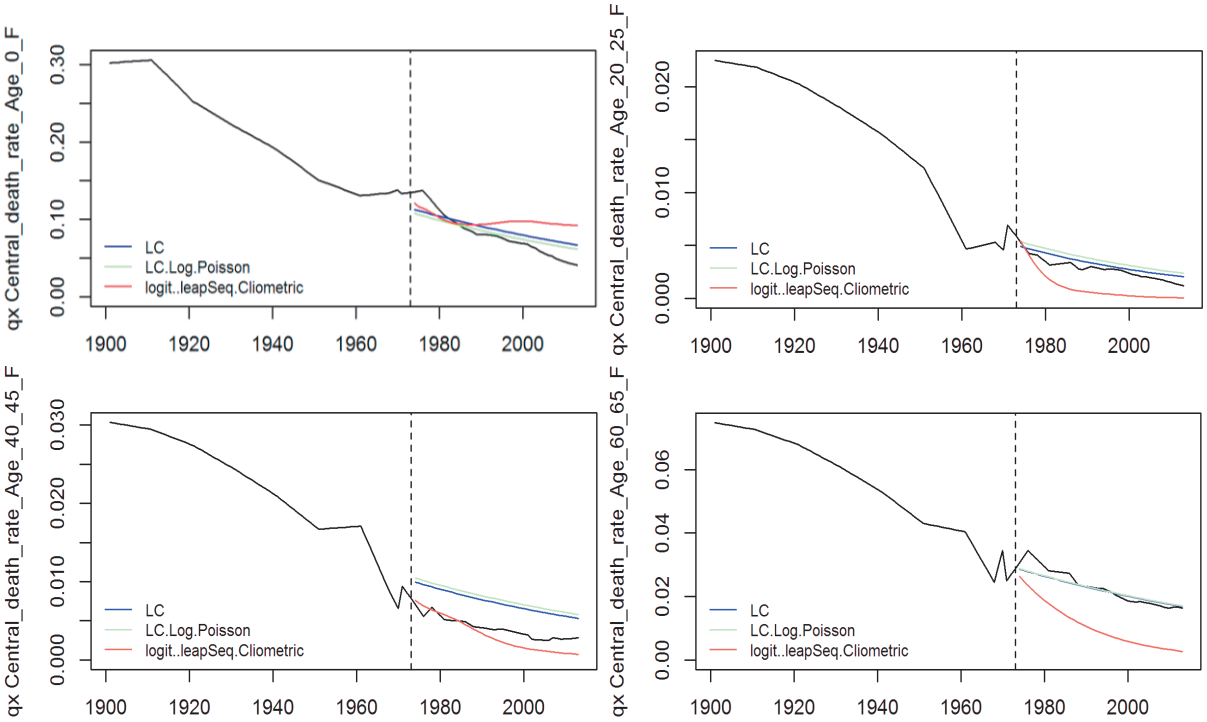
Figure 4.28 MAPE (%) selon les âges et les horizons

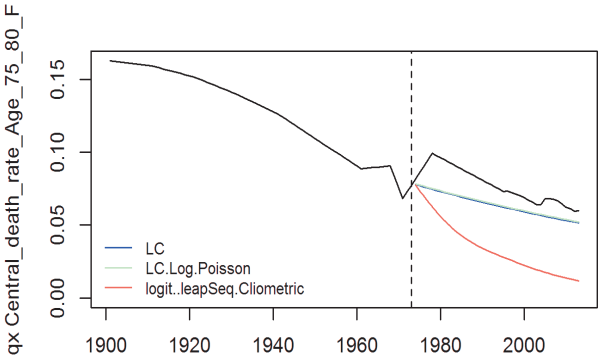


Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques tranches d'âge : 0, [20 ; 25[, [40 ; 45[, [60 ; 65[, [75 ; 80[

Figure 4.29 Prévisions comparées des modèles, pour quelques âges

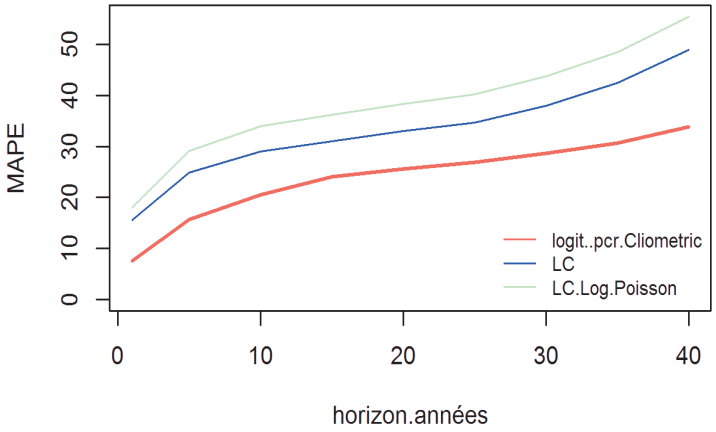




Modèle logit-PCR-Cliometric

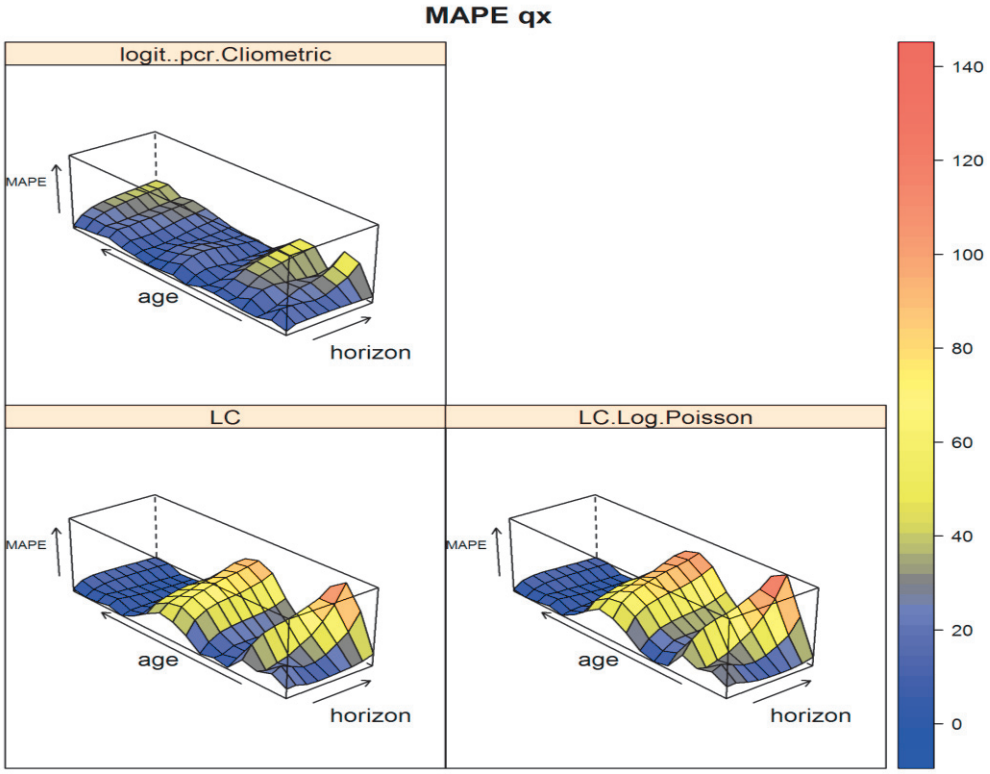
MAPE tous âges confondus, selon quelques horizons

Figure 4.30 MAPE (%) tous âges confondus, selon quelques horizons



MAPE selon les âges et les horizons

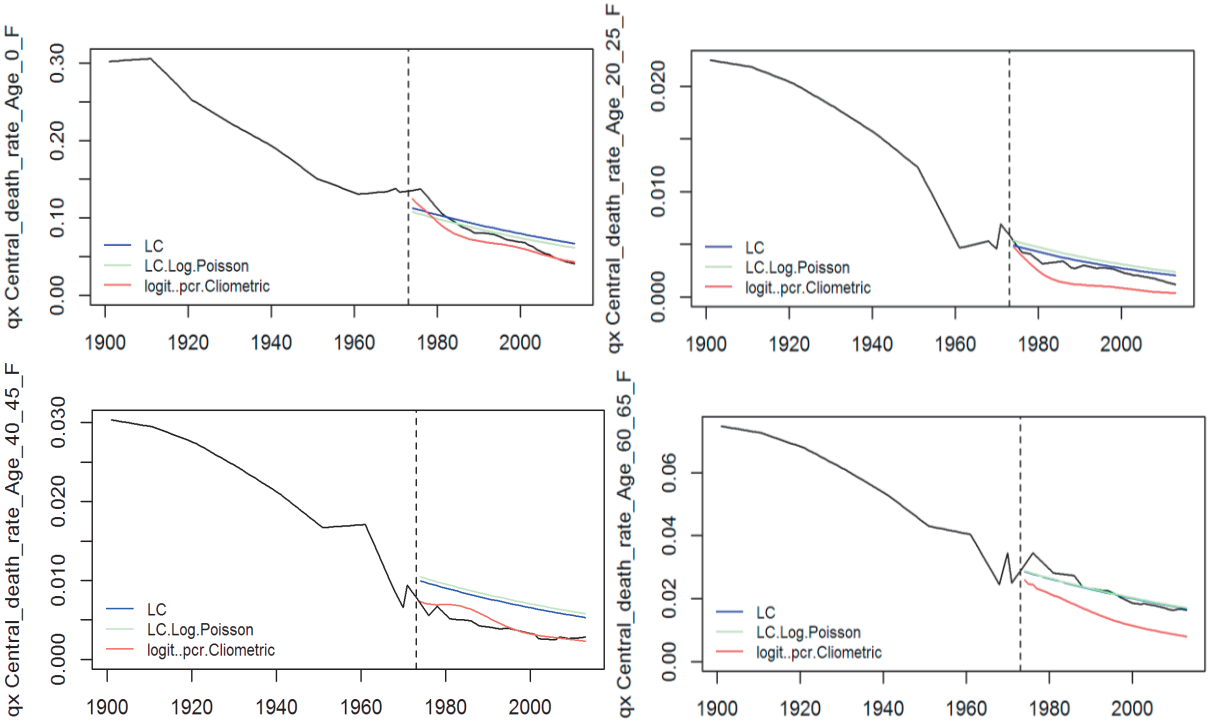
Figure 4.31 MAPE (%) selon les âges et les horizons

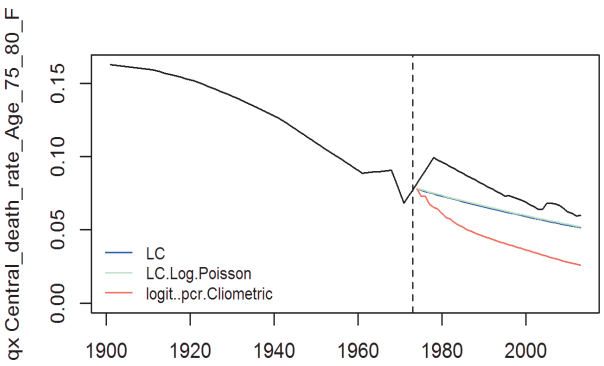


Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques tranches d'âge : 0, [20 ; 25[, [40 ; 45[, [60 ; 65[, [75 ; 80[

Figure 4.32 Prévisions comparées des modèles, pour quelques âges

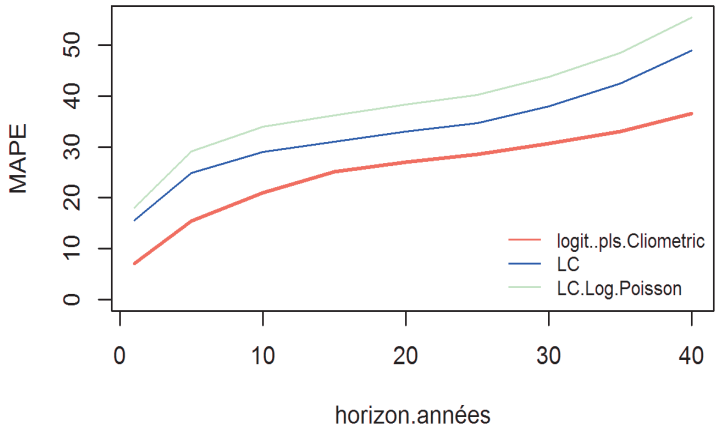




Modèle logit-PLS-Cliometric

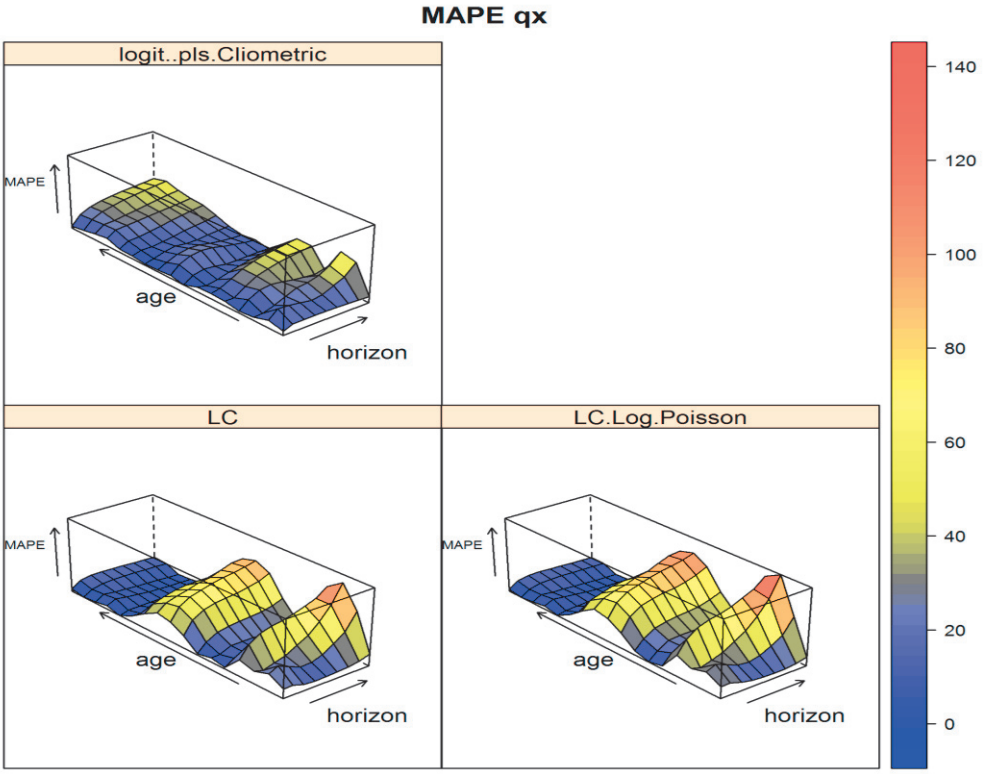
MAPE tous âges confondus, selon quelques horizons

Figure 4.33 MAPE (%) tous âges confondus, selon quelques horizons



MAPE selon les âges et les horizons

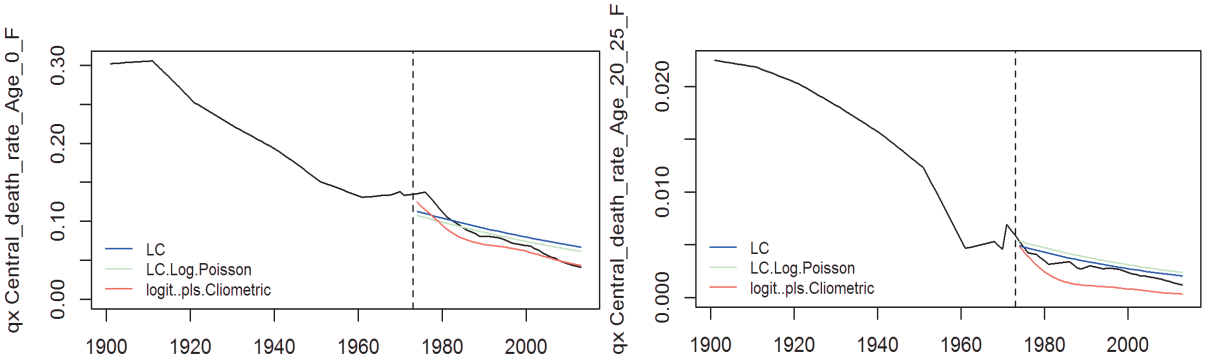
Figure 4.34 MAPE (%) selon les âges et les horizons

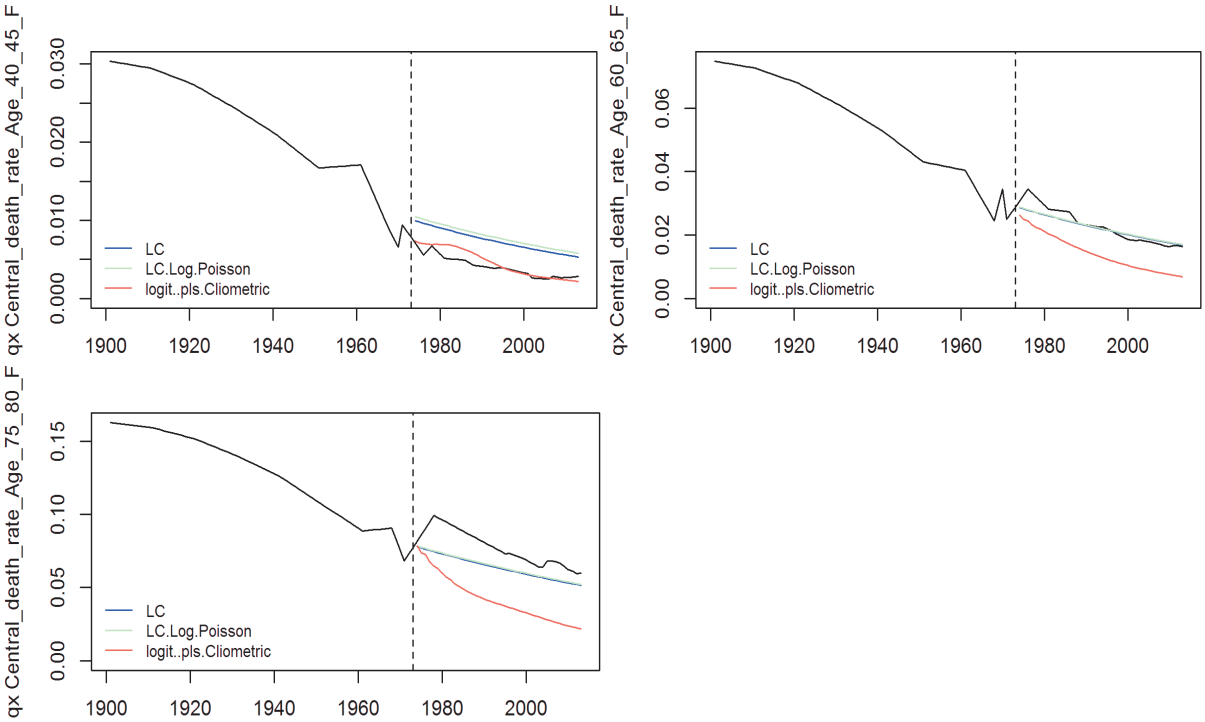


Prévisions comparées pour quelques âges

A titre illustratif, nous fournissons ci-après, les prévisions comparées pour quelques tranches d'âge : 0, [20 ; 25[, [40 ; 45[, [60 ; 65[, [75 ; 80[

Figure 4.35 Prévisions comparées des modèles, pour quelques âges



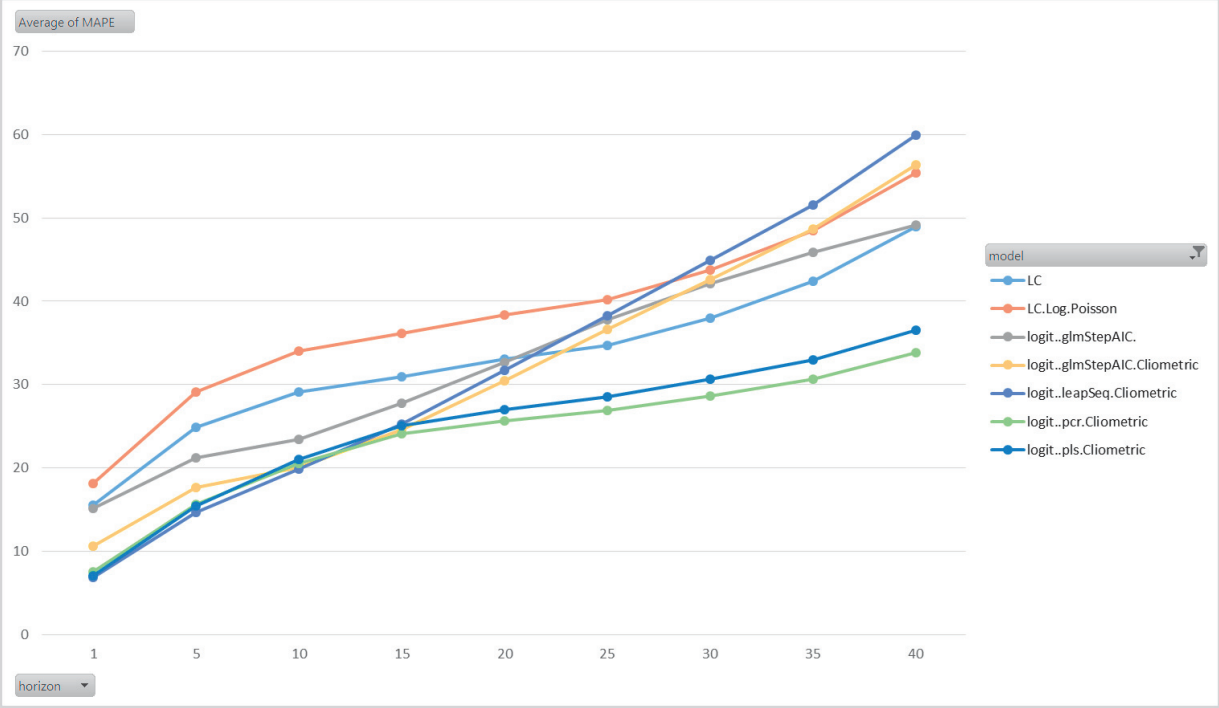


Récapitulatif des performances des modèles

	MAPE (%)								
Modèle	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**
CBD.logit.binomial	112,6	133,2	146,1	152,7	158,5	163,0	169,8	178,8	191,1
LC	15,5	24,9	29,1	31,0	33,0	34,7	38,0	42,4	48,9
LC.Log.Poisson	18,1	29,1	34,0	36,1	38,4	40,2	43,7	48,5	55,4
logit-binomial-PCR-optimal	15,2	21,2	23,4	27,8	32,7	37,8	42,1	45,9	49,1
logit-binomial-PCR-optimal Cliometric	10,6	17,7	20,1	24,6	30,5	36,6	42,6	48,6	56,3
logit-PCR-optimal-Cliometric	6,9	14,7	19,9	25,3	31,7	38,2	44,9	51,6	60,0
logit-PCR-Cliometric	7,0	15,5	21,0	25,1	27,0	28,6	30,7	33,0	36,5
logit-PLS-Cliometric	7,6	15,7	20,5	24,1	25,6	26,9	28,6	30,6	33,8

	Ecart du MAPE : logit PCR Cliometric								
par rapport à :	**1 an**	**5 ans**	**10 ans**	**15 ans**	**20 ans**	**25 ans**	**30 ans**	**35 ans**	**40 ans**
CBD.logit.binomial	-93%	-88%	-86%	-84%	-84%	-84%	-83%	-83%	-82%
LC	-51%	-37%	-29%	-22%	-23%	-23%	-25%	-28%	-31%
LC.Log.Poisson	-58%	-46%	-40%	-33%	-33%	-33%	-35%	-37%	-39%
logit-binomial-PCR-optimal	-50%	-26%	-12%	-13%	-22%	-29%	-32%	-33%	-31%
logit-binomial-PCR-optimal Cliometric	-29%	-11%	2%	-2%	-16%	-27%	-33%	-37%	-40%
logit-PCR-optimal-Cliometric	10%	7%	3%	-5%	-19%	-30%	-36%	-41%	-44%
logit-PLS-Cliometric	8%	1%	-2%	-4%	-5%	-6%	-7%	-7%	-7%

Figure 4.36 Récapitulatif des performances des modèles



Pour ce qui concerne l'Inde (en considérant les âges de manière globale), et quel que soit l'horizon (de 1 à 40 ans) les 2 meilleurs modèles sont le modèle logit-PCR-Cliometric suivi de près par le modèle logit-PLS-Cliometric.

Avec ces deux modèles, les gains de précision obtenus par rapport aux modèles de référence sont assez importants tous âges confondus.

Toutefois, il faut nuancer ces résultats par rapport aux âges : en effet, nous avons pu remarquer sur les graphiques de performances en 3D, que sur les âges les plus élevés les modèles logit-PCR-Cliometric et logit-PLS-Cliometric sont moins performants que les modèles classiques LC et LC-Log-Poisson, contrairement aux autres âges.

14 Conclusion

Dans ce chapitre nous avons pu atteindre les objectifs théoriques et empiriques que nous-nous sommes fixés dans l'introduction, pour les prévisions actuarielles de long-terme des taux de mortalité par âge, dans un contexte de transition démographique par cohortes de pays à l'échelle mondiale.

Le premier objectif était de généraliser le modèle classique de Lee-carter (1992) en y intégrant d'autres composantes (principales ou PLS) de manière différenciée (optimale selon la variance ACP expliquée, ou le MSE cross-validée du modèle) pour chaque âge dont la mortalité est modélisée. Nous appellerons ces extensions, des modèles respectivement, de type PCR-optimal et PLS-optimal.

Sur le plan empirique, l'application empirique du modèle généralisé et optimisé de type PCR-optimal à la France et l'Italie a permis de confirmer leur supériorité de performance sur les modèles classiques testés (Lee-Carter, Lee-Carter Log-Poisson, CBD-logit-binomial). Mais l'horizon de cette supériorité et son ampleur varie selon le pays (jusqu'à l'horizon maximal testé i.e. 50 ans pour la France, et pas au-delà de 25 ans pour l'Italie) l'âge dont la mortalité est modélisée, et l'historique utilisé.
Par conséquent le praticien doit définir les modalités d'usage en fonction des critères ci-dessus énumérés lors des tests des modèles sur son jeu de données.
Par ailleurs, notre modèle généralisé a également montré sa robustesse en présence de valeurs atypiques dans l'historique, ou d'historique tronqué à 20 ans.

Le second objectif était d'utiliser l'information contenue dans l'historique d'un ou de plusieurs pays transitionnellement plus « âgés », pour améliorer les modèles généralisés de l'étape précédente, dans le cadre de la modélisation des mortalités par âge d'un pays transitionnellement plus « jeune » comme les pays émergents ou en développement.
La variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est une variable créée à partir des données historiques du pays A (transitionnellement « plus âgé » que le pays R), pour jouer un rôle de variable exogène dans les modèles de prévisions de mortalité du pays R.
Nous avons proposé deux méthodes originales pour estimer la variable cliométrique ayant les propriétés souhaitées : allure de la mortalité du pays A, mais adaptée à la vitesse (pente) de transition du pays R et défini dans la période de transition et de prévision du pays R.
Nous nommons les modèles utilisant la variable cliométrique, modèles cliométriques : nous en proposons deux types qui sont les modèles PCR-optimal-cliometric et PLS-optimal-cliometric.

Nous avons effectué des tests empiriques des modèles cliometriques à un grand pays émergent (l'Inde) et les résultats ont montré que ces nouveaux modèles ont permis d'améliorer de manière importante les prévisions de long terme.
Pour ce qui concerne l'Inde (en considérant les âges de manière globale), et quel que soit l'horizon (de 1 à 40 ans), les deux meilleurs modèles sont le modèle logit-PCR-Cliometric suivi de près par le modèle logit-PLS-Cliometric.
Avec ces deux modèles, les gains de précision obtenus par rapport aux modèles de référence sont assez importants, soit des réductions du MAPE *(mean absolute percentage error)* d'au moins -31% sur les modèles classiques de référence à un horizon de prévision de 40 ans (tous âges confondus).

Des travaux complémentaires pourront permettre de tester empiriquement une deuxième méthode d'estimation des variables cliométriques que nous avons également proposée, et comparer ainsi les performances avec la méthode qui a été mise en œuvre dans les travaux empiriques présentés.

L'essentiel des contributions techniques est synthétisé dans les sections qui suivent.

La transition démographique moderne est le processus historique par lequel une population passe d'un régime démographique caractérisé par un taux de mortalité global et un taux de natalité élevés, à un nouveau régime caractérisé par des valeurs nettement plus faibles (division par 3, voire 4) de ces 2 taux. Certains pays sont donc plus « *précoces* » dans la transition démographique que d'autres qui sont alors considérés comme étant en situation de « *rattrapage* » relatif.
L'idée est d'utiliser l'histoire quantitative des mortalités et d'autres variables démographiques des pays « transitionnellement plus âgés» pour tenter d'améliorer les prévisions de mortalité dans des pays en « transitionnellement moins âgés».
En pratique, ces améliorations peuvent s'appliquer aux études prospectives de la mortalité dans les pays émergents ou en développement.

Nous avons démonté que le modèle de Lee-Carter [$ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$]
est une régression sur la 1ère composante principale (à facteur multiplicatif près) de la matrice $Z = ln(m_{x,t}) - \hat{a}_x$
Le modèle CBD est basé sur les $\text{logit}\, q_x\,(t)$, aussi bien pour la variable à modéliser, que pour l'estimation des variables explicatives.
En combinant les approches Lee-Carter (et sa relation avec la régression PCR) et CBD, nous pouvons formuler un modèle de régression PCR emboité avec les k premières composantes principales et basée sur les $\text{logit}\, q_x\,(t)$: Nous désignons ce modèle par *Logit-PCR.*

Par ailleurs, nous remarquons empiriquement des décalages temporels, des différences d'allure dans la dynamique des taux de mortalité selon les tranches d'âge.
Si une composante temporelle unique reflète la transition démographique du point de vue de la mortalité globale, la prise en compte de la transition épidémiologique et sanitaire qui affecte les âges de manière différentiée (en débutant par les plus jeunes âges) nécessite des composantes temporelles différentiées par tranche d'âge.
L'objectif étant d'améliorer le modèle Logit-PCR en utilisant une (ou "plusieurs" selon le nombre d'observations i.e. années d'historique disponibles) composante principale optimale, pour chaque tranche d'âge (ou âge) modélisée. Nous appelons ce modèle ***Logit-PCR-optimal***.

Cherchons à présent à améliorer le modèle Logit-PCR en utilisant la régression PLS et en choisissant de manière optimale (MSEP par validation croisée) les composantes principales PLS pour chaque tranche d'âge (ou âge) modélisée. On appellera ce modèle ***logit-PLS***.
La variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$) est une variable crée à partir des données historiques du pays A (transitionnellement « plus âgé » que le pays R), pour jouer un rôle de variable exogène dans les modèle de prévisions de mortalité du pays R.
L'idée est d'utiliser l'histoire quantitative des mortalités (et éventuellement d'autres variables démographiques) du pays A pour tenter d'améliorer les prévisions de mortalité dans le pays R.
Une telle variable peut être utilisée comme variable active explicative supplémentaire dans les modèles prospectifs de mortalité du pays R.
La variable de temps transitionnelle est une variable $\mathrm{m}_{\mathrm{x}_0}$ (taux de mortalité à l'âge x_0) dont la symétrique (par rapport à toute axe parallèle à l'axe horizontal du temps) est un temps transitionnelle mesurable dans les deux pays A ($\mathrm{m}_{\mathrm{A},\,\mathrm{x}_0}$) et R ($\mathrm{m}_{\mathrm{R},\,\mathrm{x}_0}$), et ayant les propriétés suivantes : La monotonie en tendance (nécessaire au rôle de temps transitionnel tel que préalablement défini). Et des valeurs dans un même intervalle dans les deux pays (pour permettre les appariements entre les points temporels des 2 processus liés aux 2 pays).
Cette variable $\mathrm{m}_{\mathrm{x}_0}$, du fait de sa monotonie statistique par rapport au temps, nous permet de faire un appariement entre les points du processus transitionnel du pays A ($\mathrm{m}_{\mathrm{A,x}}(\mathrm{t}_\mathrm{A})$)et celui du pays R ($\mathrm{m}_{\mathrm{R,x}}(\mathrm{t}_\mathrm{R})$) .
Pour créer la variable *cliométrique de mortalité à l'âge* x du pays R par rapport au pays A ($\mathcal{C}_{R,A,x}$), nous proposons deux méthodes distinctes : dans la première méthode, nous cherchons à déterminer deux

fonctions déterministes monotones qui modélisent au mieux les courbes des variables $\mathrm{m}_{\mathrm{A},\mathrm{x}_0}$ et $\mathrm{m}_{\mathrm{R},\mathrm{x}_0}$. Dans la deuxième méthode d'estimation de la variable cliométrique nous utilisons les données sans faire l'hypothèse d'une fonction explicite du taux de mortalité à l'âge x_0.

Nous souhaitons améliorer le modèle Logit-PCR, en y intégrant la variable cliométrique correspondante à chaque tranche d'âge (ou âge) modélisée. On appellera ce modèle **Logit-PCR-cliométrique**. Cette intégration de la variable cliométrique peut se faire **après** le calcul et la sélection des composantes principales. Le risque de colinéarité est gérable par le recours à une méthode d'estimation adaptée (régressions de Ridge, Lasso, elastic-net, sélection *stewise* ou *recherche exhaustive* des variables, PCR, PLS), le cas échéant.

Résultats empiriques

Etude de cas : France

La variable étudiée est la probabilité de décès (hommes et femmes) de tous les âges de 0 à 100 ans.
L'historique d'apprentissage va de 1880 à 1964, et l'horizon de prévision va de 1965 à 2014 (50 ans).
Les valeurs atypiques des deux Guerres Mondiale sont corrigées par imputation avec interpolation linéaire.
Modèles comparés : Lee-carter (LC), LC Log-Poisson, CBD logit-binomial, logit PCR optimal.
Nous constatons un avantage systématique pour le modèle logit PCR optimal, quel que soit l'horizon de prévision étudié.
Le modèle logit PCR optimal conserve son avantage relatif en présence de valeurs atypiques (guerres mondiales) et d'un historique tronqué à 20 ans.

Etude de cas : Italie

La variable étudiée est la probabilité de décès (hommes et femmes) de tous les âges de 0 à 100 ans.
L'historique d'apprentissage va de 1872 à 1964, et l'horizon de prévision va de 1965 à 2014 (50 ans).
Les valeurs atypiques des deux Guerres Mondiale sont corrigées par imputation avec interpolation linéaire.
Modèles comparés : Lee-carter (LC), LC Log-Poisson, CBD logit-binomial, logit PCR optimal (sélection *stepwise*)
Comme pour la France, le modèle CBD se révèle être le moins performant des modèles testés.
Mais, nous constatons donc que, contrairement au cas Français où le modèle logit PCR optimal (sélection stepwise) est le plus performant jusqu'à 50 ans d'horizon de prévision, cet avantage n'excède pas l'horizon prédictif de 25 ans face au modèle LC.
Toutefois, l'avantage systématique de performance modèle logit PCR optimal est conservé vis-à-vis des deux autres modèles testés, i.e. LC.Log.Poisson et CBD.logit.binomial.
Autre différence avec le cas Français, le modèle LC performe systématiquement mieux que LC.Log.Poisson, quel que soit l'horizon prédictif testé (maximum 50 ans).
En présence de valeurs atypiques, le modèle logit-PCR-optimal perd son cet avantage par rapport au LC (après l'horizon prédictif de 15 ans) et le LC Log-Poisson (dès l'horizon prédictif de 25 ans).
Par ailleurs, l'historique d'apprentissage de 20 ans se révèle moins prédictif que celui de 93 ans précédemment testé. Toutefois contrairement aux modélisations basées sur l'historique d'apprentissage de 93 ans, le modèle logit PCR optimal devient à présent (20 ans d'historique) le meilleur quel que soit l'horizon de prévision.

Conclusion sur le modèle non-cliométrique gaussien logit PCR optimal

La supériorité empirique de performance du modèle logit-PCR-optimal par rapport aux modèles de référence LC et LC-Log-Poisson, est toute relative selon le pays, l'horizon de prévision et l'âge (on peut conjecturer sur la période étudiée également). Ainsi, il revient à l'utilisateur de sélectionner les conditions d'utilisation du modèle logit-PCR-optimal, selon les résultats des prévisions tests.

Modèles cliométriques : Application à l'Inde

La variable étudiée est la probabilité de décès (femmes) de 0 à 80 ans par tranches de 5 ans.
Le pays transitionnellement « moins âgé » à modéliser, pays « R » est Inde. L'historique d'apprentissage va de 1901 à 1973. Et l'horizon de prévision va de 1974 à 2013 (40 ans)
Le pays transitionnellement « plus âgé », pays « A» utilisé est Italie.
Nous utilisons comme symétrique de la variable de temps transitionnel, le taux de mortalité à l'âge 0.
Modèles de référence : Lee-Carter, Lee-Carter Log-Poisson, CBD logit binomial
Modèles alternatifs (5) testés : logit-binomial-PCR-optimal, logit-binomial-PCR-optimal Cliometric, logit-PCR-optimal-Cliometric, logit-PCR-Cliometric, logit-PLS-Cliometric.
Pour ce qui concerne l'Inde (en considérant les âges de manière globale), et quel que soit l'horizon (de 1 à 40 ans) les 2 meilleurs modèles sont le modèle logit-PCR-Cliometric suivi de près par le modèle logit-PLS-Cliometric.
Avec ces deux modèles, les gains de précision obtenus par rapport aux modèles de référence sont assez importants, tous âges confondus.
Toutefois, il faut nuancer ces résultats par rapport aux âges : en effet, nous avons pu remarquer sur les que sur les âges les plus élevés les modèles logit-PCR-Cliometric et logit-PLS-Cliometric sont moins performants que les modèles classiques LC et LC-Log-Poisson, contrairement aux autres âges.

Des travaux complémentaires pourront permettre de tester empiriquement la méthode 1 d'estimation des variables cliométriques et comparer ainsi les performances avec la méthode 2 qui a été mise en œuvre dans les travaux empiriques présentés.

15 Bibliographie

Booth H., Maindonald J., Smith L. (2002). Applying Lee-Carter under conditions of variable mortality decline. Population Studies, 56, 325-336.

Bourbonnais R., Terraza M. (2010). Analyse des séries temporelles, Dunod

Bourbonnais R., Usunier J. C. (2017). Prévision des ventes, Economica

Box G. E. P., Tiao G. C. (1975). Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association, Vol. 70, No. 349. (Mar., 1975), pp. 70-79

Cairns A. J.G., Blake D., Dowd K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and Calibration, J. Risk Insur., 73, 4, 687-718

Cairns A. J.G., Blake D., Dowd K. (2008). Modelling and management of mortality risk: a review, Scand. Actuar. J., 2008, 2-3, 79-113, (2008) · Zbl 1224.91048

Chesnais J-C. (1986), « La transition démographique ». Paris : PUF-INED (Cahier n°113)

Courbage Y., Todd E. (2007). « Le rendez-vous des civilisations ». Seuil

Droesbeke J. J., Saporta G. (2010). « Analyse statistique des données longitudinales ». Technip

Gourieroux C., Monfort A. (1999). « Séries temporelles et modèles dynamiques ». Economica

Human Mortality Database (2020), consulté en ligne le 28/09/2020, https://www.mortality.org/cgi-bin/hmd/hmd_download.php

Hyndman R. J. (2019). demography: Forecasting Mortality, Migration, Fertility and Population Data. (R package version 1.21).

