



HAL
open science

Capital humain et compétences de base des adultes : production et valorisation

Aline Branche-Seigeot

► **To cite this version:**

Aline Branche-Seigeot. Capital humain et compétences de base des adultes : production et valorisation. Economies et finances. Université de Bourgogne, 2013. Français. NNT : . tel-00917578

HAL Id: tel-00917578

<https://theses.hal.science/tel-00917578v1>

Submitted on 12 Dec 2013

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.

UNIVERSITÉ DE BOURGOGNE
Faculté de Sciences Économiques et de Gestion
École doctorale LISIT (Langages, Idées, Société, Institution, Territoires)
IREDU - CNRS

THÈSE
Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ DE BOURGOGNE

Discipline : Sciences Économiques (CNU 05)

Par

Aline BRANCHE-SEIGEOT

Le 26 Novembre 2013

**CAPITAL HUMAIN ET COMPÉTENCES DE BASE DES ADULTES :
PRODUCTION ET VALORISATION**

Directeur de thèse : Jean BOURDON

Co-directeur de thèse : Jean-François GIRET

JURY

M. Jean BOURDON, Directeur de Recherches émérite au CNRS

M. François GARDES, Professeur à l'Université Panthéon - Sorbonne, Paris 1, CES
(rapporteur)

M. Jean-François GIRET, Professeur à l'Université de Bourgogne, IREDU

Mme Danièle GUILLEMOT, Administrateur INSEE, Chef du département de la Formation
et de la Certification, CEREQ

Mme Mareva SABATIER, Professeur à l'Université de Savoie, IREGE (rapporteur).

L'Université de Bourgogne n'entend donner aucune approbation ou improbation aux opinions émises dans cette thèse. Elles doivent être considérées comme propres à leur auteur.

Remerciements

Jamais je n'aurais pu réaliser ce travail de thèse sans les conseils et le soutien d'un certain nombre de personnes. Leur générosité, leur enthousiasme et leur intérêt pour ma recherche m'ont permis d'avancer pleinement dans cette étape délicate « d'apprenti chercheur ». Ces remerciements sont l'occasion pour moi de leur témoigner toute ma gratitude.

En premier lieu, je souhaite remercier mes Directeurs de thèse, Jean Bourdon et Jean-François Giret, pour m'avoir accordé leur confiance en acceptant d'encadrer cette thèse. Je leur suis très reconnaissante pour leur grande disponibilité, leurs conseils éclairés, leurs encouragements et leur respect sans faille des délais serrés de relecture des documents adressés. Je mesure à quel point leur encadrement a été important dans l'aboutissement de cette thèse. Enfin, j'ai été très sensible à leurs qualités humaines d'écoute et de compréhension tout au long de ce doctorat.

Je souhaite ensuite exprimer toute ma gratitude aux membres du jury qui ont bien voulu lire et juger cette thèse. Je remercie notamment Mareva Sabatier et François Gardes pour l'intérêt qu'ils ont témoigné envers mon travail en acceptant la fonction de rapporteur.

Par ailleurs, je tiens à remercier le Ministère de l'Enseignement Supérieur et de la Recherche ainsi que le Conseil Régional de Bourgogne pour m'avoir aidée financièrement au cours de mes études supérieures. Sans ces soutiens financiers, il m'aurait peut-être été plus difficile d'accéder à un troisième cycle universitaire. De même, je remercie l'Université de Bourgogne pour avoir financé cette thèse et pour m'avoir recrutée en tant qu'Attachée Temporaire d'Enseignement et de Recherche. J'ai également une pensée particulière pour Mr Lovat P. qui, le premier, a su me transmettre son goût pour les sciences économiques et sans qui rien de tout cela n'aurait commencé... Enfin, merci à l'IREDU qui m'a permis de réaliser ce travail de thèse dans de très bonnes conditions matérielles et merci aux membres du laboratoire qui m'ont toujours soutenue. Je remercie notamment Elodie pour sa joie de vivre, sa franchise et son exaltation légendaire, ainsi que Mitra pour sa gentillesse, sa délicatesse et son sourire. Je pense aussi tristement à Aliou D. qui nous a malheureusement quittés trop vite. Mes remerciements vont également à Fabienne pour son dynamisme communicatif et son aide dans les démarches administratives.

Je remercie aussi mes amis qui m'ont épaulée durant ce doctorat. Merci à Aurore P. et Anne-Laure S. pour leur présence et pour les agréables moments passés ensemble. J'espère que nous en aurons encore de nombreux comme ceux-ci. Merci à mon acolyte de master, Mathilde L., pour l'intérêt qu'elle a régulièrement su porter à l'égard de mes travaux. Merci à Alexandra M. et Sandrine L., d'avoir toujours été là, entre fous rires et caps difficiles à passer. Votre présence et votre amitié me sont précieuses. Un très grand merci également à Amélie D. et Julien B. pour avoir pris le temps de me connaître. Je vous suis particulièrement reconnaissante de l'aide et des encouragements perpétuels que vous m'avez adressés. Notre connivence aura soudé une bien belle amitié ; je n'aurai jamais autant ri au cours de cette thèse qu'en étant avec vous. Je remercie aussi toutes celles et ceux qui m'auront régulièrement permis de déconnecter le temps d'un repas, d'une soirée

ou d'un week-end. Je pense principalement à Marie A., Thomas R., Amandine P., Stéphane P., Marie M., Bertrand M., Carine N. et Clément B.. Une pensée également pour Ingrid M. que j'ai eu le plaisir de « retrouver ». Et un grand merci à toutes celles et ceux qui se sont portés volontaires pour participer au travail de relecture et qui auront compris ce que représente réellement le travail d'un thésard.

Ces remerciements seraient incomplets si je n'adressais pas quelques mots à ma famille proche. Tout d'abord, un grand merci à mon conjoint Pierre-Yves D. qui a été d'une patience, d'un appui et d'une bienveillance inestimables ces cinq dernières années. Merci de m'avoir supportée au plus près, sans jamais rien me reprocher, et au contraire, en ne cessant de m'encourager. Merci d'avoir toujours cru en moi plus que je n'ai su le faire. Je remercie également mon frère. Notre inégalable complicité a toujours été d'une grande importance pour moi ; j'espère qu'elle perdurera encore longtemps. De même, j'adresse d'affectueuses pensées à celle qui m'a appris à lire, écrire et compter et qui me manque tant aujourd'hui. Enfin et surtout, je remercie ma courageuse mère pour son infini soutien, sa présence inconditionnelle, ses sacrifices, ses valeurs et son éducation. Elle a su me transmettre sa curiosité intellectuelle, son goût d'apprendre et sa persévérance. Sans elle à mes côtés, je n'en serais pas là aujourd'hui. Maman, merci pour tout...

*« A ma mère qui m'a fait croire
... en l'école publique »*

RÉSUMÉ

La question des compétences de base est de plus en plus présente dans les préoccupations politiques actuelles au regard des coûts importants engendrés par l'illettrisme et supportés par la société. L'objectif de cette présente thèse est alors d'apporter un nouvel éclairage sur cette thématique dans une approche microéconomique. Elle se compose de deux parties. La première traite de la valorisation des compétences de base sur le marché du travail, à travers les essais 1 et 2, alors que la seconde traite de la production de compétences en français selon l'origine migratoire, à travers les essais 3 et 4.

Dans le cadre de la première partie, le premier essai s'intéresse à la question des rendements des compétences de base sur le marché du travail, sous l'angle du déclassement/surclassement professionnel et sous l'angle de la rémunération individuelle. L'hypothèse centrale consiste à supposer que l'éducation, via le diplôme, n'est pas un bon filtre des capacités individuelles. Dès lors les situations de déclassement (*surclassement*), pour un même niveau de diplôme, peuvent résulter d'un plus faible (*fort*) niveau de compétences de base et non forcément d'un déséquilibre entre offre et demande de qualifications. De même, si pour un même niveau de diplôme il existe une certaine hétérogénéité des capacités individuelles, nous pouvons penser que l'ajout des scores en compétences de base tend à diminuer le rendement purement scolaire et à améliorer l'explication de la formation des salaires. Notre approche a pour principal avantage de tenir compte d'un éventuel biais de sélection dans l'estimation des équations de gains en appréhendant l'influence exercée par les compétences de base dans l'accès à l'emploi. Le deuxième essai tente quant à lui de mieux comprendre l'impact des compétences de base sur l'accès à l'emploi mais aussi sur la participation au marché du travail et ce, dans un contexte d'enrichissement généralisé du contenu du travail. La prise en compte de ces compétences doit améliorer l'explication des situations occupées à l'égard du marché du travail (emploi, chômage, inactivité), notamment pour des publics déjà fragilisés face à l'emploi. Plus précisément, pour une demande de travail donnée, nous supposons qu'à chaque situation correspond un niveau de compétences de base.

La lecture, l'écriture, la compréhension écrite et orale du français étant des compétences de base clés pour l'intégration sociale et professionnelle en France, la deuxième partie s'intéresse à la production de ces compétences auprès des populations d'origine immigrée. En effet, la France possède une longue tradition d'accueil des migrants. Leur place dans la société française et leurs capacités à apprendre le français suscite donc un certain intérêt, qu'ils soient primo-arrivants ou enfants d'immigrés, d'autant que peu de recherches ont été réalisées à ce sujet. Dans le troisième essai, nous cherchons à déterminer dans quelle mesure les caractéristiques liées à l'origine migratoire entrave la production de compétences en français et dans quelle mesure la part des inégalités de scores en français sont attribuables à des différences d'efficacité résultant de la distance linguistique entre la langue maternelle et le français. Le quatrième et dernier essai vise, quant à lui, à déterminer les facteurs susceptibles d'expliquer le niveau de maîtrise en français écrit des primo-arrivants (considérés comme producteurs de leurs propres compétences) en âge de travailler. L'originalité de cet essai réside dans l'estimation d'une frontière stochastique de production afin d'expliquer leurs scores en français écrit et leurs performances productives.

MOTS CLÉS : Capital humain, compétences de base des adultes, déclassement, rendements, employabilité, participation au marché du travail, production de compétences, immigrés, distance linguistique.

ABSTRACT

Basic skills are important in current policies because of high costs generated by illiteracy and supported by the society. The aim of the present thesis is to bring a new contribution on this thematic in a microeconomic approach. It consists of two parts. The first one concerns the valorization of basic skills on the labor market, through chapters 1 and 2, while the second one concerns the production of French skills according to the migratory origins, through chapters 3 and 4.

Chapter 1 is about the returns to basic skills on the labor market, in terms of overeducation/undereducation and in terms of wages. The central hypothesis is to assume that education, through the diploma, is not a good filter for individual capacity. Therefore overeducation (undereducation) for the same level of degree, can result from a low (high) level of basic skills and not necessarily an imbalance between supply and demand for skills. Similarly, for a given level of education, there is some heterogeneity in individual capabilities, so we think that the addition of basic skills in the wage function tends to reduce the return of education and to improve the explanation of the wage formation. In order to avoid a possible selection bias in the estimation of wages equation, our approach has the main advantage to take into account the impact of basic skills in job access. Chapter 2 attempts to have a better understanding of the impact of basic skills on job access, but also on the labor force participation, in a general enrichment context of the work content. The inclusion of these skills should improve the explanation of the situations occupied towards the labor market (employment, unemployment, inactivity), especially for groups already weakened in the job access. More precisely, for a given labor demand, we assume that for every situation corresponds a level of basic skills.

Reading, writing, listening and reading comprehension skills in French language being essential for a successful integration in French society and labor market. The second part of this thesis focuses on the production process of these basic skills in French among immigrants. Indeed, France has a long tradition of welcoming migrants. Their place in French society and their ability to learn its language generates some interest because little research exist about it. Chapter 3 is where we try to determine to what extent the characteristics associated with the migration background hinders the production of French skills, and to what extent the share of inequality in French scores are due to differences in efficiency because of the linguistic distance between the mother tongue and French. Chapter 4, at last, seeks to determine what are the factors that can explain the proficiency level in written French skills of newcomers (considered as producers of their own skills) in working age. The originality of this chapter is the estimation of a stochastic production frontier to explain the scores in written French and the productive performance.

KEY WORDS : Human capital, basic skills of adults, overeducation, private return to basic skills, employability, labor force participation, production skills, immigrants, linguistic distance.

SOMMAIRE

INTRODUCTION GÉNÉRALE.....	11
BIBLIOGRAPHIE DE L'INTRODUCTION GÉNÉRALE.....	23

PARTIE I. LA VALORISATION DES COMPÉTENCES DE BASE SUR LE MARCHÉ DU TRAVAIL 27

ESSAI 1

LES RENDEMENTS DES COMPÉTENCES DE BASE SUR LE MARCHÉ DU TRAVAIL 28

1. Introduction.....	28
2. Le niveau des compétences de base peut-il expliquer le déclassement ou le surclassement sur le marché du travail ?.....	30
3. La rémunération des compétences de base : fonctions de gains.....	45
4. Conclusion	68
BIBLIOGRAPHIE DE L'ESSAI 1	71
ANNEXES DE L'ESSAI 1	76

ESSAI 2

LE RÔLE DES COMPÉTENCES DE BASE DANS LA PARTICIPATION AU MARCHÉ DU TRAVAIL ET L'EMPLOYABILITÉ 97

1. Introduction.....	97
2. Recherche d'emploi et capital humain.....	99
3. Compétences de base et marché du travail	102
4. Méthodologie : modèle logistique multinomial ordonné généralisé.....	106
5. Statistiques descriptives.....	108
6. Résultats.....	113
7. Conclusion	123
BIBLIOGRAPHIE DE L'ESSAI 2	125
ANNEXES DE L'ESSAI 2	128

PARTIE II. LA PRODUCTION DE COMPÉTENCES EN FRANÇAIS SELON L'ORIGINE MIGRATOIRE.....135

ESSAI 3

CAPITAL HUMAIN ET DISTANCE LINGUISTIQUE : UNE INÉGALE ACQUISITION DES COMPÉTENCES EN FRANÇAIS SELON L'ORIGINE MIGRATOIRE.....136

1. Introduction.....	136
2. Choix migratoires et capital humain.....	138
3. Distance linguistique et inégale acquisition de la langue du pays d'accueil	141
4. Objet d'étude et hypothèses.....	143
5. Statistiques descriptives.....	144
6. Fonctions de production de compétences en français.....	149
7. Décomposition Oaxaca (1973) et Blinder (1973).....	152
8. Décomposition des scores selon Juhn, Murphy et Pierce (1993).....	158
9. Conclusion	162
BIBLIOGRAPHIE DE L'ESSAI 3	164
ANNEXES DE L'ESSAI 3	168

ESSAI 4

LA PRODUCTION DE COMPÉTENCES EN FRANÇAIS ÉCRIT PAR LES IMMIGRES DE PREMIÈRE GÉNÉRATION EN ÂGE DE TRAVAILLER.....172

1. Introduction.....	172
2. Les facteurs de production de compétences.....	175
3. Les déterminants des performances productives : hypothèses.....	178
4. Spécification du modèle de frontière stochastique de production	181
5. Statistiques descriptives.....	186
6. Résultats.....	188
7. Conclusion	197
BIBLIOGRAPHIE DE L'ESSAI 4	199
ANNEXES DE L'ESSAI 4	202

CONCLUSION GÉNÉRALE.....205

BIBLIOGRAPHIE DE LA CONCLUSION GÉNÉRALE 134

TABLE DES MATIÈRES212

LISTE DES TABLEAUX215

LISTE DES FIGURES.....218

LISTE DES ANNEXES219

INTRODUCTION GÉNÉRALE

Si de nombreux travaux ont déjà démontré l'impact du niveau de capital humain sur la croissance économique d'un pays, ils se sont généralement limités au nombre moyen d'années d'éducation. Or, comme le rappellent Hanushek et Woessmann (2012), utiliser cette variable comme proxy des capacités individuelles suppose qu'une année d'éducation équivaut à une augmentation identique du stock de savoirs et de compétences¹ d'un pays à un autre, indépendamment du système d'éducation, ce qui est loin d'être réaliste. De même, l'éducation formelle serait la principale source d'éducation et les différences de qualité des facteurs non scolaires n'auraient qu'un effet négligeable, ce dont nous pouvons douter. De nombreuses enquêtes nationales ou internationales se sont alors développées pour mesurer et comparer la qualité des apprentissages et non plus seulement la quantité d'éducation acquise par les individus (Mons, 2008). Les premières sont nées au début des années 90, notamment dans les pays de l'OCDE, et visaient à quantifier les compétences de base des adultes. Parmi elles, se trouve la « National Adult Literacy Survey » (NALS), réalisée en 1992 aux États-Unis sur un échantillon de 26 000 individus âgés de plus de 16 ans. Il est ressorti de cette enquête que près de 22% des Américains possédaient seulement le premier niveau de compétences de base, soit

¹ La notion de compétences a été initialement utilisée dans le monde du travail et de la formation professionnelle (Romainville, 2001). Présente dans les premières enquêtes sur les compétences de base des adultes, elle est ensuite montée en puissance via des enquêtes internationales comme l'enquête PISA. Selon l'OCDE, la compétence est un système spécialisé d'aptitudes, de maîtrise et de savoir-faire nécessaires pour atteindre un objectif donné (OCDE, 2001). Il existerait une complémentarité entre savoirs et compétences, en cela que les compétences intègrent la capacité à utiliser et à mobiliser des savoirs et des connaissances dans une situation donnée (Roegiers, 1999). Cette caractéristique rejoint la définition la plus courante proposée par certains sociologues du travail à la fin des années 80 (Paradeise, 1987 ; Dadoy, 1989). Ce ne sont plus tant les savoirs en eux-mêmes qui importent que l'utilisation qui en est faite. Derrière le terme compétence se cache alors l'idée de « potentiel » à agir efficacement dans un contexte parfois complexe et imprévisible, en raison des rapides mutations technologiques.

le niveau le plus faible. Au regard de ce chiffre important, ce type d'enquête a ensuite été généralisé pour être conduit simultanément dans plusieurs autres pays membres de l'OCDE. Il est notamment question des enquêtes IALS (« *International Adult Literacy Survey* ») et ALL (« *Adult Literacy and Lifeskills* »). L'importance des enjeux économiques et sociaux liés à la maîtrise d'un niveau minimal de compétences de base explique pourquoi les gouvernements des pays développés ont eu à cœur de mieux quantifier l'ampleur de l'illettrisme², notamment par une évaluation objective des compétences de base des adultes. Une meilleure appréhension de ce phénomène doit en effet permettre de mieux cibler les politiques adéquates à mettre en place pour lutter contre et ainsi diminuer les conséquences et les coûts liés à l'exclusion économique et sociale qui en résulte bien souvent.

La question des savoirs cognitifs réellement acquis par les individus a fait l'objet d'une abondante littérature dans les recherches en économie de l'éducation. Hanushek et Kimko (2000) ont été parmi les premiers à intégrer dans les modèles de croissance économique les résultats que les élèves ont eu à différents tests internationaux sur les acquis en mathématiques et en sciences³. Les deux auteurs ont construit une mesure standardisée de la qualité de la force de travail pour 31 pays, couvrant la période 1960-1990, pour mieux estimer la contribution de la qualité du capital humain à la croissance économique. Ainsi, d'après leurs résultats, le niveau d'éducation ne semble pas être une variable très stable puisqu'elle perd de son importance et de sa significativité lorsque les scores en compétences de base sont inclus. En revanche, ces scores améliorent le pouvoir explicatif du modèle de croissance économique : la quantité moyenne d'éducation reflèterait donc assez mal le niveau moyen de capital humain, ce que de nombreux autres travaux ont également mis en évidence ces dernières années (Altinok, 2007 ; Barro, 2001 ; Hanushek et Wössmann, 2008).

Dans le cadre d'économies ouvertes, Jamison et al. (2007) expliquent quant à eux que l'impact des compétences de base sur la croissance (en particulier celles en mathématiques) serait principalement dû à l'impact que ces compétences ont sur

² L'illettrisme désigne l'état d'une personne qui, malgré une scolarisation, n'a pas acquis (ou a perdu) une maîtrise suffisante de la lecture, de l'écriture, du calcul et de la compréhension écrite et orale, pour être autonome dans les situations simples de la vie quotidienne. L'illettrisme se distingue de l'analphabétisme, qui désigne une maîtrise insuffisante des compétences de base en raison d'une absence de scolarisation.

³ Il s'agit principalement des enquêtes IEA (International Association of the Evaluation of Educational Achievement) et IEAP (International Assessment of Educational Progress).

l'accumulation de progrès technologiques. Ce résultat est particulièrement intéressant au regard de l'essor rapide des nouvelles innovations technologiques dans les pays développés. D'après Eurostat, entre 1999 et 2004, l'emploi dans les services à forte intensité de connaissance de haute technologie a augmenté de 2,9% dans l'Union Européenne des 15, alors que l'emploi dans l'ensemble des services n'a augmenté que de 1,9%. Dans ce contexte de fortes innovations technologiques et compte tenu de la conclusion de Jamison et al. (2007), la lutte contre l'illettrisme doit donc permettre d'accompagner au mieux les changements structuraux de l'économie.

Les décideurs publics semblent l'avoir compris. A l'échelle européenne, plusieurs actions ont été entreprises pour consolider les compétences de base de la main d'œuvre de demain. Par exemple, les ministres de l'éducation de l'Union Européenne se sont fixés comme objectif commun de réduire la proportion de jeunes de 15 ans ayant de faibles compétences de base de 5 points de pourcentage. Il s'agit de la faire passer de 20 % en 2012 à 15 % d'ici à 2020. Si certains États membres de l'Union Européenne ont déjà œuvré pour l'amélioration des niveaux moyens de compétences de base, d'autres sont un peu plus à la traîne. La Slovénie, la Slovaquie, l'Italie, l'Allemagne et la Hongrie ont réussi à augmenter le niveau des compétences de base chez les 8-10 ans entre 2001 et 2006 (Mullis et al., 2007). De même, la Lettonie, le Portugal, le Luxembourg et la Pologne ont considérablement amélioré le niveau moyen de lecture des jeunes de 15 ans ces dix dernières années (OCDE, 2010). En revanche, la France et les Pays-Bas sont quant à eux en retard dans les actions publiques, bien qu'une prise de conscience ait récemment eu lieu. En effet, des campagnes soutenues se sont dernièrement multipliées afin de sensibiliser le public au phénomène de l'illettrisme. Suite aux derniers chiffres dévoilés en décembre 2012, ce phénomène est même devenu « *grande cause nationale 2013* » en France. Agir rapidement contre ce dernier devient pressant. Selon Hanushek et Wössmann (2012) les pays qui améliorent le niveau moyen de compétences de base de leur population noueront avec une croissance plus soutenue sur le long terme.

*

Thèse : présentation

*

Pour générer de la croissance, les pays ont davantage intérêt à améliorer le niveau de compétences de base des plus faibles que de générer des diplômés hautement qualifiés ou encore d'investir dans le capital physique. D'après Coulombe et Tremblay (2005), si le score moyen de la population canadienne en littératie augmentait de 1%, alors la productivité du travail augmenterait de 2,5% et le PIB/Habitant de 1,5%. Cela rejoint les travaux de Lucas (1988). Lucas considère en effet qu'un salarié, quel que soit son niveau de diplôme, sera plus productif dans un environnement où la dotation moyenne de la main d'œuvre en capital humain est élevée. Un niveau homogène de capital humain engendre ainsi des externalités positives sur le système productif⁴. Lucas (1988) insiste sur cet aspect : même si, selon lui, le choix de se former est individuel, investir dans la formation aurait un fort impact au niveau macro-économique via l'élévation de la productivité globale des facteurs que cela provoquerait. Dans ce contexte, une élévation générale et un stock relativement homogène du niveau de compétences de base associé à chaque niveau de diplôme doit favoriser la croissance économique.

Pour autant, dans un contexte de mondialisation des flux d'informations et d'économies du savoir, les compétences de base n'en restent pas moins indispensables dans la vie quotidienne et professionnelle des individus eux-mêmes. Parce qu'il existe des compétences générales très prisées par les employeurs et parce que celles-ci conditionnent l'acquisition de compétences plus spécifiques, elles semblent indispensables dans le potentiel d'employabilité des individus, indépendamment du niveau de diplôme. Ainsi, selon Pont et Werquin (2002) : « *les nouvelles technologies de l'information et de la communication en général, ont remis l'écrit au goût du jour [...] Avec les nouveaux modes d'organisation du travail, qui entraînent des hiérarchies plus plates par exemple, il faut un niveau moyen général homogène dans l'ensemble de l'entreprise et non plus seulement des compétences de haut niveau dans la partie supérieure de la hiérarchie. On demande partout, aujourd'hui, un élargissement des compétences de base* ». Le problème est qu'un nombre important d'individus sont sortis du système éducatif en ayant peu ou mal

⁴ Il s'agit d'une externalité car le stock de capital humain d'un individu ne peut impacter sensiblement le niveau moyen de capital humain.

acquis ces compétences de base et se retrouvent en situation d'illettrisme à l'âge adulte. D'autres ne sont aucunement illettrés mais ont des difficultés de maîtrise de la lecture, de l'écriture et de la compréhension du français de par leurs origines migratoires. En effet, les immigrés ont bénéficié d'apprentissages propices au développement de compétences de base, mais ces apprentissages n'ont peut-être pas été en langue française. Les origines migratoires expliquent en ce sens pourquoi certains individus sont moins familiarisés que d'autres à la langue française et pourquoi ils en ont une maîtrise fragile, bien que hautement qualifiés pour certains. Il s'agit pourtant de compétences de base indispensables pour une bonne intégration professionnelle et sociale en France. D'un point de vue macroéconomique, si cet enjeu relève d'une certaine importance, c'est parce que la migration semble particulièrement favorable à l'activité économique française (D'Albis et al., 2013).

Ainsi, même s'il est possible de vivre avec une faible maîtrise de la lecture, de l'écriture, de la compréhension écrite et orale et du calcul, le coût lié à une telle situation peut être élevé, tant pour la société que pour l'individu lui-même. Les coûts individuels sont souvent estimés au regard des sanctions du marché du travail, en termes de situations (déclassement, nonaccès à l'emploi, retrait du marché du travail) et de salaires. Cependant, certains groupes, bien que n'ayant pas un bas niveau de qualification ou des difficultés cognitives particulières, peuvent être davantage sanctionnés que d'autres, comme les immigrés qui n'ont pas bénéficié des mêmes contextes d'apprentissage que les natifs. Des études ont effectivement mis en évidence, toutes choses égales par ailleurs, que la maîtrise de la langue du pays d'accueil avait un impact particulièrement important sur la probabilité de trouver un emploi et sur les salaires (Dustmann et Fabbri, 2003 ; Shields et Wheatley Price, 2002). Compte tenu de ce constat et de l'enjeu que constitue la maîtrise de la langue du pays d'accueil dans le processus « d'assimilation » des migrants, le mécanisme productif de compétences en langue française revêt donc d'un certain intérêt, d'autant que rares sont les travaux qui se sont penchés sur cet aspect. De fait, quelles sont les caractéristiques individuelles qui favorisent ou non la production de compétences en français écrit et/ou oral par des populations qui n'ont pas hérité du même environnement linguistique, culturel, voire éducatif, que les français natifs ?

Dans cette thèse, nous nous interrogeons donc sur ces deux aspects différents dans le cadre d'une approche microéconomique : la valorisation des compétences de base

sur le marché du travail (partie I) et la production de compétences en français selon l'origine migratoire (partie II).

*

Partie I

*

Le niveau du diplôme est associé à un niveau de résolution des problèmes et un diplôme plus élevé semble favoriser l'adaptation des individus à la complexité et aux aléas des situations rencontrées (Bennaghmouch et Gangloff-Ziegler, 2007). Ainsi, la complexité grandissante des postes pousse les employeurs à augmenter leurs exigences en termes de diplômes, ces derniers présageant d'un certain niveau de compétences. Cela rejoint le modèle de concurrence pour l'emploi (Thurow, 1975) selon lequel les entreprises cherchent à employer les individus qui présentent les plus hautes compétences et donc les plus faibles coûts de formation dans l'emploi. Pour cela, les recruteurs font alors appel au diplôme comme indicateur de l'aptitude à être formé. De fait, dans un contexte de rationnement des emplois, les plus diplômés auront donc plus de chance d'être embauchés parmi l'ensemble des demandeurs d'emploi. Néanmoins, le lien entre diplôme et capacités individuelles est à notre avis discutable, le niveau de diplôme ne garantissant pas forcément du niveau de compétences des individus, en particulier pour les détenteurs de diplômes répandus. Certains peuvent être détenteurs d'un haut diplôme sans pour autant posséder les normes de savoirs, même de base, attendus à ce niveau d'éducation.

Dans une première partie de thèse, composée de deux essais, nous émettons donc comme hypothèse principale que, pour un niveau de diplôme donné, la prise en compte des compétences de base contribue à expliquer certaines situations vis-à-vis du marché du travail et de l'emploi. En effet, une des principales conséquences des progrès technologiques sur l'emploi a été l'élévation progressive de la structure des qualifications recherchées et demandées par les employeurs (Autor et al., 1998 ; Card et Dinardo, 2002). Pour répondre aux besoins du marché du travail et consolider les savoirs de base d'une société tout entière, la tactique des pouvoirs publics a été d'accroître le niveau moyen d'éducation de la population, en l'encourageant à se former. Toutefois, l'aspect quantitatif s'est vite révélé secondaire, l'élévation du niveau moyen d'éducation ne garantissant pas pour autant une élévation homogène du niveau de compétences à travers la population. Si

d'après les théories standards en économie de l'éducation, les compétences sont légitimement indissociables du stock d'éducation reçu, une autre hypothèse est au contraire de considérer que le diplôme filtre imparfaitement les compétences des individus sur le marché du travail. Approximer le niveau de compétences par le niveau de diplôme serait alors assez réducteur de la réalité. De surcroît, la non-prise en compte de l'hétérogénéité des compétences individuelles pour un niveau de diplôme donné entraînerait une surestimation de la contribution purement scolaire (Hause, 1972). Ainsi, la corrélation positive mais imparfaite entre le niveau de compétences de base et le niveau de diplôme (Place et Vincent, 2009) nous amène à penser que le diplôme ne filtre pas forcément bien le niveau de compétences des individus et par conséquent, que l'impact des diplômes sur les « *outcomes* » du marché du travail est sans doute surestimé. Les compétences de base seraient donc des prédicateurs de la productivité individuelle complémentaires aux diplômes. Confrontées au marché du travail, elles doivent alors contribuer à expliquer les situations de déclassement professionnel, la formation des salaires, l'accès à l'emploi ou encore la participation au marché du travail.

Si l'on suppose que l'influence de l'évolution technologique est biaisée en faveur des compétences, alors ces dernières auraient une certaine « rentabilité » sur le marché du travail. C'est l'objet de notre premier essai. L'analyse de cette rentabilité sera envisagée sous deux angles différents : les situations de déclassement et de surclassement et les salaires individuels. Concernant les situations de déclassement et de surclassement, au regard des preuves empiriques existantes et en critiquant la théorie du filtre, nous faisons l'hypothèse que l'hétérogénéité des compétences de base acquises au cours du parcours scolaire et professionnel peuvent en partie expliquer les situations de déclassement ou de surclassement, et ce, pour un même niveau de diplôme. Nous suggérons ainsi que les compétences de base peuvent être de bons indicateurs de la capacité à être formé dans un emploi et peuvent expliquer la position des individus dans les files d'attente pour accéder aux emplois les plus qualifiés. Concernant les niveaux de rémunérations, nous cherchons à savoir dans quelle mesure ils s'expliquent aussi par le niveau en littératie, numératie et compréhension orale. En effet, d'après une étude américaine de Raudenbush et Kasim (1998), les rendements salariaux de la littératie augmentent avec l'intensité du savoir inhérente aux emplois. Si les individus les moins compétents de leur catégorie de diplôme sont amenés à occuper des emplois moins qualifiés que ce que laissait pourtant présager leur niveau de diplôme, alors nous pouvons supposer que

les compétences de base auront un impact sur les salaires individuels, les emplois occupés étant moins intenses en termes de savoirs et de compétences. Notre approche aura pour principal avantage de tenir compte d'un éventuel biais de sélection dans l'estimation des équations de gains, tout en appréhendant l'influence exercée par les compétences de base dans l'accès à l'emploi.

Mais si l'accès à l'emploi peut être expliqué par différents facteurs dont le niveau de compétences de base, qu'en est-il de la participation au marché du travail en elle-même ? Dans la théorie de l'offre de travail, le chômage n'existe pas et l'inactivité résulte d'un choix rationnel des agents. En revanche, dès lors que l'on se place dans un contexte de rigidités, il est nécessaire de faire la distinction entre « chômage » et « inactivité ». En effet, des individus n'ont pas d'emploi alors qu'ils le souhaiteraient et qu'ils en recherchent un activement : de par leurs démarches pour retrouver du travail, ils sont donc encore sur le marché du travail, ce qui rejoint la définition statistique du terme « chômeur ». A l'inverse, certains individus ne recherchent plus d'emploi et se retirent du marché du travail sans que cela ne relève d'un choix rationnel : leur « inemployabilité » les a amenés à se décourager au point de devenir « inactifs ». Nous supposons que parmi eux se trouvent les moins compétents, en termes de « *basic skills* ». Dans un contexte de segmentation des emplois, les individus les moins compétents sont souvent assignés à des emplois non-qualifiés desquels il serait d'autant plus difficile de sortir, notamment en raison d'un faible accès à la formation continue. Néanmoins, les recruteurs accordent de plus en plus d'importance aux compétences de base dans le cadre d'un enrichissement généralisé du contenu du travail. De fait, les individus en emploi non-qualifié ne sont pas forcément des moins aptes. Répondant aux critères de recrutement de certains emplois, ils ne ressentent pas le même découragement à l'égard du marché du travail que pourraient au contraire ressentir les moins compétents. Ainsi, dans quelle mesure les compétences de base expliquent-elles les situations occupées à l'égard du marché du travail (emploi, chômage, inactivité), notamment pour des publics déjà fragilisés face à l'emploi ? Afin de répondre à cette question, nous supposons plus précisément qu'à chaque situation correspond un niveau de compétences de base, et ce, pour une demande de travail donnée. Les plus compétents sont donc sensés avoir une probabilité plus élevée d'être en emploi tandis que les moins compétents ont plus de risques de se retirer du marché du travail. C'est l'objet du deuxième essai.

*

Partie II

*

Dans le cadre d'un enrichissement général du niveau de compétences propre à chaque emploi, qualifié ou non, la maîtrise des compétences de base par les individus d'origine immigrée est bien entendu une question importante. En effet, la migration constitue une arrivée de populations dotées de caractéristiques différentes de celles de la population native tant en termes d'âge, de patrimoine, de culture que de capital humain ou encore de langues. Or, parmi les compétences de base se trouvent précisément l'écriture, la lecture, la compréhension écrite et orale de la langue du pays d'accueil, en l'occurrence du français dans notre cas précis. La place des immigrés en France et leurs capacités à apprendre la langue française suscitent donc un certain intérêt en particulier pour les primo-arrivants, c'est-à-dire les individus nés à l'étranger et de parents étrangers. Pour autant, les enfants d'immigrés demeurent une population d'intérêt puisqu'ils évoluent bien souvent dans un environnement familial qui, par ses caractéristiques (lieu d'habitation, langue parlée à la maison...), semble moins favorable à la réussite scolaire et au développement des compétences en français. Dans cette deuxième partie, également composée de deux essais, il s'agit donc de déterminer quels sont les facteurs qui favorisent la production de compétences en français chez les immigrés.

Dans l'essai 3, nous nous interrogerons sur l'effet des caractéristiques liées à l'origine migratoire sur la production de compétences en français écrit et oral par les immigrés, qu'ils soient primo-arrivants ou enfants d'immigrés. Nous supposons que les immigrés d'origine latine sont *a priori* plus à même de produire de la compétence en français que les immigrés d'origine non-latine, ce qui peut expliquer les inégalités de scores en français écrit et oral entre les deux groupes considérés (immigrés d'origine latine // immigrés d'origine non-latine). En effet, la transférabilité des compétences linguistiques sera d'autant plus réussie que la distance linguistique entre la langue du pays d'accueil et la langue maternelle est faible (Chiswick et Miller, 2012). De fait, nous nous demandons quelle est la part des inégalités de scores en français entre les deux groupes tantôt attribuable à une dotation différenciée en caractéristiques, tantôt attribuable à des différences d'efficacité résultant de la distance linguistique entre la langue maternelle et le français.

Si la distance linguistique peut en partie expliquer les différences de performances productives, d'autres éléments sont susceptibles d'intervenir. Nous supposons plus précisément que les primo-arrivants disposent de deux principaux facteurs de production : la durée de scolarité et la durée de temps de présence dans le pays d'accueil. Le contexte familial et l'environnement dans lequel vit l'individu doivent, en revanche, expliquer sa capacité à convertir ces « inputs-temps » en compétences en français (Becker, 1965). Le quatrième essai vise ainsi à déterminer quels sont les éléments susceptibles d'expliquer les performances productives individuelles. En d'autres termes, quels sont les déterminants plus ou moins favorables à une bonne utilisation des ressources disponibles des immigrés pour produire de la compétence en français, notamment en français écrit ? Dans cet essai 4, nous nous focalisons sur les primo-arrivants afin d'analyser le processus productif d'une population ayant hérité le moins possible de l'environnement culturel français, dont fait partie la langue française. L'originalité de ce dernier essai réside dans l'estimation d'une frontière stochastique de production afin d'expliquer leurs scores en français écrit (considérés comme producteurs de leurs propres compétences) et leurs performances productives.

*

Données

*

Afin d'explorer ces différentes questions de recherche, nous utiliserons la base de données issue de l'enquête Information et Vie Quotidienne (IVQ), réalisée en 2004 par l'INSEE. Cette enquête porte sur l'évaluation des compétences des adultes de 18 à 65 ans via des exercices et des mises en situation de la vie quotidienne. L'avantage de cette enquête réside dans ses données supposées être fiables. En effet, l'enquête IVQ a pour ambition de répondre aux failles des évaluations internationales sur les compétences des adultes en tenant compte de la plupart des critiques formulées à l'égard des enquêtes IALS et ALL⁵. L'enquête IVQ permet également de connaître de façon plus précise l'étendue de l'illettrisme en France. En effet, l'enquête IALS s'étant soldée par le retrait de la France suite à une vive polémique autour de ses « *très médiocres* » résultats, aucun chiffre fiable n'avait pu être dégagé dans l'Hexagone jusqu'au début des années 2000 (Guérin-Pace et Blum, 2000 ; Bonnet et al. 2001).

⁵ Ces évaluations internationales ont vivement été critiquées, tant du point de vue de leur conception que de leur exploitation (Vallet et al., 2002).

Cela est sans compter que les études réalisées au plan national par l'INSEE portaient sur de simples déclarations de la part des individus interrogés. Autant dire que cette approche reposait sur une trop forte subjectivité pour dégager des chiffres crédibles sur la proportion d'illettrés en France. Dans ce contexte, l'arrivée de l'enquête IVQ est alors apparue comme une véritable avancée, appuyant et poursuivant ainsi les évaluations des compétences de base des jeunes adultes faites dans le cadre des Journées d'Appel de Préparation à la Défense (JAPD), instaurées en 1998.

L'originalité et l'intérêt de l'enquête IVQ résident aussi dans son caractère adaptatif. En effet, les exercices de littératie ne sont pas communs à tous les individus, contrairement à la compréhension orale et à la numératie. Une première série d'exercices de niveau standard est soumise à tous les enquêtés. Ensuite, des exercices plus complexes sont proposés à ceux qui ont été détectés comme n'ayant aucune difficulté face à l'écrit, tandis que des exercices plus faciles sont proposés à ceux qui ont été détectés comme ayant des difficultés. Pour ceux qui n'auraient pu être affectés dans l'une ou l'autre catégorie à l'issue de la première série d'exercices, de nouveaux exercices sont proposés pour trancher la décision d'affectation au regard des nouveaux résultats obtenus. La nature adaptative de l'enquête permet ainsi d'avoir une mesure beaucoup plus fine du niveau d'illettrisme et des divers degrés de difficultés rencontrés par la population française face à l'écrit sans pour autant être en situation d'illettrisme.

Les compétences de base évaluées dans l'enquête IVQ de 2004 sont la littératie, la numératie et la compréhension orale. La littératie est « l'aptitude à comprendre et à utiliser l'information écrite dans la vie courante, à la maison, au travail et dans la collectivité en vue d'atteindre des buts personnels et d'étendre ses connaissances et ses capacités » (OCDE, 2000). Elle se distingue donc de l'alphabétisation, en cela qu'elle va au-delà de la simple question : « Savez-vous lire et écrire ? » La littératie cherche plutôt à répondre à l'interrogation suivante : « Savez-vous bien mettre en pratique vos savoirs de base dans la vie quotidienne ? ». Afin de quantifier le niveau de littératie, plusieurs sous-compétences sont mesurées : la lecture de mots, la grammaire, le vocabulaire ou encore la compréhension de textes simples susceptibles d'être rencontrés dans la vie de tous les jours. Pour cela, les supports choisis sont diversifiés (page de programme de télévision, extraits d'articles de journaux, graphiques, de cartes routières, etc. ...). La notion de numératie, quant à elle, reprend les mêmes principes, mais dans le domaine de la maîtrise des nombres

et du calcul. Dans le cadre de l'enquête IVQ, plusieurs dimensions sont ainsi estimées pendant les différentes épreuves (structures additives et multiplicatives, pourcentages, règle de trois, notions simples de la théorie des ensembles...). Enfin, les individus sont aussi testés sur leur aptitude à comprendre un message oral. Pour cela, l'enquête IVQ propose d'écouter un bulletin d'informations enregistré qui fera ensuite l'objet de plusieurs questions.

L'ensemble de cette thèse repose donc sur l'utilisation de données objectives et non plus subjectives quant au niveau de compétences de base des adultes français, ce qui confère une certaine originalité à l'approche empirique de chaque essai.

BIBLIOGRAPHIE DE L'INTRODUCTION GÉNÉRALE

- ALTINOK N. [2007], Capital humain et croissance: l'apport des enquêtes internationales sur les acquis des élèves, *Économie publique/Public economics*, (18-19), pp. 177-209.
- AUTOR D.-H., KATZ L.-F. & KRUEGER A.-B [1998], Computing Inequality : Have Computers Changed the Labor Market ?, *The Quarterly Journal of Economics*, vol. 113, n°4, pp. 1169-1213.
- BARRO R.-J. [2001], Human capital and growth, *The American Economic Review*, vol. 91, n°2, pp. 12-17.
- BECKER G.-S. [1965], "A Theory of the Allocation of Time", *The Economic Journal*, vol. 75, n°299, pp. 493-517.
- BENNAGHMOUCH S. & GANGLOFF-ZIEGLER C. [2007], Diplôme et accès à l'emploi: le point de vue des employeurs. Communication au Congrès international AREF 2007 (Actualité de la recherche en éducation et en formation), Strasbourg, 28-31 août 2007.
- BONNET G., BRAXMEYER N., HORNET S., LAPPALAINEN H.P., LEVASSEUR J., NARDI, E., RÉMOND M., VRIGNAUD P. & WHITE J. [2001], The use of national reading tests for international comparisons: ways of overcoming cultural bias, Ministère de l'Éducation nationale, DPD Édition diffusion, Paris.
- CARD D. & DINARDO J. E. [2002], Skill-Biased Technological Change and Rising Wage Inequality : Some Problems and Puzzles, *Journal of Labor Economics*, vol. 20, n°4, pp. 733-783.
- CHISWICK B.-R. & MILLER P.-W. [2012], Negative and positive assimilation, skill transferability, and linguistic distance, *Journal of Human Capital*, vol. 6, n°1, pp. 35-55.
- COULOMBE S. & TREMBLAY J.-F. [2005], *Public investment in skills: Are Canadian governments doing enough?* C.D. Howe Institute, Ottawa, ON.
- D'ALBIS H., BOUBTANE E. & COULIBALY D. [2013], *Immigration et croissance économique en France entre 1994 et 2008* (No. 1302). CEPREMAP : Paris.
- DADOY M. [1989], Le retour au métier, *Revue Française des Affaires Sociales*, n°4, pp. 69-102.
- DUSTMANN C. & FABBRI F. [2003], Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, vol. 113, n°489, pp. 695-717.

- GUÉRIN-PACE F. & BLUM A. [2000], The comparative illusion: the international adult literacy survey, *Population: An English Selection*, n°12, pp. 215-246.
- HANUSHEK E.-A. & KIMKO D.-D. [2000], Schooling, labor-force quality, and the growth of nations, *American economic review*, vol. 90, n°5, pp. 1184-1208.
- HANUSHEK E.-A., & WOESSMANN L. [2008], The role of cognitive skills in economic development, *Journal of Economic Literature*, vol. 46, n°3, pp. 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*, vol. 17, n°4, pp. 267-321.
- HAUSE J.-C. [1972], Earnings Profile: Ability and Schooling, *The Journal of Political Economy*, vol. 80, n°3, pp. 108-138.
- JAMISON E.-A., JAMISON D.-T. & HANUSHEK E.-A. [2007], The effects of education quality on income growth and mortality decline, *Economics of Education Review*, vol. 26, n°6, pp. 771-788.
- LUCAS R.-E. [1988], On the mechanics of economic development, *Journal of Monetary Economics*, n°22, pp.3-42.
- MONS N. [2008], L'évaluation des politiques éducatives. Apports, limites et nécessaire renouvellement des enquêtes internationales sur les acquis des élèves. *Revue Internationale de Politique Comparée*, vol. 14, n°3, pp. 409-423.
- MULLIS I.-V.-S., MARTIN M.-O., KENNEDY A.-M. & FOY P. [2007], PIRLS 2006 international report: IEA's Progress in International Reading Literacy Study in primary schools in 40 countries. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- OCDE [2000], *La littératie à l'ère de l'information - Rapport final de l'enquête internationale sur la littératie des adultes*, Ottawa : Statistique Canada.
- OCDE [2001], *Defining and Selecting Key Competencies*, OCDE, Paris.
- OCDE [2010], *PISA 2009 Results : What Students Know and Can Do - Student Performance in Reading, Mathematics and Science (Volume I)*, OCDE, Paris.
- PARADEISE C. [1987], Des savoirs aux compétences: qualification et régulation des marchés du travail, *Sociologie du travail*, vol. 29, n°1, pp. 35-46.
- PLACE D. & VINCENT B. [2009], L'influence des caractéristiques sociodémographiques sur les diplômes et les compétences, *Économie et statistique*, vol. 424, n°1, pp.125-147.

- PONT B. & WERQUIN P. [2002], L'apprentissage des adultes en Suisse et dans d'autres pays de l'OCDE. Une perspective comparative, *Schweizerische Zeitschrift für Bildungswissenschaften*, vol. 24, n°2, pp. 279-307.
- RAUDENBUSH S.-W. & KASIM R.-M. [1998], Cognitive skill and economic inequality: Findings from the National Adult Literacy Surve, *Harvard Educational Review*, vol. 68, n°1, pp. 33-80.
- ROEGIERS X. [1999], Savoirs, capacités et compétences à l'école: une quête de sens. *Forum-pédagogies, mars*, pp.24-31
- ROMAINVILLE M. [2001], Les implications didactiques de l'approche par compétences, *Enjeux*, n°51, pp. 199-223.
- SHIELDS M.-A. & PRICE S.-W. [2002], The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas, *Journal of population Economics*, vol. 15, n°1, pp. 137-160.
- THUROW L.-C. [1975], *Generating Inequality*, Basic Books, New York.
- VALLET L.-A., BONNET G., EMIN J.-C., LEVASSEUR J., ROCHER T., BLUM A., GUERIN-PACE F., VRIGNAUD P., D'HAULTFOEUILLE X., MURAT F., VERGER D. & ZAMORA P. [2002], Enquête méthodologique « Information et Vie Quotidienne, tome 1. Série des documents de travail de la Direction des Statistiques Démographiques et Sociales (No 0202). Paris, INSEE.

PARTIE I

**LA VALORISATION DES COMPÉTENCES DE BASE SUR LE
MARCHE DU TRAVAIL**

ESSAI 1

LES RENDEMENTS DES COMPÉTENCES DE BASE SUR LE MARCHÉ DU TRAVAIL

1. Introduction

L'augmentation de l'âge de la scolarité obligatoire, la création du bac professionnel ou la diversification de l'offre de formation dans l'enseignement supérieur sont autant de facteurs qui ont contribué à une élévation sans précédent du nombre de diplômés ces dernières décennies. En 2009, 44% des générations nées de 1979 à 1983 sont diplômées du supérieur, soit 21 points de plus que les générations nées entre 1961 et 1965 au même âge. Si le diplôme reste un « rempart » contre le chômage (Mazari et al., 2011), il semble pourtant ne plus suffire au regard des difficultés croissantes d'insertion professionnelle. En effet, pour certains économistes, l'augmentation du nombre de diplômés, parallèlement à un rationnement des offres d'emploi, a en réalité contribué à la dévalorisation de l'éducation certifiée et plus précisément des diplômes les plus répandus (Baudelot et Glaude, 1986 ; Lemistre, 2003).

Une autre hypothèse est de considérer que le diplôme filtre imparfaitement les compétences des individus sur le marché du travail, du fait que celles-ci ne s'acquièrent pas uniquement au sein du système éducatif (famille, activités associatives, loisirs, réseau social...), ni par le seul apprentissage du métier et qu'elles s'entretiennent tout au long de la vie. Duru-Bellat (2006) a d'ailleurs

souligné les difficultés croissantes du système éducatif à donner à travers le diplôme une juste idée des capacités individuelles. Si l'on peut supposer que le diplôme fonctionne plus ou moins correctement comme filtre en début de carrière, faute d'information sur les caractéristiques des jeunes, d'autres signaux vont ensuite progressivement se substituer au diplôme pour informer les employeurs des capacités des salariés à s'adapter correctement à un emploi. L'affaiblissement du diplôme en tant que signal d'employabilité (Spence, 1973) suppose effectivement qu'il est de plus en plus difficile pour les employeurs de discerner les travailleurs les plus productifs.

Dans cette perspective, ce premier essai cherche à démontrer que le niveau de compétences de base est susceptible d'améliorer l'explication des situations de déclassement et des salaires. L'entrée massive de ces jeunes diplômés sur un marché du travail assez peu porteur pour l'emploi qualifié a provoqué des phénomènes massifs de déclassement en début de carrière (Forgeot et Gautié, 1997 ; Nauze-Fichet et Tomasini, 2002 ; Guironnet, 2006 ; Giret et al., 2006). Des générations de sortants du système éducatif de plus en plus formés ont rencontré des difficultés croissantes pour obtenir un emploi, correspondant notamment à leur niveau d'études. Cela suppose que, pour un même niveau de diplôme, il existe une certaine hétérogénéité d'acquisition des compétences de base (Steedman, 1999) et que ces compétences sont justement déterminantes pour exercer un emploi. De par leur importance en termes d'accès à l'emploi, elles sont donc également supposées avoir une certaine « rentabilité salariale » sur le marché du travail.

Ce premier essai s'articulera autour de deux parties. Dans la première, coécrite avec Jean-François Giret, nous cherchons à évaluer dans quelle mesure le déficit (surplus) des compétences de base serait une des causes, pour un même niveau de diplôme, du déclassement (surclassement) professionnel. L'idée est *d'apprécier l'effet « compétences »* en termes de situations de déclassement et de surclassement des individus en emploi. Dans la seconde partie, nous cherchons à estimer l'impact du niveau de compétences de base sur le niveau de rémunération. Dès lors, l'idée est de *donner une mesure à l'effet « compétences »* sur le marché du travail en termes de sanction/récompense salariale.

2. Le niveau des compétences de base peut-il expliquer le déclassement ou le surclassement sur le marché du travail ?

A l'instar du chômage, le déclassement peut être considéré comme une forme de sous-emploi liée à un déséquilibre entre l'offre et la demande de diplômés émanant du système productif, résultant alors d'une pénurie d'emplois qualifiés et/ou d'une inflation de titres scolaires. En théorie, un tel déséquilibre ne serait que temporaire, si les générations suivantes ajustaient leur choix éducatif. Pour Freeman (1976), les anticipations de salaires des jeunes diplômés jouent un rôle décisif. Se pose également la question des liens entre les effectifs de chaque filière et des professions correspondantes afin d'éviter des phénomènes d'engorgement (Diébolt, 2001). Or, sur longue période en France, les travaux empiriques indiquent que le déclassement des jeunes progresse (Couppié et al., 2005). S'il se réduit au cours des dix premières années de carrière, chaque cohorte se trouve dans une situation plus défavorable que la précédente à ancienneté comparable.

Si l'on suppose que le diplôme filtre imparfaitement le niveau de compétences de base dans un contexte de dévalorisation des diplômes les plus répandus, une autre explication des situations de déclassement (*surclassement*) réside dans l'hétérogénéité des compétences de base. Dans cette première partie, coécrite avec Jean-François Giret, nous nous interrogeons sur cette relation. Tout d'abord, nous examinerons à partir d'une revue de la littérature, l'influence que peuvent avoir les compétences de base sur le marché du travail, notamment dans un cadre d'information imparfaite sur la productivité des salariés. Ensuite, nous testerons les liens entre compétences de base et déclassement.

2.1. Cadre théorique

Quel rôle spécifique peuvent avoir les compétences de base sur le marché du travail ? Les employeurs n'ont pas nécessairement les mêmes besoins en compétences de base et peuvent donc filtrer différemment les individus. Certaines compétences en numératie sont certainement très utiles pour des emplois très qualifiés dans l'industrie alors que d'autres, comme les compétences en expression

orale, paraissent préférables pour des emplois dans le domaine du commerce ou de la vente. Les modèles d'affectation proposés par Sattinger (1993) peuvent expliquer les difficultés d'appariement dans un cadre où l'offre de compétences proposée par la formation et les besoins en compétences des différents emplois sont hétérogènes. Il semble cependant difficile de réduire la dimension du déclassement à un problème d'allocation de compétences entre l'offre et la demande. La théorie de la file d'attente proposée par Thurow (1975) fournit un autre cadre d'analyse, en prenant en compte le rôle prépondérant de l'emploi dans un marché du travail qui n'est pas concurrentiel. Les caractéristiques des individus, comme les compétences de base, peuvent expliquer le positionnement des individus dans la file d'attente pour accéder à certains emplois ainsi que les éventuels déclassements de diplômés sans pour autant affecter directement la productivité des emplois. L'employeur s'intéresse à "l'aptitude à être formé". Le problème consiste, pour l'employeur, à recruter et à former les individus qui vont générer la productivité marginale nécessaire à l'emploi avec l'investissement minimal en coût de formation. Or, l'employeur ne dispose pas de cette information. Afin de minimiser les coûts, les employeurs trient les individus sur la base de ces coûts de formation. Cela conduit à un classement des individus par coûts de formation croissants et in fine, à une file d'attente. Les travailleurs qui possèdent les caractéristiques indiquant les coûts de formation les plus bas sont les premiers embauchés. Si la position des individus varie dans la file d'attente en fonction des emplois qui leur seront proposés, certaines caractéristiques comme le diplôme sont généralement appréciées par l'employeur et permettent aux candidats d'avoir une position privilégiée dans la file d'attente, même pour des emplois peu qualifiés. A partir de l'enquête OFER sur les recrutements réalisés en 2005 dans 4000 entreprises en France, Moncel (2008) montrait par exemple que, lorsqu'ils embauchaient un débutant diplômé du supérieur sur un poste d'ouvrier qualifié ou d'employé, plus de la moitié des recruteurs jugeaient que le niveau de diplôme de la nouvelle recrue répondait à leurs attentes. Pour les employés, seulement 22% des employeurs estimaient *a posteriori* que ce niveau de diplôme était supérieur à leurs attentes. Alors que ces jeunes seraient considérés comme déclassés dans la majorité des tables de correspondance entre niveaux de diplôme et d'emploi, l'enquête OFER montre une déconnexion assez forte entre l'appréciation *a posteriori* de l'employeur et le déclassement observé par le rapprochement des nomenclatures de diplôme et d'emploi. Dans certaines filières, le diplôme confère en France un avantage relatif pour l'accès à certaines catégories d'emplois. Il ne s'agit pas d'un avantage absolu,

sauf pour quelques professions réglementées, mais d'un avantage relatif, qui dépendra à la fois des compétences des jeunes, plus ou moins bien filtrées par le système éducatif, et des exigences des emplois. Le diplôme peut avoir une valeur d'information très forte lors de l'accès aux premiers emplois mais généralement décroissante au fur et à mesure que les employeurs observent les compétences et la productivité de l'individu (Mc Cormick, 1990 ; Farber et Gibbons, 1996).

Dans le cas d'une situation de pléthore de diplômés déconnectée des besoins du marché du travail (Béduwé et Espinasse, 1995), les individus peu ou pas diplômés, situés à la fin de la file d'attente ont de fortes chances de se trouver au chômage ou dans des trappes à emplois peu qualifiés. D'un point de vue individuel, deux stratégies sont possibles pour sortir de l'emploi non qualifié. La première est la recherche d'une mobilité professionnelle ascendante sur le marché du travail. Lorsque les salariés possèdent un socle minimum de compétences de base non sanctionné par un diplôme, l'accès à certains emplois peut donner l'opportunité d'une progression de carrière et éventuellement des situations de ce que l'on peut appeler surclassement. L'expérience professionnelle, si le contenu des emplois et le mode d'organisation le permettent, peut révéler aux employeurs ces compétences de base et donner l'opportunité au salarié d'évoluer. La poursuite d'études est une seconde stratégie pour permettre aux individus d'améliorer leur position relative dans la file d'attente. Pour Thurow (1975), dans un contexte d'allongement général du niveau d'études, l'éducation devient principalement une nécessité défensive pour maintenir sa situation dans la file d'attente. Cependant, pour que cette stratégie soit rentable, en l'absence de variation de la demande de travail qualifié, il faut qu'elle se justifie par une baisse du coût des études supporté par les individus ou par une baisse de la sélectivité du système éducatif. Il n'est pas certain que la baisse du coût d'opportunité des études en période de crise explique une augmentation des poursuites d'études (Gaini et al., 2012). En revanche, comme le montrent Magnac et Thesmar (2002), il semble que ce soit plutôt la baisse de la sélectivité qui porte en France la responsabilité de la hausse du niveau d'éducation. La persistance, voire l'augmentation du déclassement, serait la conséquence d'une élévation du niveau de diplôme alors que le niveau de compétences de base resterait plus ou moins constant. Cette hypothèse correspond cependant à une version forte de la théorie du signal (Blaug, 1985) où la hausse du niveau d'éducation n'augmenterait pas le niveau de compétences de base, ce qui est évidemment très discutable. Il est également possible que même dans le cas d'une élévation du niveau de compétences

de base, les employeurs privilégient pour chaque niveau de diplôme, les individus qui ont le niveau de compétences le plus élevé lorsqu'il n'y a pas pénurie de recrutements.

2.2. « *Much ado about nothing*⁶ » ?

Peu de travaux empiriques ont abordé dans la littérature économique les causes du déclassement professionnel sous l'angle des compétences de base. Pourtant, il semblerait que cette piste offre de nouveaux éclairages en la matière. Si les individus déclassés gagnent moins que les autres c'est certainement parce qu'ils sont moins compétents que leurs pairs qui sont dans des emplois correspondants à leur niveau de diplôme (Green et al., 2002). Cette piste est d'autant plus intéressante que le coût estimé du déclassement peut être surévalué si les travailleurs déclassés ont effectivement des niveaux moyens de compétences plus faibles que les travailleurs non-déclassés ayant le même niveau de diplôme (Verhaest et Omey, 2012). Dans ce cas, la littérature sur le déclassement ferait finalement « *beaucoup de bruit pour rien* » tel que le Sloane (2003) le stipule.

Parmi les rares études existantes, Pryor et Schaffer (1999) ont montré que les travailleurs américains déclassés ont en général des capacités cognitives inférieures, indépendamment de leur niveau de diplôme, que ceux qui ne sont pas déclassés. Green et al. (1999) ont révélé quant à eux, à partir des données de la « National Child Development Study », que les individus qui avaient de faibles scores en mathématiques à l'âge de 16 ans étaient significativement plus enclins à être déclassés plus tard.

Enfin, d'autres recherches, bien que n'intégrant pas une mesure explicite des compétences, plaident également pour un déclassement lié à l'hétérogénéité du capital humain et non pas à un déséquilibre momentané entre offre et demande de qualifications (McGuinness et al., 2007 ; Chevalier, 2003 ; Chevalier et Lindley, 2009 ; Sohn, 2010). Par exemple, Chevalier (2003) explique qu'il existe deux catégories d'individus « déclassés » : les individus réellement déclassés (*genuinely over-educated*) et les individus déclassés « en apparence » (*apparently over-*

⁶ « *Beaucoup de bruit pour rien* »

*educated*⁷). Ces derniers sont déclassés au regard de leur diplôme et de l'emploi occupé mais pas au regard de leur niveau de compétences, ce qui expliquerait justement pourquoi ils ne sont pas insatisfaits de l'emploi occupé malgré leur déclassement « théorique ». L'insatisfaction à l'égard de l'école ne semble donc pas uniquement s'expliquer par la simple situation de déclassement tel que présumé au départ : le niveau de compétences intervient également. Allen et Van der Velden (2001) soulignent d'ailleurs à ce titre que l'inadéquation entre compétences acquises et compétences requises explique beaucoup plus que le déclassement, l'insatisfaction des salariés et leur intention de rechercher un autre emploi.

Ainsi, des doutes s'élèvent quant aux causes du déclassement. Ce phénomène ne résulterait pas, comme le suggère la théorie néoclassique, d'un déséquilibre momentané entre offre et demande de qualifications. Pour certains, une forte hétérogénéité des compétences expliquerait mieux la persistance du déclassement, notamment si aucune formation continue n'est suivie par ailleurs, et la non-promotion des plus instruits pour un niveau de diplôme donné. Ces résultats ne sont donc pas neutres : les compétences semblent bel et bien être en partie à l'origine du déclassement professionnel, et les sanctions salariales qui y sont liées seraient finalement attribuables aux différents niveaux de compétences.

2.3. Modèle économétrique

Afin d'étudier l'effet des compétences de base dans l'explication du déclassement ou du surclassement, nous estimons dans un premier temps deux modèles probit, un pour le déclassement, un pour le surclassement, de la forme :

$$P\left(Décl_i = \frac{1}{X_i}\right) = F(\alpha + \beta_{LITT}LITT_i + \beta_{NUM}NUM_i + \beta_{CO}CO_i + \beta X_i)$$

$$P\left(Surcl_i = \frac{1}{X_i}\right) = F(\alpha + \beta_{LITT}LITT_i + \beta_{NUM}NUM_i + \beta_{CO}CO_i + \beta X_i)$$

⁷ La question utilisée par Chevalier (2003) concernant la satisfaction des individus à l'égard de leur travail est la suivante : 'How dissatisfied are you with the match between your work and your qualifications ?'.

Où $i = 1, \dots, n$ représentant les actifs salariés occupés de l'échantillon, F la fonction de répartition de la loi normale standard. $Décl$ ($Surcl$) représente le fait que le salarié est déclassé (*surclassé*) ou non dans son emploi, $LITT$, NUM et CO matérialisent quant à eux les scores en littératie, numératie et compréhension orale respectivement⁸. Enfin, X_i est un vecteur de variables qui tient compte à la fois des facteurs socio-démographiques, des trajectoires professionnelles sur le marché du travail ainsi que des facteurs environnementaux.

Par ailleurs, il semblerait que les individus soient d'autant moins satisfaits du rôle joué par l'école que leur situation est à « *juste titre* » déclassée (Chevalier, 2003). Afin d'étudier ces relations, il est possible d'introduire des variables croisant la situation face au déclassement ou au surclassement et le niveau de compétences comme variables explicatives du degré de satisfaction à l'égard du rôle joué par l'école dans la vie professionnelle de l'individu i . Le degré de satisfaction a été appréhendé par la question « *Considérez-vous que l'école vous a été utile pour votre vie professionnelle ?* » de l'enquête IVQ de 2004. Plusieurs réponses, de nature ordinale, étaient possibles : « *pas du tout d'accord* », « *plutôt pas d'accord* », « *plutôt d'accord* », « *tout à fait d'accord* ». Il est indispensable de se placer ici dans le cadre d'un modèle logistique multinomial ordonné, construit sur une régression latente du type :

$$y_i^* = x_i\beta + \varepsilon_i$$

y^* (la probabilité d'appartenir à une catégorie de satisfaction : tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord) n'est pas observée. Il s'agit alors de déterminer des intervalles dans lesquels la variable latente y^* peut se trouver. Comme dans le cas binomial, la modalité de y (le degré de satisfaction à l'égard de l'école) dépendrait directement de la position de y^* par rapport à deux seuils, soit :

⁸ Dans l'enquête IVQ, organisée autour de modules (cf. annexe A), nous disposons d'un score global en numératie et en compréhension orale, puisque les différentes questions posées (ou « items ») ont été communes à tous les individus de l'échantillon. Ces scores ont été ramenés sur une échelle allant de 0 à 100. Chaque item est dichotomique : il prend soit la valeur 1 (réussite), soit la valeur 0 (échec). En revanche, nous n'avons pas de score « global » en littératie à part entière. En effet, afin d'affiner la mesure de l'illettrisme, l'enquête IVQ adapte les exercices de littératie, et donc leur degré de difficulté, en fonction du niveau des individus détecté via les items d'un premier module dit d'orientation. Ce caractère adaptatif de l'enquête pose problème dans la construction d'un score global de compétence en littératie sur une échelle continue (Murat et Rocher, 2009). Afin de créer un score global de littératie, nous avons envisagé trois types d'imputations (cf. annexes B).

$$\begin{aligned}
 y_i &= 0 \quad \text{si } y_i^* \leq 0, \\
 y_i &= 1 \quad \text{si } 0 < y_i^* \leq \mu_1, \\
 y_i &= 2 \quad \text{si } \mu_1 < y_i^* \leq \mu_2, \\
 &\vdots \\
 y_i &= k \quad \text{si } \mu_{k-1} \leq y_i^*
 \end{aligned}$$

Les μ sont des seuils inconnus, définissant la frontière des intervalles, qu'il faut estimer tout comme les coefficients β . Notons qu'une fois estimés, ces derniers ne pourront cependant pas s'interpréter directement : leur signe indiquera seulement le sens dans lequel la probabilité sera modifiée. Ces derniers nous permettront alors juste de dire dans quel sens la probabilité de choisir telle ou telle alternative sera affectée.

Dans notre exemple précis, le degré de satisfaction à l'égard du rôle de l'école joué dans la vie professionnelle dépend de certains facteurs x , observables et mesurables, ainsi que d'autres facteurs inobservables, captés par le terme d'erreur aléatoire ε . Nous supposons que ce terme admet une fonction de répartition $\Lambda(\cdot)$ qui suit une loi logistique entre les observations. Les probabilités de réalisation de y sont alors liées à cette fonction de répartition $\Lambda(\cdot)$. Nous avons alors les probabilités suivantes :

$$\begin{aligned}
 \Pr (y_i = 0) &= \Lambda (\mu_0 - x_i\beta) \\
 \Pr (y_i = 1) &= \Lambda (\mu_1 - x_i\beta) - \Lambda (\mu_0 - x_i\beta) \\
 \Pr (y_i = 2) &= \Lambda (\mu_2 - x_i\beta) - \Lambda (\mu_1 - x_i\beta) \\
 &\vdots \\
 \Pr (y_i = k) &= 1 - \Lambda (\mu_{k-1} - x_i\beta)
 \end{aligned}$$

Afin que toutes les probabilités soient positives:

$$0 < \mu_0 < \mu_1 < \mu_2 < \dots < \mu_{k-1}$$

La réalisation de ce modèle logistique multinomial ordonné et de ses estimations se fera via le logiciel STATA et la commande « *ologit* ».

2.4. La recherche d'un lien empirique entre déclassement, surclassement et compétences de base.

L'objectif de cette partie empirique est de tester les liens entre l'acquisition de compétences de base et le déclassement ou le surclassement professionnel des salariés occupés.

2.4.1. Statistiques descriptives

L'originalité de l'enquête IVQ est de disposer de tests de compétences de base portant sur l'ensemble de la population française, quels que soient l'âge, le niveau de diplôme ou la situation professionnelle des enquêtés. Les données étant transversales, il n'est en revanche pas possible de disposer des effets de cohorte et des effets d'âge. Notre hypothèse est que les scores de compétences de base peuvent expliquer, toutes choses égales par ailleurs, les situations de surclassement et de déclassement que l'on observe sur le marché du travail. Les scores de trois compétences de base sont mesurés : la littératie, la numératie et l'expression orale. Bien que corrélés avec le nombre d'années d'études validées par les individus, ces scores varient assez sensiblement pour un même niveau de diplôme. Ainsi 11,5% des diplômés de CAP/BEP ont un niveau de compétences en numératie inférieur au score moyen des individus sans diplôme. Inversement, 47,9% des diplômés de CAP/BEP ont un niveau de compétences supérieur au score moyen des bacheliers. 38,1% des bacheliers ont un niveau de littératie supérieur au niveau moyen des diplômés du supérieur. Ces premières observations suggèrent que le diplôme n'est pas forcément un filtre parfait des capacités individuelles si elles sont évaluées sur ces compétences de base.

Le déclassement (*respectivement le surclassement*) se définit généralement comme une situation de personnes qui possèdent un niveau de formation supérieur (*respectivement inférieur*) à celui normalement requis pour l'emploi qu'elles occupent. Autrement dit, cela suppose qu'il existe une norme d'équivalence entre la hiérarchie des niveaux d'emploi et des niveaux de formation pour appréhender ces situations de décalage. Différentes tables de correspondance basées sur des

approches normatives, statistiques ou objectives (Hartog, 2000) donnent plusieurs mesures du déclassement. Leur comparaison permet de comprendre la forte hétérogénéité de ces situations en fonction notamment de l'environnement institutionnel, qu'elles portent sur l'ensemble d'une génération (Giret et Lemistre, 2004 ; Lemistre 2010), sur un secteur particulier (di Paola, Moullet, 2012) ou sur des types d'emplois spécifiques (Lizé, 2006). Nous avons repris la grille de correspondance « diplôme-profession » proposée par Affichard (1981), afin de privilégier une mesure relativement stabilisée du déclassement et indépendante de l'enquête IVQ. Cette grille, toujours utilisée, repose sur l'analyse du poste de travail effectuée par des spécialistes et suppose donc, dans une perspective adéquationniste, la mise en relation d'une nomenclature de niveaux de diplômes et de niveaux d'emploi⁹. La modification proposée par Lemistre (2010) a été intégrée : on considère que les bacheliers ne sont pas déclassés lorsqu'ils accèdent à des professions d'employés ou d'ouvriers qualifiés. Au total, pour l'ensemble des individus salariés, 23% des individus sont surclassés et 27,3% des individus sont déclassés.

Tableau 1 : Déclassement et surclassement des salariés à partir de l'enquête IVQ

	Cadre	Profession intermédiaire	Technicien	Employé qualifié	Employé non qualifié	Ouvrier qualifié	Ouvrier non qualifié
Supérieur long	7,9%	4,3%	0,4%	0,6%	0,4%	0,03%	0,1%
Supérieur court	2,4%	5,6%	2,5%	1,6%	1,1%	0,5%	0,1%
Baccalauréat	1,5%	3,7%	1,8%	3,0%	4,0%	1,9%	0,9%
BEP /CAP	0,9%	3,0%	2,9%	2,6%	7,6%	8,8%	3,3%
Bepc - Brevet des collèges	0,4%	1,3%	0,5%	0,7%	2,1%	0,9%	0,5%
Jamais scolarisé, sans diplôme ou CEP	0,2%	0,8%	1,1%	1,2%	6,6%	6,0%	4,0%

Source : calculs de l'auteur ; enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Note de lecture : 7,6% des individus sont détenteurs d'un BEP/CAP et employés non-qualifiés : ils sont considérés comme déclassés d'après l'approche institutionnelle d'Affichard. Inversement, 1,4% des individus ne possèdent que le BEPC ou le brevet des collèges tout en occupant une profession intermédiaire : ils sont alors considérés comme déclassés d'après la même norme.

Il est possible de calculer les scores moyens de compétences de base pour les individus déclassés, surclassés, en fonction de leur niveau de diplôme ou de leur âge.

⁹ Cette grille de correspondance, proposée en France par Affichard (1981) est présentée en annexes dans le tableau de l'annexe C.

Si l'on prend par exemple, les diplômés d'un BEP/CAP (cf. tableau 2), nous remarquons que les scores moyens sont globalement plus faibles pour les individus déclassés que pour les individus dont l'emploi correspondrait à leur niveau de diplôme. Symétriquement, les individus surclassés ont en moyenne des scores plus élevés que les individus bien appariés, ce qui apparaît cohérent avec notre hypothèse de départ.

Tableau 2 : Scores moyens en compétences de base des diplômés d'un BEP/CAP selon leur situation face au déclassement ou au surclassement

	Scores moyens en littératie	Scores moyens en numératie	Scores moyens en expression orale
Diplômés d'un BEP/CAP en état de déclassement	86,9	66,9	79,5
Diplômés d'un BEP/CAP ni déclassé, ni surclassé	86,9	70,8	80,6
Diplômés d'un BEP/CAP en état de surclassement	88,5	74,1	83,2

Source : calculs de l'auteur ; enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Il est également possible d'estimer toutes choses égales par ailleurs, pour l'ensemble de l'échantillon, puis par âge et niveau de diplôme, les déterminants du déclassement et du surclassement, par des modèles probit dichotomiques. L'intérêt est d'identifier un effet net des compétences de base. Des variables individuelles comme le sexe, le lieu d'habitation, les principaux niveaux de diplôme ou la spécialité de formation ainsi que des variables liées à la situation professionnelle sur le marché du travail ont été introduites. Le tableau D.1 porte sur l'ensemble de la population salariée, le tableau D.2 sur la population par tranche d'âge et le tableau B.3 sur la population par niveau d'études (cf. annexes D).

2.4.2. Résultats

Le tableau B.1 confirme « toutes choses égales par ailleurs » un lien entre surclassement et niveau élevé de compétences de base d'une part, et déclassement et déficit de compétences de base, d'autre part. Pour les individus surclassés, un niveau plus élevé dans les trois compétences de base¹⁰ semble avoir compensé

¹⁰ Des estimations séparées pour chaque compétence conduisent à des résultats similaires.

l'absence de diplôme suffisant. L'effet des scores de compétences de base sur le déclassement est tout aussi plus prégnant : un déficit en numératie comme en littératie conduit à un risque plus élevé d'être déclassé. L'analyse par tranches d'âge (tableau D.2) conduit à des constats similaires pour les salariés surclassés, mais des résultats moins significatifs pour le déclassement. Le niveau de numératie a toujours un impact sur le déclassement pour les moins de 40 ans, ce qui peut s'expliquer par l'évolution des emplois et le besoin de ce type de compétences pour s'adapter au NTIC comme le souligne par exemple Bynner (2004). En revanche, les compétences de base ne semblent pas avoir d'effet sur le déclassement des plus de 40 ans. Du fait de la nature transversale des données, il est difficile de trancher entre deux explications, un effet de génération et un effet d'ancienneté. D'une part, dans une logique de filtre, le diplôme des plus anciennes cohortes peut donner une meilleure indication du niveau de compétences de leurs détenteurs, qui étaient moins nombreux. D'autre part, on peut penser que pour les plus expérimentés, l'acquisition d'autres compétences a permis de compenser les déficits de compétences de base. Ces salariés sont entrés sur un marché du travail plus porteur alors que la concurrence entre diplômés était moins importante, contrairement aux dernières cohortes.

Les résultats présentés dans le tableau D.3, détaillés par niveau de diplôme, vont dans le même sens que pour l'ensemble de la population. La probabilité d'être déclassé est plus élevée lorsqu'il existe un déficit en numératie pour les diplômés du secondaire comme pour les diplômés du supérieur. Ces derniers ont également plus de risques d'être déclassés lorsqu'ils ont un niveau plus faible de littératie. Parallèlement, un surplus dans toutes les compétences de base augmente la probabilité d'être surclassé pour les diplômés du secondaire. Cependant, le surclassement des sortants de l'enseignement supérieur ne semble pas lié à un surplus de compétences de base mais peut-être à d'autres compétences comme des compétences managériales qui permettent le passage de technicien à cadre.

D'autres variables de contrôle ont également été introduites pour expliquer le déclassement et le surclassement. Au niveau des variables individuelles, être une femme augmente fortement et significativement le risque d'être déclassé tout comme le fait d'avoir connu une insertion plutôt difficile sur le marché du travail. Avoir un père cadre favorise également l'obtention d'un emploi correspondant à son diplôme, voire même au-delà, sans doute en raison du plus grand réseau

relationnel du père. Les effets de la trajectoire professionnelle passée et du secteur d'activité sont également très prégnants. Ces résultats sont d'ailleurs cohérents avec une logique de signaux permettant de classer les salariés dans une logique de file d'attente pour accéder aux emplois qualifiés. Lorsque la trajectoire professionnelle est plus précaire, les individus sont plus enclins à accepter des postes moins valorisants que ne leur laisse entrevoir leur diplôme. *A contrario*, la probabilité d'être déclassé se réduit lorsque l'expérience professionnelle augmente, ce qui est cohérent avec l'hypothèse de reclassement avec le temps, même si la nature transversale de l'enquête ne permet pas de trancher sur cette hypothèse. Être salarié du privé conduit également à une baisse du risque d'être déclassé. En effet, travailler dans la sphère publique suppose très souvent d'avoir passé et réussi des concours requérant un niveau minimum de diplôme (Di Paola et al., 2005), ce qui conduit certains jeunes à des arbitrages entre stabilité de l'emploi et déclassement. Le rôle des secteurs d'activité structure également fortement les chances de déclassement et de surclassement comme l'ont par exemple montré Nauze-Fichet et Tomasini (2002). Travailler dans l'industrie ou dans le secteur tertiaire marchand semble réduire la probabilité de déclassement par rapport au fait d'être employé dans les domaines de l'éducation, de la santé ou de l'action sociale, domaines relevant principalement de la fonction publique. En revanche, travailler dans l'agriculture, l'agroalimentaire ou le commerce augmente les risques d'un déclassement.

L'effet des compétences de base peut également être observé au sein de chaque grand secteur d'activité économique. L'hypothèse serait que certains secteurs ou certains types d'emploi auraient besoin de plus de compétences de base que d'autres, ce qui est compatible avec la théorie de la file d'attente de Thurow. Le tableau 3 présente l'effet des compétences de base sur le déclassement et le surclassement dans trois grands secteurs ainsi que pour les salariés déclarant un usage fréquent d'un ordinateur sur leur lieu de travail. Les effets des compétences de base sont assez proches selon les secteurs : le niveau en numératie est la compétence la plus déterminante dans le déclassement et le surclassement, quel que soit le secteur considéré, même s'il semble un peu plus fort dans l'industrie et surtout pour les emplois utilisant un ordinateur. Ce résultat rejoint ainsi l'idée selon laquelle la numératie permet d'accompagner les changements structureaux de l'économie, telles les innovations technologiques. La littératie a en revanche un effet plus spécifique dans le tertiaire marchand : un déficit de littératie augmente les

risques de déclassement alors qu'un excédent accroît au contraire la probabilité d'un surclassement.

Tableau 3 : Effets des compétences de base par type de secteurs

VARIABLES	Usage de l'ordinateur au travail		Industrie et construction		Tertiaire marchand		Tertiaire non-marchand	
	Surclassé Coeff.	Déclassé Coeff.	Surclassé Coeff.	Déclassé Coeff.	Surclassé Coeff.	Déclassé Coeff.	Surclassé Coeff.	Déclassé Coeff.
Scores en littératie	0,004	-0,008	0.004	-0.13*	0.010**	-0.012**	0.003	-0.008
Scores en numératie	0,01***	-0,01***	0.012***	-0.01***	0.006***	-0.006**	0.009***	-0.012***
Scores en compréhension orale	0,0009	-0,0005	0.003	-0.005*	0.002	-0.001	0.003	0.002

*** : significatif à 1% ; ** : significatif à 5% ; * : significatif à 10%

Note : Les modèles ont été contrôlés par les facteurs socio-démographiques, la trajectoire professionnelle et les facteurs introduits dans le tableau A présenté en annexes. Seule la catégorie de diplôme « BEPC/Brevet des collèges » a été ôtée des variables explicatives en raison de sa colinéarité potentielle.

Aucune mesure subjective du déclassement n'existe dans l'enquête IVQ. Or, ces mesures permettent de mieux comprendre l'hétérogénéité des phénomènes de déclassement. Comme le souligne Hartog (2000), elles peuvent en partie intégrer les caractéristiques des individus, de son emploi et de son environnement, ce que ne font pas les mesures normatives ou statistiques du déclassement. Chevalier (2003), à partir d'une enquête portant sur les diplômés britanniques¹¹, identifie ainsi deux catégories d'individus « déclassés » : les individus réellement déclassés (*genuinely over-educated*) et les individus déclassés « en apparence » (*apparently over-educated*). Ces derniers sont déclassés au regard de leur diplôme et de l'emploi occupé mais pas au regard de leur niveau de compétences, ce qui expliquerait justement pourquoi ils ne sont pas insatisfaits de l'emploi occupé malgré leur déclassement « théorique ». Dans l'enquête IVQ 2004, une variable permet d'approcher la perception qu'ont les individus par rapport à l'utilité de l'école sur le marché du travail. Contrairement à différentes mesures subjectives du déclassement, cette variable ne reflète pas uniquement la correspondance entre le niveau de formation et le niveau d'emploi : un individu peut, par exemple, être satisfait parce que l'école lui a permis d'avoir un emploi même si cet emploi ne

¹¹ La question utilisée par Chevalier (2003) concernant la satisfaction des individus à l'égard de leur travail est la suivante : 'How dissatisfied are you with the match between your work and your qualifications ?'.

correspond pas forcément à son niveau d'étude. On peut cependant faire l'hypothèse que l'utilité de l'école sur le marché du travail est en partie liée au niveau de valorisation de son diplôme.

Tableau 4 : Pourcentage de surclassés et de déclassés déclarant que l'école a été utile ou non pour leur vie professionnelle

<i>L'école a-t-elle été utile pour la vie professionnelle ?</i>	Surclassés	Correspondance diplôme-emploi	Déclassés
Pas du tout d'accord	8%	6%	4%
Plutôt pas d'accord	14%	10%	9%
Plutôt d'accord	38%	39%	41%
Tout a fait d'accord	40%	45%	46%
<i>Rhô (coefficient de Spearman)</i>	0.03**		

Source : calculs de l'auteur ; enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Le tableau 4 indique que le lien entre l'utilité de l'école dans la vie professionnelle et la correspondance diplôme-emploi est faible, mais significatif. Les salariés surclassés semblent légèrement plus réservés sur l'utilité de l'école. Globalement, 78% des salariés surclassés estiment que l'école a été utile pour leur vie professionnelle contre 87% pour les salariés déclassés et 84% pour ceux qui ont en emploi en rapport avec leur diplôme. Compte tenu du fait que les salariés surclassés sont par définition des individus ayant un niveau de diplôme inférieur à celui qui est requis dans le poste occupé, ils peuvent estimer que s'ils ont réussi à obtenir un poste à un niveau plus élevé, c'est davantage en raison de leur niveau de compétences (cognitives ou non cognitives), voire grâce à un réseau social et professionnel dense, qu'à l'école elle-même. Ce constat est cependant à nuancer puisque plus des trois quarts des salariés surclassés estiment que l'école leur a été utile pour leur vie professionnelle. Résultat plus surprenant, même si les écarts restent tenus, les salariés déclassés semblent aussi satisfaits du rôle de l'école dans leur vie professionnelle que ceux qui ont un emploi correspondant à leur niveau de qualification. Ils peuvent valoriser le fait que leur diplôme leur a permis d'accéder plus facilement à l'emploi même si cela leur a généré un déclassement. Ils peuvent également se satisfaire d'un emploi apparemment déclassé du fait d'un niveau de compétences de base plus faible par rapport à leur diplôme, ce qui serait cohérent avec l'explication proposée par Chevalier (2003).

Afin d'étudier plus précisément ces relations, il est possible d'introduire des variables croisant la situation face au déclassement ou au surclassement avec le

niveau de compétences¹² comme variables exogènes en vue d'expliquer la satisfaction à l'égard de l'école dans un modèle logit ordonné, « toutes choses égales par ailleurs ». Les résultats du tableau présenté en annexe E conduisent une interprétation un peu différente du tableau 4 : indépendamment de la situation des individus à l'égard du déclassement ou du surclassement, être en déficit ou à un niveau normal de compétences¹³ a un impact soit négatif soit neutre vis-à-vis de l'utilité perçue de l'école d'un point de vue professionnel. A l'inverse, un excédent de compétences a un effet soit positif soit non significatif. Les individus déclassés et en déficit de compétences sont parmi les plus critiques par rapport à l'utilité de l'école, ce qui nuance l'hypothèse de Chevalier appliquée aux compétences de base. Mais on observe surtout que les individus surclassés présentant un déficit en numératie sont les plus enclins à trouver l'école moins utile pour leur vie professionnelle. D'après les probabilités prédites estimées suite à ce modèle (cf. annexe F), ils sont les plus nombreux (20%) à estimer que l'école a été peu utile pour leur vie professionnelle. On peut supposer que ces derniers ont pu acquérir ou valoriser différentes compétences dans leur vie professionnelle propices à une évolution positive dans leur carrière. Cependant, ces critiques face à l'utilité de l'école peuvent également s'expliquer par le fait que des individus surclassés mais en situation de déficit de compétences sont susceptibles de se sentir plus « démunis » dans l'exercice de leur emploi. En effet, ils sont à même d'estimer que l'école ne leur a pas fourni les bases nécessaires dans des compétences pourtant indispensables pour exercer efficacement leur emploi. Cette explication rejoint également celle proposée par Allen et Van der Velden (2001) qui estiment que c'est davantage le fossé entre compétences acquises et compétences requises qui explique l'insatisfaction professionnelle ; ceci plutôt que la situation de déclassement, voire de surclassement par rapport à un niveau de diplôme. Concernant les autres variables de contrôle, la satisfaction augmente avec un niveau de rémunération élevé, l'âge et l'exercice d'un emploi dans certains secteurs comme les secteurs traditionnels de la fonction publique pourtant plus soumis au phénomène de déclassement que les autres.

¹² Pour simplifier l'analyse, seul le niveau de littératie a été retenu.

¹³ Nous utilisons ici une méthode statistique pour identifier un niveau « normal », un excédent et un déficit de compétences associé à chaque niveau de diplôme. On fait l'hypothèse que le niveau « normal » de compétences de base se définit par l'appartenance à un intervalle de confiance autour du score moyen. Les individus ayant un score inférieur de plus d'un écart-type (σ) au score moyen (m) correspondant à leur niveau de diplôme sont alors considérés en déficit de compétences. A l'inverse, les individus qui ont obtenu un score supérieur de plus d'un écart-type au score moyen correspondant à leur niveau de diplôme sont considérés comme étant en surplus de compétences.

3. La rémunération des compétences de base : fonctions de gains

Nous avons démontré dans la section précédente que l'hétérogénéité des compétences de base pouvait en partie expliquer les situations de déclassement et de surclassement. Ce résultat nous invite à nous interroger sur le lien « compétences de base - salaires ». En effet, si Boothby (2002) établit qu'à diplôme égal, les individus déclassés gagnent moins que leurs camarades recrutés à leur niveau de diplôme, Pecoraro (2011) montre, en revanche, que cette perte de salaire ne concerne que ceux qui sont « réellement » déclassés. Pour les individus qui ne le seraient qu'en « apparence », la perte salariale propre à la situation de déclassement n'est plus significative et le rendement éducatif devient identique à celui des individus non-déclassés occupant le même type d'emploi. Ainsi, selon Pecoraro (2011), les pertes salariales associées au déclassement « apparent » sont principalement dues à l'omission des compétences dans les fonctions de gains. Ces résultats démontrent donc que la productivité marginale – selon laquelle les individus sont rémunérés d'après la théorie du capital humain – semble davantage être fonction du niveau général de compétences que du niveau de diplôme. Cette deuxième sous-section de l'essai 1 s'appuie donc sur cette idée. Elle a pour ambition de déterminer dans quelle mesure les compétences de base jouent un rôle dans la formation des salaires individuels et si elles améliorent par conséquent l'explication de la variabilité des salaires.

3.1. Revue de littérature

Si les travaux empiriques sur l'impact des compétences de base sur le déclassement sont relativement récents, bien d'autres se sont développés depuis les années 1970 autour de l'effet des capacités cognitives sur les « *outcomes* » du marché du travail. Les premières analyses concernant la relation entre compétences de base et salaires ont porté sur les pays anglo-saxons, au premier rang desquels les États-Unis (Taubman et Wales, 1973 ; Willis et Rosen, 1979 ; Berlin et Sum, 1987). Toutefois, les analyses empiriques ont très vite été limitées à l'étude des populations immigrées (Kossoudji, 1988 ; Rivera-Batiz, 1992, 1999 ; Chiswick et Miller, 1995, 1996 ; Charette et Meng, 1998). Il faudra attendre l'apparition des enquêtes IALS

(International Adult Literacy Survey) au milieu des années 90, testant les compétences de base des adultes à travers les pays de l'OCDE, pour voir s'élargir les travaux de recherche à d'autres pays, parfois dans une logique comparative (Freeman et Schettkat, 2000 ; Devroyes et Freeman, 2001 ; Denny et al., 2004 ; Denny et Doyle, 2010). L'effet positif et significatif de la littératie et de la numératie sur les salaires semble être attesté tant pour les États-Unis (Sum et al., 1986), le Canada (Charette et Meng, 1998), la Grande-Bretagne (Vignoles et al., 2011), l'Australie (Chiswick et al., 2003) ; la Nouvelle-Zélande (Chapple et Maré, 2000) que pour l'Allemagne (Antoni et Heineck, 2012) ou encore le Brésil (Baldini Rocha et Ponczek, 2011). Boissière, Knight et Sabot (1985), en travaillant sur des données kenyanes et tanzaniennes, trouvent, quant à eux, que les travailleurs qui ont les plus hauts niveaux de littératie et de numératie sont les plus productifs. Ils soulignent également que les rendements directs de ces compétences sont plus importants que ceux relatifs au nombre d'années d'études. Les auteurs en concluent que les rendements des capacités cognitives ne sont en réalité qu'une récompense du capital humain, qu'il soit acquis au sein du système éducatif ou non. Plus précisément, l'éducation est précieuse pour les travailleurs, car elle peut leur faire acquérir des compétences qui augmentent leur productivité, mais elle ne semble suffire à elle-même, ce qui rejoint l'idée selon laquelle les diplômés sont des filtres imparfaits des niveaux de compétences des individus.

Parmi les travaux existants, certains ont mis en avant un impact particulièrement positif de la numératie sur les salaires et les taux d'employabilité (Dearden et al., 2002). Ainsi, McIntosh et Vignoles (2001) ont démontré, à partir de données issues de la « British National Child Development Study » et de IALS 1994, que la numératie jouait significativement et positivement sur l'accès à l'emploi des moins diplômés. Lorsque ceux-ci sont en emploi, il est également démontré que les plus compétents en numératie gagnent en moyenne entre 16% et 21% de plus que les autres, toutes choses égales par ailleurs. Parallèlement, Murnane et al. (1995), s'interrogent sur la façon dont les compétences en mathématiques des diplômés du secondaire supérieur peuvent affecter leurs salaires à l'âge de 24 ans. Pour répondre à cette question, les auteurs mènent leur analyse à partir de plusieurs enquêtes, réalisées à des dates différentes. Ils montrent, en contrôlant la réussite scolaire, que les capacités en mathématiques ont un impact positif et significatif sur les salaires, qui serait amplifié avec le nombre d'années d'études. De fait, il y aurait un effet interactif entre les capacités en mathématiques et le nombre d'années

scolaires, tout du moins pour les hommes, puisque ce résultat ne semble pas être vérifié pour les femmes. A noter que cela ne vaut que si les concepts mathématiques ont bien été appris et consolidés dans les premières années d'éducation. Il semblerait également que l'éducation scolaire soit à même de favoriser l'acquisition de compétences de base, davantage que l'expérience professionnelle (Dougherty, 2003). Cela montre l'importance des niveaux d'éducation de base. Certains pourraient expliquer ce résultat par la demande croissante d'individus qualifiés, aptes à s'adapter rapidement aux changements technologiques (Berman et al., 1998 ; Green, 1998).

Néanmoins, à notre connaissance, aucun travail de recherche n'a pu être réalisé sur la valorisation des compétences de base en France, celle-ci s'étant retirée de l'enquête IALS suite aux premiers résultats. En effet, cette enquête a rapidement fait l'objet de critiques, tant dans sa conception que dans son exploitation, remettant ainsi considérablement en cause les résultats obtenus pour certains pays, dont l'Hexagone (Dickes et Vrignaud, 1995 ; Guérin-Pace et Blum, 1999, 2000). Dans ce contexte, comment les compétences de base des adultes sont-elles valorisées en France, en termes d'accès à l'emploi et de salaires ? Tel sera notre objectif avec un jeu de données encore peu exploité jusqu'ici.

3.2 Hypothèses

La théorie du capital humain stipule que la productivité, qui détermine le niveau de salaire, dépend des compétences intrinsèques individuelles, essentiellement acquises par la formation. Ces compétences sont généralement d'ordre cognitif et non cognitif (Heckman et al., 2006). Néanmoins, au regard des preuves empiriques existantes dans la littérature, nous supposons qu'une part de la productivité peut être directement expliquée par les compétences de base des individus. Les omettre des fonctions de gains conduirait ainsi à surestimer les rendements des diplômes ou de la durée de scolarité. Récemment, Patrinos et Sakellariou (2013), ont démontré pour le cas du Chili, que l'inclusion d'une mesure directe des compétences de base réduit le rendement de la scolarité de 27%, soit l'équivalent de deux années d'études supplémentaires. A l'inverse, une augmentation d'un écart-type des scores en compétences de base augmente les salaires de 20%. Cette relation supposée positive

nous amène à supposer qu'à diplôme égal, les compétences de base telles que la numératie, la littératie et la compréhension orale, ont un impact positif sur les salaires. Ne pas les inclure dans l'équation de gains risque de surestimer le rendement associés aux diplômes alors même que ceux-ci filtrent imparfaitement le niveau de compétences.

D'après la théorie du signal (Spence, 1973) et de l'écran (Stiglitz, 1975), les employeurs utiliseront le niveau de diplôme pour avoir une indication du niveau de compétences des candidats à l'embauche. Le niveau de diplôme les aidera à décider de la personne à embaucher ainsi que du niveau de salaire attribué. Mais d'après le modèle de Farber et Gibbons (1996)¹⁴, à mesure que le temps passe, les employeurs en apprendront davantage sur les réelles capacités de l'individu embauché, notamment à travers sa productivité, élément imparfaitement observé à l'embauche. En effet, c'est à travers l'expérience de l'individu embauché que l'employeur pourra obtenir de nouvelles informations sur ses capacités et donc sa productivité, informations qui supplanteront les signaux éducatifs pour déterminer l'évolution de salaire par exemple. Ce modèle théorique nous amène ainsi à s'interroger sur la linéarité des rendements associés aux compétences de base. En effet, les individus qui ont les plus faibles scores de compétences dans leur catégorie de diplôme, seront *a priori* moins productifs que ce que laissait à penser leur niveau de diplôme. Cette moindre productivité, observée à terme par l'employeur, aura des conséquences en termes de rémunération : le salaire ne sera effectivement pas revu à la hausse ou très faiblement (principe des « martingales »¹⁵ de Farber et Gibbons, 1996). A l'inverse, les salariés ayant les plus forts scores de compétences dans leur catégorie de diplôme, seront *a priori* plus productifs que ce qu'il était attendu au regard de leur niveau de diplôme. Les employeurs, en observant ce surplus de productivité, concéderont donc des hausses de salaires. Ainsi, pour un niveau de diplôme donné, le rendement des compétences de base peut s'avérer finalement plus important pour des individus très compétents que pour des individus peu compétents ayant le niveau supérieur de diplôme. En s'appuyant sur le modèle « d'apprentissage de l'employeur » de Farber et Gibbons (1996), nous supposons donc qu'à terme, les compétences de base sont plus rentables pour les individus les plus compétents de

¹⁴ Modèle « d'apprentissage des employeurs »

¹⁵ Le principe des martingales selon Farber et Gibbons (1996) suggère que les résidus de l'équation de salaire devraient suivre une martingale au sens où ces résidus ne donnent que l'illusion d'un gain salarial supplémentaire pour les travailleurs.

leur catégorie de diplôme que pour les individus les moins compétents de la catégorie supérieure de diplôme.

Par ailleurs, selon Stiglitz (1996), lorsque les perspectives économiques sont assombries, les différences interindividuelles de dotation en compétences s'exacerbent. En effet, en situation de rationnement de l'emploi, il existerait une sorte de compétition pour certaines compétences avancées, notamment celles qui permettent d'adopter et de s'adapter aux nouvelles technologies. Dans ce contexte, les individus qui ont de faibles compétences de base seront fragilisés dans leur dotation en compétences avancées, les premières étant indispensables à l'acquisition des secondes. Lorsque la situation économique ne permet pas d'embaucher tous les travailleurs, quel que soit leur niveau de compétences, les moins compétents risquent alors d'être davantage en difficulté pour trouver un emploi et de se retrouver à la fin de la file d'attente de Thurow. Dans un contexte de rationnement des emplois, nous supposons donc que les individus ayant des difficultés en compétences de base sont plus susceptibles d'être en dehors de l'emploi. Il s'agit d'une hypothèse dont il faudra tenir compte dans l'estimation des équations de gains afin d'éviter tout biais de sélection liées au niveau de compétences.

3.3. Méthodologie

D'un point de vue méthodologique, nous allons procéder en plusieurs étapes. La première consiste à choisir un échantillon à partir duquel nous pourrions mener des analyses les moins biaisées possible. Ensuite, nous présenterons les modèles standards de gains d'après Mincer, avant de les augmenter des scores en compétences de base. Enfin, nous exposerons le modèle d'Heckman en deux étapes que nous utiliserons afin de tenir compte d'un biais de sélectivité potentiel. En effet, rien ne laisse à penser que les travailleurs rémunérés à temps plein dans notre échantillon constituent un échantillon purement aléatoire de la population active. Les individus n'ayant pas accès à l'emploi sont en effet supposés avoir des caractéristiques particulières, comme un plus faible niveau de compétences.

3.3.1 Choix de l'échantillon

Les données que nous utiliserons sont celles issues de l'enquête Information et Vie Quotidienne (IVQ), réalisée en France en 2004 par l'INSEE auprès d'adultes âgés de 18 à 65 ans¹⁶. Afin de donner un éclairage neuf sur l'accès à l'emploi et apprécier le rendement des compétences de base sur le marché du travail, il faut préalablement déterminer pour quelle quotité de travail le salaire mensuel a été déclaré. En effet, aucune variable dans l'enquête IVQ ne nous permet de savoir si les individus sont à temps plein ou à temps partiel. Les salaires constituent en ce sens une variable « problématique », puisque nous ne savons pas à quelle quotité de travail le salaire mensuel correspond. Pour dépasser cette difficulté, nous avons construit des « profils » de quotité de travail à partir de l'enquête Emploi du quatrième trimestre de l'année 2004, puisque cela correspond parfaitement à la période de réalisation de l'enquête IVQ. Les éventuels changements de la structure économique et les distorsions liées à des effets conjoncturels contrastés sont ainsi évités. Dès lors, à partir des données disponibles et complètes, des tableaux croisés dynamiques ont été réalisés pour connaître le pourcentage de personnes à temps plein, à temps partiel subi et à temps partiel choisi, en fonction du sexe, du secteur d'activité et de la qualification professionnelle (cf. annexe G.)

Les hommes cadres, ouvriers et appartenant aux professions intermédiaires ont très peu de risque d'être à temps partiel, qu'il soit subi ou choisi. Ce type de profils a donc été conservé dans l'échantillon de l'enquête IVQ. En revanche, concernant les employés travaillant dans les secteurs de l'énergie et des services, le risque d'être à temps partiel est bien plus élevé. Les services aux particuliers et aux entreprises sont les deux secteurs ayant la plus faible proportion d'employés masculins à temps plein (91% et 88% respectivement). Ces profils d'individus ont donc été exclus de l'analyse, tout comme les « agriculteurs exploitants » et les « artisans » puisqu'il s'agit essentiellement de populations non salariées¹⁷. Par ailleurs, parce que les femmes sont davantage susceptibles d'être à temps partiel subi (discrimination) et choisi (charge d'enfants, de famille...) que les hommes, des tableaux complémentaires ont été réalisés selon la charge d'enfants avec des tests de Khi-

¹⁶ Voir le descriptif détaillé de cette enquête en partie 1.

¹⁷ Les systèmes de rémunérations de ces CSP ne sont pas les mêmes que pour des individus salariés.

deux par CSP¹⁸ (cf. annexes H.). Ces analyses supplémentaires nous ont amenés à ne pas conserver les femmes dans l'enquête IVQ.

Une fois cette sélection opérée dans la base de données IVQ (pour ne conserver que les profils les plus susceptibles d'être à temps plein) nous avons constaté, via la dispersion des salaires, que certains individus avaient encore des salaires bien inférieurs au SMIC mensuel net de 2004 alors qu'ils étaient supposés être à temps plein. Nous avons alors décidé de supprimer toutes les observations dont le salaire minimum net mensuel était inférieur à 892€ par mois¹⁹.

3.3.2. Modèle économétrique

Le rendement salarial des compétences de base sur le marché du travail peut s'appréhender dans cadre des fonctions de gains de Mincer (1975). Cependant, l'estimation de simples fonctions de gains risque d'être entachée d'un biais de sélection. En effet, comme les salaires sont la variable endogène d'intérêt, cela suppose que seuls les individus en emploi seront observés. Or il est probable que les individus en emploi diffèrent dans leurs caractéristiques intrinsèques des individus en dehors de l'emploi. Parmi ces caractéristiques se trouvent notamment le niveau de compétences de base. C'est pour cette raison qu'après avoir présenté des fonctions de gains « classiques », nous détaillerons des fonctions de gains corrigées du biais de sélection, grâce au modèle d'Heckman en deux étapes.

¹⁸ Comme certains secteurs d'activités détaillés ne recueillaient aucune observation pour les femmes, nous avons décidé de réaliser les tableaux croisés à un niveau plus agrégé de secteur (industrie et services)

¹⁹ Au 1er juillet 2004, le SMIC brut a été réévalué à 7,61€ de l'heure, soit 1154€ mensuels brut pour un temps plein de 35h par semaine (Source : INSEE). Afin de déterminer le niveau du SMIC mensuel net pour la même quotité de travail, il faut multiplier le salaire brut par un coefficient correspondant au pourcentage de salaire restant après déduction des cotisations sociales. Pour les salariés de moins de 65 ans, ce coefficient était de 0,7735 en 2004. Le salaire minimum net mensuel est donc établi à 892€. (Source : URSAAF).

3.3.2.1. Le modèle standard de Mincer

Mincer (1975) explique l'évolution du logarithme des salaires par une fonction de gains dont les variables exogènes sont le nombre d'années d'études réalisées en formations initiale (EDU), le nombre d'années d'expérience professionnelle (EXP) et le nombre d'années d'expérience professionnelle au carré (EXP²) :

$$\ln(W_i) = \beta_0 + \beta_1 EDU_i + \beta_2 EXP_i + \beta_3 EXP_i^2 + \varepsilon_i$$

avec $\beta_0, \beta_1, \beta_2 > 0$ et $\beta_3 < 0$ et avec μ_i , le terme d'erreur de la fonction mesurant les chocs exogènes et normalement distribué. Le terme quadratique de la fonction de gains permet de tenir compte des rendements décroissants de l'expérience professionnelle avec l'âge.

Comme nous supposons que les compétences de base ont un impact direct sur les salaires, nous proposons d'ajouter au modèle standard de Mincer des scores en compétences de base. Le modèle à estimer devient alors :

$$\ln(W_i) = \beta_0 + \beta_1 EDU_i + \beta_2 EXP_i + \beta_3 EXP_i^2 + \beta_4 LITT_i + \beta_5 NUM_i + \beta_6 CO_i + \varepsilon_i$$

Les variables *LITT*, *NUM* et *CO* représentent respectivement les scores (allant de 0 à 100) en littératie, en numératie et en compréhension orale. Les coefficients $\beta_1, \beta_4, \beta_5$ et β_6 représentent respectivement le taux de rendement moyen, en pourcentage de salaire, d'une année d'études, d'un point en littératie, en numératie et en compréhension orale. Le taux de rendement marginal d'une année d'expérience est égal à $\beta_2 + 2\beta_3 EXP$.

Bien entendu, à ce modèle standard, il est possible d'ajouter des variables d'offre, renseignant notamment sur des caractéristiques individuelles (âge, sexe, charge familiale, situation matrimoniale...), ou des variable de demande (secteur d'activité, taille de l'entreprise...)

3.3.2.2. Correction du biais de sélection : modèle d'Heckman

Afin de d'estimer au mieux l'impact des compétences de base sur les salaires, un modèle d'Heckman avec correction du biais de sélection sera utilisé. Ce modèle permet d'appréhender simultanément l'influence exercée par les compétences de base sur l'accès à l'emploi et d'estimer des rendements non-biaisés de compétences. La procédure d'Heckman (1979) dite en deux étapes permet de corriger les biais de sélection car elle permet de tenir compte de la particularité des individus qui ont été sélectionnés. Il faut alors estimer en premier lieu la probabilité d'être en emploi.

Par exemple, si nous estimons le modèle de base pour les individus en emploi, la régression devient :

$$E(\overline{\ln W}_i / Z_i ; h = 1) = \tau Z_i + E(u_j / h) \quad (1)$$

$$\text{Avec : } Z_i = \beta_0 + \beta_1 \text{EDU}_i + \beta_2 \text{EXP}_i + \beta_3 \text{EXP}_i^2 + \beta_4 \text{LITT}_i + \beta_5 \text{NUM}_i + \beta_6 \text{CO}_i + \varepsilon_i$$

Avec $h=1$ si l'individu est en emploi, 0 sinon. Heckman (1979) a démontré que l'espérance du terme à droite de l'équation (1) n'est généralement pas nulle et est corrélée avec les variables de Z_i . Le recours aux moindres carrés ordinaires risque ainsi de donner des estimateurs biaisés, d'où le recours à la procédure d'Heckman en deux étapes. La première étape consiste à estimer l'espérance $E(u_j / h)$ grâce à un modèle probit simple duquel nous pourrions ensuite extraire l'inverse du ratio de Mills (λ). A la deuxième étape, nous régressons l'équation de salaires par les moindres carrés ordinaires en incluant l'inverse du ratio de Mills, calculé lors de la première étape.

De fait, pour un individu i , la probabilité d'être en emploi, peut-être modélisé grâce à un modèle probit dichotomique simple. Soit B_i , la variable renseignant sur le fait que l'individu est en emploi ou en dehors de l'emploi. Le modèle s'écrit alors :

$$B_i = v H_i^B + \mu_i^B$$

B_i^* représente une variable latente qui détermine le fait d'être en emploi.

$B_i^* = 1$ si l'utilité associée au fait d'être en emploi est supérieure à celle associée au fait de ne pas l'être.

$B_i^* = 0$ si l'utilité associée au fait de ne pas être en emploi est supérieure à celle associée au fait d'y être.

vH_i^B est un vecteur de variables contrôlant l'influence des caractéristiques individuelles ainsi que des facteurs environnementaux.

μ_i^B est le résidu de l'équation de sélection

Le niveau de salaire est étudié seulement si l'individu i est en emploi, c'est-à-dire si $B_i^* = 1$.

Ainsi, à partir du modèle de base, le système d'équation devient :

$$\begin{cases} B_i = vH_i^B + \mu_i^B \\ Z_i = \beta_0 + \beta_1 EDU_i + \beta_2 EXP_i + \beta_3 EXP_i^2 + \beta_4 LITT_i + \beta_5 NUM_i + \beta_6 CO_i + \lambda_{B_i}(R_i^* = 1) + \varepsilon_i \end{cases}$$

La spécification de l'équation de sélection du modèle d'Heckman se base sur l'inclusion de différentes variables, dont le niveau en compétences de base des individus. En effet, il faut s'attendre à ce que les moins compétents aient plus de risques d'être exclus de l'emploi²⁰. Les plus compétents sont davantage recherchés sur le marché du travail car ils sont considérés comme plus productifs, d'où l'idée qu'un déficit en compétences puisse se traduire par plus de difficultés d'accès à l'emploi. Nous ajoutons également des variables concernant la tranche d'âge d'appartenance des individus puisque notre société se caractérise par un taux de chômage et d'inactivité bien plus fort pour les plus jeunes et un retour à l'emploi bien plus difficile pour les plus de 50 ans. D'autres variables renseignant sur le contexte économique de résidence (taux de chômage départemental, Tranche Urbaine de Résidence) ont été ajoutées. Du point de vue éducatif, nous supposons bien entendu que la durée théorique de scolarité, mais aussi la spécialité de formation et le niveau d'études à partir duquel les premières difficultés scolaires ont été ressenties, ont un impact dans l'accès à l'emploi. Enfin, nous sommes également

²⁰ La variable « accès à l'emploi » est spécifiée comme suit : elle prend la valeur 1 pour les actifs occupés et la valeur 0 qu'il s'agisse d'actifs en recherche d'emploi ou de personnes en âge et en capacité de travailler mais en retrait du marché du travail.

tentés de croire que le fait de ne pas être né en France peut avoir un effet négatif sur l'accès à l'emploi, notamment en raison des fortes tendances à la discrimination à l'embauche en France (Cédiey et Foroni, 2007). Concernant l'équation substantielle, nos variables exogènes sont la durée de scolarité, l'expérience professionnelle, l'expérience professionnelle au carré ainsi que les scores en littératie, numératie et en compréhension orale.

3.4. Statistiques descriptives

A partir de la situation professionnelle renseignée dans l'enquête IVQ-2004, il a été possible de séparer les individus en emploi et ceux qui ne l'étaient pas. Ceux qui ne sont pas en emploi sont soit chômeurs, soit inactifs. Quoi qu'il en soit, ils sont supposés être globalement moins compétents que les individus en emploi. En effet, si les travailleurs ont de trop fortes lacunes dans ces compétences, ils sont alors susceptibles d'être sanctionnés par le marché du travail, en termes de salaires, ou pire, en termes d'employabilité, sachant que ces compétences sont aujourd'hui bien plus nécessaires dans les différents emplois, quels que soient les secteurs d'activité ou la profession exercée. Dans ce contexte, nous pouvons donc supposer que les individus en dehors de l'emploi, sont peut-être les moins compétents, bien que cela puisse également résulter d'accidents de parcours, de choix indépendants du niveau de compétences (charge familiale, projet de vie...) ou encore de la conjoncture économique (Stiglitz, 1996).

Le tableau 1 montre bien que le niveau de compétences de base, que ce soit en littératie, en numératie ou en compréhension orale, est globalement plus élevé pour les personnes en emploi que pour celles qui ne le sont pas. De même les écart-types sont également plus élevés pour les individus qui ne sont pas en emploi, ce qui suggère une plus forte hétérogénéité des compétences parmi eux.

Tableau 5 : Scores moyens en compétences de base selon que l'on soit en emploi ou non.

	Littératie		Numératie		Compréhension orale	
	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type
En emploi	89.39	9.51	76.66	14.59	84.80	18.53
En dehors de l'emploi	84.72	13.81	68.68	20.84	76.48	22.46

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Si le niveau de compétences de base est plus élevé pour les personnes en emploi, il est également supposé être d'autant plus élevé que le niveau de diplôme est important. C'est ce qu'illustre le tableau suivant :

Tableau 6 : Répartition des individus selon leur niveau de diplôme ainsi que leur appartenance à tel ou tel quartile de compétences de base.

En %	Littératie				Numératie				Compréhension orale			
	Q1	Q2-Q3	Q4	Total	Q1	Q2-Q3	Q4	Total	Q1	Q2-Q3	Q4	Total
Jamais scolarisé ou sans diplôme	66.5	30.2	3.3	100	48.4	49.1	2.5	100	48.7	39.5	11.8	100
Certificat d'Études Primaires	59.7	31.8	8.5	100	31.2	65.1	3.7	100	27.4	52.1	20.5	100
BEPC/brevet des collèges	14.8	66.8	18.4	100	8.6	78.0	13.4	100	13.7	52.2	34.1	100
BEP/CAP	27.8	61.7	10.5	100	14.2	75.7	10.1	100	15.8	42.4	41.8	100
Baccalauréat	8.3	71.7	20.0	100	3.9	72.3	23.8	100	7.2	39.6	53.2	100
...général	12.1	63.4	24.5	100	1.3	71.1	27.6	100	7.6	31.5	60.9	100
...technique	5.9	65.2	28.9	100	3.51	66.8	29.6	100	9.7	35.9	57.4	100
...professionnel	6.9	76.3	10.8	100	6.2	76.7	17.1	100	5.3	50.2	44.5	100
Bac+2	10.7	61.5	27.8	100	2.9	59.4	37.6	100	4.1	38.5	57.4	100
> Bac+2	7.2	50.9	41.8	100	0.4	46.6	53.0	100	6.8	36.8	56.4	100

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Le niveau de compétences est indéniablement corrélé au niveau de diplôme. Toutefois, cette corrélation demeure imparfaite puisqu'avoir un niveau de diplôme

équivalent à un « Bac+2 » ou au-delà, ne garantit pas de faire partie des individus ayant les meilleurs scores (4ème quartile) et avoir un faible niveau de diplôme n’empêche pas des individus de se retrouver en haut de la distribution des scores.

Ainsi, la corrélation entre diplôme et niveau de compétences étant imparfaite, nous nous attendons à ce que les rendements associés aux diplômes et plus globalement à la durée de scolarité soient biaisés si l’on ne tient pas compte des compétences de base. Plus précisément, nous nous attendons à ce que ces rendements soient plus faibles dès lors que les compétences de base seront incluses en tant que variables explicatives dans les fonctions de gains.

Avant d’exposer les résultats propres aux équations de gains, voici quelques statistiques descriptives concernant le lien entre niveau de compétences de base et salaire. Le tableau suivant présente les salaires moyens et médians selon la présence d’un déficit ou d’un surplus de compétences par rapport à son niveau de diplôme.

Tableau 7: Salaires moyens et médians selon la présence d’un déficit ou d’un surplus de compétences par rapport à son niveau de diplôme

	Déficit de compétences		Surplus de compétences	
	Moyenne	Médiane	Moyenne	Médiane
En littératie	1594€	1400€	1790€	1500€
En numératie	1549€	1417€	2251€	1600€
En compréhension orale	1881€	1500€	1641€	1459€

Source : calculs de l’auteur ; enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Les salaires moyens et médians sont globalement plus élevés pour les individus qui seraient en surplus de compétences par rapport à leur niveau de diplôme que pour ceux qui seraient en déficit, sauf pour la compréhension orale. Les individus en déficit de compétences en numératie perçoivent un salaire moyen de 1549€ tandis que les salariés en surplus ont un salaire moyen de 2251€. Leur salaire médian est également plus élevé. Il semblerait donc bien qu’il y ait un lien entre niveau en compétences de base et salaires, indépendamment du niveau de diplôme, surtout pour la numératie.

3.5. Résultats

Dans la section suivante seront présentées les différentes fonctions de gains, notamment avec et sans correction du biais de sélection.

3.5.1. Fonctions mincériennes augmentées des scores de compétences

D'après le modèle 1 du tableau I.1 (annexes I), plus le diplôme est élevé par rapport au CAP/BEP, catégorie de référence, plus le taux de rendement est grand. Ainsi, alors qu'un niveau baccalauréat, toutes filières confondues, représente un taux de rendement de 17.8% pour un individu, un diplôme supérieur à un niveau BAC+2 suppose quant à lui un taux de rendement de 54.4%.

L'expérience professionnelle est une autre composante significative expliquant les salaires. Ainsi, une année d'expérience professionnelle supplémentaire représente un taux de rendement de 2.9%. Ce chiffre ne variera guère à travers l'ensemble des modèles ci-dessus, tout comme l'expérience professionnelle au carré qui appuie la thèse des rendements décroissants de l'expérience.

Par ailleurs, dans les modèles 2, 3 et 4, où les différents scores de compétences de base ont été inclus séparément, nous constatons qu'ils ont tous un impact positif et significatif sur les salaires, bien que la compréhension orale semble être la compétence la moins valorisée sur le marché du travail en termes de gain salarial. A noter que cette dernière ne sera plus une variable significative, suite à l'ajout simultané d'autres variables. Par ailleurs, d'après le modèle 6 et les suivants, la numératie paraît être la compétence la plus « rentable », ce qui va dans le sens de notre hypothèse de départ. D'après nos modèles, 10 points supplémentaires en numératie représenterait alors 3% de gain salarial, contre 2% pour la littératie.