Hyndman R. J., Kostenko A. V. (2007). "Minimum Sample Size requirements for Seasonal Forecasting Models," Foresight: The International Journal of Applied Forecasting, International Institute of Forecasters, issue 6, pages 12-15, Spring.

Hyndman R. J., Ullah S., (2007). Robust forecasting of mortality and fertility rates: A functional data approach, Comput. Statist. Data Anal., 51, 10, 4942-4956, (2007) · Zbl 1162.62434

Lebart L., Piron M., Morineau A. (2000). « Statistique exploratoire multidimensionnelle ». Dunod

Lee R.D., Carter L.R. (1992). Modeling and forecasting US mortality. Journal of the American Statistical Association, 87, 659-671.

Maheu J. M., McCurdy T. H. (2009). How Useful Are Historical Data for Forecasting the Long-Run Equity Return Distribution?, Journal of Business & Economic Statistics, Vol. 27, No. 1 (Jan., 2009), pp. 95-112

Max Planck Institute for Demographic Research (2020), Demographic Data, consulté en ligne le 28/09/2020, https://www.demogr.mpg.de/en/research_6120/demographic_data_27/

Millossovich P., Villegas A.M., Kaishev V. K. (2018). StMoMo: An R Package for Stochastic Mortality Modelling. Journal of Statistical Software, 84(3), doi:10.18637/jss.v084.i03

Planchet F., Lelieur V., (2007). Utilisation des méthodes de Lee-Carter et Log-Poisson pour l'ajustement des tables de mortalité dans le cas de petits échantillons, Bulletin français d'actuariat, Vol. 7, n°14, juillet-décembre 2007, pp. 118-146

Saporta G. (2011). « Probabilités, Analyse des données et Statistique ». Technip

Todd E. (1999). « La Diversité du monde : Famille et modernité ». Seuil

Tufféry S. (2012). « Data Mining et Statistique Décisionnelle ». Technip.

Vidal R., Ma Y., Sastry S., (2016), Generalized Principal Component Analysis. New York, Springer-Verlag

Chapitre 5.
Analyse cliométrique de la Mortalité dans l'infanterie pendant la Grande Guerre

1. Table des matières

Chapitre 5. Analyse cliométrique de la Mortalité dans l'infanterie pendant la Grande Guerre 361

1. Table des matières 362

2. Introduction 363

3. Base de données cliométriques de la Grande Guerre 365

4. Géomatique des mouvements des RI 370

4.1. Position de casernement des RI en France métropolitaine, avant le début de la guerre 370
4.2. Projection des RI sur le front 371
4.3. Détection des concentrations géographiques et analyse des batailles associées 372
4.4. Animation quotidienne des concentrations de RI 374
4.5. Vocation de l'application internet d'animation géomatique des RI 375

5. Définition, mesure de la mortalité journalière dans les RI 376

6. Analyse multidimensionnelle ACIDE de la mortalité dans l'infanterie 378

6.1. Mortalité dans les RI selon l'origine géographique 378
6.2. Mortalité dans l'infanterie selon l'orientation politique du département d'origine du régiment .. 379
6.3. Mortalité dans les RI selon le profil économique du département d'origine du régiment 381
6.4. Mortalité dans les RI selon le profil anthropologique du département d'origine du régiment 382
6.5. Mortalité dans les RI selon le profil cognitif et sociétal du département d'origine du régiment 383
6.6. Mortalité dans les RI selon les caractéristiques militaires de l'engagement 384

7. Bilan des facteurs de la mortalité des régiments 388

8. Conclusion 390

9. Bibliographie 391

2. Introduction

Dans le présent chapitre consacré à une analyse cliométrique de la mortalité dans l'infanterie pendant la Grande Guerre, nous appliquons la grille de lecture multidimensionnelle ACIDE (Anthropologie, Cognition et société, Institutions, Démographie, Economie) des lieux d'origine des régiments d'infanterie, à l'analyse explicative de la mortalité des soldats. Grille à laquelle s'ajoute l'analyse selon des facteurs purement militaires de modalités de participation aux combats, ou d'appartenance à l'armée mobile.

Seront successivement proposés : la présentation de la base de données créée, ainsi que le module d'animation géomatique qui reproduit les mouvements des régiments d'infanterie (RI)sur le front entre juillet 1914 et novembre 1918 ; la définition de la mesure de mortalité, ainsi que ses statistiques descriptives ; l'analyse multidimensionnelle ACIDE de la mortalité selon les profils sociodémographiques des départements d'origine des RI ; enfin, l'analyse ACIDE est complétée par l'adjonction de variables explicatives purement militaires : type de régiment (mobile ou non), modalités de participation aux batailles et combats.

Ce chapitre s'inscrit dans la continuité de la littérature existante consacrée aux morts de la Grande Guerre, complète et enrichit celle-ci. Sur le sujet, les articles de références sont ceux de Guironnet et Parent (2018), Gilles, Guironnet, et Parent (2014a), Guillot et Parent (2018).

Les questions abordées dans cette littérature portent, pour les deux premiers, sur la disparité régionale des morts durant la Grande Guerre. Les auteurs s'interrogent sur le fait de savoir si ces disparités relèvent uniquement de différences démographiques (densité de population, taux de mobilisables) entre départements ou si d'autres variables économiques (chômage et catégories socio-professionnelles), politiques (vote antimilitariste, géographiques (proximité aux zones de combats, effet frontière) par département améliorent l'explication des écarts observés. L'article de Guillot et Parent (2018) est d'obédience micro-économétrique et porte, à partir d'un échantillon de la base de données « Mémoire des hommes » sur les différences individuelles de survie entre soldats et propose une explication micro-économétriques de celles-ci, sur la base de données individuelles (âge, corps, rang pour un proxy de classe sociale).

Notre chapitre se distingue à six niveaux de contributions, en ce qu'il analyse la mortalité au niveau du régiment et cherche à étendre la gamme des variables explicatives à des variables issue de la grille multidimensionnelle ACIDE augmentées de la prise en compte de la mobilité et du degré d'exposition des régiments au feu.

De plus, et de façon inédite, sont testées les hypothèses de Todd (1979) concernant l'association entre la mortalité au front et le triptyque dysfonctionnel suicide-alcoolisme-folie du lieu d'origine du RI. En particulier, nous tentons d'apprécier l'étonnante remarque de Todd (1979) selon laquelle la mortalité de la guerre de chaque pays serait « grossièrement proportionnelle à son taux de suicide en temps de paix », hypothèse qui souligne selon ses propres termes, une « continuité frappante dans l'autodestruction » (Todd, 1979). Nous serons ainsi en mesure de tester l'hypothèse d'une association entre la mortalité au front et les « pathologies » sociales du département d'origine des régiments.

Le quatrième niveau d'apport concerne bien évidemment l'utilisation des techniques statistiques adaptées pour gérer la multicolinéarité des variables explicatives, notamment, le recours aux modèles PLS (*Partial Least Square*) et aux régressions sur composantes ACP (Analyse en Composantes Principales) ou AFCM (Analyse Factorielle des Composantes Multiples), pour estimer la part de variance de mortalité expliquée par un groupe de variables spécifiques.

Le cinquième niveau de contribution concerne la création d'une base de données cliométriques de la Guerre 14-18 : elle est composée de 3 grandes catégories de données. D'abord, les données sur les positions quotidiennes des régiments d'infanterie de 1914 à 1918. Ensuite, les données descriptives de régiments d'infanterie : rattachement administratif, effectifs, origine géographique, date de mobilisation, date de dissolution, projection à l'étranger durant la guerre. Et enfin, les données propres à l'origine géographique des régiments : choix politiques aux élections législatives (de 1877 à 1914), données anthropologiques, cognitives et sociales, politiques, démographiques et économiques.
Cette base de données cliométriques de la Grande Guerre a été nécessaire à la réalisation de toutes les analyses précédemment indiquées.

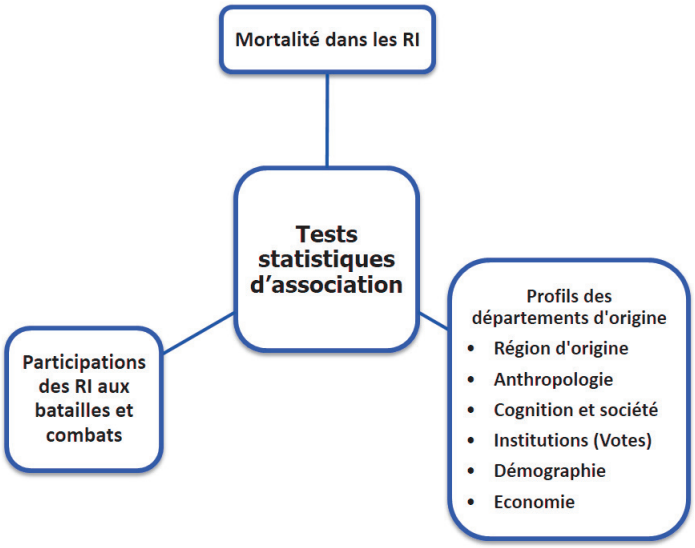
Enfin, grâce aux géolocalisations quotidiennes des régiments, un logiciel internet d'animation géomatique a été développé. Cette application internet reproduit les mouvements des régiments d'infanterie sur le front, du 27 juillet 1914 (le 28 juillet étant la date de la 1ère déclaration de la guerre, celle de l'Autriche-Hongrie à la Serbie), au 11 novembre 1918 (date de l'armistice). Elle permet de repérer les concentrations géographiques de troupes d'infanterie, d'y détecter, le cas échéant, les combats importants s'y déroulant, ainsi que les modalités d'engagement et la liste des régiments impliqués.
Ce logiciel internet à vocation, à terme, à être mis à disposition de la communauté des chercheurs de l'histoire de la Grande Guerre.

Ces six nouveautés garantissent l'originalité de la contribution proposée dans ce chapitre. A noter, ce chapitre a fait l'objet de plusieurs communications dans des colloques, notamment à l'Institut des Hautes Etudes de Défense Nationale (IHEDN) à l'occasion du centenaire de l'Armistice en 2018 à Paris, et aux Journées de l'Economie (Jéco) 2018 à Lyon, et sera soumis prochainement dans une revue internationale de cliométrie.

3. Base de données cliométriques de la Grande Guerre

L'objectif de ce chapitre est double : Engager l'analyse multidimensionnelle ACIDE (Anthropologie, Cognition et société, Institutions, Démographie, Economie) de la mortalité selon les profils des départements de casernement (i.e. d'origine) des régiments d'infanterie de l'Armée française. Cette première étape est complétée par l'analyse militaire de la mortalité selon les modalités de participation (ou non) aux batailles et combats, et d'appartenance à l'armée mobile.
L'objectif analytique de ce chapitre peut être résumé par le schéma suivant qui indique les associations à modéliser et tester statistiquement :

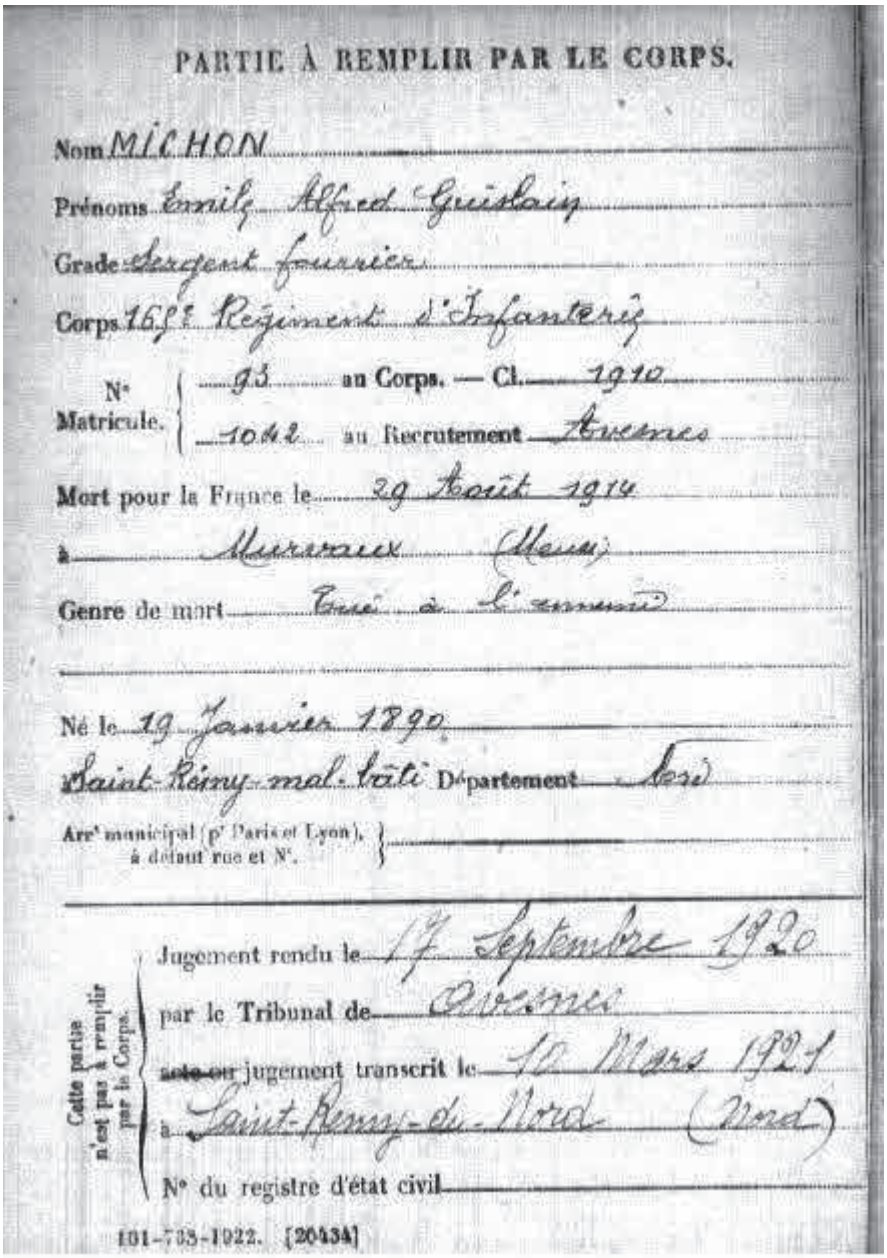
Figure 5.1 Données et objectif analytique du chapitre



Ainsi, plusieurs sources d'informations ont été exploitées pour constituer notre base de données cliométriques : selon le format de l'information, les traitements spécifiques ont été effectués, dont les saisies manuelles, les collectes automatisées, les travaux de restructuration et de normalisation d'informations textuelles.

La base de données *Mémoire des hommes* (Ministère des Armées, 2018) qui recense nominativement les *Morts pour la France* lors de la Grande Guerre, a été exploitée pour les besoins de la thèse.
Cette base a été constituée par la numérisation et la saisie des fiches élaborées par l'administration des anciens combattants. La mention « Mort pour la France » est accordée, suivant certaines conditions, en vertu du code des pensions militaires d'invalidité et des victimes de la guerre.
Plus de 1,3 million de militaires décédés au cours de la Grande Guerre et ayant obtenu la mention "Mort pour la France" figurent dans cette base de données des « Morts pour la France de la Première Guerre mondiale ».
Nous avons utilisé la base de données *Mémoire des hommes*, pour son caractère quasi-exhaustif de recensement des *Morts pour la France*, ce qui a permis un calcul précis des taux de mortalité par régiment, par jour ou période spécifique.

Figure 5.2 Exemple d'une fiche d'un soldat « Mort pour la France »



PARTIE À REMPLIR PAR LE CORPS.

Nom MICHON

Prénoms Emile Alfred Gustave

Grade Sergent fourrier

Corps 165e Régiment d'Infanterie

N° Matricule. 93 au Corps. — Cl. 1910
1042 au Recrutement Avesnes

Mort pour la France le 29 Août 1914

à Maissemy (Aisne)

Genre de mort Tué à l'ennemi

Né le 19 Janvier 1890

Saint-Remy-mal-bâti Département Nord

Arr^t municipal (p^r Paris et Lyon), à défaut rue et N°.

Cette partie n'est pas à remplir par le Corps.

Jugement rendu le 17 Septembre 1920

par le Tribunal de Avesnes

~~acte ou~~ jugement transcrit le 10 Mars 1921

à Saint-Remy-du-Nord (Nord)

N° du registre d'état civil

101-733-1922. [20434]

Source : Archives départementales du Nord, 2020

Le site internet *Mémoire des hommes* (Ministère des Armées, 2018) recense également un certain nombre de documents qui ont été des sources précieuses d'information sur le casernement d'origine, la composition initiale, le corps d'appartenance (brigade, division, corps), les évolutions (dissolutions, fusions, constitutions *ad'hoc*) et le parcours des régiments, ainsi que les évènements auxquels ces régiments ont pris part durant la Grande Guerre. Il s'agit notamment des historiques régimentaires et des Journaux des Marches et Opérations (JMO).

Les données de l'association de généalogistes FranceGenWeb (2018), membre de la Fédération Française de Généalogie (association agréée d'utilité publique, regroupant plus de 150 associations) nous ont permis de collecter les dates, les lieux, le régiment, et surtout la cause de décès de 612 052 soldats (sur plus de 1 300 000 estimés par le Ministère des Armées [2018]) de l'armée française durant la Première Guerre mondiale.

Les sources utilisées pour la construction de cette base de données sont diverses, dont notamment les relevés des monuments aux Morts, les historiques régimentaires.
Les lieux de décès et leurs causes permettent de marquer les lieux de passage des régiments et leur participation aux batailles et combats. Notamment la mention « tué à l'ennemi » est une forte présomption de présence du régiment sur le lieu du décès, contrairement à un décès dans un hôpital où le soldat a pu être évacué ou dans une ambulance lors de l'évacuation.

Ont également fait l'objet de saisies et normalisation, les données historiques collectées par Todd (1979, Todd et Le Bras (2012, 2013) sur l'alphabétisation, le suicide, les violences (homicides, coups et blessures), l'ivresse publique, les aliénés, les pratiques religieuses, les systèmes anthropologiques dans les départements français.

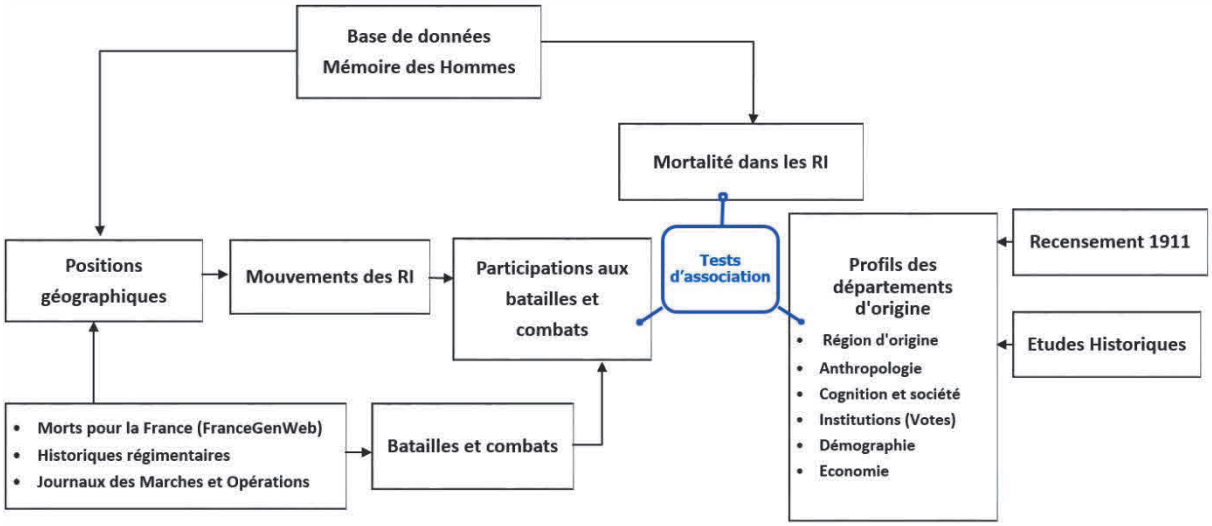
Guironnet et Parent (2014a, 2014b, 2018) nous ont fourni les données démographiques (population, population mobilisable) et socio-professionnelles issues du recensement de 1911, ainsi que les nombres de mobilisés et de morts par département.
Ces mêmes auteurs nous ont également mis à disposition des cartographies des résultats départementaux des élections législatives françaises de 1877 à 1914 : il s'agit notamment des pourcentages de votes pour un ensemble de partis catégorisés comme « extrême gauche » (Parti socialiste, Parti ouvrier, et Radicaux socialistes).
Nous avons manuellement transcrits ces résultats cartographiques dans la base de données cliométriques.

La cartographie détaillée des votes aux élections législatives de 1914 par circonscriptions électorales et départements (Salmon, 2001), a été manuellement retranscrite et codifiée dans la base de données.

La combinaison de ces multiples sources d'informations et les traitements des données intégrées, ont abouti à la création de notre base de données cliométriques, qui permet d'effectuer des tests d'association statistique entre les trois types de données suivantes : les taux de mortalité dans les régiments d'infanterie métropolitains de l'armée française durant la Grande Guerre, les participations aux batailles et combats des régiments, et les profils des départements d'origine des régiments (caractéristiques anthropologiques, cognitives et socio-professionnelles, politiques, démographiques et économiques)

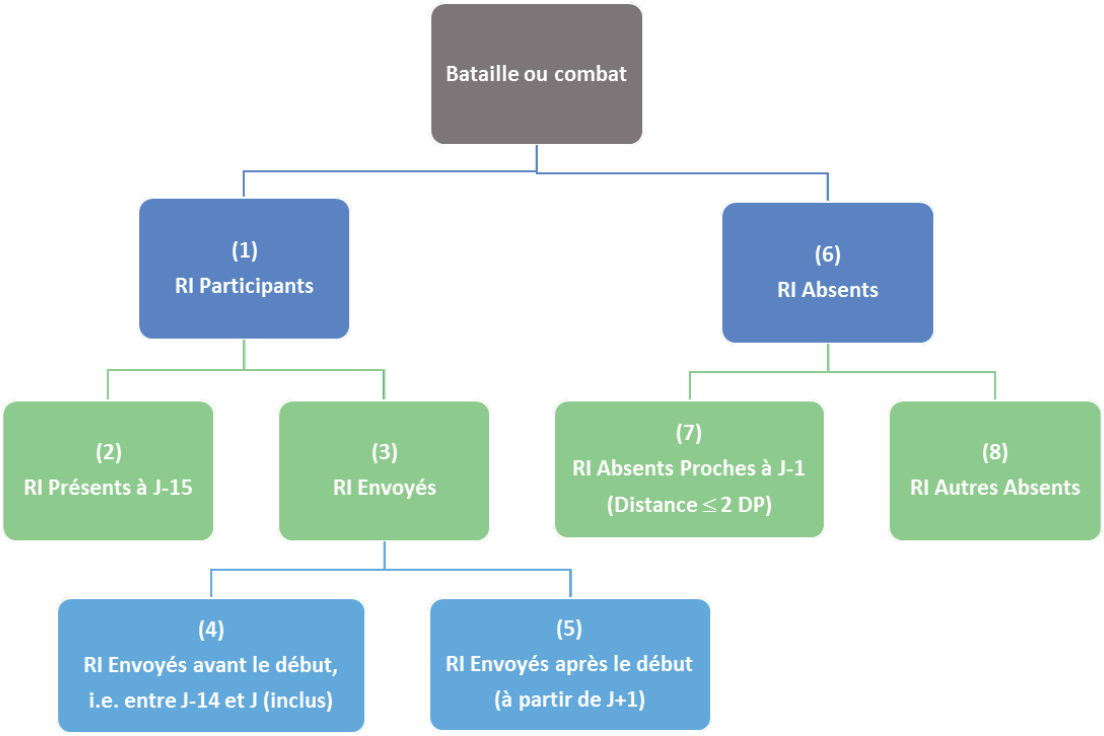
La base de données cliométriques a donc été constituée selon l'architecture ci-après :

Figure 5.3 Architecture de la base de données cliométriques, et son module d'analyse statistique



Les modalités de « participations des régiments aux batailles et combats » sont codifiées dans la base de données d'après la nomenclature suivante :

Figure 5.4 Les modalités de participations (ou non) des régiments aux batailles et combats



Une bataille est considérée comme composée de plusieurs combats dispersés dans le temps et l'espace. Les batailles durent par conséquent plus longtemps et impliquent un plus grand nombre de régiments, qu'un combat qui peut être une simple offensive ponctuelle : ces batailles et combats sont évoqués dans les historiques régimentaires et les Journaux de Marches et Opérations (JMO).

La date de début d'une bataille ou combat est désignée par « J » dans la figure ci-dessus.

Participation d'un régiment d'infanterie (RI) à une bataille : Un RI positionné sur les lieux (départements) et la période d'une bataille est considéré comme participant à ladite bataille.

Notion de distance départementale : l'unité géographique utilisée étant le département, la distance utilisée préférentiellement est le nombre de départements (DP). Cette distance est égale à 1, entre 2 départements contigus. Une base de données des distances entre couples de départements a été constituée afin de calculer les distances entre la position de chaque régiment et les lieux de chaque combat ou bataille, pour chaque jour.
Nous avons également introduit les distances entre chefs-lieux de départements (en Km), et les distance entre barycentres de départements (en Km) : les analyses sont donc possibles selon ces deux distances également.

Distance entre RI et lieux de batailles : chaque bataille se déroule dans 1 ou plusieurs départements. Par conséquent, la distance entre un RI et une bataille est la distance départementale entre le RI et le département de bataille le plus proche de l'emplacement du RI.

Il est ainsi possible de calculer la distance parcourue pour les RI envoyés pour participer à une bataille.
Il est également possible de calculer la distance où se situent les RI ne participant pas à la bataille :
Et cela avant et pendant la bataille.
De même il est possible de déterminer si les RI ne participant pas à une bataille sont éventuellement impliqués au même moment dans une autre bataille : cela permettra de les écarter dans certaines analyses.
RI présents sur les lieux avant le début d'une bataille : Il est possible de déterminer la présence sans interruption, durant N (paramètre) jours avant la date de début d'une bataille.

Informations complémentaires sur les régiments d'infanterie lors de la Grande Guerre
Durant la Grande Guerre, nous avons recensé 370 régiments d'infanterie cantonnés en France métropolitaine. 50 de ces RI sont cantonnés sur plusieurs départements.
Nous avons compté 176 régiments d'infanterie *d'active*, 173 RI créés à partir de réserves et 21 RI créés par ordonnance au cours de la guerre.
En 1914, un régiment d'infanterie typique est composé de 2.000 à 4.000 soldats répartis dans 2 à 4 bataillons (d'environ 1.000 soldats chacun).
Chaque bataillon est composé de 4 compagnies. Une compagnie est composée de 2 pelotons, qui à leur tour sont subdivisés en 2 sections. Chaque section est composée de 2 demi-sections, elles-mêmes subdivisées en 2 escouades. Une escouade typique est composée de 15 soldats.

Synthèse de la base de données cliométriques de la Guerre 14-18
La base de données cliométriques de la Guerre 14-18, est composée de trois grandes catégories de données : D'abord, les données sur les positions quotidiennes des régiments d'infanterie de 1914 à 1918 : Lieu, Département, Date de début de position, Date de fin de position, Type de position (casernement, combats, autres).
Ensuite, les données descriptives de régiments d'infanterie : division, corps d'armée, brigade, appartenance à l'armée mobile, nombre initial de bataillons, lieu de casernement, date de mobilisation, date de dissolution (éventuellement), projection à l'étranger durant la guerre (en Belgique, Italie, Grèce par exemple), et enfin, les données descriptives des départements de casernement des régiments : Votes aux élections législatives de 1877 à 1914, données démographiques et économiques du recensement de 1911, données départementales anthropologiques, cognitives et sociales, politiques, démographiques et économiques

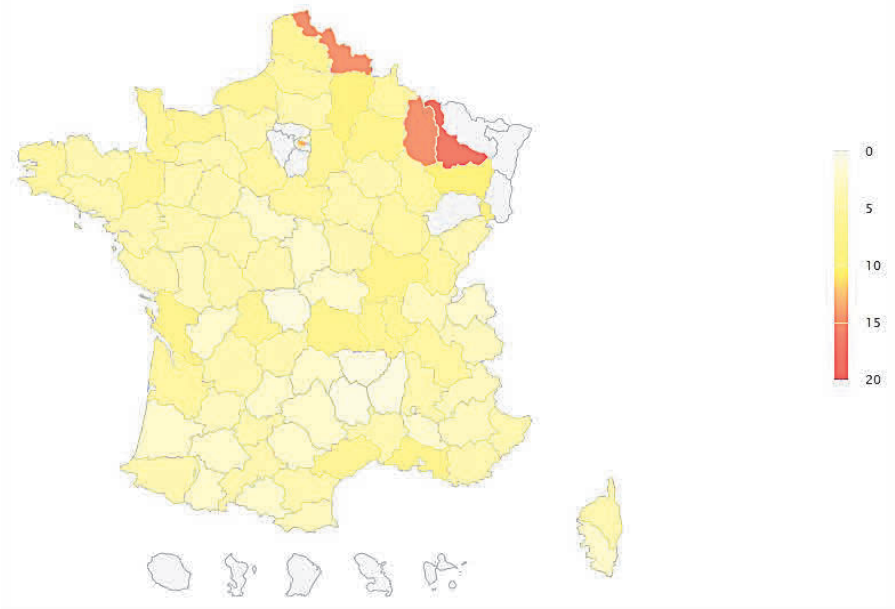
4. Géomatique des mouvements des RI

La base de données cliométriques créé permet de positionner quotidiennement tous les régiments d'infanterie actifs en France métropolitaine.
Cette géolocalisation nous a donc permis de développer une application internet d'animation géomatique qui reproduit les mouvements et concentrations des régiments d'infanterie sur le front entre juillet 1914 et novembre 1918

4.1. Position de casernement des RI en France métropolitaine, avant le début de la guerre

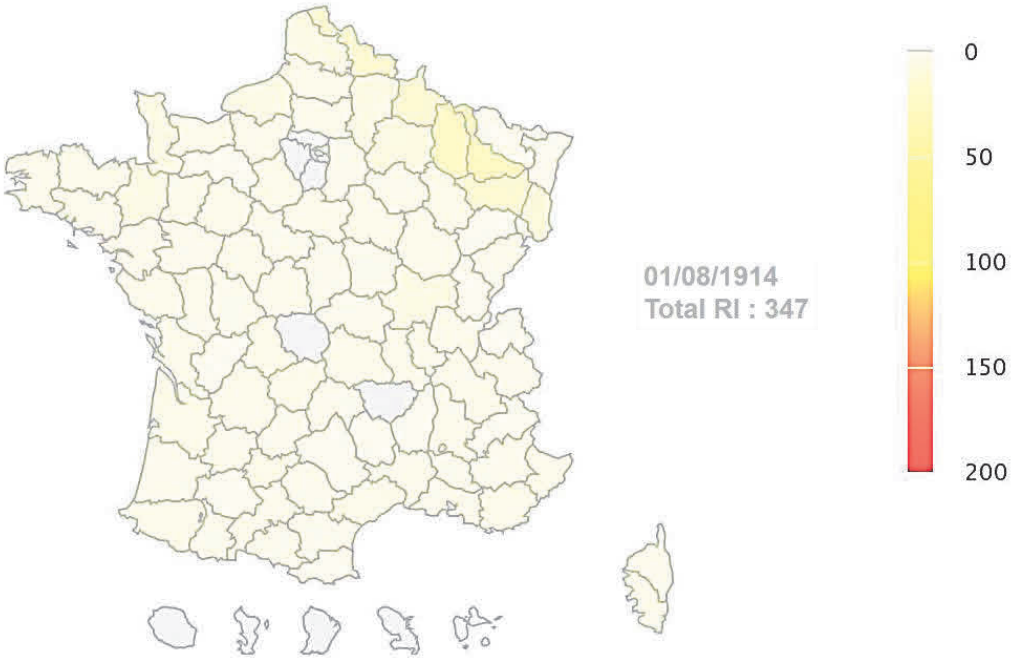
Cette première carte permet de visionner le positionnement des RI français début août 1914, avant la déclaration de la guerre. Nous y voyons les RI cantonnés, dans l'ensemble, dans leurs départements d'origine, avec toutefois de légères renforcements de la ligne Maginot aux frontières belge, Luxembourgeoise et surtout allemande.

Figure 5.5 Nombre de régiments d'infanterie en casernement au début août 1914



En nous mettant sur l'échelle des grandes concentrations géographiques de régiments qui vont avoir lieu durant la guerre (avec parfois plus de cent RI par département), nous constatons ci-après que ces effectifs de RI par département sont faibles et diffuses, en cohérence avec une situation de paix *antebellum :*

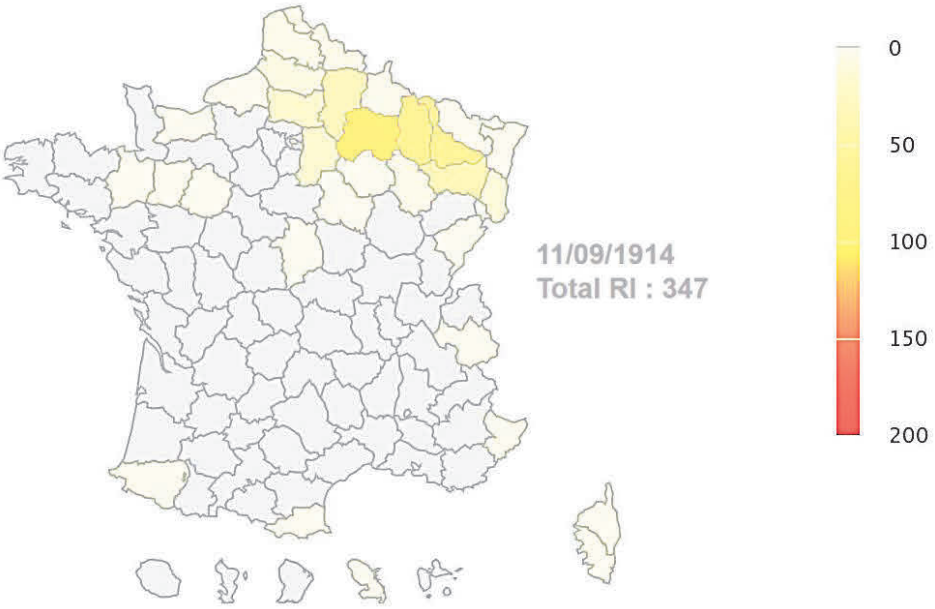
Figure 5.6Nombre de RI par département au 1er aout 1914, avant le début de la guerre



4.2. Projection des RI sur le front

Le 3 août 1914 l'Allemagne déclare la guerre à la France et le 11 août 1914 la France déclare la guerre à l'Autriche-Hongrie. Un mois plus tard (11 septembre 1914), nous pouvons observer que l'essentiel des RI sont remontés vers le nord-est aux frontières belge, Luxembourgeoise et allemande.

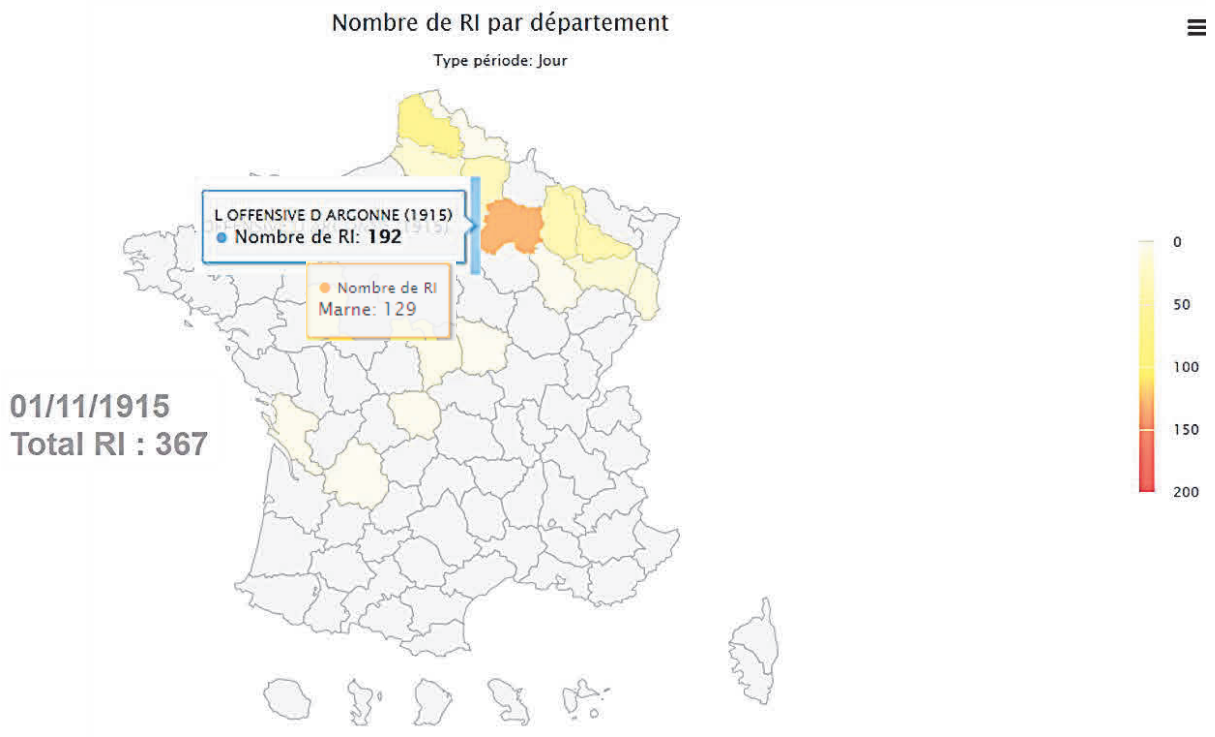
Figure 5.7 Projection des RI sur le front, un mois après le début de la guerre. Nombre de RI par département.



4.3. Détection des concentrations géographiques et analyse des batailles associées

Avec le différentiel de coloration des départements, il est possible de repérer les concentrations géographiques de troupes. L'application affiche alors les batailles et combats se déroulant dans le département sélectionné.
Nous pouvons ainsi observer au 1er novembre 1915, que l'offensive d'Argonne (qui a mobilisé en tout 192 RI dans la Marne) est en cours dans le département de la Marne, avec 129 RI qui y sont mobilisés en ce jour spécifique.
Nous constatons également que le nombre total de RI mobilisés sur le territoire est de 367 en hausse par rapport aux 347 présents au début de la guerre.