Jusqu'ici, nous avons considéré le diplôme comme mesure du capital humain dans les équations de gains. Or, celles-ci peuvent aussi intégrer le nombre d'années d'étude en lieu et place du diplôme. Nous décidons alors de réaliser la même analyse

que précédemment sur les déterminants des salaires, en remplaçant le diplôme par le nombre d'années d'études correspondant à la plus haute classe atteinte, afin de voir si nous retrouvons bien des résultats convergents (tableau I.2 en annexes I).

Dans ces modèles incluant cette fois la durée théorique de scolarité selon la plus haute classe atteinte par les individus de notre échantillon, les coefficients associés à l'expérience professionnelle sont toujours aussi stables et significatifs. Nous pouvons également dégager le taux de rendement d'une année d'études supplémentaire : selon les variables introduites simultanément, ce taux varie entre 5,7% et 6,6%. Lorsque la littératie, la numératie et la compréhension orale sont introduites séparément dans l'équation de gains, elles sont toutes les trois significatives, avec toujours une plus faible valorisation de la compréhension orale. Néanmoins, dès lors que les trois scores sont simultanément introduits, seul le coefficient de la numératie reste significativement positif²¹. Là encore, le taux de rendement associé à un point supplémentaire de numératie est de 0,3%. 20 points supplémentaires en numératie apporteraient ainsi un rendement légèrement supérieur à une année d'étude (6% contre 5,7% dans le modèle 6).

Ainsi, d'après l'ensemble de ces résultats, il existerait bien un lien non négligeable entre le niveau de maîtrise des compétences de base, notamment en numératie, et la détermination du niveau des salaires. La montée en puissance du concept de compétence doit alors supplanter les références traditionnelles de l'éducation certifiée, comme le diplôme ou le nombre d'années d'études. Une étude longitudinale devrait permettre d'appréhender cette question par le biais d'une confrontation des taux de rendement du nombre d'années d'études par exemple et ceux des compétences de base.

3.5.2. Les rendements privés des compétences sont-ils non-linéaires ?

D'après le modèle de Farber et Gibbons (1996), déjà évoqué précédemment, à mesure que le temps passe, les employeurs en apprennent davantage sur les réelles capacités des individus embauchés, notamment à travers leur productivité, élément

²¹ Soulignons toutefois que le score en littératie est presque significatif au seuil de 10% ($P > |t| = 0.109$)

imparfaitement observé à l'embauche. C'est à travers l'expérience de l'individu que l'employeur pourra obtenir de nouvelles informations sur leurs capacités et donc leur productivité qui en dépend. Ces nouvelles informations supplanteront alors les signaux éducatifs traditionnels tel que le diplôme, pour fixer les niveaux futurs de rémunération. Cette approche remet en cause la théorie du capital humain, selon laquelle les gains sont d'autant plus élevés que le niveau de diplôme est important. Le lien n'est pas aussi mécanique comme le révèlent Farber et Gibbons (196). L'exercice même d'un emploi révèle aux employeurs les réelles capacités des individus. Il est donc possible d'imaginer qu'un employeur rémunère bien mieux un individu particulièrement compétent pour son niveau de diplôme, qu'un individu particulièrement incompétent pour un niveau de diplôme supérieur. Cette stratégie permettrait en outre de faire perdurer la motivation et donc la forte productivité du travailleur très compétent. Ce modèle nous amène ainsi à s'interroger sur la linéarité des rendements associés aux compétences de base.

Afin de mieux appréhender cette question, de nouvelles équations de gains sont estimées en introduisant cette fois-ci une variable polytomique particulière. Celle-ci prend plusieurs modalités selon le plus haut diplôme en possession et le quartile de scores en compétences dans lequel se trouve l'individu au regard de son diplôme. En réalisant de telles estimations, l'idée est de déterminer s'il y a bien un effet « niveau de diplôme » qui améliorerait le rendement des compétences. Autrement dit, est-ce qu'un individu qui appartiendrait au premier quartile des scores en compétences, peut espérer un meilleur rendement de celles-ci si son diplôme est relativement plus élevé que la catégorie de référence ? Si tel est le cas, cela signifie qu'il y a bien un effet « diplôme » : le diplôme permettrait alors de masquer le réel niveau de compétence de son détenteur auprès des employeurs au moment du recrutement.

Tableau 8 : Équations de gains selon le diplôme et le quartile d'appartenance de la distribution des scores

	Variables	LITTÉRATIE	NUMÉRATIE
	Durée de scolarité théorique	0.034***	0.029***
	Expérience professionnelle	0.029***	0.029***
	(Expérience professionnelle) ²	-0.0004***	-0.0004***
	Score en compétences	0.003*	0.003***
Sans diplôme ou Certificat d'Études Primaires	Q1_score	0.003	0.079
	Q2_Q3_score	-0.047	0.011
	Q4_score	0.088	0.050
Brevet des collèges - BEPC	Q1_score	0.005	0.144
	Q2_Q3_score	0.099	0.015
	Q4_score	-0.015	0.127
CAP / BEP	Q1_score	-0.056	0.031
	Q2_Q3_score	-0.042	Référence
	Q4_score	Référence	0.016
Niveau Bac.	Q1_score	0.119**	0.095
	Q2_Q3_score	0.066*	0.116***
	Q4_score	0.109**	0.108
Niveau BAC + 2 (BTS, DEUG, DUT...)	Q1_score	0.219***	0.200***
	Q2_Q3_score	0.176***	0.249***
	Q4_score	0.273***	0.210***
Niveau supérieur à BAC+2	Q1_score	0.136**	0.246***
	Q2_Q3_score	0.345***	0.316***
	Q4_score	0.299***	0.449***
	Constante	6.312***	6.318***
	Observations	1877	1877
	R ² ajusté	0.32	0.33

* significatif au seuil de 10% ; ** significatif au seuil de 5% ; *** significatif au seuil de 1%

Note : la compréhension orale étant une compétence dont les scores sont saturés à droite de la distribution, nous ne mènerons l'étude de la non-linéarité des rendements uniquement sur la littératie et la numératie.

D'après le tableau 8, outre les résultats déjà mis en avant dans les premières estimations d'équations de gains (stabilité des coefficients liés à l'expérience professionnelle, rendement du diplôme à peine moins fort lorsque l'on considère la numératie...), nous pouvons remarquer que seuls les coefficients relatifs aux catégories supérieures à celle de référence sont significatifs, quelle que soit la compétence considérée. D'après ces coefficients, autant il semble effectivement

exister un effet diplôme pour la numératie, autant les résultats sont plus contrastés pour la littératie. En effet, concernant la numératie, il vaut mieux faire partie des moins compétents des détenteurs d'un diplôme supérieur à une BAC+2 que de faire partie des plus compétents des détenteurs d'un diplôme de niveau BAC+2. C'est l'inverse pour la littératie, ce qui signifie qu'il y aurait davantage un effet « compétences en littératie » qu'un effet « diplôme ». Ce résultat pourrait laisser entrevoir certaines préconisations en termes de politiques éducatives. En effet, cela ne sert à rien d'augmenter le niveau d'études des individus si cela ne se traduit pas par une élévation du niveau de compétences. Répondre aux besoins du marché du travail en augmentant la quantité de diplômés ne suffit donc pas. Il faut au contraire s'assurer que la poursuite des études s'accompagne bien d'un développement certain des niveaux de compétences. Sans cela, la massification de l'éducation ne saurait être efficace et pertinente, tant pour l'individu que pour la société. Pour autant, cet effet « compétences en littératie » doit être pris avec recul dans la mesure où nous ne le retrouvons nulle part ailleurs, ni pour la littératie, ni pour la numératie.

Ce qui semble en revanche être plus certain, c'est qu'il existe un effet conjugué entre niveau d'études et niveau de compétence. En effet, en considérant un même niveau de numératie, par exemple en se focalisant sur le 4^{ème} quartile, nous constatons que le rendement augmente beaucoup à mesure que le diplôme s'élève. La tendance est la même pour les 2^{ème} et 3^{ème} quartiles ainsi que pour le 1^{er} quartile. Globalement, cela signifie bien qu'il y a un effet « signal » avec le diplôme. Mais cet effet se conjugue forcément avec un effet compétence puisque les variations du rendement escompté selon le niveau de diplôme, par rapport à la catégorie de référence, sont bien plus fortes pour les individus qui sont en haut de la distribution des scores en numératie.

3.5.3. Fonctions de gains corrigées du biais de sélection

Les résultats du modèle d'Heckman sont présentés ci-après. Pour des raisons de présentation, nous révélerons dans un premier temps l'estimation de l'équation de sélection visant à expliquer la probabilité d'être en emploi :

Tableau 9 : Probabilité d'être en emploi (versus être sans emploi) : modèle probit

<i>Accès à l'emploi (N=2401)</i>	<i>Coefficients</i>	<i>Effets marginaux (variation en points de %)</i>
Durée de scolarité théorique	0.011	0.0029
Score en littératie	0.009***	0.0023***
Score en numératie	0.007***	0.0021***
Score en compréhension orale	0.003**	0.0009**
<i>Tranche d'âge d'appartenance</i>		
18-30 ans	-0.224***	-0.065***
31-50 ans	<i>Référence</i>	
Plus de 50 ans	-0.449***	-0.135***
Origine étrangère	-0.081	-0.023
ZUS	-0.315***	-0.094***
Taux de chômage départemental	-0.033**	-0.009**
<i>Spécialités de formation</i>		
Formation de base (diplôme < BEPC)	-0.567***	-0.183***
<i>Formations générales</i>	<i>Référence</i>	
Formations industrielles	-0.168**	-0.048**
Formations tertiaires	-0.202***	-0.057***
<i>Niveau d'études des débuts de difficultés scolaires</i>		
<i>Pas de difficultés scolaires</i>	<i>Référence</i>	
A partir du primaire	-0.242***	-0.070***
A partir du secondaire inférieur	-0.236***	-0.069***
A partir du secondaire supérieur	-0.326**	-0.100**
Constante	-0.012	--

* significatif au seuil de 10% ; ** significatif au seuil de 5% ; *** significatif au seuil de 1%

D'après l'équation de sélection, il semblerait qu'il soit plus difficile d'accéder à l'emploi pour les individus ayant entre 18 et 30 ans et particulièrement pour les plus de 50 ans, d'où le signe négatif et significatif des coefficients. Cela rejoint l'idée selon laquelle, en France, le chômage touche davantage les jeunes, en manque

d'expérience, et les plus âgés, souvent perçus par les employeurs comme moins productifs et donc moins rentables.

Autre résultat classique, le lieu d'habitation peut également être déterminant dans l'accès à l'emploi. Par exemple, habiter à Paris augmente significativement la probabilité d'occuper un emploi par rapport au fait d'habiter dans une ville provinciale de plus de 100 000 habitants. De surcroît, plus le département du lieu de résidence est touché par le chômage, plus la probabilité de trouver un emploi diminue.

Si l'on considère les variables éducatives, plusieurs constatations peuvent encore être faites. Tout d'abord, il est surprenant de remarquer que la durée de scolarité a un impact non-significatif dans l'accès à l'emploi, contrairement aux compétences. En effet, alors que nous ne pouvons rien conclure quant à l'influence du nombre d'années d'études sur la probabilité d'être en emploi, nous pouvons dire que celle-ci augmente de 2,2 points de pourcentage pour 10 points supplémentaires en littératie. Ce chiffre est respectivement de 1,7 et de 0,8 pour la numératie et la compréhension orale. L'impact plus élevé de la littératie peut s'expliquer par le besoin grandissant des entreprises à avoir des salariés maîtrisant la transmission d'informations écrites. Les enquêtes « Techniques de l'Organisation du Travail » (TOTTO), menées entre 1987 et 1993, ont effectivement montré que les instructions orales de travail avaient chuté de 38% en 1987 et de 29% en 1993 alors qu'en parallèle la part des informations transmises par écrit ont augmenté de 47% en 1987 et de 55% en 1993. L'intensification de l'utilisation de l'écrit dans le monde professionnel provient certainement de la mise en place croissante de normes (chartes de qualité, certifications ISO, renforcement des règlements intérieurs, traçabilité de certains produits...) qui complexifie le travail par l'abondance de documents écrits (procédures, consignes de sécurité, notices d'utilisation des machines...). Ainsi, les moins compétents en littératie peuvent vite être fragilisés face à l'emploi.

Toutefois, un déficit en compétences²² de base n'exclut pas forcément systématiquement de l'emploi ceux qui en sont porteurs. En effet, deux tiers d'entre eux sont malgré tout en emploi dans notre échantillon, même si globalement ils le

²² Les individus de l'échantillon ont été considérés en déficit de compétences dès lors que chacun de leurs scores était inférieur de plus d'un écart-type (σ) au score moyen (m) correspondant à leur niveau de diplôme.

sont proportionnellement moins que ceux n'ayant aucun déficit. Ce constat peut résulter de deux phénomènes : soit le travail en lui-même n'implique pas de recourir particulièrement aux compétences de base, soit il est possible pour certains individus de compenser un déficit partiel, c'est-à-dire dans une seule compétence, par une compétence bien mieux maîtrisée, sous réserve que l'emploi occupé permette une telle compensation.

Alors qu'un déficit dans une seule compétence peut ne pas être sanctionné dans l'accès à l'emploi, nous pensons à l'inverse qu'un déficit total, c'est-à-dire touchant les trois compétences simultanément, peut avoir des répercussions plus importantes. Afin de tester cette hypothèse, nous avons calculé à partir du modèle probit précédent les probabilités prédites²³ d'être en emploi ou non selon que le déficit était inexistant, partiel ou total (tableau) :

Tableau 10 : Probabilités prédites d'être en emploi ou non selon l'importance du déficit de compétences

		En emploi	Sans emploi
Aucun déficit de compétences		0,92	0,08
Uniquement un déficit de compétences en...	...littératie	0,82	0,18
	...numératie	0,81	0,19
	...compréhension orale	0,88	0,12
<i>Déficit de compétences dans les trois domaines</i>		0,72	0,28

Source : calculs de l'auteur ; enquête Information et Vie Quotidienne 2004, INSEE

La probabilité d'un individu donné correspondant au profil retenu a une probabilité de 0,92 d'être en emploi s'il n'a aucun déficit en compétences de base. A l'inverse, un déficit partiel réduit cette probabilité de près de 10 points de pourcentage, voire jusqu'à 20 points lorsque le déficit en compétences est total. Les déficits en compétences de base pénalisent ainsi d'autant plus l'employabilité qu'ils sont généralisés.

Concernant la filière de formation, nous observerons que les filières générales sont celles qui permettent *a priori* une meilleure insertion par rapport aux autres types

²³ Ces probabilités ont été calculées pour le profil suivant : Homme - Durée de scolarité moyenne de 11 ans - Excès en compétences de base - A entre 31 et 50 ans - Français de souche - Taux de chômage départemental moyen de 10,2% - A suivi une formation industrielle. Notons que pour des individus n'ayant aucun déficit, nous avons supposé qu'aucune difficulté scolaire n'avait été ressentie au cours de la scolarité. Pour ceux ayant un déficit partiel ou total, nous avons au contraire supposé que les difficultés sont respectivement apparues à partir du secondaire inférieur et du primaire.

de formations. Ceci est sans doute dû au fait que les formations générales en question sont globalement des formations longues. Or, d'après la théorie du capital humain, plus la durée d'étude d'un individu sera importante, plus celui-ci sera considéré productif. Ce meilleur niveau de productivité, certifié par le diplôme, doit ainsi favoriser l'insertion professionnelle. Nous remarquerons en outre que les individus qui n'auraient pas un niveau d'études supérieur au Brevet des Collèges (BEPC) seront particulièrement sanctionnés : leur probabilité d'accès à l'emploi sera diminué de 18,1 points de pourcentage par rapport aux individus qui auraient un niveau de formation plus élevé.

Le fait d'avoir connu des difficultés scolaires durant son parcours peut également se révéler néfaste dans l'accès à l'emploi, signe de difficultés d'acquisition de certains savoirs. Toutefois, le fait d'avoir rencontré ses premières difficultés scolaires à partir du secondaire supérieur est bien plus pénalisant en termes d'accès à l'emploi (-9,3 points de pourcentage) que si elles avaient été ressenties pour la première fois au niveau primaire (-6,6 points de pourcentage) ou secondaire inférieur (-6,1 points de pourcentage). Ce résultat est à la fois étonnant – nous nous attendions initialement à l'inverse – et rassurant. En effet, ceci signifierait que notre système éducatif est suffisamment efficace pour détecter et prendre en charge les élèves qui seraient en difficultés d'apprentissage dès leur début de scolarité. Autrement dit, plus tôt ces difficultés seraient détectées, mieux le système éducatif serait en mesure de les « corriger » en vue de consolider les parcours de formation et d'améliorer les perspectives d'insertion professionnelle.

Grâce à ce modèle probit simple, constituant la première étape du modèle d'Heckman, nous avons pu estimer l'inverse du ratio de Mills qui servira de facteur explicatif du gain en étant réintroduit dans l'équation substantielle. Nous présentons également les résultats du modèle d'Heckman en deux étapes par secteurs d'activités. Sachant que l'échantillon initial est déjà assez restreint, l'industrie et la construction ont été regroupés ensemble. Les résultats sont les suivants :

Tableau 11 : Équations de gains avec et sans correction du biais de sélection

	ENSEMBLE		INDUSTRIE ET CONSTRUCTION		TERTIAIRE	
	Sans correction	Avec correction	Sans correction	Avec correction	Sans correction	Avec correction
<i>Logarithme du salaire mensuel (N=1877)</i>	<i>Coefficients</i>	<i>Coefficients</i>	<i>Coefficients</i>	<i>Coefficients</i>	<i>Coefficients</i>	<i>Coefficients</i>
Durée de scolarité théorique	0,057***	0.058***	0.053***	0.056***	0.059***	0.060***
Expérience professionnelle	0,022***	0.025***	0.017***	0.020***	0.026***	0.028***
Expérience professionnelle au carré	-0.0002***	-0.0003***	-0.00009	-0.0002*	-0.0003***	-0.0003***
Score en littératie	0.0014	0.0023**	0.0002	0.001	0.0037**	0.0043**
Score en numératie	0.0030***	0.0037***	0.0035***	0.0041***	0.0026***	0.0032***
Score en compréhension orale	0.0005	0.0008	0.0005	0.0008	0.0003	0.0052
<i>Mills</i>	--	0.172**	--	0.16*	--	0.14
Constante	5,99***	5.75***	6.16***	5.92***	5.77***	5.58***
N	1877	1877	748	748	1088	1088
R ² ajusté	0.29	0.29	0.29	0.30	0.29	0.29

* significatif au seuil de 10% ; ** significatif au seuil de 5% ; *** significatif au seuil de 1%

Dans le modèle d'ensemble, le fait que le coefficient associé à l'inverse du ratio de Mills soit significatif au seuil de 5%, signifie que l'équation de gains ne peut se passer de l'estimation préalable d'une équation d'accès à l'emploi, au risque que ces résultats soient biaisés. Ainsi, à titre de comparaison, même si de toute évidence le biais est relativement faible, ne pas le corriger amène à une légère sous-estimation du rendement de la numératie (3% au lieu 3,7% pour 10 points) et surtout à la non-significativité de la littératie (cf. colonne 1 du tableau). Par ailleurs, le rendement associé à une année d'étude supplémentaire est quant à lui de 5,8% dans le modèle corrigé. Ce rendement aurait été plus élevé si les compétences de base n'avaient pas été incluses. Parallèlement, d'après les coefficients relatifs à chacune des compétences, la numératie est la mieux valorisée sur le marché du travail. Ainsi, le rendement de 10 points supplémentaires en littératie et en numératie est de 2,7% et de 3,7% respectivement. Bien qu'il soit difficile d'interpréter ce meilleur rendement de la numératie, on peut supposer que la diffusion des innovations technologiques

conduit à une mutation perpétuelle de la structure des emplois en faveur de travailleurs qualifiés et compétents, notamment dans des domaines plus « scientifiques » tel que la R&D, l'ingénierie, la médecine ou encore l'informatique. Un bon niveau de maîtrise en numératie étant nécessaire pour l'acquisition de compétences plus complexes dans ces domaines, il peut ainsi paraître plus compréhensible que les rendements associés à la numératie soient plus élevés.

Les résultats par secteurs d'activités montrent que les salaires dans l'industrie, la construction et les activités tertiaires sont tout aussi déterminés par la numératie. Cette compétence semble juste avoir plus d'importance dans la détermination des salaires dans les secteurs de l'industrie et la construction que dans les activités tertiaires, davantage sensibles à la littératie.

4. Conclusion

Dans ce premier essai, nous avons cherché à étudier les rendements des compétences de base sur le marché du travail, tant du point de vue du déclassement professionnel que du point de vue salarial.

Les différents résultats présentés dans la première partie de cet essai montrent qu'il existe effectivement un lien entre le déclassement, le surclassement et le niveau de compétences de base des salariés. Les individus ayant un déficit de compétences par rapport à leur niveau de diplôme seraient justement déclassés en raison de ce déficit et moins bien rémunérés. Notons que le niveau de compétences en numératie semble particulièrement structurant dans les parcours professionnels. A diplôme égal, les salariés qui ont les niveaux les plus élevés de compétences de base ont en général plus de chances d'occuper un niveau d'emploi au dessus de leur niveau de diplôme. *A contrario*, un niveau compétences plus faible, notamment en numératie, conduit plus fréquemment à un déclassement, notamment pour les plus jeunes, dans l'industrie et dans les emplois où ils sont obligés d'utiliser un ordinateur.

Dans la seconde partie de cet essai, nous avons cherché à appréhender dans quelle mesure les compétences de base pouvaient justement expliquer le niveau de salaire des individus. D'après nos résultats, il semblerait qu'il existe bel et bien un lien entre

le niveau de maîtrise des compétences de base et la détermination des salaires. De plus, la numératie et la littératie sont les deux compétences dont l'effet sur les salaires est le plus fort. Leur impact est toutefois différent selon les secteurs d'activités. Autant la numératie est la compétence la mieux valorisée dans les secteurs de l'industrie et de la construction en termes de salaires (ce qui rejoint d'ailleurs les résultats sur le déclassement), autant il s'agit de la littératie dans le secteur tertiaire.

L'ensemble des résultats présentés dans ce premier essai implique donc de porter l'effort des politiques éducatives en amont dans les parcours de formation, afin de consolider un socle de compétences, nécessaire dans la sécurisation des trajectoires sur le marché du travail. Pour autant ces résultats restent limités. L'absence de données longitudinales ne permet pas de mettre en évidence des effets de génération et d'expérience, ce qui constitue une limite à notre analyse. Outre le fait que les compétences de base n'ont certainement pas été produites de la même manière dans chaque formation et chaque niveau de diplôme ces dernières années, elles peuvent être acquises également sur le marché du travail en fonction des emplois qui sont proposés aux salariés. Certains emplois non qualifiés représentent de véritables trappes alors que d'autres, plus vertueux permettent de conserver ou d'accroître le stock de compétences de base et ouvrent ainsi des perspectives d'évolution plus favorables. On peut penser que l'accès à la formation continue par exemple peut jouer un rôle déterminant dans cette évolution. Or, Kapsalis (1997) constate une certaine disparité dans l'accès aux systèmes d'éducation et de formation des adultes en fonction de leur niveau de compétences. La formation continue est alors complémentaire de la formation initiale puisqu'un faible niveau de compétences de base risque de réduire les chances d'accès et de réussite dans la formation continue (Gleeson, 2005).

Enfin, avoir un niveau plus faible de compétences de base, pour un diplôme donné, ne suffit pas nécessairement à conclure à un déficit général de compétences. Dans une logique de signal, les employeurs ont une préférence à diplôme égal, pour les salariés qui ont le plus de compétences de base, surtout dans des domaines techniques où la numératie est très importante. Cela ne veut cependant pas dire que ceux qui ont un niveau de compétences de base inférieur représentent réellement un problème pour l'employeur. En effet, d'autres compétences peuvent être tout autant voir davantage recherchées, comme les compétences sociales (« *soft skills* »).

Bowles et Gintis (1976) ont été parmi les premiers à avancer que ces compétences sont d'importants prédicateurs de la réussite scolaire et du succès professionnel. Plus récemment, Lleras (2008), a montré que les compétences sociales expliquaient davantage de variance des salaires que ne le faisaient les compétences cognitives. Il est donc possible d'envisager une complémentarité entre compétences de base et compétences sociales. La maîtrise des secondes peut même contrebalancer un déficit des premières, ce qui explique pourquoi un faible niveau en compétences de base n'est pas forcément problématique sur le marché du travail, notamment dans les secteurs et les emplois où les compétences sociales sont centrales. Preuve en est, 70% des individus de notre échantillon, pourtant en déficit total de compétences, sont en emploi.

BIBLIOGRAPHIE DE L'ESSAI 1

- AFFICHARD J. [1981], Quels emplois après l'école : la valeur des titres scolaires depuis 1973, *Économie et Statistique*, n° 134, pp. 7-26.
- ALLEN J. & VAN DER VELDEN R. [2001], Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search, *Oxford Economic Papers*, vol. 53, n°3, pp. 434-452.
- ANTONI M. & HEINECK G. [2012], *Do Literacy and Numeracy Pay Off ? On the Relationship between Basic Skills and Earnings* (No. 6882). Institute for the Study of Labor (IZA).
- BALDINI ROCHA M.-S. & PONCZEK V. [2011], The effects of adult literacy on earnings and employment, *Economics of Education Review*, vol. 30, n°4, pp. 755-764.
- BAUDELLOT C. & GLAUDE M. [1986], Les diplômes se dévaluent-ils en se multipliant ?, *Économie et Statistique*, n° 225, pp. 3-16.
- BÉDUWÉ C. & ESPINASSE J.-M. [1995], France : politique éducative, amélioration des compétences et absorption des diplômés par l'économie, *Sociologie du Travail*, vol. 37n°4, pp. 527-556.
- BERLIN G. & SUM A. [1987], *Toward a More Perfect Union : Basic Skills, Poor Families and Our Economic Future*, Research report, New-York : Ford Foundation.
- BERMAN E., BOUND J. & MACHIN S. [1998], Implications of Skill-Biased Technological Change: International Evidence, *Quarterly Journal of Economics*, vol. 113, n° 4, pp. 1245-1279.
- BLAUG M. [1985], Where Are We in Economics of Education ?, *Economics of Education Review*, vol. 4, n°1. pp. 17-28.
- BOISSIÈRE M., KNIGHT J. B. & SABOT R.-H. [1985], Earnings, schooling, ability, and cognitive skills, *The American Economic Review*, vol. 75, n°5, pp. 1016-1030.
- BOOTHBY D. [2002], *Literacy Skills, Occupational Assignment and the Returns to Over- and Under-Education*, Statistics Canada, Ottawa.
- BOWLES S. & GINTIS H. [1976], Schooling in capitalist America: Education and the contradictions of economic life. *New York : Basic*.
- BOWLES S. & GINTIS H. [2002], Schooling in capitalist America, *Sociology of Education*, vol. 75, n°1, pp. 1-18.
- BYNNER J. [2004], Literacy, numeracy and employability : Evidence from the British birth cohort studies, *Literacy and Numeracy Studies*, vol. 13, n°31, pp. 45.
- CEDIEY E. & FORONI F. [2007], *Les discriminations à raison de l'origine dans les embauches en France. Une enquête nationale par tests de discrimination selon la méthode du BIT*, Genève : BIT.

- CHAPPLE S. & MARÉ D. [2000], Literacy formation and its influence on earnings and jobs. *Unpublished research paper, Department of Labour, Wellington.*
- CHARETTE M.-F. & MENG R. [1998], The determinants of literacy and numeracy, and the effect of literacy and numeracy on labour market outcomes, *Canadian Journal of Economics*, vol. 31, n°3, pp. 495-517
- CHEVALIER A. [2003], Measuring over-education, *Economica*, vol. 70, n°3, pp. 509-531.
- CHEVALIER A. & LINDLEY J. [2009], Overeducation and the skills of UK graduates, *Journal of the Royal Statistical Society : Series A (Statistics in Society)*, vol. 172, n°2, pp. 307-337.
- CHISWICK B.-R. & MILLER P.-W. [1996], Ethnic networks and language proficiency among immigrants, *Journal of Population Economics*, vol. 9, n°1, pp. 19-35.
- CHISWICK B.-R. & MILLER P.-W. [1995], The Endogeneity between Language and Earnings : International Analyses, *Journal of Labor Economics*, vol. 13, n°2, pp. 246-288.
- CHISWICK B.-R., LEE, Y.-L. & MILLER P.-W. [2003], Schooling, literacy, numeracy and labour market success, *Economic Record*, vol. 70, n°242, pp. 33-58.
- COUPPIÉ T., GIRET J.-F & LOPEZ A. [2005], Des formations initiales aux premiers emplois : une correspondance rarement assurée, In Giret J.-F., Lopez A. et Rose J. (dir.), *Des formations pour quels emplois ?*, Paris, La Découverte, Collection Recherche, pp.79-96.
- DI PAOLA V. & MOULLET S. [2012], Peut-on en finir avec le déclassement ?, *Formation Emploi*, n°118, pp. 83-101.
- DI PAOLA V., MOULLET S. & VERO J. [2005], Le déclassement dans les Fonctions publiques d'Etat et territoriales, In Giret J.-F., Lopez A. et Rose J. (dir.), *Des formations pour quels emplois ?*, Paris, La Découverte, Collection Recherche, pp.79-96.
- DEARDEN L., MCINTOSH S., MYCK M. & VIGNOLES A. [2002], The Returns to Academic, Vocational and Basic Skills in Britain, *Bulletin of Economic Research*, vol. 54, n°3, pp. 249-274.
- DEGORRE A. & MURAT F. [2009], La mesure des compétences des adultes, un nouvel enjeu pour la statistique publique, *Économie et Statistique*, n°424-425, pp. 3-30.
- DENNY K. & DOYLE O. [2010], Returns to basic skills in central and eastern Europe, *Economics of Transition*, vol. 18, n°1, pp. 183-208.
- DENNY K., HARMON C. & O'SULLIVAN V. [2004], *Education, earnings and skills : A multi-country comparison*, The Institute for Fiscal Studies, Dublin.
- DEVROYE D. & FREEMAN R. [2001], *Does Inequality in skills Explain Inequality of Earnings Across Advanced Countries ?* CEPDP, 552. Centre for Economic Performance, London School of Economics and Political Science, London, UK.

- DICKES P. & VRIGNAUD P. [1995], *Rapport sur les traitements des données françaises de l'enquête internationale sur la littératie*, Ministère de l'Éducation Nationale, Direction de l'Évaluation et de la Prospective, Paris: MEN/DEPP.
- DIÉBOLT C. [2001], La théorie de l'engorgement, *Économie Appliquée*, vol. 54, n°4, pp. 7-31.
- DOUGHERTY C. [2003], Numeracy, literacy and earnings: evidence from the National Longitudinal Survey of Youth, *Economics of Education Review*, vol. 22, n°5, pp. 511-521.
- DURU-BELLAT M. [2006], *L'inflation scolaire – Les désillusions de la méritocratie*, Paris, Éditions du Seuil & La République des Idées.
- FARBER H.-S. & GIBBONS R. [1996], Learning and wage dynamics, *The Quarterly Journal of Economics*, vol. 111, n°4, pp. 1007-1047.
- FORGEOT G. & GAUTIE J. [1997], Insertion professionnelle des jeunes et processus de déclassement, *Économie et statistique*, vol. 304-305, n°1, pp. 53-74.
- FREEMAN R.B. [1976], *The Overeducated American*, Academic Press, New-York.
- FREEMAN R. & SCHETTKAT R. [2000], Skill Compression, wage differentials and employment : Germany vs. the US, *Oxford Economic Papers*, vol. 53, n°3, pp. 582-603.
- GAINI M., LEDUC A. & VICARD A. [2012], *School as a shelter ? School leaving-age and the business cycle in France*, INSEE Série des documents de travail de la Direction des Études et Synthèses Économiques, G 2012 / 04.
- GIRET J.-F. & LEMISTRE P. [2004], Le déclassement à l'embauche des jeunes : vers un changement de la valeur des diplômes ?, *Cahiers Économiques de Bruxelles – Brussels Economic Review, special issue in Economics of Education and Human Resources*, vol. 47, n° 3, pp. 483-503.
- GIRET J.-F., NAUZE-FICHET E. & TOMASINI M. [2006], Le déclassement des jeunes sur le marché du travail, *Données sociales – La société française*, pp. 307-314.
- GLEESON L. [2005], Economic returns to education and training for adults with low numeracy skills, Adelaide, South Australia: NCVER, 29 p.
- GREEN, F. [1998], The value of skills, *Studies in Economics* 9819.
- GREEN F., MCINTOSH S. & VIGNOLES A. [1999], *Overeducation and skills-clarifying the concepts* (No. dp0435), Centre for Economic Performance, LSE.
- GREEN F., MCINTOSH S. & VIGNOLES A. [2002], The utilization of education and skills: evidence from Britain, *The Manchester School*, vol. 70, n°6, pp. 792-811.
- GUÉRIN-PACE F. & BLUM A. [1999], L'illusion comparative. Les logiques de l'élaboration et d'utilisation d'une enquête internationale sur l'illettrisme, *Population*, vol. 54, n°2, pp. 271-302.
- GUÉRIN-PACE F. & BLUM A. [2000], *Des lettres et des chiffres : Des tests d'intelligence*

à l'évaluation du «savoir lire», un siècle de polémiques, Paris, Fayard, 191 p.

GUIRONNET J.-P. [2006], Le déclassement en France : vers une dévalorisation des diplômes du supérieur ?, *Économie Appliquée*, vol. 59, n°1, pp. 93-120.

HARTOG J. [2000], Overeducation and earnings: where we are and where we should go, *Economics of Education Review*, vol. 19, n°2, pp. 131-147.

HECKMAN J.-J. [1979], Sample Selection Bias as a Specification Error, *Econometrica*, vol. 47, n°1, pp. 153-161.

HECKMAN J.-J., STIXRUD, J. & URZUA S. [2006], *The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior* (No. w12006), National Bureau of Economic Research.

KAPSALIS C. [1997], *Formation des employés : Une perspective internationale*, Statistique Canada et Développement des ressources humaines Canada, No 89-552-MIF au catalogue, n°2.

KOSSOUDJI S.-A. [1988], English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men, *Journal of Labor Economics*, vol. 6, n°2, pp. 202-228.

LEMISTRE P. [2003], Dévalorisation des diplômes et accès au premier emploi. *Revue d'Économie Politique*, vol. 113, n°1, pp. 37-58.

LEMISTRE P. [2009], Faut-il poursuivre l'expansion scolaire ?, *Économie et Société*, n°31, pp. 1469-1501.

LEMISTRE P. [2010], *La formation initiale, une valeur sûre pour les jeunes ?*, Toulouse, Presses universitaires du Capitole.

LIZE L. [2006], Facettes du déclassement, quel rôle pour les politiques de l'emploi ?, *Travail et Emploi*, n°107, pp. 33-45.

LLERAS C. [2008], Do skills and behaviors in high school matter ? The contribution of noncognitive factors in explaining differences in educational attainment and earnings, *Social Science Research*, vol. 37, n°3, pp. 888-902.

MAGNAC T. & THESMAR D. [2002], Analyse économique des politiques éducatives: l'augmentation de la scolarisation en France de 1982 à 1993, *Annales d'Économie et de Statistique*, n°65, pp. 1-33.

MAZARI Z., MEYER V., ROUAUD P., RYK F. & WINNICKI P. [2011], Le diplôme : un atout gagnant pour les jeunes face à la crise, *Bref*, n°283, 4 p.

MCCORMICK B. [1990], A theory of signalling during job search, employment efficiency, and stigmatised job , *Review of Economics Studies*, vol. 57, n°2, pp. 299-313.

MCGUINNESS S. & BENNETT J. [2007], Overeducation in the graduate labour market: A quantile regression approach, *Economics of Education Review*, vol. 26, n°5, pp. 521-531.

- MCINTOSH S. & VIGNOLES A. [2001], Measuring and assessing the impact of basic skills on labour market outcomes, *Oxford Economic Papers*, vol. 53, n°3, pp. 453-481.
- MINCER J. [1975], Schooling, experience, and earnings, *The Journal of Political Economy*, vol. 83, n°2, pp. 444-446.
- MONCEL N. [2008], Recrutement en entreprise: les débutants sont-ils victimes d'un tri trop sélectif, *Bref*, n°250, CEREQ, 4 p.
- MURAT F. & ROCHER T. [2009], Création d'un score global dans le cadre d'une épreuve adaptative, *Économie et statistique*, vol. 424, n°1, pp. 149-178.
- MURNANE R.-J., WILLET J.-B. & LEVY F. [1995], *The growing importance of cognitive skills in wage determination* (No. w5076). National Bureau of Economic Research.
- NAUZE-FICHET E. & TOMASINI M. [2002], Diplôme et insertion sur le marché du travail: approches socioprofessionnelle et salariale du déclassement, *Économie et statistique*, vol. 354, n°1, pp. 21-48.
- PATRINOS H. A. & SAKELLARIOU C. [2013], Adult literacy, heterogeneity and returns to schooling in Chile, *Education Economics*, (ahead-of-print), pp. 1-15.
- PECORARO M. [2011], *Is there still a wage penalty for being overeducated but well-matched in skills ? A panel data analysis of a Swiss graduate cohort*. Institut de Recherches Économiques et Sociales, Louvain.
- PRYOR F.-L. & SCHAFFER D.-L. [1999], *Who's not working and why: Employment, cognitive skills, wages, and the changing US Labor market*, Cambridge University Press : Cambridge UK.
- RIVERA-BATIZ F. [1992], Quantitative literacy and the likelihood of employment among young adults in the United States, *Journal of Human Resources*, vol. 27, n°2, pp. 313-328.
- RIVERA-BATIZ F. [1999], Undocumented workers in the labor market: An analysis of the earnings of legal and illegal Mexican immigrants in the United States, *Journal of Population Economics*, vol. 12, n°1, pp. 91-116.
- SATTINGER M. [1993], Assignment models of the distribution of earnings, *Journal of Economic Literature*, vol. 31, n°2, pp. 831-880.
- SLOANE P. J. [2003], Much ado about nothing ? What does the overeducation literature really tell us, In Büchel F., de Grip A. & Mertens A. (Eds.). *Overeducation in Europe. Current Issues in Theory and Policy*, Cheltenham: Edward Elgar, pp. 11-45.
- SOHN K. [2010], The role of cognitive and noncognitive skills in overeducation, *Journal of Labor Research*, vol. 31, n°2, pp. 124-145.
- STEEDMAN H. [1999], Measuring the Quality of Educational Outputs: Some Unresolved Problems, In Alexander R., Broadfoot P. & Phillips D. (eds) *Learning*

from *Comparing: New Directions in Comparative Education Research*, vol. 1, Oxford: Symposium Books, 304 p.

STIGLITZ E. [1975], The theory of screening, education and the theory of income, *American Economic Review*, vol. 65, n°3, pp. 283-300.

STIGLITZ J.-E. [1996], Some lessons from the East Asian miracle. *The World Bank Research Observer*, vol. 11, n°2, pp. 151-177.

SUM A., HARRINGTON P. & W. GOEDICKE. [1986], "Basic Skills, High-School Diplomas, and the Labor Force, Employment, Unemployment, and Earnings Experiences of Young Adults in the U.S." Research Report. New York: Ford Foundation.

TAUBMAN P.-J. & WALES T.J. [1973], Higher Education, Mental Ability and Screening, *Journal of Political Economy*, vol. 81, n°1, pp. 25-55.

THURLOW L.-C. [1975], *Generating Inequality*, Basic Books, New York.

VERHAEST D. & OMEY E. [2012], Overeducation, undereducation and earnings: Further evidence on the importance of ability and measurement error bias, *Journal of Labor Research*, vol. 33, n°1, pp. 76-90.

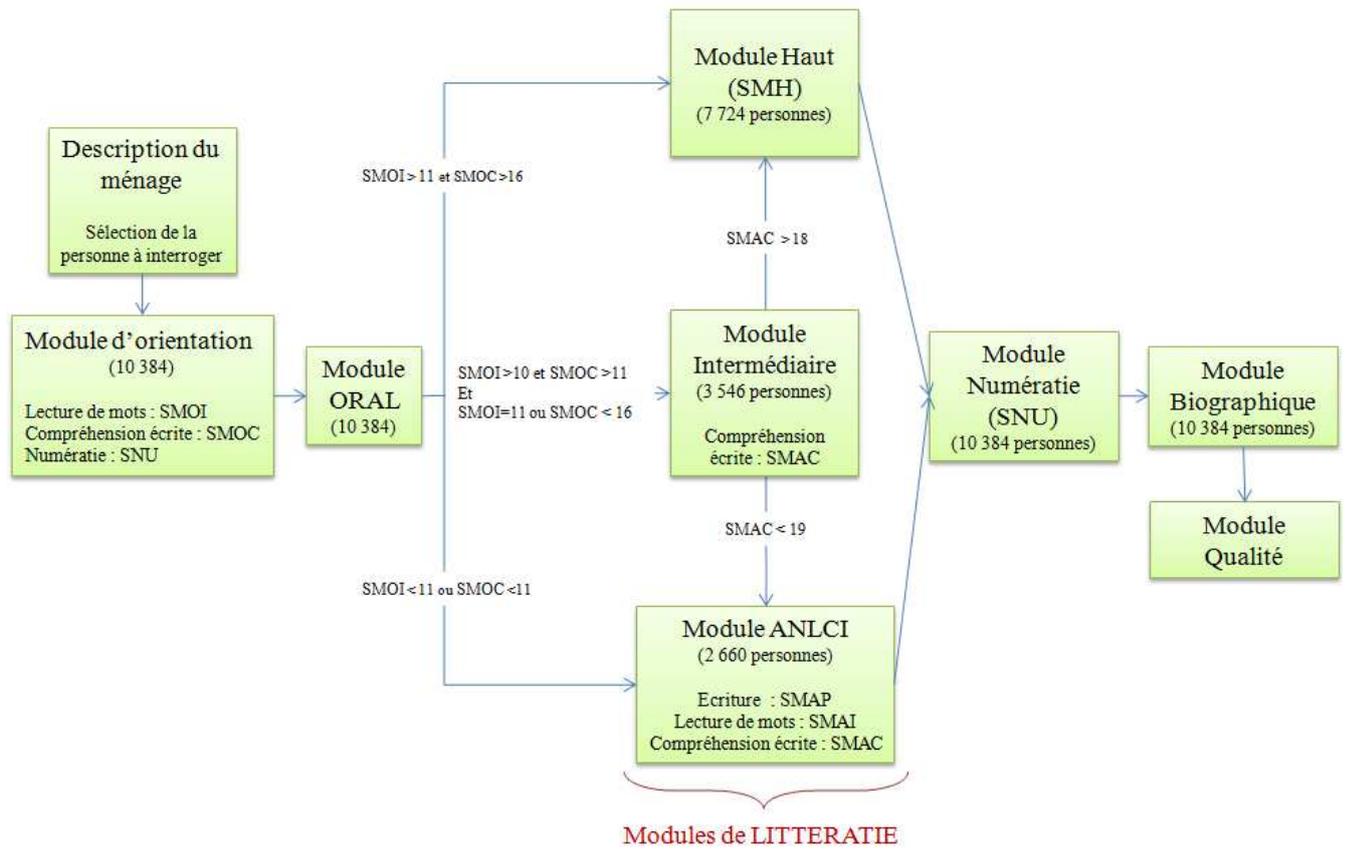
VIGNOLES A., DE COULON A. & MARCENARO-GUTIERREZ, O. [2011], The value of basic skills in the British labour market, *Oxford Economic Papers*, vol.63, n°1, pp.27-48.

WILLIS R. & ROSEN S. [1979], Education and self-selection, *The Journal of Political Economy*, vol. 87, n°5, pp. 7-36.

ZORA M., MEYER V., ROUAUD P., RYK F. & WINNICKI P. [2011], Le diplôme : un atout gagnant pour les jeunes face à la crise, *Bref*, n° 283 , CEREQ, 4 p.

ANNEXES DE L'ESSAI 1

ANNEXE A : ARCHITECTURE DU QUESTIONNAIRE IVQ 2004.



Imputation arbitraire

La première imputation repose sur l'hypothèse selon laquelle les personnes ayant été orientées vers le module bas n'auraient pas eu vraisemblablement les capacités nécessaires de répondre juste à des items plus difficiles, comme ceux du module haut. A l'inverse, les personnes du module haut auraient répondu juste à des items plus faciles, comme ceux du module bas. De fait, dans le premier cas, les valeurs manquantes sont imputées par 0 et dans le second cas, par 1. Toutefois, cette imputation apparaît comme trop rigide puisque les individus sont rarement divisés en deux groupes, entre ceux qui possèdent un très faible niveau de compétences et ceux qui en possèdent un très élevé (Murat et Rocher, 2009).

Imputation économétrique

La deuxième imputation est une imputation économétrique, également proposée par Murat et Rocher en 2009. L'idée est de modéliser les scores aux modules haut et bas, à partir des scores obtenus au module d'orientation, sachant que pour ce dernier, l'ensemble des individus de l'enquête y ont été soumis. Ceci doit alors permettre de relier l'ensemble des scores entre eux.

Ainsi, dans un premier temps, sur les groupes ayant passé la partie de compréhension écrite du module bas²⁴, on réalise une régression linéaire entre le score obtenu à celui-ci et le score obtenu au module d'orientation. Ceci doit permettre d'extrapoler les réponses des individus directement orientés vers le module haut comme des réponses attendues au module bas. Ensuite, selon le même principe, on réalise une régression linéaire sur les groupes ayant passé le module haut, entre le score à ce module et le score obtenu au module d'orientation. Ces deux régressions permettent alors d'établir les équations suivantes :

$$SMA = 1,23 + 0,73.SMO \quad \text{et} \quad SMH = 5,31 + 0,76.SMO$$

Nous pouvons dès lors imputer des scores pour les items qui n'ont pas été renseignés en raison de l'aiguillage des individus. Le score en littérature sera donné selon l'équation suivante :

$$SCLITT = SMO + SMA + (1,23 + 0,73.SMO) \times I_{indiv_mod_H} + SMH + (5,31 + 0,76.SMO) \times I_{indiv_mod_B}$$

Même si ce type d'imputation semble déjà plus adapté, car moins restrictif de la réalité que le premier, il possède aussi des limites. La première extrapolation repose sur une hypothèse facilement acceptable, vu qu'un nombre important de personnes ayant passé le module d'orientation avec un relatif succès est malgré tout amené à passer partiellement le module ANLCI²⁵. Néanmoins, la deuxième extrapolation repose sur une hypothèse beaucoup plus forte et risque de biaiser l'imputation, puisque peu d'individus du module haut direct ont obtenu un score inférieur à 7 sur 10 au module d'orientation. De fait, l'imputation économétrique est plus fragile, vu que les personnes du groupe ANLCI se concentrent justement en dessous des 7 points sur 10. Il existe donc un fort risque pour que les scores imputés pour les individus n'ayant pas passé le module haut soit surévalués par rapport à leurs réelles capacités.

²⁴ Les items de compréhension écrite du module bas sont aussi ceux du module intermédiaire. Le module intermédiaire sert à répartir de façon plus fiable les individus qui auraient eu un score moyen au module d'orientation, empêchant toute orientation directe vers le module haut ou le module bas.

²⁵ Il s'agit des items de compréhension écrite.

Imputation logistique basée sur le Modèle de Réponse à l'Item (MRI)

La dernière méthode suppose d'imputer des valeurs aux données manquantes par un modèle logistique, où la variable dépendante est l'item comportant des valeurs manquantes en raison de l'adaptation de l'épreuve au niveau de capacité des individus.

Les variables indépendantes utilisées sont les items du module d'orientation, communs à tous, pour lesquels nous avons des alphas de Cronbach supérieurs à 0,7²⁶ ainsi que des paramètres de difficulté²⁷ proches de l'item dont on cherche justement à imputer les données manquantes. En réalité, cela revient à réaliser un modèle de réponse à l'item (MRI) qui explique la probabilité de réussite à un item par le niveau de compétence du répondant et par les caractéristiques propres à l'item. Ainsi si on note θ la variable latente représentant le niveau de compétences des individus et $(\delta_1, \dots, \delta_j)$ un ensemble de paramètres de difficulté²⁸ pour chaque item, alors le MRI consiste à déterminer par la fonction logistique la probabilité $P(\theta, \delta_k)$ de répondre juste à l'item k sachant que le niveau de compétences de l'individu est θ .

Graphiquement, cette relation correspond à une courbe en forme de S (cf. figure ci-dessous). La difficulté de l'item est alors définie comme étant la valeur du trait latent à laquelle correspond une probabilité de réussite exactement égale à 0,5. C'est précisément cette valeur du trait latent que l'on appelle paramètre de difficulté. Dans le cadre de la figure proposé ci-dessous, plus la valeur du trait latent est proche de -4, plus l'item est facile et inversement. Il est effectivement logique que pour avoir une probabilité de réussite égale à 0,5 à un item plus difficile, le niveau de compétence, représenté par le trait latent, soit plus élevé.

L'avantage de ces modèles est de dissocier l'aptitude d'un individu de la difficulté de l'épreuve. L'aptitude d'une personne ne sera pas ainsi définie en fonction de la difficulté d'un item et inversement. Un item du module haut représenterait le même niveau de difficulté pour quelqu'un qui a été orienté vers ce module que pour quelqu'un qui ne l'a pas été. Seul le niveau de compétence entre les deux personnes diffère. Ce point constitue une hypothèse centrale des modèles de réponses à l'item (hypothèse d'indépendance locale). Or, il est clair que dans l'enquête IVQ, enquête adaptative, la difficulté des items de littérature est fonction du niveau de compétences détecté au préalable. L'idée est alors de figer l'élément difficulté de l'item en ne considérant que des items ayant le même degré de complexité. En procédant de la sorte, nous retombons bien sur l'idée que la probabilité de réussite à un item ne dépend que du niveau de compétences des individus, ce qui correspond à un MRI à un paramètre (niveau de capacité), c'est-à-dire à un modèle de Rasch. Pour réaliser ce type d'imputation, nous avons utilisé la commande `Imputeitems` sous le logiciel STATA, proposée et développée par Jean-Benoît Hardouin²⁹. Seules les variables exogènes significatives au seuil de 5% ont été conservées dans l'analyse. De fait, cette méthode est bien moins biaisée que dans l'imputation économétrique où l'on surévaluait le niveau de littérature d'une partie des personnes enquêtées.

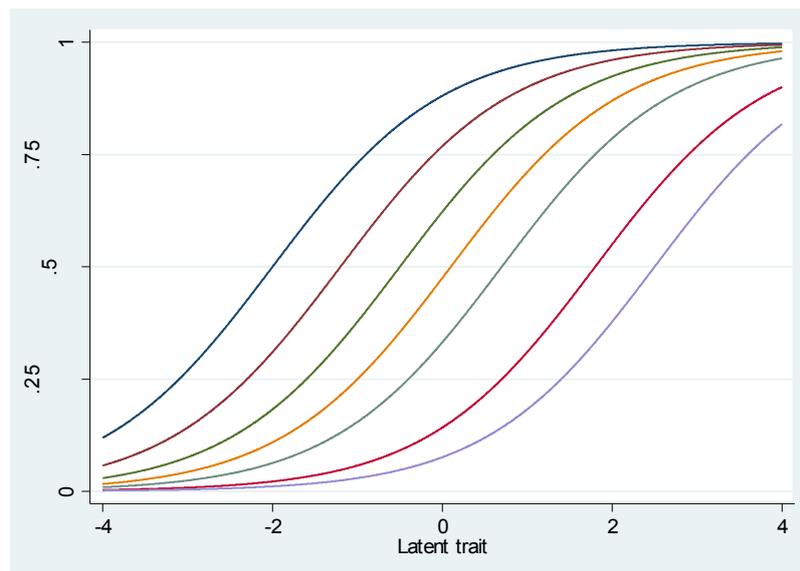
²⁶ Assurant ainsi une bonne cohérence interne des items

²⁷ Paramètres obtenus à posteriori de l'enquête, grâce à la commande `KR20` de STATA et correspondant aux coefficients de Kuder-Richardson.

²⁸ Les paramètres de difficulté sont déterminés à posteriori via le degré de réussite à l'item.

²⁹ Site de Jean-Benoît Hardouin : <http://www.anaqol.org/> Dernière connexion le 13/04/2013.

Illustration du Modèle de Réponse à l'Item à un paramètre (Modèle de Rasch³⁰)



Source : Jean-Benoît Hardouin (2005)

La création d'un score global de compétences en littératie peut donc être envisagée de trois manières différentes, même si la dernière méthode d'imputation semble être plus fiable que les deux autres : elle n'est pas empreinte d'hypothèses fortes et elle utilise une méthode déterministe rigoureuse. En effet, si d'après le modèle logistique, la probabilité de répondre juste à un item est inférieure à 0,5, alors la valeur manquante est imputée par 0 (échec). À l'inverse, si la probabilité est supérieure à 0,5, elle sera imputée par 1 (réussite). Cette méthode sera privilégiée pour construire le score global de littératie. La construction d'un score global nous permet d'obtenir une variable continue qu'il sera ensuite facile d'intégrer dans nos modèles statistiques visant à expliquer l'accès à l'emploi et les salaires individuels. Notons que l'ensemble des scores ont été ramenés sur une échelle pouvant aller de 0 à 100.

³⁰ Le modèle de Rasch est un modèle à un paramètre : la difficulté des items. Les MRI peuvent en effet se réaliser en tenant compte de plusieurs paramètres (hasard de la réponse donnée, pouvoir discriminant de l'item...).

ANNEXE C : TABLE DE CORRESPONDANCE D’AFFICHARD (1981) :

	Cadre, professeur, ingénieur	Profession intermédiaire	Technicien	Employé qualifié	Employé non qualifié	Ouvrier qualifié	Ouvrier non qualifié
Supérieur long	Normal	Déclassé	Déclassé	Déclassé	Déclassé	Déclassé	Déclassé
Supérieur court	Surclassé	Normal	Normal	Déclassé	Déclassé	Déclassé	Déclassé
Baccalauréat	Surclassé	Normal	Normal	Déclassé	Déclassé	Déclassé	Déclassé
BEP /CAP	Surclassé	Surclassé	Surclassé	Normal	Déclassé	Normal	Déclassé
BEPC - Brevet des collèges	Surclassé	Surclassé	Surclassé	Normal	Déclassé	Surclassé	Normal
Sans diplôme ou CEP	Surclassé	Surclassé	Surclassé	Surclassé	Normal	Surclassé	Normal

ANNEXES D : RÉSULTATS DES MODÉLISATIONS PROBIT ...

TABLEAU D.1 : ... DE LA PROBABILITÉ D'ÊTRE SURCLASSÉ/DÉCLASSÉ SUR LA POPULATION TOUT ENTIÈRE

VARIABLES	Surclassé	Déclassé
	Coefficients	Coefficients
Facteurs socio-démographiques		
Être une femme	-0,73***	0,61***
Vivre en ZUS	-0,03	0,04
Scores en littératie	0,006**	-0,011***
Scores en numératie	0,008***	-0,008***
Scores en compréhension orale	0,003**	-0,0008
Père cadre	0,23**	-0,30***
Spécialité de formation		
Formation disciplinaire	0,336***	-0,189**
Formation industrielle	-0,471***	-0,300***
Formation tertiaire	<i>Référence</i>	
Plus haut diplôme obtenu³¹		
Jamais scolarisé, CEP ou sans diplôme	1,09***	--
BEPC ou Brevet des collèges	0,98***	0,23**
BEP/CAP	0,75***	0,43***
Baccalauréat	<i>Référence</i>	
Bac+2	0,58***	-0,11
> Bac+2	--	0,62***
Trajectoires professionnelles sur le marché du travail		
Période de chômage de plus d'un an par le passé	-0,16***	0,27***
En emploi précaire	-0,20***	0,25***
Expérience professionnelle	0,02***	-0,02***
Facteurs environnementaux		
Taux de chômage départemental	0,0008	-0,004
Salarié du privé	0,197**	-0,314***
Secteur d'activité de l'entreprise		
Agriculture, sylviculture et pêche	-0,74***	0,87***
Industrie	0,34***	-0,42***
Construction et énergie	0,41***	-0,47***
Commerce	0,11	0,21***
Transports	0,30**	-0,39***
Activités financières et immobilières	0,29**	-0,26**
Services aux entreprises et aux particuliers	-0,05	0,21***
Éducation, santé, action sociale	<i>Référence</i>	
Administration	-0,004	0,10
Constante	-3,21***	1,45***
Observations	4632	4300
Pseudo R ²	22,03%	17,26%

*** : significatif à 1% ; ** : significatif à 5% ; * : significatif à 10%

³¹ Les individus sans diplôme ou avec un Certificat d'Études Primaires ne peuvent être considérés comme déclassés. À l'inverse, les individus ayant un diplôme supérieur ou égal à BAC+3 ne peuvent être considérés comme surclassés. Les premiers sont donc omis de l'échantillon sur le déclassement et les seconds de l'échantillon sur le surclassement.

TABLEAU D.2: ... DE LA PROBABILITÉ D'ÊTRE SURCLASSÉ/DÉCLASSÉ SELON LA TRANCHE D'ÂGE D'APPARTENANCE

VARIABLES	SURCLASSÉ		DÉCLASSÉ	
	De 18 à 40 ans	Plus de 40 ans	De 18 à 40 ans	Plus de 40 ans
	Coefficients	Coefficients	Coefficients	Coefficients
Facteurs socio-démographiques				
Être une femme	-0,62***	-0,75***	0,63***	0,61***
Vivre en ZUS	0,06	-0,12	-0,06	0,16
Scores en littératie	0,019***	0,002	0,001	0,003
Scores en numératie	0,008***	0,007***	-0,004**	-0,003
Scores en compréhension orale	0,0002	0,004***	0,0005	-0,001
Père cadre	0,19	0,08	-0,24**	-0,30**
Spécialité de formation				
Formation disciplinaire	0,42***	-0,04	-0,51***	-0,52***
Formation industrielle	-0,32***	-0,52***	-0,31***	-0,27***
Formation tertiaire	<i>Référence</i>			
Plus haut diplôme obtenu				
Jamais scolarisé, CEP ou sans diplôme	1,78***	1,30***	--	--
BEPC ou Brevet des collèges	1,44***	1,26***	0,65***	1,19***
BEP/CAP	1,10***	0,76***	-0,63***	1,07***
Baccalauréat	0,98***	0,73***	0,09	0,13
Bac+2	--	--	0,83***	1,07***
Trajectoires professionnelles sur le marché du travail				
Période de chômage de plus d'un an par le passé	-0,11	-0,25***	0,31***	0,12
En emploi précaire	-0,18*	-0,16	0,17**	0,28**
Expérience professionnelle	0,04***	0,009**	-0,03***	-0,025**
Facteurs environnementaux				
Taux de chômage départemental	-0,01	0,01	0,004	-0,017
Salarié du privé	0,11	0,27***	-0,30***	-0,35***
Secteur d'activité de l'entreprise				
Agriculture, sylviculture et pêche	-0,62**	-0,90***	0,90***	0,52*
Industrie	0,25*	0,41***	-0,33***	-0,78***
Construction et énergie	0,21	0,57***	-0,34***	-0,85***
Commerce	0,08	0,23***	0,24**	0,03
Transports	-0,005	0,63***	-0,30***	-0,77***
Activités financières et immobilières	0,41**	0,25	-0,25	-0,31**
Services aux entreprises et aux particuliers	0,22*	-0,19*	0,06	0,23**
<i>Éducation, santé, action sociale</i>				
Administration	-0,04	0,07	-0,01	0,13
Constante	-4,11***	-2,62***	-0,015	-0,43
Observations	2731	2389	2878	2152
Pseudo R ²	26,91%	20,83%	18,45%	26,11%

*** : significatif à 1% ; ** : significatif à 5% ; * : significatif à 10%

TABLEAU D.3 : ... DE LA PROBABILITÉ D'ÊTRE SURCLASSÉ/DÉCLASSÉ SELON LE NIVEAU DE DIPLÔME

VARIABLES	SURCLASSÉ		DÉCLASSÉ	
	Secondaire	Supérieur	Secondaire	Supérieur
	Coefficients	Coefficients	Coefficients	Coefficients
Facteurs socio-démographiques				
Être une femme	-0,49***	-0,17	0,702***	0,460***
Vivre en ZUS	-0,025	-0,28	0,049	0,079
Scores en littératie	0,0154***	0,0006	0,0065	-0,0303***
Scores en numératie	0,0090***	0,0053	-0,0072***	-0,0111***
Scores en compréhension orale	0,0038**	-0,0023	-0,0033**	0,0038
Père cadre	0,229	-0,219	0,223	-0,352***
Spécialité de formation				
Formation disciplinaire	0,359**	-0,270*	-0,277**	-0,088
Formation industrielle	-0,301***	0,072	-0,199**	-0,471***
Formation tertiaire	<i>Référence</i>		<i>Référence</i>	
Plus haut diplôme obtenu³²				
BEPC ou Brevet des collèges	1,01***	--	0,29**	--
BEP/CAP	0,80***	--	0,41***	--
Baccalauréat	<i>Référence</i>		<i>Référence</i>	
Bac+2	--	--	--	-0,833***
Bac et 3 et plus	--	--	--	<i>Référence</i>
Trajectoires professionnelles sur le marché du travail				
Période de chômage de plus d'un an par le passé	-0,19**	-0,18	0,20**	0,321***
En emploi précaire	-0,27**	0,076	0,291***	0,216**
Expérience professionnelle	0,03***	0,032***	-0,015***	-0,026***
Facteurs environnementaux				
Taux de chômage départemental	-0,003	0,011	-0,02	0,009
Salarié du privé	0,06	0,22	-0,67***	0,06
Secteur d'activité de l'entreprise				
Agriculture, sylviculture et pêche	-0,80**	-- ³³	10,01***	0,39
Industrie	0,26**	0,31	-0,55***	-0,26*
Construction et énergie	0,15	0,40*	-0,83***	0,17
Commerce	0,16	0,01	0,121	0,42***
Transports	0,05	0,02	-0,97***	0,42**
Activités financières et immobilières	0,40**	0,30	-0,80***	-0,10
Services aux entreprises et aux particuliers	-0,03	0,58***	0,034***	-0,29**
<i>Éducation, santé, action sociale</i>	<i>Référence</i>			
Administration	-0,07	0,09	-0,16	0,13
Constante	-4,34***	-2,59**	1,53***	3,55***

³² La variable « Plus haut diplôme obtenu » est une variable polytomique pour les analyses menées sur les diplômés du secondaire, la référence étant le baccalauréat. Inversement, il s'agit d'une variable dichotomique pour les analyses menées sur les diplômés du supérieur. Toutefois, elle ne sera pas incluse dans le modèle de surclassement des diplômés du supérieur en raison de sa colinéarité potentielle : en effet, la comparaison se ferait par rapport aux individus ayant un niveau supérieur ou égal à Bac+3, or ces individus ne peuvent être surclassés compte tenu de leur niveau de diplôme déjà élevé.

³³ Le secteur d'activité « agriculture, sylviculture et pêche » a également été ôté des variables explicatives pour ce modèle en raison de problème de colinéarité.

Observations	2761	1539	2761	1539
Pseudo R ²	17,99%	10,81%	23,01%	16,33%

*** : significatif à 1% ; ** : significatif à 5% ; * : significatif à 10%

ANNEXE E: RÉSULTATS DE LA MODÉLISATION LOGISTIQUE ORDONNÉE SUR LE SENTIMENT D'UTILITÉ DE L'ÉCOLE DANS LA VIE PROFESSIONNELLE.

VARIABLES	Coefficients
Individus déclassés et étant en déficit en numératie	-0,352***
Individus déclassés et n'étant ni en déficit ni en surplus de numératie	-0,255***
Individus déclassés et étant en surplus en numératie	0,249
Individus ni déclassés ni surclassés et étant en déficit en numératie	-0,161
Individus ni déclassés ni surclassés et n'étant ni en déficit ni en surplus en numératie	<i>Référence</i>
Individus ni déclassés ni surclassés et étant en surplus en numératie	0,305**
Individus surclassés et étant en déficit en numératie	-0,644**
Individus surclassés et n'étant ni en déficit ni en surplus de numératie	-0,387***
Individus surclassés et étant en surplus en numératie	-0,017
Être une femme	0,098
Vivre en ZUS	0,217**
Age	0,014***
Être d'origine étrangère	0,232**
Spécialité de formation	
Formation disciplinaire	-0,154**
Formation industrielle	0,056
Formation tertiaire	<i>Référence</i>
Période de chômage de plus d'un an par le passé	-0,233***
Tranches de salaire mensuel	
< 500€ par mois	-0,218
Entre 500€ et 1000€ par mois	-0,237***
Entre 1000€ et 1500€ par mois	<i>Référence</i>
Entre 1500€ et 2000€ par mois	0,119
Plus de 2000€ par mois	0,377***
Secteur d'activité de l'entreprise	
Industrie	-0,149
Construction et énergie	-0,295***
Commerce	-0,286**
Transports	-0,580***
Activités financières et immobilières	-0,712***
Services aux entreprises et aux particuliers	-0,193**
<i>Éducation, santé, action sociale</i>	<i>Référence</i>
Administration	-0,417***
	<i>Cut1</i> -2,81
	<i>Cut2</i> -1,56
	<i>Cut3</i> 0,54
Observations	4300
Pseudo R ²	1,96%

*** : significatif à 1% ; ** : significatif à 5% ; * : significatif à 10%

ANNEXE F : PROBABILITÉS CONDITIONNELLES DE CONSIDÉRER QUE L'ÉCOLE A ÉTÉ PLUS OU MOINS UTILE POUR LA VIE PROFESSIONNELLE

	Pas du tout d'accord	Plutôt pas d'accord	Plutôt d'accord	Tout à fait d'accord
Individus déclassés et étant en déficit en numératie	0.05	0.11	0.44	0.40
Individus déclassés et n'étant ni en déficit ni en surplus de numératie	0.05	0.11	0.44	0.40
Individus déclassés et étant en surplus en numératie	0.03	0.07	0.37	0.53
Individus ni déclassés ni surclassés et étant en déficit en numératie	0.04	0.09	0.41	0.46
Individus ni déclassés ni surclassés et n'étant ni en déficit ni en surplus en numératie	0.04	0.08	0.39	0.49
Individus ni déclassés ni surclassés et étant en surplus en numératie	0.03	0.06	0.34	0.58
<i>Individus surclassés et étant en déficit en numératie</i>	0.07	0.13	0.45	0.35
Individus surclassés et n'étant ni en déficit ni en surplus de numératie	0.05	0.10	0.43	0.42
Individus surclassés et étant en surplus en numératie	0.04	0.08	0.38	0.50

ANNEXE G : QUOTITÉ DE TRAVAIL (EN %) DANS L'ENQUÊTE EMPLOI 2004 – 4^{ÈME} TRIMESTRE (H=HOMMES ; F=FEMMES)

	Agriculteurs		Artisans		Cadres		Professions		Employés		Ouvriers		% moyen par NAF
	H	F	H	F	H	F	H	F	H	F	H	F	
Agriculture, sylviculture et pêche	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	94,66	75,00	90,91	80,00	83,33	50,0	86,21	70,00	100,0	62,7	92,91	64,2	84,60
<i>Temps partiel subi</i>	0,36	1,14	0,00	0,00	0,00	0,00	0,00	5,00	0,00	9,30	1,57	11,4	2,20
<i>Temps partiel choisi</i>	4,98	23,86	9,09	20,00	16,67	50,0	13,79	25,00	0,00	27,9	5,51	24,2	13,20
Industries agricoles et alimentaires	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	100,0	97,92	87,50	100,0	65,2	96,15	86,05	100,0	73,2	97,87	75,0	88,37
<i>Temps partiel subi</i>	0,00	0,00	0,00	0,00	0,00	8,70	0,00	2,33	0,00	15,1	0,53	7,50	3,83
<i>Temps partiel choisi</i>	0,00	0,00	2,08	12,50	0,00	26,0	3,85	11,63	0,00	11,6	1,60	17,5	7,81
Industrie des biens de	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	100,0	91,30	87,50	91,80	82,2	91,67	78,31	100,0	60,0	97,06	80,8	85,63
<i>Temps partiel subi</i>	0,00	0,00	0,00	0,00	1,64	0,00	2,78	2,41	0,00	6,15	0,00	4,17	2,14
<i>Temps partiel choisi</i>	0,00	0,00	8,70	12,50	6,56	17,7	5,56	19,28	0,00	33,8	2,94	15,0	12,23
Industrie automobile	100	100	100	Pas d'obs.	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	100,0	100,0		100,0	73,6	100,0	75,00	93,75	64,5	100,0	77,5	89,70
<i>Temps partiel subi</i>	0,00	0,00	0,00		0,00	5,26	0,00	0,00	3,13	22,5	0,00	5,00	3,33
<i>Temps partiel choisi</i>	0,00	0,00	0,00		0,00	21,0	0,00	25,00	3,13	12,9	0,00	17,5	6,97
Industrie des biens d'équipement	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	75,00	95,45	86,67	98,84	67,1	95,39	75,00	98,52	68,9	98,94	75,5	87,78
<i>Temps partiel subi</i>	0,00	0,00	0,00	6,67	0,00	1,43	1,32	1,67	0,00	5,43	0,53	8,15	2,35
<i>Temps partiel choisi</i>	0,00	25,00	4,55	6,67	1,16	31,4	3,29	23,33	1,48	25,5	0,53	16,3	9,88
Industrie des biens intermédiaires	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	75,00	95,45	86,67	98,84	67,1	95,39	75,00	98,52	68,9	98,94	75,5	87,52
<i>Temps partiel subi</i>	0,00	0,00	0,00	6,67	0,00	1,43	1,32	1,67	0,00	5,43	0,53	8,15	2,53
<i>Temps partiel choisi</i>	0,00	25,00	4,55	6,67	1,16	31,4	3,29	23,33	1,48	25,5	0,53	16,3	9,96
Énergie	Pas	Pas	Pas	Pas	100	100	100	100	100	100	100	100	100

<i>Temps plein</i>	d'obs.	d'obs.	d'obs.	d'obs.	90,00	62,5	94,92	80,00	93,75	78,2	100,0	80,0	87,10
<i>Temps partiel subi</i>					0,00	0,00	1,69	0,00	6,25	4,35	0,00	20,0	2,15
<i>Temps partiel choisi</i>					10,00	37,5	3,39	20,00	0,00	17,3	0,00	0,00	10,75
Construction	100		100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	Pas d'obs.	98,50	46,15	100,0	73,3	97,71	80,46	97,66	50,3	97,72	66,4	84,04
<i>Temps partiel subi</i>	0,00		0,00	3,85	0,00	3,33	0,00	2,30	0,58	17,1	0,76	14,7	4,51
<i>Temps partiel choisi</i>	0,00		1,50	50,00	0,00	23,3	2,29	17,24	1,75	32,5	1,52	18,8	11,45
Commerce	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	80,00	97,54	81,87	95,29	72,0	97,40	79,05	93,80	63,2	96,49	70,9	83,78
<i>Temps partiel subi</i>	0,00	0,00	0,99	1,04	0,00	4,07	0,52	4,47	2,13	13,8	1,41	12,3	5,01
<i>Temps partiel choisi</i>	0,00	20,00	1,48	17,10	4,71	23,8	2,08	16,48	4,07	22,9	2,11	16,7	11,21
Transports	100		100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	Pas d'obs.	91,67	60,00	97,83	86,0	95,42	79,61	94,89	76,7	94,82	68,7	86,73
<i>Temps partiel subi</i>	0,00		0,00	6,67	2,17	2,33	0,00	3,88	1,46	6,16	2,79	15,6	4,39
<i>Temps partiel choisi</i>	0,00		8,33	33,33	0,00	11,6	4,58	16,50	3,65	17,1	2,39	15,6	8,88
Activités financières	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	75,00	50,00	100,0	100,0	99,10	75,0	93,52	84,13	98,88	75,6	94,44	75,0	87,29
<i>Temps partiel subi</i>	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,59	0,00	2,61	0,00	6,25	0,85
<i>Temps partiel choisi</i>	25,00	50,00	0,00	0,00	0,90	25,0	6,48	14,29	1,12	21,7	5,56	18,7	11,86
Activités immobilières			100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	Pas d'obs.	Pas d'obs.	100,0	75,00	91,67	71,4	97,37	76,79	95,59	77,2	90,63	57,8	84,38
<i>Temps partiel subi</i>			0,00	0,00	4,17	0,00	0,00	3,57	0,00	7,58	3,13	15,7	3,41
<i>Temps partiel choisi</i>			0,00	25,00	4,17	28,5	2,63	19,64	4,41	15,1	6,25	26,3	12,22
Services aux entreprises	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	60,00	100,0	76,32	98,55	78,8	96,34	77,81	91,10	72,8	94,26	62,5	85,26
<i>Temps partiel subi</i>	0,00	0,00	0,00	0,00	0,00	1,71	1,57	5,21	2,88	9,48	2,73	16,4	4,36
<i>Temps partiel choisi</i>	0,00	40,00	0,00	23,68	1,45	19,4	2,09	16,99	6,02	17,7	3,01	21,0	10,38
Service aux particuliers	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Temps plein</i>	100,0	85,71	94,97	85,45	87,21	73,0	91,91	71,35	88,70	53,4	93,85	61,8	77,26
<i>Temps partiel subi</i>	0,00	14,29	1,26	1,21	4,65	2,25	1,73	5,26	3,77	19,2	3,69	20,4	8,32
<i>Temps partiel choisi</i>	0,00	0,00	3,77	13,33	8,14	24,7	6,36	23,39	7,54	27,2	2,46	17,6	14,42

Éducation, santé, action sociale	100												
<i>Temps plein</i>	100,0	45,45	96,08	71,43	92,28	75,2	94,46	76,78	95,69	69,1	92,19	68,4	82,27
<i>Temps partiel subi</i>	0,00	0,00	1,96	11,43	1,25	2,91	0,41	3,83	1,69	12,3	2,40	13,6	4,78
<i>Temps partiel choisi</i>	0,00	54,55	1,96	17,14	6,47	21,8	5,14	19,39	2,62	18,4	5,41	17,9	12,95
Administration	100												
<i>Temps plein</i>	100,0	50,00	93,33	80,00	95,31	71,3	94,29	71,35	94,74	74,1	92,15	66,1	83,04
<i>Temps partiel subi</i>	0,00	0,00	0,00	0,00	0,47	4,37	1,50	5,34	2,44	8,95	3,14	13,8	4,86
<i>Temps partiel choisi</i>	0,00	50,00	6,67	20,00	4,23	24,2	4,20	23,31	2,82	16,9	4,71	20,0	12,11
TOTAL	100												
<i>Temps plein</i>	95,36	73,66	96,83	76,92	95,37	74,9	95,31	76,72	94,43	67,1	95,71	69,4	83,92
<i>Temps partiel subi</i>	0,29	1,34	0,63	2,15	0,74	2,57	0,80	4,07	1,96	11,8	1,64	12,4	4,48
<i>Temps partiel choisi</i>	4,35	25,00	2,54	20,92	3,88	22,4	3,88	19,21	3,61	21,0	2,65	18,1	11,60

ANNEXES H : TESTS DU KHI-DEUX

TABLEAUX H.1 : ... PAR CSP

Tests du Khi-deux - INDUSTRIE				
Catégories socioprofessionnelles		Valeur	ddl	Signification asymptotique (bilatérale)
Cadres	Khi-deux de Pearson	26,643 ^a	2	,000
	Rapport de vraisemblance	27,716	2	,000
	Association linéaire par linéaire	26,203	1	,000
	Nombre d'observations valides	395		
Professions Intermédiaires	Khi-deux de Pearson	29,514 ^b	2	,000
	Rapport de vraisemblance	30,404	2	,000
	Association linéaire par linéaire	23,433	1	,000
	Nombre d'observations valides	792		
Employés	Khi-deux de Pearson	25,147 ^c	2	,000
	Rapport de vraisemblance	25,037	2	,000
	Association linéaire par linéaire	20,806	1	,000
	Nombre d'observations valides	1109		
Ouvriers	Khi-deux de Pearson	32,797 ^d	2	,000
	Rapport de vraisemblance	33,342	2	,000
	Association linéaire par linéaire	26,563	1	,000
	Nombre d'observations valides	949		
a. 1 cellules (16,7%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 4,78.				
b. 0 cellules (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 13,51.				
c. 0 cellules (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 61,10.				
d. 0 cellules (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 44,30.				

Tests du Khi-deux - SERVICES				
Catégories socioprofessionnelles		Valeur	ddl	Signification asymptotique (bilatérale)
Cadres	Khi-deux de Pearson	43,035 ^a	2	,000
	Rapport de vraisemblance	43,298	2	,000
	Association linéaire par linéaire	41,917	1	,000
	Nombre d'observations valides	1363		
Professions Intermédiaires	Khi-deux de Pearson	88,314 ^b	2	,000
	Rapport de vraisemblance	87,751	2	,000
	Association linéaire par linéaire	87,077	1	,000
	Nombre d'observations valides	2087		
Employés	Khi-deux de Pearson	50,568 ^c	2	,000
	Rapport de vraisemblance	49,868	2	,000
	Association linéaire par linéaire	42,840	1	,000
	Nombre d'observations valides	2751		
Ouvriers	Khi-deux de Pearson	21,927 ^d	2	,000
	Rapport de vraisemblance	22,117	2	,000
	Association linéaire par linéaire	9,673	1	,002
	Nombre d'observations valides	1002		
a. 0 cellule (,0%) n'a un effectif théorique inférieur à 5. L'effectif théorique minimum est de 16,04.				
b. 0 cellule (,0%) n'a un effectif théorique inférieur à 5. L'effectif théorique minimum est de 38,51.				
c. 0 cellule (,0%) n'a un effectif théorique inférieur à 5. L'effectif théorique minimum est de 131,32.				
d. 0 cellule (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 76,95.				

TABLEAUX H.2 : ... SELON LA CHARGE D'ENFANTS

Tests du Khi-deux - INDUSTRIE				
Charge d'enfant(s)		Valeur	ddl	Signification asymptotique (bilatérale)
Pas de charge d'enfant(s)	Khi-deux de Pearson	62,487 ^a	6	,000
	Rapport de vraisemblance	66,698	6	,000
	Association linéaire par linéaire	2,530	1	,112
	Nombre d'observations valides	1681		
Charge d'enfant(s)	Khi-deux de Pearson	50,392 ^b	6	,000
	Rapport de vraisemblance	54,402	6	,000
	Association linéaire par linéaire	,510	1	,475
	Nombre d'observations valides	1564		
a. 0 cellules (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 17,99.				
b. 0 cellules (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 14,23.				

Tests du Khi-deux - SERVICES				
Charge d'enfant(s)		Valeur	ddl	Signification asymptotique (bilatérale)
Pas de charge d'enfant(s)	Khi-deux de Pearson	157,469 ^a	6	,000
	Rapport de vraisemblance	162,557	6	,000
	Association linéaire par linéaire	24,044	1	,000
	Nombre d'observations valides	4043		
Charge d'enfant(s)	Khi-deux de Pearson	72,510 ^b	6	,000
	Rapport de vraisemblance	78,358	6	,000
	Association linéaire par linéaire	,006	1	,936
	Nombre d'observations valides	3160		
a. 0 cellule (,0%) n'a un effectif théorique inférieur à 5. L'effectif théorique minimum est de 46,02.				
b. 0 cellule (,0%) n'a un effectif théorique inférieur à 5. L'effectif théorique minimum est de 36,75.				

ANNEXES I : ESTIMATION DES FONCTIONS DE GAINS PAR LA MÉTHODE DES MOINDRES CARRÉS ORDINAIRES (MCO).

TABLEAU I.1 : ... AVEC LE NIVEAU DE DIPLÔME

<i>Variable dépendante : logarithme du salaire mensuel</i>	Modèle 1	Modèle 2	Modèle 3	Modèle 4	Modèle 5	Modèle 6
<i>Réf. : CAP/BEP</i>						
Jamais scolarisé ou sans diplôme	-0.157 (4.78)***	-0.102 (2.92)***	-0.095 (2.82)***	-0.127 (3.76)***	-0.072 (2.06)**	-0.069 (1.95)
Certificat d'études primaires	-0.135 (3.14)***	-0.111 (2.60)***	-0.111 (2.61)***	-0.122 (2.84)***	-0.101 (2.37)**	-0.099 (2.33)**
BEPC/brevet des collèges	0.086 (1.75)	0.074 (1.52)	0.067 (1.39)	0.082 (1.68)	0.063 (1.30)	0.064 (1.31)
Baccalauréat général	0.210 (5.13)***	0.188 (4.61)***	0.175 (4.30)***	0.196 (4.79)***	0.168 (4.11)***	0.166 (4.07)***
Baccalauréat professionnel	0.1251 (3.57)***	0.113 (3.23)***	0.105 (3.03)***	0.118 (3.39)***	0.101 (2.91)***	0.106 (2.89)***
Baccalauréat technique	0.200 (4.83)***	0.178 (4.31)***	0.167 (4.06)***	0.187 (4.51)***	0.159 (3.86)***	0.158 (3.83)***
Bac+2	0.333 (12.75)***	0.311 (11.81)***	0.295 (11.20)***	0.321 (12.26)***	0.288 (10.88)***	0.288 (10.85)***
Au dessus de Bac+2	0.544 (23.05)***	0.515 (21.19)***	0.493 (20.09)***	0.532 (22.44)***	0.483 (19.49)***	0.484 (19.51)***
Expérience professionnelle	0.0291 (10.75)***	0.0294 (10.92)***	0.0293 (10.98)***	0.0291 (10.82)***	0.0295 (11.04)***	0.0295 (11.04)***
Expérience professionnelle ²	-0.00040 (6.39)***	-0.00041 (6.50)***	-0.00041 (6.57)***	-0.00040 (6.43)***	-0.00041 (6.61)***	-0.00041 (6.60)***
Score en littérature		0.005 (4.62)***			0.003 (2.51)**	0.002 (2.02)**
Score en numératie			0.004 (6.61)***		0.003 (5.33)***	0.003 (4.99)***
Score en compréhension orale				0.002 (3.79)***		0.001 (1.24)
Constante	6.879 (236.50)***	6.479 (70.99)***	6.608 (132.12)***	6.739 (143.54)***	6.416 (70.23)***	6.413 (70.18)***
Observations (Hommes)	1877	1877	1877	1877	1877	1877
R-squared	0.29	0.30	0.31	0.30	0.31	0.31
Absolute value of t statistics are in parentheses * significatif au seuil de 10% ; ** significatif au seuil de 5% ; *** significatif au seuil de 1%						

TABLEAU I.2 : ... AVEC LE NOMBRE D'ANNÉES D'ÉTUDES

<i>Variable dépendante : logarithme du salaire mensuel</i>	Modèle 1	Modèle 2	Modèle 3	Modèle 4	Modèle 5	Modèle 6
Durée de scolarité théorique	0.066 (24.84)***	0.061 (21.11)***	0.058 (19.93)***	0.064 (23.26)***	0.057 (18.84)***	0.057 (18.82)***
Expérience professionnelle	0.022 (8.24)***	0.022 (8.35)***	0.022 (8.54)***	0.022 (8.27)***	0.022 (8.55)***	0.022 (8.54)***
Expérience professionnelle au carré	-0.0002 (3.17)***	-0.0002 (3.19)***	-0.0002 (3.41)***	-0.0002 (3.13)***	-0.0002 (3.39)***	-0.0002 (3.37)***
Score en littératie		0.004 (4.00)***			0.002 (1.76)	0.001 (1.35)
Score en numératie			0.004 (6.36)***		0.0032 (5.23)***	0.0030 (4.91)***
Score en compréhension orale				0.002 (3.48)***		0.0005 (1.00)
Constante	6.281 (145.6)***	5.9931 (71.48)***	6.1023 (119.3)***	6.1797 (118.8)***	5.986 (71.88)***	5.9867 (71.89)***
Observations (hommes)	1877	1877	1877	1877	1877	1877
R-squared	0.27	0.28	0.29	0.28	0.29	0.29

* significatif au seuil de 10% ; ** significatif au seuil de 5% ; *** significatif au seuil de 1%

ESSAI 2

LE RÔLE DES COMPÉTENCES DE BASE DANS LA PARTICIPATION AU MARCHÉ DU TRAVAIL ET L'EMPLOYABILITÉ

1. Introduction

Dans la théorie de l'offre de travail, le chômage n'existe pas et l'inactivité résulte d'un choix rationnel des agents. Mais si l'on se place dans un contexte de rigidités, une nécessaire distinction entre chômage et inactivité doit être faite. En effet, des personnes sont sans emploi alors qu'elles souhaiteraient vivement en avoir un et qu'elles ont entamé des recherches actives pour en trouver un. Ainsi, de par leurs démarches, elles sont donc encore sur le marché du travail, ce qui rejoint la définition statistique du terme « chômeur ». Inversement, certains ne recherchent plus de travail et se retirent du marché sans que cela ne relève d'un choix rationnel : inemployables, ils se sont découragés au point de devenir inactifs.

De nombreux facteurs tangibles (l'état de santé, les raisons familiales, le niveau de diplôme ou encore la mobilité géographique) peuvent expliquer les trajectoires professionnelles des individus. Mais qu'en est-il des compétences de base ? Peu de travaux se sont penchés sur l'effet qu'elles pouvaient avoir sur l'inemployabilité et donc l'inactivité, malgré leur hétérogénéité pour un même niveau de diplôme et pour une même expérience professionnelle. Dans la lignée des travaux empiriques

existants que nous détaillerons ultérieurement, cet essai a donc pour objectif d'apporter un éclairage empirique sur le rôle joué par les compétences dans la participation au marché du travail et l'accès à l'emploi. Plus spécifiquement, cet essai s'intéresse à la valorisation des compétences de base des adultes en termes d'accès à l'emploi et de participation au marché du travail. En effet, comme il est difficile d'appréhender le niveau général de compétences d'un individu, celles-ci pouvant être de natures différentes (compétences, cognitives, sociales, transférables, spécifiques, professionnelles, techniques...), nous nous focalisons sur les compétences de base, puisqu'elles sont supposées conditionner et favoriser l'apprentissage ultérieur de toute autre compétence plus avancée (Pont et Werquin, 2001). De plus, le niveau de compétences de base devient un critère de plus en plus important aux yeux des recruteurs (Brown et Hesketh, 2004). Dans cette optique, les compétences de base des adultes sont indispensables pour qu'un individu s'intègre avec succès sur le marché du travail tout en restant indépendant dans sa vie quotidienne et personnelle. Parce qu'elles sont supposées prémunir du chômage dès lors qu'elles sont élevées ou tout au moins demandées dans l'emploi d'après les théories standards, elles peuvent améliorer la compréhension des positions individuelles occupées face au marché du travail (emploi, chômage, inactivité). Dans un contexte d'enrichissement des emplois en termes de qualifications et de compétences, les compétences de base paraissent en ce sens indispensables pour augmenter son potentiel d'employabilité. Néanmoins, il existe de nombreuses autres variables pouvant expliquer pourquoi les individus sont plus ou moins intégrés dans l'emploi. Ces variables sont principalement l'âge, la charge familiale, les problèmes de santé ou encore les contraintes liées à l'emploi.

Cumulé à ces facteurs déterminants, un déficit de compétences de base peut contribuer à expliquer les trajectoires professionnelles individuelles. Si dans nos sociétés développées des individus sortent du système scolaire en ne maîtrisant pas ou en maîtrisant mal les compétences de base, c'est que les inégalités sociales et familiales n'ont pu être enrayerées par le système éducatif, voire pire, se sont creusées en son sein. De cette accumulation des inégalités naît des différences interindividuelles de dotations en capital humain, différences pouvant expliquer en partie la diversité des situations occupées et qui risquent de s'exacerber dans un contexte de marché du travail segmenté.

2. Recherche d'emploi et capital humain

D'après les théories orthodoxes, l'individu, en tant qu'agent rationnel maximisateur de son utilité, a un rapport essentiellement financier à l'égard du travail. En effet, à revenu égal entre salaire et indemnités chômage ou allocations, sa préférence ira pour le loisir plutôt que pour le travail. Une trappe d'inactivité existe dès lors qu'un individu dépourvu d'emploi n'a financièrement pas intérêt à accepter les emplois auxquels il peut prétendre. En s'appuyant sur les principes de la théorie de recherche d'emploi qui s'inscrit dans la pensée orthodoxe, il serait même possible d'envisager un certain ordre dans le choix des situations occupées face au marché du travail (emploi, chômage, inactivité). Cet ordre dépendrait essentiellement de caractéristiques personnelles telles que le niveau de diplôme. Ainsi, les individus les moins qualifiés auraient des coûts assez élevés de recherche d'emploi, que ce soit à l'entrée du marché du travail ou suite à un licenciement dont ils seront les premières victimes. Leur faible niveau de diplôme les amènerait à une recherche longue et laborieuse avec une faible probabilité d'être embauché puisque les employeurs auront une faible utilité à le faire, surtout pour une quantité de travail demandée restreinte. Ces derniers préféreront se tourner vers des individus *a priori* plus compétents de par les signaux qu'ils renvoient, notamment le diplôme, car ils sont supposés plus productifs, plus aptes à apprendre et par conséquent moins coûteux pour l'entreprise. Les coûts de la recherche d'emploi pour les personnes ayant un faible niveau d'éducation seraient alors supérieurs aux bénéfices qu'elles pourraient retirer d'un éventuel emploi, d'où leur retrait du marché du travail. A l'inverse, les plus éduqués n'auraient que peu de risques d'être au chômage ou de se retirer du marché du travail, car leur niveau de diplôme et de compétences intéressera toujours au moins un employeur prêt à leur offrir un salaire qu'ils jugent suffisamment élevé pour qu'ils restent incités à travailler, même en cas de ralentissement de l'économie. Enfin, il existerait une situation intermédiaire où les individus ayant un niveau moyen de diplôme seraient dans un premier temps soumis à une insertion difficile et précaire sur le marché du travail par manque d'expérience professionnelle, puis à une alternance de périodes d'emploi et de chômage dépendant essentiellement de l'activité cyclique de l'économie. Ainsi, même s'il existe bien d'autres facteurs affectant la prise d'emploi, les individus ayant un faible niveau de diplôme sont supposés se décourager davantage et se retirer du

marché du travail alors que les individus ayant un haut niveau de diplôme sont supposés minimiser ces risques (Barham, 2002 ; Gürbüz et al., 2013).

Plusieurs critiques peuvent être émises à l'égard de cette théorie. La première concerne l'hypothèse de chômage volontaire sur laquelle elle s'appuie. Dans un contexte de crise et de chômage de masse, beaucoup ne sont pas en mesure de choisir entre emploi, chômage et inactivité. Les individus ne font pas face à un arbitrage entre travailler ou ne pas travailler sur la base d'un simple calcul financier. Certains, généralement les moins diplômés, risquent d'être assignés à des emplois peu qualifiés, voire au chômage, surtout dans un contexte de segmentation du marché du travail (Doeringer et Piore, 1971). En effet, en cherchant à stabiliser une partie de leur main d'œuvre, les entreprises constituent des marchés internes, où les modes de gestion et de sélection des travailleurs peuvent être responsables de l'inemployabilité de travailleurs (Delattre et Salognon, 2008).

Toutefois, si pour Doeringer et Piore (1971) la segmentation du marché du travail s'explique au regard de la nature des emplois eux-mêmes, il semble pourtant que le marché externe relève de plus en plus des caractéristiques individuelles (Lemistre, 2012), dont fait partie le niveau de capital humain. En effet, aujourd'hui, le marché du travail français peut être scindé en deux groupes : les individus à carrières, généralement les plus qualifiés, et les individus fragilisés face au marché du travail (récurrence des périodes de chômage, précarité de l'emploi...), de par leur faible niveau de diplôme. En d'autres termes, l'appartenance des individus au marché externe est de moins en moins temporaire, la segmentation entre emplois rejoignant celle entre individus à long terme (Lemistre, 2012). Pour certains, le marché externe serait même constitué de deux segments : un segment précaires et un segment inemployables (Castel, 2009). L'appartenance à l'un ou l'autre de ces deux segments du marché externe peut être partiellement conditionnée par le niveau de qualifications et de compétences des individus compte tenu de l'importance du capital humain dans le potentiel d'employabilité.

De plus, les individus ayant un faible niveau d'éducation ne sont pas nécessairement exclus de l'emploi. Beaucoup ont un emploi malgré leur bas niveau de qualification, même s'il s'agit souvent d'un emploi non-qualifié. Cependant, ces situations ont été encouragées par les politiques publiques qui ont vu au travers de ces emplois un moyen d'insertion des populations les moins qualifiées et les plus à même d'être

exclues de l'emploi. Ces politiques participent donc au maintien d'un niveau important d'emplois non-qualifiés dans l'emploi total et par conséquent, à la segmentation du marché du travail. Indirectement, elles maintiennent aussi dans ces emplois les individus ayant au final peu de possibilités d'accéder à la qualification et donc d'évoluer sur le marché du travail (Pierron, 2012). En effet, le taux d'accès à la formation continue des moins qualifiés est près de trois à quatre fois plus faible que celui des plus qualifiés (Bonaïti et al. 2006). Il en résulte que les salariés de ce type d'emploi ont un risque accru d'obsolescence de leurs compétences, déjà supposées faibles. Leurs perspectives professionnelles restent ainsi figées et les inégalités sur le marché du travail ne peuvent que s'accroître en leur défaveur. L'emploi non qualifié constituerait ainsi même une trappe d'où il serait difficile de sortir, notamment en raison de ce moindre accès à la formation continue.

Pour autant, il existerait une certaine déconnexion entre emploi non qualifié et travail non qualifié. Ce dernier se serait enrichi et les tâches simples, répétitives et prescrites qui le caractérisent semblent désormais dépassées. En effet, la correspondance entre un travail considéré comme non qualifié et le contenu réel des tâches n'est plus aussi évidente. Malgré l'augmentation de la part de l'emploi non-qualifié dans l'emploi total depuis les années 90, il semblerait que les tâches qui incombent aux salariés occupant ce type d'emploi requièrent globalement plus de compétences et de savoirs, notamment en raison de la tertiarisation des emplois non qualifiés (Rose, 2012). Ceci serait donc le signe d'une réelle évolution du contenu des tâches dans certains secteurs d'activité, d'ailleurs marqués par des formes innovantes d'organisation. A partir de la troisième enquête européenne sur les conditions de travail des salariés, Lorenz et Valeyre (2005) ont mis en évidence que les organisations qui se caractérisent par un fort contenu cognitif et qui nécessitent de fait des salariés qualifiés, voire hautement qualifiés, se sont substituées, avec le temps et à travers l'Europe, aux organisations à faible contenu cognitif. Selon les deux auteurs, les organisations dites apprenantes et en *lean* production regrouperaient à elles deux 67% des salariés alors que les organisations tayloriennes et de structure simple ne rassemblent que 33% des salariés.

Ainsi, même si la part de l'emploi non-qualifié reste importante dans l'emploi total, il semble néanmoins qu'il y ait un enrichissement général du contenu du travail en termes de qualifications et de compétences. Dans ce contexte, dans quelle mesure les compétences de base peuvent-elles expliquer pourquoi certains individus ont

aujourd'hui plus ou moins de chances de trouver un emploi que d'autres ? Dans quelle mesure cette nouvelle variable apporte-t-elle un éclairage supplémentaire quant aux situations occupées à l'égard du marché du travail ? Les individus peu compétents en lecture, écriture, compréhension écrite et orale ainsi qu'en calculs simples, sont-ils effectivement plus sanctionnés par ce dernier compte tenu de l'enrichissement général des emplois en termes de qualifications et de compétences ?

3. Compétences de base et marché du travail

Si le niveau de compétences ne peut à lui seul expliquer la situation occupée sur le marché du travail, de nombreux travaux empiriques ont pourtant dégagé un lien significatif entre ces deux variables, ce qui justifie l'intérêt de cette recherche. Malgré la corrélation indéniable entre éducation et compétences de base, le niveau de littératie³⁴ aurait même parfois un impact plus grand sur le potentiel d'employabilité que le niveau de diplôme (Finnie et Meng, 2007). Toutes choses égales par ailleurs, les individus ayant de faibles niveaux de compétences de base (littératie et numératie confondues) ont effectivement plus de risques d'être au chômage ou encore d'occuper des emplois pas ou peu qualifiés. Ils ont également une plus faible probabilité d'utiliser un ordinateur dans leur travail, de recevoir une formation et de progresser au sein de leur entreprise (Bynner et al., 2001). Ces compétences ne sont donc pas seulement les principaux déterminants des progrès scolaires et de la poursuite des études (Mangini, 2012). Elles semblent aussi conditionner l'insertion et la carrière professionnelle, tant pour les hommes que pour les femmes (Charrette et Meng, 1998 ; McIntosh et Vignoles, 2001 ; Chiswick et al., 2003 ; De Coulon et al., 2007)³⁵, d'autant plus que si l'on se place du point de vue des employeurs, il semblerait que le niveau individuel de compétences de base prenne de plus en plus de poids dans les décisions de recrutement (Brown et Hesketh, 2004). D'après une enquête menée en 2007 par l'Institut anglais des Administrateurs (*Institute of Directors, IoD*), sur une liste de 28 compétences clés pour l'emploi, la littératie, la compréhension orale et la numératie étaient

³⁴ La littératie regroupe les compétences en lecture, en écriture et en compréhension écrite. La numératie quant à elle désigne les compétences de base en mathématiques.

³⁵ Pour une revue détaillée sur l'impact des compétences de base, se référer à Ananiadou K. et al., 2004.

respectivement classées comme étant la deuxième, la troisième et la sixième compétence la plus importante.

3.1. Les compétences de base des non-diplômés et des faiblement diplômés

Dans un contexte de changements technologiques et organisationnels rapides, la consolidation des compétences de base semble, d'après la littérature existante, prendre de l'importance de par son rôle dans l'employabilité de publics déjà fragilisés et davantage sanctionnés sur le marché du travail. Parmi ces publics se trouvent notamment les individus peu diplômés, les seniors, les habitants de ZUS ou encore les individus d'origine immigrée. Concernant les individus peu diplômés, d'après une enquête canadienne³⁶ menée en 1989 auprès de décrocheurs du secondaire, un socle plus solide de compétences en littératie et en numératie améliore en effet considérablement et significativement la position individuelle sur le marché du travail, et ce, indépendamment de l'impact de l'éducation scolaire. Tyler (2004) estime même que l'augmentation d'un écart-type dans les scores aux tests de compétences cognitives augmente le salaire moyen des décrocheurs de 6,5% sur les trois premières années passées sur le marché du travail. Cela prouve que malgré un faible niveau de diplôme, les individus décrocheurs ne sont pas systématiquement relégués à des postes à faible contenu cognitif. Ces résultats rejoignent ceux de Patrinos et Sakellariou (2013). Pour ces deux auteurs, l'éducation contribuerait assez peu à expliquer les niveaux de salaires des plus faibles contrairement aux compétences de base, qui constitueraient un élément bien plus déterminant.

Au regard de ces éléments, nous supposons donc que les compétences de base des individus ayant un diplôme proportionnellement très répandu sont un élément explicatif supplémentaire de la situation occupée face au marché du travail parce que leur diplôme ne signale que très mal leur niveau de compétences de base.

³⁶ Enquête sur les Capacités de Lecture et d'Écriture Utilisées Quotidiennement.

3.2. L'obsolescence des compétences de base dans un contexte marqué par de fortes innovations technologiques

L'importance prise par les compétences de base sur le marché du travail peut poser de sérieux problèmes à certains publics fragilisés, tels que les peu ou pas diplômés mais aussi les seniors, soumis à un risque d'obsolescence des compétences principalement acquises durant la scolarité. L'effritement des compétences de base provient d'un manque d'entretien de celles-ci (Coles et Masters, 2000 ; Decreuse et Di Paola, 2002) qui peut aussi bien concerner les individus qui les utilisent peu dans leur emploi que les individus en arrêt maladie, au chômage, en inactivité ou en fin de carrière. Ces individus sont alors soumis à l'obsolescence technique des compétences (De Grip et Van Loo, 2002). Elle correspond à la dégradation des compétences provoquée par les effets de l'âge³⁷ ou par l'éloignement temporaire ou prolongé de l'emploi, voire de certaines conditions de travail. L'obsolescence des compétences peut aussi résulter des changements technologiques et organisationnels qui rendent obsolètes certaines compétences par rapport à d'autres en particulier chez les seniors : il s'agit de l'obsolescence économique des compétences qui, en somme, affecte la valeur du capital humain du travailleur (Rosen, 1975 ; Welch et Ureta, 2002 ; Allen et De Grip, 2007).

Les seniors sont particulièrement exposés à ce phénomène d'obsolescence, qu'il soit technique ou économique, notamment en raison d'une perte des fonctions exécutives et de la mémoire avec l'âge. Les fonctions exécutives renvoient à un ensemble assez hétérogène de processus cognitifs de haut niveau dont le rôle majeur est de favoriser l'adaptation d'un individu face à des situations nouvelles, conflictuelles ou complexes³⁸. De fait, dans un contexte de fortes mutations technologiques et organisationnelles, les plus de 50 ans en âge de travailler sont moins enclins à s'adapter facilement aux innovations en raison de ces fonctions cognitives qui se perdent grandement avec l'âge (Souhay et al., 2000 ; Prull et al., 2000). Ils seraient ainsi bien moins nombreux dans les entreprises innovantes ou utilisant les NTIC que les plus jeunes générations (Aubert et al., 2006). Dans cet environnement, les employeurs seront prêts à recruter des individus de plus de 50

³⁷ Ce qui rejoint le modèle de Mincer (1974) augmenté d'un freinage, par la prise en compte du coefficient négatif sur le terme quadratique de l'expérience.

³⁸ Les fonctions exécutives rassemblent les capacités liées à l'anticipation, l'organisation, l'apprentissage, la motivation, l'initiative ou encore la résolution de problèmes.

ans uniquement si ceux-ci présentent des compétences propres à soutenir les changements structuraux de l'économie. Parmi ces compétences se trouve la numératie : elle favoriserait en effet l'accompagnement de ces changements (Atkinson et Splisbury, 1993 ; Berman et al., 1998) et donc l'accès à l'emploi (Dearden et al., 2002 ; Chiswick et al., 2003). Bynner (2004) par exemple, dénote que dans les entreprises utilisant fortement les NTIC, le calcul est un attribut essentiel pour obtenir et conserver un emploi dans ces entités. De même, Koedel et Tyhurst (2012), en élaborant une expérimentation à partir de curriculum-vitae affectés aléatoirement comme réponses à des offres d'emploi, ont clairement mis en évidence que les compétences en mathématiques recevaient plus d'attention de la part des employeurs, toutes choses égales par ailleurs. Une faible maîtrise des compétences en numératie limiterait l'accès à l'emploi, indépendamment du niveau de diplôme, surtout dans des secteurs d'activités recrutant des diplômés de tout horizon disciplinaire (Durrani et Tariq, 2012).

Nous supposons donc que les employeurs utilisent principalement le niveau de numératie dans la sélection à l'embauche des individus, en particulier des plus âgés, le niveau de numératie servant de signal de compétences de base et de capacités d'adaptation dans un contexte d'innovations technologiques.

3.3. Marché du travail et pratiques discriminatoires

Se pose aussi la question de la valorisation des caractéristiques individuelles des immigrés sur le marché du travail. Par exemple, la population vivant en Zone Urbaine Sensible (ZUS) est globalement moins bien dotée en capital humain et social que la population vivant hors-ZUS, ce qui expliquerait en partie les difficultés des individus habitant des quartiers défavorisés sur le marché du travail. Il en est de même à l'égard des populations immigrées, qui occupent d'ailleurs une part importante au sein des ZUS. Pour autant, d'autres obstacles, indépendants des individus eux-mêmes, affectent la valorisation des potentiels. D'après la théorie de la discrimination statistique développée par Phelps (1972), l'information est imparfaite sur la productivité des candidats à l'embauche, ce qui amène les employeurs à sélectionner les individus sur la base de caractéristiques facilement observables et corrélées à la performance (sexe, origine, niveau de diplôme, âge...).

L'employeur pourra ainsi sélectionner les individus en fonction de ses observations et de ses jugements statistiques quant à la productivité moyenne du groupe associé à chaque caractéristique. Une discrimination territoriale a lieu dès lors que les employeurs attribuent la productivité moyenne des travailleurs habitant dans une ZUS aux individus dont la productivité est inobservée. La discrimination liée à l'origine migratoire suit le même processus pour les personnes issues de la première, voire de la deuxième génération d'immigration.

Dans ce cadre-ci, nous supposons alors qu'un surplus de compétences pour les individus vivant en ZUS risque d'être bien moins valorisé que pour ceux vivant hors-ZUS, sous prétexte que la productivité moyenne des premiers est jugée moins bonne que celle des seconds. Idem pour les personnes issues de l'immigration.

4. Méthodologie : modèle logistique multinomial ordonné généralisé

Indépendamment de la précarité des emplois ou des formes de chômage, les compétences de base semblent faire partie intégrante des caractéristiques individuelles expliquant la position objective des individus sur le marché du travail en termes d'emploi, de chômage ou d'inactivité. Plus précisément, nous pensons que le marché du travail sanctionne les individus selon leur niveau de compétences de base, compte tenu du fait qu'elles sont hétérogènes pour un même niveau de diplôme, que les recruteurs semblent y être sensibles et que le contenu du travail en général en requiert davantage, quel que soit le type d'emploi. De fait, dans un contexte de chômage de masse et cumulé à d'autres facteurs, un déficit de compétences de base peut contribuer à l'exclusion de l'emploi, voire au retrait du marché du travail dans sa forme la plus radicale, et ce, en dehors des raisons liées à l'état de santé ou des raisons familiales. A l'inverse, un surplus de compétences est supposé contribuer à la protection des individus des épisodes de chômage et de l'inactivité, toutes choses égales par ailleurs. La variable dépendante, que constitue la situation occupée face au marché du travail, est ainsi envisagée comme étant de nature ordinale, comme dans la théorie de la recherche d'emploi, puisqu'à chacune d'elle correspond un niveau de compétences de base :

	<i>En emploi</i>	=	1
<i>Situation marché du travail (Sit_{prof-i}) :</i>	<i>Chômage</i>	=	2
	<i>Inactif</i>	=	3

Bien que le résultat soit discret, le modèle logistique multinomial seul ne permet pas de prendre en compte cette idée d'ordre. Il est indispensable de se placer dans le cadre d'un modèle logistique polytomique ordonné. Néanmoins, ce type de modèle implique de respecter une hypothèse fondatrice : celle de l'égalité des pentes ou encore de régression parallèle. Cette hypothèse suppose que l'effet des variables exogènes soit quasi-identique pour les trois modalités possibles de la variable endogène. Dans le cas présent, suite à l'utilisation de tests de Brant, cette hypothèse s'est révélée être violée, principalement en raison de certaines variables de contrôle. Comme le rappellent Freese et Long (2006), cette hypothèse n'est d'ailleurs que très rarement vérifiée. Une solution pour contourner cet obstacle est alors de faire appel à un modèle logistique polytomique ordonné généralisé³⁹, mieux adapté aux données. Grâce à ce type de modèle, il est possible de contraindre la relation entre les variables exogènes et endogènes satisfaisant à l'hypothèse des rapports de cotes proportionnels de la régression logistique, tout en permettant aux relations ne satisfaisant pas cette hypothèse de varier. Ce modèle est moins restrictif qu'un modèle logistique ordonné mais plus parcimonieux qu'un modèle logistique multinomial (Freese et Long, 2006). Ce modèle peut alors être formalisé de la manière suivante :

$$P(\text{Sit}_{prof_i} > j) = g(X\beta_j) = \frac{\exp(\alpha_j + X_i\beta_j)}{1 + [\exp(\alpha_j + X_i\beta_j)]}$$

où i est l'individu et $j = 1, 2, \dots, J-1$ avec J , le nombre de catégories. L'équation permet de déterminer les probabilités suivantes d'appartenance à telle ou telle catégorie :

$$P(\text{Sit}_{prof_i} = 1) = 1 - g(X_i\beta_1)$$

$$P(\text{Sit}_{prof_i} = j) = g(X_i\beta_{j-1}) - g(X_i\beta_j) \quad \text{avec } j = 2, \dots, J - 1$$

$$P(\text{Sit}_{prof_i} = J) = g(X_i\beta_{J-1})$$

³⁹ Ce modèle est également appelé « régression logistique ordonnée proportionnelle partielle »

Afin de s'assurer que $P(\text{Sit}_{\text{prof}_i} = j)$ est bien compris entre 0 et 1, la condition suivante doit être respectée :

$$(\alpha_j - X_i\beta_j) \geq (\alpha_{j-1} - X_i\beta_{j-1})$$

Comme il y a trois catégories ordonnées (emploi, chômage, inactivité), c'est-à-dire $J=3$, alors l'estimation d'un modèle logistique polytomique ordonné généralisé⁴⁰ revient à estimer deux régressions logistiques dichotomiques où les modalités de la variable endogène sont combinées. Ainsi, quand $j=1$, la catégorie 1 (individus en emploi) est opposée aux catégories 2 (individus au chômage) et 3 (individus retirés du marché du travail) simultanément. A l'inverse, lorsque $j=2$, la catégorie 1 et la catégorie 2 sont opposées ensemble à la catégorie 3. Cependant, l'avantage de ces modèles logistiques polytomiques ordonnés généralisés par rapport à deux analyses dichotomiques séparées est qu'ils permettent de bien appréhender les variables qui impactent significativement la situation vis-à-vis du marché du travail (respect de l'hypothèse d'égalité des pentes) et celles qui auraient un effet différencié. Les premières supposent par conséquent qu'elles jouent bel et bien un rôle dans l'ordre établi des degrés de compétences alors que ce n'est pas le cas des secondes. Passer par des analyses dichotomiques séparées ne permettrait donc pas de dégager un effet global sur les différentes situations occupées face au marché du travail.

5. Statistiques descriptives

Les données exploitées sont également issues de l'enquête Information et Vie Quotidienne réalisée en France en 2004 (IVQ) par l'INSEE. Comme il est supposé que les individus les plus compétents ont une probabilité plus importante d'être en emploi, les scores moyens en compétences de base devraient donc être relativement plus élevés pour eux. Les scores moyens en compétences de base par situation occupée face au marché du travail⁴¹ et selon certaines caractéristiques (niveau de

⁴⁰ L'estimation est réalisable sous STATA avec la commande « gologit2 », développée par Williams R. (2006).

⁴¹ Au regard de l'ensemble des informations disponibles sur la situation des individus face au marché du travail et de leurs comportements sur le marché du travail, il a été nécessaire de revoir préalablement la constitution des groupes de ces situations (cf. Annexe A).

diplôme, tranche d'âge, lieu de résidence, origine migratoire) sont présentés dans le tableau 1.

Tableau 1 : Scores moyens en compétences de base selon la situation occupée face au marché du travail et autres caractéristiques⁴²

	Littératie			Numératie			Compréhension orale		
	E	C	I	E	C	I	E	C	I
Niveau de diplôme									
Sans diplôme ou CEP	81,3	81,0	80,4	62,4	57,7	59,7	73,3	71,1	69,9
Diplômés du secondaire	89,4	89,0	87,9	74,4	73,3	71,5	84,3	82,0	81,5
Diplômés du supérieur	93,3	92,9	91,4	81,9	78,7	78,5	89,3	84,1	84,6
Tranche d'âge d'appartenance									
18-30 ans	91,2	89,5	88,7	76,8	73,0	69,0	87,1	82,1	83,3
31-50 ans	89,1	86,8	84,9	74,2	68,6	66,2	83,4	79,0	77,8
51-65 ans	86,3	84,9	84,2	72,0	66,0	67,4	79,4	74,1	74,5
Lieu de résidence									
ZUS	84,4	81,9	79,3	67,4	64,1	56,8	76,1	71,4	65,4
Hors-ZUS	89,2	88,1	85,6	74,6	70,4	68,1	84,1	80,3	78,0
Origine migratoire									
1 ^{ère} génération d'immigrés	81,6	80,9	80,8	67,9	62,5	64,6	73,3	71,1	68,3
2 ^{ème} génération d'immigrés	89,8	87,6	86,3	72,6	69,3	69,2	82,9	74,6	79,5
Natifs	89,7	89,0	85,5	75,3	71,7	66,8	85,0	82,7	78,4

E = Emploi ; C = Chômage ; I = Inactivité.

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Les scores moyens en compétences de base sont globalement plus élevés pour les individus en emploi que pour les autres. Cependant, les écarts de scores moyens sont parfois très faibles entre les individus au chômage et les inactifs. Les personnes en inactivité ne seraient pas forcément moins compétentes en moyenne que celles se trouvant au chômage. D'autres déterminants pourraient alors expliquer le retrait du marché du travail (facteurs culturels, religieux, découragement suite à des discriminations, raisons familiales et/ou de santé...).

Le niveau moyen de compétences de base des individus sans diplôme est bien plus faible que celui des diplômés du secondaire et du supérieur, qu'ils soient en emploi ou non. Ce résultat s'explique au regard de la corrélation positive entre niveau de

⁴² Les moyennes de scores sont significativement différentes selon le niveau de diplôme et la tranche d'âge d'appartenance par situation occupée face au marché du travail.

diplôme et compétences de base. Toutefois, les personnes en emploi ont un score moyen légèrement plus élevé que celles étant au chômage ou en inactivité. Cela rejoint en partie les résultats de l'enquête ELCA (Enquête sur la Littératie et les Compétences des Adultes), réalisée en 2003 au Canada, bien que les écarts ne soient pas particulièrement importants. Les individus ayant un emploi ont obtenu un score moyen de littératie plus élevé que les individus sans emploi. Ce constat peut faire l'objet d'une double interprétation. Tout d'abord, les plus diplômés ont plus de chances de trouver un emploi parce que le niveau d'éducation signale le niveau de compétences de base, même si la relation reste imparfaite. Secondement, il est possible que les personnes en poste aient des scores plus élevés que ceux qui n'ont pas d'emploi, tout simplement parce qu'ils sont amenés à utiliser et donc entretenir leurs compétences en littératie dans leur travail, ce qui assure un certain maintien de leur niveau de compétences de base. Les scores moyens par tranche d'âge vont d'ailleurs également dans ce sens. Il est également très probable que ces deux effets se conjuguent pour donner lieu à des scores en compétences de base assez élevés. Néanmoins, d'après ces statistiques descriptives, il est assez délicat de dire quel élément, entre le niveau de diplôme et le niveau de compétences de base, est le plus explicatif des situations occupées face au marché du travail, le lien entre les deux étant complexe et difficile à appréhender. L'appel aux outils économétriques s'en trouve ainsi justifié.

De plus, quelles que soient la compétence considérée et la situation occupée face au marché du travail, le score moyen diminue avec l'âge, signe possible d'une obsolescence technique des compétences de base, principalement acquises et développées durant la scolarité. Alors que la numératie est devenue une compétence clé dans l'employabilité d'après la revue de littérature, il semblerait qu'elle soit la compétence la moins bien maîtrisée par les jeunes et les moins jeunes. Notons que le niveau moyen de numératie est d'ailleurs bien plus faible pour les plus de 50 ans. Quel impact la numératie pourrait-elle donc avoir sur la situation occupée par les seniors sur le marché du travail ?

Les individus vivant en ZUS ont, pour leur part, un moins bon niveau en compétences de base que les individus vivant hors-ZUS, quelle que soit la situation considérée face au marché du travail. Ce plus faible niveau de compétences peut être le résultat d'une caractéristique importante de la population habitant en ZUS, celle d'avoir en moyenne un niveau de diplôme plus bas que la population vivant hors-

ZUS. D'après l'INSEE, entre 2004 et 2007, 28,8% des individus vivant en ZUS ne possédaient aucun diplôme, contre 14,1% pour ceux vivant hors-ZUS. Toutefois, le niveau d'éducation ne suffit pas pour expliquer un tel écart puisqu'à niveau de diplôme équivalent, les personnes habitant en ZUS ont malgré tout des scores moyens plus faibles en compétences de base (cf. tableau 2).

Tableau 2 : Scores moyens en compétences de base selon le lieu de résidence (ZUS ou hors-ZUS) et le niveau de diplôme

		Littératie	Numératie	Compréhension orale
ZUS	Sans diplôme ou CEP	76,5	54,9	63,4
	Secondaire	86,6	70,4	79,4
	Supérieur	91,8	78,6	85,6
Hors-ZUS	Sans diplôme ou CEP	81,6	62,1	73,5
	Secondaire	89,4	74,2	84,2
	Supérieur	93,2	81,6	88,9

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Ce résultat peut s'expliquer par un plus faible capital social, par un éloignement des zones plus dynamiques en matière d'emploi, par une plus faible mobilité ou encore par la forte proportion d'individus d'origine immigrée au sein des ZUS. D'après l'INSEE, en 2007, 24% des habitants vivant en ZUS étaient issus de la première génération d'immigration contre 4% des habitants vivant hors-ZUS. Or, à niveau de diplôme équivalent, la maîtrise des compétences de base en français est certainement plus fragile pour des individus ayant appris à lire, à écrire et à compter dans leur pays d'origine et donc dans une langue différente du français (cf. tableau 3).