Figure 5.8 Analyse de la concentration de troupes d'infanterie, dans la Marne, au 1er novembre 1915



Il est également possible d'extraire la liste exhaustive des RI présents dans la Marne le 1er novembre 1915, avec les principales caractéristiques : appartenance (brigade, division, corps d'armée), lieux de casernement d'origine, nombre initiale de bataillons et effectif estimé initial.

Figure 5.9 Exemple de liste de RI ayant participé à un combat ou bataille (ici l'offensive d'Argonne 1915)

Département : MARNE (51) Période : 01/11/1915 Nombre de RI : 131

Liste des RI | Position de casernement | Casernement comparé | Grandes Batailles | Batailles & Combats

Afficher 10 enregistrements Search:

Export to xlsx | Export to csv

Dénomination	Corps d'armée	Division	Brigade	Casernement (Département)	Nombre intial de bataillons	Effectif initial estimé
99 ème RI	14 ème Corps d'armée	28 ème Division	55 ème Brigade	ISERE (38)	3	3000
99 ème RI	14 ème Corps d'armée	28 ème Division	55 ème Brigade	RHONE (69)	3	3000
96 ème RI	16 ème Corps d'armée	31 ème Division	61 ème Brigade	HERAULT (34)	3	3000
94 ème RI	6 ème Corps d'armée	42 ème Division	83 ème Brigade	MEUSE (55)	3	3000
93 ème RI	11 ème Corps d'armée	21 ème Division	42 ème Brigade	VENDEE (85)	3	3000
87 ème RI	2 ème Corps d'armée	3 ème Division	6 ème Brigade	AISNE (02)	3	3000

Une étude plus approfondie de la bataille identifiée (ici l'offensive d'Argonne) est également possible. En effet, nous sommes en mesure de caractériser la bataille par ses dates estimées de début et de fin, le nombre de RI impliqués, la liste des départements où ont lieu les combats, la liste des RI présents avant le début des hostilités, les RI envoyés sur site en cours de bataille, les RI absents mais proches (avec une option permettant de lister les RI impliqués ou pas dans d'autres combats).

Figure 5.10 Informations des RI, concernant l'offensive d'Argonne de 1915

1ère Guerre Mondiale : Grandes Batailles

Nom: **L OFFENSIVE D ARGONNE (1915)**

Du: **01-03-1915** Au: **30-11-1915** Nombre de RI impliqués: **236**
Département(s) : **3 : MARNE(51) ARDENNES(08) MEUSE(55)**

Régiments d'Infanterie Présents : 236 | Régiments d'Infanterie Absents : 134 | TABLEAU RECAPITULATIF

Régiment(s) présent(s) dans l'un des départements 15 jour(s) avant le début de la grande bataille

RI Présents **15** jour(s) avant le début : **191** | RI Envoyés entre le jour **15** avant le début et le début : **6** | RI Envoyés après le début : **39**

Pour chacun des régiments identifiés dans les cas ci-dessus, l'application fournit les informations suivantes :
Numéro du régiment d'infanterie, Armée mobile (oui ou non), nombre de bataillons initial, effectif initial estimé, nombre de Morts pour la France, taux de mortalité calculé en nombre de soldats par jour, taux pondéré (par la dangerosité des jours) de mortalité, le lieu de casernement initial (département), le pourcentage des votes pour différentes orientations politiques lors des législatives de 1877 à 1914.

Figure 5.11 Exemple de tableau extractible de données sur les RI listés selon le mode de participation à un combat ou bataille

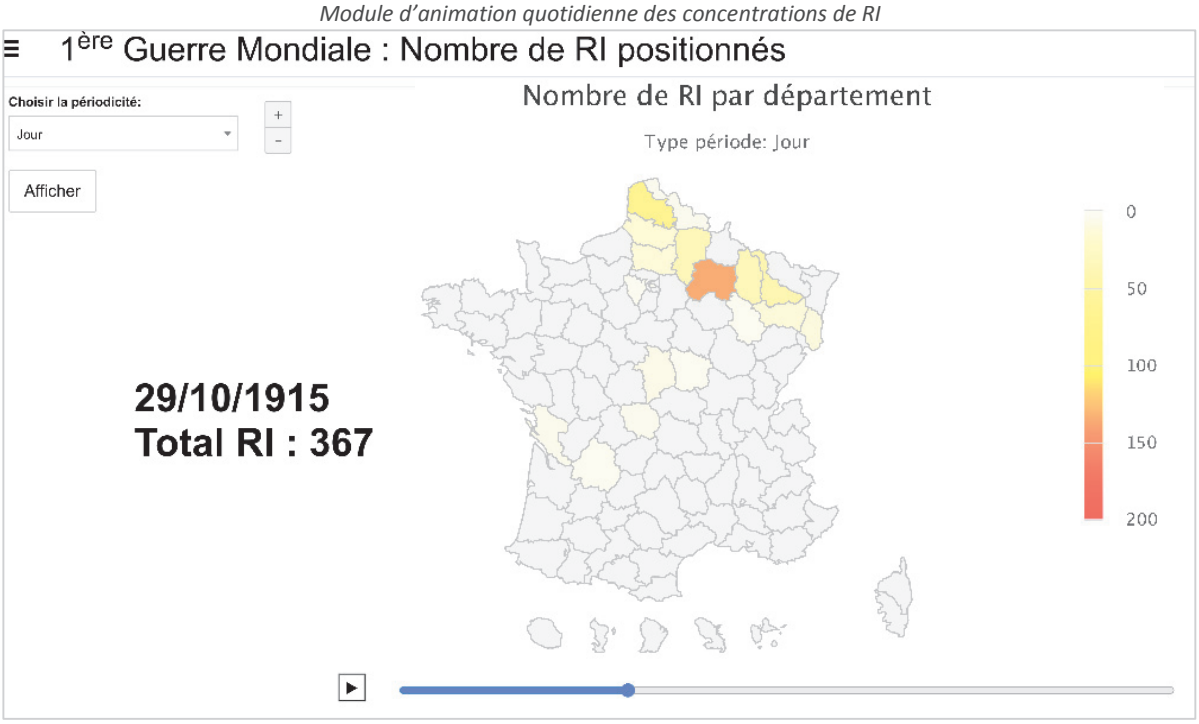
Export to xlsx | Export to csv

Régiment d'infanterie	Armée Mobile	Nombre de bataillons	Effectif initial estimé	Morts pour la France	Morts calculés en nombre de soldats par jour	Taux pondéré de morts	Casernement (Département)	Législatives d'Octobre 1877 Vote Extrême-Gauche	Législatives d'Octobre 1885 Vote Extrême-Gauche	Législatives d'Octobre 1889 Vote Extrême-Gauche	Législatives d'Août-Septembre 1893 Vote Extrême-Gauche
366 ème RI	NON	2	2000	1603	0.04	0.80	MEUSE (55)	1 - Aucun élu d'Extrême-Gauche	1 - Aucun élu d'Extrême-Gauche	2 - Un seul élu d'Extrême-Gauche	1 - Néant
365 ème RI	NON	2	2000	1832	0.05	0.92	NORD (59)	2 - Un seul élu d'Extrême-Gauche	1 - Aucun élu d'Extrême-Gauche	3 - Elus d'Extrême-Gauche en minorité	4 - De 10% à 15% des électeurs inscrits
364 ème RI	NON	2	2000	447	0.59	0.22	MEUSE (55)	1 - Aucun élu d'Extrême-Gauche	1 - Aucun élu d'Extrême-Gauche	2 - Un seul élu d'Extrême-Gauche	1 - Néant
362 ème RI	NON	2	2000	880	0.96	0.44	NORD (59)	2 - Un seul élu d'Extrême-Gauche	1 - Aucun élu d'Extrême-Gauche	3 - Elus d'Extrême-Gauche en minorité	4 - De 10% à 15% des électeurs inscrits
361 ème RI	NON	2	2000	954	1.36	0.48	MARNE (51)	1 - Aucun élu d'Extrême-Gauche	1 - Aucun élu d'Extrême-Gauche	3 - Elus d'Extrême-Gauche en minorité	4 - De 10% à 15% des électeurs inscrits

4.4. Animation quotidienne des concentrations de RI

Du 27 juillet 1914 (le 28 juillet étant la date de la 1ère déclaration de la guerre, celle de l'Autriche-Hongrie à la Serbie), au 11 novembre 1918 (date de l'armistice), il est possible de visualiser sur l'écran ci-après, l'animation quotidienne des concentrations de RI :

Module d'animation quotidienne des concentrations de RI



4.5. Vocation de l'application internet d'animation géomatique des RI

L'application internet d'animation géomatique des RI a vocation, à terme, à être mise à disposition de la communauté des chercheurs de l'histoire de la Grande Guerre.
Cette application web permettra l'exploration et l'analyse des données des régiments d'infanterie françaises avant et pendant la Guerre 14-18, et sera composé des modules suivants : module de cartographie dynamique, permettant de suivre les mouvements au jour le jour des régiments d'infanterie de 1914 à 1918 ; modules d'analyse statistiques de la mortalité basés sur le *framework* Shiny, permettant d'utiliser des fonctions du logiciel R pour l'analyse des données en ligne.

5. Définition, mesure de la mortalité journalière dans les RI

Comme dans la littérature précitée sur la Grande Guerre, la mortalité a été étudiée au niveau des départements et régions (est calculé dans ce cas le nombre total de morts par zone, le nombre de morts par habitant de la zone, ou encore le nombre de morts par mobilisable de la zone), nous proposons un nouvel indicateur qui est le taux de mortalité par régiment.

La question centrale que nous traitons ici est celle de la mortalité, une fois que les régiments sont sur le front. Comme il nous faut un indicateur comparable entre régiments, il a donc fallu pour ce faire neutraliser les facteurs suivants : (i) Effectifs initiaux des régiments ; (ii) Durées de vie des régiments ; (iii) Mortalité (i.e. dangerosité ou létalité) des périodes d'existence des régiments.

En effet, dans les faits : tous les régiments n'ont pas nécessairement le même nombre de soldats ; tous les régiments ne sont pas constitués (ou éventuellement dissous) au même moment de la guerre, autrement dit tous les régiments n'ont pas la même durée de vie ; tous les régiments n'ont pas non plus le même degré d'exposition aux périodes les plus meurtrières, et de ce fait, tous les régiments ne sont pas exposés au même risque de mortalité. Comme les régiments n'ont pas la même durée d'existence, il faut calculer une mortalité quotidienne. Comme tous les jours n'ont pas la même dangerosité, il faut pondérer le nombre de morts par jour par la mortalité de la période d'existence du régiment. Cette pondération consiste à diviser le nombre de morts par jour par un indice de dangerosité de la période d'existence du régiment. Puisque les régiments n'ont pas le même effectif initial, il faut encore diviser la valeur obtenue par l'effectif initial du régiment afin de rendre tous les régiments comparables.

Au final nous mesurons **le nombre de MPF (Morts Pour la France) par jour d'existence du RI, pondéré par la mortalité quotidienne de la période d'existence du régiment pour 1000 soldats initiaux** (comme calculé ci-dessous).

$$\mathrm{m_i} = \frac{\mathrm{MPF_i}}{(\sum_{\mathrm{j}\epsilon[\mathrm{C_i};\mathrm{D_i}]} \mathrm{pct_MPF_j}) \times \mathrm{Duree_Guerre}} \times \frac{1000}{\mathrm{Effectif_initial_i}}$$

Avec :

$\mathrm{MPF_i} = \text{Morts pour la France du RI i}$

$\mathrm{C_i} = \text{Date de 1er casernement du RI i (si RI créé en cours de guerre) ou Date de début de guerre}$

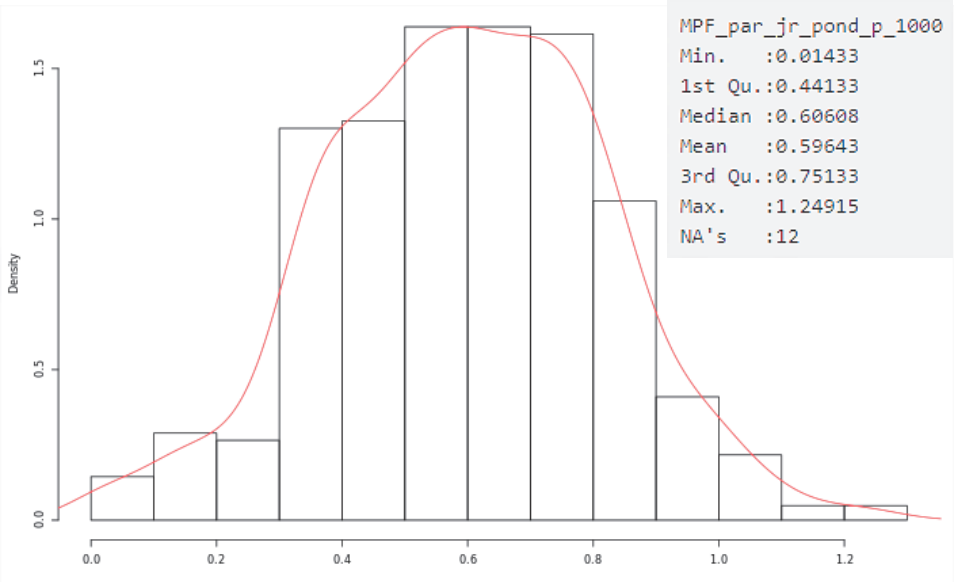
$\mathrm{D_i} = \text{Date de dissolution du RI i (si RI dissout au cours de la guerre) ou Date de fin de guerre}$

$\mathrm{Duree_Guerre} = \text{Date de début guerre (28 août 1914) - Date de fin guerre (11 nov. 1918) } +1 = 1\,537 \text{ jours}$

$$\mathrm{pct_MPF_j} = \frac{\text{Morts pour la France au jour j}}{\text{Morts pour la France durant la guerre}}$$

$\mathrm{Effectif_initial_i} = \text{Effectif initial du RI i}$

Figure 5.12 Nombre de MPF (Morts Pour la France) par jour d'existence du RI, pondéré par la mortalité quotidienne de la période d'existence du régiment pour 1000 soldats initiaux



Les statistiques descriptives sur l'ensemble des régiments durant les 1537 jours observés indiquent un taux moyen de mortalité quotidienne pour 1000 soldats initiaux de 0,6, le minimum étant de 0,014 et le maximum de 1,25. La distribution paraît équilibrée des deux côtés de la moyenne, au vu de l'histogramme.

La stratégie de test d'association statistique utilisée dans ce chapitre est la suivante :

Tableau 5-1 Stratégie de test d'association statistique

	Input	
Target	***Numérique***	***Texte***
Numérique	Test de corrélation de Pearson	ANOVA
Texte	ANOVA	Test exact de Fisher + V-Cramer

Target désigne la variable expliquée ; Input désigne la variable explicative)

Chaque association est systématiquement testée linéairement (Test de corrélation de Pearson) et non linéairement (Test exact de Fisher et V de Cramer, après discrétisation des variables numériques), ce qui signifie que lorsque nous concluons à l'absence d'association, cela vaut aussi bien linéairement que non linéairement. Ceci est fondamental pour l'analyse historique.
Les tests d'association entre la mortalité et une modalité d'une variable qualitative est effectuée en dichotomisant la mortalité deux classes (mortalité supérieure au 3e quartile [donc très forte], et mortalité inférieure au 3e quartile), puis en réalisant le test exact de Fisher ; la force et le sens de l'association est mesurée par le V de Cramer associé, qui varie alors entre -1 et 1 comme un coefficient de corrélation de Pearson.

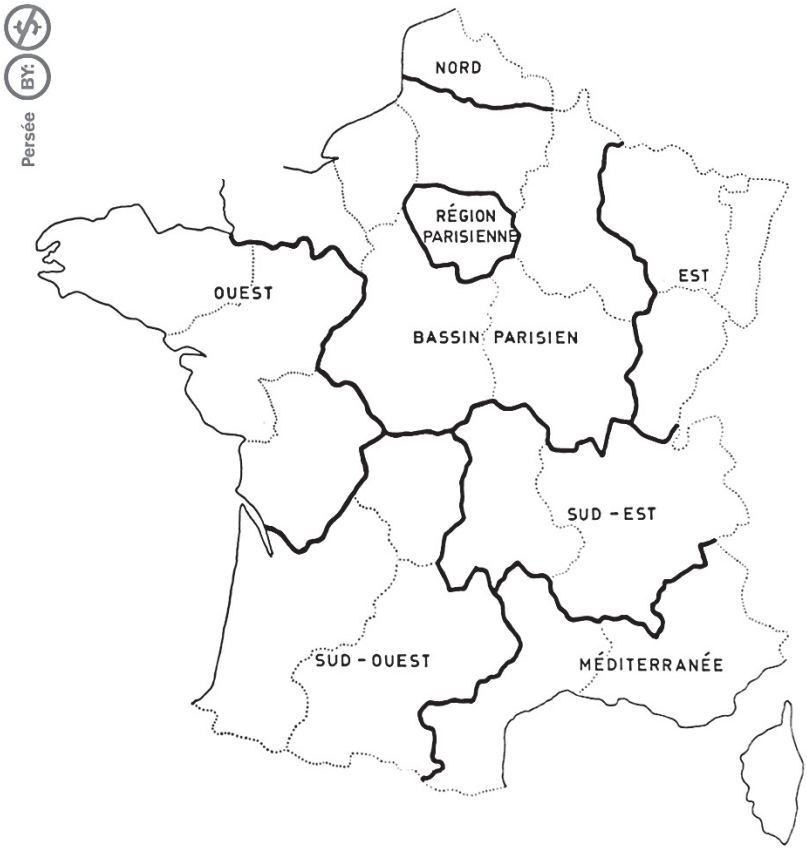
De plus, afin de calculer la part de variance de mortalité expliquée simultanément par un groupe de variables explicatives, nous utilisons un modèle PLS ou une régression sur composantes ACP (Analyse en Composantes Principales) ou AFCM (Analyse Factorielle des Composantes Multiples), afin de dépasser l'obstacle de colinéarité des variables explicatives utilisées ensemble.

6. Analyse multidimensionnelle ACIDE de la mortalité dans l'infanterie

6.1. Mortalité dans les RI selon l'origine géographique

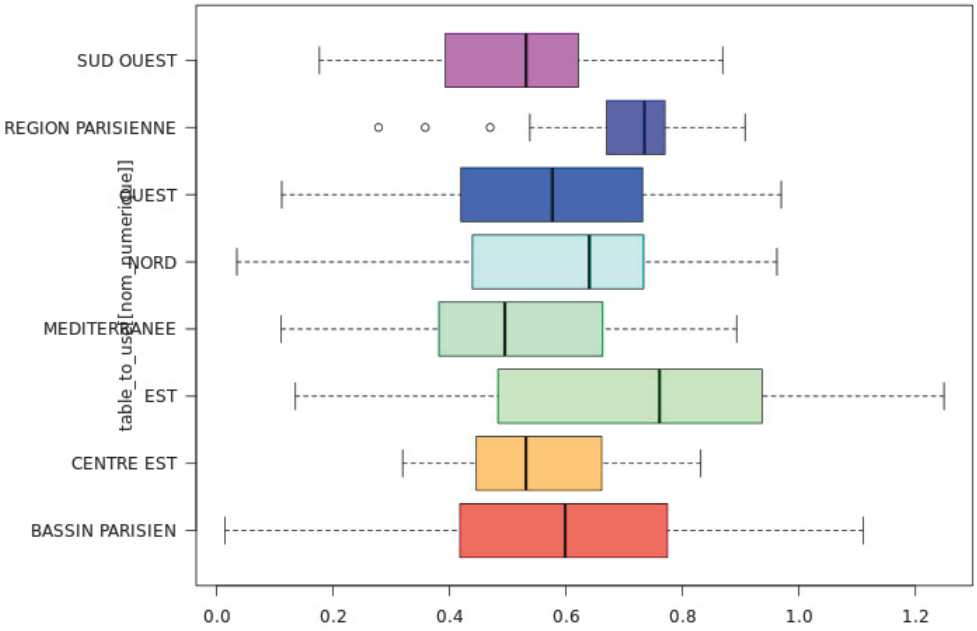
L'origine géographique d'un RI désigne son département de cantonnement initiale.
Le découpage retenu (car ayant peu de modalités) reprend les 8 Zones d'Etude et d'Aménagement du Territoire (ZEAT) de l'INSEE (1969, 2020), illustrées par la carte ci-dessous. La zone Est a par ailleurs, l'avantage de couvrir la frontière allemande.

Figure 5.13 Les 8 Zones d'Etude et d'Aménagement du Territoire (ZEAT) de l'INSEE



Les statistiques descriptives de la mortalité sont indiquées dans le graphique ci-après :

Figure 5.14 Taux de mortalité quotidienne par Zones d'Etude et d'Aménagement du Territoire



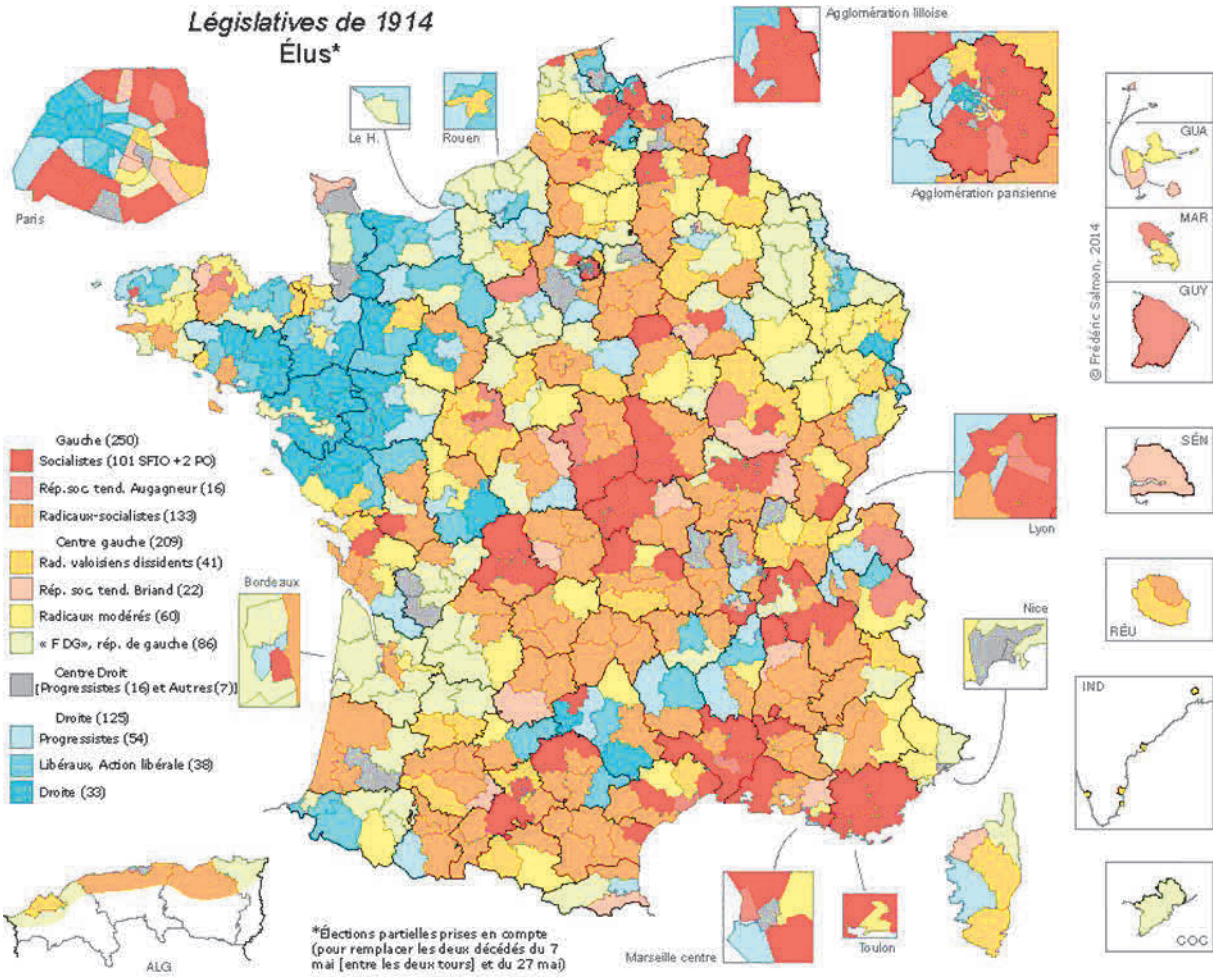
Les tests statistiques indiquent une association positive significative dans le cas de la région Est (V-Cramer de 27%) et une association négative faible pour la Méditerranée et le Sud-Ouest (V-Cramer de -13%). L'analyse de variance (modèle ANOVA utilisé ici puisque la variable dépendante est quantitative - le taux de mortalité quotidienne, alors que la variable indépendante est qualitative - zones géographiques) indique que l'origine géographique explique, toutes choses égales par ailleurs, 9% de la variance de mortalité entre régions.

Ces résultats indiquent un possible effet frontière à l'Est et pour la région parisienne, un effet proximité avec les lieux de combat (voir Guillot et Parent, 2018 ; Gilles, Guironnet et Parent, 2014 ; Guironnet et Parent, 2018).

6.2. Mortalité dans l'infanterie selon l'orientation politique du département d'origine du régiment

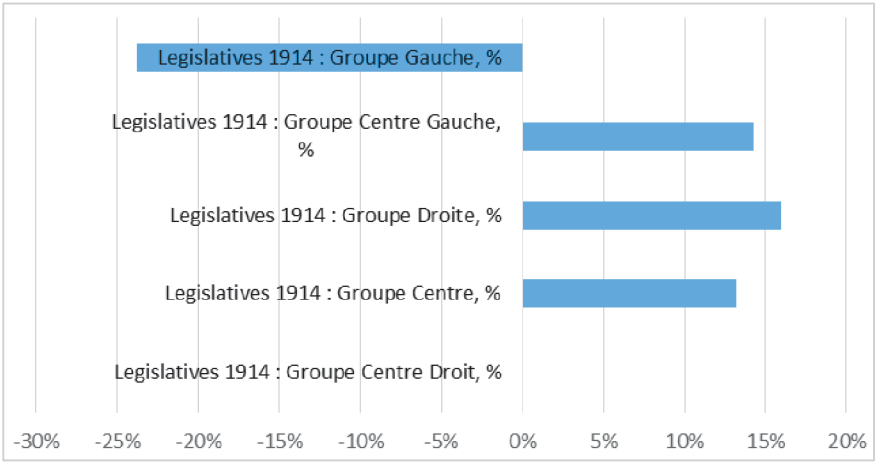
Pour creuser une possible dimension politique dans la mort au front, nous nous sommes appuyés sur une cartographie des résultats des élections législatives de 1914, circonscription par circonscription. Nous avons reproduit une échelle de la Gauche à la Droite selon les partis politiques de l'époque (Salmon, 2001) en regroupant les partis de l'époque en trois grands groupes (Gauche, Centre, et Droite), avec une subdivision du Centre entre Centre Gauche et Centre Droite. Les résultats des votes par circonscription sont ensuite agrégés au niveau départemental.

Figure 5.15 Votes aux élections législatives de 1914 par circonscriptions électorales (Salmon, 2001)



Le modèle utilisé ici est le PLS, puisque non seulement les variable exogènes et endogènes sont toutes deux quantitatives, mais aussi que les variables exogènes sont potentiellement corrélées entre elles puisqu'il existe des affinités ou des oppositions de vote entre partis politiques.

Figure 5.16 Corrélation de la mortalité dans les régiments d'infanterie avec les votes aux législatives de 1914 dans les départements d'origine



La dimension politique explique là encore 7% de la variance de mortalité observée. Nous trouvons une association positive faible entre taux de mortalité et importance du groupe de Droite ; une association positive faible avec le groupe Centre et le groupe Centre gauche ; une association négative plus marquée avec le groupe de la Gauche.

Si l'on admet le découpage arbitraire entre une Droite « plus patriotique » (que la Gauche, selon les catégories de l'époque) et une Gauche « plus pacifiste » que la Droite, alors la mortalité aurait davantage touchée les départements « patriotes », avec toutes réserves afférentes à ce type d'association. Le graphique précédent est tout à fait parlant à cet égard.

Ces résultats sont à mettre en parallèle avec les résultats de Guironnet et Parent (2018) et Guillot et Parent (2018) qui trouvent un effet semblable mesuré, pour les premiers avec le vote d'extrême gauche (mais sujet à un biais d'endogénéité avec le critère de la jeunesse du département), pour les seconds avec le taux d'abstention du département (proxy d'une « distance au nationalisme).

Nous ne pouvons donc, et c'était déjà le sens de la conclusion de ces deux articles, ignorer une association faible mais significative du facteur politique avec le taux de mortalité quotidien au front. Ici, le niveau de découpage plus fin du vote rend l'analyse plus complète au niveau des partis politiques.

6.3. Mortalité dans les RI selon le profil économique du département d'origine du régiment

L'analyse de la dimension économique du taux de mortalité au front fait ressortir deux facteurs significatifs : Le taux de chômage positivement (mais faiblement) corrélé avec la mortalité journalière du RI. Le taux de population active dans l'agriculture est, quant à lui, corrélé négativement avec la mortalité journalière du RI. Au total, l'analyse PLS estime que la dimension économique représente 3% de la variance de la mortalité expliquée.

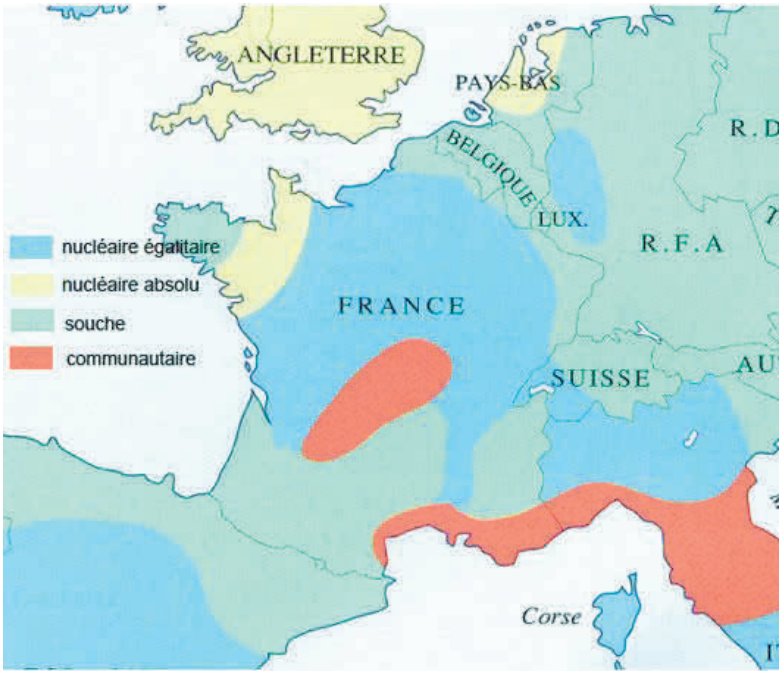
Ces résultats sont extrêmement intéressants car les variables économiques sont peu nombreuses à ressortir comme significatives. Ces résultats corroborent les résultats trouvés par Guironnet et Parent (2018) et Gilles et al. (2014) et déjà évoqués dans le rapport Marin (1921) : l'idée intéressante est que le monde ouvrier a pu être « relativement protégé » dans le cadre de l'effort industriel de guerre. On trouve ici que le fait d'être originaire d'un département à fort taux de chômage est associé à un taux de mortalité quotidienne plus élevé. Concernant le facteur rural, Guillot et Parent (2018) trouvent également que le fait d'être issu de départements agricoles « riches » favorise une durée de survie plus longue au front.

Si ces trois dimensions (géographiques, politiques et économiques) confortent les résultats d'études antérieures (et à ce titre, nos résultats peuvent se lire comme des tests de robustesse de ces travaux), l'innovation majeure de ce chapitre concerne l'adjonction de deux nouvelles dimensions, l'une appelée anthropologique, l'autre sociétale.

6.4. Mortalité dans les RI selon le profil anthropologique du département d'origine du régiment

L'analyse s'appuie ici sur les types familiaux catégorisés par Todd (1983) : famille communautaire, famille autoritaire, nucléaire égalitaire, nucléaire absolue.

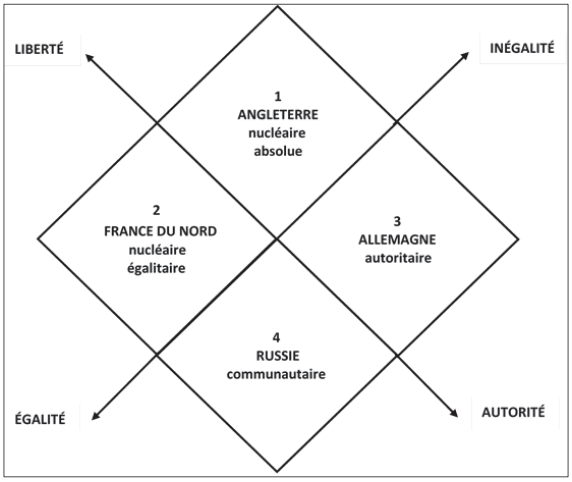
Figure 5.17 Types familiaux en France pré-industrielle, Todd (1983)



Communautaire : famille communautaire exogame ; Souche : famille souche ou autoritaire ; Nucléaire absolue : famille nucléaire absolue ; Nucléaire égalitaire : famille nucléaire égalitaire.

Ce découpage de Todd (1983) organise les types familiaux en France et dans le monde, autour de deux axes : un axe qui oppose liberté à autorité ; un axe qui oppose égalité à inégalités.

Figure 5.18 Parangons européens des types familiaux



Todd (1983, p.36)

Dans le cas de la France, cette grille de lecture appliquée à la mortalité des régiments définis sur une base territoriale départementale s'avère extrêmement riche d'enseignements : tout d'abord, elle

explique à elle seule 8% de la variance de la mortalité (par le modèle PLS comme avec le modèle de régression sur composantes principales).

Pour rappel les l'intensité et le sens des associations sont mesurés par le V-Cramer qui se lit ici comme un coefficient de corrélation de Pearson (cf. section 5).

Nous trouvons une association positive significative (V-Cramer) entre certaines valeurs anthropologiques des départements d'origine, et le taux de mortalité quotidienne au front, notamment le fait de provenir d'un département ayant des valeurs de liberté (+25 %) et d'égalité (+17%).
A contrario, d'autres valeurs anthropologiques sont associées négativement avec le taux de mortalité quotidien au front : il s'agit des valeurs d'autorité ou verticalité (discipline, autorité forte) à -25% et d'inégalités à -17%.

Nous constatons également des associations positives et négatives de la mortalité d'infanterie avec les types familiaux : famille nucléaire-égalitaire (liberté, égalité) à + 26%, famille souche (autoritaire et inégalitaire) à -17%, famille communautaire exogame (égalitaire, autoritaire et exogame) à -11%.

Ce résultat enrichit les grilles d'analyse traditionnelle et augmente significativement le pourcentage de variance totale expliquée. Selon ce découpage, les départements de type familial nucléaire et égalitaire ont une prédisposition à être plus touchées que les autres : ses deux composantes (liberté et égalité) sont positivement associées à la mortalité quotidienne au front ; sont concernées les départements du bassin et région parisienne essentiellement.

A contrario, la famille souche (autoritaire et inégalitaire), qui prédispose à une sujétion au commandement et à la hiérarchie, est associée négativement au taux de mortalité. Les résultats sont très clairs sur un plan statistique. Sont concernés, la pointe bretonne, le pays basque et la Gascogne, et la zone frontalière avec l'Allemagne et la Suisse.

A mi-chemin entre ces deux cas très marqués, se trouvent les régions de type familial communautaire exogame : le Limousin et le pourtour méditerranéen, avec une composante « mortifère » liée aux valeurs d'égalité, surcompensée par des valeurs de « survie » liées aux valeurs autoritaires. Au final, ce type familial a une association légèrement négative avec la mortalité.

Si l'on extrapole à partir de ces résultats, on peut dire qu'un Etat-major soucieux de préserver ses troupes aura tendance à les recruter dans les départements de valeurs « inégalitaires » et/ou « autoritaires ».

6.5. Mortalité dans les RI selon le profil cognitif et sociétal du département d'origine du régiment

Ici nous reprenons et testons une autre hypothèse de Todd (1979), cette fois-ci relative à la période de la Grande guerre, qui conjecture l'existence, parmi les pays belligérants, d'une corrélation positive et importante entre les taux de mortalité nationaux à la guerre et les niveaux de « pathologies sociales » nationales mesurés avant-guerre par le triptyque suicide-alcoolisme-folie. Ces dimensions sont renseignées dans notre base de données.

Nous trouvons effectivement systématiquement des corrélations positives et significatives entre la mortalité au front et les variables de pathologies sociales mesurées par le suicide (19%), la folie (19%), l'alcoolisme mesuré l'ivresse publique (14%). Ce triptyque mesure un comportement autodestructeur à

travers l'alcoolisme et le suicide, et nos tests statistiques montrent que ces comportements autodestructeurs sont corrélés positivement avec la mort au front.

Au-delà de ce triptyque, nous trouvons que les comportements criminels et violents dans les départements d'origine sont également associés positivement (quoique faiblement) à la mortalité au front : viol (12%), coups et blessures (12%).

Nous avons également analysé l'association de la mortalité au front avec deux autres variables sociales structurantes : l'éducation (taux d'alphabétisation) et pratique religieuse (taux d'ordination des prêtres) : les coefficients se sont révélés positifs pour ces deux variables, avec respectivement 18% et 10%.
En première analyse, le lien positif et significatif entre taux d'alphabétisation et mortalité au front peut paraître surprenant, voire paradoxal. Cependant, nous notons aussi un lien positif entre taux de suicide et taux d'alphabétisation. Ce résultat est commun à de nombreuses études statistiques et éclaire l'association positive entre mortalité au front et taux d'alphabétisation.

Au total la dimension sociale ici analysée (pathologies sociales, éducation, religion) expliquent globalement 13% de la variance de la mortalité, soit plus que les facteurs anthropologiques (8%), géographiques (9%), politiques (7%) et économiques (3%).
Il s'agit d'un apport notable par rapport aux études existantes sur la mortalité au front (Gilles et al., 2014 ; Guironnet et Parent, 2018 ; Guillot et Parent, 2018) qui ne prenaient pas en compte ces facteurs, en se limitant aux aspects démographiques, géographiques, économiques, et politiques pour les deux dernières références citées.

Nous concluons l'analyse de ce chapitre par l'analyse des facteurs strictement militaires dans la mort au front.

6.6. Mortalité dans les RI selon les caractéristiques militaires de l'engagement

Parmi ces critères strictement militaires, nous avons distingué :

- le fait pour un régiment d'appartenir (ou non) à l'armée mobile (IVème corps d'armée, comprenant les régiments de la région militaire XII – Limoges et XVII Toulouse).
- les modalités de participation et d'engagement des régiments aux batailles et combats (cf. *Figure 5*.4) à savoir si les régiments étaient pré-positionnés ou envoyés avant le début ou durant les hostilités, ou si le régiment était absent des combats (en distinguant deux types d'absence, l'absence proche à une distance de moins de deux départements et l'absence pour position lointaine)
- la mortalité selon le type de confrontation : grandes batailles *versus* autres batailles et combats de moindre envergure

Les tests d'association entre le taux de mortalité au front et le statut d'armée mobile ne révèlent pas d'associations significatives. Une des hypothèses évoquée dans Gonzalez-Feliu et Parent (2017) et Gonzalez-Feliu et al. (2018) n'est pas vérifiée ici, puisque l'on trouve que le fait, pour un régiment, d'appartenir à l'armée mobile n'est pas discriminant pour la mortalité au front.

Nous rappelons, pour mémoire, que l'armée mobile, cantonnée en retrait de la ligne de front au déclenchement des hostilités a absorbé les premiers chocs, les régiments les composant étant envoyés en soutien face aux premières percées allemandes.

Nous rappelons aussi que les deux premières années de guerre ont été les plus meurtrières. Ceci constitue des faits stylisés reconnus, mais l'indicateur calculé de mortalité quotidienne au front corrige de cet effet. Si le facteur « armée mobile » ne ressort pas statistiquement ici, cela tient à la construction du taux de mortalité ici analysé, et ne remet pas en cause le fait qu'au début de la guerre, le fait pour un régiment d'appartenir à l'armée mobile (Corps d'armée XII et XVII) et de participer aux épisodes les plus meurtriers de la guerre a contribué, toutes choses égales par ailleurs à exposer le régiment à une mortalité plus forte.

Cette statistique nous renseigne cependant sur le fait que les régiments de l'armée mobile ayant subi le choc des batailles les plus meurtrières n'ont pas un taux de mortalité au front différent des autres régiments ayant connu les mêmes épisodes meurtriers.
Pour ce qui est des autres caractéristiques strictement militaires, l'analyse statistique révèle logiquement un taux de mortalité exposé de façon hiérarchique croissante au degré d'exposition du régiment ; le type bataille (grande ou simple combat) s'avère également discriminant.

6.6.1. Mortalité dans les grandes batailles

Les grandes batailles sont caractérisées par le nombre de régiments mobilisés engagés et la durée de l'épreuve. Ces caractéristiques permettent d'isoler les grandes batailles dont les plus connues sont, par ordre chronologique : la bataille des frontières, la bataille de la Marne, la bataille de Champagne, la bataille de la Somme, l'offensive du Chemin des Dames, les batailles de l'Aisne. Nos critères statistiques pour désigner une « Grande bataille » permettent d'en recenser 41, d'une durée moyenne de 46 jours et engageant en moyenne 117 régiments.

En moyenne un régiment d'infanterie a participé à 36% de grandes batailles dont 24% pré-positionné depuis au moins 15 jours avant le déclenchement des hostilités. Pour 6% des batailles la participation s'est faite en étant envoyé dans les deux semaines précédant le début des combats ; 6% étant envoyé après le début des combats.
Ci-dessous les résultats de calcul des corrélations et des tests de significativité montrent que :

Tableau 5-2 Test d'association entre le taux de mortalité par jour pondéré (par la mortalité) pour 1000 soldats initiaux d'un RI et les modalités de Participation des RI aux grandes batailles

Target	Input	p-value	Correlation	Confidence Band	Conclusion
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Participants (% Batailles)	4,35E-38	0,582	0,514; 0,643	hypothèse H0 rejetée au seuil de 5%
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Presents a J-15 (% Batailles)	4,60E-33	0,548	0,476 ; 0,612	hypothèse H0 rejetée au seuil de 5%
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Envoyes Avant (entre J-14 et J) (% Batailles)	1,10E-04	0,191	0,095 ; 0,283	hypothèse H0 rejetée au seuil de 5%
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Envoyes a Partir du Debut (a partir de J+1) (% Batailles)	2,82E-01	0,054	-0,044 ; 0,150	hypothèse H0 rejetée au seuil de 5%

Hypothèse H_0 : Absence de corrélation ; Probabilité de H_0 : p-value

Certaines corrélations sont naturellement fortes : la participation aux grandes batailles a une corrélation de 58% avec le taux de mortalité quotidienne.
Nous observons un gradient croissant de corrélation selon l'ancienneté de présence sur zone : Présence à J-15 (+55%), Envoyé entre J et J-14 (+19%), Envoyé à partir de J+1 (0%). Au total, la participation aux grandes batailles explique 30% de la variance du taux de mortalité (modèle de régression PLS).

Ensuite, on observe que la participation d'un régiment aux autres combats et batailles est plus meurtrière encore:

6.6.2. Mortalité dans les autres combats et batailles :

En moyenne un régiment d'infanterie a participé à 22% des autres batailles et combats dont 15% pré-positionné depuis au moins 15 jours avant le déclenchement des hostilités. Pour 5% des autres batailles et combats, la participation s'est faite en étant envoyé dans les deux semaines précédant le début des combats ; 2% étant envoyé après le début des combats. Ces chiffres sont naturellement plus faibles que pour les grandes batailles qui mobilisent plus de régiments. En revanche, ces combats de moindre ampleur sont plus meurtriers comme le montrent les résultats ci-dessous.

Tableau 5-3 Test d'association entre le taux de mortalité par jour pondéré (par la mortalité) pour 1000 soldats initiaux d'un RI et les modalités de Participation des RI aux batailles et combats

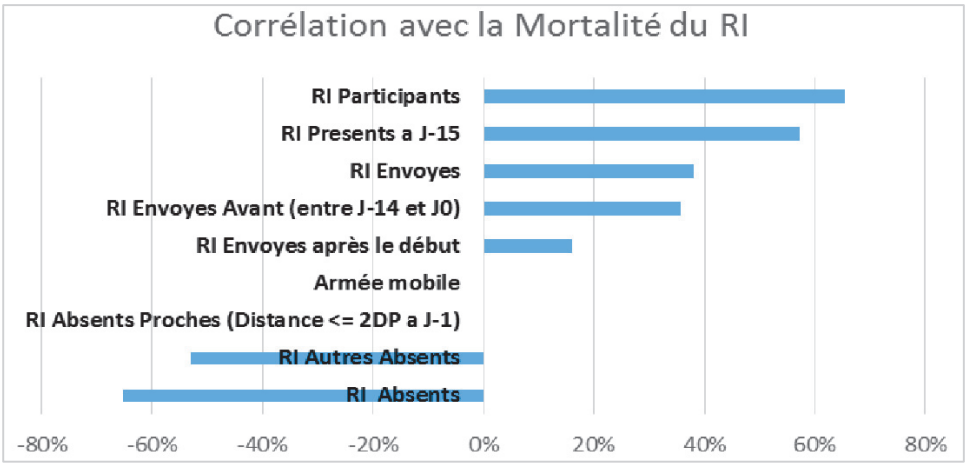
Target	Input	p-value	Correlation	Confidence Band	Conclusion
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Participants (% Batailles)	8,35E-42	0,605	0,539 ; 0,664	hypothese H0 rejetee au seuil de 5%
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Presents a J-15 (% Batailles)	1,61E-26	0,496	0,419 ; 0,567	hypothese H0 rejetee au seuil de 5%
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Envoyes Avant (entre J-14 et J) (% Batailles)	5,78E-14	0,362	0,274 ; 0,444	hypothese H0 rejetee au seuil de 5%
MPF par jour pondere (par la mortalité) pour 1000 soldats initiaux du RI	RI Envoyes a Partir du Debut (a partir de J+1) (% Batailles)	3,27E-05	0,205	0,110 ; 0,296	hypothese H0 rejetee au seuil de 5%

Hypothèse H_0 : Absence de corrélation ; Probabilité de H_0 : p-value

La participation aux autres batailles et combats a une corrélation de 61% avec le taux de mortalité quotidienne, corrélation un peu plus forte qu'avec les grandes batailles, ce qui signifie donc que ces engagements sont légèrement plus meurtriers.
Nous observons toujours un gradient croissant de corrélation selon l'ancienneté de présence sur zone : Présence à J-15 (+50%), Envoyé entre J et J-14 (+36%), Envoyé à partir de J+1 (20%). Au total, la participation aux autres batailles et combats explique 35% de la variance du taux de mortalité (modèle de régression PLS).

6.6.3. Mortalité du régiment tout type de bataille et combat confondu :

Figure 5.19 Modalités de Participation (ou non) des RI aux batailles et combats



Lorsqu'on fait une régression large incluant l'ensemble des variables militaires (en utilisant le modèle PLS), nous expliquons 42 % de la variance du taux de mortalité. Il ressort donc que la mortalité au front reste principalement déterminée par les engagements des régiments dans les combats. Mais si ce chiffre est important, il n'explique pas toute la variance du taux de mortalité. L'apport de ce travail est aussi d'avoir su déterminer la part de la variance expliquée par d'autres facteurs non négligeables et tout aussi significatifs. Le bilan récapitulatif de l'analyse des contributions est proposé dans la section suivante.

7. Bilan des facteurs de la mortalité des régiments

Les graphiques ci-dessous expriment les contributions une à une des différentes dimensions passées en revue dans l'explication de la mortalité quotidienne au front. Il ressort que, massivement, les facteurs militaires (engagement des régiments) prédominent dans l'explication de la mortalité avec 42% de variance de mortalité expliquée.
Ensuite, viennent, par ordre d'importance les facteurs sociétaux (13% de variance expliquée), géographiques (9%), anthropologiques (8%), politiques (7%), économiques (3%). Cette hiérarchie a son importance : l'ensemble de ces facteurs non militaires expliquent 20% (dans un modèle conjoint PLS) de la variance de mortalité, soit près de la moitié du la variance expliquée par les facteurs purement militaires ce qui est considérable.

Figure 5.20 Part de variance de mortalité quotidienne expliquée, selon les dimensions analysées

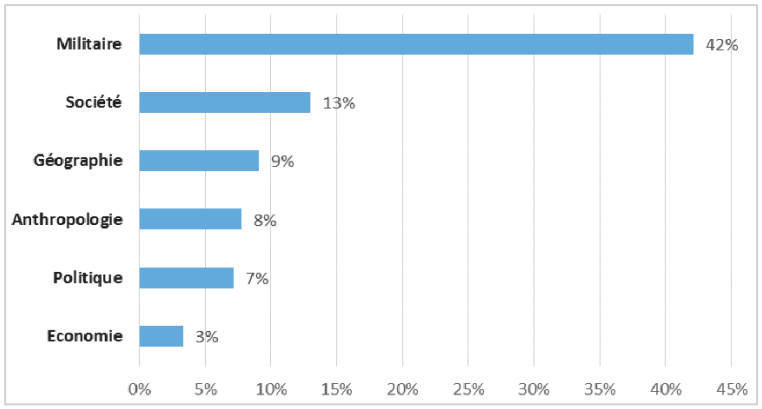
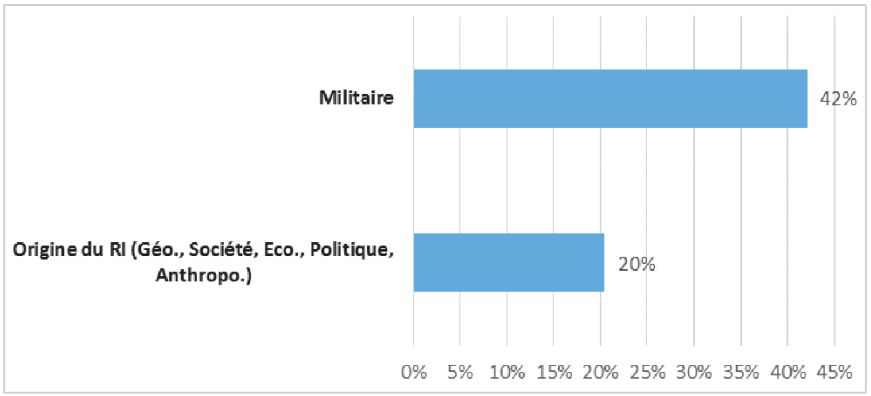


Figure 5.21 Part de variance de mortalité quotidienne expliquée, selon les dimensions militaire et origine du RI



Nous constatons également que la variance expliquée par l'ensemble des facteurs non militaires liés aux départements d'origine des régiments est largement inférieur à ce que serait l'addition des variances expliquées prises séparément. Ceci signifie qu'il y a entre ces variables non militaires une très forte colinéarité. Il est de ce point vue surprenant de voir en comparant entre elles les variances individuelles, la prééminence des facteurs sociétaux et anthropologiques sur les facteurs économiques et politiques. Avec toutes les limites afférentes à ce type de comparaison, la multicolinéarité trouvée suggère une imbrication et redondance d'information entre l'ensemble des facteurs liés à l'origine départemental du régiment.

Enfin, lorsqu'on considère l'ensemble de tous les facteurs testés (militaires et non militaires), dans un modèle explicatif de type PLS qui neutralise les effets de la multicolinéarité, la part de variance de mortalité expliquée est de 54%, ce qui signifie que d'autres facteurs, en plus des origines et de l'engagement étudiés ici, pourraient expliquer une partie de la variance restante (on pense notamment aux maladies, aux accidents, aux morts en prison, notamment).

8. Conclusion

Ce chapitre est une contribution à la recherche sur la Grande guerre de plusieurs manières. Il innove par :
La définition d'un indicateur de mortalité original qui corrige des effets de taille, de durée d'existence, d'exposition au risque de mortalité des régiments. Cet indicateur inédit, a vocation à être généralisé à tout type de conflit à travers l'histoire, ce qui est un apport pour une cliométrie des conflits armés.

La constitution d'une base de données unique, historique, multidimensionnelle, pluridisciplinaire.
Il s'agit d'une base de données cliométriques de la Guerre 14-18 : elle est composée de 3 grandes catégories de données. Premièrement, les données sur les positions quotidiennes des régiments d'infanterie durant la Grande Guerre. Deuxièmement, les données descriptives de régiments : rattachement administratif, effectifs, origine géographique, date de mobilisation, date de dissolution, projection à l'étranger durant la guerre. Et troisièmement, les données propres à l'origine géographique des régiments : choix politiques aux élections législatives sur presque 1 siècle (1818 à 1914), données anthropologiques, cognitives et sociales, politiques, démographiques et économiques.
Cette base de données cliométriques de la Grande Guerre a été nécessaire à la réalisation de toutes les analyses statistiques et historiques réalisées dans le présent chapitre.

Le développement d'un logiciel internet d'animation géomatique a été développé grâce aux géolocalisations quotidiennes des régiments. Cette application internet reproduit les mouvements des régiments d'infanterie sur le front durant toute la guerre. Elle permet de détecter les concentrations géographiques de troupes d'infanterie, d'y identifier le cas échant les combats célèbres s'y déroulant, ainsi que les modalités et la liste des régiments impliqués.
Ce logiciel internet vocation à terme à être mis à disposition de la communauté des chercheurs de l'histoire de la Grande Guerre.

La conception et le développement d'une application de mesure et de tests d'associations linéaires et non-linéaires, quelque-soit le type de variable testée, qualitative ou quantitative.

Sur un plan méthodologique, afin d'estimer la part de variance de mortalité expliquée, nous avons systématiquement corrigé les effets de colinéarité des variables explicatives en utilisant des modèles de régression PLS (Partial Least Square) et régressions sur axes factoriels issus d'ACP ou d'ACM : cela a permis d'estimer des modèles valides qui donnent les parts de variance de mortalité expliquées pour les groupes de variables testés.

Sur un plan stratégique et organisationnel, l'analyse des valeurs anthropologiques d'une part, et des modalités associées au « timing » de l'engagement des régiments au combat d'autre part, est entièrement inédite et parvient à réduire la part traditionnellement dévolue au hasard dans la mort au front. Des choix anthropologiques judicieux, pourraient potentiellement permettre de réduire le risque de mortalité, notamment par un recrutement préférentiel dans les régions à valeurs « inégalitaires » et/ou « autoritaires » dans un cadre très restrictif.

Sur le plan de la connaissance historique, nous avons pu mesurer et comparer les contributions à la mortalité des facteurs militaires et d'autres facteurs, au sein desquels, les aspects sociaux et anthropologiques l'emportent sur les facteurs économiques. En ce sens, cette thèse enrichit l'approche cliométrique existante de variables anthropologiques et sociétales.

9. Bibliographie

Archives départementales du Nord (2020), La Première Guerre mondiale : les soldats "Morts pour la France", consulté en ligne le 22/08/2018, https://archivesdepartementales.lenord.fr/?id=388

Fiévet G. (1992). De la stratégie militaire à la stratégie d'entreprise, InterEditions

FranceGenWeb (2018), Base recoupements 1914-1918, Recoupements par régiments et par lieux 1914-1918, consulté en ligne le 01/10/2018, http://www.francegenweb.org/b1914-1918/

Gilles H., Guironnet J.P., Parent A. (2014a): « Géographie économique des morts de 14-18 en France », Revue économique, 2014/3 Vol. 65, p. 519-532

Gilles H., Guironnet J.P., Parent A. (2014b): « La géographie des Morts pour la France », Journal Le Monde, Supplément « Le journal du centenaire » du mardi 13 mai, double page

Gonzalez-Feliu J., Ovtracht N., Parent, A. (2018). Gare et Front: la logique de répartition des troupes dans le plan XVII de Joffre.

Gonzalez-Feliu J., Parent, A. (2016). “Clio-Combinatorics: A Novel Framework to Analyze Military Logistics Choices Using Operations Research Techniques”, Handbook of Research on Military, Aeronautical, and Maritime Logistics and Operations, p. 79-101

Guillot O., Parent A., (2018), “Farewell life, farewell love’: Analysis of survival inequalities among soldiers who “Died for France” during World War I”, Population, 73(3), 411–444.

Guironnet J.P., Parent, A. (2018), “Morts pour la France: Demographic or Economic Factors?”, Defence and Peace Economics

INSEE (2020), Zone d'études et d'aménagement du territoire ZEAT, consulté en ligne le 16/11/2020, https://www.insee.fr/fr/metadonnees/definition/c1910

INSEE (1969), Les huit zones d'études et d'aménagement du territoire. In: Economie et statistique, n°5, Octobre 1969. pp. 45-48.

Ministère des Armées (2018). Mémoire des hommes, Morts pour la France de la Première Guerre mondiale, consulté en ligne le 09/05/2018, https://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?larub=24&titre=morts-pour-la-france-de-la-premiere-guerre-mondiale

Salmon F. (2001). « Atlas électoral de la France, 1848-2001 », Seuil

Saporta G. (2011). « Probabilités, Analyse des données et Statistique ». Technip

Saporta, G. (1975). « Liaisons entre plusieurs ensembles de variables et codage de données Qualitatives ». Thèse de 3° cycle, Paris VI : Université Pierre et Marie Curie.

Todd E. (1979). Le fou et le prolétaire, Robert Laffont

Todd E. (1996). « L'invention de l'Europe ». Seuil

Todd E. (1998). L'Illusion économique : Essai sur la stagnation des sociétés développées, Paris, Gallimard

Todd E. (1999). « La Diversité du monde : Famille et modernité ». Seuil

Todd E., Le Bras H. (2012). L'invention de la France. Atlas anthropologique et politique, Gallimard, coll. «NRF Essais»

Todd E., Le Bras H. (2013). Le Mystère français, Seuil

Conclusion

La première originalité de notre approche cliométrique de l'histoire globale, est de proposer une grille de lecture quantitative, multidisciplinaire et systémique, basée sur des définitions concises et des indicateurs mesurables.
Notre grille quantitative et multidisciplinaire est composée de cinq domaines interdépendants, désignés par l'acronyme ACIDE : Anthropologie, Cognition, Institutions, Démographie et Economie.
La grille ACIDE a pour vocation à être systématiquement utilisé pour l'analyse des phénomènes économiques, afin d'éviter les tautologies réductrices (expliquer l'économie par l'économie) et permettre d'avoir une vision large et profonde de l'économie, étant donné que cette dernière est une activité humaine complètement encastrée dans la société qui la porte.
De manière plus large, la grille cliométrique multidisciplinaire est destinée à un usage analytique en histoire globale des domaines ACIDE.

Dans le domaine institutionnel, nous avons défini deux indices originaux : l'un mesurant l'inclusivité économique et l'autre la non-inclusivité économique. Il s'agit ici d'une des originalités de notre approche institutionnelle. Ce qui nous a permis de montrer de manière inattendue que la Chine a été plus inclusive économiquement que la moyenne de l'OCDE et la moyenne mondiale de 1995 à 2010. Même si l'inclusivité économique chinoise a chutée depuis lors, elle reste relativement proche de la moyenne mondiale.
Une autre originalité de notre approche institutionnelle consiste à caractériser les institutions via, d'une part, des indicateurs « externes » liés à la souveraineté et la puissance, et d'autre part, des indicateurs « internes » de gouvernance et de respect des droits de l'individu. Nous définissons ainsi une matrice 2x2 de mesure des institutions regroupées en quatre catégories : les institutions politiques internes, les institutions politiques externes, les institutions économiques internes et les institutions économiques externes. Notre approche quantitative et matricielle réhabilite la téléologie bicéphale (prospérité et puissance) de l'économie politique par Smith (1776, tome 2, P.11), trop souvent oubliée depuis l'Ecole classique.

L'analyse des dynamiques transitionnelles ACIDE, nécessite également de développer un temps transitionnel qui évolue au rythme de la transition étudiée et non du temps calendaire. Ce temps transitionnel permet de faire une analyse comparative de trajectoires transitionnelles des pays (au niveau global) dont les transitions se déroulent à des époques, à des rythmes et à des intensités hétérogènes.

Dans un domaine ACIDE donné, il est également nécessaire de détecter les phases historiques des trajectoires nationales des pays du monde (vue globale), et d'utiliser ces phases pour faire une analyse séquentielle des trajectoires nationales.
L'analyse séquentielle des trajectoires nationales, consiste à mettre en œuvre les méthodologies d'analyse des séquences d'états, ainsi que l'analyse des séquences d'événements majeurs : visualisations de séquences, statistiques descriptives des phases, typologies des trajectoires, modèles explicatifs des séquences d'états ou d'évènements, prospectives basées sur les règles d'associations séquentielles.

Une des limites des analyses longitudinales (comme l'analyse séquentielle) est la disponibilité de données quantitatives riches sur le temps long (plusieurs générations).
Notre solution originale à cette limite a été d'inférer les trajectoires moyennes (globale ou par groupes de pays) sur le temps long, via une coupe transversale à une date donnée.
Cela suppose de définir (nous avons déterminé pour cela un test statistique permettant de sélectionner les séries temporelles pouvant jouer ce rôle) un temps transitionnel pour le domaine étudié, qui permet alors de pallier l'absence du temps calendaire dans la coupe transversale. Avec le temps transitionnel ainsi créé, il est possible de caractériser une trajectoire transitionnelle moyenne (globale ou par groupe de pays) en l'absence de temps calendaire.

Il est intéressant d'utiliser l'aspect global (universel) et propagatoire de la transition dans un domaine disciplinaire ACIDE donné, pour effectuer des prévisions actuarielles de long terme pour les pays à transition plus tardive, en prenant en compte, non seulement leurs trajectoires propres, mais également l'historique des trajectoires des pays à transition précoce.
Notre solution originale à ce besoin passe par la création de variables exogènes dites « variables cliométriques » qui permettent de faire le lien statistique entre les deux types de pays.
Nous appliquons notre technique dans le domaine démographique pour les prévisions actuarielles à long terme des taux de mortalité (donc également de l'espérance de vie) par âge et sexe.
Les modèles généralisés et optimisés que nous avons introduits ont permis d'améliorer les modèles actuariels classiques à l'horizon de 50 ans. Pour le pays émergent étudié (Inde), la prise en compte des « variables cliométriques » a abouti à des prévisions actuarielles beaucoup plus précises à l'horizon de 40 ans.

Toutefois, notre méthode de prévisions cliométriques à long terme se généralise également à d'autres variables démographiques (fécondité ou même solde migratoire par exemple), et à de nombreuses variables dans les autres domaines de la grille ACIDE.

Enfin, nous appliquons la grille cliométrique multidisciplinaire ACIDE à l'étude des déterminants multidisciplinaires de la mortalité dans l'infanterie française, au cours de la première Guerre mondiale. Pour cela il nous a fallu développer une mesure originale de la mortalité en conflit armé, qui tient compte évidement des différences d'effectifs entre unités, mais aussi de la durée de mobilisation des régiments et de l'hétérogénéité du risque de mortalité auquel sont soumises les différentes unités.
Nous avons ainsi pu montrer que durant le Grande Guerre de 14-18, la part de mortalité dans l'infanterie expliquée par les facteurs anthropologiques, cognitifs, et socio-économiques est approximativement égale à la moitié de la part de mortalité dus aux causes purement militaires.
Notre démarche historique utilisée pour la mortalité dans l'infanterie durant la première Guerre mondiale, peut se généraliser à l'analyse des déterminants multifactoriels ACIDE d'une variable à expliquer dans un épisode historique quelconque.

Les perspectives de recherches futures varient selon les thématiques traitées.

Le premier sujet de recherche future concerne ce que nous pourrions appeler le « piège post-industriel ».
La transition économique aboutit dans sa phase d'achèvement à une forte baisse de la part de l'emploi et de la valeur ajoutée industrielle et manufacturière, au profit du secteur des services. L'amplitude de cette perte du tissu industriel varie selon la tradition industrielle et l'anthropologie des pays (Todd, 1998).
Il serait très utile d'investiguer sur les conséquences de cette transformation structurelle qui constitue une perte de capacités techniques et technologiques, sur les performances macroéconomiques (croissance, inégalités par exemple), ainsi que sur divers domaines comme les institutions (libertés civiles, prélèvements obligatoires par exemple), la cognition (scores dans les tests cognitifs) et la démographie (notamment l'espérance de vie).
Pour ne pas rester dans une logique purement « spectatrice », il pourrait également être intéressant de faire une revue historique pour proposer des pistes qui ont montré une certaine efficacité par le passé et qui pourraient aider à sortir du piège post-industriel.

Dans toutes les transitions étudiées, les résultats obtenus ne s'étendent pas automatiquement aux « petits » Etats qui ont été préalablement exclus de l'étude : ces derniers pourront faire l'objet d'une étude spécifique intéressante.
Dans la même veine, il y a un intérêt théorique à décliner les théories transitionnelles ACIDE pour les pays dont une part importante de la valeur ajoutée intérieure provient de l'exploitation de ressources naturelles (pêche, forêts, mines, charbon, pétrole, gaz), et qui souffrent traditionnellement de la

« maladie hollandaise » (*dutch disease*, Ebrahim-zadeh [2003]) nocive pour l'industrie manufacturière et dont la trajectoire politique est susceptible d'ingérences et de perturbations spécifiques (Acemoglu et Robinson, 2015, p.426, 544) relevant de la « malédiction des ressources naturelles » *(natural resource curse)*.

En ce qui concerne l'analyse séquentielle longitudinale, l'absence de données économiques structurelles sur le temps long a été une limite à l'exploration historique de cette étude. Le plus pénalisant est le manque de données nationales de long terme du secteur industriel, qui est une composante fondamentale de la structure économique d'un pays. Nous pourrions tenter une analyse séquentielle de temps long, c'est-à-dire depuis 1875, sans la part des emplois industriels et en nous restreignant à seulement aux deux composantes dont les données sont disponibles : La part d'emplois agricoles et le PIB par habitant.

Enfin, pour ce qui est de la prospective actuarielle des taux de mortalité et espérances de vie, des travaux complémentaires pourront permettre de tester empiriquement notre seconde méthode d'estimation des variables cliométriques et en comparer les performances avec la méthode qui a été mise en œuvre dans les travaux empiriques présentés.
D'autres méthodes actuarielles de référence pourront également être testées, contre les nouvelles méthodes que nous proposons.
Notre méthode de prospective cliométrique et actuarielle pourrait également s'étendre à d'autres variables démographiques ou économiques nécessaires à la gestion des risques.

Bibliographie

Acemoglu D. (2012), “Institutions, Political Economy and Growth”, Nobel Symposium, Stockholm, September 4, 2012, http://nobel2012.iies.su.se/frontpage.html

Acemoglu D., Autor D., (2012), What Does Human Capital Do? A Review of Goldin and Katz's Race between Education and Technology, June 2012, Journal of Economic Literature, 50(2), pp. 426–463.

Acemoglu D., Robinson J.A., (2012), Why Nations Fail : The Origins of Power, Prosperity, and Poverty, Crown Business, 546p.

Acemoglu D., Robinson J.A., (2015), Prospérité, puissance et pauvreté : Pourquoi certains pays réussissent mieux que d'autres, Markus Haller, 637 p.

Acemoglu, D. ; Robinson, J. (2001). "A Theory of Political Transitions". American Economic Review. 91 (4): 938–963.

Acemoglu D., Johnson S., Robinson J. A. (2001b), ‘Colonial origins of comparative development : an empirical investigation’, American Economic Review, 91(5), 1369–401.

Afonso A., Rault C. (2008). « Bootstrap Panel Granger-Causality Between Government Spending and Revenue ». Working Papers WP 39/2008/DE/UECE, School of Economics and Management : Technical University of Lisbon

Afonso A., Rault C. (2009). « Bootstrap Panel Granger-Causality Between Government Budget and External Deficits for the EU ». Working Papers WP 02/2009/DE/UECE, School of Economics and Management : Technical University of Lisbon

Aghion P., Cohen E. (2004), "Education et croissance", La Documentation Française

Aghion P., Howitt P. (2006). "Joseph Schumpeter Lecture Appropriate Growth Policy: A Unifying Framework", Journal of the European Economie Association, 4, 269-314.

Aghion P., Howitt P. (2010)., L'économie de la croissance, Economica

Allais M. (1999). « La mondialisation: La destruction des emplois et de la croissance : l'évidence empirique ». Clément Juglar

Altinok N., Aydemir A. (2017), Does one size fit all? The impact of cognitive skills on economic growth, Journal of Macroeconomics, 2017, vol. 53, issue C, 176-190

American Psychological Association (2019). Glossary of psychological terms, https://dictionary.apa.org/cognition, consulté le 28/09/2019

Archives départementales du Nord (2020), La Première Guerre mondiale : les soldats "Morts pour la France", consulté en ligne le 22/08/2018, https://archivesdepartementales.lenord.fr/?id=388

Artus P., VIRARD M.-P. (2015), « Croissance zéro, comment éviter le chaos? », Fayard

Artus P., Cette G. (2004). Productivité et croissance. Conseil d'Analyse Economique, La Documentation française. Paris.

Artus P., García-Peñalosa C., Mohnen P., (2014), Redresser la croissance potentielle de la France, Notes du conseil d’analyse économique 2014/6 (n° 16), pages 1 à 12

Avdeev A. (2014). « Histoire de la population mondiale et la transition démographique ». Ressources ouvertes de l’Institut de démographie de l’Université de Paris 1, Année académique 2014/2015. (document consulté le 05/06/2015). Disponible sur : http://dmo.econ.msu.ru/teaching/Histpop/

Azar, A., Staudinger, U. M., Slachevsky, A., Madero-Cabib, I., & Calvo, E. (2019). From Snapshots to Movies: The Association Between Retirement Sequences and Aging Trajectories in Limitations to

Perform Activities of Daily Living. Journal of Aging and Health, 31(2), 293-321. https://doi.org/10.1177/0898264318782096

Bairoch P. (1997). « Victoires et déboires I ». Folio

Bairoch P. (1997). « Victoires et déboires III: histoire économique et sociale du monde au XVIe siècle à nos jours ». Folio

Bairoch P. (2005). « Mythes et paradoxes de l'histoire économique ». La découverte

Bandyopadhyay, D. (1997). "Distribution of Human Capital and Economic Growth," Working Papers 157, Department of Economics, The University of Auckland.

Barro R.J., Lee J-W, (2015), Education Matters: Global Schooling Gains from the 19th to the 21st Century, Oxford University Press

Barro, R. (2001). "Human Capital and Growth," American Economic Review, American Economic Association, vol. 91(2), pages 12-17, May.

Baten J. et al. 2016. A History of the Global Economy: 1500 to the Present, Cambridge University Press

Baudelle G. (2003). « Géographie du peuplement ». Armand Colin

Baudellle G. (2003). « Géographie du peuplement ». Armand Colin - 2ème édition

Bazil B. et Fourastié J. (1980). « Le jardin du voisin, les inégalités en France »

BBC (2018), Ulster University withdraws status from Prof Richard Lynn, 14 April 2018, consulté en ligne le 13/03/2020, https://www.bbc.com/news/uk-northern-ireland-43768132

Beaujard P., Berger L., Norel p. (2009). Histoire globale, mondialisations et capitalisme. La Découverte

Becker G. (1964) : Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, New York, Columbia University Press

Becker G. S. (1962), "Investment in Human Capital: A Theoretical Analysis," Journal of Political Economy 70, no. 5, Part 2 (Oct., 1962): 9-49.

Becker G. S. (1994), "Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education", Third Edition, The University of Chicago Press

Becker G. S. (2002),The age of human capital

Benoît Prevost (2010), « Douglass North : hétérodoxie néo-institutionnelle versus néolibéralisme ? », Revue de la régulation [En ligne], 7 | 1er semestre / Spring 2010, mis en ligne le 18 juin 2010, consulté le 05 août 2019. URL : http://journals.openedition.org/regulation/7719 ; DOI : 10.4000/regulation.7719

Biraben J.-N. (2003). « L'évolution du nombre des hommes », Population & Sociétés, n°394, octobre 2003, INED

Blanc L. (2000). « Données spatio-temporelles en écologie et analyses multitableaux : examen d'une relation », Thèse de 3° cycle, Lyon I : Université Claude Bernard

Blanchet D., Simon O., Sylvander M. (2007). L'évolution des niveaux de vie, productivitéet bien-être en longue période : la France et les principauxpays développés, L'économie française, édition 2007, INSEE

Booth H., Maindonald J., Smith L. (2002). Applying Lee-Carter under conditions of variable mortality decline. Population Studies, 56, 325-336.

Boqueho V. (2012). « Les civilisations à l'épreuve du climat ». Dunod

Boserup E. (1965). The conditions of Agricultural Growth. The Economics of Agrarian Change under Population Pressure. London and New York 1965.

Bosserelle E. (2005). « Dynamique économique : Croissance, crises, cycles ». Gualino

Boubtane E., Coulibaly D., Rault C. (2013). « Immigration, unemployment and GDP in the host country : Bootstrap panel Granger causality analysis on OECD countries ». Documents de Travail du Centre d'Economie de la Sorbonne – 2013, 14

Boumard P. « Institution ». In: Recherche & Formation, N°23, 1996. Pratiques de formation initiale et continue desenseignants. pp. 151-161

Bourbonnais R., Terraza M. (2010). Analyse des séries temporelles, Dunod

Bourbonnais R., Usunier J. C. (2017). Prévision des ventes, Economica

Bourdin J. (2007). La productivité et le niveau de vie. Rapport d'information N°189, Sénat, France. Consulté en ligne le 04/11/2020, https://www.senat.fr/rap/r06-189/r06-1891.pdf

Box G. E. P., Tiao G. C. (1975). Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association, Vol. 70, No. 349. (Mar., 1975), pp. 70-79

Box George E.-P., Jenkins G.-M., Reinsel G.-C. (2008). « Time Series Analysis : Forecasting and Control». Wiley-Blackwell

Bresson G., Pirotte A. (1995). « Econométrie des séries temporelles : Théorie et applications ». PUF

Brondex F. (1999). « Evolution : Synthèse des faits et théorie - Cours et questions de révision». Dunod

Bruegel (2013), « Manufacturing Europe's future », BRUEGEL BLUEPRINT SERIES, Volume XXI, Reinhilde Veugelers (editor)

Cairns A. J.G., Blake D., Dowd K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and Calibration, J. Risk Insur., 73, 4, 687-718

Cairns A. J.G., Blake D., Dowd K. (2008). Modelling and management of mortality risk: a review, Scand. Actuar. J., 2008, 2-3, 79-113, (2008) · Zbl 1224.91048

Capelle-Blancard G. et Coulibaly D. (2011). « Index Trading and Agricultural Commodity Prices : A Panel Granger Causality Analysis ». CEPII, WP No 2011 – 28

Carré J. J., Dubois P., Malinvaud E. (1984). « Abrégé de la croissance française ». Seuil

Casin P., Stachowiak C., Marque F. (2011). « L'analyse en composantes principales de variables non stationnaires ». Math. Sci. hum / Mathematics and Social Sciences, 49e année, n° 196, 2011 (4), 27-40

Casper, G., & Wilson, M. (2015). Using Sequences to Model Crises. Political Science Research and Methods, 3(2), 381-397. https://doi.org/10.1017/psrm.2014.27

Castello-Climent, A. and Domenech, R., (2012), Human Capital and Income Inequality: Some Facts and Some Puzzles, No 1201, Working Papers, International Economics Institute, University of Valencia

Cauvin J. (2010). « Naissance des divinités, naissance de l'agriculture : La révolution des symboles au Néolithique ». CNRS

Cazes P., Lise-Ceremade (2004). « Quelques méthodes d'analyse factorielle d'une série de tableaux de données ». Revue MODULAD, Numéro 31

Central Intelligence Agency (2015a). "GDP - Composition, By sector of origin", The World Factbook. [en ligne] (page consultée le 15/10/2015). https://www.cia.gov/library/publications/the-world-factbook/fields/2012.html

Central Intelligence Agency (2015b). “Labor force - by occupation”, The World Factbook. [en ligne] (page consultée le 15/10/2015). https://www.cia.gov/library/publications/the-world-factbook/fields/2048.html

Cette G., Strauss-Kahn M.-O. (2003). Productivité horaire et PIB par tête aux États-Unis eten France - Comparaisons et recommandations. Bulletin de la Banque de France – N° 120 – Décembre 2003

Champault P. (1910). « Les types familiaux », dans les colonnes de La Science Sociale, Décembre 1910

Chauvel L. (2011). « Le Destin des générations : Structure sociale et cohortes en France du XXe siècle aux années 2010 ». PUF

Chesnais J-C. (1986), « La transition démographique ». Paris : PUF-INED (Cahier n°113)

Chevrier S. (2013). « Le management interculturel ». PUF

Clark, C., (1940). The Conditions of Economic Progress. London: Macmillan. Retrieved 2019-07-13.