Tableau 3: Scores moyens en compétences de base selon l'origine migratoire et le niveau de diplôme

		Littératie	Numératie	Compréhension orale
1^{ère} génération d'immigrés	Sans diplôme ou CEP	73,6	57,6	62,8
	Secondaire	85,2	70,6	75,3
	Supérieur	89,8	77,4	85,1
2^{ème} génération d'immigrés	Sans diplôme ou CEP	82,7	60,4	70,7
	Secondaire	89,5	71,8	83,4
	Supérieur	93,4	80,3	86,4
Natifs	Sans diplôme ou CEP	82,8	62,4	75,4
	Secondaire	89,6	74,7	84,8
	Supérieur	93,5	82,2	89,7

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

En outre, selon le pays d'origine des individus, les niveaux de diplôme sont difficilement comparables entre eux compte tenu des caractéristiques propres à chaque système éducatif. Par exemple, le niveau baccalauréat d'un individu d'origine italienne sera certainement plus fiable pour signaler un niveau de compétences de base que celui d'un individu d'origine angolaise. Le niveau de diplôme ne saurait donc être un bon indicateur du niveau de la littératie et de la compréhension orale du français pour les individus ayant suivi tout ou partie de leur scolarité dans un autre pays et dans une autre langue que le français.

6. Résultats

6.1. Situation occupée à l'égard du marché du travail et scores de compétences⁴³

Différentes variables explicatives ont été introduites dans le modèle logistique polytomique ordonné généralisé, dont les scores en compétences de base afin d'expliquer la probabilité d'être dans l'une ou l'autre des situations occupée sur le marché du travail. D'autres variables ont été ajoutées afin de contrôler le modèle par certaines caractéristiques individuelles (âge, origine migratoire, sexe) ainsi que par les contextes éducatif et familial. Les modèles ont été réalisés séparément pour les sans diplôme, les diplômés du secondaire et les diplômés du supérieur afin de permettre aux effets des différentes variables de varier selon le niveau de diplôme. Il en a été de même pour les trois tranches d'âge (18-30 ans ; 31-50 ans ; 51-65 ans)⁴⁴. Les résultats des différentes estimations se trouvent en annexe B, tableaux B.1. et B.2.

6.1.1. Par niveau de diplôme

Le diplôme étant un signal imparfait du niveau de compétences de base des individus, il est possible que pour les peu ou pas diplômés, le diplôme n'intègre que peu de capital humain. L'estimation du modèle par niveau de diplôme doit permettre de mieux appréhender le rôle joué par les compétences de base dans la participation au marché du travail des individus mal signalés et dans leur accès à l'emploi. Si le diplôme signale et filtre correctement les individus en fonction de leur

⁴³ Dans les modèles logistiques ordonnés généralisés, dès lors que les coefficients associés aux scores en compétences sont identiques dans les deux combinaisons de catégories, il est alors possible de conclure à un effet indifférencié de ces variables d'intérêt sur les différentes situations occupées face au marché du travail (variable endogène). L'interprétation se fera alors de la même manière qu'un modèle ordonné classique. La variable endogène cherchant à confronter deux modalités de situations occupées sur le marché du travail, un coefficient négatif supposera que c'est pour la première situation occupée face au marché du travail que la probabilité sera influencée à la hausse et inversement pour un coefficient positif.

⁴⁴ Les tranches d'âge correspondent respectivement à une période d'insertion sur le marché du travail (18-30 ans), à une période de maturité des compétences (31-50 ans) et à une période d'obsolescence des compétences (51-65 ans).

niveau de compétences de base, ce qui est mis en doute dans cet essai, alors ces dernières n'auront pas d'impact significatif sur les situations occupées face au marché du travail.

Dans le premier modèle (cf. annexes, tableau B.1), concernant les individus sans diplôme, seul le score en compréhension orale a un effet positivement significatif. 10 points supplémentaires en compréhension orale diminuent de 4,82% la probabilité d'être au chômage ou inactif. Ce résultat assez modeste laisse à penser qu'un faible niveau d'éducation signale relativement bien le niveau (faible) de compétences de base des individus.

Le modèle 2 fournit des résultats bien plus intéressants pour les diplômés du secondaire, puisque les trois compétences de base ont un impact positif et significatif au seuil de 5% et moins sur la situation occupée face au marché du travail. La littératie semble d'ailleurs avoir un impact plus fort que les autres compétences puisque 10 points supplémentaires diminuent la probabilité d'être au chômage ou inactif de 11,84%. Ceci est tout à fait logique avec l'idée de *barrière d'entrée* sur le marché du travail pour ceux qui ont des difficultés avec le français écrit. Les démarches de recherches d'emploi et de recrutement (ciblage des offres d'emploi, construction d'un CV et d'une lettre de motivation, entretiens...) nécessitent en effet un minimum de maîtrise en lecture, écriture et en compréhension écrite, or, de fortes lacunes ne permettent pas *a priori* de rechercher efficacement un emploi. Les moins compétents en termes de *basic skills* sont alors plus susceptibles, toutes choses égales par ailleurs, de se retirer du marché du travail. Ceci rejoint les chiffres de l'INSEE concernant l'illettrisme : 40% des illettrés sont en emploi tandis que 15% et 45% sont au chômage et en inactivité respectivement. *A contrario*, avoir un très bon niveau en compétences de base, pour des niveaux de diplômes relativement répandus comme ceux du secondaire, semble être un facteur jouant positivement sur la participation au marché du travail et l'accès à l'emploi. Cela rejoint l'idée que les diplômés proportionnellement plus répandus signalent et filtrent incorrectement les individus en fonction de leurs compétences de base. Par conséquent, les décisions d'embauche de la part des recruteurs peuvent davantage se prendre sur la base du niveau de ces compétences, comme cela semble d'ailleurs être de plus en plus le cas au regard des travaux empiriques existants. Une forte demande pour de solides compétences de base

amènerait ainsi les individus diplômés du secondaire à être moins découragés sur le marché du travail s'ils peuvent effectivement répondre aux attentes des employeurs.

A l'inverse, concernant les diplômés du supérieur, le modèle 3 révèle que les compétences de base n'ont aucunement d'impact significatif sur la situation occupée face au marché du travail. L'écart-type de chaque score en compétences de base est d'ailleurs beaucoup plus faible pour les diplômés du supérieur que pour les autres, quelle que soit la compétence considérée (cf. tableau 4). Un diplôme du supérieur semble donc signaler assez convenablement le niveau de compétences de base des individus, contrairement à un diplôme du secondaire. Dans ce contexte, les efforts des politiques pour l'emploi doivent se concentrer sur le renforcement des compétences de base des individus ayant un très faible niveau d'éducation. Renforcer les compétences de base doit en réalité permettre à des individus pas ou très peu qualifiés d'accéder ultérieurement à des formations plus spécifiques, nécessaires au retour à l'emploi (Acemoglu et Pischke, 1999 ; De Grip et Zwick, 2005). D'après les résultats obtenus, les efforts semblent également devoir se tourner vers les diplômés du secondaire qui auraient des niveaux particulièrement faibles en compétences de base, et ce, afin d'améliorer leur potentiel d'employabilité et réduire leur risque de découragement en raison de la sanction du marché du travail à l'égard des moins compétents.

6.1.2. Par tranche d'âge

Les compétences de base étant sujettes au phénomène d'obsolescence, il est possible qu'à partir d'un certain âge, elles deviennent un critère plus important de sélection par les employeurs, voire même d'auto-sélection. Estimer le modèle pour chaque tranche d'âge doit permettre de dégager une telle tendance.

Ne pas avoir de diplôme pour un individu de moins de 30 ans augmente considérablement et significativement la probabilité d'être sans emploi et de ne pas participer au marché du travail. Cela est également vrai pour les individus âgés entre 31 et 50 ans, mais dans une moindre mesure. Globalement, pour les plus jeunes, plus le diplôme est élevé, plus la probabilité d'être en emploi est grande, ce qui est tout à fait cohérent avec la littérature existante. Pour les seniors, ce résultat n'est vérifié

que pour un niveau de diplôme sanctionnant plus de deux années d'enseignement supérieur. Cela suppose que pour les plus de 30 ans, l'expérience professionnelle peut réduire l'effet de l'éducation comme déterminant des situations occupées face au marché du travail au profit des compétences de base. En effet, les compétences de base n'impactent pas significativement la situation occupée face au marché du travail des moins de 30 ans alors que c'est le cas pour les tranches d'âge supérieures. Si l'on peut supposer que le rôle de filtre attribué au diplôme fonctionne plus ou moins correctement en début de carrière, faute d'information(s) sur les caractéristiques des jeunes, d'autres signaux, comme le niveau de compétences de base, vont ensuite progressivement se substituer au diplôme pour informer les employeurs des capacités des salariés. Ainsi, pour les plus de 30 ans, les compétences de base jouent significativement un rôle dans les positions occupées sur le marché du travail. Plus précisément, pour les individus âgés entre 31 et 50 ans, la littératie a un très fort impact. 10 points en littératie, toutes choses égales par ailleurs, diminuent la probabilité d'être inactif de 14,04% alors qu'ils augmentent parallèlement celle d'être en emploi de 7,35%. La numératie et la compréhension orale ont aussi un effet significatif sur la probabilité d'être au chômage ou inactif mais cet effet est moins important. Ces deux compétences sont d'ailleurs les deux seules qui ont un effet significatif pour les personnes de plus de 50 ans, encore en âge de travailler. L'impact de la numératie est particulièrement fort. Par exemple, 10 points diminuent la probabilité d'être au chômage ou inactif de 10,15% pour les seniors. Ce pourcentage n'est que de 5,10% pour les 31-50 ans, ce qui est presque moitié moins. Ce résultat rejoint les éléments mis en avant par la littérature : la numératie, dans un fort contexte d'innovation, de production et d'utilisation des NTIC est devenue un élément central dans l'employabilité et la participation au marché du travail des individus. De ces résultats, il ressort que de hauts niveaux de compétences de base chez les seniors sont toujours valorisés en termes d'employabilité malgré leur âge, pouvant être également source de discrimination. Retarder l'obsolescence des compétences de base au-delà de 50 ans, voire même avant, par le biais de formations continues est donc une piste possible pour améliorer le potentiel d'employabilité des seniors.

Enfin, les seniors qui ont une suivi une formation industrielle au cours de leur scolarité sont significativement défavorisés dans la participation au marché du travail et dans l'accès à l'emploi. Ceci peut provenir du fait que les formations

industrielles de l'époque, c'est-à-dire avant les années 1970, étaient principalement destinées à la manufacture, contrairement à aujourd'hui.

6.2. Situation occupée à l'égard du marché du travail et indicatrices de déficit / de surplus de compétences

D'après les modèles précédents, les scores en compétences peuvent impacter la situation occupée par les individus face au marché du travail, indépendamment du diplôme. Toutefois, dans quelle mesure cet impact est-il important pour les individus ayant un déficit ou un surplus de compétences par rapport à ceux qui auraient un niveau normal⁴⁵ ? Plus précisément, quel est le degré de sanction d'un déficit ou le degré de récompense d'un surplus, sur des populations déjà soumises à un risque discriminatoire (population vivant en ZUS, populations issues de l'immigration) ? Afin de répondre à cette question, de nouveaux modèles logistiques polytomiques ordonnés généralisés ont été réalisés avec non plus l'inclusion de scores en compétences, mais avec l'inclusion d'indicatrices de déficit ou de surplus de compétences de base, construites à cet effet. La construction de ces indicatrices est relativement simple. En premier lieu, il a fallu calculer la moyenne et l'écart-type de chaque score en compétences de base par niveau de diplôme. (cf. tableau 4)

⁴⁵ Dans une perspective adéquationniste, jugée comme étant « *conceptuellement supérieure* » par Hartog (2000) dans l'étude du déclassement professionnel, un niveau « *normal* » doit se comprendre au regard d'une nomenclature de niveaux de diplômes et de niveaux de compétences de base. Cette méthode définit un niveau « *normal* » de compétences de base pour chaque niveau de diplôme et un intervalle de confiance autour de ce seuil, conduisant mécaniquement à une symétrie dans les résultats de l'impact du déficit de compétences d'une part et du surplus d'autre part.

Tableau 4 : Moyennes et écarts-types des scores en compétences de base par niveau de diplôme.

Diplômes	Littératie		Numératie		Compréhension orale⁴⁶	
	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type
Sans diplôme	79,71	14,77	58,94	22,30	70,41	23,42
Certificat d'Études Primaires	83,49	9,58	65,44	18,59	75,96	20,77
Brevet/BEPC	89,13	6,41	74,08	12,87	84,33	17,03
BEP/CAP	87,95	7,49	72,18	14,22	82,27	19,25
Baccalauréat général	92,04	5,29	77,45	12,12	85,32	19,08
Baccalauréat professionnel	91,01	7,25	75,76	10,01	86,99	17,42
Baccalauréat technologique	92,11	6,32	79,05	9,95	88,67	16,16
Bac +2	92,36	5,20	79,80	10,73	88,85	15,04
Supérieur à Bac+2	93,96	4,68	83,20	9,99	88,85	16,08

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Les individus ayant un score inférieur de plus d'un écart-type au score moyen correspondant à leur niveau de diplôme sont considérés en déficit de compétences. A l'inverse, les individus qui ont obtenu un score supérieur de plus d'un écart-type (σ) au score moyen (m) correspondant à leur niveau de diplôme sont considérés comme étant en surplus de compétences (cf. tableau 5). Pour ces individus, en déficit ou en surplus, le diplôme en tant que signal ne fonctionnerait *a priori* que mal. Puisque pour chaque diplôme, il existe une norme de niveau de compétences de base, il est difficile de distinguer les personnes qui seraient en déficit ou en surplus de compétences de base au regard du seul élément *diplôme*. Ce dernier tendrait finalement à *flouter* les niveaux réels de compétences de base de ces individus et ne constituerait pas un bon filtre de ces niveaux. De fait, il est peut-être plus intéressant d'avoir un niveau moyen de diplôme avec un niveau normal, voire un surplus de compétences de base, qu'un niveau de diplôme plus élevé avec un déficit de compétences de base. Pour des populations soumises à un risque de

⁴⁶ La structure des données sur les scores en compréhension orale est telle qu'il n'a pas été possible de construire d'indicateur de surplus sur le même schéma que les deux autres compétences ($m+\sigma$). En effet, l'écart-type du score en compréhension orale par niveau de diplôme est élevé, ce qui aurait conduit à considérer des individus ayant un score de plus de 100/100, ce qui est impossible, le score ne pouvant prendre que des valeurs dans l'intervalle [0 ; 100]. C'est pourquoi ont été considérés comme étant en surplus uniquement les individus qui avaient un score de plus de 95/100.

discrimination sur le marché du travail (population vivant en ZUS, population d'origine immigrée...), ceci est d'autant plus délicat que les déficits de compétences de base seront *a priori* plus pénalisés et les surplus moins valorisés, d'après notre hypothèse.

Tableau 5: Part des individus présentant des déficits ou des surplus de compétences dans l'échantillon total

	Déficit de plus d'un écart-type par rapport à la moyenne de compétences par classe de diplômes	Surplus de plus d'un écart-type par rapport à la moyenne de compétences par classe de diplômes (sauf pour la compréhension orale : surplus si score $\geq 95/100$)
Littératie	12,69%	8,85%
Numératie	16,59%	9,79%
Compréhension orale	20,14%	38,31%

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Une très forte proportion d'individus est en situation de surplus de compétences en compréhension orale. Ce pourcentage est très important et montre en réalité qu'il y a une saturation dans la mesure de cette compétence⁴⁷.

Tableau 6: Part des individus présentant un déficit et un surplus de compétences selon la situation occupée à l'égard du marché du travail et le lieu résidence (ZUS et hors-ZUS).

		ZUS		Hors-ZUS	
		DÉFICIT	SURPLUS	DÉFICIT	SURPLUS
En emploi	Littératie	60,96%	73,28%	78,61%	83,86%
	Numératie	62,31%	83,82%	77,48%	88,15%
	Compréhension orale	65,35%	78,24%	81,28%	86,43%
Au chômage	Littératie	24,59%	10,49%	9,24%	8,19%
	Numératie	18,13%	11,50%	9,50%	7,19%
	Compréhension orale	17,61%	16,12%	8,27%	6,63%
Inactif	Littératie	14,46%	16,23%	12,15%	7,95%
	Numératie	19,57%	4,67%	13,03%	4,66%
	Compréhension orale	17,04%	5,64%	10,4%	6,93%

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Lecture : 73,28% des individus vivant en ZUS et ayant un surplus en littératie sont en emploi

⁴⁷ Cf. Annexe C. Histogrammes et densités des scores en compétences de base

D'après le tableau 6, près de 61% des individus ayant un déficit en littératie et résidant dans une ZUS sont en emploi contre 78,61% pour ceux qui auraient également un déficit dans cette compétence mais qui résideraient en dehors d'une ZUS. La même tendance est à observer pour la numératie et la compréhension orale. Globalement, la part des individus au chômage, voir en inactivité est particulièrement plus élevée pour les individus vivant en ZUS que pour les autres, qu'ils aient un déficit ou un surplus de compétences de base. Vivre dans un quartier défavorisé peut potentiellement être un mauvais signal pour les employeurs, susceptibles d'adopter un comportement discriminatoire. Toutefois, à cette discrimination territoriale possible peuvent s'ajouter des facteurs religieux ou culturels mais aussi et surtout le manque de capital social, caractérisant la population vivant en ZUS, afin d'expliquer les difficultés d'insertion dans l'emploi, le découragement vis à vis du marché du travail et donc le retrait de celui-ci.

Qu'ils vivent en ZUS ou non, les individus qui auraient un surplus de compétences en numératie sont plus souvent en emploi que ceux qui auraient un surplus dans une autre compétence. L'écart entre les deux groupes est d'ailleurs beaucoup plus faible pour cette compétence, ce qui suppose que la numératie est mieux valorisée sur le marché du travail. Mais vu la forte proportion d'individus au chômage parmi les plus compétents vivant en ZUS, n'existerait-il pas une perte importante liée à une éventuelle discrimination territoriale ?

D'après les résultats des modèles logistiques polytomiques ordonnés généralisés à partir des indicatrices (cf. tableau D.1. en annexe D), être en déficit ou en surplus de compétences de base pour les individus vivant en ZUS n'a pas d'impact significatif sur les situations occupées à l'égard du marché du travail. A l'inverse, les déficits de compétences sont sanctionnés pour les individus qui résident dans un quartier hors-ZUS. Un déficit en littératie ou en numératie augmente respectivement leur probabilité d'être au chômage ou en inactivité de 33% et de 29%. Même si le diplôme constitue un signal de productivité pour les employeurs, celui-ci demeure imprécis et peut, au contraire, camoufler des déficits de compétences de base, qui ont pourtant leur importance dans l'explication de la participation au marché du travail et de l'accès à l'emploi.

Parallèlement, être sans diplôme ou être uniquement titulaire du Certificat d'Études Primaires a un impact négatif sur la situation occupée face au marché du travail, tant

pour ceux vivant en ZUS que pour les autres. Toutefois, même si les résultats sont convergents, les individus sans diplôme vivant en ZUS ont, toutes choses égales par ailleurs, deux fois plus de risques d'être au chômage ou en inactivité qu'un individu sans diplôme résidant en dehors d'une ZUS. Ceci peut s'expliquer par le fait que les individus vivant en ZUS sont moins mobiles, soit parce que leur capital social est plus faible, soit parce qu'ils sont bel et bien victimes de préjugés de la part des employeurs sur le marché du travail, ce qui leur bloque l'accès aux emplois. Ces facteurs sont d'autant plus plausibles que rares sont les variables de capital humain (diplômes et compétences de base) qui, d'après les estimations, ont un impact significatif sur la situation occupée à l'égard du marché du travail des individus résidant en ZUS, contrairement à ceux résidant en dehors. Une analyse sur les populations d'origine immigrée a pu mettre en avant des résultats assez similaires.

Tableau 7 : Part des individus présentant un déficit et un surplus de compétences selon la situation occupée à l'égard du marché du travail et de l'origine migratoire.

		Individus issus de l'immigration : 1 ^{ère} génération		Individus issus de l'immigration : 2 ^{ème} génération		Natifs	
		DÉFICIT	SURPLUS	DÉFICIT	SURPLUS	DÉFICIT	SURPLUS
En emploi	<i>Littératie</i>	71,82%	64,98%	65,17%	77,22%	80,24%	85,60%
	<i>Numératie</i>	62,51%	60,54%	77,19%	81,93%	79,09%	91,37%
	<i>Compréhension orale</i>	68,37%	77,17%	78,29%	82,61%	83,23%	87,29%
Au chômage	<i>Littératie</i>	17,84%	3,42%	20,37%	14,18%	6,93%	7,54%
	<i>Numératie</i>	22,95%	19,90%	14,34%	11,47%	6,19%	5,63%
	<i>Compréhension orale</i>	13,72%	9,44%	15,72%	9,05%	6,50%	6,52%
Inactif	<i>Littératie</i>	10,34%	31,58%	14,46%	8,60%	12,83%	6,86%
	<i>Numératie</i>	14,53%	19,55%	8,46%	6,60%	14,72%	3,00%
	<i>Compréhension orale</i>	17,91%	13,38%	6,00%	8,34%	10,26%	6,19%

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Lecture : 14,53% des individus issus de la première génération d'immigration et ayant un déficit en numératie sont inactifs

Proportionnellement, les individus issus de la première génération d'immigration et ayant un déficit en littératie, voire même en numératie, sont plus nombreux à être en emploi que ceux ayant un surplus. Le réel niveau de compétences de ces derniers semble être mal signalé sur le marché du travail, bloquant ainsi l'accès à l'emploi des plus compétents. Un certain découragement peut alors naître chez ces individus, ce

qui expliquerait en partie leur retrait du marché du travail et ainsi la part élevée de ces individus en situation d'inactivité.

Cette situation ne se retrouve pas vraiment pour les individus issus de la deuxième génération d'immigration et encore moins pour les natifs. Toutefois, le surplus de compétences des natifs semble être mieux valorisé que celui des individus issus de la deuxième génération d'immigration, proportionnellement plus nombreux à être au chômage. A l'inverse un déficit de compétences de base, surtout en littératie, est davantage sanctionné pour les seconds que pour les premiers. Ceci confirmerait alors qu'un déficit serait moins sanctionné en termes de participation au marché du travail et d'employabilité pour les natifs que pour les individus d'origine immigrée, en raison d'une discrimination à l'égard de ces derniers.

D'après les résultats des modèles logistiques polytomiques ordonnés généralisés à partir des indicatrices (cf. tableau D.2. en annexe D), un individu issu de la première génération d'immigration et en surplus de littératie a 3,4 fois plus de risques d'être inactif qu'un individu ayant un niveau *normal*. Il en est de même pour un surplus de compétences en numératie. Chez des individus particulièrement compétents mais *a priori* plus susceptibles d'être victimes de discrimination à l'embauche, peut naître un fort sentiment d'injustice et de découragement les amenant ainsi à se retirer du marché du travail. De plus, connaissant leur réel niveau de compétences, même si ces individus accédaient à l'emploi, ils pourraient estimer que cet emploi est bien trop précaire, peu rémunérateur et peu valorisant par rapport à leur potentiel, également source de découragement. Pour les français natifs, un déficit en numératie augmente significativement la probabilité d'être au chômage ou en inactivité de 34% tandis qu'un surplus la diminue de 26%. Dans un contexte de mutations par l'innovation technologique, la numératie est donc bien une compétence déterminante pour participer au marché du travail et accéder à l'emploi. Un déficit en littératie augmente quant à lui cette même probabilité de 36%, ce qui est nettement moins que pour les individus issus de la deuxième génération d'immigration (+101%), toutes choses égales par ailleurs. Être né en France mais être d'origine étrangère de par ses parents semble susciter la méfiance des employeurs, qui d'après la théorie de la discrimination statistique, privilégieront les Français natifs, groupe pour lequel ils estimeront une productivité moyenne plus élevée. Notons enfin, qu'un individu issu d'une ZUS a, toutes choses égales par ailleurs, significativement plus de risques d'être au chômage ou en inactivité, surtout

s'il est issu de la première génération d'immigration, ce qui tend à confirmer notre hypothèse de départ.

7. Conclusion

Les scores en compétences de base donnent un éclairage supplémentaire à l'analyse des trajectoires professionnelles, notamment pour les individus qui n'auraient pas un diplôme filtrant fidèlement les capacités individuelles. Les compétences de base complètent en ce sens l'information révélée par le diplôme sur la productivité individuelle et sur la capacité à être formé dans le métier. Elles font ainsi partie intégrante des variables objectivant la position occupée sur le marché du travail. En d'autres termes, elles font partie, pour certaines populations, des éléments qui expliquent pourquoi des individus sont plus susceptibles que d'autres d'être en emploi, au chômage ou en inactivité, indépendamment des formes existantes d'emploi et des différents types de chômage. Ne pas maîtriser les compétences de base essentielles pour s'insérer professionnellement peut se révéler être un des nombreux facteurs d'éloignement, si ce n'est d'exclusion, du marché du travail. Cela est d'autant plus alarmant que les compétences de base sont également nécessaires pour gérer les tâches de la vie quotidienne, de plus en plus complexes avec le développement des NTIC. Ne pas les maîtriser pourrait donc accélérer le processus d'exclusion professionnelle et sociale.

Afin d'accompagner sur le marché du travail les individus fragilisés dans leur rapport aux changements technologiques et organisationnels, les stratégies, qu'elles relèvent de politiques d'État ou d'entreprise, doivent s'axer sur la consolidation des compétences de base et de toutes autres compétences recherchées par les employeurs, via la formation continue par exemple. Ces stratégies doivent non seulement permettre de remédier aux méfaits d'une trop faible scolarité mais aussi de retarder l'obsolescence des compétences chez les personnes de plus de 50 ans, notamment de celles recherchées sur le marché du travail comme la numérisation. Sans cela, les peu diplômés, les décrocheurs et les seniors conserveraient une situation précaire face à l'emploi avec le risque de se décourager au point de se retirer du marché du travail. Cependant, en se retirant de ce dernier par découragement, qu'il soit le fruit de discriminations ou de caractéristiques individuelles, les individus

n'entreprendront plus beaucoup leurs compétences existantes, quelles qu'elles soient. Ce manque d'entretien risque alors d'accélérer l'érosion d'un niveau de compétences déjà fragile. À terme, ces individus s'exposent donc à un phénomène de trappe à inactivité que seules les formations continues pourraient casser pour rendre réversibles les trajectoires professionnelles. Une enquête longitudinale sur la valorisation des compétences de base sur le marché du travail devrait ouvrir le champ de la recherche en permettant d'appréhender ce phénomène sur le long terme.

BIBLIOGRAPHIE DE L'ESSAI 2

- ACEMOGLU D. & PISCHKE J.-S. [1999], Beyond Becker : training in imperfect labour markets, *Economic Journal*, vol. 109, n°453, pp. 112-142.
- ALLEN J. & DE GRIP A. [2007], *Skill obsolescence, lifelong learning and labor market participation*, Research Centre for Education and the Labour Market (ROA), Faculty of Economics and Business Administration, Maastricht University.
- ANANIADOU K., JENKINS A. & WOLF A. [2004], *The benefits to employers of raising workforce basic skills levels: A review of the literature*. National Research and Development Centre for Adult Literacy and Numeracy, Institute of Education, University of London.
- ATKINSON J. & SPILSBURY M. [1993], *Basic skills and jobs*. Adult Literacy and Basic Skills Unit (ALBSU) and Institute of Manpower Studies (IMS), London.
- AUBERT P., CAROLI E. & ROGER M. [2006], New technologies, organization and age: firm-level evidence, *Economic Journal*, vol. 116, n°509, pp. 73-93.
- BARHAM C. [2002], Economic inactivity and the labour market. *Labour Market Trends*, vol. 110, n°2, pp. 69-78.
- BERMAN E., BOUND J. & MACHINES S. [1998], Implications of Skill-Biased Technological Change : International Evidence, *Quarterly Journal of Economics*, vol. 113, n°4, pp. 1245-1279.
- BONAÏTI C., FLEURET A., POMMIER P. & ZAMORA P. [2006], *Pourquoi les moins qualifiés se forment-ils moins ?* Document d'études DARES, n°116
- BROWN P. & HESKETH A. [2004], *The mismanagement of talent: Employability and jobs in the knowledge economy*. Oxford University Press.
- BYNNER J., DEARDEN H., MCINTOSH S., REED J., VAN REENEN J. & VIGNOLES A. [2001], *Improving Adult Basic Skills: Benefits to the Individual and to Society*, DfEE Research Brief n°251.
- BYNNER J. [2004], Literacy, numeracy and employability : Evidence from the British birth cohort studies, *Literacy and Numeracy Studies*, vol. 13, n°1. pp.31-48.
- CASTEL R. [2009], *La montée des incertitudes. Travail, protection, statut des individus*. Éditions du Seuil, La couleur des idées :Paris, 457 p.
- CHARETTE M.-F. & MENG R. [1998], The determinants of literacy and numeracy, and the effect of literacy and numeracy on labour market outcomes, *Canadian Journal of Economics*, vol. 31, n°3, pp. 495-517.
- CHISWICK B.-R., LEE Y.-L. & MILLER P.-W. [2003], Schooling, literacy, numeracy and labour market success, *Economic Record*, vol. 79, n° 245, pp. 165-181.
- COLES A. & MASTERS M. [2000], Retraining and long-term unemployment in a model of unlearning by not doing, *European Economic Review*, vol. 44, n°9, pp. 1801-1822.

- DE COULON A., MARCENARO-GUTIERREZ O. & VIGNOLES A. [2007], *The Value of basic Skills in the British Labour Market*, Centre for Economic Performance, London School of Economics.
- DE GRIP A. & VAN LOO J. [2002], The Economics of Skills Obsolescence : A Review. In A. De Grip, J. Van Loo & K. Mayhew (Eds.), *The Economics of Skills Obsolescence*, Research in Labor Economics, vol. 21, JAI Press, pp. 1-26.
- DE GRIP A. & ZWICK T. [2005], *The employability of low-skilled workers in the knowledge economy*, Maastricht, The Netherlands: Maastricht University.
- DEARDEN L., MCINTOSH S., MYCK M. & VIGNOLES A. [2002], The Returns to Academic, Vocational and Basic Skills in Britain, *Bulletin of Economic Research*, vol. 54, n°3, pp. 249-274.
- DECREUSE B. & DI PAOLA V. [2002], L'employabilité des chômeurs de longue durée : mise en perspective des littératures théoriques et empiriques, *Revue d'Économie Politique*, vol. 112, n°2, pp. 197-227.
- DELATTRE E. & SALOGNON M. [2008], Entreprises, modes de gestion de la main-d'œuvre et allongement de la durée de chômage : une analyse économétrique, *Recherches économiques de Louvain*, vol. 74, n°3, pp. 299-325.
- DOERINGER P.-B. & PIORE M.-J. [1971], *Internal labor markets and manpower analysis*, Heath, Lexington, Massachusetts.
- DURRANI N. & TARIQ V.N. [2012], The role of numeracy skills in graduate employability, *Education and Training*, vol. 54, n°5, pp. 419-434.
- FINNIE R. & MENG R. [2007], Literacy and Employability, *Perspectives on Labour and Income*, vol. 19, n°2, pp. 44-52.
- FREESE J. & LONG J.-S. [2006], *Regression Models for Categorical Dependent Variables Using Stata*, 2nd edition, College Station, Texas, Stata Press.
- GÜRBÜZ A.-A., POLAT S. & ULUS M. [2013], *In Limbo: Exploring Transition to discouragement*. Galatasaray University, GIAM.
- KOEDEL C. & TYHURST E. [2012], Math Skills and Labor-Market Outcomes: Evidence from a Resume-Based Field Experiment, *Economics of Education Review*, vol. 31, n°1, pp. 131-140.
- LEMISTRE P. [2012], *La segmentation actuelle du marché du travail est-elle soluble dans la flexicurité ?*, CÉREQ, Net Doc, n°88.
- LORENZ E. & VALEYRE A. [2005], Les formes d'organisation du travail dans les pays de l'Union européenne, *Travail et emploi*, n°102, pp. 91-105.
- MANGINI J.-K. [2012], *Dropout: Students leaving urban high schools prior to graduation*, Counselor Education Master's Theses, (paper n°132), Brockport College, New-York.

- MCINTOSH S. & VIGNOLES A. [2001], Measuring and assessing the impact of basic skills on labour market outcomes, *Oxford Economic Papers*, vol. 53, n°3, pp. 453-481.
- PATRINOS H.-A. & SAKELLARIOU C. [2013], Adult literacy, heterogeneity and returns to schooling in Chile, *Education Economics*, pp. 1-15.
- PHELPS E. [1972], The Statistical Theory of Racism and Sexism, *American Economic Review*, vol. 62, n°4, pp. 659-661.
- PIERRON, R. [2012], La non-qualification: révélateur ultime ou masque multiforme de la nature du travail? Recension de l'ouvrage de José Rose, Qu'est-ce que le travail non qualifié ?, *Formation Emploi*, n°3, pp. 101-108.
- PONT B. & WERQUIN P. [2001], Les compétences pour l'économie du savoir in OCDE, *Analyse des politiques d'éducation 2001*, Paris : OCDE.
- PRULL M.-W., GABRIELI J.-D.-E. & BUNGE S.-A. [2000], Age-Related Changes in Memory : A Cognitive Neuroscience Perspective, in *Handbook of Aging and Cognition II* , Craik F.I.M. et Salthouse T.A. (éds.), pp. 91-153, Mahwah, NJ : Lawrence Erlbaum Associates: Mahwah (NJ).
- ROSEN S. [1975], *Measuring the obsolescence of knowledge*, New-York, National Bureau of Economic Research.
- ROSE J. [2012], *Qu'est ce que le travail non qualifié ?*, Paris, La Dispute, coll. Travail et salariat , 180 p.
- SOUCHAY C., ISINGRINI M. & ESPAGNET L. [2000], Aging, Episodic Memory, Feeling-of-Knowing, and Frontal Functioning, *Neuropsychology*, vol. 14 , n°2, pp. 299-309.
- SPENCE M. [1973], Job market signaling, *The quarterly journal of Economics*, vol. 87, n°3, pp. 355-374.
- TYLER J.-H. [2004], Basic skills and the earnings of dropouts, *Economics of Education Review*, vol. 23, n°3, pp. 221-235.
- WELCH F. & URETA M. [2002], The Obsolescence of Skill. In De Grip A., Van Loo J. & Mayhew K. (Eds.), *The Economics of Skills Obsolescence*, Research in Labor Economics, vol. 21, JAI Press, pp. 1-26.
- WILLIAMS R. [2006], Generalized Ordered Logit/ Partial Proportional Odds Models for Ordinal Dependent Variables, *The Stata Journal*, vol. 6, n°1, pp. 58-82.

ANNEXES DE L'ESSAI 2

ANNEXE A : CONSTRUCTION DES GROUPES DE SITUATIONS OCCUPÉES FACE AU MARCHÉ DU TRAVAIL

Au moment de l'enquête, par rapport au marché du travail, les individus se sont déclarés comme étant :

- 1) Actif occupé
- 2) Chômeur
- 3) Étudiant, en stage ou en formation
- 4) Militaire du contingent
- 5) Retraité ou préretraité
- 6) Retiré des affaires
- 7) Femme ou homme au foyer
- 8) Autre inactif

L'analyse portant sur les individus en âge de travailler (hors étudiants), les catégories 3, 4, 5 et 6 ont été ôtées. Les individus en préretraite ou en chômage partiel qui ont encore un pied dans l'emploi, sont considérés comme en emploi. En revanche, les étudiants qui travaillent pour financer leurs études sont considérés comme inactifs et sont donc ôtés de l'analyse puisque seuls les profils ayant fini leur formation et en âge de travailler sont considérés. Le reste des catégories peut être départagé en trois groupes : les actifs occupés, les actifs inoccupés (ceux qui se sont déclarés comme étant chômeurs) et les inactifs dont les femmes et hommes au foyer. Cependant, le statut auto-déclaré n'était pas toujours en phase avec le comportement décrit par ces mêmes personnes sur le marché du travail. Par exemple, des individus déclaraient ne pas rechercher d'emploi alors qu'ils avaient entrepris une ou plusieurs démarches en ce sens. A l'inverse, nombre d'individus se sont dits au chômage alors qu'ils exprimaient clairement leur souhait de ne pas travailler et de ne pas rechercher d'emploi. Officiellement, ces individus peuvent effectivement être au chômage, en raison de la perception d'indemnités, toutefois, ils sont dans une logique de retrait du marché du travail. C'est pourquoi ils ont été considérés comme inactifs. Plus généralement, voici le reclassement effectué selon le comportement constaté au regard du marché du travail :

<i>Statut déclaré :</i>	<i>Comportement sur le marché du travail :</i>	<i>Classé comme :</i>
Actifs occupés		En emploi (N=6 020)
Actifs inoccupés	Déclare ne pas rechercher d'emploi	Sans emploi mais en recherche active (N=824)
	Déclare rechercher un emploi mais n'a entrepris aucune démarche	
Inactifs	Femme au foyer, étudiant sortant du système éducatif et autres inactifs déclarant rechercher un emploi et ayant entrepris au moins une démarche	Sans emploi et pas de recherche active (retrait du marché du travail) (N=848)
	Déclare ne pas rechercher d'emploi mais désire travailler et a entrepris au moins une démarche	

ANNEXE B : RÉSULTATS DES MODÈLES LOGISTIQUES POLYTOMIQUES ORDONNÉS GÉNÉRALISÉS AVEC INTRODUCTION DE SCORES EN COMPÉTENCES

TABLEAU B.1 : IMPACT DES SCORES DE COMPÉTENCES SUR LA PROBABILITÉ D'APPARTENIR À TELLE OU TELLE SITUATION OCCUPÉE FACE AU MARCHÉ DU TRAVAIL SELON LE NIVEAU DE DIPLÔME

Situation face au marché du travail Catégorie 1 = en emploi Catégorie 2 = au chômage Catégorie 3 = inactif	MODÈLE 1 Non diplômés ou CEP (N= 1969)				MODÈLE 2 Diplômés du secondaire (N= 3822)				MODÈLE 3 Diplômés du supérieur (N= 1901)			
	Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)	
	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO
Durée théorique de scolarité	0,005	1,00	0,005	1,00	0,006	1,00	0,006	1,00	-0,055	0,95	-0,055	0,95
Score en littératie	-0,006	-5,73	-0,006	-5,73	-0,013**	-11,84	-0,013**	-11,84	-0,005	-4,43	-0,005	-4,43
Score en numératie	-0,004	-3,54	-0,004	-3,54	-0,008**	-7,42	-0,008**	-7,42	-0,01	-9,14	-0,01	-9,14
Score en compréhension orale	-0,005**	-4,82	-0,005**	-4,82	-0,005**	-5,11	-0,005**	-5,11	-0,003	-3,26	-0,003	-3,26
A connu des difficultés scolaires	0,097	1,10	0,097	1,10	0,128	1,14	0,128	1,14	0,439***	1,55	0,439***	1,55
Être une femme	0,799***	2,22	0,799***	2,22	1,068***	2,91	1,446***	4,24	0,700***	2,01	0,700***	2,01
Avoir au moins un enfant	0,017	1,02	0,017	1,02	0,115***	1,12	0,303	1,35	0,589***	1,80	0,589***	1,80
Tranche d'âge d'appartenance (réf : 31-50 ans)												
18 - 30 ans	0,904***	2,46	0,164	1,18	0,640***	1,90	0,168	1,18	0,491***	1,63	-0,313	0,73
51 - 65 ans	0,243**	1,27	0,743***	2,10	0,666***	1,95	1,223***	3,40	1,076***	2,93	1,076***	2,93
Origine migratoire (Réf : français natifs)												
Personnes issues de la première génération d'immigration	0,022	1,02	-0,212	0,81	0,863***	2,37	0,392**	1,48	1,012***	2,75	1,012***	2,75
Personnes issues de la seconde génération d'immigration	0,104	1,11	0,104	1,11	0,445***	1,56	0,445***	1,56	0,333	1,39	0,333	1,39
Constante	-0,38	--	-1,23***	--	-0,417	--	-1,67***	--	-0,952	--	-1,796	--
R ² ajusté	5.64%				8.32%				6.11%			

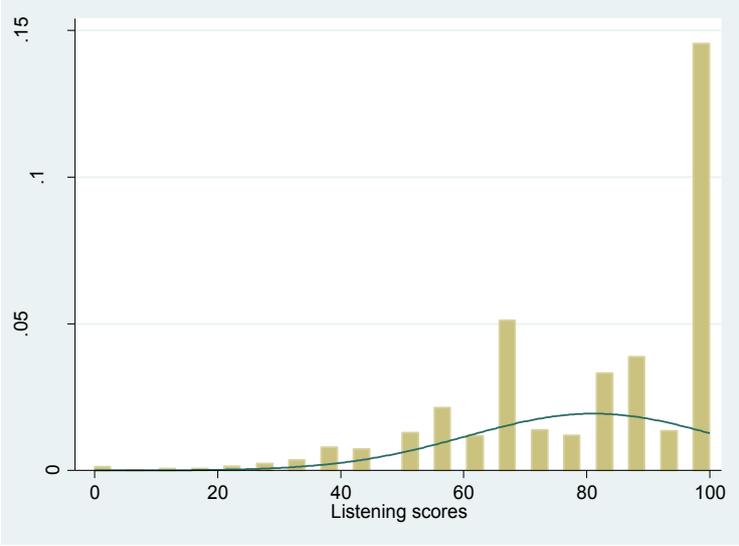
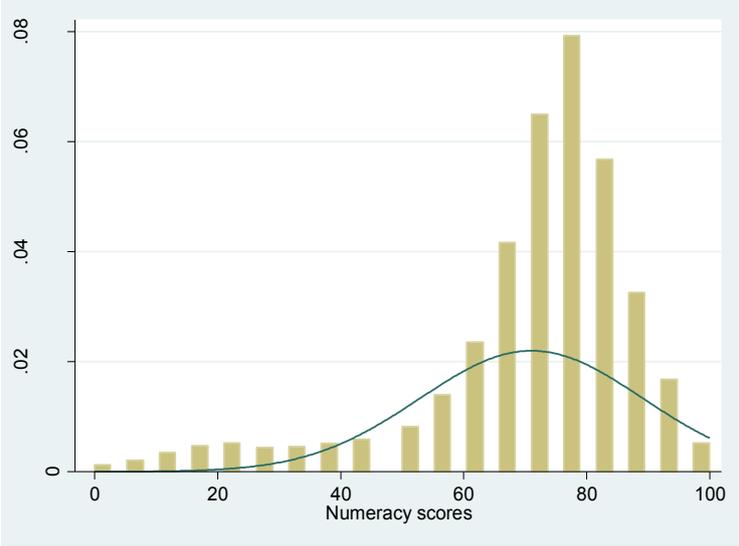
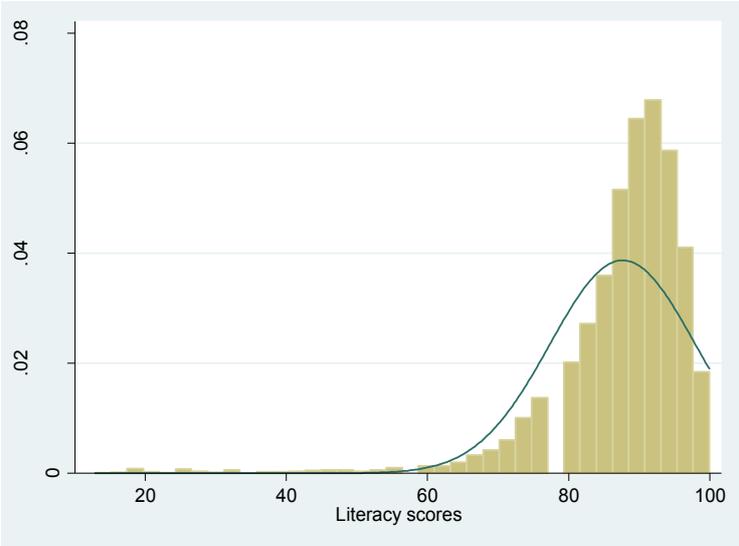
*** : significatif au seuil d'1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10% - Les odds-ratios des scores en compétences ont été calculés pour 10 points.

TABLEAU B.2 : IMPACT DES SCORES DE COMPÉTENCES SUR LA PROBABILITÉ D'APPARTENIR À TELLE OU TELLE SITUATION OCCUPÉE FACE AU MARCHÉ DU TRAVAIL SELON LA TRANCHE D'ÂGE D'APPARTENANCE

Situation face au marché du travail Catégorie 1 = en emploi Catégorie 2 = au chômage Catégorie 3 = inactif	MODÈLE 1 Personnes âgées de 18 à 30 ans (N=1569)				MODÈLE 2 Personnes âgées de 31 à 50 ans (N=4312)				MODÈLE 3 Personnes âgées de 51 à 65 ans (N=1811)			
	Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)	
	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO
Niveau de diplôme (réf: CAP/BEP)												
Sans diplôme	0.79***	2,20	0.79***	2,20	0.45***	1,57	0.45***	1,57	0.09	1,09	0.09	1,09
Certificat d'Études Primaires	0.75	2,12	0.75	2,12	0.43**	1,53	0.43**	1,53	0.27	1,31	0.27	1,31
Brevet des collèges	0.33	1,39	0.33	1,39	0.11	1,12	0.11	1,12	0.18	1,20	0.18	1,20
Baccalauréat	-0.33*	0,72	-0.33*	0,72	-0.07	0,93	-0.07	0,93	-0.38	0,68	-0.38	0,68
Bac +2	-1.00***	0,37	-1.00***	0,37	-0.49***	0,61	-0.49***	0,61	0.10	1,11	0.10	1,11
> Bac+2	-1.09***	0,34	-1.09***	0,34	-0.63***	0,53	-0.63***	0,53	-0.72***	0,49	-0.72***	0,49
Score en littératie	-0.007	-6,60	-0.007	-6,60	-0.008*	-7,35	-0.02***	-14,04	0.001	1,22	0.001	1,22
Score en numératie	0.002	2,33	0.002	2,33	-0.005**	-5,10	-0.005**	-5,10	-0.01***	-10,15	-0.01***	-10,15
Score en compréhension orale	0.001	1,05	0.001	1,05	-0.005**	-4,75	-0.005**	-4,75	-0.007***	-7,19	-0.007***	-7,19
Formations industrielles	0.045	1,05	0.045	1,04	-0.052	0,95	-0.052	0,95	0.382**	1,47	0.382**	1,47
Être une femme	1.056***	2,86	1.546***	4,69	0.986***	2,68	1.346***	3,84	0.684***	1,98	0.684***	1,98
Avoir au moins un enfant	0.488***	1,63	1.056***	2,83	0.0004	1,00	0.222*	1,24	-0.042	0,96	-0.042	0,96
Origine migratoire (Réf: français natifs)												
Personnes issues de la première génération d'immigration	0.412*	1,51	0.412*	1,51	0.601***	1,82	0.271*	1,31	0.439***	1,55	0.075	1,08
Personnes issues de la seconde génération d'immigration	0.412**	1,51	0.412**	1,51	0.404***	1,50	0.404***	1,50	0.122	1,13	0.122	1,13
Constante	-1.491**	--	-0.74***	--	-0.761**	--	-1.37***	--	-0.227	--	-0.615	--
R ² ajusté	10.91%				7.22%				4.70%			

*** : significatif au seuil de 1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10% - Les odds-ratios des scores en compétences ont été calculés pour 10 points.

ANNEXE C : HISTOGRAMMES ET DENSITÉS DES SCORES EN COMPÉTENCES DE BASE



ANNEXE D: RÉSULTATS DES MODÈLES LOGISTIQUES POLYTOMIQUES ORDONNÉS GÉNÉRALISÉS AVEC INTRODUCTION D'INDICATRICES DE DÉFICIT ET SURPLUS.

TABLEAU D.1 : IMPACT DES DÉFICITS ET SURPLUS DE COMPÉTENCES SUR LA PROBABILITÉ D'OCCUPER TELLE OU TELLE SITUATION SUR LE MARCHÉ DU TRAVAIL SELON LE LIEU DE RÉSIDENCE (ZUS - HORS ZUS)

Situation face au marché du travail Catégorie 1 = en emploi Catégorie 2 = au chômage Catégorie 3 = inactif	MODÈLE 1 ZUS (N=997)				MODÈLE 2 HORS ZUS (N=5850)			
	Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)	
	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO
Niveau de diplôme (réf: CAP/BEP)								
Sans diplôme	1.100***	3.00	1.100***	3.00	0.417***	1.52	0.417***	1.52
Certificat d'Études Primaires	0.516*	1.67	0.516*	1.67	0.387***	1.47	0.387***	1.47
Brevet des collèges	0.366	1.44	0.366	1.44	0.054	1.05	0.054	1.05
Baccalauréat	-0.024	0.98	-0.024	0.98	-0.39***	0.68	-0.39***	0.68
Bac +2	-0.283	0.75	-0.283	0.75	-0.77***	0.46	-0.77***	0.46
> Bac+2	-0.508	0.60	-0.508	0.60	-0.93***	0.39	-0.93***	0.39
État du niveau de compétences (réf: Niveau normal pour chacune des compétences)								
Déficit en littératie	0.247	1.28	0.247	1.28	0.286***	1.33	0.286***	1.33
Surplus en littératie	0.352	1.42	0.352	1.42	-0.036	0.96	-0.036	0.96
Déficit en numératie	0.127	1.14	0.127	1.14	0.252***	1.29	0.252***	1.29
Surplus en numératie	-0.409	0.66	-0.409	0.66	-0.134	0.87	-0.134	0.87
Déficit en compréhension orale	0.194	1.21	0.194	1.21	0.086	1.09	0.086	1.09
Surplus en compréhension orale	-0.191	0.83	-0.191	0.83	0.026	1.03	0.026	1.03
Avoir connu des difficultés scolaires	-0.116	0.89	0.405**	1.50	0.137*	1.15	0.137*	1.15
Être une femme	0.771***	2.16	1.481***	4.40	0.804***	2.23	0.997***	2.71
Avoir au moins un enfant	0.141	1.15	0.141	1.15	0.145*	1.16	0.145*	1.16
Formations industrielles	0.593**	1.81	0.593**	1.81	-0.085	0.92	-0.085	0.92
Tranche d'âge d'appartenance (réf: 31-50 ans)								
18 - 30 ans	0.512***	1.67	-0.230	0.79	0.663***	1.94	-0.021	0.98
51 - 65 ans	0.565***	1.76	1.378***	3.97	0.716***	2.05	1.318***	3.74
Origine migratoire (Réf: français natifs)								
Immigrés de première génération	0.301*	1.35	0.301*	1.35	0.384***	1.46	0.067	1.07
Immigrés de seconde génération	0.092	1.10	0.092	1.10	0.356***	1.43	0.356***	1.43
Constante	-0.03***	--	-0.03***	--	-0.41***	--	-0.48***	--
R ² ajusté	9.63%				8.95%			

*** : significatif au seuil d'1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10% - Les odds-ratios des scores en compétences ont été calculés pour 10 points, tout comme pour le tableau D.2.

TABLEAU D.2 : IMPACT DES DÉFICITS ET SURPLUS DE COMPÉTENCES SUR LA PROBABILITÉ D'OCCUPER TELLE OU TELLE SITUATION SUR LE MARCHÉ DU TRAVAIL SELON L'ORIGINE MIGRATOIRE

Situation face au marché du travail Catégorie 1 = en emploi Catégorie 2 = au chômage Catégorie 3 = inactif	MODÈLE 1 Première génération d'immigration (N= 730)				MODÈLE 2 Seconde génération d'immigration (N=842)				MODÈLE 3 Français natifs (N=5274)			
	Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)		Catégorie 1 (=0) VS catégories 2 et 3 (=1)		Catégories 1 et 2 (=0) VS catégorie 3 (=1)	
	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO	Coeff.	ODD-RATIO
Niveau de diplôme (réf: CAP/BEP)												
Sans diplôme	0.045	1.05	0.045	1.05	0.650**	1.91	0.650**	1.91	0.674***	1.96	0.674***	1.96
Certificat d'Études Primaires	-0.239	0.79	-0.239	0.79	0.443	1.56	0.443	1.56	0.513***	1.67	0.513***	1.67
Brevet des collèges	0.131	1.14	0.131	1.14	0.091	1.09	0.091	1.09	0.073	1.08	0.073	1.08
Baccalauréat	0.096	1.10	0.096	1.10	-0.640**	0.53	-0.640**	0.53	-0.351***	0.70	-0.351***	0.70
Bac +2	-0.487	0.61	-0.487	0.61	-0.902**	0.41	-0.902**	0.41	-0.723***	0.48	-0.723***	0.48
> Bac+2	-0.511	0.60	-0.511	0.60	-1.312***	0.27	-1.312***	0.27	-0.924***	0.40	-0.924***	0.40
État du niveau de compétences (réf: Niveau normal pour chacune des compétences)												
Déficit en littératie	0.031	1.03	0.031	1.03	0.697***	2.01	0.697***	2.01	0.309***	1.36	0.309***	1.36
Surplus en littératie	0.089	1.09	1.234***	3.43	-0.121	0.88	-0.121	0.88	-0.008	0.99	-0.008	0.99
Déficit en numératie	0.258	1.29	0.258	1.29	-0.190	0.83	-0.190	0.83	0.294***	1.34	0.294***	1.34
Surplus en numératie	0.725**	2.07	0.725**	2.07	-0.314	0.73	-0.314	0.73	-0.295*	0.74	-0.295*	0.74
Déficit en compréhension orale	0.312	1.36	0.312	1.36	0.324	1.38	0.324	1.38	0.039	1.04	0.039	1.04
Surplus en compréhension orale	0.096	1.10	0.096	1.10	0.171	1.19	0.171	1.19	-0.032	0.97	-0.032	0.97
A connu des difficultés scolaires	0.094	1.10	0.094	1.10	0.075	1.08	0.075	1.08	0.094	1.10	0.094	1.10
Être une femme	0.770***	2.16	0.770***	2.16	0.965***	2.63	0.965***	2.63	0.808***	2.24	1.044***	2.84
Avoir au moins un enfant	0.166	1.18	0.166	1.18	0.221	1.25	0.221	1.25	0.084	1.09	0.247**	1.28
ZUS	0.650***	1.91	0.650***	1.91	0.225	1.25	0.225	1.25	0.449***	1.57	0.449***	1.57
Formation industrielle	-0.262	0.78	-0.262	0.78	0.557**	1.74	0.557**	1.74	-0.029	0.97	-0.029	0.97
Tranche d'âge d'appartenance (réf: 31-50 ans)												
18 - 30 ans	0.499**	1.65	-0.409	0.66	0.499**	1.62	-0.146	0.86	0.700***	2.01	0.022	1.02
51 - 65 ans	0.831***	2.30	1.424***	4.15	0.831***	1.56	1.305***	3.69	0.700***	2.01	1.346***	3.84
Constante	-1.957***	--	-3.300***	--	-1.957***	--	-3.339***	--	-2.425***	--	-3.641***	--
R ² ajusté	8,49%				10,38%				9,76%			

*** : significatif au seuil d'1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10%

PARTIE II

**LA PRODUCTION DE COMPÉTENCES EN FRANÇAIS SELON
L'ORIGINE MIGRATOIRE**

ESSAI 3

CAPITAL HUMAIN ET DISTANCE LINGUISTIQUE : UNE INÉGALE ACQUISITION DES COMPÉTENCES EN FRANÇAIS SELON L'ORIGINE MIGRATOIRE

1. Introduction

L'immigration qualifiée a des effets vertueux sur l'économie du pays d'accueil. Les travailleurs qualifiés ont une productivité supérieure à la moyenne qui accroît le potentiel d'offre de l'économie et à terme le PIB/habitant, surtout si les compétences des migrants sont complémentaires à celles des natifs. Elle permet également aux pays développés de répondre à des besoins spécifiques propres aux secteurs en tension et de bénéficier des compétences d'une main d'œuvre qualifiée supplémentaire (Domingues Dos Santos, 2006). De plus, d'après les théories économiques, un stock global de capital humain élevé conditionne la rapidité de la croissance économique pour diverses raisons. Par exemple, l'affluence de travailleurs qualifiés incite les entreprises à se spécialiser dans les industries à fort contenu en travail qualifié, favorisant ainsi l'innovation et l'adaptation du système productif aux nouvelles technologies. De même, le niveau de qualification assure une meilleure insertion professionnelle et une meilleure adaptation aux évolutions et aux aléas du marché du travail. Par ailleurs, un migrant qualifié contribue à la réduction des dépenses publiques, surtout dans un contexte de vieillissement de la population (Saint Paul et Cahuc, 2009). Quel que soit l'âge, la contribution nette

d'impôts d'un travailleur qualifié est presque toujours positive et plus importante que celle d'un travailleur peu ou pas qualifié (Plassard et Desbarats, 2008).

Face à ces effets positifs non exhaustifs, l'immigration de travailleurs qualifiés constitue un enjeu macroéconomique pour les pays développés, comme la France, qui cherchent à attirer cette main d'œuvre compétente par le biais de leur politique d'immigration. En effet, pour augmenter l'offre de travailleurs qualifiés, soit les pays investissent en formation, soit ils importent de la main d'œuvre étrangère hautement qualifiée. Le recours à cette dernière est généralement privilégié si le besoin de travailleurs qualifiés dans un secteur particulier revêt un caractère urgent et temporaire. En outre, ce recours est bien moins coûteux pour les finances publiques puisque le pays d'accueil bénéficie des investissements éducatifs déjà financés par les pays d'émigration.

Si le recours à une main d'œuvre qualifiée représente un intérêt économique et financier, c'est aussi parce que cette main d'œuvre est parallèlement plus à même d'être familiarisée avec la langue du pays d'accueil ou a minima de l'acquérir avec plus de facilités. Son apprentissage à des fins d'intégration sociale et professionnelle est aussi devenu un enjeu politique, économique et social dans un nombre croissant de pays, tel la France, comme en témoigne la loi de 2007 relative à la maîtrise de l'immigration, à l'intégration et à l'asile⁴⁸. Dans une perspective plus économique, recourir à une main d'œuvre étrangère familiarisée avec la langue française évite aussi au pays d'accueil des coûts de formations linguistiques à l'arrivée des migrants. C'est ainsi que se pose naturellement la question de l'acquisition du français par les migrants. Cette question concerne également les enfants d'immigrés, soit les immigrés de deuxième génération, puisqu'ils sont plus enclins à évoluer dans un contexte pouvant être marqué par l'origine migratoire des parents. Comme le rappelle l'OCDE, d'après l'enquête PISA de 2009, en France, 17% des autochtones n'ont pas atteint le niveau minimal de compétences pour participer efficacement à la vie courante (niveau 2), contre 35% des élèves de la deuxième génération d'immigration.

⁴⁸ Depuis le 1er décembre 2008, toute personne âgée de 17 à 65 ans désireuse de s'installer en France par la voie du regroupement familial, doit passer un test dans son pays d'origine pour justifier d'une bonne maîtrise de la langue française et de sa connaissance des valeurs de la République⁴⁸. Seuls les individus ayant étudié dans une école francophone ou ayant suivi des études supérieures en France pendant au moins un an en sont dispensés. Les autres doivent, selon le niveau décelé lors du premier test, suivre jusqu'à deux mois de formation avant de passer un nouveau test.

Ainsi cet essai explore plus précisément les données de l'enquête IVQ de 2004 pour répondre à deux questions principales. Tout d'abord, les immigrants originaires d'une aire linguistique latine réussissent-ils significativement mieux à acquérir la langue française que ceux issus d'une aire linguistique non-latine, toutes choses égales par ailleurs ? Ensuite, quelle partie des écarts de scores en français écrit et oral est tantôt attribuée à des différences de caractéristiques moyennes, tantôt attribuée à des différences de performances entre les deux groupes d'individus considérés ? Nous supposons en effet que les écarts de scores entre les deux populations ne sont pas uniquement attribuables à des différences de caractéristiques, mais aussi à des différences d'efficacité résultant de la distance linguistique entre la langue maternelle et le français. Pour répondre à ces questions, nous estimerons des fonctions de production et utiliserons les techniques de décomposition d'Oaxaca (1973) et Blinder (1973), ainsi que celles de Juhn, Murphy et Pierce (1993). Avant cela, nous présenterons l'approche théorique sur laquelle repose cet essai. Elle consiste à montrer que si le choix de migrer dépend fortement du salaire espéré dans le pays de destination, la forte ségrégation des emplois, notamment envers les immigrants, peut, ex-post, décevoir ce choix et rendre moins indispensable la maîtrise de la langue du pays d'accueil, surtout si l'apprentissage de cette dernière représente un coût élevé.

2. Choix migratoires et capital humain

D'après l'approche néoclassique, le fait d'émigrer de son pays d'origine relève de l'individu lui-même qui, en tant qu'agent rationnel, cherche à minimiser ses coûts d'émigration et à maximiser son revenu en migrant notamment vers les marchés du travail qui offrent les meilleures perspectives de salaires. Un point essentiel de l'approche néoclassique est qu'elle considère la migration comme un facteur d'équilibrage des inégalités géographiques. En effet, les individus migreront des zones à bas salaires vers les zones à hauts salaires, modifiant ainsi la structure de l'offre et de la demande de travail, jusqu'à ce que le différentiel, net des coûts de migration, soit résorbé (Lewis, 1954). Cependant, selon le modèle d'Harris et Todaro (1970), ce n'est pas seulement la différence de salaires entre deux zones géographiques qui incite les individus à migrer, c'est aussi le salaire espéré par le migrant potentiel, compte tenu de ses caractéristiques et de ses coûts liés à la

migration. A l'inverse de l'approche néoclassique pure, ce modèle élargi relâche l'hypothèse de plein emploi et estime que les choix migratoires se font au regard du différentiel de salaires espérés plutôt que sur le différentiel de salaires observés. Concrètement, cela signifie que la variable déterminante du choix migratoire devient le salaire pondéré par la probabilité de trouver un emploi dans la région de destination.

D'après le modèle de Sjaastad (1962), les opportunités offertes sur les marchés du travail, dépendent essentiellement des caractéristiques propres aux migrants. En effet, ces dernières doivent contribuer à expliquer pourquoi certains sont plus averses au risque que d'autres, pourquoi ils ont une utilité et des coûts liés à la migration qui diffèrent et pourquoi ils ont un accès inégal à l'information sur les marchés du travail. A cet égard, le niveau de capital humain joue un rôle central dans les décisions individuelles de migrer en cela qu'il accroît le salaire espéré et la probabilité de trouver un emploi (Sjaastad, 1962). En fonction de leur niveau d'éducation, les candidats à la migration calculent la valeur actuelle des revenus futurs de leur capital humain dans différentes zones géographiques. La migration aura lieu si les rendements de ce capital, nets des coûts migratoires, sont supérieurs dans une région autre que celle d'origine. Les plus instruits sont aussi plus à même d'accéder à l'information, de la collecter et de la traiter de façon pertinente pour réduire les risques et les coûts liés à la migration (De Jong et Fawcett, 1981). Dans cette perspective, le niveau de capital humain influence donc le ratio coûts/bénéfices associé à la migration et le choix migratoire, au même titre que d'autres caractéristiques (âge, genre...), encore faut-il que le lieu d'arrivée envisagé valorise ce type de capital.

Toutefois, même si un niveau de capital humain élevé laisse entrevoir de meilleures opportunités professionnelles, la ségrégation des emplois, en termes d'accès, de statut ou encore de contenu du travail, peuvent vite les bloquer. En effet, outre les difficultés d'accès à l'emploi, notamment à l'emploi stable, une grande partie des immigrés et de leurs descendants reste cantonnée dans des emplois peu qualifiés et dans des secteurs d'activités soumis à de fortes variations cycliques de l'économie (Tanay et Audirac, 2000 ; Meurs et al., 2005) et où une bonne maîtrise de la langue du pays d'accueil n'est pas forcément nécessaire. Dayant, Echardour et Glaude (1997) observent que 75% des immigrés restent ouvriers après 20 ans de carrière alors que c'est le cas pour seulement 30% des français natifs. Certains groupes

migratoires auraient toutefois de meilleures perspectives professionnelles que d'autres, comme les Espagnols. A l'inverse, les Turcs et les Marocains seraient fortement soumis à une immobilité professionnelle. Les enfants d'immigrés auraient quant à eux profité d'une modification substantielle des formes d'activités, bien que leurs difficultés à entrer sur le marché du travail soient toujours aussi fortes, surtout pour les descendants de marocains et d'africains (Meurs et al., 2007). Mais malgré le *processus de diffusion dans l'espace des professions*, la ségrégation des emplois perdure pour certains groupes migratoires. Selon Meurs, Pailhé et Simon (2007), les enfants d'immigrés portugais restent surreprésentés dans le bâtiment et la petite industrie alors que les enfants d'immigrés maghrébins sont surreprésentés dans l'action sociale.

L'une des conséquences de la ségrégation professionnelle, assez marquée en France, est la faible incitation des migrants à apprendre la langue du pays d'accueil s'ils sont relégués dans des emplois ne nécessitant pas forcément une bonne maîtrise de cette dernière. Cela s'explique soit en raison du faible contenu cognitif de l'emploi occupé, soit parce que d'autres compétences sont recherchées dans cet emploi, comme les compétences sociales (Lleras, 2008). Ces compétences peuvent d'ailleurs compenser un faible niveau de maîtrise du français. Si la faible valorisation économique des compétences dans la langue du pays d'accueil désincite les migrants à investir dans l'acquisition de cette langue, notamment en raison de la ségrégation professionnelle, ils seront d'autant plus réfractaires à le faire que les coûts d'apprentissage seront élevés. En effet, apprendre et développer ses connaissances dans la langue du pays d'accueil constitue un investissement en capital humain que les individus cherchent à optimiser au regard de ses bénéfices mais aussi de ses coûts (Chiswick et Miller, 2001). Or, comme nous le montrerons dans la section suivante, les compétences propres à la langue du pays d'accueil seront d'autant plus coûteuses⁴⁹ à acquérir que la langue maternelle du migrant est structurellement éloignée de la langue de destination, toutes choses égales par ailleurs. En conséquence, si les individus sont financièrement moins enclins à investir dans l'acquisition de la langue du pays d'accueil que d'autres, ils risquent davantage de tomber dans une trappe à emplois peu-qualifiés. En effet, si une pauvre maîtrise de la langue du pays d'accueil n'est pas nécessairement un handicap sur le marché du travail, une bonne maîtrise reste indispensable pour accéder à des emplois plus qualifiés et plus stables.

⁴⁹ Ces coûts peuvent être directs et indirects. Directs de par les investissements purement matériels et indirects de par le manque à gagner en raison du temps octroyé à l'apprentissage de la langue du pays d'accueil.

3. Distance linguistique et inégale acquisition de la langue du pays d'accueil

Selon le pays d'origine et autres caractéristiques socio-économiques et démographiques, certains immigrés ont plus de difficultés que d'autres à acquérir la langue du pays d'accueil. Ceci peut être dû à différents facteurs (migration temporaire, nonaccès à une formation linguistique avant ou après la migration...) affectant inégalement les incitations à investir dans l'apprentissage de la langue du pays d'accueil. La distance linguistique fait partie de ces facteurs. Elle affecte l'efficacité d'apprentissage des langues et les coûts d'acquisition de celle du pays d'accueil (Esser, 2006). Selon Crystal (1987, p. 371), *Si la langue étrangère (L2) est structurellement similaire à la langue d'origine (L1), l'apprentissage est supposé être plus facile que dans les cas où la L2 est très différente de la L1*. Ainsi, si l'anglais est linguistiquement plus proche des langues d'Europe occidentale (français, allemand...) que des langues asiatiques (chinois, coréen...), alors les immigrés d'un pays anglophone originaires d'Europe occidentale atteindront un niveau supérieur de compétences en anglais, ou atteindront plus rapidement un niveau de compétences donné, que les immigrés originaires d'Asie de l'Est (Corder, 1981).

Un bon indicateur de distance linguistique est précieux pour mieux comprendre les écarts de capacités d'acquisition de la langue du pays d'accueil, le recours inégal à la formation linguistique ou encore le choix du pays de destination (Chiswick et Miller, 1994 ; Adsera et Pytlikova 2012). Cependant, la construction d'un tel indicateur est délicate en raison des différences structurelles profondes entre deux langues. Selon Snow (1998), la distance entre l'anglais et les autres langues par exemple augmente dans l'ordre suivant : le néerlandais (1), l'allemand (2), le scandinave (3) les langues latines (4), les langues slaves (5), les langues indoeuropéennes (6) et autres (7). Shields et Price (2002), pour le Royaume-Uni, se sont quant à eux limités à l'utilisation de variables dichotomiques renseignant sur le pays de naissance des immigrés en guise de mesure de la distance linguistique. D'autres ont proposé une échelle quantitative de la distance linguistique (Chiswick et Miller, 2004).

L'importance de cette dernière dans l'acquisition de la langue du pays d'accueil reste un domaine récent encore peu exploité par les sciences économiques. Pourtant, elle ne peut-être ignorée dans l'examen des déterminants de l'acquisition de la langue du pays d'accueil (Xi, 2013). Isphording et Otten (2011), en utilisant une mesure

sophistiquée de la distance linguistique⁵⁰, ont montré, à partir de données de panel allemandes, que la distance linguistique est le facteur le plus déterminant dans l'acquisition de la langue du pays d'accueil et le plus explicatif des écarts de compétences linguistiques entre les immigrés. D'après ces travaux, plus la distance linguistique augmente, plus la probabilité d'avoir de bonnes compétences dans la langue du pays d'accueil diminue. Ceci contraste avec ce qu'avancait Carliner (2000), pour qui la distance linguistique n'est pas aussi importante dans l'acquisition de la langue du pays d'accueil comparé à d'autres facteurs. Par ailleurs, Beenstock, Chiswick et Repetto (2001) ont montré que parmi les immigrés juifs en Israël, ceux dont la langue d'origine est l'arabe sont plus compétents en hébreu que les autres, toutes choses égales par ailleurs. Ils suggèrent que cela est dû à la faible distance linguistique qui sépare l'hébreu de l'arabe, notamment de par leur appartenance à la même branche linguistique (Crystal 1987 ; Grimes et Grimes, 1993).

La théorie du capital humain (Becker, 1964) permet d'éclairer ce phénomène de distance linguistique, par la distinction qu'elle opère entre compétences générales et compétences spécifiques. Les premières sont celles qui peuvent être transférées d'une entreprise à une autre, tandis que les secondes renvoient à celles acquises sur le tas et qui ne peuvent être valorisées que dans l'entreprise où elles ont été apprises. A cet égard, les compétences linguistiques peuvent être perçues comme des compétences générales transférables. Les capacités en langues sont principalement acquises durant la scolarité et sont mobilisables dans n'importe quelle entreprise. Mais lorsqu'il s'agit de populations immigrées, la maîtrise de la langue du pays d'accueil peut tantôt être perçue comme une compétence générale, tantôt comme une compétence spécifique. En effet, la maîtrise de la langue du pays d'accueil peut être une compétence spécifique si la distance linguistique est importante : certes, elle peut s'acquérir avant le processus de migration via de la formation, mais elle ne pourra vraiment se développer et se valoriser que sur le territoire d'accueil ou dans des pays partageant la même langue. La distance linguistique constitue ainsi un bon indicateur du degré de transférabilité et donc de généralité des compétences en langues (Chiswick et Miller, 2012).

Les recherches sur la distance linguistique visent principalement le public des primo-arrivants. En revanche, pour les enfants d'immigrés, la question de la distance

⁵⁰ Il s'agit d'une mesure récente développée par l'Institut Max Planck d'anthropologie évolutionniste. Basée sur un algorithme automatique visant à comparer la prononciation et le vocabulaire de deux langues, elle doit permettre de mieux expliquer l'hétérogénéité des compétences linguistiques des immigrés.

linguistique ne semble pas se poser. Nés en France, ces individus ont plus naturellement acquis la langue française, que la très grande majorité d'entre eux considère d'ailleurs comme leur langue maternelle. Mais le rythme d'assimilation de la langue française par les parents migrants eux-mêmes impacte la maîtrise des compétences en français écrit et oral de leurs enfants (Djajić, 2003). Les différentes caractéristiques sociomigratoires des parents peuvent donc entraver le processus d'acquisition du français, en particulier si l'origine migratoire est très éloignée de la culture et de la langue française et si elle est encore très présente dans la vie quotidienne des enfants. De fait, l'étude des écarts de compétences dans la langue écrite et orale du pays d'accueil selon l'origine migratoire doit autant concerner les primo-arrivants que les enfants d'immigrés.

4. Objet d'étude et hypothèses

Au regard de ces éléments théoriques et empiriques, dans quelle mesure l'origine migratoire affecte-t-elle les scores en français écrit et oral ? Nous ferons une distinction entre les immigrés d'origine latine et les immigrés d'origine non-latine, ceci en guise d'appréciation de la distance linguistique entre les langues d'immigration et le français. Cette distinction se justifie au regard des racines latines de la langue française. Nos données ne nous permettent pas de construire et d'utiliser une échelle continue pour quantifier la distance linguistique, c'est pourquoi cette dernière sera appréhendée de cette façon.

Secondement, dans quelle mesure les scores en français des immigrés d'origine latine diffèrent-ils de ceux des immigrés d'origine non-latine et quelle est l'ampleur des inégalités de scores entre ces deux groupes ? Plus précisément, ces inégalités sont-elles dues à des différences de caractéristiques ou bien à des différences de performances, dues à la distance linguistique (effet rendement) ? Cette question se justifie au regard des travaux de Chiswick et Miller (2001). En effet, l'apprentissage de la langue du pays d'accueil étant un investissement en capital humain, divers facteurs peuvent affecter ses coûts et ses bénéfices et ainsi même les décisions d'investissement. Chiswick et Miller (2001) entrevoient quatre groupes de facteurs : ceux liés à l'exposition de l'individu à la langue du pays d'accueil (*Exposure*), ceux liés aux incitations économiques à investir dans cette langue (*Economic Incentive*),

ceux liés à la richesse (*Wealth*) et ceux liés à l'efficacité productive de compétences linguistiques (*Efficiency*). La production de compétences dans la langue du pays d'accueil est alors fonction réduite de ces quatre types de facteurs, d'où :

$$LANG = f(Exposure, Efficiency, Economic Incentive, Wealth)$$

La distance linguistique évoquée précédemment est un élément qui, selon les deux auteurs, affecterait directement l'efficacité de production, tout comme les raisons de la migration ou encore l'âge.

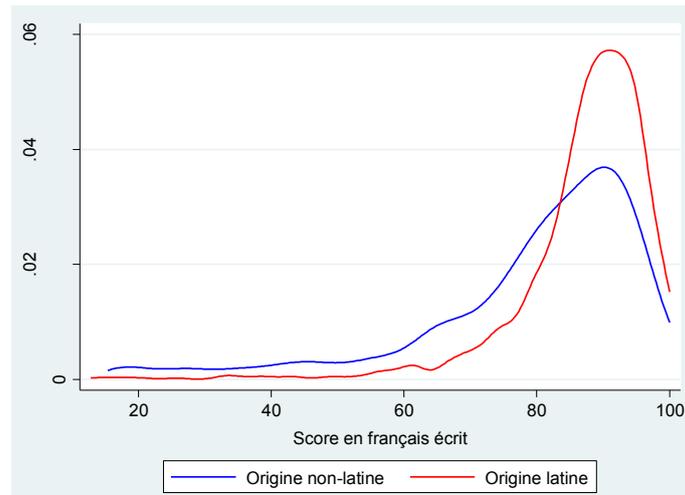
Compte tenu de cet objet d'étude et de la littérature existante, nous émettons donc deux hypothèses principales. Tout d'abord, nous supposons que la distance linguistique a un impact sur la production de compétences en français écrit et oral, tout comme d'autres variables sociomigratoires telle que la langue parlée à la maison. Secondement, au regard des travaux de Chiswick et Miller (2001), les inégalités de scores en français écrit et oral entre les deux groupes considérés (immigrés d'origine latine // immigrés d'origine non-latine) peuvent en partie s'expliquer par des différences d'efficacité en raison de la distance linguistique entre la langue maternelle et le français.

5. Statistiques descriptives

5.1. Les écarts de scores selon l'origine migratoire

Les données sont issues de l'enquête ménage Information et Vie Quotidienne réalisée par l'INSEE en 2004 afin de quantifier le niveau de compétences de base de la population française. Les scores en français écrit et oral sont sur une échelle allant de 0 à 100. La figure 1 montre une densité des scores en français écrit différente entre les deux populations d'immigrés.

Figure 1 : Distribution des scores en français écrit selon l'origine migratoire



Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

D'après la figure 1 et le tableau 1, les immigrés d'origine latine ont des scores en français écrit et oral plus élevés que ceux issus d'un pays non-latin. La proximité linguistique des langues latines (français, espagnol, italien, portugais) explique sans doute ce constat. Notons qu'il y a saturation des scores en compréhension orale à droite de la distribution, pouvant rendre délicate l'analyse des résultats de la décomposition à partir du 90ème centile.

Tableau 1 : Distribution des scores en français oral selon l'origine migratoire

	Population : individus issus de l'immigration (N=1717)		Écart
	Origine non-latine (n=492)	Origine latine (n=1225)	
Q10	33,3	50,00	17,0
Q25	50,0	66,7	16,7
Q50	66,7	83,3	16,3
Q75	88,9	100,0	11,1
Q90	100,0	100,0	00,0
Moyenne ⁵¹	68,4	79,2	10,8

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Les scores des immigrés d'origine non-latine étant plus faibles que ceux des individus étant d'origine latine, nous nous interrogeons sur les facteurs pouvant altérer les scores en français écrit et oral pour chacun de ces groupes. Le tableau 3

⁵¹ Les tests d'égalités de moyennes révèlent qu'il y a des différences significatives de scores moyens en français écrit et oral entre les immigrés d'origine latine et non-latine [$\Pr(|T|>|t|)=0.0000$].

présentera les résultats des fonctions de production de compétences en français écrit et oral selon l'origine migratoire.

5.2. Les variables de contrôle

Dans les modèles économétriques suivants, plusieurs variables de contrôle ont été intégrées afin de tenir compte de l'impact contextuel et des dimensions socio-économiques. Le tableau 2 donne un aperçu de la répartition des individus selon ces variables et selon l'origine migratoire. Premièrement, parmi les 22,8% des plus de 50 ans, 32,7% sont d'origine non-latine, ce qui est proportionnellement plus que pour les autres tranches d'âge. Les plus de 50 ans sont soumis au risque d'obsolescence technique de leurs compétences, d'où le nécessaire contrôle de nos modèles par cette variable. L'annexe B (cf. annexes) montre effectivement que les scores en français écrit et oral sont globalement plus faibles pour les plus de 50 ans composant notre échantillon, surtout s'ils sont d'origine non-latine.

Par ailleurs, parmi les 56,2% d'enfants d'immigrés de l'échantillon, seulement 9,7% sont d'origine non-latine. Cette faible part s'explique certainement par l'histoire de l'immigration en France. Jusqu'en 1975, les populations venues s'installer dans l'Hexagone étaient principalement originaires d'Europe latine (Cf. Pan Ké Shon, 2013), expliquant ainsi la forte proportion d'enfants d'immigrés d'origine latine (90,3%). Le fait que les immigrés de deuxième génération sont assez peu nombreux à être d'origine non-latine peut être un élément explicatif des inégalités de scores en français entre les deux populations. En effet, la durée de présence dans le pays d'accueil, égale à l'âge pour les immigrés de deuxième génération, joue un rôle important dans l'acquisition de la langue du nouveau pays. Ainsi, d'après l'annexe B (cf. annexes), les scores moyens en français écrit et oral des enfants d'immigrés sont supérieurs à ceux des primo-arrivants, quelle que soit leur origine. Contrôler par cette autre variable semble indispensable.

Tableau 2 : Statistiques descriptives selon l'origine migratoire (en %)

	Origine non-latine	Origine latine	TOTAL
Genre			
Homme (52,3)	28,4	71,6	100
Femme (47,7)	24,5	75,5	100
Age			
Moins de 30 ans (21,7)	26,7	73,3	100
Entre 30 et 50 ans (55,6)	24,4	75,6	100
Plus de 50 ans (22,8)	32,7	68,3	100
Enfant d'immigrés			
Oui (56,2)	9,7	90,3	100
Non (43,8)	48,2	51,8	100
Durée de scolarité théorique			
Entre 0 et 5 ans (10,7)	30,3	69,7	100
Entre 5 et 10 ans (17,6)	39,7	60,3	100
Entre 10 et 15 ans (60,0)	21,6	78,4	100
Entre 15 et 20 ans (11,7)	28,6	71,4	100
Nombre de redoublement(s) dans le cursus scolaire			
0 (15,6)	40,0	60,0	100
1 (11,7)	40,8	59,2	100
2 (14,7)	21,7	78,3	100
3 (16,9)	22,0	78,0	100
4 (19,0)	22,0	78,0	100
Plus de 4 (22,1)	20,2	79,8	100
Appartenance à une ZUS			
Oui (12,1)	51,8	48,2	100
Non (87,9)	23,1	76,9	100
Score de pauvreté			
0 - 5 (73,0)	21,3	78,7	100
6 - 10 (20,8)	37,7	62,3	100
11 - 15 (5,7)	49,3	50,7	100
16 - 20 (0,5)	72,2	27,8	100
Intensité d'usage de la langue française à la maison			
Seul le français est parlé (69,9)	11,7	88,3	100
Le français est la langue principale parlée (19,2)	54,2	45,8	100
Le français est une langue secondaire (6,3)	75,4	24,6	100
Le français n'est pas du tout parlé (4,6)	71,3	28,7	100
Fréquence de lecture de la mère quand l'enquête avait entre 8 et 12 ans			
La mère ne lisait jamais (39,6)	38,1	61,9	100
La mère lisait de temps en temps (23,9)	19,7	80,3	100
La mère lisait régulièrement (16,8)	16,2	83,8	100
La mère lisait quotidiennement (19,7)	20,7	79,3	100

Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Note : Les pourcentages entre parenthèses sont les pourcentages pour l'ensemble de la population

En outre, la part des immigrés d'origine latine augmente avec le nombre de redoublement(s) alors que c'est l'inverse pour les immigrés d'origine non-latine. En

effet, parmi les 15,6% d'individus n'ayant jamais redoublé, 60% sont d'origine latine alors que parmi les 22,1% d'individus à avoir redoublé au moins 4 fois, ils sont près de 80%. Ceci est certainement le résultat d'une pratique du redoublement largement plus répandue au sein des systèmes éducatifs latins (Paul et Troncin, 2004). L'OCDE, en 2009, soulignait que la proportion d'élèves de 15 ans ayant redoublé au moins une fois au collège était de 32% pour l'Espagne, 23,5% pour la France et 20,9% pour le Portugal tandis que cette proportion est en moyenne de 10,4% pour l'ensemble des pays de l'Union Européenne. Dans une logique plus internationale, l'Amérique Latine et l'Afrique subsaharienne, font partie des régions du monde ayant le plus recours au redoublement par rapport aux autres régions du monde (cf. annexes, Annexe A). Les langues latines étant surreprésentées dans ces régions (l'espagnol et le portugais pour l'Amérique latine, le français pour l'Afrique subsaharienne), on comprend mieux pourquoi la part des immigrés d'origine latine est d'autant plus élevée que le nombre de redoublement est grand.

D'autres caractéristiques sont assez intéressantes. Tout d'abord, les immigrés d'origine non-latine sont proportionnellement plus nombreux à avoir un faible niveau de scolarité (moins de 10 années d'éducation). Parmi les 28,3% d'individus ayant moins de 10 ans de scolarité, 35% sont d'origine non-latine alors que parmi les 71,7% d'individus ayant plus de 10 ans de scolarité, ils sont 25,1%. De même, parmi les 73% d'individus ayant les plus faibles scores de pauvreté en conditions de vie, 21,3% sont d'origine non-latine. Cette part augmente à mesure que le score en pauvreté s'élève. Ainsi, parmi les 0,5% d'individus particulièrement pauvres, 72,2% sont d'origine non-latine, ce qui est 2,6 fois plus que les immigrés d'origine latine. Les immigrés d'origine non-latine semblent globalement moins bien dotés en capital humain et plus sujets à la pauvreté en France, ce qui peut expliquer les inégalités de scores entre les deux populations. D'après l'annexe B (cf. annexes), en moyenne, les scores de compétences en français des immigrés d'origine non-latine augmentent (*diminuent*) avec la durée de la scolarité (*avec le niveau de pauvreté*) mais demeurent inférieurs à ceux des immigrés d'origine latine pour le même niveau d'éducation (*de pauvreté*) considéré.

Enfin, l'intensité du recours au français à la maison est une autre caractéristique essentielle pour tenter d'expliquer l'écart de scores entre les individus d'origine latine et les individus d'origine non-latine. En effet, parmi les 39,6% d'immigrés qui parlent uniquement le français à la maison, seulement 11,7% sont d'origine non-

latine. En revanche, ils sont proportionnellement plus nombreux que les immigrés latins à parler le français comme langue principale (54,2%) ou comme langue secondaire (75,4%). Ils sont également proportionnellement plus nombreux à ne jamais parler le français à la maison (71,3%). Ne jamais parler le français à la maison ou ne le parler qu'occasionnellement peut nuire aux compétences en français écrit et oral, comme semblent en témoigner les scores moyens présentés dans l'annexe B en annexes. Ces différences peuvent être source d'inégalités dans la maîtrise du français écrit et oral. La décomposition de ces derniers nous permettra de dire si oui ou non, les écarts sont effectivement le résultat de différences moyennes émanant de ces caractéristiques.

Pour estimer la performance productive de compétences en français et décomposer le différentiel des scores observés entre les deux groupes d'immigrés, nous allons, dans un premier temps, considérer deux fonctions de production de compétences (une en français écrit, une en français oral) pour identifier les caractéristiques qui ont un impact sur les scores individuels. Ensuite, nous utiliserons respectivement la méthode d'Oaxaca (1973) et Blinder (1973) et la méthode de Juhn, Murphy et Pierce (1993) afin de décomposer les différentiels de scores. Ces méthodes, relativement récentes en économie de l'éducation, sont aujourd'hui principalement appliquées aux données de l'enquête PISA (Ammermueller, 2007 ; Schneeweiss, 2011 ; Duque et al., 2013). L'originalité de cet essai est de les appliquer à un public migrant adulte à partir d'une source statistique encore inutilisée dans cette perspective.

6. Fonctions de production de compétences en français

Une fonction de production de compétences mesure l'impact des caractéristiques individuelles, familiales et contextuelles sur les compétences en français au niveau de l'immigré. Elle se modélise de la manière suivante :

$$Y_i = \beta X_i + u_i \quad (1)$$

Y_i est le score en français (écrit ou oral) du i -ème individu ($i = 1, \dots, n$), β , le vecteur des coefficients, X_i , les variables relatives aux caractéristiques de l'individu, de son milieu familial et de son contexte et u_i , le terme d'erreur.

Pour tester l'hypothèse selon laquelle deux groupes distincts ne sont pas soumis à la même performance productive, il est nécessaire de réaliser des estimations séparées de l'équation (1) pour chacun des sous-groupes considérés :

$$\begin{aligned} Y_i^n &= \beta^n X_i^n + u_i^n \\ Y_i^m &= \beta^m X_i^m + u_i^m \end{aligned} \tag{2}$$

Dans l'équation (2), les exposants font la distinction entre les deux groupes d'individus. Les exposants n et m désignent respectivement le groupe des immigrants d'origine latine et le groupe des immigrants d'origine non-latine, qu'ils soient primo-arrivants ou enfants d'immigrés. Le tableau suivant présente l'ensemble des fonctions de production de compétences en français, écrit et oral, à la fois pour l'échantillon total et les sous-échantillons (selon l'origine migratoire).

Tableau 3 : Fonctions de production de français écrit et oral selon l'origine migratoire (latine ou non)

	Français écrit			Français oral		
	<i>Total</i>	<i>Origine non-latine</i>	<i>Origine latine</i>	<i>Total</i>	<i>Origine non-latine</i>	<i>Origine latine</i>
Femme	-0,023 (-0,04)	-0,62 (0,46)	0,32 (0,66)	-1,64 (-1,60)	-3,68* (1,71)	-0,61 (0,53)
Age	-0,003 (-0,11)	0,03 (0,40)	-0,03 (1,35)	-0,04 (-0,85)	-0,03 (0,31)	-0,09 (1,45)
Durée de scolarité théorique	1,58*** (19,47)	2,10*** (11,01)	1,30*** (16,15)	1,97*** (12,38)	2,47*** (8,13)	1,66*** (8,89)
Nombre de redoublements	0,41 (3,48)	0,63** (2,26)	0,28** (2,35)	-0,10 (-0,42)	0,53 (1,19)	-0,48* (1,75)
Enfant d'immigrés	1,41** (2,22)	3,79** (2,12)	0,85 (1,45)	1,92 (1,54)	5,00* (1,75)	1,12 (0,82)
Appartenance à une ZUS	-1,68*** (-2,60)	-2,23 (1,57)	-1,09 (1,64)	-2,62** (-2,06)	-2,88 (1,27)	-2,11 (1,36)
Score de pauvreté	-0,14* (-1,87)	0,001 (0,01)	-0,20*** (2,71)	-0,45*** (-2,96)	-0,25 (0,85)	-0,53*** (3,03)
Intensité d'usage de la langue française à la maison						
Seul le français est parlé	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>
Le français est la langue principale parlée	-2,27*** (-3,14)	-1,83 (1,13)	-2,70*** (3,64)	-4,27*** (-3,02)	-1,62 (0,63)	-5,99*** (3,47)
Le français est une langue secondaire	-4,98*** (-4,21)	-3,60* (1,74)	-7,95*** (4,71)	-7,37*** (-3,18)	-4,04 (1,22)	-13,19*** (3,35)
Le français n'est pas du tout parlé	-7,43*** (-6,05)	-6,24*** (2,82)	-8,02*** (4,92)	-9,10*** (-3,78)	-6,20* (1,75)	-9,74** (2,56)
Fréquence de lecture de la mère quand l'enquêté avait entre 8 et 12 ans						
La mère ne lisait jamais	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>
La mère lisait quotidiennement	0,37 (0,51)	1,33 (0,66)	0,24 (0,35)	2,76** (1,90)	5,28* (1,65)	2,43 (1,51)
La mère lisait régulièrement	-1,01 (-1,29)	-5,26** (2,38)	0,23 (0,32)	1,04 (0,68)	-5,36 (1,52)	2,91* (1,73)
La mère lisait de temps en temps	-0,32 (-0,46)	-2,99 (1,62)	0,47 (0,73)	0,57 (0,42)	-2,17 (0,74)	1,49 (0,99)
Être d'origine latine	3,98*** (5,58)	--	--	4,74*** (3,40)	--	--
Constante	67,11*** (33,57)	57,34*** (12,19)	74,25*** (37,98)	56,61*** (14,02)	47,30*** (6,28)	67,87*** (14,88)
Observations	1717	492	1225	1717	492	1225
R ² ajusté	0,33	0,30	0,33	0,19	0,22	0,16

*** : significatif au seuil d'1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10%

D'après le tableau 3, sur l'ensemble de l'échantillon, être d'origine latine augmente les scores en français écrit et oral de 4 et 5 points respectivement, toutes choses égales par ailleurs. Les modèles par origine migratoire montrent que les variables

liées à l'immigration ont des impacts assez différents selon le groupe considéré. Ainsi, l'impact de la durée de scolarité est plus fort pour les immigrés d'origine non-latine, et ce, quelle que soit la compétence considérée. Être enfant d'immigrés améliore en revanche significativement les scores en français écrit et oral des individus qui ne sont pas d'origine latine. Leur présence en France depuis leur naissance, contrairement aux primo-arrivants, joue un rôle essentiel dans leur acquisition d'une bonne maîtrise du français, surtout oral. Ce rôle semble plus incertain pour les enfants d'immigrés non-latins compte tenu de la non-significativité des coefficients. La durée de présence paraît être d'autant plus déterminante des scores en français que les individus ont une langue maternelle structurellement éloignée du français. De plus, ne jamais parler le français à la maison, par rapport au fait de parler exclusivement cette langue, réduit les scores en français écrit d'un peu plus de 6 points pour les immigrés d'origine non-latine et de 8 points pour les immigrés d'origine latine. Ces chiffres sont respectivement de 6,2 points et de 9,7 points pour l'oral. L'impact est plus fort pour les immigrés d'origine latine et sur leur compréhension orale du français. La fréquence de lecture de la mère quand l'enquêté avait entre 8 et 12 ans ainsi que le score de pauvreté sont également deux variables ayant un impact sur l'acquisition de compétences en français mais dans une bien moindre mesure. Les variables liées au contexte migratoire sont donc susceptibles de mieux appréhender le processus productif des individus d'origine étrangère en matière de compétences en français.

7. Décomposition Oaxaca (1973) et Blinder (1973)

En reprenant les fonctions de production de compétences, l'estimation de la différence totale de scores moyens entre deux groupes différents, est telle que :

$$\bar{Y}^n - \bar{Y}^m = \hat{\beta}^n \bar{X}^n - \hat{\beta}^m \bar{X}^m \quad (3)$$

\bar{Y}^n et \bar{Y}^m sont les scores moyens des immigrés d'origine latine et non-latine respectivement, \bar{X}^n et \bar{X}^m , les vecteurs des valeurs moyennes des régresseurs à ces groupes et $\hat{\beta}^n$ et $\hat{\beta}^m$, les vecteurs correspondants des coefficients estimés.

La méthode développée par Oaxaca (1973) et Blinder (1973) permet de décomposer le différentiel total de scores en deux composantes : la composante des caractéristiques et la composante des coefficients. L'écart de score moyen entre les deux groupes est alors :

$$\bar{Y}^n - \bar{Y}^m = (\hat{\beta}^n - \hat{\beta}^m) \bar{X}^m - \hat{\beta}^n (\bar{X}^n - \bar{X}^m) \quad (4)$$

La composante des caractéristiques $\hat{\beta}^n (\bar{X}^n - \bar{X}^m)$ est dite expliquée, en cela qu'elle est attribuée aux différences de caractéristiques X . Elle représente la partie de l'écart total qui peut être expliquée par la variété de caractéristiques moyennes. Il s'agit de l'écart qui perdurerait entre les immigrés d'origine latine et non-latine si les premiers avaient les mêmes rendements de leurs caractéristiques en termes de résultats, soit la même performance ou efficacité, que les seconds. Il en résulte un écart entre les deux groupes simplement dû à des différences au niveau des variables explicatives, soit à une dotation différente en caractéristiques. A l'inverse, la composante des coefficients $(\hat{\beta}^n - \hat{\beta}^m) \bar{X}^m$, est dite inexpliquée car elle résulte de différences dans les coefficients β . L'écart inexpliqué capture l'écart de scores qui perdurerait si les immigrés d'origine latine avaient les mêmes caractéristiques moyennes que les immigrés d'origine non-latine. Il en résulte un écart dû à des différences de performances, toutes choses égales par ailleurs.

Cette décomposition pose cependant deux problèmes (Neumark, 1988 ; Jann, 2008 ; Fortin, 2008). Premièrement, le modèle proposé en (4) utilise comme score contrefactuel le score moyen qu'un immigré d'origine latine aurait obtenu s'il avait été soumis aux performances d'un immigré d'origine non-latine. L'inverse est également possible mais le score contrefactuel à partir duquel la décomposition s'effectue peut être légèrement différent, pouvant ainsi altérer partiellement les résultats. Selon Boudarbat et Pray (2011), le choix du score contrefactuel peut être contourné en appliquant la méthode de Neumark (1988). Elle consiste à utiliser les coefficients issus d'une fonction de production où les immigrés d'origine non-latine et latine sont regroupés mais où toutefois une variable binaire renseignant sur l'appartenance à tel ou tel groupe a été ajoutée comme régresseur (*pooled regression*). Nous emprunterons cette méthode. Le second problème concerne les variables catégorielles. La contribution de ces variables à la performance productive est difficilement identifiable car, une fois transformée en autant de variables binaires que de modalités possibles, le choix de la variable de référence altère les

résultats. Afin d'éviter ce problème, nous utiliserons la procédure *oaxaca* sous STATA (Jann, 2008). Elle limite les coefficients des variables binaires afin que leur somme soit nulle, ce qui revient à les exprimer en termes de déviations par rapport à la moyenne de la catégorie.

La décomposition des scores à partir de l'équation (4) peut ainsi se faire en estimant l'apport de chaque caractéristique du vecteur X , à l'écart expliqué et inexpliqué. L'ensemble des résultats est présenté dans le tableau 4 suivant :

Tableau 4: Décomposition de l'écart de scores à la moyenne entre les immigrés d'origine latine et ceux d'une origine non-latine

	<i>Français écrit</i>		<i>Français oral</i>	
Groupe n : immigrés d'origine latine [N=1225]	87,31***		79,19***	
Groupe m : immigrés d'origine non-latine [N=492]	79,32***		68,52***	
	<i>Coefficient</i>	<i>%</i>	<i>Coefficient</i>	<i>%</i>
Différence	7,99***	100	10,66***	100
Expliquée	4,04***	50,6	5,92***	55,5
Inexpliquée	3,94***	49,4	4,74***	44,5
EXPLIQUÉE				
Femme	-0,002	0,06	-0,154	2,4
Age	-0,001	0,04	-0,018	0,3
Durée de scolarité théorique	0,419	9,9	0,523	8,1
Nombre de redoublements dans le cursus scolaire	0,319***	7,5	-0,075	1,2
Enfant d'immigrés	0,713**	16,8	0,970	15,2
Appartenance à une ZUS	0,389**	9,2	0,606**	9,4
Score de pauvreté	0,256*	6,0	0,797***	12,4
Intensité d'usage de la langue française à la maison				
Seul le français est parlé	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>
Le français est la langue principale parlée	0,532***	12,6	1,002***	15,6
Le français est une langue secondaire	0,685***	16,2	1,015***	15,8
Le français n'est pas du tout parlé	0,806***	19,0	0,986***	15,4
Fréquence de lecture de la mère quand l'enquêté avait entre 8 et 12 ans				
La mère ne lisait jamais	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>	<i>Référence</i>
La mère lisait quotidiennement	0,021	0,5	0,157*	2,4
La mère lisait régulièrement	-0,071	1,7	0,072	1,1
La mère lisait de temps en temps	-0,023	0,5	0,042	0,7
INEXPLIQUÉE				
Femme	0,473	2,7	1,522	6,7
Age	-2,511	14,6	-2,287	10,1

Durée de scolarité théorique	-8,843***	51,2	-8,938**	39,2
Nombre de redoublements dans le cursus scolaire	-0,921	5,3	-2,660**	11,6
Enfant d'immigrés	-0,810*	4,7	-1,097	4,8
Appartenance à une ZUS	0,327	1,9	0,197	0,9
Score de pauvreté	-1,131	6,6	-1,556	6,8
Intensité d'usage de la langue française à la maison				
Seul le français est parlé	Référence	Référence	Référence	Référence
Le français est la langue principale parlée	0,229	1,3	-1,260	5,5
Le français est une langue secondaire	0,289	1,7	-0,669**	2,9
Le français n'est pas du tout parlé	0,171	1,0	-0,398	1,7
Fréquence de lecture de la mère quand l'enquête avait entre 8 et 12 ans				
La mère ne lisait jamais	Référence	Référence	Référence	Référence
La mère lisait quotidiennement	-0,175	1,0	-0,460	2,0
La mère lisait régulièrement	0,699*	4,1	1,056*	4,6
La mère lisait de temps en temps	0,677**	3,9	0,723	3,2
Constante	16,848***		20,568**	

*** : significatif au seuil d'1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10%

Lecture : 44,5% des écarts de scores en français écrit sont dus à des différences de performances, soit d'efficacité productive, entre les deux groupes. Autrement dit, les immigrés d'origine latine ont des caractéristiques qui, par rapport aux immigrés d'origine non-latine, leur procurent un avantage de compétences en français de l'ordre de 44,5%. S'en suit la part (in)expliquée de chaque variable rapportée à la part totale (in)expliquée (=100%). Elles ont été calculées à partir de la somme des valeurs absolues des coefficients propres à la part (in)expliquée.

D'après le tableau 4, l'écart de points entre les immigrés d'origine latine et les autres est plus important pour les scores en compréhension orale. La différence totale de points est de 8 points pour l'écrit et de 10,6 points pour l'oral. L'écart de scores à l'écrit s'explique autant par des différences de caractéristiques que de performances alors que l'écart de scores en compréhension orale est davantage expliqué par des différences de caractéristiques. En effet, 55,5% (5,92 points) des différences de scores observés à l'oral entre les immigrés d'origine latine et les immigrés d'origine non-latine sont attribuables à une dotation inférieure de caractéristiques moyennes des immigrés d'origine non-latine. La part inexpliquée, soit celle relative à une différence de performances, s'élève à hauteur de 44,5% (4,74 points). Les écarts d'efficacité jouent tous significativement en défaveur des immigrés d'origine non-latine tant à l'écrit qu'à l'oral.

De surcroît, les variables enfant d'immigrés, nombre de redoublement(s) dans le cursus scolaire, appartenance à une ZUS, score de pauvreté et intensité d'usage de la langue française à la maison favorisent les immigrés d'origine latine dans leurs résultats en français écrit. Cela signifie que si les immigrés d'origine latine avaient les mêmes dotations moyennes sur ces cinq caractéristiques que les immigrés

d'origine non-latine, ils auraient en moyenne des résultats inférieurs en français écrit. Par exemple, si la dotation moyenne en caractéristiques des enfants d'immigrés était la même entre les deux groupes, la part expliquée des inégalités de scores en français écrit diminuerait de 16,8%. Cette diminution serait de 19% si la dotation moyenne en caractéristiques des immigrés ne parlant jamais le français à la maison était identique entre les deux groupes. Notons aussi que plus le français est parlé à la maison, plus l'écart expliqué des scores attribuables à cette variable diminue. Ces deux caractéristiques sont les plus déterminantes des inégalités expliquées de scores en français écrit.

Concernant la part expliquée des inégalités de scores en français oral, seules les variables appartenance à une ZUS, score de pauvreté et intensité d'usage de la langue française à la maison jouent un rôle significatif en faveur des immigrés d'origine latine. Les deux derniers facteurs sont les plus importants. La part expliquée des inégalités de scores à l'oral diminuerait de 12,4% si la dotation moyenne en caractéristiques des individus ayant le même score de pauvreté était identique entre les deux groupes d'immigrés. Elle serait réduite de 15,4% à 15,8% si celle des immigrés ayant la même intensité d'usage du français à la maison était similaire. Plus précisément, si celle des immigrés ne parlant que le français à la maison était identique entre les deux populations, la part expliquée des inégalités de scores en français écrit diminuerait de 53,2% (3,4 points)⁵². Cette variable compose la part la plus importante de l'écart expliqué. Son effet sur la part inexpliquée est en revanche plus incertain vu la non-significativité globale des coefficients.

Par ailleurs, la durée de scolarité est une variable affectant très fortement les performances productives de compétences en français écrit et oral : toutes choses égales par ailleurs, elle compose la plus grande part de l'écart inexpliqué. Ainsi, si les performances des immigrés ayant un même niveau d'éducation étaient les mêmes entre les deux groupes, la part inexpliquée des inégalités de scores en français écrit et oral baisserait respectivement de 8,4 points (51,2%) et de 8,8 points (39,2%). Soulignons que l'efficacité estimée de la durée d'éducation est plus élevée pour les immigrés d'origine non-latine. Pour une durée d'éducation donnée et pour un niveau identique de caractéristiques moyennes, les immigrés d'origine non-latine mettent mieux à profit leur niveau d'études pour produire de la compétence en français écrit et oral que les immigrés d'origine latine.

⁵² 53,2% car on retranche la somme des parts propres à chacune des modalités de la variable (46,8%) de l'écart expliqué total (100%).

A l'inverse, les pratiques de lecture de la mère impactent d'autant plus les performances d'acquisition de la langue française, surtout à l'écrit, que les individus sont d'origine latine. Ce constat est cohérent avec l'importance de la durée de scolarité pour ceux d'origine non-latine. Compte tenu des racines linguistiques structurellement différentes de celles de la langue française et donc des langues latines en général, ils ne peuvent pas se *nourrir* de certains environnements, contrairement aux immigrés d'origine latine, pour acquérir et consolider leurs connaissances d'une langue latine. Plus limités dans leurs sources de familiarisation avec la structure de ce type de langue, ils sont mécaniquement plus enclins à être performants dans leur utilisation des capacités et des savoirs principalement acquis et développés durant la scolarité, en tant que principale source d'appropriation d'une langue étrangère. Cette explication semble d'autant plus cohérente que les primo-arrivants migrent généralement vers le pays dont ils ont étudié la langue (Chiswick et Miller, 1994 ; Wolton, 2008).

La variable enfant d'immigrés est significative dans la part inexpliquée des inégalités des scores, mais à un faible seuil de significativité (10%). Pour un niveau de caractéristiques moyennes identique, les enfants d'immigrés sont légèrement plus performants que les immigrés d'origine non-latine. De plus, outre la durée de scolarité théorique, le nombre d'années redoublées pendant la scolarité affecte aussi significativement les performances productives de compétences en français oral. Si les performances des immigrés ayant un même nombre total de redoublement(s) étaient semblables entre les deux populations, la part inexpliquée des inégalités de scores en français oral diminuerait de 2,7 points (11,6%). Notons que l'effet efficacité associé au(x) redoublement(s) est plus favorable aux immigrés d'origine non-latine, ce qui renforce l'idée que la formation initiale joue un rôle important dans les performances en français, en particulier pour ceux qui ont une langue maternelle éloignée de la branche des langues latines.

8. Décomposition des scores selon Juhn, Murphy et Pierce (1993)

Un des inconvénients de la décomposition proposée par Oaxaca-Blinder est qu'elle ne considère que la moyenne de la variable d'intérêt. La décomposition proposée par Juhn, Murphy et Pierce (1993) considère l'ensemble de sa distribution. De plus, elle se justifie car la décomposition proposée par Oaxaca-Blinder peut surestimer l'effet efficacité. Contrairement à celle de Juhn, Pierce et Murphy (1993), elle ne considère pas l'effet des compétences non-observées, soit généralement des compétences non-cognitives. Or certains travaux dans la littérature ont montré l'importance de ces compétences sur l'acquisition de compétences cognitives (Borghans et al., 2008 ; Behncke, 2009). Le modèle de Juhn, Pierce et Murphy (1993) doit donc donner un éclairage supplémentaire sur la décomposition des scores de compétences de base selon l'origine migratoire, d'autant que les écarts de scores s'amenuisent le long de la distribution. Reprenons les fonctions de production de compétences des immigrés d'origine latine (n) et des immigrés d'origine non-latine (m) :

$$\begin{aligned} Y_i^n &= \beta^n X_i^n + u_i^n \\ Y_i^m &= \beta^m X_i^m + u_i^m \end{aligned} \quad (5)$$

u_i^n , le résidu du groupe n , regroupe deux composantes : un centile individuel dans la distribution du résidu (θ_i^n) et la fonction de distribution des résidus de la fonction de production de compétences, $F^n(\cdot)$. La fonction de distribution cumulative des résidus peut alors être ainsi exprimée :

$$u_i^n = F^{n(-1)}(\theta_i^n | X_i^n) \quad (6)$$

$F^{n(-1)}(\cdot | X_i^n)$ correspond à la fonction de distribution cumulative inverse du résidu. Les changements dans les écarts entre des deux groupes peuvent être dus à des changements dans la distribution des caractéristiques Z , à des changements d'efficacité β , ou à des changements dans la distribution des résidus, c'est-à-dire dans l'influence des facteurs inobservables qui relèvent principalement des compétences non-observables selon Juhn, Murphy et Pierce (1993).

Suite aux fonctions de production (5), il est possible d'utiliser les coefficients estimés pour définir la distribution actuelle des scores en français (écrit ou oral) ainsi que deux distributions hypothétiques pour chacun des deux groupes étudiés. Les distributions actuelles pour ces derniers correspondent aux scores avec des caractéristiques, des performances et une distribution des résidus qui varient, d'où :

$$Y_i^n(Act) = \hat{\beta}^n X_i^n F^{n(-1)}(\theta_i^n | X_i^n) \quad (7a)$$

$$Y_i^m(Act) = \hat{\beta}^m X_i^m F^{m(-1)}(\theta_i^m | X_i^m) \quad (7b)$$

Les distributions hypothétiques avec des caractéristiques fluctuantes entre les groupes, des performances (β) et des résidus fixés, sont ainsi exprimées⁵³ :

$$Y_i^n(hyp\ 1) = \hat{\beta}^n X_i^n + F^{n(-1)}(\theta_i^n | X_i^n) \quad (8a)$$

$$Y_i^m(hyp\ 1) = \hat{\beta}^m X_i^m + F^{m(-1)}(\theta_i^m | X_i^m) \quad (8b)$$

Les distributions hypothétiques avec des caractéristiques et des performances fluctuantes et des résidus fixés s'expriment ainsi :

$$Y_i^n(hyp\ 2) = \hat{\beta}^n X_i^n + F^{n(-1)}(\theta_i^n | X_i^n) \quad (9a)$$

$$Y_i^m(hyp\ 2) = \hat{\beta}^m X_i^m + F^{m(-1)}(\theta_i^m | X_i^m) \quad (9b)$$

L'équation (8b) révèle quelle serait la distribution des scores des immigrés d'origine non-latine si ces derniers utilisaient le processus de production de compétences des immigrés d'origine latine. Elle donne également une idée sur ce que seraient les résidus correspondant à la distribution des résidus des immigrés d'origine latine. L'équation (9b), quant à elle, donne une indication sur ce que serait la distribution des scores des immigrés d'origine non-latine si la distribution des résidus était la même que celle des immigrés d'origine latine. Soit Y , une statistique résumant la distribution d'une variable. Y serait la moyenne ou la dispersion interquartile de la distribution de Y_i . L'écart $Y^n - Y^m$ peut alors être décomposé de la façon suivante :

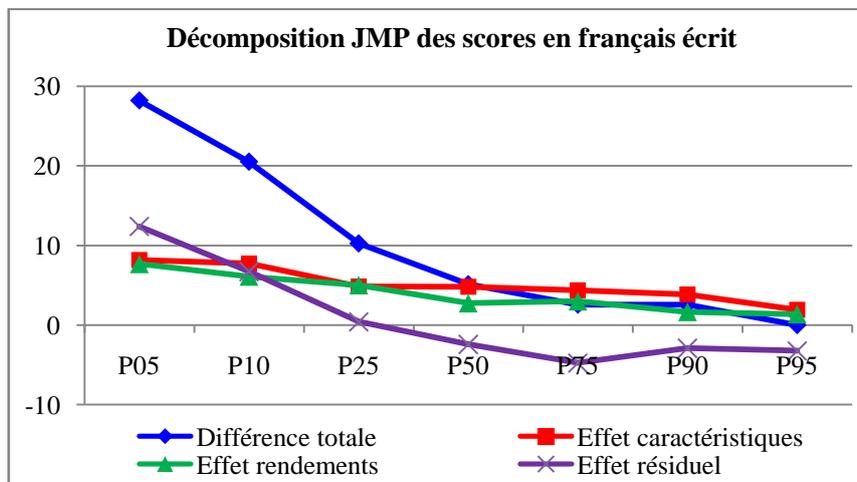
⁵³ Les immigrés d'origine latine constituent le groupe de référence. Les coefficients et les résidus de l'estimation réalisée pour ce groupe sont donc respectivement utilisés comme prix et distribution résiduelle de référence.

$$\begin{aligned}
 Y^n - Y^m &= [Y^n(\text{Hyp1}) - Y^m(\text{Hyp1})] \\
 &+ [(Y^n(\text{Hyp2}) - Y^m(\text{Hyp2})) - (Y^n(\text{Hyp1}) - Y^m(\text{Hyp1}))] \\
 &+ [(Y^n(Y) - Y^m(Y)) - (Y^n(\text{Hyp2}) - Y^m(\text{Hyp2}))]
 \end{aligned}
 \tag{10}$$

La différence totale (T) est alors décomposée en trois effets : l'effet des caractéristiques (Q), l'effet des rendements (P) et l'effet des résidus (U).

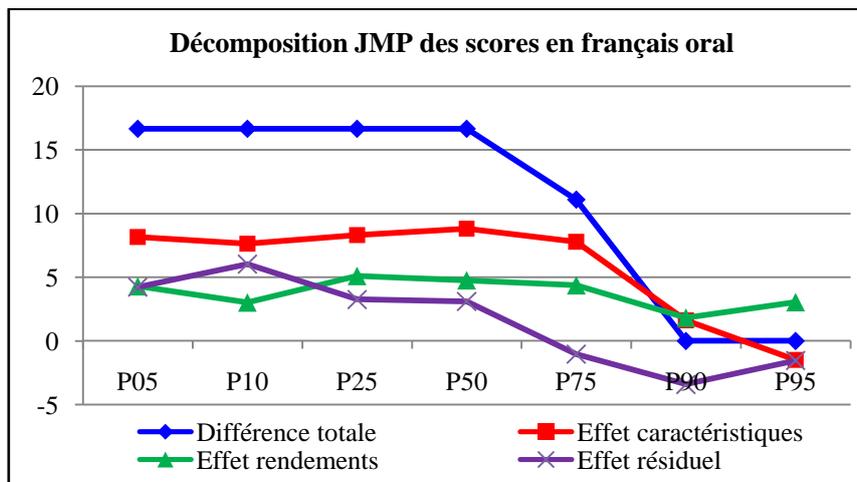
D'après l'annexe C (cf. annexes), les inégalités de scores en français écrit sont particulièrement importantes au 5ème centile, le différentiel total étant de 28,2 points. Ces inégalités diminuent cependant jusqu'à devenir nulles au 95ème centile. La courbe associée à l'effet des caractéristiques est décroissante à mesure que les scores s'élèvent, signe que les caractéristiques des immigrés d'origine non-latine s'améliorent tout au long de la distribution. De même les inégalités de caractéristiques sont plus fortes au début de la distribution (5ème centile) qu'à la fin. Les performances en français expliquent quant à elles presque tout autant les écarts de scores que les caractéristiques, quel que soit le point de la distribution. Cet effet s'affaiblissant à mesure que les scores augmentent, les inégalités de performances sont aussi les plus élevées pour les immigrés les plus faibles en français. De fait, l'origine migratoire est d'autant moins pénalisante que les individus sont compétents en français. Enfin, l'effet résiduel, décroissant du 5ème au 75ème centiles, et négatif à partir du 25ème centile, suggère que les inégalités de facteurs inobservables sont de moins en moins déterminants des inégalités de scores entre les deux groupes considérés. Néanmoins, l'effet résiduel est plus important que l'effet des caractéristiques et des performances pour les individus du 5ème centile de scores à l'écrit. Les fortes inégalités de scores constatées pour ces derniers sont surtout dues à une différence de dotation en compétences non-cognitives. Ceci serait alors cohérent avec les différents travaux montrant l'impact des compétences non-cognitives sur le niveau de compétences cognitives et le niveau d'éducation.

Figure 2: Effets caractéristiques et effets rendements le long de la distribution des scores en français écrit selon l'origine migratoire (latine ou non)



Les inégalités de scores à l'oral sont globalement moins importantes que celles à l'écrit, sauf entre le 25ème et le 75ème centile, preuve que les écarts de scores en français écrit diminuent plus rapidement que ceux en français oral. Elles restent toutefois stables jusqu'à la médiane avant de diminuer brusquement et devenir nulles du 90ème au 95ème centile. Contrairement au français écrit, l'effet des caractéristiques est plus important que l'effet rendements sur la quasi-totalité de la distribution. Les caractéristiques des immigrés d'origine non-latine se détériorent cependant dans la première partie de la distribution et s'améliorent dans la seconde. Les écarts de scores dus à des différences de caractéristiques sont donc les plus élevés au niveau de la médiane et les plus faibles sur les 90ème et 95ème centiles. L'effet des caractéristiques devient même négatif au 95ème centile. Cela suppose que les immigrés d'origine non-latine qui ont les meilleurs scores en français oral ont des caractéristiques supérieures à celles des immigrés d'origine latine. En revanche, ces derniers ont de meilleures performances, pouvant s'expliquer par la distance linguistique entre la langue maternelle et le français. Celle-ci paraît d'autant moins pénalisante que les individus sont très faibles ou au contraire très bons en français oral.