Claudia Goldin (2014), Human Capital, Handbook of Cliometrics, Claude Diebolt and Michael Haupert, editors, Springer-Verlag

Cliometrics and complexity CAC (2020). The team CAC, hosted by Complex Systems Institute in Rhône-Alpes IXXI – ENS Lyon, [On-line], (page accessed on 02/03/2020), http://www.ixxi.fr/recherche/cliometrie-et-complexite?set_language=en&cl=en

Coase R. (1960), “The Problem of Social Cost”, Journal of Law and Economics, 3(1), 1-44

Cohen E., Buigues P.-A. (2014), Le décrochage industriel, Fayard, 438 pages

Cohen J. (1988), «Statistical power analysis for the behavioral sciences ». 2nd Ed. Lawrence Erlbaum Associates

Colletis G. (2012). « L'urgence industrielle ! ». Le Bord de l'eau

Commission européenne (2011), « Nouveau «tableau de bord de l’Union de l’innovation»: l’UE est dépassée par ses principaux concurrents malgré des progrès dans de nombreux États membres », IP/11/114, Bruxelles, le 1er février 2011, https://europa.eu/rapid/press-release_IP-11-114_fr.htm?locale=FR , consulté le 21/08/2019

Commission européenne (2014), « Tableau de bord de l’Union de l’innovation 2014 : Le tableau de bord de l’Union de l’innovation pour la recherche et l’innovation », Entreprises et industrie

Conrad S., 2017. What Is Global History ?, Princeton University Press

Cornwell, B. (2018). Network Analysis of Sequence Structures. In G. Ritschard & M. Studer (Éds.), Sequence Analysis and Related Approaches: Innovative Methods and Applications (p. 103-120). Springer International Publishing. https://doi.org/10.1007/978-3-319-95420-2_7

Courbage Y., Todd E. (2007). « Le rendez-vous des civilisations ». Seuil

Cronk L. (Sous la direction de). Chagnon N. (Sous la direction de). Irons W. (Sous la direction de) (2002). « Adaptation and Human Behavior : An Anthropological Perspective (Evolutionary Foundations of Human Behavior) »

Dahlman C. J. (1979), « The Problem of Externality », The Journal of Law and Economics, vol. 22, n° 1, avril 1979, p. 148 cité par Coase R., op. cit. p. 23.

Danchin E., Giraldeau L. A., Cézilly F. (2005). « Ecologie comportementale : Cours et sujets corrigés ». Dunod

Das K., Quirk T., (2016). Which Institutions Promote Growth? Revisiting the Evidence. Economic Papers, 2016, vol. 35, issue 1, 37-58

Daudin G. (2007). Histoire Économique. Dictionnaire de l’économie, Albin Michel, pp.1-10, 2007

Dawkins R. (1999). « L'Horloger aveugle ». Robert Laffont

Dawkins R. (2003). « Le Gène égoïste ». Odile Jacob

Dawkins R. (2011). « Il était une fois nos ancêtres : Une histoire de l'évolution ». Fayard/Pluriel

Dazy F., Lavallard F., Saporta, G. (2000). « L'analyse des données évolutives : Méthodes et applications ». Technip

de Rosnay J. (1977). « Les Origines de la vie : de l'atome à la cellule ». Seuil.

de Rosnay J. (1988). « L'aventure du vivant ». Seuil.

Deubel P. (2008). Analyse economique et historique des sociétés contemporaines, Pearson Education France

Diard M. C., Grelet-Terriou C. (2005). « La nouvelle comptabilité nationale ». Vuibert

d'Iribarne Ph., Henry A., Segal J.-P., Chevrier S., Globokar T. (2002). « Cultures et mondialisation : Gérer par-delà les frontières ». Seuil

Diebolt C., Parent A. (2011). Essais Cliometriques : Capital Humain, Monnaie Et Finance, Croissance Et Cycles. Editions Peter Lang, Frankfurt am Main

Douglass N. C. (1994), “Economic Performance Through Time”, The American economic review, june 1994, 84:3, p. 359-368.

Doyle, S. ; Meier zu Selhausen, F. ; Weisdorf, J. (2019), The Blessings of Medicine? Patient Characteristics and Health Outcomes in a Ugandan Mission Hospital, 1908–1970, Social History of Medicine, 2019/03/15

Droesbeke J. J., Saporta G. (2010). « Analyse statistique des données longitudinales ». Technip

Drummond V. (2014). « Le management interculturel ». Gereso

Durand M.-F., Copinschi P., Martin B., Placidi-Frot D., (2008). Atlas de la mondialisation 2008, Presses de Sciences Po

Durkheim E. (1894), Les Règles de la méthode sociologique, Paris, Payot, coll. "Petite Bibliothèque Payot", 2009

Ebrahim-zadeh C. (2003). Dutch Disease: Too much wealth managed unwisely. Finance and Development, IMF, March 2003, Volume 40, Number 1

Emmenegger, P., Häusermann, S., Palier, B., & Seeleib-Kaiser, M. (2013). Structural change and the politics of dualization. Rassegna Italiana di Sociologia, 2/2013. https://doi.org/10.1423/74142

European Commission (2013), “Comparing the innovation performance of EU Member States”, MEMO, Brussels, 26 March 2013, https://europa.eu/rapid/press-release_IP-13-270_fr.htm , consulté le 21/08/2019

Ferrari, G., & Pailhé, A. (2017). Transition to adulthood in France: Do children of immigrants differ from natives? Advances in Life Course Research, 31, 34-56. https://doi.org/10.1016/j.alcr.2016.10.001

Fiévet G. (1992). De la stratégie militaire à la stratégie d'entreprise, InterEditions

Fisher, A., (1939). "Production, Primary, Secondary And Tertiary," The Economic Record, The Economic Society of Australia, vol. 15(1), pages 24-38, June.

Fishlow A., Gwin C., Haggard A., Rodrik D., Wade R. (1994). Miracle or Design ? Lessons From The East Asian Experience, Overseas Development Council - Washington DC, Policy Essay No. 11

Flatt T., Heyland A. (2011). « Mechanisms of Life History Evolution : The Genetics and Physiology of Life History Traits and Trade-Offs «. Oxford University Press

Flaux D. (2019), Classement de 188 pays : des économies les plus prédatrices aux plus redistributrices, Ecofin Hebdo, 21 août 2019

Fourastié J. (1949). « Le Grand espoir du XXe siècle ». (Editions 1949, 1963, 1989) : Gallimard

Fourastié J. (1965). « Les 40000 heures »

Fourastié J. (1975). « Pourquoi nous travaillons ». PUF

Fourastié J. (1977). « Pouvoir d'achat, prix et salaires ». Gallimard

Fourastié J. (1980). « Les Trente glorieuses ou la Révolution invisible de 1946 à 1975 »

Fourastié J. (1982). « La Civilisation de 2001 ». PUF

Fourastié J. et Fourastié J. (1986). « La Réalité économique ». Hachette Littérature

Fourastié J. et Fourastié J. (1987). « D'une France à une autre : Avant et après les Trente Glorieuses »

Fourastié J., Harouel J. L. (Sous la direction de) (2005). « Productivité et richesse des nations ». Gallimard

FranceGenWeb (2018), Base recoupements 1914-1918, Recoupements par régiments et par lieux 1914-1918, consulté en ligne le 01/10/2018, http://www.francegenweb.org/b1914-1918/

Fuller, S. (2015). Do Pathways Matter? Linking Early Immigrant Employment Sequences and Later Economic Outcomes: Evidence from Canada. International Migration Review, 49(2), 355-405. https://doi.org/10.1111/imre.12094

Gabadinho, A., & Ritschard, G. (2016). Analysing state sequences with probabilistic suffix trees: the PST R library. Journal of Statistical Software, 72(3), 1-39. https://doi.org/10.18637/jss.v072.i03

Gabadinho, A., Ritschard, G., Müller, N.S. & Studer, M. (2011), "Analyzing and visualizing state sequences in R with TraMineR", Journal of Statistical Software. Vol. 40(4), pp. 1-37.

Galor O. (2011). « Unified Growth Theory ». Princeton University Press

Gallup J. L., Sachs J. D., Mellinger A. D. (1998), 'Geography and economic development', NBER Working Paper No. 6849.

Gershenkron A. (1962), Economic Backwardness in Historical Perspective: A Book of Essays, Belknap Press of Harvard University Press, 1962.

Gilles H., Guironnet J.P., Parent A. (2014a): « Géographie économique des morts de 14-18 en France », Revue économique, 2014/3 Vol. 65, p. 519-532

Gilles H., Guironnet J.P., Parent A. (2014b): « La géographie des Morts pour la France », Journal Le Monde, Supplément « Le journal du centenaire » du mardi 13 mai, double page

Giraud R., Chaix N. (1994). « Econométrie ». Presses Universitaires de France - PUF

Goldfarb B., Pardoux C. (2001). « Étude de données multidimensionnelles évolutives et comparaison de codages par l'analyse factorielle multiple ». Revue de Statistique Appliquée, 49 (1), 97-117

Goldin C. (2016), Human Capital. In: Handbook of Cliometrics. Heidelberg, Germany: Springer Verlag

Goldin, C., and Lawrence F. K. (2008). The Race between Education and Technology. Cambridge, MA: Belknap Press.

Gonzalez-Feliu J., Ovtracht N., Parent, A. (2018). Gare et Front: la logique de répartition des troupes dans le plan XVII de Joffre.

Gonzalez-Feliu J., Parent, A. (2016). "Clio-Combinatorics: A Novel Framework to Analyze Military Logistics Choices Using Operations Research Techniques", Handbook of Research on Military, Aeronautical, and Maritime Logistics and Operations, p. 79-101

Gourieroux C., Monfort A. (1999). « Séries temporelles et modèles dynamiques ». Economica

Grataloup C. (2010). « Géohistoire de la mondialisation : Le temps long du monde ». Armand Colin

Griffith R., Redding S. et Van Rennan J. (2004). "Mapping the Two Faces of R&D : Productivity Growth in a Panel of OECD Industries", Review of Economies and Statistics, 86, 883-895.

Guillot O., Parent A., (2018), “Farewell life, farewell love’: Analysis of survival inequalities among soldiers who “Died for France” during World War I”, Population, 73(3), 411–444.

Guironnet J.P., Parent, A. (2018), “Morts pour la France: Demographic or Economic Factors?”, Defence and Peace Economics

Hafer R. W. and Jones G. (2015), Are entrepreneurship and cognitive skills related? Some international evidence, Small Business Economics, Vol. 44, No. 2 (February 2015), pp. 283-298

Hall E.T. (1971), La Dimension cachée [« The Hidden Dimension »], Paris, Points, 1971, 1978 (1966 en anglais),

Hampden-Turner C. et Trompenaars A. (1991), The Seven Cultures of Capitalism, New York, Doubleday, 1993.

Hampden-Turner Ch., Trompenaars A., (1993). « The Seven Cultures of Capitalism: Value Systems for Creating Wealth in the United States, Japan, Germany, France, Britain, Sweden, and the Netherlan ». First Edition First Printing

Hanushek E.A. (2016), Will more higher education improve economic growth?, Oxford Review of Economic Policy, Volume 32, Number 4, 2016, pp. 538–552

Hanushek E.A., Ruhose J. and Woessmann L. (2017). Economic Gains from Educational Reform by US States, Journal of Human Capital, 2017, vol. 11, no. 4, University of Chicago

Hanushek E.A., Woessmann L. (2008), The Role of Cognitive Skills in Economic Development, Journal of Economic Literature 2008, 46:3, 607–668

Hanushek E.A., Woessmann L. (2012), Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation, Journal of Economic Growth, December 2012, Volume 17, Issue 4, pp 267–321

Hanushek E.A., Woessmann L. (2017), "School Resources and Student Achievement: A Review of Cross-Country Economic Research", Cognitive Abilities and Educational Outcomes, pp. 149-171, M. Rosén et al. (eds.), Springer International Publishing AG 2017

Hearns, Baten et Crayen (2009), Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital, The Journal of Economic History, Vol. 69, No. 3

Helske S., Helske J. (2019). Mixture Hidden Markov Models for Sequence Data: The seqHMM Package in R, Journal of statistical software, january 2019

Heyer E., Plane M., Timbeau X. (2007). Les Français sont-ils toujours productifs ?, E-book de campagne à l'usage des citoyens, OFCE, pp.12-15, 2007.

Hofstede G., Minkov M. (2010). « Cultures et organisations ». Pearson

Human Mortality Database (2020), consulté en ligne le 28/09/2020, https://www.mortality.org/cgi-bin/hmd/hmd_download.php

Hyndman R. J. (2019). demography: Forecasting Mortality, Migration, Fertility and Population Data. (R package version 1.21).

Hyndman R. J., Kostenko A. V. (2007). "Minimum Sample Size requirements for Seasonal Forecasting Models," Foresight: The International Journal of Applied Forecasting, International Institute of Forecasters, issue 6, pages 12-15, Spring.

Hyndman R. J., Ullah S., (2007). Robust forecasting of mortality and fertility rates: A functional data approach, Comput. Statist. Data Anal., 51, 10, 4942-4956, (2007) · Zbl 1162.62434

Imbert J., Legohérel H. (2004). « Histoire de la vie économique ancienne, médiévale et moderne ». Cujas

Inglehart R. (1997). « Modernization and postmodernization : Cultural, economic, and political change in 43 societies », Princeton University Press.

Inglehart R. and Welzel C. (2019), Cultural map – World Value Survey wave 6 (2010-2014), http://www.worldvaluessurvey.org/WVSContents.jsp, consulté le 03/09/2019

Inglehart R., and Welzel C. (2005), Modernization, Cultural Change and Democracy New York, Cambridge University Press, 2005

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin & B. Puranen et al. (eds.), (2014). World Values Survey: Round Six - Country-Pooled Datafile Version: http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp. Madrid: JD Systems Institute.

INSEE (1969), Les huit zones d'études et d'aménagement du territoire. In: Economie et statistique, n°5, Octobre 1969. pp. 45-48.

INSEE (2020), Zone d'études et d'aménagement du territoire ZEAT, consulté en ligne le 16/11/2020, https://www.insee.fr/fr/metadonnees/definition/c1910

Jablonka E., Lamb M.- J. (2006). « Evolution in Four Dimensions : Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life (Life and Mind : Philosophical Issues in Biology and Psychology) »

Jessua C. (1991), Histoire de la théorie économique, Presses Universitaires de France

"Johnston, L. D. (2012), ""Distribution of the output amoung sectors, 1840-2010"", MinnPost. [On-line], (page accessed on 02/03/2017), https://www.minnpost.com/macro-micro-minnesota/2012/02/history-lessons-understanding-decline-manufacturing#sourcenote

"

Kang, T. (2019). The Transition to Adulthood of Contemporary Delinquent Adolescents. Journal of Developmental and Life-Course Criminology, 5(2), 176-202. https://doi.org/10.1007/s40865-019-00115-6

Kennett D. J. (Sous la direction de). Winterhalder B. (Sous la direction de) (2006). « Behavioral Ecology And the Transition to Agriculture ». University of California Press

Kluckhohn F.L., Strodtbeck F.R. (1961). « Variations in Value Orientations ». Row, Peterson and Company

Konya L. (2005). « Saving and Growth : Granger Causality Analysis with Bootstrapping On Panels of Countries ». Journal of Economic Research, 10 (2005), 231–260

Laborit H. (1968). « Biologie et structure ». Gallimard.

Laborit H. (1974). « La nouvelle grille ». Robert Laffont.

Laborit H. (1980). « L'inhibition de l'Action : biologie comportementale et de physio-pathologie ». Masson & Presses Universitaires de Montréal.

Laborit H. (1997). « La légende des comportements ». Flammarion.

Landry A. (1934). « La Révolution démographique ». Paris : Librairie du Recueil Sirey (Ed. INED, Paris, 1982)

Le Bras H. (1994). « Les limites de la planète : Mythe de la nature et de la population ». Ed. Flamarion

Le Bras H. (2003), « Malthus ou Boserup : validité et continuité historique des modèles démo-économiques », Mathématiques et sciences humaines, no 164, Hiver 2003,

Le Guen M. (2002). La boîte a moustaches pour sensibiliser a la statistique. Bulletin de Methodologie Sociologique / Bulletin of Sociological Methodology, SAGE Publications, 2002, pp.43-64.

Le Meur, N., Vigneau, C., Lefort, M., Lebbah, S., Jais, J.-P., Daugas, E., & Bayat, S. (2019). Categorical state sequence analysis and regression tree to identify determinants of care trajectory in chronic disease: Example of end-stage renal disease. Statistical Methods in Medical Research, 28(6), 1731-1740. https://doi.org/10.1177/0962280218774811

Le Play F. (1871). « L'Organisation de La Famille »

Le Roy Ladurie E. (1997). « L'Historien, le chiffre et le texte «. – avril 1997

Le Roy Ladurie E. (2004). « Histoire humaine et comparée du climat : Tome 1, Canicules et glaciers XIIIe-XVIIIe siècles »

Le Roy Ladurie E. (2006). « Histoire humaine et comparée du climat : Tome 2, Disettes et révolutions (1740-1860) »

Le Roy Ladurie E. (2009). « Histoire du climat depuis l'an mil ». Tome 1, Tome 2, Tome 3, Flammarion

Le Roy Ladurie E., Séchet G. (2009). « Histoire humaine et comparée du climat : Tome 3, Le réchauffement de 1860 à nos jours »

Lebart L., Piron M., Morineau A. (2000). « Statistique exploratoire multidimensionnelle ». Dunod

Lee J-W, Lee H, (2016), "Human Capital in the Long Run," Journal of Development Economics, vol. 122, pp. 147-169.

Lee R.D., Carter L.R. (1992), Modeling and Forecasting U.S. Mortality, Journal of the American Statistical Association, Vol. 87, No. 419 (Sep., 1992), pp. 659-671

Lewis B., Jurmain R., Kilgore L. (2012). « Understanding Humans: An Introduction to Physical Anthropology and Archaeology »

List F. (1841), Troisième réédition en français : « Système national d'économie politique », Gallimard, 1998

Losos J. B. (Sous la direction de). Ricklefs R.-E. (Sous la direction de) (2009). « The Theory of Island Biogeography Revisited ». Princeton University Press

Lucas, R. (1988), « On the Mechanisms of Economic Growth », Journal of MonetaryEconomics, Vol.22, N°. 1, pp. 3-42.

Lynn R., Vanhanen T. (2002), IQ and the wealth of nations, Praeger/Greenwood

Lynn R., Vanhanen T. (2006), IQ and Global inequality, Washington Summit, Atlanta, Georgia

Lynn R., Vanhanen T. (2012), Intelligence: an unifying factor for the social science, Ulster Institute for Social Research, London

MacArthur R. H., Wilson E. O. (2001). « The Theory of Island Biogeography : Princeton Landmarks in Biology ». Princeton University Press

Maddison A. (2001). « L'économie mondiale. Une perspective millénaire ». OCDE

Maheu J. M., McCurdy T. H. (2009). How Useful Are Historical Data for Forecasting the Long-Run Equity Return Distribution?, Journal of Business & Economic Statistics, Vol. 27, No. 1 (Jan., 2009), pp. 95-112

Mankiw N.G., Romer D. et Weil D.N., (1992), « A contribution to the empirics of economic growth", Quarterly Journal of Economics, vol.107-2

Manyika J. et al., (2012), « Manufacturing the Future ; The Next Area of Global Growth and Innovation», McKinsey Global Institute

Matteudi E. (1997), Structures familiales et développement local, Editions L'Harmattan

Matteudi E. (2011), Du nord au sud. Parcours et réflexions sur le développement local et la lutte contre la pauvreté. Géographie. Université de Provence - Aix-Marseille I, 2011.

Maurau G. (2002). « Démographie économique ». Bréal

Max Planck Institute for Demographic Research (2020), Demographic Data, consulté en ligne le 28/09/2020, https://www.demogr.mpg.de/en/research_6120/demographic_data_27/

Mazerolle F. (2005). « Démographie économique ». Vuibert

McMillan M., Rodrik D. (2014). Globalization, Structural Change, and Productivity Growth, with an Update on Africa. World DevelopmentVol. 63, pp. 11–32, 2014

McVicar D., Anyadike-Danes M. (2002). Predicting successful and unsuccessful transitions from school to work by using sequence methods. Journal of the Royal Statistical Society. Series A (Statistics in Society) 165(2), 317-334.

Ménesguen A., Laurec A. (1984). « Analyse statistique de l'évolution temporelle de paramètres hydrobiologiques dans les bassins du port du Havre ». Oceanologica Acta 1984, 7 (1)

Michel Albert (1991), Capitalisme contre Capitalisme, Paris, Seuil, 1991.

Milanović B. (2019), Inégalités mondiales. Le destin des classes moyennes, les ultra-riches et l'égalité des chances, Paris, La Découverte

Millossovich P., Villegas A.M., Kaishev V. K. (2018). StMoMo: An R Package for Stochastic Mortality Modelling. Journal of Statistical Software, 84(3), doi:10.18637/jss.v084.i03

Mincer, J. (1970), “The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach.” Journal of Economic Literature, 8(1): 1–26.

Mincer, J. (1974), "Schooling, Experience, and Earnings". New York: National Bureau of Economic Research.

Mincer, J. (1996). "Changes in Wage Inequality, 1970-1990," NBER Working Papers 5823, National Bureau of Economic Research, Inc.

Mincer, J. (1997), The Production of Human Capital and the Life Cycle of Earnings: Variations on a Theme, Journal of Labor Economics, 15, (1), S26-47

Mincer, J. 1991. "Education and Unemployment," NBER Working Papers 3838, National Bureau of Economic Research, Inc.

Ministère des Armées (2018). Mémoire des hommes, Morts pour la France de la Première Guerre mondiale, consulté en ligne le 09/05/2018, https://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?larub=24&titre=morts-pour-la-france-de-la-premiere-guerre-mondiale

Müller N, Studer M, Ritschard G, Gabadinho A (2010). Extraction de règles d'association séquentielle à l'aide de modèles semi-paramétriques à risques proportionnels. Revue des nouvelles technologies de l'information (RNTI), E-19, 25-36.

Naizot T., Auda Y., Dervieux A., Thioulouse J. et Bellan M. F. (2004). « Une nouvelle analyse multi-temporelle d’images satellitales, les résidus de l’Analyse en Composantes Principales. Un cas

d'étude : une série d'images Landsat Thematic Mapper de la Camargue, France », INT. Remote Sensing J., 25, (10) 1925–1938

North D. C., (1996). "Economic Performance Through Time: The Limits to Knowledge," Economic History 9612004, University Library of Munich, Germany.

North, D.C. (1990), Institutions, Institutional Change and Economic Performance, Cambridge: Cambridge University Press.

Notestein F. (1945). « Population : the Long View ». In: Food for the World (ed. By T.W Schultz). Chicago: Univ. of Chicago Press

OECD Programme for International Student Assessment (2019), Upcoming - Results from PISA 2018, page internet : http://www.oecd.org/pisa/, consultée le 29/09/2019

Parent A. (2017). Introduction to the Special Issue on colonial institutions and African development. Journal of Institutional Economics: pp. 1-9, Millennium Economics Ltd 2017 doi:10.1017/S1744137417000510

Parente S. L., Prescott E. C., (2002). Les richesses défendues. Conférences Walras - Pareto. Editions Payot Lausanne

Paulus F. (2004). « Coévolution dans les systèmes de villes : croissance et spécialisation des aires urbaines françaises de 1950 à 2000 ». Thèse de doctorat en Géographie, Paris 1 : Université Panthéon-Sorbonne

Perhoniemi, R., Blomgren, J., & Laaksonen, M. (2019). Sources of income following a rejected disability pension application: a sequence analysis study. Disability and Rehabilitation, 0(0), 1-9. https://doi.org/10.1080/09638288.2018.1555619

Petri P.A. (1993), The lessons of East Asia : common foundations of East Asian success, The World Bank, WVashington, D.C.

Piersma T., van Gils J.-A. (2010). « The Flexible Phenotype : A Body-Centred Integration of Ecology, Physiology, and Behaviour ». Oxford Biology

Piriou P., Bournay J. (2012). « La comptabilité nationale ». La Découverte

Planchet F., Lelieur V., (2007). Utilisation des méthodes de Lee-Carter et Log-Poisson pour l'ajustement des tables de mortalité dans le cas de petits échantillons, Bulletin français d'actuariat, Vol. 7, n°14, juillet-décembre 2007, pp. 118-146

Poingt G. (2018), « La France peut-elle protéger ses secteurs stratégiques ? », Le Figaro, Publié le 10/07/2018 à 17:15, consulté en ligne le 27/0/2019, https://amp.lefigaro.fr/conjoncture/2018/07/10/20002-20180710ARTFIG00219-la-france-peut-elle-proteger-ses-secteurs-strategiques.php

Polity IV (Marshall et Jaggers, 2015). Publié par Université de Sherbrooke (Juin 2019) Disponible sur http://perspective.usherbrooke.ca/

Porter M. E. (1993), The Competitive Advantage of Nations, New York, The Free Press, 1990

Raab, M., & Struffolino, E. (2019). The Heterogeneity of Partnership Trajectories to Childlessness in Germany. European Journal of Population, 1–18.

Rezzouk M. (2011), "Secteurs stratégiques, Concept et politique : Une perspective économique", Document de travail, Centre européen de droit & d'économie, ESSEC Business School

Ritschard G., Gabadinho A., Muller N.S., Studer M. (2009), Données séquentielles: concepts et principe d'anaylse, Conference EGC 2009, Strasbourg, 27-30 janvier 2009

Ritschard, G., & Studer, M. (2018b). Sequence Analysis: Where Are We, Where Are We Going? (p. 1-11). https://doi.org/10.1007/978-3-319-95420-2_1

Ritschard, G., Bürgin, R. & Studer, M. (2013), "Exploratory Mining of Life Event Histories", In McArdle, J.J. & Ritschard, G. (eds) Contemporary Issues in Exploratory Data Mining in the Behavioral Sciences. Series: Quantitative Methodology, pp. 221-253. New York

Robette (2014). L'étude des parcours de vie, démarche descriptive ou causale ?, Idées économiques et sociales 2014/3 (N° 177), pages 8 à 15

Rodrik D. (1999). Institutions For High-Quality Growth : What They Are and How to Acquire Them. IMF Conference on Second Generation Reforms. [en ligne], (page consultée le 11/06/2020). https://www.imf.org/external/pubs/ft/seminar/1999/reforms/rodrik.htm#II

Rodrik D. (2001), Les mirages de l'ouverture extérieure, Altern. économiques , « L'Économie politique », 2001/2 no 10 , pages 44 à 54

Rodrik D. (2002). Le commerce des illusions, De Boeck Supérieur, Reflets et perspectives de la vie économique, 2002/2 Tome XLI, pages 41-51

Rodrik D., Subramanian A., (2003). La primauté des institutions (ce que cela veut dire et ce que cela ne veut pas dire), Finances & Développement, Juin 2003

Roff A.-D. (2001, 2002). « Life History Evolution ». Sinauer Associates Inc.,U.S

Romer, P.M. (1986), « Increasing Returns and Long-Run Growth », Journal of Political Economy, Vol.94, N°. 5, pp. 1002-1037.

Roser M. (2015) – 'Number of world citizens living under different political regimes'. Published online at OurWorldInData.org. Retrieved from:https://ourworldindata.org/democracy/ [Online Resource]. Polity IV dataset and I have added the information on colonial regimes based on Wimmer and Min

Roser M., Ritchie H. , Ortiz-Ospina E., (2019) - "World Population Growth". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/world-population-growth' [Online Resource], 2019/12/09

Rostow W.W. (1959). « The Stages of Economic Growth », The economic history review, New series, vol. 12, N°1, 1959.

Rostow W.W. (1997). « Les étapes de la croissance économique ». Economica, 3ème Edition 1997

Sabéran S. (2008), La notion d'intérêt général chez Adam Smith : de la richesse des nations à la puissance des nations, Géoéconomie 2008/2 (n° 45), pages 55 à 71

Safitri L., Mardiyati S., Rahim H. (2018). Estimation of mortality rate in Indonesia with Lee-Carter model, AIP Conference Proceedings 2023, 020210 (2018); https://doi.org/10.1063/1.5064207, Published Online: 23 October 2018

Saint-Etienne C. (2010). « Guerre et Paix au XXIe siècle ». François Bourin

Salmon F. (2001). « Atlas électoral de la France, 1848-2001 », Seuil

Salvà-Mut, F., Thomás-Vanrell, C., & Quintana-Murci, E. (2016). School-to-work transitions in times of crisis: the case of Spanish youth without qualifications. Journal of Youth Studies, 19(5), 593-611. https://doi.org/10.1080/13676261.2015.1098768

Saporta G. (2011). « Probabilités, Analyse des données et Statistique ». Technip

Saporta, G. (1975). « Liaisons entre plusieurs ensembles de variables et codage de données Qualitatives ». Thèse de 3° cycle, Paris VI : Université Pierre et Marie Curie.

Sauvy A. (1980). « La machine et le chômage : les progrès techniques et l'emploi ». Dunod/Bordas

Schoffeniels E. (1973). « L'anti-hasard ». Gauthier-Villars.

Schumpeter J. (1943). Traduction française 1951 « Capitalisme, socialisme et démocratie », Paris, Payot.

Sen A. (2003), Un nouveau modèle économique. Développement, justice, liberté, Odile Jacob, 2003

Sequence Analysis Association (2020). SAA Bibliography, Sequence analysis. Consulté en ligne le 22/11/2020,https://www.zotero.org/groups/2268769/saa_bibliography/collections/T33535YZ/items/CB7J48ZW/collection

Smith A. (1776). La richesse des nations, GF-Flammarion, 1991

Stearns S.-C. (1992). « The Evolution of Life Histories ». OUP Oxford

Stone L. (1964), «The Educational Revolution in England, 1560-1640 », Past and Present, juillet 1964.

Stone L. (1969), « Literacy and Education in England, 1640-1900 », Past and Present, février 1969.

Structures familiales et développement local, E Matteudi - 1997 - Editions L'Harmattan

Studer M. (2012), Étude des inégalités de genre en début de carrière académique à l'aide de méthodes innovatrices d'analyse de données séquentielles., Thèse de doctorat : Univ. Genève, 2012, no. SES 777

Studer M., Ritschard G., Gabadinho A., Müller N. S. (2011). Discrepancy analysis of state sequences. Sociological Methods and Research 40 (3), 471–510.

Studer, M., Ritschard, G., Gabadinho, A. & Müller, N.S. (2011), "Discrepancy Analysis of State Sequences", Sociological Methods and Research. Vol. 40(3), pp. 471-510.

Taylor F. W. (1967). « La direction scientifique des entreprises ». Dunod

Teixeira, A. and Queirós, A, (2016). "Economic growth, human capital and structural change: A dynamic panel data analysis," Research Policy, Elsevier, vol. 45(8), pages 1636-1648.

Testot L. et al. (2015). Histoire globale, Un autre regard sur le monde, Sciences Humaines Éditions

The World Bank (2019), "The World Bank In Small States", page internet : https://www.worldbank.org/en/country/smallstates/overview, consulté le 12/10/2019

Thioulouse J. et Chessel D. (1987). « Les analyses multitableaux en écologie factorielle ». Acta Oecologica, Oecologia Generalis, 8 (4), 463-480.

Thompson W. (1929). « Population ». Amer J Sociology 34(6): 959-975

Todd E. (1976). « La chute finale : Essai sur la décomposition de la sphère soviétique », Robert Laffont

Todd E. (1979). Le fou et le prolétaire, Robert Laffont

Todd E. (1983). « La Troisième planète : Structures familiales et systèmes idéologiques ». Seuil

Todd E. (1984). « L'enfance du monde : Structures familiales et développement ». Seuil

Todd E. (1996). « L'invention de l'Europe ». Seuil

Todd E. (1998). L'Illusion économique : Essai sur la stagnation des sociétés développées, Paris, Gallimard

Todd E. (1999). « La Diversité du monde : Famille et modernité ». Seuil

Todd E., Le Bras H. (2012). L'invention de la France. Atlas anthropologique et politique, Gallimard, coll. «NRF Essais»

Todd E., Le Bras H. (2013). Le Mystère français, Seuil

Trompenaars F., Hampden-Turner C. (2008). « L'entreprise multiculturelle ». Maxima Laurent du Mesnil

Tufféry S. (2012). « Data Mining et Statistique Décisionnelle ». Technip.

Ubertalli B., Permin M. O. (1990). « Utilisation d'une méthode multi-tableaux en sciences sociales. Une étude longitudinale de carrières : les 12 premières promotions de l'école d'infirmières de Roanne ». In : Population, 45e année, (6), 1092-1100.

Vidal R., Ma Y., Sastry S., (2016), Generalized Principal Component Analysis. New York, Springer-Verlag

Weber M. (1921), Économie et société, tome 1 - Les catégories de la sociologie (1921), p. 94 pour l'édition Pocket de 1995

Weinhold D. (1996). « Tests de causalité sur données de panel : une application à l'étude de la causalité entre l'investissement et la croissance ». In : Économie & prévision. Numéro 126, 1996-5. Analyse des comportements économiques à partir de données de panel, 163-175.

Wilson, R. A. and Briscoe, G. (2004) The impact of human capital on economic growth: a review Impact of education and training, Cedefop Reference Series.

Winterhalder B. (Sous la direction de). Smith E. A. (Sous la direction de) (1992) « Evolutionary Ecology and Human Behavior : Foundations of Human Behavior ».

World Bank (2015), “World Development Indicators: Contribution of natural resources to gross domestic product”, [en ligne] (page consultée le 15/10/2015). http://wdi.worldbank.org/table/3.15

World Bank (2017), World Development Indicators, [en ligne] (page consultée le 22/11/2017), https://databank.worldbank.org/source/world-development-indicators

World Bank (2018), “Doing Business 2019: Training for Reform. Region Pro le of OECD High Income”, A World Bank Group Flagship Report, 16Th edition

World Bank (2018), Development Research Group, disponible sur http://data.worldbank.org/data-catalog/world-development-indicators

World Bank (2020), Gapminder, consulté en ligne le 13/03/2020, https://www.gapminder.org/tools/#$chart-type=linechart

Zalc C., Lemercier C. (2008). « Méthodes quantitatives pour l'historien ». La Découverte

Zhelyazkova, N., & Ritschard, G. (2018). Parental Leave Take-Up of Fathers in Luxembourg. Population Research and Policy Review, 1-25. https://doi.org/10.1007/s11113-018-9470-8

Annexes

1. Annexes

1.1. Mesures de la qualité de la partition selon le nombre de classes

	1.Ngroups	1.stat	2.Ngroups	2.stat	3.Ngroups	3.stat
PBC	4	8.731755e-01	5	8.518742e-01	6	8.515732e-01
HG	30	9.999948e-01	29	9.999914e-01	28	9.999880e-01
HGSD	30	9.999733e-01	29	9.999572e-01	28	9.999420e-01
ASW	30	9.217391e-01	29	9.086957e-01	28	9.014493e-01
ASWw	30	9.608696e-01	29	9.521739e-01	28	9.473430e-01
CH	30	1.087563e+03	29	8.988022e+02	28	7.782234e+02
R2	30	9.973122e-01	29	9.965944e-01	28	9.958766e-01
CHsq	30	2.299537e+04	29	1.901138e+04	28	1.646889e+04
R2sq	30	9.998726e-01	29	9.998385e-01	28	9.998044e-01
HC	30	5.932289e-07	28	7.871918e-07	29	7.890772e-07
	4.Ngroups	**4.stat**	**5.Ngroups**	**5.stat**	**6.Ngroups**	**6.stat**
PBC	8	8.490462e-01	7	8.486501e-01	9	7.886325e-01
HG	27	9.999331e-01	26	9.999046e-01	25	9.998177e-01
HGSD	27	9.998845e-01	26	9.998527e-01	25	9.997323e-01
ASW	27	8.884058e-01	25	8.838528e-01	26	8.833333e-01
ASWw	27	9.386473e-01	26	9.314010e-01	25	9.298304e-01
CH	27	6.946890e+02	26	6.163829e+02	25	5.186745e+02
R2	27	9.951515e-01	26	9.942575e-01	25	9.928219e-01
CHsq	27	1.468688e+04	26	1.161038e+04	25	9.019304e+03
R2sq	27	9.997696e-01	26	9.996935e-01	25	9.995844e-01
HC	27	1.870816e-06	26	5.457371e-05	25	1.600220e-04
	7.Ngroups	**7.stat**	**8.Ngroups**	**8.stat**	**9.Ngroups**	**9.stat**
PBC	10	7.531739e-01	11	7.525391e-01	12	7.510521e-01
HG	24	9.997745e-01	23	9.996987e-01	22	9.995057e-01
HGSD	24	9.996839e-01	23	9.995875e-01	22	9.993897e-01
ASW	24	8.767663e-01	22	8.733974e-01	23	8.718388e-01
ASWw	24	9.212547e-01	22	9.121473e-01	23	9.119068e-01
CH	24	4.604782e+02	23	4.197276e+02	22	3.903834e+02
R2	24	9.914810e-01	23	9.901351e-01	22	9.887831e-01
CHsq	24	7.464042e+03	23	6.480911e+03	22	5.662724e+03
R2sq	24	9.994702e-01	23	9.993552e-01	22	9.992186e-01
HC	24	2.634458e-04	23	3.670579e-04	22	5.752149e-04
	10.Ngroups	**10.stat**	**11.Ngroups**	**11.stat**	**12.Ngroups**	
PBC	13	7.476275e-01	14	7.429638e-01	3	
HG	21	9.994111e-01	20	9.993310e-01	19	
HGSD	21	9.992678e-01	20	9.991830e-01	19	
ASW	20	8.699969e-01	17	8.666713e-01	21	
ASWw	20	9.006292e-01	21	8.978819e-01	19	
CH	21	3.615727e+02	20	3.500251e+02	19	
R2	21	9.871680e-01	20	9.859165e-01	19	
CHsq	21	5.106818e+03	20	4.660259e+03	19	
R2sq	21	9.990805e-01	20	9.989282e-01	19	
HC	21	7.266739e-04	20	8.779353e-04	19	
	12.stat	**13.Ngroups**	**13.stat**	**14.Ngroups**	**14.stat**	
PBC	7.215996e-01	15	7.060724e-01	16	7.058555e-01	
HG	9.991880e-01	18	9.987710e-01	8	9.986820e-01	
HGSD	9.990341e-01	8	9.986266e-01	18	9.985623e-01	
ASW	8.652393e-01	18	8.606889e-01	16	8.602531e-01	
ASWw	8.940557e-01	17	8.924425e-01	18	8.903666e-01	