Figure 3: Effets caractéristiques et effets rendements le long de la distribution des scores en français oral selon l'origine migratoire (latine ou non)



9. Conclusion

Les principaux résultats de cet essai montrent que les immigrés d'origine latine sont plus à même de produire de la compétence en français. Les immigrés d'origine non-latine semblent aussi utiliser plus efficacement leur durée moyenne d'éducation que les immigrés d'origine non-latine, expliquant ainsi partiellement les écarts de scores en français. Enfin, l'origine migratoire est d'autant moins pénalisante que les individus sont performants en français. Mais elle reste moins pénalisante que le niveau de compétences non-cognitives pour les individus les plus faibles, surtout à l'écrit.

Concernant les enfants d'immigrés, comment est-il possible d'atténuer les effets liés à l'origine migratoire de leurs parents pour favoriser leur bonne maîtrise de la langue française, indispensable pour une intégration professionnelle et sociale réussie ? Le système éducatif joue ici un rôle important. Néanmoins certaines études montrent que les moindres performances des immigrés sont non seulement le résultat d'une dotation plus faible en caractéristiques, mais également le résultat d'une ségrégation scolaire (Meunier, 2011). Dans ce contexte, comment enrayer la propagation des inégalités issues de l'origine migratoire face à la langue du pays d'accueil d'une génération à une autre ? Cette question est centrale car la persistance de ces inégalités, sans compter les pratiques de discrimination à

l'embauche, peut entraver la participation au marché du travail, l'employabilité et par conséquent la réussite du processus d'intégration. Chercher à évaluer l'impact de la ségrégation scolaire à l'égard des populations immigrées sur les *outcomes* du marché du travail ouvrirait ainsi le champ de recherches.

En outre, comme la maîtrise du français par les migrants diminue les coûts de formations linguistiques pour le pays d'accueil et comme elle conditionne l'intégration professionnelle et sociale, les politiques d'immigration se durcissent et la connaissance de la langue française devient une condition. En conséquence, il se peut que le niveau moyen en français des migrants augmente à terme en raison de la sélectivité des migrants qui se fera de plus en plus en faveur des plus compétents en français et des plus qualifiés en général. De plus, les formations linguistiques imposées par la loi du 20 novembre 2007 aux demandeurs de regroupement familial ayant un faible niveau en français, peuvent contribuer à l'élévation du niveau général en langue française du public migrant. Mais le critère de réussite au test final pour l'obtention d'un visa de long séjour n'est pas tant le niveau atteint lors de l'examen que l'assiduité⁵⁴. Ce constat doit susciter le débat autour des politiques de formations à destination du public migrant. Si l'offre de formations linguistiques est dorénavant relativement bien structurée, qu'en est-il de son efficacité ? A ce jour, aucune évaluation d'envergure n'a été réalisée à ce sujet, ce qui laisse entrevoir de nouvelles pistes de recherche. Néanmoins, ceci n'est pas simple puisque l'analyse des situations d'apprentissage en formation ne peut être dissociée de l'analyse des situations d'apprentissage dans la vie sociale de tous les jours. Les formations linguistiques à destination des migrants s'inscrivent en effet dans un processus plus large d'appropriation de la langue où *l'essentiel des acquis se constitue au contact direct des locuteurs natifs, au cours des interactions quotidiennes* (Adami, 2012).

⁵⁴ Article 1 de la loi n°2007-1631 du 20 novembre 2007

BIBLIOGRAPHIE DE L'ESSAI 3

- ADAMI H. [2012], La formation linguistique des migrants adultes, *Savoirs*, n°2, pp. 9-44.
- ADSERA A. & PYTLIKOVA M. [2012], *The role of language in shaping international migration* (No. 2012014). Norface Research Programme on Migration, Department of Economics, University College London.
- AMMERMUELLER A. [2007], Poor background or low returns ? Why immigrant students in Germany perform so poorly in the programme for international student assessment, *Education Economics*, vol. 15, n°2, pp. 215-230.
- BECKER G.-S. [1964], *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, by Gary S. Becker, reprinted 1993. University of Chicago Press.
- BEENSTOCK M., CHISWICK B.-R. & REPETTO G.-L. [2001], The effect of linguistic distance and country of origin on immigrant language skills : application to Israel, *International Migration*, vol. 39, n°3, pp. 33-60.
- BEHNCKE S. [2009], *How Do Shocks to Non-Cognitive Skills Affect Test Scores?* (No. 4222). Institute for the Study of Labor (IZA): Bonn.
- BLINDER A.-S. [1973], Wage discrimination: reduced form and structural estimates, *Journal of Human resources*, vol. 8, n°4, pp. 436-455.
- BORGHANS L., MEIJERS H. & TER WEEL B. [2008], The role of non cognitive skills in explaining cognitive test scores, *Economic Inquiry*, vol. 46, n°1, pp. 2-12.
- BOUDARBAT B. & PRAY M. C. [2011], *L'écart salarial entre les sexes chez les nouveaux diplômés postsecondaires*. CIRANO: Montréal.
- CARLINER G. [2000], "The language ability of U.S. immigrants : assimilation and cohort effects, *International Migration Review*, vol. 34, n°1, pp. 158-182.
- CHISWICK B.-R. & MILLER P.-W. [1994], Language choice among immigrants in a multi-lingual destination, *Journal of Population Economics*, vol. 7, n°2, pp. 119-131.
- CHISWICK B.-R. & MILLER P.-W. [2001], A model of destination-language acquisition: Application to male immigrants in Canada, *Demography*, vol. 38, n°3, pp. 391-409.
- CHISWICK B.-R. & MILLER P.-W. [2004], *Linguistic distance: A quantitative measure of distance between English and other Languages*. IZA Discussion Paper n° 1246.
- CHISWICK B.-R. & MILLER P.-W. [2012], Negative and positive assimilation, skill transferability, and linguistic distance, *Journal of Human Capital*, vol. 6, n°1, pp. 35-55.
- CORDER S. P. [1981], *Error analysis and interlanguage*. Oxford : Oxford university press.

- CRYSTAL D. [1987], *The encyclopedia of language*. Cambridge : CUP.
- DE JONG G.-F. & FAWCETT J.-T. [1981], Motivations for migration: An assessment and a valueexpectancy research model. In De Jong, G.F. & Gardner, R.W. (eds.), *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries* (pp. 13-58). New York: Pergamon Press.
- DAYAN J.-L., ECHARDOUR A. & GLAUDE M. [1997], Le parcours professionnel des immigrés en France: une analyse longitudinale, *Old and New Minorities*, pp. 113-146.
- DJAJIĆ S. [2003], Assimilation of immigrants: implications for human capital accumulation of the second generation, *Journal of population economics*, vol. 16, n°4, pp. 831-845.
- DOMINGUES DOS SANTOS M. [2006], Attraction des élites et exode des cerveaux: les enjeux économiques d'une concertation entre pays d'origine et pays d'accueil, *Horizons Stratégiques*, n°1, pp. 18-27.
- DUQUE J. -C., RAMOS R. & NIETO S. [2013], *Decomposing the Rural-Urban Differential in Student Achievement in Colombia Using PISA Microdata*. University of Barcelona, Research Institute of Applied Economics, n° 201223.
- ESSER H. [2006], *Migration, language and integration*. WZB: Berlin.
- FORTIN N.-M. [2008], The Gender Wage Gap among Young Adults in the United States The Importance of Money versus People, *Journal of Human Resources*, vol. 43, n°4, pp. 884-918.
- GRIMES J.-E. & GRIMES, B. F. [1993], *Ethnologue: Language Family Index*. Summer Institute of Linguistics: Dallas.
- HARRIS J.-R. & TODARO M.-P. [1970], Migration, unemployment and development: a two-sector analysis, *The American Economic Review*, vol. 60, n°1, pp. 126-142.
- ISPHORDING I.-E. & OTTEN S. [2011], *Linguistic Distance and the Language Fluency of Immigrants*. Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen, n°0274.
- JANN B. [2008], A Stata implementation of the Blinder-Oaxaca decomposition, *Stata Journal*, vol. 8, n°4, pp. 453-479.
- JUHN C., MURPHY K.-M. & PIERCE B. [1993], Wage inequality and the rise in returns to skill, *Journal of political Economy*, pp. 410-442.
- LEWIS W.-A. [1954], Economic development with unlimited supplies of labour, *The manchester school*, vol. 22, n°2, pp. 139-191.
- LLERAS C. [2008], Do skills and behaviors in high school matter ? The contribution of noncognitive factors in explaining differences in educational attainment and earnings, *Social Science Research*, vol. 37, n°3, pp. 888-902.

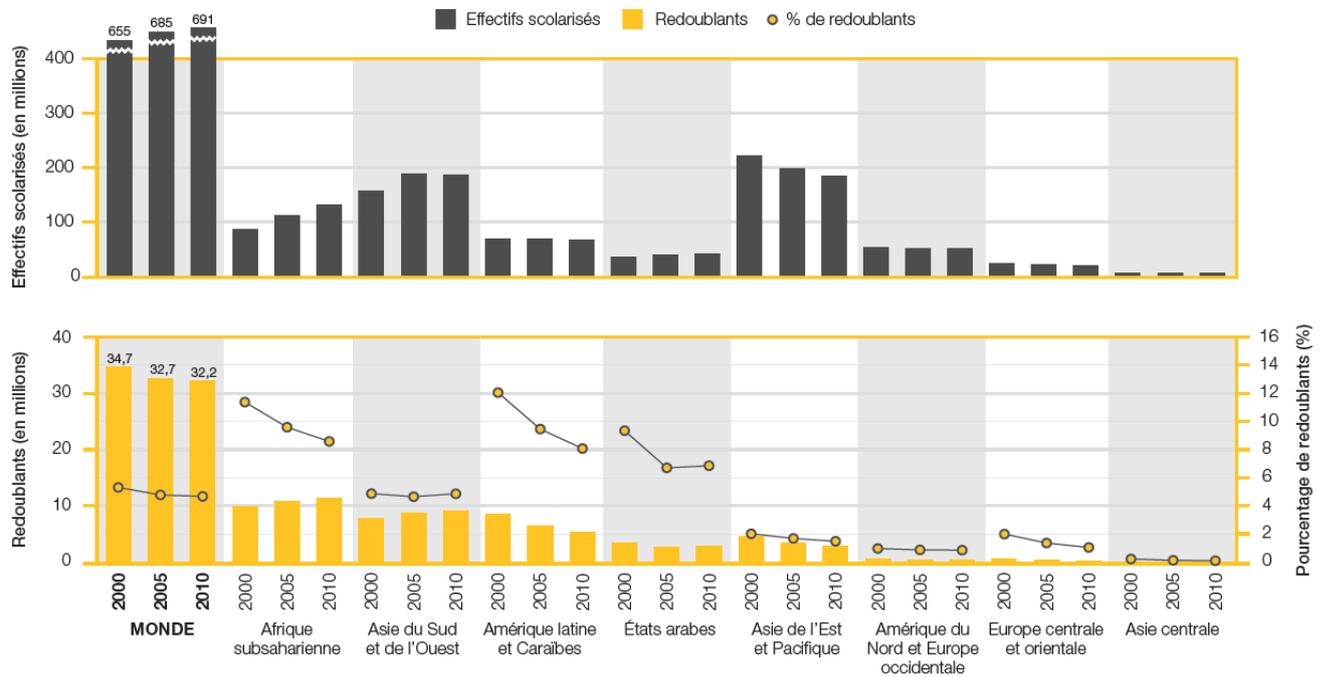
- MEUNIER M. [2011], Immigration and student achievement : Evidence from Switzerland, *Economics of Education Review*, vol. 30, n°1, pp. 16-38.
- MEURS D., PAILHÉ A. & SIMON P. [2005], Immigrés et enfants d'immigrés sur le marché du travail: une affaire de génération, *Histoires Familiales : Les Résultats de L'enquête Famille 1999*, pp. 461-82.
- MEURS D., PAILHÉ A. & SIMON P. [2007], Persistance des inégalités entre générations liées à l'immigration: l'accès à l'emploi des immigrés et de leurs descendants en France, *Population*, vol. 61, n°5, pp. 763-801.
- NEUMARK D. [1988], Employers' discriminatory behavior and the estimation of wage discrimination, *Journal of Human resources*, vol. 23, n°3, pp. 279-295.
- OAXACA R. [1973], Male-female wage differentials in urban labor markets, *International Economic Review*, vol. 14, n°3, pp. 693-709.
- PAN KÉ SHON J.-L. [2013], *Quarante ans de ségrégation ... et d'incorporation des immigrés en France, 1968-2007*. Rapport à l'Agence nationale pour la cohésion sociale et l'égalité des chances et au Plan Urbain Construction Architecture du ministère du Logement, 162 pages
- PAUL J.-J. & TRONCIN T. [2004], *Les apports de la recherche sur l'impact du redoublement comme moyen de traiter les difficultés scolaires au cours de la scolarité obligatoire*. Rapport établi à la demande du Haut conseil de l'évaluation de l'école. DEP/ bureau de l'édition, Paris.
- PLASSARD J.-M. & DESBARATS I. [2008], Étrangers et marché du travail français : droit et analyse économique en harmonie ? In : *Au-delà des droits économiques et des droits politiques, les droits sociaux ?*, éd. par Domin (J.-P.), Maric (M.), Delabruyère (S.) & Hédoïn (C.), XXVIIIe journées de l'Association d'Économie Sociale, L'Harmattan, vol. 2, pp. 9-21. – Paris, 2008.
- SAINT-PAUL G. & CAHUC P. [2009], *Immigration, qualifications et marché du travail*. La Documentation Française: Paris.
- SCHNEEWEIS N. [2011], Educational institutions and the integration of migrants, *Journal of Population Economics*, vol. 24, n°4, pp. 1281-1308.
- SHIELDS M.-A. & PRICE S.-W. [2002], The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas, *Journal of Population Economics*, vol. 15, n°1, pp. 137-160.
- SJAASTAD L.-A. [1962], The costs and returns of human migration, *The Journal of Political Economy*, vol. 70, n°5, pp. 80-93.
- SNOW M.-S. [1998], Economic, statistical, and linguistic factors affecting success on the test of English as a foreign language (TOEFL), *Information, Economics and Policy*, vol. 10, n°2, pp. 159-172.
- TANAY A. & AUDIRAC P.-A. [2000], Les salariés étrangers: loin des secteurs porteurs et des positions valorisées, *Premières Synthèses*, n°46.1.

WOLTON D. [2008], *Francophonie et migrations internationales*. Cellule de réflexion stratégique de la francophonie. www.francophonie.org

XI J. [2013], English Fluency of the US Immigrants: Assimilation Effects, Cohort Variations and Periodical Changes, *Social Science Research*, vol. 42, n°4, pp. 1092-1108.

ANNEXES DE L'ESSAI 3

ANNEXE A : REDOUBLANTS ET EFFECTIFS SCOLARISÉS DANS L'ENSEIGNEMENT PRIMAIRE PAR RÉGION - 2000, 2005 ET 2010.



Source : Base de données de l'Institut Statistiques de l'UNESCO.

ANNEXE B : SCORES MOYENS EN FRANÇAIS ÉCRIT ET ORAL SELON LES CARACTÉRISTIQUES INDIVIDUELLES ET L'ORIGINE MIGRATOIRE.

	<i>Français écrit</i>			<i>Français oral</i>		
	<i>ONL</i>	<i>OL</i>	<i>Ensb.</i>	<i>ONL</i>	<i>OL</i>	<i>Ensb.</i>
Genre						
Homme	78.4	88.1	85.4	70.7	81.2	78.2
Femme	80.6	88.2	86.3	67.1	79.8	76.7
Age						
Moins de 30 ans	83.4	89.9	88.1	76.2	80.6	79.43
Entre 30 et 50 ans	79.3	88.7	86.4	69.4	81.8	78.8
Plus de 50 ans	76.2	84.9	82.1	62.9	77.2	72.6
Enfant d'immigrés						
Oui	85.4	89.4	89.0	73.2	82.2	72.7
Non	77.8	85.4	81.7	68.0	76.9	81.3
Durée de scolarité théorique						
Entre 0 et 5 ans	63.1	76.7	72.5	48.9	67.3	61.7
Entre 5 et 10 ans	73.9	84.6	80.4	63.1	76.0	70.9
Entre 10 et 15 ans	83.3	89.8	88.4	73.3	82.9	80.8
Entre 15 et 20 ans	91.1	93.5	92.8	85.2	85.2	85.2
Nombre de redoublement(s) dans le cursus scolaire						
0	73.1	84.3	79.8	63.5	76.4	71.3
1	81.9	87.1	85.0	65.9	76.6	72.2
2	78.0	88.7	86.4	65.8	85.4	81.1
3	80.6	90.0	87.9	71.9	84.6	81.8
4	80.5	88.4	86.7	77.7	81.0	80.3
Plus de 4	84.2	88.6	87.7	72.4	77.7	76.7
Appartenance à une ZUS						
Oui	74.1	85.2	79.5	60.9	74.1	78.9
Non	81.0	88.4	86.7	71.6	81.1	67.3
Score de pauvreté						
0 - 5	81.9	88.8	87.3	72.6	81.7	79.8
6 - 10	77.1	86.0	82.7	66.1	76.2	72.4
11 - 15	74.2	85.3	79.8	63.1	77.4	70.3
16 - 20	57.0	78.4	63.0	26.4	62.3	36.3
Intensité d'usage de la langue française à la maison						
Seul le français est parlé	84.3	89.2	88.6	74.0	82.6	81.6
Le français est la langue principale	78.7	84.2	81.2	68.4	71.9	70.0
Le français est secondaire	76.0	76.2	76.0	66.5	60.6	65.0
Le français n'est pas du tout parlé	74.0	77.8	75.1	63.0	66.9	64.1
Fréquence de lecture de la mère quand l'enquêté avait entre 8 et 12 ans						
Jamais	77.2	86.4	82.9	66.4	76.0	72.3
De temps en temps	82.5	88.4	87.2	75.1	81.6	80.4
Régulièrement	77.1	89.3	87.3	63.4	83.9	80.6
Quotidiennement	85.7	89.6	88.8	76.8	82.8	81.6

ONL = Origine non-latine ; OL : Origine latine ; Ensb. : Ensemble
 Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

ANNEXE C : DÉCOMPOSITIONS JUHN, MURPHY ET PIERCE

ORIGINE LATINE - ORIGINE NON-LATINE				
<i>Français écrit</i>				
	Différence totale	Effet caractéristiques	Effet rendements	Effet résiduel
Moyenne	7,99	4,19	3,71	0,09
P05	28,2	8,17	7,67	12,36
P10	20,51	7,72	6,09	6,70
P25	10,25	4,85	5,01	0,39
P50	5,12	4,81	2,74	-2,43
P75	2,57	4,35	2,99	-4,77
P90	2,56	3,82	1,65	-2,91
P95	0	1,89	1,36	-3,26
<i>Français oral</i>				
Moyenne	10,66	6,58	3,95	0,13
P05	16,66	8,14	4,29	4,23
P10	16,66	7,62	3,01	6,03
P25	16,66	8,30	5,10	3,26
P50	16,66	8,80	4,76	3,10
P75	11,1	7,77	4,37	-1,04
P90	0	1,60	1,81	-3,41
P95	0	-1,50	3,04	-1,54

Source : enquête Information et Vie Quotidienne 2004, INSEE.

ESSAI 4

LA PRODUCTION DE COMPÉTENCES EN FRANÇAIS ÉCRIT PAR LES IMMIGRES DE PREMIÈRE GÉNÉRATION EN ÂGE DE TRAVAILLER

1. Introduction

Historiquement, économiquement, socialement et humainement, la France est un pays d'immigration. Elle s'est construite avec des vagues successives de personnes venues de l'étranger. La France métropolitaine comptait un million d'immigrés au début du 20^{ème} siècle, 2,3 millions en 1954, et 4.9 millions courant 2004. L'origine de ces immigrés, toujours plus nombreux, est de plus en plus variée. D'après l'INSEE, les immigrés arrivés avant 1974 sur le sol français et résidant aujourd'hui sur le territoire sont principalement d'origine européenne (57%) et maghrébine (34%). Seulement 4% et 5 % sont respectivement originaires d'Asie et d'Afrique Subsaharienne. Preuve que les flux migratoires se sont diversifiés parmi les immigrés arrivés en France après 1974 : la part des originaires d'Asie et d'Afrique Subsaharienne s'est fortement accrue avec 18% et 17% des immigrés respectivement.

Bien que les individus venus d'Afrique subsaharienne et du Maghreb parlent généralement français couramment en raison du passé colonial, ceci est loin d'être le cas pour ceux originaires d'Europe, d'Amérique, d'Océanie ou encore d'Asie. Selon

l'INSEE, parmi la très grande majorité des immigrés qui n'ont pas le français comme langue maternelle (84%), beaucoup estiment avoir des difficultés à lire un journal, à écrire une lettre ou à s'exprimer dans cette langue, sans pour autant être en situation d'illettrisme. Une bonne connaissance de la langue du pays d'accueil est pourtant essentielle pour l'intégration des primo-arrivants⁵⁵ (Alba et Nee, 2003 ; Bean et Stevens, 2003). Pourtant, beaucoup d'entre eux parlent peu, ou avec difficulté, la langue du pays vers lequel ils migrent (Portes et Rumbaut, 2006), ce qui peut nuire à leur carrière professionnelle, tant en termes d'emploi que de salaires (Chiswick et Miller, 1995 ; Dustmann et Fabbri, 2003 ; Berman et al., 2003). Il est compréhensible que des personnes dont le français n'est pas la langue maternelle, et venant souvent de milieux sociaux défavorisés dans leur pays d'origine, ne parlent pas ou peu le français en arrivant en France. Or, l'apprentissage du français pose la question des savoirs de base pour les plus défavorisés et introduit de fait un rapport inégalitaire avec les natifs. Ce rapport s'inscrit dans un contexte où la maîtrise imparfaite du français, surtout écrit, devient de plus en plus un handicap en raison des mutations technologiques et organisationnelles qui s'opèrent dans la société et dans le monde du travail. D'après les enquêtes Techniques et Organisations du Travail (TOTTO), réalisées en 1987 et en 1993, l'écrit a pris un poids considérable en milieu professionnel avec la multiplication des chartes qualité, des règlements intérieurs, des consignes de sécurité... La robotisation croissante ainsi que la dématérialisation du travail et le développement de l'informatique contribuent à l'augmentation de la contrainte linguistique dans le contenu du travail et nécessitent un véritable effort sur la compréhension et l'expression écrites.

De nombreuses recherches examinent les capacités linguistiques des immigrés, notamment en anglais, afin d'expliquer leurs rémunérations ou leur situation professionnelle (Akresh et Frank, 2011 ; Carnevale et al., 2001 ; Chiswick et Miller, 1995 ; Tainer, 1988), mais peu tentent en réalité d'analyser leurs déterminants. D'après les travaux empiriques existants, la capacité à apprendre la langue du pays d'accueil, vue comme une fonction de production de compétences, dépend principalement de caractéristiques individuelles (l'âge au moment de la migration, le pays d'origine, le niveau d'études, les raisons de la migration...). Néanmoins, à notre connaissance, aucune recherche ne fait vraiment la distinction entre ce qui relève des ressources dont disposent les primo-arrivants pour produire de la

⁵⁵ Nous désignons par « primo-arrivants » ou encore « immigrés de première génération », les individus qui sont nés à l'étranger et de parents étrangers.

compétence en français écrit et les facteurs qui peuvent entraver une utilisation efficace de ces ressources pour optimiser le processus productif. La revue de littérature proposée ici offre pourtant des éléments forts intéressants pour mener une analyse en ce sens. En effet, deux facteurs de production semblent se dégager des travaux empiriques : le niveau d'éducation et le temps de présence dans le pays d'accueil. En parallèle, le contexte dans lequel vit le primo-arrivant et les caractéristiques qui lui sont propres peuvent expliquer pourquoi il utilise plus ou moins efficacement ses ressources (durée d'éducation et temps de présence dans le pays d'accueil). Il s'agit donc ici de déterminer les facteurs susceptibles d'expliquer l'efficacité d'une production, à savoir le niveau de maîtrise en français écrit des immigrés de première génération en âge de travailler. Pour cela, nous utiliserons des scores individuels de mesure des compétences en français écrit, obtenus grâce à l'enquête française Information et Vie Quotidienne (2004).

L'intérêt de cet essai est triple. Tout d'abord, alors que la majorité des travaux de recherche porte sur la compréhension orale de la langue du pays d'accueil, cet essai s'intéresse plus spécifiquement à la compréhension écrite, tout aussi importante (Chiswick, 2008). De plus, à notre connaissance, aucune étude scientifique n'a été réalisée en ce sens auprès des adultes migrants de première génération venus s'installer en France. Enfin, la plupart des recherches sur les déterminants de la maîtrise de la langue du pays d'accueil a été menée à partir de données auto déclarées, avec un risque potentiel de biais, et non à partir de scores individuels issus de tests standardisés de mesure des compétences. L'originalité de cet essai réside par ailleurs dans l'utilisation d'un modèle de frontière stochastique de production afin d'expliquer les scores en français écrit des primo-arrivants, considérés comme producteurs de leurs propres compétences. En effet, d'un point de vue microéconomique, les compétences en français écrit peuvent être perçues comme un processus productif, dans lequel interviennent différents inputs individuels. Ces inputs détermineraient alors un output multidimensionnel qui prendrait la forme de satisfaction tirée des dites compétences produites⁵⁶. Dans ce contexte, comment les primo-arrivants allouent-ils leurs acquis scolaires, approximés par la durée d'éducation, et leur temps de présence en France pour produire des compétences en français écrit ?

⁵⁶ La satisfaction proviendrait à la fois de l'intégration sociale et culturelle (participation civique, reconnaissance sociale...) ainsi que de l'intégration professionnelle (accès au marché du travail et à l'emploi, conditions de travail, salaires...), les deux étant liées. Le degré d'intégration dépendra essentiellement du niveau de compétences atteint par l'individu lui-même.

Bien que de nombreux travaux se soient penchés sur la production d'éducation, peu ont mis en évidence le mécanisme même d'acquisition de compétences, c'est-à-dire le mécanisme qui permet à un individu de transformer des heures d'études et d'exposition à la langue du pays d'accueil en connaissance de cette dernière. Le recours à un modèle de frontière stochastique de production vise ainsi à mettre en lumière ce mécanisme d'apprentissage individuel. En effet, il présente l'avantage de tester l'effet de l'environnement et des caractéristiques propres aux primo-arrivants sur les distances individuelles par rapport à la frontière d'efficacité, c'est-à-dire par rapport au score maximum qu'il est possible d'atteindre en français écrit étant donnée sa dotation en facteurs de production. Une telle analyse doit éclairer le débat public autour des politiques éducatives/de formation à mettre en place à destination des immigrés de première génération, cela afin de favoriser leur apprentissage du français, apprentissage indispensable afin de s'insérer dans la vie active et dans la société.

2. Les facteurs de production de compétences

Si les fonctions de production en éducation ont pris de plus en plus de place dans la littérature économique, elles se restreignent bien souvent à analyser les déterminants des acquisitions scolaires sans pour autant faire la distinction entre facteurs de production et déterminants .

Un individu peut être considéré comme étant producteur de ses propres compétences. Il dispose d'un certain nombre de ressources, également appelées facteurs, dont l'utilisation efficace lui permettra d'optimiser son processus productif. Si les facteurs de production sont indispensables à la fabrication d'un stock donné de compétences, d'autres éléments, surtout contextuels, détermineront principalement l'efficacité d'utilisation des facteurs disponibles. En effet, ces éléments, par leur seule présence, peuvent ralentir ou accélérer la production de compétences en agissant comme des « *catalyseurs* ». En d'autres termes, ils influencent le stock d'une compétence, non parce qu'ils participent à sa fabrication proprement dite, mais parce qu'ils déterminent la capacité d'un individu à utiliser efficacement ses ressources.

Dans le cadre de cet essai qui s'inscrit dans cette approche, nous supposerons qu'il existe deux principaux facteurs de production de compétences en français écrit : la durée d'éducation et la durée de présence en France. La durée d'éducation⁵⁷ d'une part car au cours d'une longue scolarité, les individus ont étudié de façon approfondie la grammaire, le vocabulaire et la structure propre à leur langue maternelle. Or, une bonne connaissance de sa langue maternelle doit favoriser le transfert des compétences linguistiques d'une langue vers une autre (Chiswick, 1998). De plus, les individus ayant de hauts niveaux d'éducation ont une probabilité plus élevée d'avoir été formés à une langue étrangère pendant leur cursus, langue qui correspond bien souvent à celle du pays d'immigration (Chiswick et Miller, 1994 ; Wolton, 2008). En ce sens, la durée d'éducation apparaît donc comme étant un facteur de production. D'autre part, le temps de présence donne une indication sur les opportunités que les immigrés ont eues d'apprendre et de se perfectionner dans la langue du pays d'accueil. Ce temps correspond à du temps disponible pour les primo-arrivants afin d'acquérir leur nouvelle langue (Johnson et Newport, 1989 ; Birdsong, 2005). Dans ce cas, ce n'est pas tant le temps de présence dans le pays d'accueil qui compte que la capacité effective à utiliser cette durée de façon optimale pour produire de la compétence dans la langue du pays d'accueil. Il faut alors distinguer deux aspects : le nombre d'années de présence et l'intensité des liens créés avec la langue du pays d'accueil pendant ce nombre d'années (Chiswick et Miller, 2001), sachant que sans temps de présence dans le pays d'accueil, aucun lien n'aurait pu être créé⁵⁸. Cette distinction est alors fondamentale car elle permet de comprendre ce qui est du ressort des ressources en elles-mêmes d'une part, et ce qui est du ressort des déterminants de l'efficacité d'utilisation des facteurs d'autre part. Finalement, cela revient à répondre à la question suivante : dans quelle mesure les individus ont-ils utilisé la période depuis leur installation dans le pays d'accueil pour se confronter à la langue locale et ainsi optimiser la production de compétences dans cette langue ?

⁵⁷ Nous nous référons plus précisément à la durée théorique de formation initiale correspondant au plus haut niveau de diplôme obtenu (déclaré par l'individu), indépendamment du lieu d'obtention de ce dernier. Bien entendu, la durée d'éducation des primo-arrivants ne peut refléter la qualité de l'enseignement reçu, qualité supposée d'autant plus hétérogène que les individus ont été sous l'influence de systèmes éducatifs aucunement comparables. Ceci constitue une limite à notre recherche.

⁵⁸ Cela ne vaut pas pour les primo-arrivants originaires d'une zone géographique francophone qui ont pu se familiariser avec la langue française bien avant leur arrivée en France. A cet égard, le temps de présence dans un environnement francophone aurait été un « input » préférable, mais dans l'enquête utilisée nous ne disposons ni de cette variable, ni d'informations qui nous auraient permis de l'approximer.

Ainsi, la durée d'éducation initiale apparaît comme indispensable à la fabrication des compétences à l'écrit, voire même dans une langue étrangère écrite si celle-ci a été étudiée durant la scolarité, tandis que la durée de présence en France apparaît comme indispensable à la fabrication de compétences en français plus spécifiquement, en particulier pour les individus qui ne sont pas originaires d'un pays francophone. Ces deux facteurs semblent donc complémentaires dans le cadre d'une production spécifique de compétences en français écrit.

Par ailleurs, la durée d'éducation et la durée de présence en France peuvent être perçus comme des inputs temporels, qui, selon Becker (1965), sont à prendre en compte au même titre que les inputs matériels dans la fonction de production de capital humain. Les compétences ne peuvent effectivement être considérées comme un stock de savoirs inerte qui n'évoluerait pas avec le temps. Plus un individu octroie du temps à des activités susceptibles de développer ses compétences, plus celles-ci seront acquises et maîtrisées. Toutefois, le temps étant un input rare dont les individus ne disposent pas en grande quantité, il doit faire l'objet d'une allocation optimale entre ses différents usages.

Si l'on cherche, plus spécifiquement, à appliquer l'approche de Becker à la production de compétences en français écrit par les primo-arrivants, l'immigré de première génération est alors considéré comme un producteur, disposant d'un certain nombre de ressources (en termes de temps et de savoirs déjà acquis), et dont l'utilisation efficace lui permettrait de progresser dans son acquisition du français écrit. Toutes choses égales par ailleurs, un primo-arrivant combinerait alors un ensemble de savoirs, principalement acquis durant sa scolarité, ainsi que du temps pour produire de la compétence en français écrit.

Certains primo-arrivants auraient des productivités plus faibles que d'autres en termes de production de compétences en français écrit. Ces différences de productivités peuvent s'expliquer en partie par les choix d'allocation des ressources, c'est-à-dire du temps affecté à l'acquisition du français écrit, et par une efficacité différente dans l'utilisation des ressources disponibles. On peut alors faire l'hypothèse que le niveau de compétences en français écrit (P) est conditionné par le temps passé en France (T), par la durée d'éducation (S) mais aussi et surtout par la capacité individuelle à produire efficacement de la compétence (E), d'où:

$$P = P(T, S, E)$$

où $P_T > 0$, $P_S > 0$ et $P_E > 0$.

L'intérêt ici est donc d'expliquer pourquoi certains primo-arrivants arrivent à avoir de meilleures compétences en français écrit que d'autres, alors qu'ils ont consacré moins de temps à leur apprentissage. L'explication peut résider dans le fait que ces mêmes primo-arrivants ont été plus efficaces dans l'utilisation de leurs « *ressources-temps* », pour acquérir le français écrit. Toutefois, cette situation correspond à un cas particulier, puisqu'en général, le choix optimal de la durée d'éducation et du temps consacré à l'apprentissage du français lors de sa présence en France dépend à la fois de sa capacité à convertir efficacement ce temps en compétences ainsi que de ses propres préférences.

3. Les déterminants des performances productives : hypothèses

Un certain nombre de variables peuvent expliquer que les immigrés de première génération acquièrent moins facilement que d'autres la langue du pays d'accueil, malgré un niveau d'éducation ou un temps de présence en France similaire. Ces variables, essentiellement de contexte et de caractéristiques propres aux immigrés, semblent conditionner leur capacité à convertir le temps d'éducation et de présence en France en compétences en français écrit.

3.1. L'âge au moment de la migration

L'âge durant lequel le processus d'émigration du pays natal s'est effectué a également son importance (Jasso et Rosenzweig, 1990 ; Stevens, 1994) en raison de la plasticité cérébrale⁵⁹ qui se perd avec l'âge (Long, 1990). En termes plus économiques, cela rejoint le concept d'obsolescence technique des compétences (De Grip et Van Loo, 2002). Cette obsolescence correspond principalement à la

⁵⁹ Désigne la capacité du système nerveux central à modifier sa structure et sa fonction de réponse à des contraintes.

dégradation des compétences provoquée par les effets de l'âge⁶⁰. De nombreux auteurs ont ainsi montré que la capacité à apprendre du vocabulaire, à comprendre des règles de grammaire et de syntaxe, ne se perd pas uniquement au passage de l'enfance à l'âge adulte : elle se perd tout au long de la vie avec les effets de l'âge (Veltman, 1988 ; Service et Craik, 1993). Si l'implantation dans le nouveau pays s'est faite jeune ou très jeune, l'individu pourra également suivre tout ou partie de sa scolarité dans le pays d'accueil : dans ces conditions, l'acquisition de la langue sera particulièrement favorisée. Espenshade et Fu (1997) ont d'ailleurs montré qu'une année d'éducation supplémentaire aux États-Unis en tant que pays d'accueil, avait un effet beaucoup plus important sur l'acquisition de l'anglais par les immigrés qu'une année d'éducation supplémentaire dans leur pays d'origine. Enfin, l'âge au moment de la migration est particulièrement intéressant car il marque le début d'une immersion dans cette langue étrangère (Birdsong et Molis, 2001). Ainsi, nous supposons que plus un individu immigré jeune, plus il sera efficace dans son acquisition du français écrit pour deux raisons : une meilleure capacité d'adaptation (pas d'obsolescence technique des compétences) et une scolarisation tout ou partie dans la langue du pays d'accueil qui démultiplie les effets producteurs de l'éducation

3.2. La langue maternelle, la langue parlée à la maison, les origines sociales et les inégalités territoriales

De même, si l'on se réfère aux études linguistiques et aux preuves empiriques détaillées notamment dans l'essai 3, il semble plus délicat d'apprendre une langue étrangère dont la structure est très éloignée de sa propre langue maternelle. En revanche, un usage fréquent de la langue du pays d'accueil à la maison peut rendre plus efficace son acquisition malgré sa distance linguistique avec la langue maternelle de l'individu.

Par ailleurs, on ne peut oublier l'influence du facteur social dans les pratiques linguistiques. Les travaux de Bernstein (1975) sont très intéressants à ce sujet car ils mettent bien en lumière le lien entre l'appartenance sociale et les pratiques

⁶⁰ Ce qui rejoint le modèle de Mincer (1974) augmenté d'un freinage, par la prise en compte du coefficient négatif sur le terme quadratique de l'expérience.

linguistiques. Par exemple, il est à noter des différences importantes entre les enfants de classe supérieure et les enfants de classe populaire en ce qui concerne l'aptitude à changer de grammaire et de vocabulaire en fonction du contexte (Hawkins, 1969).

D'un point de vue plus économique, cela fait référence au capital humain intergénérationnel, qui a fait l'objet d'un certain nombre de travaux. Par exemple, dans le modèle de Becker et Tomes (1986), les parents peuvent impacter les résultats économiques de leurs enfants en investissant dans le capital humain de ces derniers. Toutefois, compte tenu des imperfections du marché, certains parents, les plus défavorisés, souvent les moins éduqués, ne sont pas toujours en mesure d'investir au mieux dans l'éducation de leur progéniture, ce qui fait perdurer les inégalités à travers les générations. Selon l'approche traditionnelle du capital humain, la principale explication à la persistance intergénérationnelle des inégalités de revenus concerne les contraintes de crédits qui peuvent entraver ou rationner le financement de l'investissement en capital humain. Nous émettons donc comme autre hypothèse qu'un immigré dont le capital humain des parents est élevé (approximé par le niveau de diplôme par exemple) sera plus performant dans son acquisition d'une langue étrangère. Des parents plus éduqués sont supposés avoir de meilleurs revenus pour pouvoir investir dans le capital humain de leurs enfants⁶¹.

Par ailleurs, parmi les variables de contexte, le fait d'habiter dans des zones à forte concentration de populations immigrées affecte négativement les incitations à utiliser son niveau d'études ou son temps de présence pour acquérir le plus de compétences possible dans la langue du pays d'accueil. Les individus résidant et/ou travaillant au sein de ces zones ont effectivement moins besoin d'une bonne maîtrise de la langue du pays d'accueil puisqu'ils ont moins de contacts avec les natifs. Florax et al. (2005) ont démontré, à partir de données hollandaises, que les immigrés résidant dans des zones urbaines à forte densité de populations immigrées avaient moins d'incitations à investir dans l'apprentissage de la langue du pays d'accueil que ceux vivant dans des zones à plus faible densité. Cela rejoint l'idée de Lazear (1999) selon laquelle les migrants auraient tendance à apprendre la langue du pays

⁶¹ Dans cet essai, nous appréhenderons le niveau de capital humain des parents par le plus haut niveau de diplôme obtenu par la mère. Le plus haut diplôme obtenu par le père a également été ajouté dans le modèle, or, cette variable à l'exemple de l'essentiel de la littérature s'est révélée non significative et rendait instable les coefficients associés au plus haut diplôme obtenu par la mère lorsque les deux variables étaient introduites ensemble, là aussi du fait d'une colinéarité bien identifiée dans la littérature.

d'accueil plus rapidement si leur communauté linguistique est restreinte. A contrario, un migrant qui résiderait au sein d'une communauté linguistique assez répandue accorderait moins d'importance à l'apprentissage de la langue du pays d'accueil, d'où l'intérêt des politiques publiques de lutter contre les inégalités territoriales et ainsi favoriser la diversification des zones de résidence, d'autant que les migrants ont tendance à se rassembler là où ils sont bien représentés (Chiswick et Miller, 1994 ; Wolton, 2008). Selon Waldorf et al. (2010), le degré d'acquisition de la langue du pays d'accueil est également plus grand pour les migrants résidant dans des zones à faible densité de population étrangère et ce degré d'acquisition augmenterait à mesure que le niveau d'éducation s'élève. De fait, nous nous attendons en dernier lieu à ce que vivre en Zone Urbaine Sensible (ZUS) affecte négativement l'efficacité d'acquisition du français écrit, compte tenu de la forte concentration de populations d'origine immigrée dans ce type de quartiers et par conséquent, de la faible incitation à mettre à profit ses acquis pour utiliser cette langue.

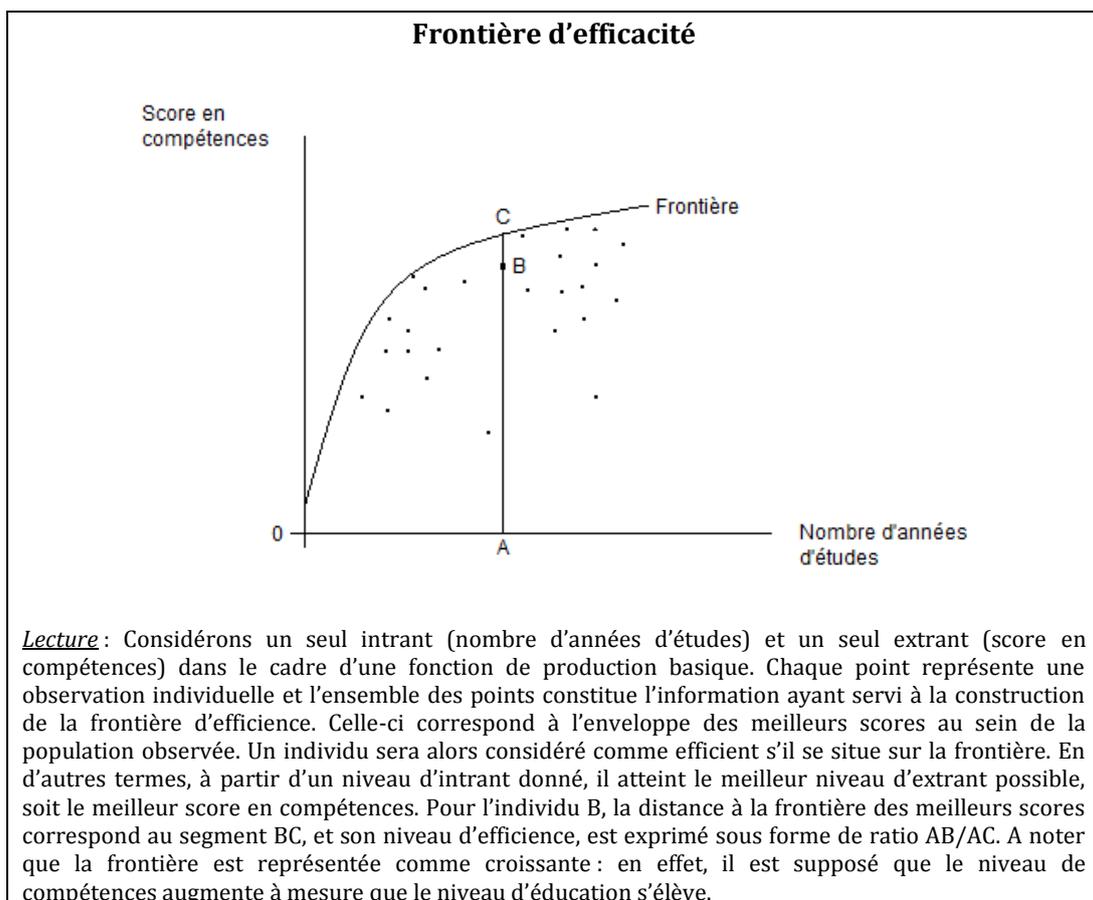
4. Spécification du modèle de frontière stochastique de production

4.1. Concepts généraux

Une fonction de production peut être représentée par une isoquante qui exprime la frontière de production, c'est-à-dire l'extrait maximal qu'il est possible d'atteindre, compte tenu du niveau des intrants. Ce type de fonction est souvent utilisé afin d'étudier la production d'un bien ou d'un service par une entité quelconque, telle une firme. Toutefois, les concepts généraux d'une fonction de production sont tout à fait transposables aux compétences. Les individus peuvent très bien être perçus comme des producteurs de compétences, dont la quantité est exprimée sous forme de scores (Hanushek, 1979).

L'intérêt de cet essai est alors d'estimer le score maximum en français écrit que les primo-arrivants peuvent atteindre étant donnée leur dotation en facteurs (durée d'éducation et temps de présence en France) : cela correspond communément à la

performance productive ou encore l'efficacité technique⁶². Cette notion a été introduite par Farrell, en 1957, à la suite des travaux de Debreu (1951) et Koopmans (1951). L'idée de Farrell est d'estimer l'efficacité technique par une frontière mesurant les performances des entreprises les unes par rapport aux autres. Ce concept peut-être illustré de façon très simple dans le cadre de notre objet de recherche (cf. encadré ci-dessous). L'efficacité technique représenterait ici l'aptitude d'un individu à combiner de façon optimale ses facteurs de production (la durée de scolarité et le temps de présence en France), indépendamment de leur prix, afin de produire au maximum des compétences en français écrit. L'avantage de ce type de modèle est qu'il permet ainsi de prendre en compte des éléments qui ne rentrent pas directement dans la fonction de production, mais qui peuvent expliquer pourquoi certains immigrants sont plus ou moins éloignés de la frontière de production.



⁶² Bien que la notion d'efficacité soit proche de celle d'efficacité, elle s'en distingue. L'efficacité fait référence au niveau de réalisation d'un but désiré. A l'inverse, l'efficacité désigne le rapport entre ce qui est réalisé et les ressources mobilisées pour y parvenir.

4.2. Modèle de frontière stochastique de production

Afin d'estimer l'inefficience technique par le biais d'un modèle de frontière, nous optons pour une approche paramétrique. Parmi les approches paramétriques, nous avons le choix entre une frontière déterministe ou une frontière stochastique. Nous retenons la seconde sachant que la première risque de surestimer les niveaux d'inefficience technique (Amara et Romain, 2000). En effet, le recours à des frontières stochastiques permet d'isoler le terme d'erreur purement aléatoire du terme d'inefficience technique et doit ainsi fournir une mesure plus précise de l'efficacité technique. La frontière paramétrique stochastique peut s'écrire de la façon suivante :

$$\ln Y_i = f(X_i, \beta) + \varepsilon_i \quad (1)$$

Où i, \dots, N est un indicateur de l'individu considéré, en tant que producteur de compétences.

Et avec :

Y_i , le score obtenu par l'individu i au test de français écrit.

X_i , le vecteur de facteurs de production associés au i -ème individu.

β , le vecteur de paramètres inconnus à estimer

ε_i , le terme d'erreur composé de la forme $\varepsilon_i = V_i - U_i$

V_i est un terme d'erreurs aléatoires supposé suivre une distribution normale $N(0; \sigma_V^2)$. Il permet de capter les erreurs de mesure sur les variables, les effets des chocs aléatoires non contrôlables par les producteurs (conditions économiques, apparition d'une maladie, d'un handicap etc.), ainsi que les erreurs d'approximation de la compétence par la relation fonctionnelle choisie.

U_i , un terme d'erreurs indépendant de V_i , suivant une loi de distribution semi-normale $N(0; \sigma_U^2)$. Dans la littérature consacrée aux frontières d'efficience, le terme U_i correspond à la distance par rapport au score maximum, distance représentée par la frontière stochastique. Le terme U_i représente donc une mesure de l'inefficience technique par opposition à V_i , l'erreur stochastique.

Le modèle de frontière stochastique de production permettant d'isoler les facteurs susceptibles d'influencer l'inefficience technique, le terme U_i est donc défini de la manière suivante :

$$U_i = \delta_0 + Z_i \delta + \omega_i \quad (2)$$

Z_i sont les variables explicatives de l'inefficience technique des producteurs, δ est un vecteur de paramètres inconnus qu'il faudra estimer et ω_i est un terme résiduel. Dans le cadre d'une approche par compétence, ω_i reflèterait l'efficience technique purgée de l'influence des facteurs exogènes.

4.3. L'indice d'efficience technique

Alors que nous avons identifié comme principaux facteurs de production la durée d'éducation et le temps de présence en France, d'autres éléments semblent davantage jouer sur l'utilisation efficace de ces facteurs que sur la production de compétences en français écrit. Ces éléments font référence aux variables de contexte et aux caractéristiques des primo-arrivants mentionnées ci-avant. Au regard des données disponibles dans l'enquête utilisée, nous décidons d'introduire comme variables Z_i , les variables suivantes : la langue maternelle, le fait d'être arrivé avant la fin de l'âge de scolarité obligatoire, le fait que la langue d'apprentissage de la lecture soit différente de la langue maternelle, le diplôme de la mère, le fait de résider en ZUS ou encore la fréquentation du monde extérieur⁶³.

L'indice d'efficience technique du i -ème individu est alors défini par le ratio d'efficacité suivant :

⁶³ Le descriptif de ces variables est donné en annexes (annexes A et B).

$$TE_i = \frac{Y_i \text{ réalisé}}{Y_i \text{ max}} = \frac{f(X_i, \beta) \exp(V_i - U_i)}{f(X_i, \beta) \exp(V_i)} = \exp(-U_i) = \exp(-\delta_0 - Z_i \delta - \omega_i) \quad (3)$$

L'estimation de la fonction de production de compétences en français écrit se fera par la fonction translogarithmique dont la flexibilité et le calcul d'élasticités de substitution sont les principaux avantages. Cette fonction correspond à une approximation de second ordre de la fonction sous-jacente basée sur deux facteurs de production. Ainsi, en considérant le temps de présence ($\ln x_1$) et la durée d'éducation ($\ln x_2$) comme facteurs de production, la fonction Translog s'écrit de la façon suivante :

$$\ln Y_i = \beta_0 + \beta_1 \ln x_{1,i} + \beta_2 \ln x_{2,i} + 0.5\beta_3 (\ln x_{1,i})^2 + 0.5\beta_4 (\ln x_{2,i})^2 + \beta_5 (\ln x_{1,i})(\ln x_{2,i}) + (v_i - u_i) \quad (4)$$

où $\beta_k = 0, 1, \dots, 5$ sont les paramètres de la frontière à estimer.

En incluant un terme quadratique pour chaque facteur de production, la fonction Translog permet d'estimer s'il existe des effets non linéaires des durées d'éducation et de présence. Ceci est d'autant plus intéressant que d'après la revue de littérature, les effets du temps de présence, bien qu'encore présents, ont tendance à s'essouffler avec le temps. Par ailleurs, le terme croisé entre les deux facteurs de production doit permettre de dire si oui ou non il existe une substitution possible entre les deux facteurs de production. En d'autres termes, un faible temps de présence sur le territoire d'accueil peut-il être compensé par une durée plus importante d'éducation et inversement ?

D'autres méthodes auraient pu être utilisées, comme la méthode de l'enveloppe (DEA⁶⁴). Or, contrairement aux approches déterministes, dont fait partie la méthode de l'enveloppe, le modèle de frontière stochastique n'attribue pas tout l'écart à la frontière à de l'inefficience technique, ce qui serait trop restrictif. L'utilisation d'une approche stochastique permet d'attribuer la variance observée du modèle à la fois à la capacité des primo-arrivants à convertir leurs ressources en compétences mais

⁶⁴ Data Envelopment Analysis

aussi à l'hétérogénéité aléatoire entre les individus. En d'autres termes, l'utilisation d'un modèle stochastique plutôt que déterministe se justifie au regard de l'hétérogénéité non-observée qui peut parfois être bien plus grande comparée à l'hétérogénéité observée. Les résultats de l'estimation de la frontière stochastique de production seront présentés dans la section 5.

5. Statistiques descriptives

Les données exploitées sont toujours issues de l'enquête Information et Vie Quotidienne réalisée en France en 2004 (IVQ) par l'INSEE. En mettant en relation les scores individuels obtenus avec la durée d'éducation et de présence en France, nous observons que les scores en français écrit⁶⁵ augmentent conjointement avec les années d'éducation et avec le temps de présence en France (cf. tableau 1).

Tableau 1 : Scores moyens en français écrit selon le nombre d'années d'éducation et le nombre d'années de présence en France

Catégories	N	Français écrit
Année d'éducation		
0-5 ans	156	69.09
6-9 ans	147	76.48
10-11 ans	191	84.88
12-14 ans	167	86.49
15-20 ans	121	90.16
Temps de présence		
0-5 ans	74	75.05
6-10 ans	59	74.95
11-15 ans	78	77.98
16-20 ans	71	81.01
21-25 ans	83	84.18
26-30 ans	87	84.82
31-35 ans	111	78.10
36-40 ans	80	83.78
41-45 ans	90	85.67
Plus de 45 ans	49	86.94
ENSEMBLE	782	81.44

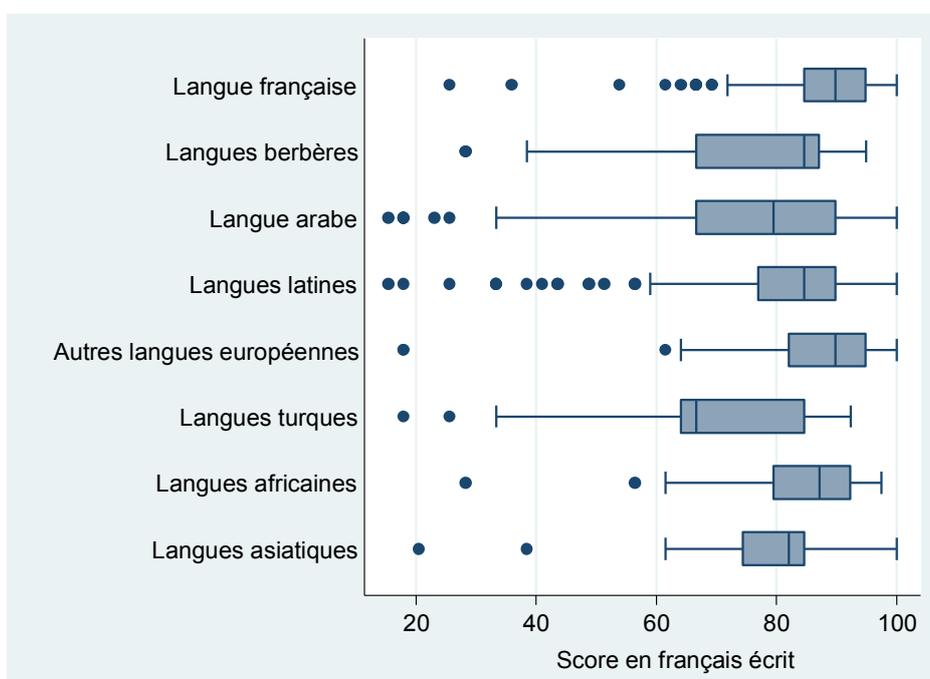
Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

⁶⁵ Cf. Annexes, Annexe B « Histogramme et densité du score en français écrit chez les migrants de première génération en âge de travailler »

Bien que ces tendances ne s'appuient que sur des données en coupe transversale, elles illustrent assez convenablement le rôle joué par le niveau d'éducation et le temps de présence dans le pays d'accueil sur le degré d'acquisition du français écrit.

De plus, d'après la figure 1, les primo-arrivants ayant une langue maternelle assez éloignée du français, comme les langues turques, berbères et arabe, ont des scores en français écrit beaucoup plus dispersés que les autres, ce qui va dans le sens de notre hypothèse de départ. La distance linguistique entre la langue du pays d'accueil et celle du pays d'origine semble en effet expliquer en partie qu'il soit plus difficile pour certains migrants d'acquérir la langue française que pour d'autres.

Figure 1: Distribution des scores en français écrit selon la langue maternelle



Source : enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

Les langues latines (espagnol, portugais, italien) et les autres langues européennes (allemand, anglais, langues slaves occidentales...) étant proches du français, les migrants ayant l'une ou l'autre de ces langues maternelles ont à la fois des scores plus élevés et moins dispersés que les autres migrants (cf. figure 1). Toutefois, les scores des primo-arrivants ayant pour langue maternelle une langue africaine sont assez peu dispersés et assez élevés. L'explication la plus plausible réside certainement dans l'héritage du passé colonial des pays africains. Idem pour les primo-arrivants asiatiques, bien que leurs scores moyens soient plus faibles.

6. Résultats

6.1. Paramètres de la frontière et de l'inefficience technique

Les résultats de l'estimation sont donnés dans le tableau 2⁶⁶. Tout d'abord, concernant l'estimation de la frontière en elle-même, notons que le temps de présence et de scolarité ont un impact positif et significatif sur l'estimation du score maximum qu'un individu est censé pouvoir atteindre étant donné sa dotation en facteurs. Autrement dit, plus un immigré a fait d'études et/ou plus il réside sur le territoire français depuis longtemps, plus le score maximum qu'il sera en mesure d'atteindre en français écrit sera élevé. Le coefficient négatif associé au terme croisé suppose par ailleurs que le temps de présence en France est d'autant plus important que le niveau d'éducation est faible, ce qui traduit effectivement un effet significatif de substitution. De plus, il semblerait que 18% de la variance totale de la production de compétences en français écrit soit due à de l'inefficience technique, c'est-à-dire à une utilisation sous-optimale des facteurs de production. L'étude des facteurs pouvant expliquer cette inefficience technique s'en trouve ainsi justifiée.

⁶⁶ Le modèle de frontière stochastique de production avec les effets d'inefficience technique a ici été estimé à l'aide de la commande « *frontier* » du logiciel Stata.