	12.stat	13.Ngroups	13.stat	14.Ngroups	14.stat	
CH	3.232964e+02	18	3.033998e+02	17	2.686775e+02	
R2	9.837710e-01	18	9.815407e-01	17	9.777113e-01	
CHsq	3.869872e+03	18	3.353898e+03	17	2.979305e+03	
R2sq	9.986237e-01	18	9.983016e-01	17	9.979484e-01	
HC	1.080855e-03	8	1.395763e-03	18	1.628837e-03	
	15.Ngroups	15.stat	16.Ngroups	16.stat	17.Ngroups	
PBC	17	7.050453e-01	18	6.899798e-01	19	
HG	17	9.974658e-01	16	9.971829e-01	15	
HGSD	17	9.972678e-01	16	9.969796e-01	15	
ASW	19	8.592437e-01	15	8.413955e-01	8	
ASWw	16	8.863969e-01	15	8.769414e-01	14	
CH	16	2.577108e+02	15	2.508019e+02	14	
R2	16	9.750294e-01	15	9.723086e-01	14	
CHsq	16	2.473979e+03	15	2.148446e+03	14	
R2sq	16	9.973393e-01	15	9.966864e-01	14	
HC	17	2.707726e-03	16	2.989998e-03	15	
	17.stat	18.Ngroups	18.stat	19.Ngroups	19.stat	
PBC	6.863121e-01	20	6.856689e-01	21	6.845509e-01	
HG	9.969978e-01	14	9.959160e-01	13	9.954061e-01	
HGSD	9.967942e-01	14	9.953633e-01	13	9.948479e-01	
ASW	8.300734e-01	14	8.237298e-01	7	8.229885e-01	
ASWw	8.562664e-01	13	8.490916e-01	8	8.451538e-01	
CH	2.177087e+02	13	2.065472e+02	12	2.037978e+02	
R2	9.655433e-01	13	9.604738e-01	12	9.560726e-01	
CHsq	1.837238e+03	13	1.641301e+03	12	1.502723e+03	
R2sq	9.957891e-01	13	9.948479e-01	12	9.938075e-01	
HC	3.135612e-03	14	3.965830e-03	13	4.590846e-03	
	20.Ngroups	20.stat	21.Ngroups	21.stat	22.Ngroups	
PBC	22	6.821611e-01	23	6.806705e-01	24	
HG	12	9.939709e-01	7	9.931231e-01	11	
HGSD	12	9.934194e-01	7	9.930673e-01	11	
ASW	13	8.180467e-01	6	8.108047e-01	9	
ASWw	7	8.381321e-01	12	8.297699e-01	11	
CH	11	2.035635e+02	10	2.026976e+02	9	
R2	11	9.513936e-01	10	9.455755e-01	9	
CHsq	11	1.349872e+03	10	1.186420e+03	9	
R2sq	11	9.923545e-01	10	9.902623e-01	9	
HC	12	5.465419e-03	11	6.385563e-03	7	
	22.stat	23.Ngroups	23.stat	24.Ngroups	24.stat	
PBC	6.790743e-01	25	6.774757e-01	26	6.750248e-01	
HG	9.929682e-01	10	9.922696e-01	9	9.921590e-01	
HGSD	9.924217e-01	6	9.919910e-01	10	9.917255e-01	
ASW	8.064152e-01	12	8.058952e-01	11	8.044674e-01	
ASWw	8.281476e-01	9	8.264397e-01	6	8.222755e-01	
CH	1.925786e+02	6	1.894049e+02	7	1.833164e+02	
R2	9.356261e-01	8	9.212191e-01	7	9.105885e-01	
CHsq	1.106644e+03	8	1.045103e+03	7	8.701502e+02	
R2sq	9.881685e-01	8	9.855848e-01	7	9.797332e-01	
HC	7.090006e-03	10	7.223242e-03	6	8.005476e-03	

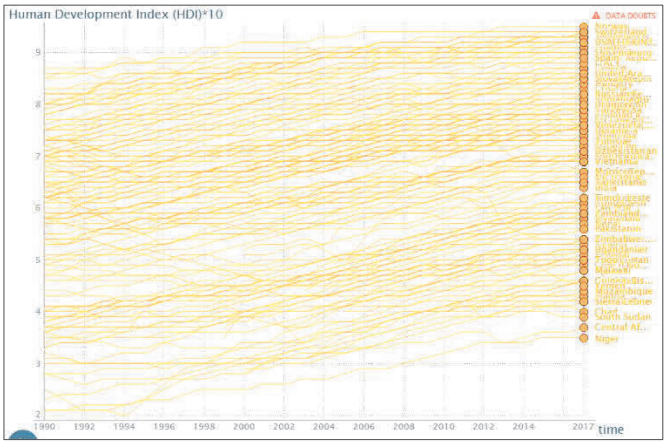
	25.Ngroups	25.stat	26Ngroups	26.stat	27.Ngroups	
PBC	27	6.742238e-01	28	0.67337197	29	
HG	6	9.920689e-01	5	0.97583515	4	
HGSD	9	9.916302e-01	5	0.97563613	4	
ASW	10	7.894070e-01	5	0.77620153	4	
ASWw	10	8.150309e-01	5	0.78562691	4	
CH	8	1.787424e+02	5	148.18435492	4	
R2	6	8.967827e-01	5	0.84346927	4	
CHsq	6	7.733188e+02	5	417.46090579	4	
R2sq	6	9.725827e-01	5	0.93819682	4	
HC	9	8.889776e-03	5	0.02189567	4	
	27.stat	28.Ngroups	28.stat	29.Ngroups	29.stat	
PBC	0.67251864	30	0.67166483	2	0.6064304	
HG	0.96941917	3	0.86813954	2	0.7460159	
HGSD	0.96888909	3	0.86787487	2	0.7457927	
ASW	0.74890009	3	0.63833824	2	0.4829338	
ASWw	0.75642694	3	0.64425386	2	0.4919162	
CH	131.64983127	3	73.31738290	2	54.2961510	
R2	0.78061051	3	0.56695690	2	0.3245511	
CHsq	343.81880586	3	108.97748553	2	71.3532064	
R2sq	0.90284093	3	0.66055974	2	0.3870462	
HC	0.02474494	3	0.09593861	2	0.1025364	
	30.Ngroups	30.stat				
PBC	NA	NA				
HG	NA	NA				
HGSD	NA	NA				
ASW	NA	NA				
ASWw	NA	NA				
CH	NA	NA				
R2	NA	NA				
CHsq	NA	NA				
R2sq	NA	NA				
HC	NA	NA				

1.2. Mesures de la qualité de partition selon le nombre de classes

	PBC	HG	HGSD	ASW	ASWw	CH	R2	CHsq	R2sq	HC
cluster2	0,60643	0,74602	0,74579	0,48293	0,49192	54,29615	0,32455	71,35321	0,38705	0,10254
cluster3	0,72160	0,86814	0,86787	0,63834	0,64425	73,31738	0,56696	108,97749	0,66056	0,09594
cluster4	0,87318	0,96942	0,96889	0,74890	0,75643	131,64983	0,78061	343,81881	0,90284	0,02474
cluster5	0,85187	0,97584	0,97564	0,77620	0,78563	148,18435	0,84347	417,46091	0,93820	0,02190
cluster6	0,85157	0,99207	0,99199	0,81080	0,82228	189,40495	0,89678	773,31875	0,97258	0,00801
cluster7	0,84865	0,99312	0,99307	0,82299	0,83813	183,31637	0,91059	870,15018	0,97973	0,00709
cluster8	0,84905	0,99868	0,99863	0,83007	0,84515	178,74236	0,92122	1045,10348	0,98558	0,00140
cluster9	0,78863	0,99216	0,99163	0,80642	0,82644	192,57861	0,93563	1106,64353	0,98817	0,00889
cluster10	0,75317	0,99227	0,99173	0,78941	0,81503	202,69765	0,94558	1186,42041	0,99026	0,00722
cluster11	0,75254	0,99297	0,99242	0,80447	0,82815	203,56351	0,95139	1349,87226	0,99235	0,00639
cluster12	0,75105	0,99397	0,99342	0,80590	0,82977	203,79777	0,95607	1502,72250	0,99381	0,00547
cluster13	0,74763	0,99541	0,99485	0,81805	0,84909	206,54721	0,96047	1641,30135	0,99485	0,00459
cluster14	0,74296	0,99592	0,99536	0,82373	0,85627	217,70873	0,96554	1837,23823	0,99579	0,00397
cluster15	0,70607	0,99700	0,99679	0,84140	0,87694	250,80188	0,97231	2148,44560	0,99669	0,00314
cluster16	0,70586	0,99718	0,99698	0,86025	0,88640	257,71078	0,97503	2473,97873	0,99734	0,00299
cluster17	0,70505	0,99747	0,99727	0,86667	0,89244	268,67749	0,97771	2979,30491	0,99795	0,00271
cluster18	0,68998	0,99877	0,99856	0,86069	0,89037	303,39985	0,98154	3353,89795	0,99830	0,00163
cluster19	0,68631	0,99919	0,99903	0,85924	0,89406	323,29638	0,98377	3869,87226	0,99862	0,00108
cluster20	0,68567	0,99933	0,99918	0,87000	0,90063	350,02515	0,98592	4660,25944	0,99893	0,00088
cluster21	0,68455	0,99941	0,99927	0,86524	0,89788	361,57267	0,98717	5106,81753	0,99908	0,00073
cluster22	0,68216	0,99951	0,99939	0,87340	0,91215	390,38339	0,98878	5662,72383	0,99922	0,00058
cluster23	0,68067	0,99970	0,99959	0,87184	0,91191	419,72758	0,99014	6480,91076	0,99936	0,00037
cluster24	0,67907	0,99977	0,99968	0,87677	0,92125	460,47817	0,99148	7464,04195	0,99947	0,00026
cluster25	0,67748	0,99982	0,99973	0,88385	0,92983	518,67450	0,99282	9019,30428	0,99958	0,00016
cluster26	0,67502	0,99990	0,99985	0,88333	0,93140	616,38291	0,99426	11610,37571	0,99969	0,00005
cluster27	0,67422	0,99993	0,99988	0,88841	0,93865	694,68903	0,99515	14686,88351	0,99977	0,00000
cluster28	0,67337	0,99999	0,99994	0,90145	0,94734	778,22345	0,99588	16468,89092	0,99980	7,87191769025649e-7
cluster29	0,67252	0,99999	0,99996	0,90870	0,95217	898,80223	0,99659	19011,38184	0,99984	7,89077207225427e-7
cluster30	0,67166	0,99999	0,99997	0,92174	0,96087	1087,56260	0,99731	22995,36789	0,99987	5,93228891484931e-7

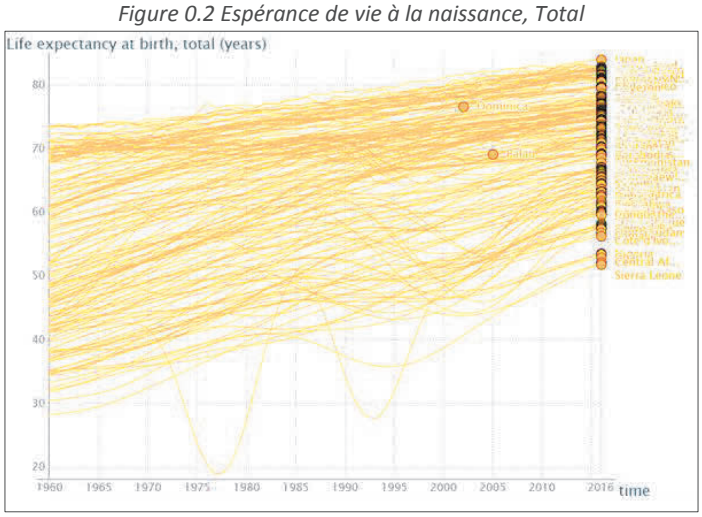
2. Courbes longitudinales des variables passives

*Figure 0.1 Indice de développement humain (HDI*10)*



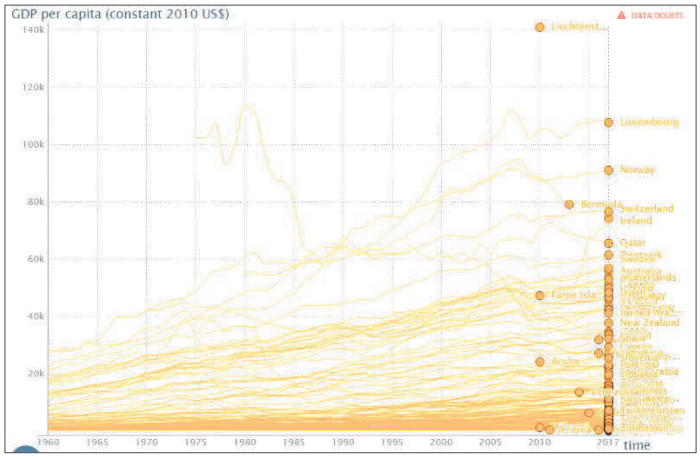
Source : Perspective Monde (2018)

Figure 0.2 Espérance de vie à la naissance, Total



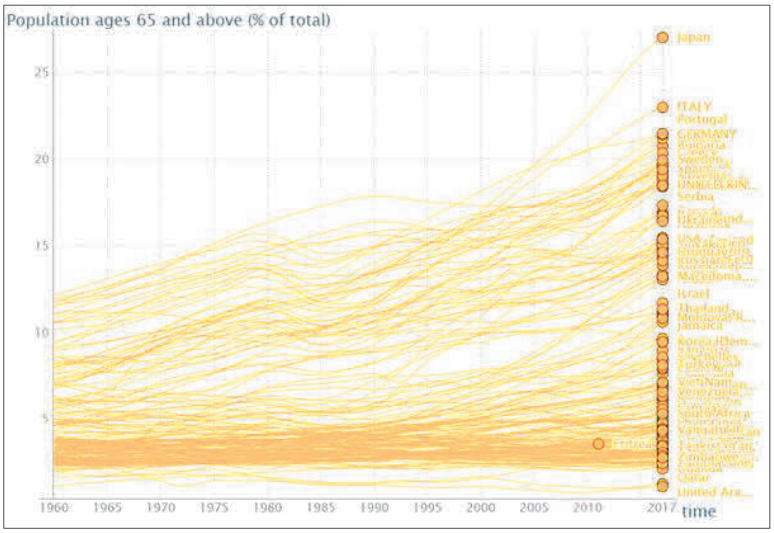
Source : World Bank (2018)

Figure 0.3 PIB par habitant (USD constants de 2010)



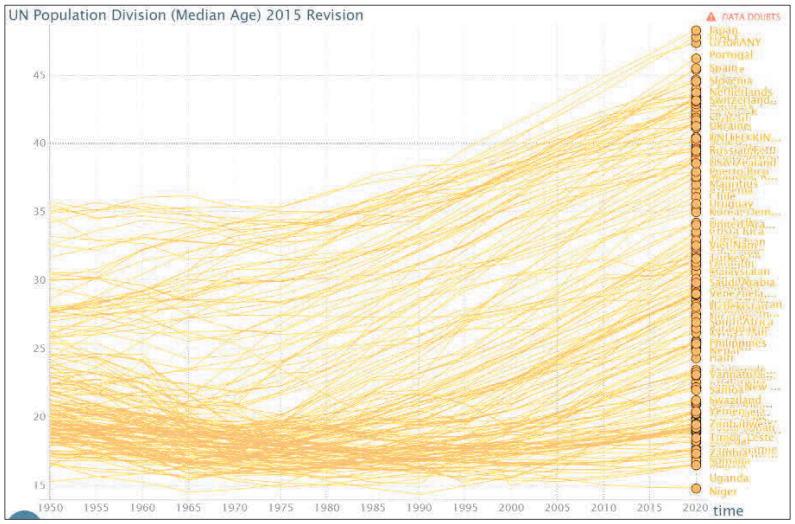
Source : World Bank (2018)

Figure 0.4 Population âgée de 65 ans et plus (% de la population totale)



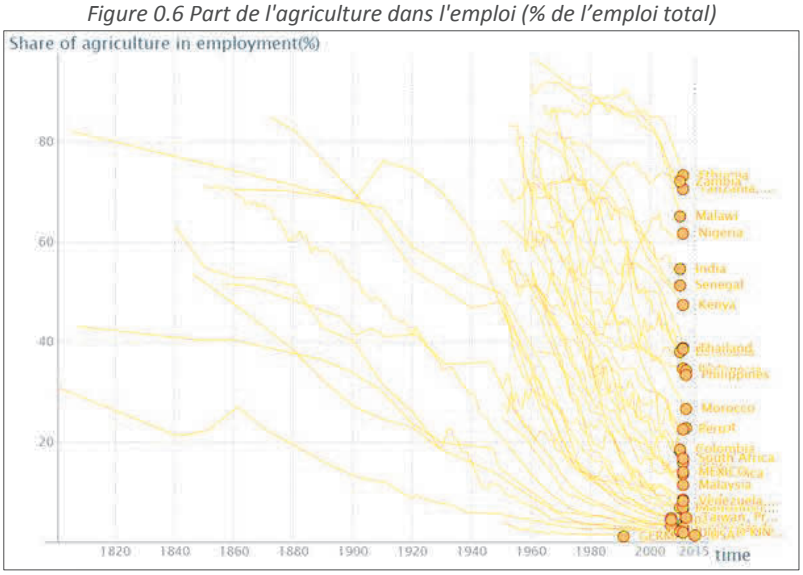
Source : World Bank (2018)

Figure 0.5 Division de la population des Nations Unies (âge médian)



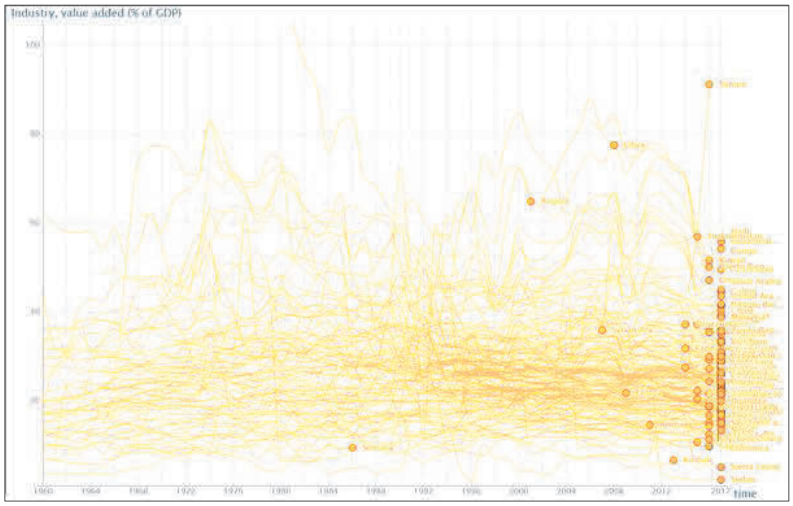
Source : Roser Max (2015)

Figure 0.6 Part de l'agriculture dans l'emploi (% de l'emploi total)



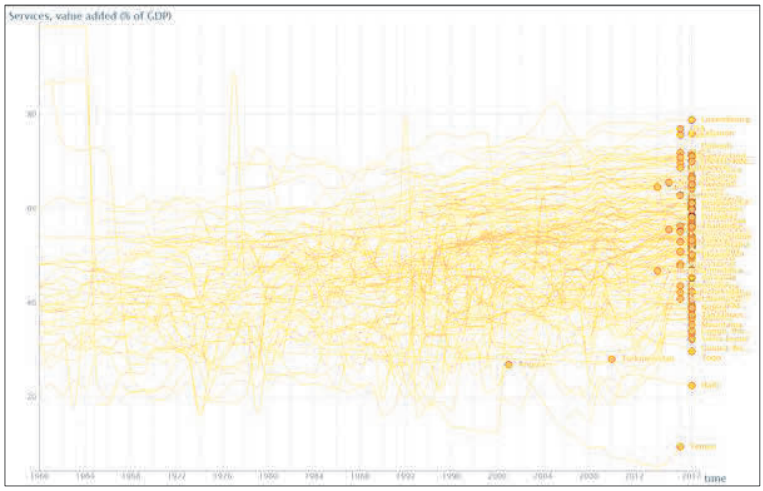
Source : Roser Max (2015)

Figure 0.7 Industrie, valeur ajoutée (% du PIB)



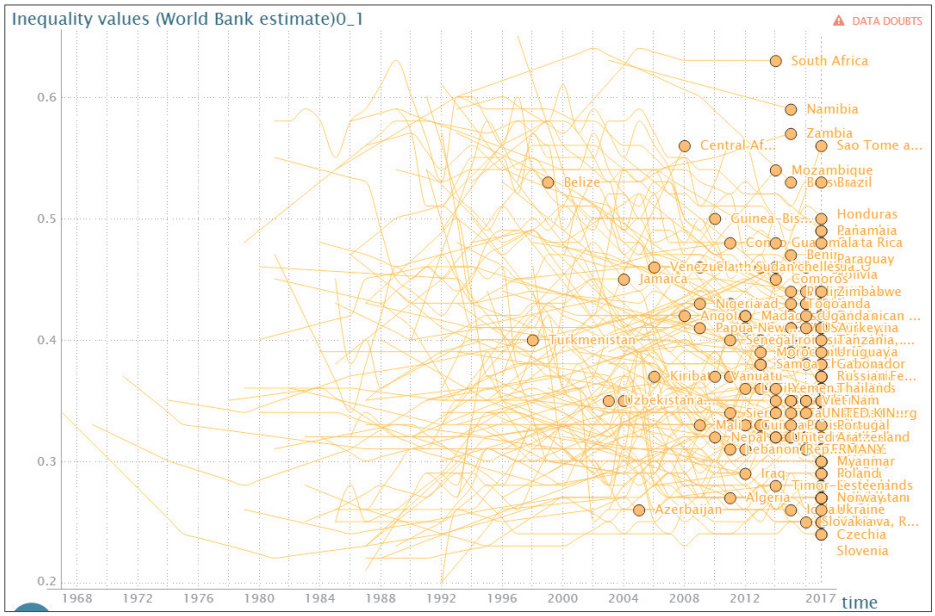
Source : World Bank (2018)

Figure 0.8 Services, valeur ajoutée (% du PIB)



Source : World Bank (2018)

Figure 0.9 Valeurs inégalitaires (estimation Banque mondiale)

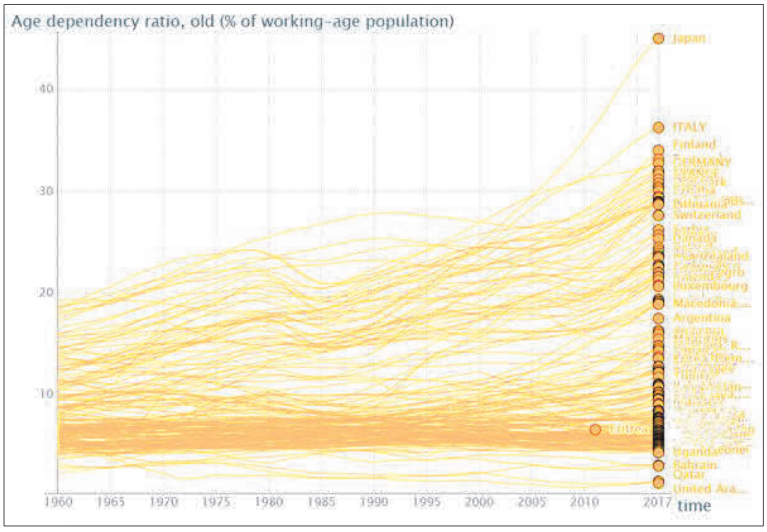


Source : World Bank (2018)

Figure 0.10 Ratio de dépendance à l'âge (% de la population en âge de travailler)

Source : World Bank (2018)

Figure 0.11 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler)

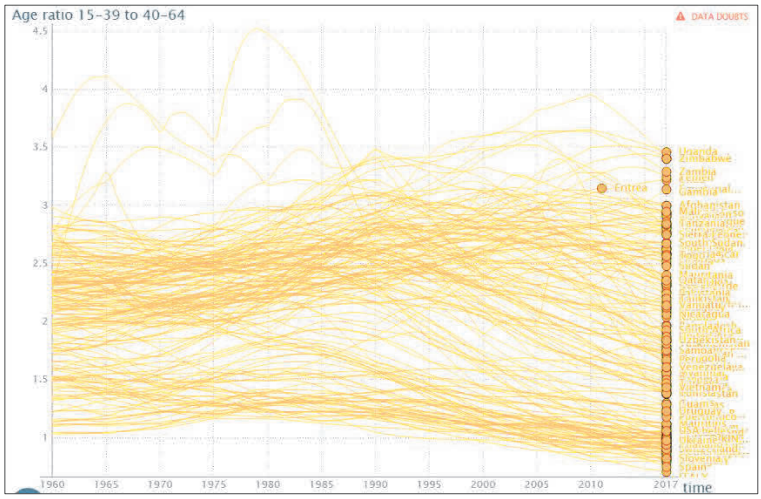


Source : World Bank (2018)

Figure 0.12 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler)

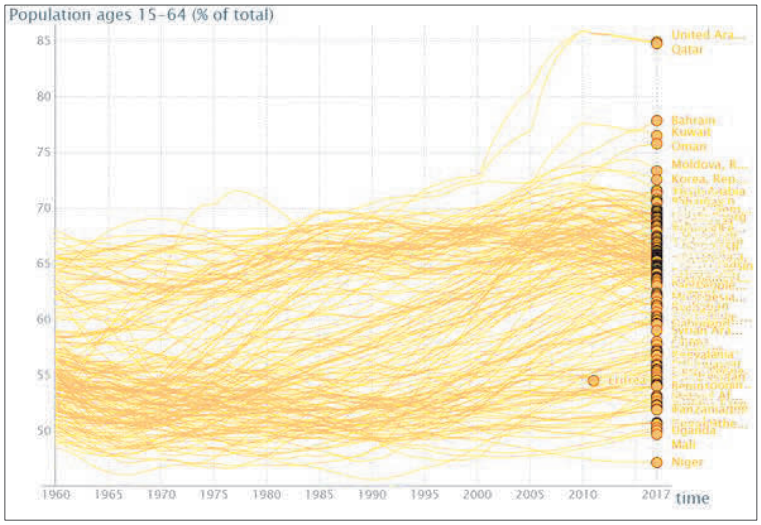
Source: World Bank (2018)

Figure 0.13 Ratio d'âge15-39 à 40-64



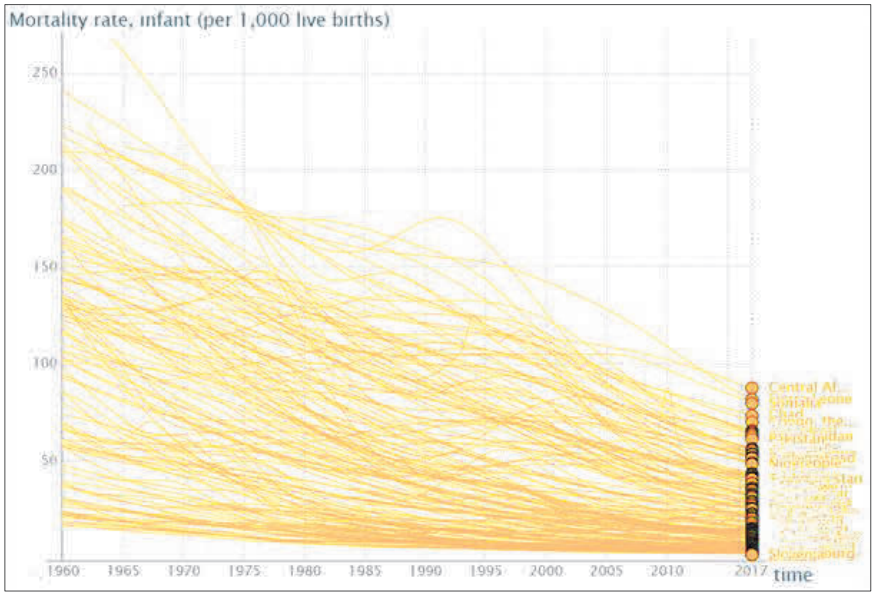
Source : World Bank (2018)

Figure 0.14 Population âgée de 15 à 64 ans (% du total)



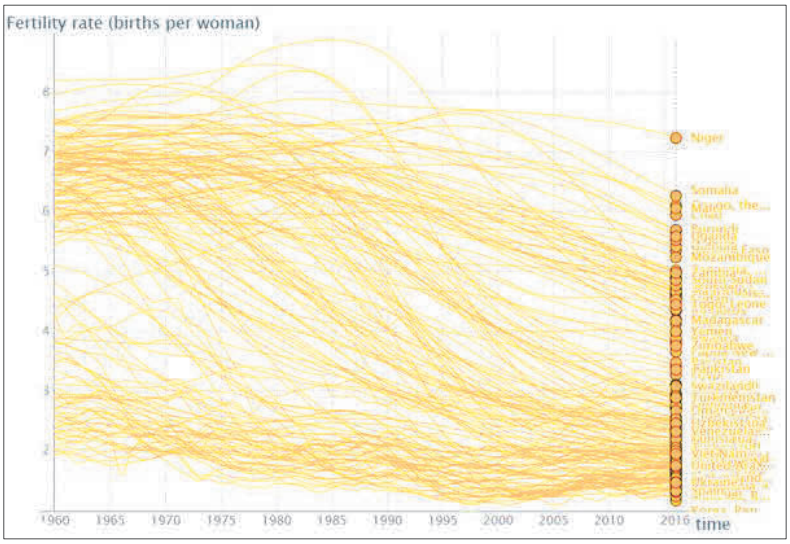
Source : World Bank (2018)

Figure 0.15 Taux de mortalité infantile (pour 1 000 naissances vivantes)



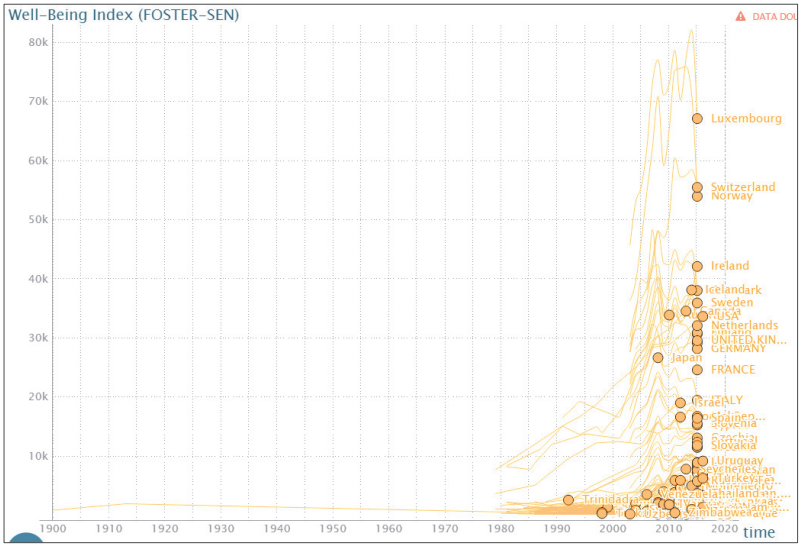
Source : World Bank (2018)

Figure 0.16 Taux de fécondité, total (naissances par femme)



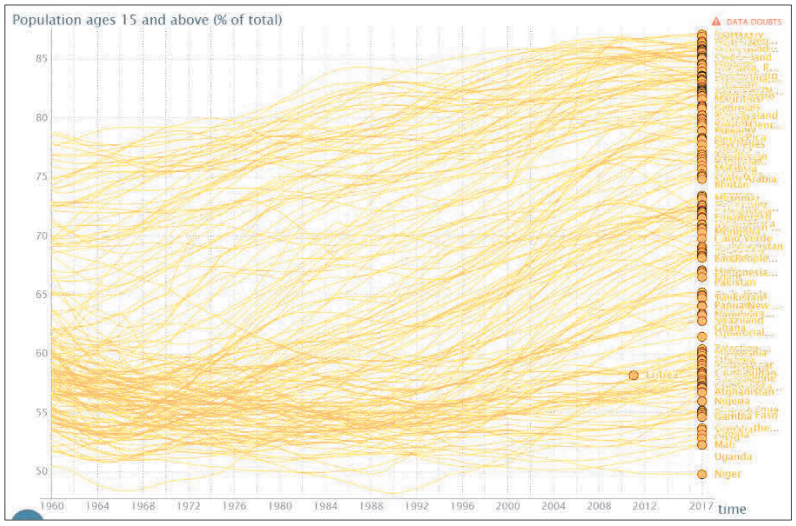
Source : Perspective Monde (2018)

Figure 0.17 Indice de bien-être (FOSTER-SEN)



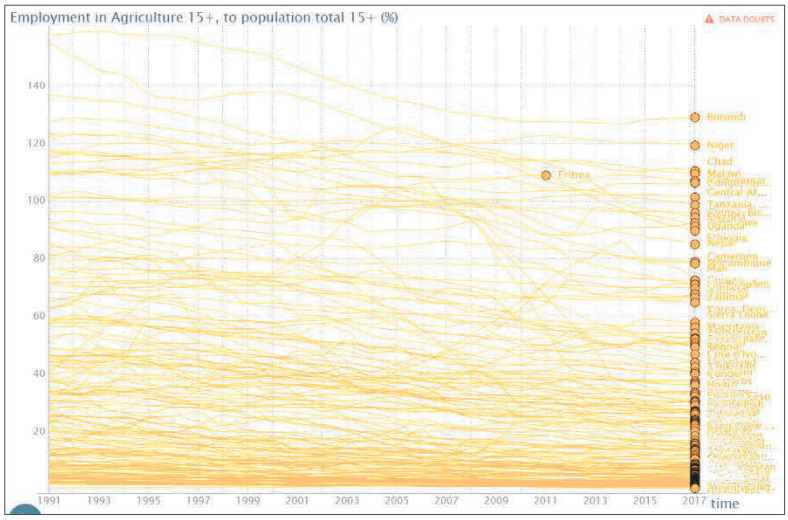
Source : World Bank (2018)

Figure 0.18 Population de 15 ans et plus (% du total)



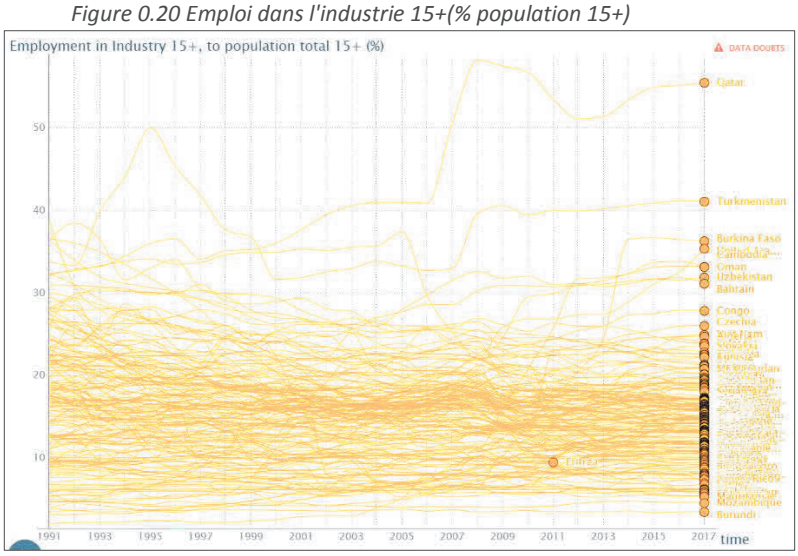
Source : World Bank (2018)

Figure 0.19 Emploi dans l'agriculture 15 ans et plus, (% population 15 ans et plus)



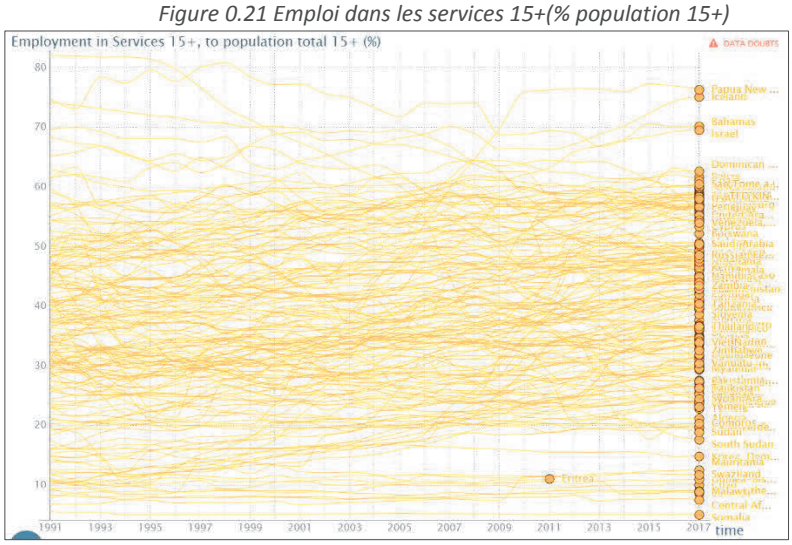
Source : World Bank (2018)

Figure 0.20 Emploi dans l'industrie 15+(% population 15+)



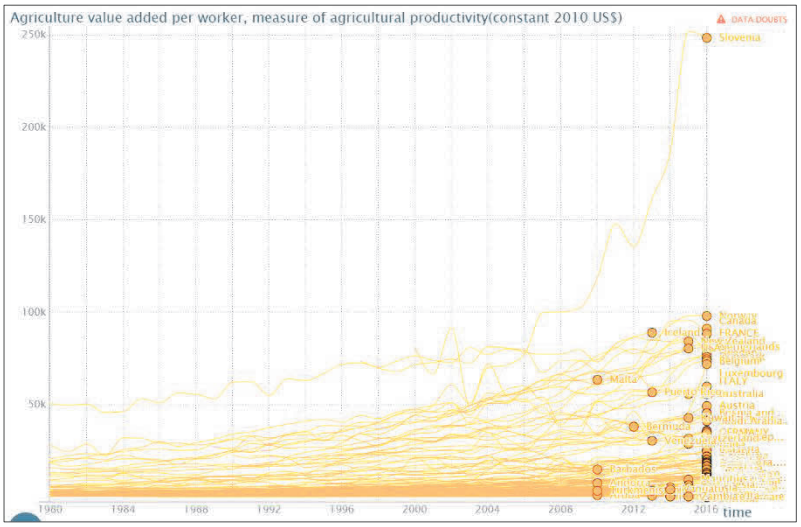
Source : World Bank (2018)

Figure 0.21 Emploi dans les services 15+(% population 15+)



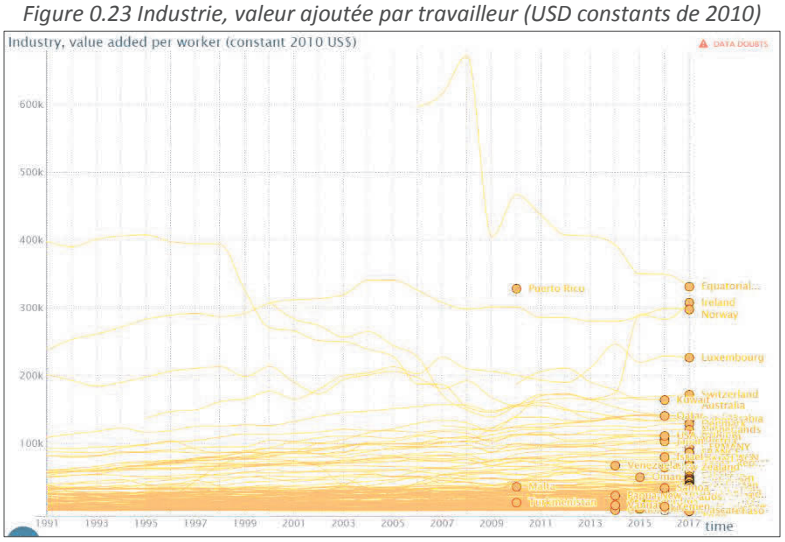
Source : World Bank (2018)

Figure 0.22 Valeur ajoutée agricole par travailleur, mesure de la productivité agricole ($ US constants de 2010)



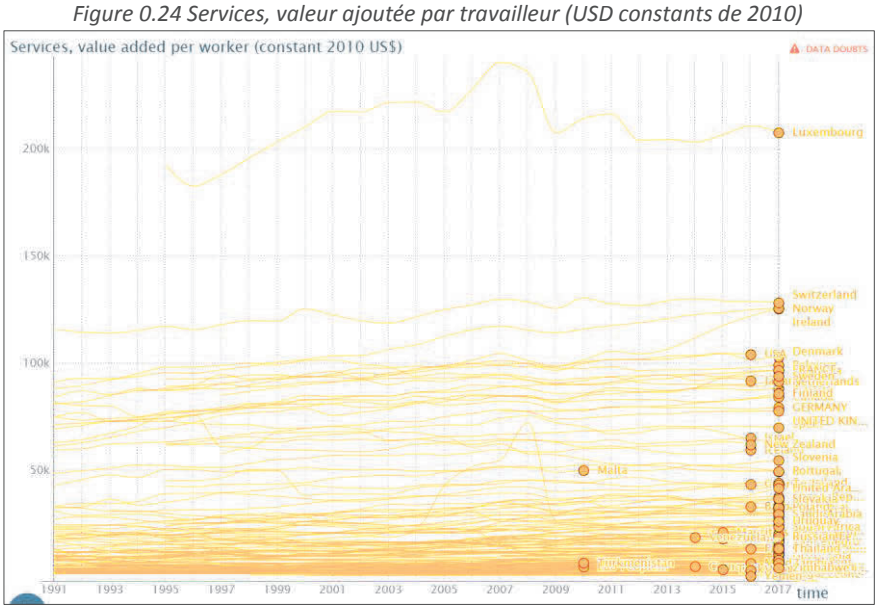
Source : World Bank (2018)

Figure 0.23 Industrie, valeur ajoutée par travailleur (USD constants de 2010)



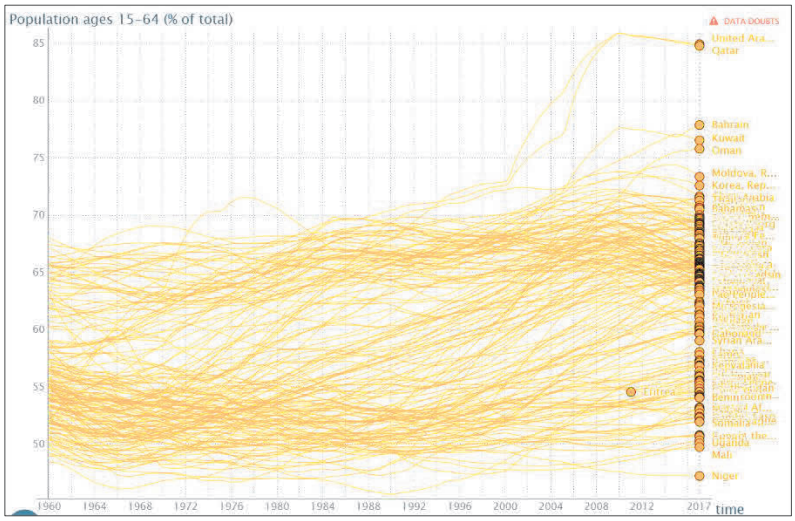
Source : World Bank (2018)

Figure 0.24 Services, valeur ajoutée par travailleur (USD constants de 2010)



Source : World Bank (2018)

Figure 0.25 Population âgée de 15 à 64 ans (% du total)

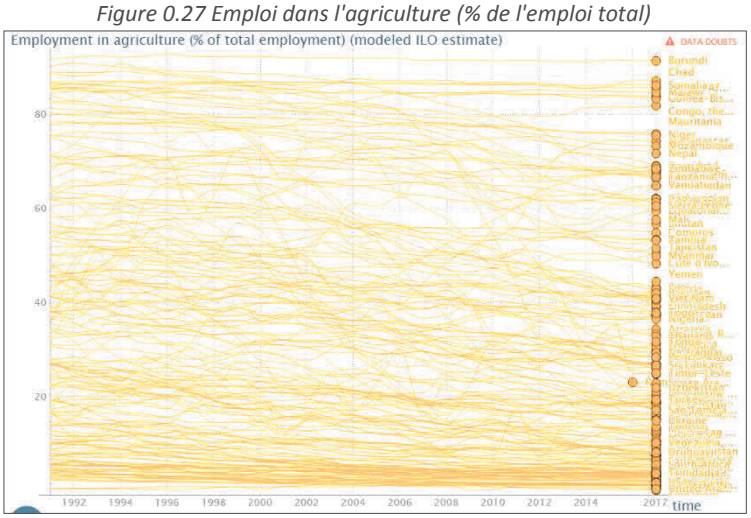


Source : World Bank (2018)

Figure 0.26 Taux de natalité, brut (pour 1 000 personnes)

Source : World Bank (2018)

Figure 0.27 Emploi dans l'agriculture (% de l'emploi total)

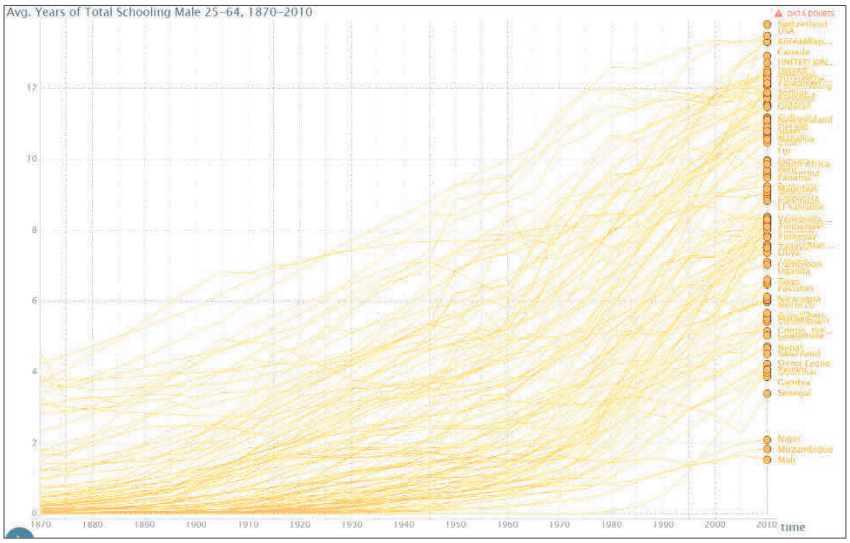


Source : World Bank (2018)

Figure 0.28 Emploi dans les services (% de l'emploi total)

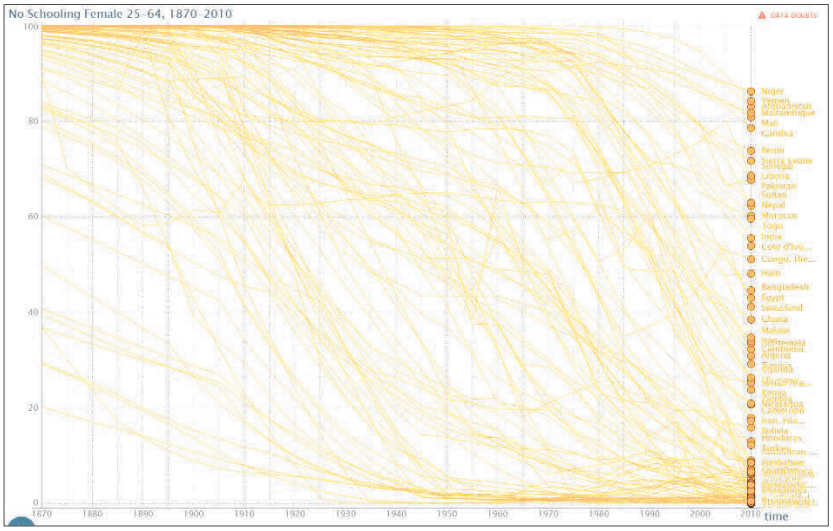
Source : World Bank (2018

Figure 0.29 Nombre moyen d'années de scolarisation totale des hommes de 25 à 64 ans



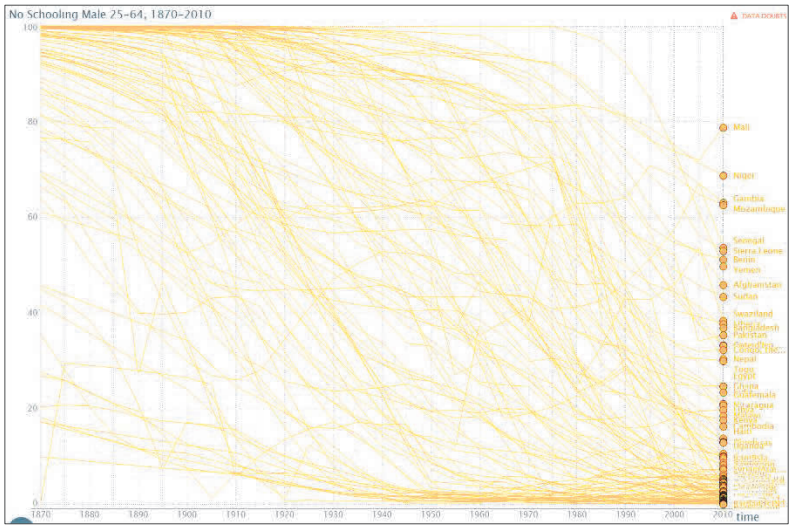
Source : Barro-Lee (2015)

Figure 0.30 Taux (%) des Femmes non scolarisées de 25 à 64 ans



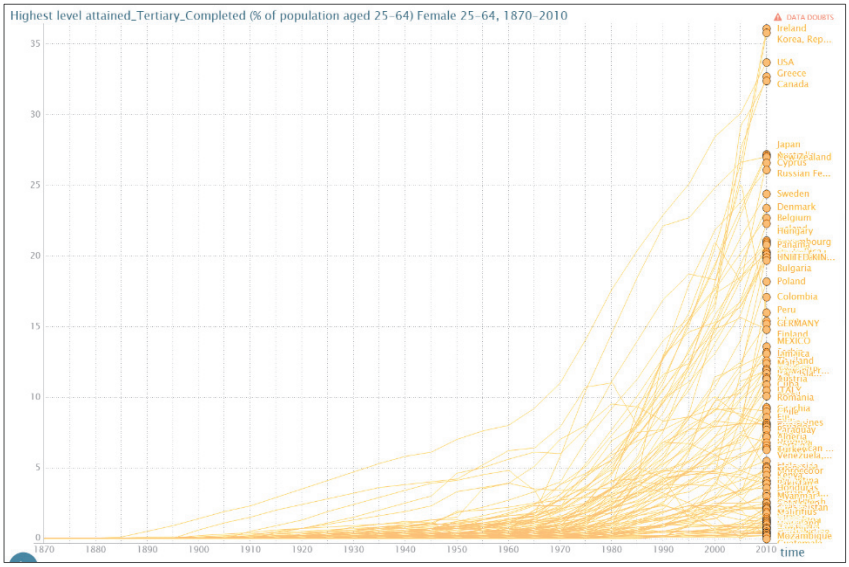
Source : Barro-Lee (2015)

Figure 0.31 Taux (%) des Hommes non scolarisés de 25 à 64 ans



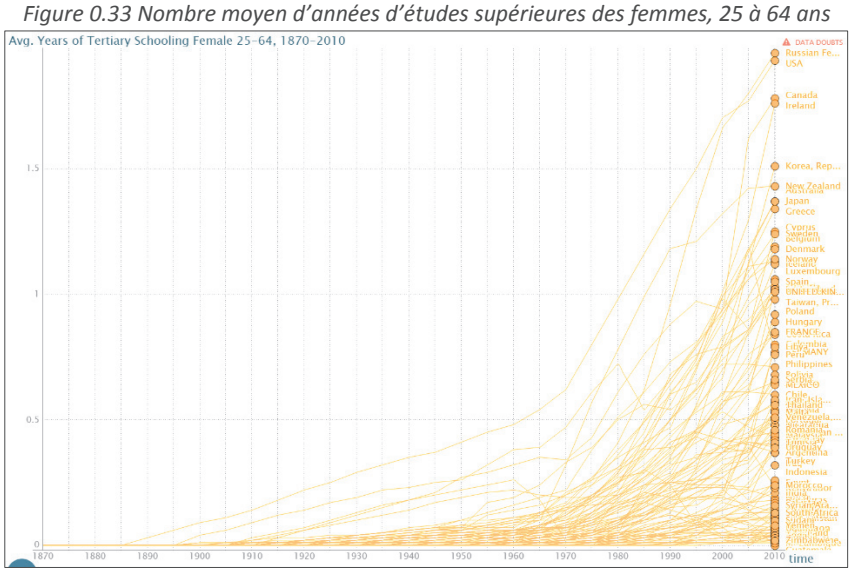
Source : Barro-Lee (2015)

Figure 0.32 Taux (%) des femmes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint



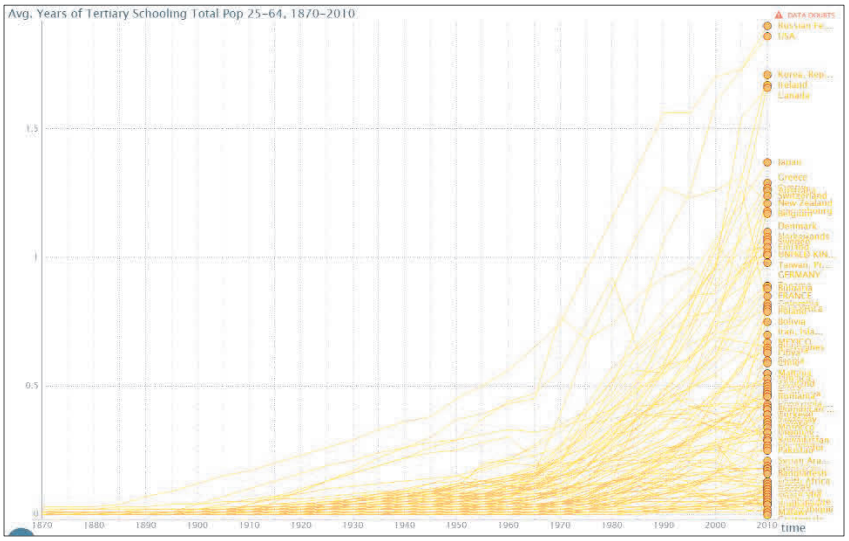
Source : Barro-Lee (2015)

Figure 0.33 Nombre moyen d'années d'études supérieures des femmes, 25 à 64 ans



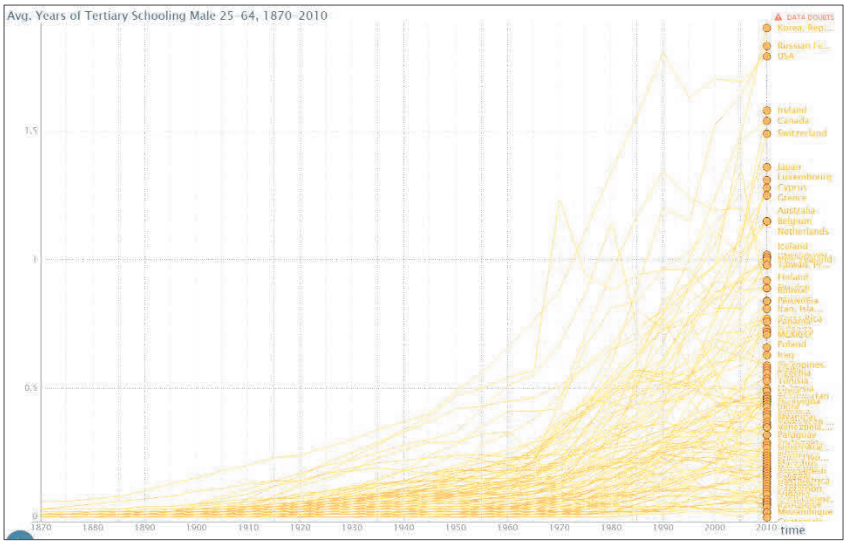
Source : Barro-Lee (2015)

Figure 0.34 Nombre moyen d'années d'études supérieures, population totale, 25 à 64 ans



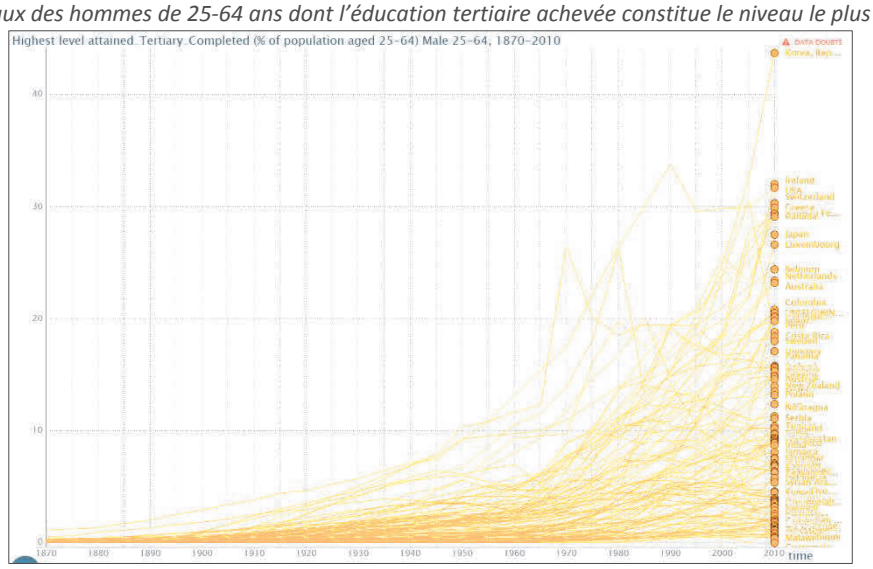
Source : Barro-Lee (2015)

Figure 0.35 Nombre moyen d'années d'études supérieures, des hommes, 25 à 64 ans



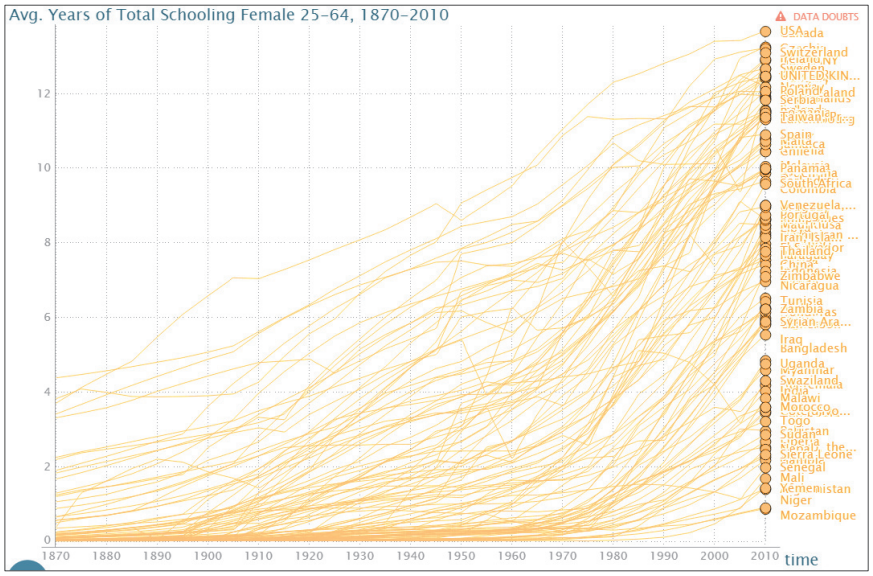
Source : Barro-Lee (2015)

Figure 0.36 Taux des hommes de 25-64 ans dont l'éducation tertiaire achevée constitue le niveau le plus élevé atteint



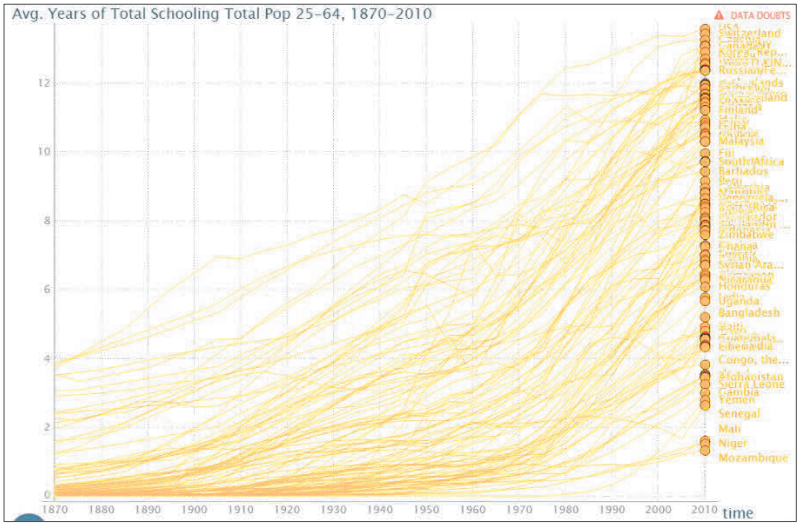
Source : Barro-Lee (2015)

Figure 0.37 Nombre moyen d'années de scolarisation des Femmes de 25 à 64 ans



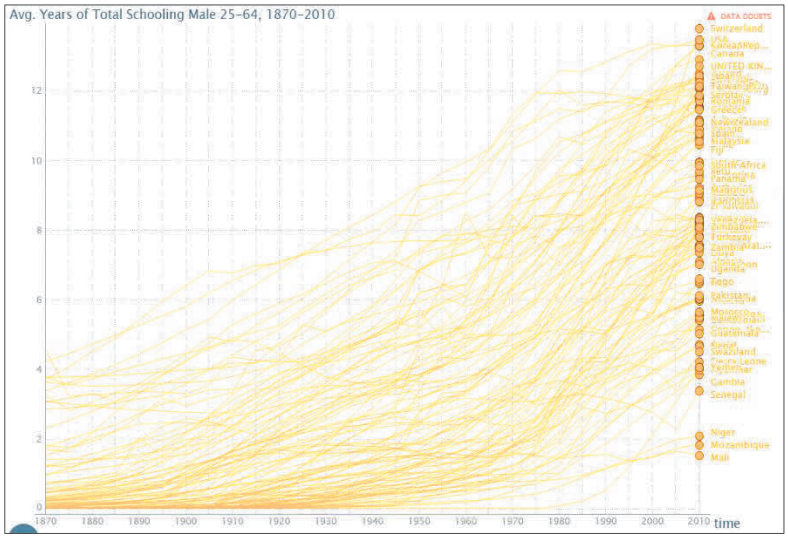
Source : Barro-Lee (2015)

Figure 0.38 Nombre moyen d'années de scolarisation (Population totale) 25 à 64 ans



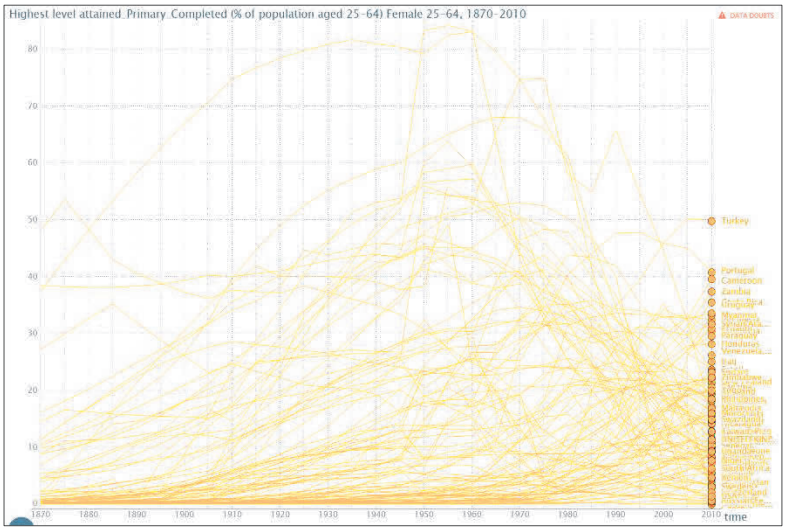
Source : Barro-Lee (2015)

Figure 0.39 Nombre moyen d'années de scolarisation des hommes, 25 à 64 ans



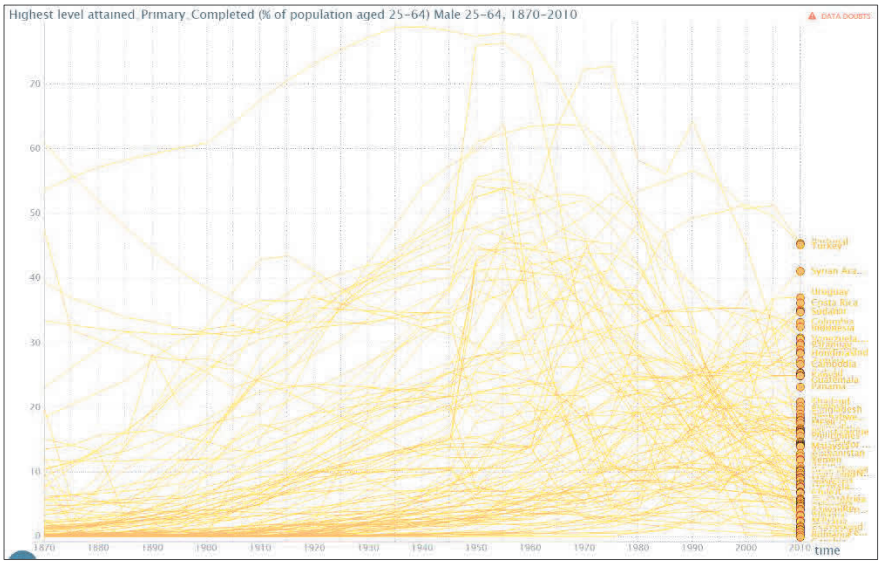
Source : Barro-Lee (2015)

Figure 0.40 Taux des femmes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint,



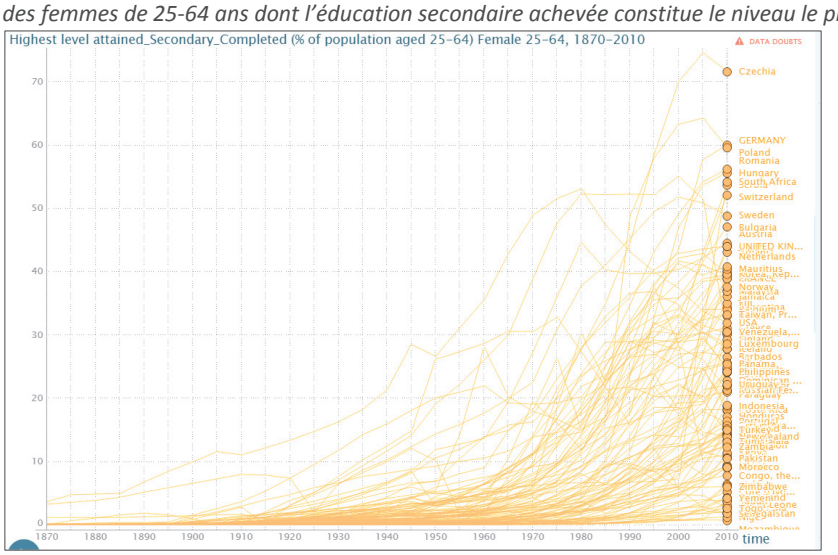
Source : Barro-Lee (2015)

Figure 0.41 Taux des hommes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint



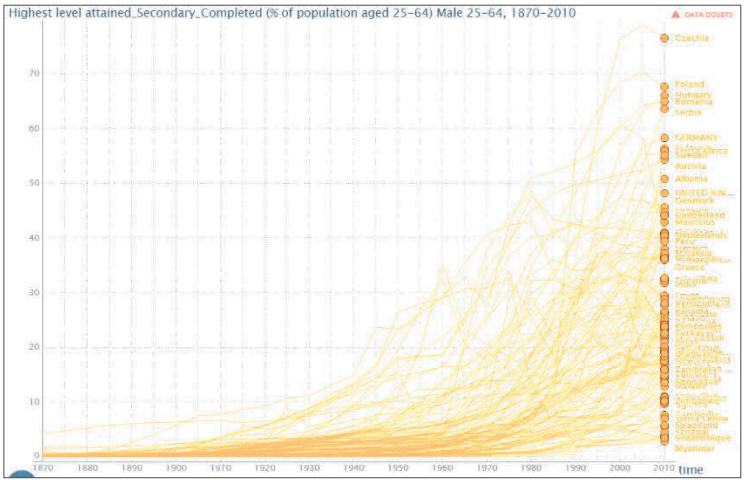
Source : Barro-Lee (2015)

Figure 0.42 Taux des femmes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint



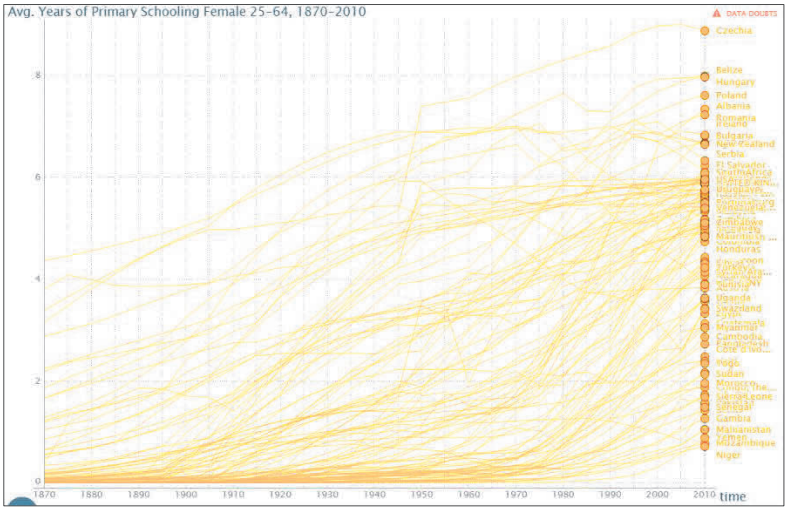
Source : Barro-Lee (2015)

Figure 0.43 Taux des hommes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint



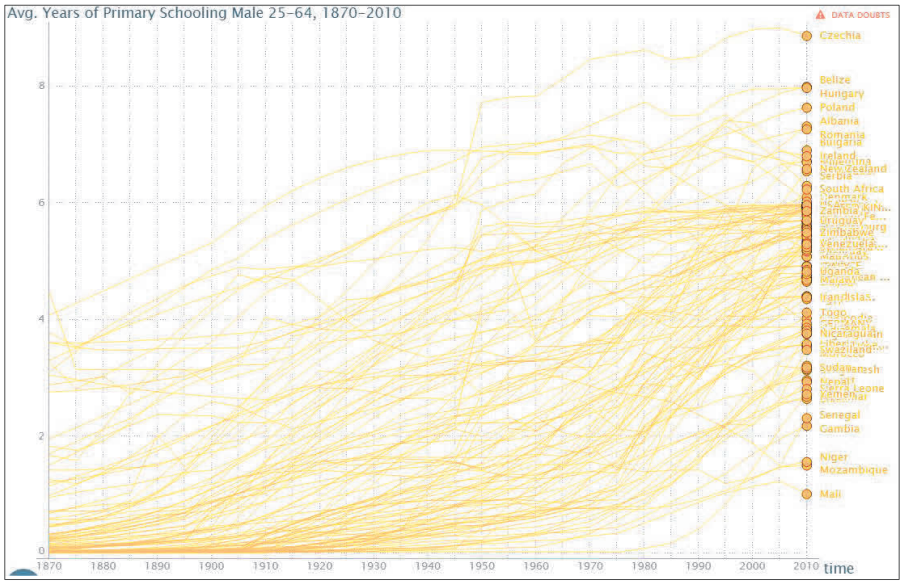
Source : Barro-Lee (2015)

Figure 0.44 Nombre moyen d'années d'études primaires des Femmes



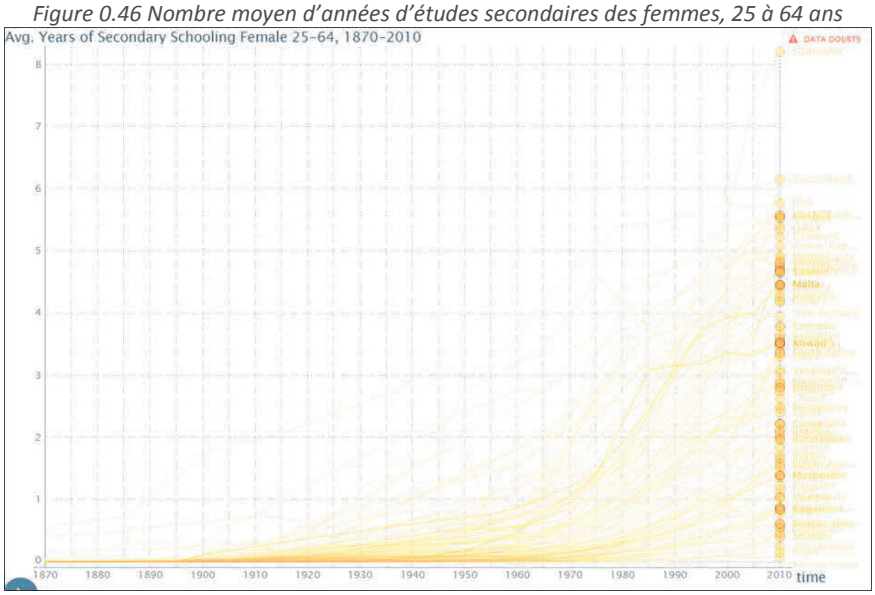
Source : Barro-Lee (2015)

Figure 0.45 Nombre moyen d'années d'études primaires des Hommes



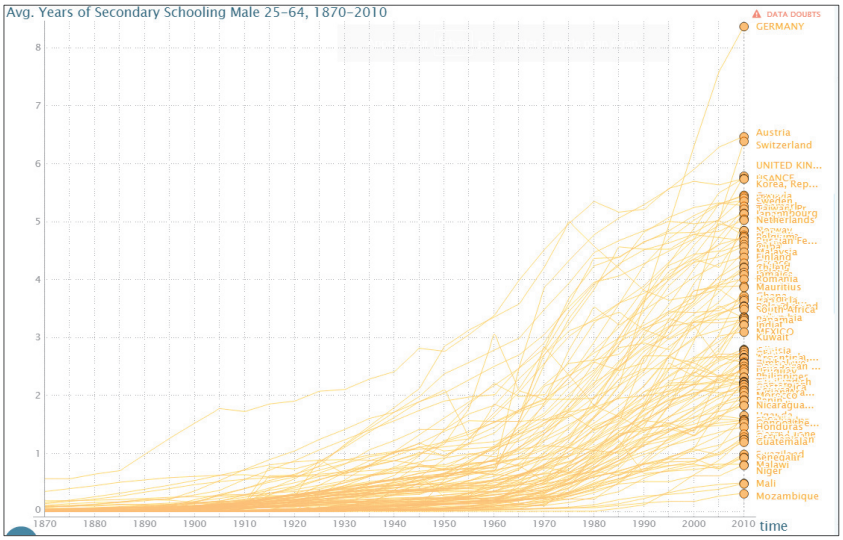
Source : Barro-Lee (2015)

Figure 0.46 Nombre moyen d'années d'études secondaires des femmes, 25 à 64 ans



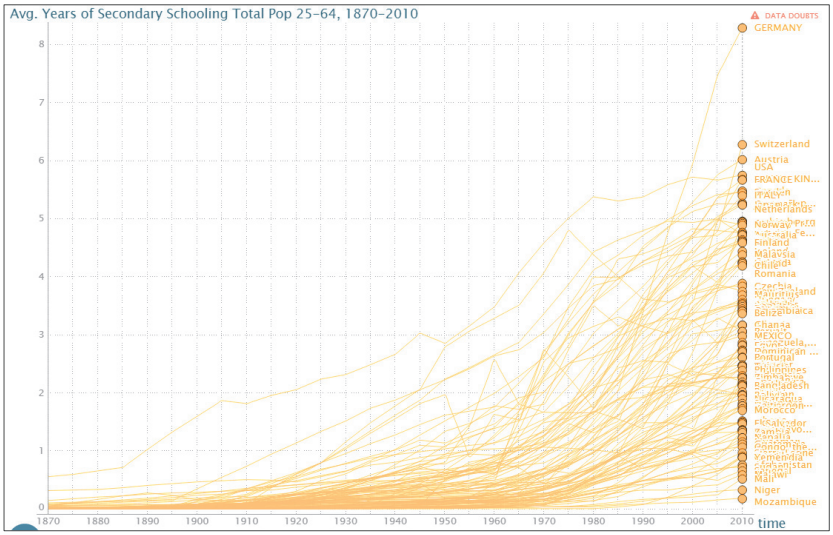
Source : Barro-Lee (2015)

Figure 0.47 Nombre moyen d'années d'études secondaires des hommes, 25 à 64 ans



Source : Barro-Lee (2015)

Figure 0.48 Nombre moyen d'années d'études secondaires, population totale, 25 à 64 ans



Source : Barro-Lee (2015)

Figure 0.49 Ratio femmes vs hommes, nombre moyen d'années d'études (% population totale 25-64 ans)

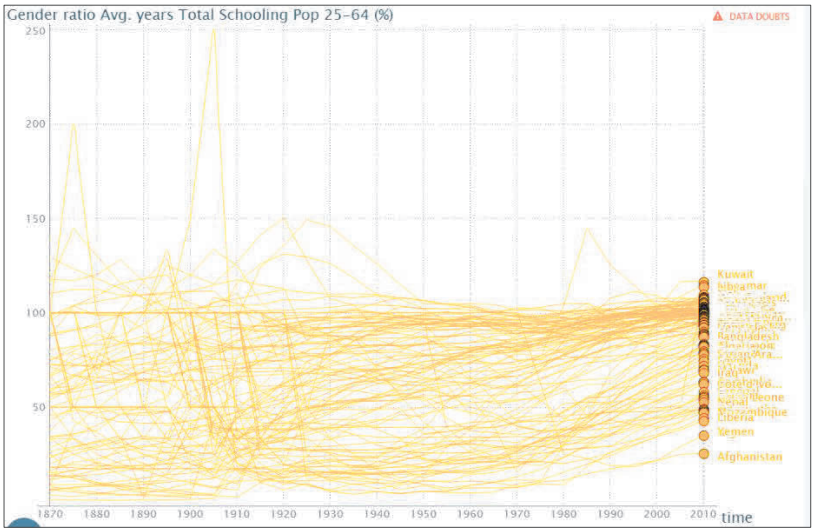


Table des figures

Figure 1.1 Schéma transitionnel longitudinal théorique 30
Figure 1.2 Evolution de la population mondiale 31
Figure 1.3 Exemple historique : Les USA, un cas avant-gardiste 32
Figure 1.4 Répartition de la population active par secteur, 1840-2010 33
Figure 1.5 Répartition du produit entre les secteurs, 1840-2010 33
Figure 1.6 Exemple historique : Le cas de l'Inde 33
Figure 1.7 Carte mondiale indiquant la proportion d'emplois dans le secteur de l'agriculture (% de l'emploi total) 2015 34
Figure 1.8 Répartition de la population active dans quelques pays vers 1985 34
Figure 1.9 Schéma transitionnel, phases et prospective 36
Figure 1.10 Schéma général de la transition démographique 36
Figure 1.11 Mouvement naturel de la population, Suède 1720-1980 37
Figure 1.12 Migration nette et accroissement naturel, Suède 1851-1980 37
Figure 1.13 Mouvement naturel de la population, Italie 1861-1980 38
Figure 1.14 Migration nette et accroissement naturel, Italie 1869-1980 39
Figure 1.15 La transition démographique en Angleterre 39
Figure 1.16 La transition démographique en Allemagne 40
Figure 1.17 Mouvement naturel de la population, Inde 1960-2010 40
Figure 1.18 Croissance annuelle de la population (en % de la population totale), Inde 41
Figure 1.19 Mouvement naturel de la population, Tunisie 1960-2010 41
Figure 1.20 Croissance annuelle de la population (en % de la population totale), Tunisie 42
Figure 1.21 Taux de fertilité (naissances par femme), Tunisie 42
Figure 1.22 Espérance de vie à 10 ans, 1750 à 2015 43
Figure 1.23 Âge médian, 1950 à 2100 43
Figure 1.24 Structure en pelure d'oignon (Hofstede et Minkov, 2010) : Valeurs et Pratiques « identitaires » 45
Figure 1.25 Ages et lieux d'acquisitions des valeurs et pratiques sociales (Hofstede et Minkov, 2010) . 45
Figure 1.26 Les 3 niveaux de « programmation mentale humaine » (Hofstede et Minkov, 2010) 46
Figure 1.27 Cartographies mondiales des dimensions anthropologiques au sens de Hofstede-Minkov (2010) 48
Figure 1.28 Parangon européen des types familiaux 52
Figure 1.29 . Cartographie mondiale des types familiaux 53
Figure 1.30 Carte européenne des types familiaux 53
Figure 1.31 Carte culturelle mondiale 56
Figure 1.32 Structure familiale et efficacité du travail 57
Figure 1.33 Relation entre structure familiale et efficacité du travail 58
Figure 1.34 Relation entre structure familiale, alphabétisation et chute de fécondité 58
Figure 1.35 Expériences naturelles de l'histoire : par exemple la division de la Corée en Nord et Sud .. 67
Figure 1.36 Primauté des institutions 68
Figure 1.37 Indice d'inclusivité des institutions économiques, 2015 78

Figure 1.38 Indice d'inclusivité des institutions économiques, Chine vs moyenne OCDE et moyenne Monde 79
Figure 2.1 Gain d'inertie en fonction du nombre de classes 105
Figure 2.2 Représentation des classes d'âge transitionnel de pays sur les deux premiers axes factoriels 106
Figure 2.3. Visualisation de tous les pays selon leur état initial sur la période étudiée 109
Figure 2.4. Visualisation par zone géographique (ou autre caractéristique de pays), des pays selon leur état initial sur la période étudiée 110
Figure 2.5. Visualisation des 10 séquences d'états les plus fréquentes 110
Figure 2.6. Visualisation des 10 séquences d'états les plus fréquentes, par zone géographique (ou autre caractéristique de pays) 111
Figure 2.7. Distributions transversales des états par année (par tranches verticales) 113
Figure 2.8. Indice d'entropie transversale 113
Figure 2.9. Mode des états dans les distributions transversales (une colonne par année) 114
Figure 2.10. Durée moyenne passée dans chaque état de l'alphabet 116
Figure 2.11. Dendrogramme de la classification des pays selon leurs trajectoires 117
Figure 2.12. Courbes des mesures de qualité des partitions selon le nombre de classes 118
Figure 2.13. Courbes des mesures standardisées de qualité des partitions selon le nombre de classes 119
Figure 2.14. Séquences (ou trajectoires) de chaque classe, avec quadrillage selon le pays et l'année 119
Figure 2.15. Séquence représentative de chaque classe. Années représentées par des colonnes. 120
Figure 2.16. Pseudo R2 et statistique de Levene entre les trajectoires et la variable explicative (ici zone géographique des pays), sur fenêtres de 5 années mobiles 123
Figure 2.17. Evolution des discrépances intra-groupes, sur fenêtres de 5 années mobiles 124
Figure 2.18. Les 15 sous-séquences les plus fréquentes 125
Figure 2.19. Sous-séquences discriminantes des classes de trajectoires. Fréquence relative. 127
Figure 2.20. Sous-séquences discriminantes des classes de trajectoires 127
Figure 3.1 Le taux d'emplois agricoles (% de l'emploi total), estimation BIT 153
Figure 3.2 Le taux d'emplois agricoles (% de l'emploi total), historique long 154
Figure 3.3 Part de l'agriculture dans le PIB (% du PIB), historique long 154
Figure 3.4 Le taux d'emplois agricoles (non-lissé) des pays de l'OCDE 155
Figure 3.5 Le taux d'emplois agricoles (%) des pays de l'OCDE. Historique long. 155
Figure 3.6 Le taux d'emplois agricoles (non-lissé) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) 156
Figure 3.7 Le taux d'emplois agricoles (%) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) 156
Figure 3.8 Le taux d'emplois agricoles (non-lissé) des pays à faible revenu par habitant (au sens de la Banque Mondiale) 157
Figure 3.9 Le taux d'emplois agricoles (%) des pays à faible revenu par habitant, (au sens de la Banque Mondiale) 157
Figure 3.10 Exemple de la France. Taux d'emplois agricoles (lissé) 159
Figure 3.11 Exemple du Ghana. Taux d'emplois agricoles (lissé) 159
Figure 3.12 Pourcentage de variance expliquée par les axes principaux (ou dimensions) 160
Figure 3.13 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives 160

Figure 3.14 Cercle de corrélation entre axes principaux (Dim) et variables actives (avec leur contribution à la création des axes) 161
Figure 3.15 Gain d'inertie en fonction du nombre de classes 162
Figure 3.16 Dendrogramme de la classification ascendante hiérarchique 162
Figure 3.17 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels 163
Figure 3.18 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes 164
Figure 3.19 Cartographie des Classes temporelles de Pays 165
Figure 3.20 Emploi dans l'agriculture (% de l'emploi total) 166
Figure 3.21 Emploi dans l'industrie (% de l'emploi total) 166
Figure 3.22 Emploi dans les services (% de l'emploi total) 167
Figure 3.23 Population âgée de 15 à 64 ans (% de la population totale) 168
Figure 3.24 Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010) 168
Figure 3.25 Industrie, valeur ajoutée par travailleur (USD constants de 2010) 169
Figure 3.26 Services, valeur ajoutée par travailleur (USD constants de 2010) 170
Figure 3.27 Indice de développement humain (HDI*10) 171
Figure 3.28 Espérance de vie à la naissance, Total 171
Figure 3.29 PIB par habitant (USD constants de 2010) 171
Figure 3.30 Population âgée de 65 ans et plus (% de la population totale) 172
Figure 3.31 Division de la population des Nations Unies (âge médian) 172
Figure 3.32 La part de l'agriculture dans l'emploi (% de l'emploi total) 172
Figure 3.33 Industrie, valeur ajoutée (% du PIB) 173
Figure 3.34 Services, valeur ajoutée (% du PIB) 174
Figure 3.35 Valeurs des inégalités (estimation de la Banque mondiale) 174
Figure 3.36 Ratio de dépendance à l'âge (% de la population en âge de travailler) 175
Figure 3.37 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler) 175
Figure 3.38 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler) 175
Figure 3.39 Ratio d'âge 15-39 à 40-64 176
Figure 3.40 Population âgée de 15 à 64 ans (% du total) 176
Figure 3.41 Taux de mortalité infantile (pour 1 000 naissances vivantes) 176
Figure 3.42 Taux de fécondité, total (naissances par femme) 177
Figure 3.43 Indice de bien-être (FOSTER-SEN) 177
Figure 3.44 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne par classe d'âge transitionnel 178
Figure 3.45 Schéma transitionnel transversal selon l'âge transitionnel. Médiane par classe d'âge transitionnel 178
Figure 3.46 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse 179
Figure 3.47 Schéma transitionnel longitudinal théorique 179
Figure 3.48 Correspondance entre Classes d'âge transitionnelles et Phases théoriques de la transition économique 180
Figure 3.49 Mortalité infantile (non-lissée), ensemble des pays du monde 182
Figure 3.50 Mortalité infantile de 1751 à 2013 182

Figure 3.51 Mortalité infantile des pays de l'OCDE 183
Figure 3.52 Mortalité infantile de 1751 à 2013 183
Figure 3.53 Mortalité infantile (non-lissée) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) 184
Figure 3.54 Mortalité infantile (non lissée) des pays à faible revenu par habitant (au sens de la Banque Mondiale) 184
Figure 3.55 Exemple de la France d'après-guerre (taux de mortalité infantile pour 1000 naissances vivantes lissé) 186
Figure 3.56 Exemple de l'Afghanistan ((taux de mortalité infantile pour 1000 naissances vivantes lissé) 186
Figure 3.57 Pourcentage de variance expliquée par les axes principaux (ou dimensions) 187
Figure 3.58 Corrélation entre axes principaux (Dim) et variables actives 187
Figure 3.59 Cercle de corrélation entre axes principaux (Dim) et variables actives 188
Figure 3.60 Gain d'inertie en fonction du nombre de classes 188
Figure 3.61 Dendrogramme de la classification ascendante hiérarchique 189
Figure 3.62 Représentation en polygones des classes temporelles de pays sur les deux premiers axes factoriels 189
Figure 3.63 Représentation en ellipse des classes temporelles de pays sur les deux premiers axes 190
Figure 3.64 Cartographies des Classes Transitionnelles de Pays 191
Figure 3.65 Taux de natalité, brut (pour 1 000 personnes) 191
Figure 3.66 Taux de mortalité, brut (pour 1 000 personnes) 192
*Figure 3.67 Indice de développement humain (HDI*10)* 193
Figure 3.68 Espérance de vie à la naissance 193
Figure 3.69 PIB par habitant (USD constants de 2010) 194
Figure 3.70 Population âgée de 65 ans et plus (% population totale) 194
Figure 3.71 Division de la population des Nations Unies (âge médian) 194
Figure 3.72 La part de l'agriculture dans l'emploi (% de l'emploi total) 195
Figure 3.73 Industrie, valeur ajoutée 195
Figure 3.74 Services, valeur ajoutée (% du PIB 196
Figure 3.75 Taux de mortalité infantile (pour 1 000 naissances vivantes) 197
Figure 3.76 Population de 15 ans et plus (% du total) 197
Figure 3.77 Emploi dans l'agriculture 15 ans et plus, (% population 15 ans et plus) 197
Figure 3.78 Emploi dans l'industrie 15+ (% population 15+) 198
Figure 3.79 Emploi dans les services 15+ (% population 15+) 198
Figure 3.80 Valeur ajoutée agricole par travailleur, mesure de la productivité agricole (USD constants de 2010) 198
Figure 3.81 Industrie, valeur ajoutée par travailleur (USD constants de 2010) 199
Figure 3.82 Services, valeur ajoutée par travailleur (USD constants de 2010) 199
Figure 3.83 Valeurs des inégalités (estimation de la Banque mondiale) 199
Figure 3.84 Ratio de dépendance à l'âge (% de la population en âge de travailler) 200
Figure 3.85 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler) 200
Figure 3.86 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler) 200
Figure 3.87 Taux de fécondité, total (naissances par femme) 201
Figure 3.88 Ratio d'âge 15-39 à 40-64 201
Figure 3.89 Population âgée de 15 à 64 ans (% du total) 201

Figure 3.90 Indice de bien-être (FOSTER-SEN) 202
Figure 3.91 Schéma transitionnel transversal selon le temps transitionnel. Moyennes par classes d'âge transitionnel 202
Figure 3.92 Schéma transitionnel transversal selon le temps transitionnel. Médianes par classes d'âge transitionnel 203
Figure 3.93 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse 203
Figure 3.94 Schéma démographique transitionnel longitudinal théorique 204
Figure 3.95 Schéma démographique transitionnel longitudinal théorique, avec valeurs de référence . 204
Figure 3.96 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (non-lissée), ensemble des pays du monde 207
Figure 3.97 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans des pays de l'OCDE 208
Figure 3.98 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (non-lissée) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) 208
Figure 3.99 Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (non lissée) des pays à faible revenu par habitant (au sens de la Banque Mondiale) 208
Figure 3.100 Exemple de la France d'après-guerre (Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans lissé) 210
Figure 3.101 Exemple de l'Afghanistan (Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans lissé) 210
Figure 3.102 Algérie : Ratio Femmes vs Hommes nombre moyen d'années d'étude, 25-34 ans (lissé)210
Figure 3.103 Pourcentage de variance expliquée par les axes principaux (ou dimensions) 211
Figure 3.104 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives 212
Figure 3.105 Cercle de corrélation entre axes principaux (Dim) et variables actives (avec leur contribution à la création des axes) 212
Figure 3.106 Gain d'inertie en fonction du nombre de classes 213
Figure 3.107 Dendrogramme de la classification ascendante hiérarchique 213
Figure 3.108 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels 214
Figure 3.109 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes 215
Figure 3.110 Cartographie des Classes temporelles de Pays 216
Figure 3.111 Valeurs d'expression personnelle (vs Valeurs de survie) 216
Figure 3.112 Valeurs laïques-rationnelles (vs Valeurs traditionnelles) 217
Figure 3.113 Valeurs inégalitaires (estimation Banque mondiale) 217
Figure 3.114 Taux de natalité, brut (pour 1 000 personnes) 218
Figure 3.115 Emploi dans l'agriculture (% de l'emploi total) 218
Figure 3.116 Emploi dans l'industrie (% de l'emploi total) 218
Figure 3.117 Emploi dans les services (% de l'emploi total) 219
Figure 3.118 Taux de fécondité, total (naissances par femme) 219
Figure 3.119 PIB par habitant (USD constants de 2010) 219
Figure 3.120 Indice de développement humain (HDI*10) 220
Figure 3.121 Taux de mortalité infantile (pour 1 000 naissances vivantes) 220
Figure 3.122 Indice de bien-être (FOSTER-SEN) 220

Figure 3.123 Ratio d'âge15-39 à 40-64 221
Figure 3.124 Division de la population des Nations Unies (âge médian) 221
Figure 3.125 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne par classe d'âge transitionnel 222
Figure 3.126 Schéma transitionnel transversal selon l'âge transitionnel. Médiane par classe d'âge transitionnel 223
Figure 3.127 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse 223
Figure 3.128 Indice de développement humain (HDI*10) 226
Figure 3.129 Indice de développement humain (HDI*10) des pays de l'OCDE 227
Figure 3.130 Indice de développement humain (HDI*10) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) 227
Figure 3.131 Indice de développement humain (HDI*10) des pays à faible revenu par habitant (au sens de la Banque Mondiale) 227
Figure 3.132 Exemple de la France d'après-guerre, indice de développement humain (IDH) lissé 229
Figure 3.133 Exemple de l'Afghanistan, indice de développement humain (IDH) lissé 229
Figure 3.134 Pourcentage de variance expliquée par les axes principaux (ou dimensions) 230
Figure 3.135 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives 230
Figure 3.136 Cercle de corrélation entre axes principaux et variables actives (avec leur contribution à la création des axes) 231
Figure 3.137 Gain d'inertie en fonction du nombre de classes 231
Figure 3.138 Dendrogramme de la classification ascendante hiérarchique 232
Figure 3.139 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels 233
Figure 3.140 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes 233
Figure 3.141 Cartographie des Classes temporelles de Pays 235
Figure 3.142 Indice global des libertés civiles (1=libre, 7=répression). Courbe transversale. 235
Figure 3.143 Indice global des libertés civiles (1=libre, 7=répression) 236
Figure 3.144 Indice global de démocratie (Polity IV-Roser). Courbe transversale. 236
Figure 3.145 Indice global de démocratie (Polity IV-Roser) 236
Figure 3.146 Indice de développement humain (HDI*10). Courbe transversale. 239
Figure 3.147 Espérance de vie à la naissance. Moyennes par classe d'âge transitionnel 240
Figure 3.148 PIB par habitant (USD constants de 2010). Moyennes par classe d'âge transitionnel 240
Figure 3.149 Population âgée de 65 ans et plus. Moyennes par classe d'âge transitionnel 241
Figure 3.150 Division de la population des Nations Unies (âge médian). Moyennes par classe d'âge transitionnel 241
Figure 3.151 La part de l'agriculture dans l'emploi (% de l'emploi total), historique long). Moyenne par classe d'âge transitionnel 241
Figure 3.152 Industrie, valeur ajoutée (% du PIB) 242
Figure 3.153 Emploi dans les services (% de l'emploi total). Moyenne par classe d'âge transitionnel 242
Figure 3.154 Taux de mortalité infantile (pour 1 000 naissances vivantes). Moyenne par classe d'âge transitionnel 242
Figure 3.155 Ratio de dépendance à l'âge (% de la population en âge de travailler). Moyenne par classe d'âge transitionnel 243

Figure 3.156 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler). Moyenne par classe d'âge transitionnel 243
Figure 3.157 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler). Moyenne par classe d'âge transitionnel 243
Figure 3.158 Taux de fécondité, total (naissances par femme). Moyenne par classe d'âge transitionnel 244
Figure 3.159 Ratio d'âge15-39 à 40-64. Moyenne par classe d'âge transitionnel 244
Figure 3.160 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne par classe d'âge transitionnel 245
Figure 3.161 Schéma transitionnel transversal selon l'âge transitionnel. Médiane par classe d'âge transitionnel 246
Figure 3.162 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse 246
Figure 3.163 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, 1870-2010 249
Figure 3.164 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, des pays à revenu élevé de 1870 à 2010 249
Figure 3.165 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, (non-lissé) des pays à moyen revenu par habitant (au sens de la Banque Mondiale) de 1870 à 2010 250
Figure 3.166 Nombre moyen d'années de scolarisation, hommes de 25-64 ans, (non lissé) des pays à faible revenu par habitant (au sens de la Banque Mondiale) de 1870 à 2010 250
Figure 3.167 Exemple du Sénégal. Nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé) 252
Figure 3.168 Exemple de la France. Nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé) 252
Figure 3.169 Exemple du Ghana. Nombre moyen d'années de scolarisation des hommes de 25-64 ans (lissé) 252
Figure 3.170 Pourcentage de variance expliquée par les axes principaux (ou dimensions) 253
Figure 3.171 Corrélation entre les 2 premiers axes principaux (Dim) et variables actives 254
Figure 3.172 Cercle de corrélation entre axes principaux et variables actives (avec leur contribution à la création des axes) 254
Figure 3.173 Gain d'inertie en fonction du nombre de classes 255
Figure 3.174 Dendrogramme de la classification ascendante hiérarchique 255
Figure 3.175 Représentation en polygones des classes d'âge transitionnel de pays sur les deux premiers axes factoriels 256
Figure 3.176 Représentation en ellipse des classes d'âge transitionnel de pays sur les deux premiers axes 257
Figure 3.177 Cartographie des Classes temporelles de Pays 258
Figure 3.178 Population totale non scolarisée de 25 à 64 ans 259
Figure 3.179 Taux (%) d'adultes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint 259
Figure 3.180 Taux (%) d'adultes de 25-64 ans, dont l'éducation primaire achevée, constitue le niveau le plus élevé atteint 260
Figure 3.181 Taux (%) d'adultes de 25-64 ans, dont l'éducation secondaire achevée, constitue le niveau le plus élevé atteint 261
Figure 3.182 compétences cognitives 261

Figure 3.183 Indice de développement humain (HDI*10) 262
Figure 3.184 Espérance de vie à la naissance, Total 263
Figure 3.185 PIB par habitant (USD constants de 2010) 263
Figure 3.186 Population âgée de 65 ans et plus, total 263
Figure 3.187 Division de la population des Nations Unies (âge médian) 264
Figure 3.188 Part de l'agriculture dans l'emploi (% de l'emploi total) 264
Figure 3.189 Industrie, valeur ajoutée (% du PIB) 264
Figure 3.190 Services, valeur ajoutée (% du PIB) 265
Figure 3.191 Taux de mortalité infantile (pour 1 000 naissances vivantes) 265
Figure 3.192 Ratio de dépendance à l'âge (% de la population en âge de travailler) 265
Figure 3.193 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler) 266
Figure 3.194 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler) 266
Figure 3.195 Taux de fécondité, total (naissances par femme) 266
Figure 3.196 Ratio d'âge15-39 à 40-64 267
Figure 3.197 Nombre moyen d'années de scolarisation totale des hommes de 25 à 64 ans 267
Figure 3.198 Taux (%) des Femmes non scolarisées de 25 à 64 ans 267
Figure 3.199 Taux (%) des Hommes non scolarisés de 25 à 64 ans 268
Figure 3.200 Taux (%) des femmes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint 268
Figure 3.201 Nombre moyen d'années d'études supérieures des femmes, 25 à 64 ans 268
Figure 3.202 Nombre moyen d'années d'études supérieures, population totale, 25 à 64 ans 269
Figure 3.203 Nombre moyen d'années d'études supérieures, des hommes, 25 à 64 ans 269
Figure 3.204 Taux des hommes de 25-64 ans dont l'éducation tertiaire achevée constitue le niveau le plus élevé atteint 269
Figure 3.205 Nombre moyen d'années de scolarisation des Femmes de 25 à 64 ans 270
Figure 3.206 Nombre moyen d'années de scolarisation (Population totale) 25 à 64 ans 270
Figure 3.207 Nombre moyen d'années de scolarisation des hommes, 25 à 64 ans 271
Figure 3.208 Taux des femmes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint, ... 271
Figure 3.209 Taux des hommes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint 271
Figure 3.210 Taux des femmes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint 272
Figure 3.211 Taux des hommes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint 272
Figure 3.212 Nombre moyen d'années d'études primaires des Femmes 272
Figure 3.213 Nombre moyen d'années d'études primaires des Hommes 273
Figure 3.214 Nombre moyen d'années d'études secondaires des femmes, 25 à 64 ans 273
Figure 3.215 Nombre moyen d'années d'études secondaires des hommes, 25 à 64 ans 273
Figure 3.216 Nombre moyen d'années d'études secondaires, population totale, 25 à 64 ans 274
Figure 3.217 Ratio femmes vs hommes, nombre moyen d'années d'études (% population totale 25-64 ans) 274
Figure 3.218 Taux de natalité, brut (pour 1 000 personnes) 274
Figure 3.219 Schéma transitionnel selon l'indice d'âge transitionnel. Moyenne par classe d'âge transitionnel 276

Figure 3.220 Schéma transitionnel selon l'indice d'âge transitionnel. Médiane par classe d'âge transitionnel 277
Figure 3.221 Schéma transitionnel transversal selon l'âge transitionnel. Moyenne d'âge transitionnel des classes en abscisse 277
Figure 3.222 Economie vs démographie. Appariement descriptif des classes d'âge transitionnel des pays 279
Figure 3.223 Economie vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays 280
Figure 3.224 Economie vs cognition. Appariement descriptif des classes d'âge transitionnel des pays 280
Figure 3.225 Economie vs Institutions. Appariement descriptif des classes d'âge transitionnel des pays 280
Figure 3.226 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel économiques et classes des autres transitions 281
Figure 3.227 Démographie vs économie. Appariement descriptif des classes d'âge transitionnel des pays 283
Figure 3.228 Démographie vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays 283
Figure 3.229 Démographie vs cognition. Appariement descriptif des classes d'âge transitionnel des pays 283
Figure 3.230 Démographie vs institutions. Appariement descriptif des classes d'âge transitionnel des pays 284
Figure 3.231 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel démographiques et classes des autres transitions 284
Figure 3.232 Anthropologie vs économie. Appariement descriptif des classes d'âge transitionnel des pays 286
Figure 3.233 Anthropologie vs démographie. Appariement descriptif des classes d'âge transitionnel des pays 286
Figure 3.234 Anthropologie vs cognition. Appariement descriptif des classes d'âge transitionnel des pays 286
Figure 3.235 Anthropologie vs institutions. Appariement descriptif des classes d'âge transitionnel des pays 287
Figure 3.236 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel anthropologiques et classes des autres transitions 287
Figure 3.237 Institutions vs économie. Appariement descriptif des classes d'âge transitionnel des pays 289
Figure 3.238 Institutions vs démographie. Appariement descriptif des classes d'âge transitionnel des pays 289
Figure 3.239 Institutions vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays 289
Figure 3.240 Institutions vs cognition. Appariement descriptif des classes d'âge transitionnel des pays 290
Figure 3.241 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel institutionnelles et classes des autres transitions 290
Figure 3.242 Cognition vs économie. Appariement descriptif des classes d'âge transitionnel des pays 292

Figure 3.243 Cognition vs démographie. Appariement descriptif des classes d'âge transitionnel des pays 292
Figure 3.244 Cognition vs anthropologie. Appariement descriptif des classes d'âge transitionnel des pays 292
Figure 3.245 Cognition vs institutions. Appariement descriptif des classes d'âge transitionnel des pays 293
Figure 3.246 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel cognitives et classes des autres transitions 293
Figure 4.1 Décalages temporels et différences d'allure des taux mortalité selon les tranches d'âge. France 304
Figure 4.2 Profils principaux de transition démographique 305
Figure 4.3 Courbes de mortalité des pays A (bleue) et R (noire) 321
Figure 4.4 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision 329
Figure 4.5 MAPE (%) des probabilités de décès selon les âges et les horizons 330
Figure 4.6 Prévisions comparées des modèles pour quelques âges 331
Figure 4.7 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Présence de valeurs atypiques 332
Figure 4.8 MAPE (%) des probabilités de décès selon les âges et les horizons. Présence de valeurs atypiques 332
Figure 4.9 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Historique d'apprentissage de 20 ans au lieu de 85 ans 333
Figure 4.10 MAPE (%) des probabilités de décès selon les âges et les horizons. Historique d'apprentissage de 20 ans au lieu de 85 ans 334
Figure 4.11 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision 335
Figure 4.12 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Sans le modèle CBD 335
Figure 4.13 MAPE (%) des probabilités de décès selon les âges et les horizons 336
Figure 4.14 Prévisions comparées des modèles pour quelques âges 337
Figure 4.15 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Présence de valeurs atypiques 338
Figure 4.16 MAPE (%) des probabilités de décès selon les âges et les horizons. Présence de valeurs atypiques 338
Figure 4.17 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision. Historique d'apprentissage de 20 ans au lieu de 93 ans 339
Figure 4.18 MAPE (%) des probabilités de décès selon les âges et les horizons. Historique d'apprentissage de 20 ans au lieu de 93 ans 340
Figure 4.19 Taux de mortalité à l'âge 0 des femmes. Quelques pays européens vs Inde 341
Figure 4.20 Courbes de variables cliométriques selon l'âge Inde 341
Figure 4.21 Performance des modèles (tous âges confondus) selon l'horizon maximal de prévision 344
Figure 4.22 MAPE des modèles selon les âges et les horizons 344
Figure 4.23 Prévisions comparées des modèles, pour quelques âges 345
Figure 4.24 MAPE (%) tous âges confondus, selon quelques horizons 346
Figure 4.25 MAPE (%) selon les âges et les horizons 346
Figure 4.26 Prévisions comparées des modèles, pour quelques âges 347
Figure 4.27 MAPE (%) tous âges confondus, selon quelques horizons 347

Figure 4.28 MAPE (%) selon les âges et les horizons....348
Figure 4.29 Prévisions comparées des modèles, pour quelques âges....348
Figure 4.30 MAPE (%) tous âges confondus, selon quelques horizons....349
Figure 4.31 MAPE (%) selon les âges et les horizons....350
Figure 4.32 Prévisions comparées des modèles, pour quelques âges....350
Figure 4.33 MAPE (%) tous âges confondus, selon quelques horizons....351
Figure 4.34 MAPE (%) selon les âges et les horizons....352
Figure 4.35 Prévisions comparées des modèles, pour quelques âges....352
Figure 4.36 Récapitulatif des performances des modèles....354
Figure 5.1 Données et objectif analytique du chapitre....365
Figure 5.2 Exemple d'une fiche d'un soldat « Mort pour la France »....366
Figure 5.3 Architecture de la base de données cliométriques, et son module d'analyse statistique....368
Figure 5.4 Les modalités de participations (ou non) des régiments aux batailles et combats....368
Figure 5.5 Nombre de régiments d'infanterie en casernement au début août 1914....370
Figure 5.6Nombre de RI par département au 1er aout 1914, avant le début de la guerre....371
Figure 5.7 Projection des RI sur le front, un mois après le début de la guerre. Nombre de RI par département....371
Figure 5.8 Analyse de la concentration de troupes d'infanterie, dans la Marne, au 1er novembre 1915....372
Figure 5.9 Exemple de liste de RI ayant participé à un combat ou bataille (ici l'offensive d'Argonne 1915)....373
Figure 5.10 Informations des RI, concernant l'offensive d'Argonne de 1915....373
Figure 5.11 Exemple de tableau extractible de données sur les RI listés selon le mode de participation à un combat ou bataille....374
Figure 5.12 Nombre de MPF (Morts Pour la France) par jour d'existence du RI, pondéré par la mortalité quotidienne de la période d'existence du régiment pour 1000 soldats initiaux....377
Figure 5.13 Les 8 Zones d'Etude et d'Aménagement du Territoire (ZEAT) de l'INSEE....378
Figure 5.14 Taux de mortalité quotidienne par Zones d'Etude et d'Aménagement du Territoire....379
Figure 5.15 Votes aux élections législatives de 1914 par circonscriptions électorales (Salmon, 2001) 380
Figure 5.16 Corrélation de la mortalité dans les régiments d'infanterie avec les votes aux législatives de 1914 dans les départements d'origine....380
Figure 5.17 Types familiaux en France pré-industrielle, Todd (1983)....382
Figure 5.18 Parangons européens des types familiaux....382
Figure 5.19 Modalités de Participation (ou non) des RI aux batailles et combats....387
Figure 5.20 Part de variance de mortalité quotidienne expliquée, selon les dimensions analysées....388
Figure 5.21 Part de variance de mortalité quotidienne expliquée, selon les dimensions militaire et origine du RI....388
Figure 0.1 Indice de développement humain (HDI*10)....415
Figure 0.2 Espérance de vie à la naissance, Total....415
Figure 0.3 PIB par habitant (USD constants de 2010)....415
Figure 0.4 Population âgée de 65 ans et plus (% de la population totale)....416
Figure 0.5 Division de la population des Nations Unies (âge médian)....416
Figure 0.6 Part de l'agriculture dans l'emploi (% de l'emploi total)....416
Figure 0.7 Industrie, valeur ajoutée (% du PIB)....417
Figure 0.8 Services, valeur ajoutée (% du PIB)....417

Figure 0.9 Valeurs inégalitaires (estimation Banque mondiale) 417
Figure 0.10 Ratio de dépendance à l'âge (% de la population en âge de travailler) 418
Figure 0.11 Ratio de dépendance à l'âge, adultes (% de la population en âge de travailler) 418
Figure 0.12 Ratio de dépendance à l'âge, jeunes (% de la population en âge de travailler) 418
Figure 0.13 Ratio d'âge15-39 à 40-64 419
Figure 0.14 Population âgée de 15 à 64 ans (% du total) 419
Figure 0.15 Taux de mortalité infantile (pour 1 000 naissances vivantes) 419
Figure 0.16 Taux de fécondité, total (naissances par femme) 420
Figure 0.17 Indice de bien-être (FOSTER-SEN) 420
Figure 0.18 Population de 15 ans et plus (% du total) 420
Figure 0.19 Emploi dans l'agriculture 15 ans et plus, (% population 15 ans et plus) 421
Figure 0.20 Emploi dans l'industrie 15+(% population 15+) 421
Figure 0.21 Emploi dans les services 15+(% population 15+) 421
Figure 0.22 Valeur ajoutée agricole par travailleur, mesure de la productivité agricole ($ US constants de 2010) 422
Figure 0.23 Industrie, valeur ajoutée par travailleur (USD constants de 2010) 422
Figure 0.24 Services, valeur ajoutée par travailleur (USD constants de 2010) 422
Figure 0.25 Population âgée de 15 à 64 ans (% du total) 423
Figure 0.26 Taux de natalité, brut (pour 1 000 personnes) 423
Figure 0.27 Emploi dans l'agriculture (% de l'emploi total) 423
Figure 0.28 Emploi dans les services (% de l'emploi total) 424
Figure 0.29 Nombre moyen d'années de scolarisation totale des hommes de 25 à 64 ans 424
Figure 0.30 Taux (%) des Femmes non scolarisées de 25 à 64 ans 424
Figure 0.31 Taux (%) des Hommes non scolarisés de 25 à 64 ans 425
Figure 0.32 Taux (%) des femmes de 25-64 ans, dont l'éducation tertiaire achevée, constitue le niveau le plus élevé atteint 425
Figure 0.33 Nombre moyen d'années d'études supérieures des femmes, 25 à 64 ans 425
Figure 0.34 Nombre moyen d'années d'études supérieures, population totale, 25 à 64 ans 426
Figure 0.35 Nombre moyen d'années d'études supérieures, des hommes, 25 à 64 ans 426
Figure 0.36 Taux des hommes de 25-64 ans dont l'éducation tertiaire achevée constitue le niveau le plus élevé atteint 426
Figure 0.37 Nombre moyen d'années de scolarisation des Femmes de 25 à 64 ans 427
Figure 0.38 Nombre moyen d'années de scolarisation (Population totale) 25 à 64 ans 427
Figure 0.39 Nombre moyen d'années de scolarisation des hommes, 25 à 64 ans 427
Figure 0.40 Taux des femmes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint, ... 428
Figure 0.41 Taux des hommes de 25-64 ans dont l'éducation primaire achevée constitue le niveau le plus élevé atteint 428
Figure 0.42 Taux des femmes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint 428
Figure 0.43 Taux des hommes de 25-64 ans dont l'éducation secondaire achevée constitue le niveau le plus élevé atteint 429
Figure 0.44 Nombre moyen d'années d'études primaires des Femmes 429
Figure 0.45 Nombre moyen d'années d'études primaires des Hommes 429
Figure 0.46 Nombre moyen d'années d'études secondaires des femmes, 25 à 64 ans 430

Figure 0.47 Nombre moyen d'années d'études secondaires des hommes, 25 à 64 ans 430
Figure 0.48 Nombre moyen d'années d'études secondaires, population totale, 25 à 64 ans 430
Figure 0.49 Ratio femmes vs hommes, nombre moyen d'années d'études (% population totale 25-64 ans) 431

Liste des tableaux

Tableau 1.1 Caractéristiques des types familiaux 51
Tableau 1.2 Types familiaux et phénomènes sociaux 52
Tableau 1.3 Quelques statistiques des types familiaux 54
Tableau 1.4 Verticalité (ou Autorité) versus Liberté 54
Tableau 1.5 Endogamie versus Exogamie 55
Tableau 1.6 Statut des femmes 55
Tableau 1.7 Egalité versus Inégalité 55
Tableau 1.8 Quelques corrélations concernant le tiers monde (d'après les données statistiques de l'annexe 1) 58
Tableau 1.9 Structures familiales dominantes et taux d'inflation dans les pays de l'OCDE en 1980 59
Tableau 1.10 Inégalités de revenu mesurées par le coefficient de Gini 61
Tableau 1.11 Inégalités de revenu : Corrélations et variance expliquée par les variables anthropologiques 62
Tableau 1.12 Revue des mesures des institutions 74
Tableau 1.13 Indice d'inclusivité des institutions économiques, Chine vs moyenne OCDE et moyenne Monde 79
Tableau 1.14 Matrice IEPE (Interne, Externe, Politique, Economique) de mesure des institutions 81
Tableau 1.15 Sigles liés aux mesures des compétences cognitives et du niveau d'éducation 87
Tableau 1.16 Revue des mesures des compétences cognitives et du niveau d'éducation 87
Tableau 1.17 Variables cognitives corrélées au QI national 92
Tableau 1.18 Niveau d'éducation corrélé au QI national 92
Tableau 2.1 Variables actives 104
Tableau 2.2 Description des classes selon les variables actives 107
Tableau 2.3. Nombre et Pourcentage d'occurrences de toutes les séquences d'états 112
Tableau 2.4 Temps passé dans les différents états, pour chaque pays 115
Tableau 2.5. Statistiques descriptives du Temps passé dans les différents états 116
Tableau 2.6. Taux de transition entre états de l'alphabet 116
Tableau 2.7. Mesures de la qualité d'une partition (clustering) 118
Tableau 2.8. ANOVA-généralisé pour m groupes (ou modalités) 122
Tableau 2.9. Estimation et test de la part de discrépance des séquences expliquée par la zone géographique du pays 122
Tableau 2.10. Discrépance des séquences par modalité de la variable explicative 123
Tableau 2.11. Sous-séquences les plus fréquentes 126
Tableau 2.12 règles d'associations séquentielles obtenues avec un support relatif supérieur ou égal à 0.01% 129
Tableau 3.1 Variables actives 152
Tableau 3.2 Description des classes selon les variables actives 164
Tableau 3.3 Variables actives ou de transition 181
Tableau 3.4 Description des classes selon les variables actives 190

Tableau 3.5 Correspondance entre Classes d'âge transitionnel et Phases théoriques de la transition démographique 205
Tableau 3.6 Variables actives ou de transition..... 206
Tableau 3.7 Description des classes selon les moyennes des variables actives 215
Tableau 3.8 Valeurs moyennes des variables transitionnelles par classes d'âge transitionnel..... 224
Tableau 3.9 Variables actives ou transitionnelles 225
Tableau 3.10 Description des classes selon les moyennes des variables actives 234
Tableau 3.11 Valeurs moyennes des variables transitionnelles par classes d'âge transitionnel..... 247
Tableau 3.12 Variables actives ou transitionnelles 248
Tableau 3.13 Description des classes selon les variables actives..... 257
Tableau 3.14 Valeurs moyennes des variables transitionnelles par classes d'âge transitionnel..... 278
Tableau 3.15 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel des transitions ACIDE 295
Tableau 3.16 Intensité des liens entre les 5 transitions ACIDE, selon l'échelle de Cohen 295
Tableau 3.17 V de Cramer (significatifs au seuil de 5%) entre les classes d'âge transitionnel des transitions ACIDE. Modalités croisées. 295
Tableau 4.1 Comparaison théorique des modèles Logit-PCR et Logit-PCR-Optimal 317
Tableau 4.2 Comparaison théorique des modèles Logit-PCR-Optimal et Logit-PLS 320
Tableau 4.3 Tableau du pays A..... 323
Tableau 4.4 Tableau d'appariement des tables du pays A et du pays R 324
Tableau 4.5 Tableau du pays A. Méthode 2 327
Tableau 4.6 Tableau d'appariement des tables du pays A et du pays R 328
Tableau 5-1 Stratégie de test d'association statistique 377
Tableau 5-2 Test d'association entre le taux de mortalité par jour pondéré (par la mortalité) pour 1000 soldats initiaux d'un RI et les modalités de Participation des RI aux grandes batailles..... 385
Tableau 5-3 Test d'association entre le taux de mortalité par jour pondéré (par la mortalité) pour 1000 soldats initiaux d'un RI et les modalités de Participation des RI aux batailles et combats..... 386