Tableau 2 : Résultats de l'analyse en frontière stochastique de production dans le cas d'une fonction Translog

VARIABLES		Coefficients	Z	
Paramètres de la frontière stochastique				
Nombre d'années de présence en France ($\ln x_1$)		0,050**	1,96	
Nombre d'années d'études ($\ln x_2$)		0,067**	2,21	
$(\ln x_1)^2$		-0,003	-0,94	
$(\ln x_2)^2$		0,006	0,94	
$(\ln x_1)(\ln x_2)$		-0,012*	-1,65	
<i>Sigma_u</i>		0,019		
<i>Sigma_v</i>		0,35		
λ		17,99		
Paramètres de l'inefficience technique				O.R.
Arrivée avant l'âge de scolarité obligatoire en France		-0,62***	-4,53	0,54
La langue d'apprentissage de la lecture est la langue maternelle		0,96***	6,91	2,62
Le lieu de résidence se trouve en ZUS		0,45***	3,54	1,58
Langue maternelle (réf: le français)	Langues berbères	1,99***	6,72	7,29
	Langue arabe	1,40***	7,13	4,05
	Langues latines	0,67***	3,70	1,95
	Autres langues européennes	0,55**	2,16	1,73
	Langues turques	1,40***	4,44	4,06
	Dialectes africains	0,43*	1,72	1,53
	Langues asiatiques	1,51***	5,33	4,50
Fréquentation de la famille (réf: souvent)	Occasionnellement	-0,18	-1,39	0,83
	Jamais	-0,02	0,09	1,02
Fréquentation des amis (réf: souvent)	Occasionnellement	0,45***	3,53	1,57
	Jamais	-0,02	-0,09	0,98
Niveau de diplôme de la mère (réf: jamais scolarisée ou sans diplôme)	Diplôme de niveau primaire	-0,34*	-1,82	0,71
	Diplôme du secondaire	-0,70***	-3,83	0,50
	Diplôme du supérieur	-1,46***	-5,04	0,23
Intensité d'usage de la langue française à la maison (réf: seul le français est parlé)	Le français est la langue principale parlée	0,40***	2,90	1,50
	Le français est une langue secondaire	0,83***	4,14	2,28
	Le français n'est pas parlé	1,48***	7,26	4,39
Constante		-4,36	-18,4	

*** : significatif au seuil d'1% ; ** : significatif au seuil de 5% ; * : significatif au seuil de 10%

Champ : 782 primo-arrivants en France en âge de travailler en 2004

A noter, en termes d'interprétation des paramètres de l'inefficacité technique, qu'un coefficient positif traduit une mobilisation inefficace des facteurs de production. Cela suggère un éloignement de la frontière de production, donc du score maximal. A l'inverse, un coefficient négatif supposerait une utilisation efficace des facteurs de production et donc un rapprochement du score maximal.

L'étude de l'inefficacité est intéressante et semble dans l'ensemble confirmer nos hypothèses. Être arrivé avant la fin de l'âge de scolarité obligatoire rend incontestablement un immigré de première génération plus efficace dans l'utilisation de ses facteurs de production. Cela rejoint l'idée que plus un individu immigré jeune, plus il pourra acquérir facilement la langue du pays d'accueil (Long, 1990 ; Service et Craik, 1993). Il n'est effectivement pas soumis au phénomène d'obsolescence des compétences qui pourrait nuire à une utilisation efficace des facteurs de production. De plus, les primo-arrivants qui sont arrivés jeunes en France sont plus à même d'utiliser efficacement la durée de scolarité dans la production de compétences en français écrit, cela puisqu'une partie de leurs études s'est elle-même déroulée en français.

Globalement, lorsque la langue maternelle n'est pas le français, les individus s'éloignent du score maximal qu'ils sont censés pouvoir atteindre. Plus spécifiquement, le fait d'avoir comme langue maternelle une langue berbère, arabe, turque ou encore asiatique éloigne particulièrement de la frontière de production. Un primo-arrivant ayant pour langue maternelle le berbère a 7,29 fois plus de risque d'être plus inefficace dans son utilisation des facteurs de production pour produire de la compétence en français. Ce risque s'élève à 4,05 pour les immigrés arabes, 4,06 pour les immigrés turcs et 4,50 pour les immigrés d'origine asiatique. Cela s'explique certainement par des différences linguistiques profondes tant structurellement qu'historiquement. Les langues berbères, arabe, turques ou asiatiques s'appuient historiquement sur des alphabets très différents de l'alphabet latin sur lequel repose le français. Il semblerait toutefois que l'apprentissage de la lecture dans une autre langue que la langue maternelle soit un atout pour l'utilisation efficace de ces facteurs de production. Les individus étant amenés à s'adapter à une autre langue très jeune, ils sont plus enclins à apprendre une langue étrangère, surtout si la langue d'apprentissage de la lecture est finalement la même que celle du pays d'immigration.

De plus, fréquenter de façon modérée sa famille et fréquenter très régulièrement ses amis semble positif pour une acquisition efficace du français écrit. En résumé, avoir une vie sociale active favorise l'apprentissage du français écrit⁶⁷. A l'inverse, se renfermer chez soi et rester dans son cercle familial paraît plus pénalisant. Un immigré qui vit en ZUS risque également de se retrouver au sein d'enclaves ethniques⁶⁸, ce qui l'éloignera potentiellement du score maximum qu'il lui est possible d'atteindre en français écrit, compte tenu de son niveau d'éducation et du temps passé en France. Cela rejoint en partie les résultats sur la sphère familiale et la littérature concernant l'impact des enclaves ethniques sur l'acquisition de la langue du pays d'accueil. Toutefois, au regard des rapports de chance, le fait de vivre en ZUS est un facteur bien moins important que la langue maternelle par exemple.

Enfin, plus la mère du migrant est diplômée, plus ce dernier est doté de capacités propices à utiliser efficacement ses facteurs de production pour s'approcher de son score d'efficacité maximum. Il en va de même pour l'intensité d'usage du français à la maison : plus le français y est parlé, plus le migrant se rapproche de la frontière de production. Notons que ne jamais parler le français à la maison est tout aussi pénalisant que d'avoir comme langue maternelle l'arabe, le turc ou encore une langue asiatique. Un migrant qui aurait à la fois une forte distance linguistique entre sa langue maternelle et la langue française et qui ne parlerait pas du tout cette dernière à la maison perd ainsi grandement en efficacité d'apprentissage du français, ce qui peut expliquer les inégalités de scores entre les primo-arrivants.

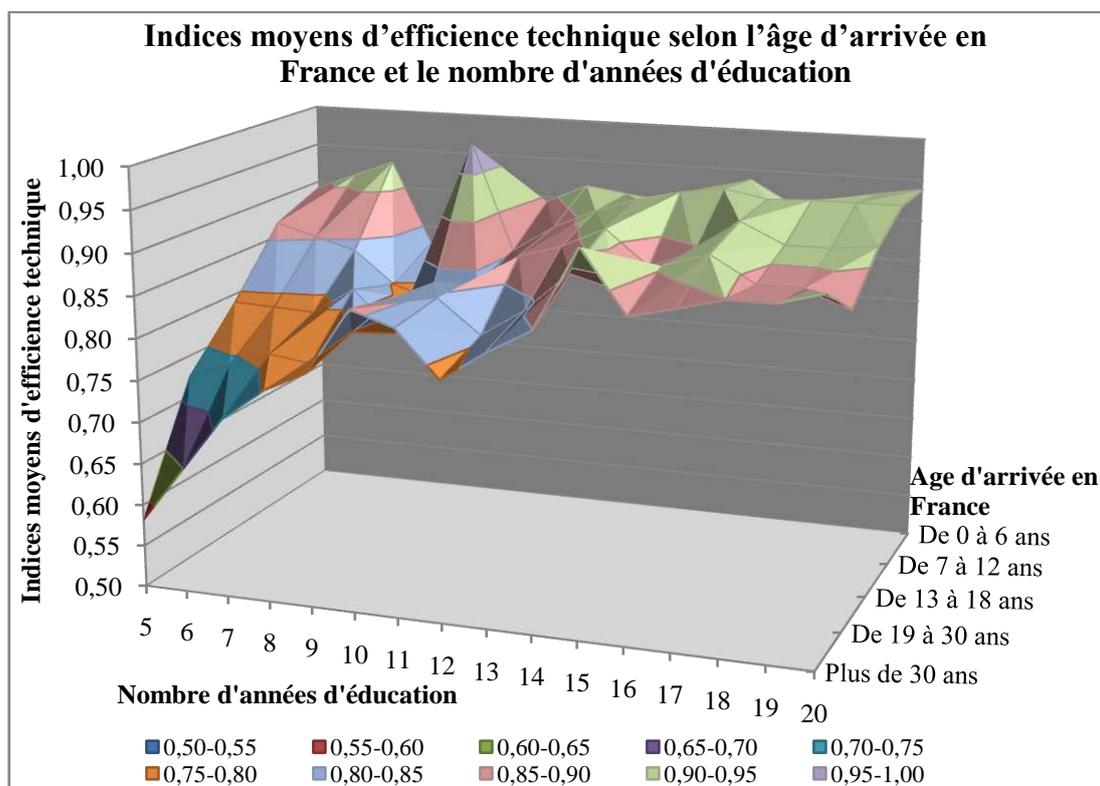
6.2. Indices d'efficience technique

Par ailleurs, suite à l'estimation de la frontière de production, nous avons pu dégager un indice d'efficience technique, autrement dit, un indice de performance productive. La figure suivante met en évidence l'augmentation progressive de la performance productive avec l'âge d'arrivée et le niveau d'éducation.

⁶⁷ La fréquentation des amis est supposée être plus efficace si les amis d'un primo-arrivant ne parlent pas sa langue d'origine.

⁶⁸ Cette notion d'enclave ethnique est toutefois beaucoup plus forte aux États-Unis par exemple qu'en France.

Figure 2: Indices d'efficiency technique moyens selon l'âge d'arrivée en France et le nombre d'années d'éducation⁶⁹.



Source: calculs de l'auteur

Quel que soit le niveau de diplôme, plus un individu arrive tardivement en France, plus son indice d'efficiency technique diminue. Ceci s'explique par l'obsolescence technique des compétences avec l'âge. Néanmoins, les primo-arrivants qui sont arrivés tardivement en France peuvent contrecarrer cet effet grâce notamment à un niveau d'éducation élevé, comme en témoigne la figure précédente. Ceux qui sont arrivés après 30 ans et qui ont une scolarité courte (moins de sept années d'éducation) ont des indices moyens d'efficiency technique inférieurs à 0.70. A l'inverse, ceux qui ont une scolarité longue (plus de 14 années d'éducation) ont des indices moyens supérieurs à 0.80. De même, les primo-arrivants qui sont arrivés très jeunes en France (avant 6 ans), mais qui ont un faible niveau d'éducation, ont des indices moyens d'efficiency technique presque tout aussi élevés que les individus ayant un haut niveau d'éducation.

⁶⁹ Rares étant les individus de l'échantillon à n'avoir qu'entre 1 et 4 années d'éducation, l'échelle du nombre d'années d'éducation commence dès 5 années.

6.3. Robustesse des résultats et élasticités

Afin de vérifier que le modèle économétrique est bien spécifié, trois hypothèses principales ont été testées. La première vise à savoir si la fonction de production Translog est bien appropriée. En d'autres termes, cela revient tester s'il y a bien un écart à la frontière dû à de l'inefficience technique ou non. Dans le cas où il n'y aurait pas d'inefficience technique $H_0 : \sigma_u = 0$, l'estimation pourrait alors se réduire à une estimation par les Moindres Carrés Ordinaires (MCO). La deuxième hypothèse testée a, quant à elle, pour objectif de savoir si la spécification Cobb-Douglas est préférable à la spécification Translog ou non. Si $H_0 : \beta_{ik} = 0$ est vérifiée alors cela suppose qu'une fonction Cobb-Douglas est préférable à une fonction translogarithmique. Enfin, la dernière hypothèse testée vise à savoir s'il est légitime d'étudier les effets liés à l'inefficience technique. Autrement dit, si $H_0 : \delta_0 = \delta_1 = \dots = \delta_n = 0$ est vérifiée alors cela signifie que les caractéristiques propres aux immigrants de première génération et les variables de contexte n'ont pas d'influence sur la performance productive en français écrit.

Le premier test du rapport de vraisemblance (T_1) indique que l'on peut rejeter l'hypothèse nulle, selon laquelle il n'y a pas d'inefficience technique au seuil de 1% de risque : l'estimation d'une frontière de production ne peut donc pas se réduire à une régression MCO. L'approche par une frontière stochastique de production s'en trouve ainsi justifiée. Les deuxième et dernier tests du rapport de vraisemblance (T_2 et T_3) permettent également de rejeter les hypothèses H_0 . D'après ces tests, la fonction Translog semble ainsi être la plus appropriée et la spécification de l'inefficience technique par les variables Z_i présentées précédemment semble bonne.

Tableau 3: Tests d'hypothèses pour la fonction Translog

Hypothèse nulle	LR	Valeur critique $\chi^2_{0,99}$	Décision
$T_1 : H_0 : \sigma_u = 0$	480	6.65	H_0 rejetée
$T_2 : H_0 : \beta_{ik} = 0$	70.29	15.09	H_0 rejetée
$T_3 : H_0 : \delta_0 = \delta_1 = \dots = \delta_n = 0$	377.83	33.41	H_0 rejetée

Source: calculs de l'auteur

L'ensemble de ces tests ainsi que les résultats obtenus semblent donc bien confirmer nos hypothèses de départ.

La fonction Translog ne nous permet pas de lire les élasticités des facteurs de production directement à partir des coefficients, comme c'est le cas pour la fonction Cobb-Douglas. Afin d'estimer les élasticités directes, il faut les calculer de la façon suivante :

$$\varepsilon_j = \frac{\partial \ln(y)}{\partial \ln(x_i)} = \beta_i + \sum_j \beta_{ij} \ln(x_j)$$

Le tableau 4 donne les valeurs moyennes minimales puis maximales des élasticités directes des deux facteurs de production. Dans l'interprétation des résultats, seules les élasticités moyennes seront considérées.

L'un des avantages de la fonction Translog étant le calcul d'élasticités de substitution, nous avons également inclus ces dernières après avoir appliqué la formule suivante ayant servi à leur estimation :

$$DES_{ij} = \frac{x_i f_i + x_j f_j}{x_i x_j} = \frac{f_i f_j}{2f_{ij} f_i f_j - f_{ii} f_j^2 - f_{jj} f_i^2}$$

Avec :

$$f_i = \frac{\partial f}{\partial x_i} = \frac{y}{x_i} \left(\beta_i + \sum_j \beta_{ij} \ln(x_j) \right)$$

$$f_{ij} = \frac{y}{x_i x_j} (\beta_{ij} + \varepsilon_i \varepsilon_j)$$

$$f_{ii} = \frac{y}{x_i^2} (\beta_{ii} + \varepsilon_i^2 - \varepsilon_i)$$

Tableau 4 : Valeurs des élasticités directes et de substitution des facteurs de production.

Facteurs de production	Minimum	Moyenne	Maximum
Élasticités directes			
Durée de scolarité	0.0238	0.044	0.087
Temps de présence en France	0.0023	0.014	0.047
Rendement d'échelle	0.0261	0.058	0.134
Élasticité de substitution directe			
Durée de scolarité*temps de présence	0.699	1.513	2.211

Source : calculs de l'auteur

Pour les deux facteurs de production, tous deux significatifs, il est possible d'observer que la production de compétences dépend davantage de la durée d'éducation (coefficient de 0.044) que du temps de présence dans le pays d'accueil. Une augmentation de 10% du temps de scolarité entraîne une augmentation de la production de compétences en français écrit de 0,44%. Une augmentation de 10% du temps de présence en France élève cette même production de 0,14%. Ainsi pour une durée moyenne de présence en France de 25 ans, une augmentation de 2,5 ans entraîne une augmentation de 0,14% du score en français écrit. De plus, pour une durée moyenne de 10 ans de scolarité, une année d'éducation supplémentaire entraîne une hausse du score en compétences de 0.44%.

Les élasticités semblent toutes deux assez faibles, mais cela tient davantage à la nature même des variables (dont l'amplitude est assez restreinte) qu'à une inélasticité des facteurs de production, comme en témoigne le tableau 5. Ce tableau illustre bien le fait qu'1% d'une durée d'éducation en moyenne plus élevée a un plus grand impact sur les scores en français écrit que 1% d'une durée d'éducation plus dispersée et donc en moyenne plus faible au sein de la population étudiée. Ce tableau témoigne aussi de l'importance d'une éducation de base conséquente auprès du public migrant en vue d'acquérir plus la langue du pays d'accueil. En effet, si le public migrant avait au minimum un niveau d'éducation plus ou moins équivalent au brevet des collèges, l'augmentation de 10% du temps moyen de scolarité aurait permis une augmentation de 2,7% du score en français écrit.

Tableau 5 : Valeurs des élasticités directes des facteurs de production sous certaines hypothèses de durée d'éducation ou de temps de présence en France

	Minimum	Moyenne	Maximum
Élasticités directes de la durée de scolarité...			
... si la durée minimale de scolarité des primo-arrivants était de 9 ans (= niveau Brevet des collèges)	0.21	0.27	0.35
... si la durée minimale de scolarité des primo-arrivants était de 12 ans (= niveau Baccalauréat)	0.39	0.47	0.56
... si la durée minimale de scolarité des primo-arrivants était de 14 ans (= niveau Bac+2)	0.68	0.87	1.16
Élasticités directes du temps de présence...			
... si la durée minimale de présence en France était de 5 ans	0.014	0.026	0.057
... si la durée minimale de présence en France était de 10 ans	0.044	0.065	0.112
... si la durée minimale de présence en France était de 20 ans	0.16	0.21	0.29

Source : calculs de l'auteur

Sur la base des résultats présentés, les primo-arrivants ne semblent pas non plus bénéficier d'économies d'échelle associées à leurs facteurs (cf. tableau 4). En effet, les rendements d'échelle⁷⁰ dans la production sont inférieurs à l'unité : les rendements sont donc décroissants. En conséquence, la production de compétences varie de façon moins importante que la variation des facteurs de production utilisés. Ceci est cohérent, par exemple, avec l'idée que l'acquisition de la langue du pays d'accueil se fait principalement lors des premières années de présence sur la terre d'immigration. Ce résultat justifie par ailleurs l'approche retenue pour estimer la frontière de production (approche stochastique). En effet, une approche déterministe dite par enveloppe n'aurait pas permis de rendre compte de la non-linéarité des rendements des facteurs de production contrairement à l'approche stochastique⁷¹. Par ailleurs, les élasticités directes de substitution directe⁷² montrent qu'il y a bel et bien substituabilité entre le temps de scolarité et le temps de présence en France, les élasticités étant positives. En d'autres termes, pour une

⁷⁰ Les rendements d'échelle sont calculés en faisant la somme des élasticités directes des facteurs de production : $\varepsilon = \sum_i \varepsilon_i$

⁷¹ Les rendements sont supposés constants dans les modèles DEA (Data Envelopment Analysis).

⁷² Cette mesure de l'élasticité a toutefois pour inconvénient, notamment lorsqu'il y a plus de deux facteurs de production, de considérer par hypothèse que les autres facteurs sont fixes.

durée moyenne d'études et de présence en France, une année d'éducation équivaut à une année et demie de présence sur le territoire d'accueil. Ces élasticités de substitution justifient à leur tour le choix d'une spécification Translog : une fonction Cobb-Douglas ne nous aurait effectivement pas permis de mesurer la substituabilité des deux facteurs de production.

7. Conclusion

L'intérêt de cet essai a été d'estimer les facteurs environnementaux et les caractéristiques propres aux primo-arrivants susceptibles d'expliquer leurs performances dans la production de compétences en français écrit. Les principaux résultats mettent en avant que la langue maternelle, l'âge au moment de l'immigration, le diplôme de la mère et la vie sociale jouent un rôle essentiel dans la capacité à utiliser efficacement son temps de présence en France et son temps d'éducation scolaire. Par ailleurs, plus un individu immigré jeune et plus son niveau d'éducation est élevé, meilleure sera alors sa capacité à utiliser les ressources qu'il a à disposition pour produire de la compétence en français écrit. Ces résultats devraient être pris en compte dans le débat autour des politiques éducatives et de formation à l'égard des primo-arrivants, en vue de favoriser leur apprentissage du français, indispensable pour s'insérer dans la vie active et dans la société. Les politiques publiques devraient ainsi mieux cibler les actions à mettre en place selon les caractéristiques propres aux primo-arrivants et notamment selon leur langue maternelle.

Bien que cet essai soit particulièrement éclairant sur les déterminants de l'acquisition du français écrit par les immigrés de première génération installés en France, elle ne permet cependant pas d'étudier l'existence d'un éventuel *effet de cohorte* (Borjas, 1985 ; Carliner, 2000). De nombreuses études se sont effectivement limitées au temps de présence sur le territoire d'accueil sans tenir compte de l'évolution du niveau de qualifications des migrants à travers le temps au moment de leur arrivée (Chiswick et Miller, 1994 ; Espinosa et Massey, 1997). Dans des recherches ultérieures, il faudrait donc ainsi prendre en considération ce point, en menant par exemple une analyse sur données de panel. Or, de telles données n'existent pas à ce jour pour la France. Cependant, le renouvellement de l'enquête

Information et Vie Quotidienne (IVQ) devrait apporter de nouveaux éclairages à long terme.

Des données plus précises sur les trajectoires professionnelles auraient également été les bienvenues pour étudier le phénomène de production de compétences. En effet, l'étude de ces trajectoires et du recours à la formation continue formelle ou informelle approfondirait sans doute ce présent essai. Ceci semble d'autant plus important que dans une perspective économique, les incitations à produire de la compétence dépendent principalement des sanctions infligées par le marché du travail, tant en termes d'emploi occupé que de salaires. Nous pouvons alors supposer que les individus n'ont peut-être pas tous besoin de bien maîtriser le français écrit en raison de leur situation occupée face au marché du travail. Par exemple, si un individu décide de se retirer du marché du travail et de vivre en autarcie, sa maîtrise du français importera peu car il n'y aura aucune sanction économique d'un déficit de compétences en français écrit.

Cette faible maîtrise de la langue peut d'ailleurs être la cause exacte du retrait du marché du travail. En effet, tel que nous l'avons démontré dans l'essai n°273, un individu peut se retirer du marché du travail parce qu'il est conscient que son niveau de compétences est trop faible pour qu'il lui soit rentable de s'insérer sur le marché du travail, au regard de ses coûts d'apprentissage et de ses bénéfices. Bien entendu, la question est délicate en raison d'un potentiel biais d'endogénéité. En effet, est-ce parce que les individus sont retirés du marché du travail qu'ils ont un faible niveau de compétences ou bien est-ce parce qu'en tant qu'agents rationnels, ils savent qu'il ne leur sera pas rentable d'entrer dans la vie active, étant donné leur pauvre niveau de compétences ? L'enquête IVQ apparaît assez limitée pour étudier ces relations assez complexes entre trajectoires professionnelles et production de compétences, mais la mise en place du Programme pour l'Évaluation Internationale des Compétences des Adultes (PIAAC) devrait cependant s'avérer plus instructive.

⁷³ « Le rôle des compétences de base dans la participation au marché du travail et l'employabilité »

BIBLIOGRAPHIE DE L'ESSAI 4

- AKRESH I.-R. & FRANK R. [2011], At the intersection of self and other: English language ability and immigrant labor market outcomes, *Social Science Research*, vol. 40, n°5, pp. 1362-1370.
- ALBA R. D. & NEE V. [2003], *Remaking the American mainstream: Assimilation and the new immigration*, Cambridge, MA: Harvard University Press
- AMARA N. & ROMAIN R. [2000], *Mesure de l'efficacité technique: revue de la littérature*, Québec: Centre de recherche en économie agroalimentaire, Université Laval.
- BEAN F.-D. & STEVENS G. [2003], *America's Newcomers and the Dynamics of Diversity*, New-York, Russell Sage Foundation Publications.
- BECKER G.-S. [1965], A Theory of the Allocation of Time, *The Economic Journal*, vol. 75, n°299, pp. 493-517.
- BECKER G. & TOMES N. [1986], Human Capital and the Rise and Fall of Families, *Journal of Labor Economy*, vol. 4, n°3, pp. 1-39.
- BERMAN E., LANG K. & SINIVER E. [2003], Language-skill complementarity: returns to immigrant language acquisition, *Labour Economics*, vol. 10, n°3, pp. 265-290.
- BERNSTEIN B. [1975], *Langage et classes sociales: codes sociolinguistiques et contrôle social*, Présentation de J.-C., Chamboredon, Paris : Minuit
- BIRDSONG D. [2005], Interpreting age effects in second language acquisition. In Kroll J.-F. & de Groot A.-M.-B. (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 109-127). New York: Oxford University Press.
- BIRDSONG D. & MOLIS M. [2001], On the evidence for maturational effects in second language acquisition, *Journal of Memory and Language*, vol. 44, n°2, pp. 235-249.
- BORJAS G.-J. [1985], Assimilation, changes in cohort quality, and the earnings of immigrants, *Journal of Labor Economics*, vol. 3, n°4, pp. 463-489.
- CARLINER G. [2000], The language ability of US immigrants: Assimilation and cohort effects, *International Migration Review*, vol. 34, n°129, pp. 158-182.
- CARNEVALE A.-P., FRY R.-A. & LOWEEL B.-L. [2001], Understanding, speaking, reading, writing, and earnings in the immigrant labor market, *The American Economic Review*, vol. 91, n°2, pp. 159-163.
- CHISWICK, B.-R. [1998], Hebrew language usage : determinants and effects on earnings among immigrants in Israel, *Journal of Population Economics*, vol. 11, n°2, pp. 253-271.
- CHISWICK B.-R. & MILLER P.-W. [1994], Language choice among immigrants in a multi-lingual destination, *Journal of Population Economics*, vol. 7, n°2, pp. 119-131.

- CHISWICK B.-R. & MILLER P.-W. [1995], The endogeneity between language and earnings: International analyses, *Journal of Labor Economics*, vol. 13, n°2, pp. 246-288.
- CHISWICK B.-R. & MILLER P.-W. [2001], A model of destination-language acquisition: Application to male immigrants in Canada, *Demography*, vol. 38, n°3, pp. 391-409.
- CHISWICK B.-R. [2008], *The Economics of Language: An Introduction and Overview*, Institute for the Study of Labor (IZA).
- CRYSTAL D. [1987], *The encyclopedia of language*, Cambridge : CUP.
- DE GRIP A. & VAN LOO J. [2002], The economics of skills obsolescence : a review, *Research in Labor Economics*, n°21, pp. 1-26.
- DEBREU G. [1951], The coefficient of resource utilization, *Econometrica : Journal of the Econometric Society*, vol. 19, n°3, pp. 273-292.
- DUSTMANN C. & FABBRI F. [2003], Language proficiency and labour market performance of immigrants in the UK, *The Economic Journal*, vol. 113, n°489, pp. 695-717.
- ESPENSHADE T.-J. & FU H. [1997], An analysis of English-language proficiency among US immigrants, *American Sociological Review*, vol. 62, n°2, pp. 288-305.
- ESPINOSA K.-E. & MASSEY D.-S. [1997], Determinants of English proficiency among Mexican migrants to the United States, *International Migration Review*, vol. 31, n°1, pp. 28-50.
- FARRELL M.-J. [1957], The measurement of productive efficiency, *Journal of the Royal Statistical Society, Series A (General)*, vol. 120, n°3, pp. 253-290.
- FLORAX R.-T., DE GRAAFF & WALDORF B.-S. [2005], A spatial economic perspective on language acquisition: segregation, networking, and assimilation of immigrants. *Environment and Planning A*, vol. 37, n°10, pp. 1877-1897.
- HANUSHEK E.-A. [1979], Conceptual and empirical issues in the estimation of educational production functions, *Journal of Human Resources*, vol. 14, n°3, pp. 351-388.
- HAWKINS P.-R. [1969], Social class, the nominal group and reference, *Language and Speech*, vol. 12, n°2, pp. 125-135.
- JASSO G. & ROSENZWEIG M.-R. [1990]. *The new chosen people: Immigrants in the United States*, National Committee for Research.
- JOHNSON J.-S. & NEWPORT E.-L. [1989], Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language, *Cognitive psychology*, vol. 21, n°1, pp. 60-99.
- KOOPMANS T.-C. [1951], Analysis of production as an efficient combination of activities. *Activity Analysis of Production and Allocation*, n°13, pp. 33-37.

- LAZEAR E. [1999], Language and culture. *Journal of Political Economy*, vol. 107, n°6, pp. 95-126.
- LONG M. [1990], Maturational constraints on language development, *Studies in Second Language Acquisition*, vol. 12, n°3, pp. 251-285.
- PORTES A. & RUMBAUT R.-G. [2006], *Immigrant America : A portrait*, 3rd (eds) Berkley : University of California Press
- SERVICE E. & CRAIK F.-I.-M. [1993], Differences Between Young and Older Adults in Learning A Foreign Vocabulary, *Journal of Memory and Language*, vol. 32, n°5, pp. 608-623.
- STEVENS G. [1994]. Immigration, emigration, language acquisition, and the English language proficiency of immigrants in the United States. *Immigration and ethnicity: The integration of America's newest arrivals*, pp. 163-185.
- TAINER E. [1988], English language proficiency and the determination of earnings among foreign-born men, *Journal of Human Resources*, vol. 23, n°1, pp. 108-122.
- VELTMAN C. [1988], Modelling the language shift process of Hispanic immigrants, *International Migration Review*, vol. 22, n°4, pp. 545-562.
- WALDORF B.-S., BECKHUSEN J., FLORAX R.-J & DE GRAAFF T. [2010], The role of human capital in language acquisition among immigrants in US metropolitan. *Regional Science Policy and Practice*, vol. 2, n°1, pp. 39-49.
- WOLTON D. [2008], *Francophonie et migrations internationales*. Cellule de réflexion stratégique de la francophonie. www.francophonie.org

ANNEXES DE L'ESSAI 4

ANNEXE A : DESCRIPTION DES VARIABLES

VARIABLES Z_i	Définition
Y : Score en français écrit	Output : Score allant de 0 à 100
X_i	Facteurs de production
Durée de scolarité	Nombre d'années d'études
Temps de présence	Nombre d'années de présence en France
Z_i	Variables retenues pour l'efficienc technique
Arrivée avant l'âge de scolarité obligatoire en France	0= non ; 1= oui
La langue d'apprentissage de la lecture est la même que la langue maternelle	0= non ; 1= oui
La langue maternelle	1= français ; 2= langues berbères ; 3= arabe ; 4= langues latines ; 5= autres langues européennes ; 6= langues turques ; 7= langues africaines ; 8= langues asiatiques
La fréquence de fréquentation de la famille	1= souvent ; 2= occasionnellement ; 3= jamais
La fréquence de la fréquentation des amis	1= souvent ; 2= occasionnellement ; 3= jamais
Le lieu de résidence	0= hors-ZUS ; 1= ZUS
Le niveau de diplôme de la mère	1= jamais scolarisée ou sans diplôme ; 2= diplôme de niveau primaire ; 3= diplôme du secondaire ; 4= diplôme du supérieur

ANNEXE B : STATISTIQUES DESCRIPTIVES

VARIABLES Z_i	Observations	Moyenne	Écart-Type	Min.	Max.
Score en français écrit	782	81,44	15,40	15,35	100
Durée de scolarité	782	10,45	3,76	1	20
Temps de présence en France	782	27,14	13,54	1	55
La langue d'apprentissage de la lecture est la même que la langue maternelle	782	0,65	0,48	0	1
Arrivée avant l'âge de scolarité obligatoire en France	782	0,47	0,50	0	1
La langue maternelle	782	3,37	2,07	1	8
La fréquence de fréquentation de la famille	782	1,51	0,63	1	3
La fréquence de la fréquentation des amis	782	1,51	0,61	1	3
Le lieu de résidence	782	0,16	0,37	0	1
Le niveau de diplôme de la mère	782	1,44	0,81	1	4

Source : calculs de l'auteur ; enquête Information et Vie Quotidienne 2004, INSEE (données pondérées)

CONCLUSION GÉNÉRALE

Au regard des théories du capital humain et du filtre, le niveau d'études, souvent matérialisé par un diplôme, apparaît tel un signal important et relativement fiable de la productivité des individus et plus généralement, de leur niveau d'aptitudes. Or, face à la massification de l'enseignement supérieur, les individus qui entrent sur le marché du travail sont de plus en plus nombreux à présenter des niveaux d'études élevés et plus ou moins similaires. L'homogénéisation de cette offre de travail amène donc à s'interroger sur la place que tient encore le diplôme comme signal auprès des employeurs. En effet, dans ce contexte, comment peuvent-ils s'assurer qu'ils recrutent toujours les meilleurs individus et comment ces derniers peuvent-ils encore se différencier sur le marché du travail, si ce n'est par le diplôme ? Le recours à la notion de compétence apparaît alors indispensable, comme en témoigne d'ailleurs la fréquence à laquelle ce terme est utilisé dans la vie quotidienne, éducative et professionnelle. Il est fréquemment rappelé que les employeurs sont aujourd'hui à la recherche d'individus dits compétents, comme s'il s'agissait d'un meilleur gage de qualité, dans nos économies du savoir et de la connaissance. Se pose alors naturellement la question de la définition et de la mesure de ces compétences. A partir d'un travail sur les compétences de base, cette présente thèse repose ainsi sur l'idée qu'il existe une déconnexion partielle entre les différentes échelles de mesure du capital humain et que la mesure du capital humain via la seule durée d'éducation (ou le seul niveau de diplôme) ne saurait se suffire à elle-même.

C'est dans ce cadre que cette thèse a ainsi cherché à montrer que la prise en compte des compétences de base des adultes dans la mesure du niveau de capital humain améliorerait l'explication de certains « *outcomes* » du marché du travail. Mais la

déconnexion partielle des mesures du capital humain n'est pas forcément due à l'homogénéisation de l'offre de travail. Elle peut également être liée à l'origine migratoire de l'individu ou de sa famille et à son intégration dans la société. En effet, il semble naturel de penser que les individus d'origine étrangère se trouvent plus en mal de valoriser leurs compétences de base (écriture, lecture, compréhension écrite et orale) face à une demande de travail pour laquelle ils ne sont pas *a priori* préparés, surtout lorsque la production même de ces compétences s'est faite dans une langue différente. La maîtrise des compétences en français écrit et oral relève donc plutôt de compétences linguistiques indispensables pour une intégration durable et réussie en France. Or, la société et l'environnement général dans lesquels s'intègre l'individu a également un rôle actif dans la production des compétences linguistiques. C'est pourquoi cette présente thèse a parallèlement cherché à montrer que certaines populations migrantes étaient plus à même de produire et d'acquérir la langue française que d'autres, en raison de leurs caractéristiques propres. Notre travail se situe donc à la croisée de l'économie de l'éducation, du travail et de l'immigration.

Si nos travaux ont permis d'apporter quelques éclairages dans ces domaines, les perspectives de recherches sur les compétences de base restent pourtant nombreuses. Elles le sont d'autant plus que sur le plan national et international, de nouvelles données sont ou seront très prochainement disponibles. En effet, l'enquête IVQ a été renouvelée en 2011 et ses premiers résultats ont été dévoilés à la fin de l'année 2012. Le Programme pour l'évaluation internationale des compétences des adultes (PIAAC) devrait quant à lui rendre publiques ses données courant octobre 2013. Ces nouvelles enquêtes, couplées à celles qui existent déjà, donneront des mesures plus régulières du niveau de compétences de base des adultes et permettront ainsi de mieux comprendre les effets du vieillissement en séparant les effets d'âge des effets de génération dans l'étude des trajectoires professionnelles et salariales. Dans la mesure où les profils de compétences établis en fonction de l'âge à partir de données transversales sont difficiles à interpréter, le renouvellement de ce type d'enquête est prometteur de ce point de vue et devrait prolonger à terme les recherches menées dans la première partie de cette thèse.

De même, en croisant les données de PIAAC avec celles de PISA, il serait possible de mesurer l'impact des performances en littératie et en numératie au sortir de l'enseignement secondaire sur le niveau de ces mêmes compétences à l'âge adulte.

Par exemple, la cohorte de PISA 2000 avait entre 27 et 28 ans au moment de la collecte de données de PIAAC, soit en 2011-2012. De fait, il serait tout à fait envisageable, grâce à cette toute nouvelle enquête, d'évaluer dans quelle mesure le passage du système scolaire au monde professionnel impacte le niveau de compétences de base par exemple.

Les données longitudinales s'avèrent tout aussi intéressantes concernant l'étude de l'acquisition du français par le public migrant. En effet, si certains expliquent qu'il est plus facile d'acquérir une langue étrangère lorsque l'on est jeune, d'autres considèrent qu'il existe aussi un effet de génération. De nombreuses études se sont limitées à la durée d'éducation comme principal élément explicatif des compétences dans la langue du pays d'accueil (Chiswick et Miller, 1994 ; Espinosa et Massey, 1997) et ce, sans tenir compte de l'évolution du niveau de qualifications des migrants à travers le temps, au moment de leur arrivée. Pourtant, Borjas (1985) a démontré que le niveau d'instruction et de qualification des migrants arrivant aux États-Unis avait légèrement diminué, notamment depuis les années 40. La scolarité moyenne des immigrants arrivés avant les années 50 était plus élevée que celle des migrants arrivés dans les années 80. Carliner (2000) ira également dans ce sens en expliquant que ce constat provient principalement de la diversification des flux migratoires et de l'arrivée de migrants dont la langue maternelle est très éloignée de l'anglais. Ces éléments sont à prendre en compte dans les politiques de formation à l'égard des immigrants de façon à homogénéiser le niveau de capital humain des nouveaux arrivants et plus généralement de la population tout entière.

L'homogénéisation du niveau de capital humain des individus revêt d'un certain enjeu pour l'équité et l'efficacité du système éducatif et par conséquent pour l'économie. Selon Heckman (2006), les inégalités de scores proviennent essentiellement du contexte familial et l'école ne contribuerait que très peu à leur formation. En effet, les enfants issus de familles pauvres ont les plus faibles scores en mathématiques entre 6 et 12 ans et la plus grande partie de l'écart de scores qui demeure à 12 ans avec les enfants issus d'un milieu plus aisé apparaissait déjà à 6 ans, soit à l'âge de la première scolarisation. De fait, la société se doit de compenser le plus rapidement possible les effets négatifs d'un contexte familial défavorisé sur les acquis scolaires. Attendre trop longtemps pour le faire suppose des coûts plus élevés et il serait dès lors moins efficient d'investir dans la formation des plus *désavantagés*. Consolider les compétences de base de ces derniers est important

dans les premières années de scolarité pour deux principales raisons. Tout d'abord, l'apprentissage des premières années de scolarité confère de la valeur aux compétences de base acquises, valeur qui renforce la motivation à apprendre davantage. En outre, plus la maîtrise de compétences cognitives et non-cognitives se fait jeune, plus le développement d'autres compétences se fera facilement à des âges plus avancés. Pour Heckman (2006), investir dans la formation des enfants et des adolescents issus de milieux défavorisés est donc une piste de politique publique rarement envisagée, qui favoriserait pourtant l'équité et la justice sociale en même temps qu'elle stimulerait la productivité économique et de fait la croissance. Il faudrait donc élever la productivité moyenne du travail non pas en poussant la main d'œuvre à être plus productive qu'elle ne l'est déjà, mais en cherchant à rapprocher de la productivité médiane la tranche de la population active la moins productive. Selon Carneiro et Heckman (2003), il s'agirait même d'une piste plus rentable que celle qui consiste à diminuer la taille des classes par exemple.

Pour autant, même si les investissements en formation générale sont moins rentables lorsqu'ils s'effectuent à l'âge adulte, ils restent néanmoins bénéfiques. En effet, les bienfaits obtenus d'une intervention publique auprès d'élèves défavorisés seront d'autant plus durables et décuplés que de nouvelles formations de qualité seront suivies par la suite. En d'autres termes, les formations suivies tout au long de la vie d'un individu doivent permettre de maximiser la rentabilité des premiers investissements en éducation générale (Heckman, 2006).

C'est dans ce schéma que s'inscrit une recommandation européenne sur les compétences clés pour l'éducation et la formation tout au long de la vie. Faite par le Parlement européen et le Conseil européen à la fin de l'année 2006, elle cherche à ce que les États membres permettent aux jeunes et aux moins jeunes, principalement ceux en difficulté, de développer et de mettre à jour des compétences clés, nécessaires à la vie d'adulte. Suite à cette recommandation, la France, en 2009, a ainsi lancé un programme de formation nommé Compétences clés, financé conjointement par les Régions, l'État et le Fonds Social Européen (FSE). Au cours de l'année 2011, 50 100 personnes ont participé à ce programme. 92% des participants étaient alors sans emploi et 34% inscrits à Pôle Emploi depuis plus d'un an (DARES, 2013). De futures recherches pourraient alors s'axer sur l'évaluation d'impact de ce type de formation visant le développement des compétences clés sur les carrières professionnelles, à l'instar des travaux de Fougère et al. (2001) concernant la

formation continue. Cela aurait d'ailleurs l'intérêt d'élargir le champ de l'évaluation de la formation en France, champ dans lequel cette dernière aurait accumulé un certain retard par rapport à la plupart des pays industrialisés (Crépon et Ferracci, 2009).

Alors que les politiques de consolidation des compétences de base revêtent un enjeu éthique, social et économique, elles restent encore bien trop limitées comme l'illustre l'exemple précédent. Pourtant, il devient urgent d'investir dans ces compétences, surtout dans nos économies axées sur le Savoir et la Connaissance et résolument tournées vers les NTIC. Cette question est d'autant plus vive que pour le cas français, les résultats de PISA 2009 mettent en avant un creusement des inégalités scolaires d'origine sociale au cours de ces dix dernières années (Meuret et Lambert, 2012). Ce creusement serait notamment dû à la difficulté croissante de la France et de son système éducatif à intégrer ses populations d'origine immigrée. L'idée d'une intervention publique auprès des jeunes issus de milieux défavorisés n'a donc jamais autant été d'actualité pour réduire les inégalités de compétences au sein de la main d'œuvre de demain et renouer avec la croissance économique qui nous fait tant défaut aujourd'hui.

BIBLIOGRAPHIE DE LA CONCLUSION GÉNÉRALE

- BORJAS G.-J. [1985], Assimilation, changes in cohort quality, and the earnings of immigrants, *Journal of Labor Economics*, vol. 3, n°4, pp. 463-489.
- CARLINER G. [2000], The language ability of US immigrants: Assimilation and cohort effects, *International Migration Review*, vol. 34, n°1, pp. 158-182.
- CARNEIRO P. & HECKMAN J.-J. [2003], in *Inequality in America: What Role for Human Capital Policies ?* Heckman J.-J., Krueger A.-B. & Friedman B., Eds. (MIT Press, Cambridge, MA, 2003), ch.2, pp. 77-237.
- CHISWICK B.-R. & MILLER P.-W. [1994], Language choice among immigrants in a multi-lingual destination, *Journal of Population Economics*, vol. 7, n°2, pp. 119-131.
- CRÉPON B., & FERRACCI M. [2009], Le cas français: approfondir les connaissances empiriques pour mieux cibler la formation, *Travail et emploi*, n°2, pp. 107-111.
- DARES [2013], Les bénéficiaires de la formation « compétences clés », Publication de la Direction de l'Animation de la Recherche, des Études et des Statistiques (DARES), n°044.
- ESPINOSA K.-E. & MASSEY D.-S. [1997], Determinants of English proficiency among Mexican migrants to the United States, *International Migration Review*, vol. 31, n°1, pp. 28-50.
- FOUGÈRE D., GOUX, D. & MAURIN E. [2001], Formation continue et carrières salariales: une évaluation sur données individuelles, *Annales d'Économie et de Statistique*, n°62, pp. 49-69.
- HECKMAN J.-J. [2006], Skill formation and the economics of investing in disadvantaged children, *Science*, vol. 312, n°5782, pp. 1900-1902.
- MEURET D. & LAMBERT M. [2012], La dégradation de l'équité de l'école française dans les années 2000, *Revue française de pédagogie*, vol. 177, n°4, pp. 85-100.

TABLE DES MATIÈRES

INTRODUCTION GÉNÉRALE.....	11
BIBLIOGRAPHIE DE L'INTRODUCTION GÉNÉRALE.....	23

PARTIE I. LA VALORISATION DES COMPÉTENCES DE BASE SUR LE MARCHÉ DU TRAVAIL 27

ESSAI 1

LES RENDEMENTS DES COMPÉTENCES DE BASE SUR LE MARCHÉ DU TRAVAIL 28

1. Introduction.....	28
2. Le niveau des compétences de base peut-il expliquer le déclassement ou le surclassement sur le marché du travail ?.....	30
2.1. Cadre théorique.....	30
2.2. <i>Much ado about nothing</i> ?.....	33
2.3. Modèle économétrique.....	34
2.4. La recherche d'un lien empirique entre déclassement, surclassement et compétences de base.....	37
2.4.1. Statistiques descriptives.....	37
2.4.2. Résultats.....	39
3. La rémunération des compétences de base : fonctions de gains.....	45
3.1. Revue de littérature.....	45
3.2. Hypothèses.....	47
3.3. Méthodologie.....	49
3.3.1. Choix de l'échantillon.....	50
3.3.2. Modèle économétrique.....	51
3.3.2.1. Le modèle standard de Mincer.....	52
3.3.2.2. Correction du biais de sélection : modèle d'Heckman.....	53
3.4. Statistiques descriptives.....	55
3.5. Résultats.....	58
3.5.1. Fonctions mincériennes augmentées des scores de compétences.....	58
3.5.2. Les rendements privés des compétences sont-ils non-linéaires ?.....	59
3.5.3. Fonctions de gains corrigées du biais de sélection.....	63
4. Conclusion.....	68
BIBLIOGRAPHIE DE L'ESSAI 1.....	71
ANNEXES DE L'ESSAI 1.....	76

ESSAI 2

LE RÔLE DES COMPÉTENCES DE BASE DANS LA PARTICIPATION AU MARCHÉ DU TRAVAIL ET L'EMPLOYABILITÉ..... 97

1. Introduction.....	97
2. Recherche d'emploi et capital humain.....	99
3. Compétences de base et marché du travail	102
3.1. Les compétences de base des non-diplômés et des faiblement diplômés	103
3.2. L'obsolescence des compétences de base dans un contexte marqué par de fortes innovations technologiques.....	104
3.3. Marché du travail et pratiques discriminatoires.....	105
4. Méthodologie : modèle logistique multinomial ordonné généralisé.....	106
5. Statistiques descriptives.....	108
6. Résultats.....	113
6.1. Situation occupée à l'égard du marché du travail et scores de compétences....	113
6.1.1. Par niveau de diplôme.....	113
6.1.2. Par tranche d'âge	115
6.2. Situation occupée à l'égard du marché du travail et indicatrices de déficit / de surplus de compétences.....	117
7. Conclusion	123
BIBLIOGRAPHIE DE L'ESSAI 2	125
ANNEXES DE L'ESSAI 2	128

PARTIE II. LA PRODUCTION DE COMPÉTENCES EN FRANÇAIS SELON L'ORIGINE MIGRATOIRE.....135

ESSAI 3

CAPITAL HUMAIN ET DISTANCE LINGUISTIQUE : UNE INÉGALE ACQUISITION DES COMPÉTENCES EN FRANÇAIS SELON L'ORIGINE MIGRATOIRE.....136

1. Introduction.....	136
2. Choix migratoires et capital humain.....	138
3. Distance linguistique et inégale acquisition de la langue du pays d'accueil	141
4. Objet d'étude et hypothèses.....	143
5. Statistiques descriptives.....	144
5.1. Les écarts de scores selon l'origine migratoire	144
5.2. Les variables de contrôle.....	146
6. Fonctions de production de compétences en français.....	149
7. Décomposition Oaxaca (1973) et Blinder (1973).....	152
8. Décomposition des scores selon Juhn, Murphy et Pierce (1993).....	158
9. Conclusion	162
BIBLIOGRAPHIE DE L'ESSAI 3	164
ANNEXES DE L'ESSAI 3	168

ESSAI 4	
LA PRODUCTION DE COMPÉTENCES EN FRANÇAIS ÉCRIT PAR LES IMMIGRES DE PREMIÈRE GÉNÉRATION EN AGE DE TRAVAILLER.....	172
1. Introduction.....	172
2. Les facteurs de production de compétences.....	175
3. Les déterminants des performances productives : hypothèses.....	178
3.1. L'âge au moment de la migration.....	178
3.2. La langue maternelle, la langue parlée à la maison, les origines sociales et les inégalités territoriales.....	179
4. Spécification du modèle de frontière stochastique de production.....	181
4.1. Concepts généraux.....	181
4.2. Modèle de frontière stochastique de production.....	183
4.3. L'indice d'efficacité technique.....	184
5. Statistiques descriptives.....	186
6. Résultats.....	188
6.1. Paramètres de la frontière et de l'inefficacité technique.....	188
6.2. Indices d'efficacité technique.....	191
6.3. Robustesse des résultats et élasticités.....	193
7. Conclusion.....	197
BIBLIOGRAPHIE DE L'ESSAI 4.....	199
ANNEXES DE L'ESSAI 4.....	202
CONCLUSION GÉNÉRALE.....	205
BIBLIOGRAPHIE DE LA CONCLUSION GÉNÉRALE.....	210
TABLE DES MATIÈRES.....	212
LISTE DES TABLEAUX.....	215
LISTE DES FIGURES.....	218
LISTE DES ANNEXES.....	219

LISTE DES TABLEAUX

ESSAI 1

Tableau 1 : Déclassement et surclassement des salariés à partir de l'enquête IVQ.....	38
Tableau 2 : Scores moyens en compétences de base des diplômés d'un BEP/CAP selon leur situation face au déclassement ou au surclassement	39
Tableau 3 : Effets des compétences de base par type de secteurs	42
Tableau 4 : Pourcentage de surclassés et de déclassés déclarant que l'école a été utile ou non pour leur vie professionnelle	43
Tableau 5 : Scores moyens en compétences de base selon que l'on soit en emploi ou non.	56
Tableau 6 : Répartition des individus selon leur niveau de diplôme ainsi leur appartenance à tel ou tel quartile de compétences de base.....	56
Tableau 7 : Salaires moyens et médians selon la présence d'un déficit ou d'un surplus de compétences par rapport à son niveau de diplôme.....	57
Tableau 8 : Équations de gains selon le diplôme et le quartile d'appartenance de la distribution des scores	61
Tableau 9 : Probabilité d'être en emploi (versus être sans emploi) : modèle probit	63
Tableau 10 : Probabilités prédites d'être en emploi ou non selon l'importance du déficit de compétences	65
Tableau 11 : Équations de gains avec et sans correction du biais de sélection	67

ESSAI 2

Tableau 1 : Scores moyens en compétences de base selon la situation occupée face au marché du travail et autres caractéristiques..... 109

Tableau 2 : Scores moyens en compétences de base selon le lieu de résidence (ZUS ou hors-ZUS) et le niveau de diplôme..... 111

Tableau 3 : Scores moyens en compétences de base selon l'origine migratoire et le niveau de diplôme 112

Tableau 4: Moyennes et écarts-types des scores en compétences de base par niveau de diplôme. 118

Tableau 5 : Part des individus présentant des déficits ou des surplus de compétences dans l'échantillon total 119

Tableau 6 : Part des individus présentant un déficit et un surplus de compétences selon la situation occupée à l'égard du marché du travail et le lieu résidence (ZUS et hors-ZUS). 119

Tableau 7 : Part des individus présentant un déficit et un surplus de compétences selon la situation occupée à l'égard du marché du travail et de l'origine migratoire. 121

ESSAI 3

Tableau 1 : Distribution des scores en français oral selon l'origine migratoire 145

Tableau 2 : Statistiques descriptives selon l'origine migratoire (en %) 147

Tableau 3 : Fonctions de production de français écrit et oral selon l'origine migratoire (latine ou non)..... 151

Tableau 4 : Décomposition de l'écart de scores à la moyenne entre les immigrés d'origine latine et ceux d'une origine non-latine 154

ESSAI 4

Tableau 1 : Scores moyens en français écrit selon le nombre d'années d'éducation et le nombre d'années de présence en France	186
Tableau 2 : Résultats de l'analyse en frontière de production stochastique dans le cas d'une fonction Translog	189
Tableau 3 : Tests d'hypothèses pour la fonction Translog	193
Tableau 4 : Valeurs des élasticités directes et de substitution des facteurs de production.....	195
Tableau 5 : Valeurs des élasticités directes des facteurs de production sous certaines hypothèses de durée d'éducation ou de temps de présence en France.....	196

LISTE DES FIGURES

ESSAI 3

Figure 1 : Distribution des scores en français écrit selon l'origine migratoire 145

Figure 2 : Effets caractéristiques et effets rendements le long de la distribution des scores en français écrit selon l'origine migratoire (latine ou non)..... 161

Figure 3 : Effets caractéristiques et effets rendements le long de la distribution des scores en français oral selon l'origine migratoire (latine ou non)..... 162

ESSAI 4

Figure 1 : Distribution des scores en français écrit selon la langue maternelle 187

Figure 2 : Indices d'efficience technique moyens selon l'âge d'arrivée en France et le nombre d'années d'éducation..... 192

LISTE DES ANNEXES

ESSAI 1

Annexe A : Architecture du questionnaire IVQ 2004.....	78
Annexe B : Construction d'un score global en littératie.....	79
Annexe C : able de correspondance d'affichard (1981) :.....	82
Annexes D : Résultats des modélisations probit	83
Tableau D.1 : ... de la probabilité d'être surclassé/déclassé sur la population toute entière.....	83
Tableau D.2 : ... de la probabilité d'être surclassé/déclassé selon la tranche d'âge d'appartenance.....	84
Tableau D.3 : ... de la probabilité d'être surclassé/déclassé selon le niveau de diplôme.	85
Annexe E. : Résultats de la modélisation logistique ordonnée sur le sentiment d'utilité de l'école dans la vie professionnelle.....	87
Annexe F : Probabilités conditionnelles de considérer que l'école a été plus ou moins utile pour la vie professionnelle.....	88
Annexe G . Quotité de travail (en %) dans l'enquête Emploi 2004 - 4 ^{ème} trimestre (H=hommes ; F=femmes)	89
Annexes H. Tests du Khi-deux	92
Tableaux H.1. ... par CSP	92
Tableaux H.2 : ... selon la charge d'enfants	94
Annexes I : Estimation des fonctions de gains par la méthode des moindres carrés ordinaires (MCO).....	95
Tableau I.1. : ... avec le niveau de diplôme.....	95
Tableau I.2. : ... avec le nombre d'années d'études	96

ESSAI 2

Annexe A : Construction des groupes de situations occupées face au marché du travail	129
Annexe B : Résultats des modèles logistiques polytomiques ordonnés généralisés avec introduction de scores en compétences	130
Tableau B.1. : Impact des scores de compétences sur la probabilité d'appartenir à telle ou telle situation occupée face au marché du travail selon le niveau de diplôme	130
Tableau B.2. : Impact des scores de compétences sur la probabilité d'appartenir à telle ou telle situation occupée face au marché du travail selon la tranche d'âge d'appartenance.....	131
Annexe C : Histogrammes et densités des scores en compétences de base	132
Annexe D: Résultats des modèles logistiques polytomiques ordonnés généralisés avec introduction d'indicatrices de déficit et surplus.....	133
Tableau D.1. : Impact des déficits et surplus de compétences sur la probabilité d'occuper telle ou telle situation sur le marché du travail selon le lieu de résidence (ZUS - Hors ZUS).....	133
Tableau D.2. : Impact des déficits et surplus de compétences sur la probabilité d'occuper telle ou telle situation sur le marché du travail selon l'origine migratoire	134

ESSAI 3

Annexe A : Redoublants et effectifs scolarisés dans l'enseignement primaire par région - 2000, 2005 et 2010.....	169
Annexe B : Scores moyens en français écrit et oral selon les caractéristiques individuelles et l'origine migratoire.	170
Annexe C : Décompositions Juhn, Murphy et Pierce.....	171

ESSAI 4

Annexe A : Description des variables	203
Annexe B : Statistiques descriptives	203

CAPITAL HUMAIN ET COMPÉTENCES DE BASE DES ADULTES : PRODUCTION ET VALORISATION

Résumé : La question des compétences de base est de plus en plus présente dans les préoccupations politiques actuelles au regard des coûts importants engendrés par l'illettrisme et supportés par la société. L'objectif de cette présente thèse est alors d'apporter un nouvel éclairage sur cette thématique dans une approche microéconomique. Elle se compose de deux parties. La première traite de la valorisation des compétences de base sur le marché du travail, à travers les essais 1 et 2, alors que la seconde traite de la production de compétences en français selon l'origine migratoire, à travers les essais 3 et 4.

Dans le cadre de la première partie, le premier essai s'intéresse à la question des rendements des compétences de base sur le marché du travail, sous l'angle du déclassement/surclassement professionnel et sous l'angle de la rémunération individuelle. L'hypothèse centrale consiste à supposer que l'éducation, via le diplôme, n'est pas un bon filtre des capacités individuelles. Dès lors les situations de déclassement (*surclassement*), pour un même niveau de diplôme, peuvent résulter d'un plus faible (*fort*) niveau de compétences de base et non forcément d'un déséquilibre entre offre et demande de qualifications. De même, si pour un même niveau de diplôme il existe une certaine hétérogénéité des capacités individuelles, nous pouvons penser que l'ajout des scores en compétences de base tend à diminuer le rendement purement scolaire et à améliorer l'explication de la formation des salaires. Notre approche a pour principal avantage de tenir compte d'un éventuel biais de sélection dans l'estimation des équations de gains en appréhendant l'influence exercée par les compétences de base dans l'accès à l'emploi. Le deuxième essai tente quant à lui de mieux comprendre l'impact des compétences de base sur l'accès à l'emploi mais aussi sur la participation au marché du travail et ce, dans un contexte d'enrichissement généralisé du contenu du travail. La prise en compte de ces compétences doit améliorer l'explication des situations occupées à l'égard du marché du travail (emploi, chômage, inactivité), notamment pour des publics déjà fragilisés face à l'emploi. Plus précisément, pour une demande de travail donnée, nous supposons qu'à chaque situation correspond un niveau de compétences de base.

La lecture, l'écriture, la compréhension écrite et orale du français étant des compétences de base clés pour l'intégration sociale et professionnelle en France, la deuxième partie s'intéresse à la production de ces compétences auprès des populations d'origine immigrée. En effet, la France possède une longue tradition d'accueil des migrants. Leur place dans la société française et leurs capacités à apprendre le français suscite donc un certain intérêt, qu'ils soient primo-arrivants ou enfants d'immigrés, d'autant que peu de recherches ont été réalisées à ce sujet. Dans le troisième essai, nous cherchons à déterminer dans quelle mesure les caractéristiques liées à l'origine migratoire entrave la production de compétences en français et dans quelle mesure la part des inégalités de scores en français sont attribuables à des différences d'efficacité résultant de la distance linguistique entre la langue maternelle et le français. Le quatrième et dernier essai vise, quant à lui, à déterminer les facteurs susceptibles d'expliquer le niveau de maîtrise en français écrit des primo-arrivants (considérés comme producteurs de leurs propres compétences) en âge de travailler. L'originalité de cet essai réside dans l'estimation d'une frontière stochastique de production afin d'expliquer leurs scores en français écrit et leurs performances productives.

MOTS CLÉS : Capital humain, compétences de base des adultes, déclassement, rendements, employabilité, participation au marché du travail, production de compétences, immigrés, distance linguistique.

HUMAN CAPITAL AND BASIC SKILLS : PRODUCTION AND VALORIZATION

Abstract : Basic skills are important in current policies because of high costs generated by illiteracy and supported by the society. The aim of the present thesis is to bring a new contribution on this thematic in a microeconomic approach. It consists of two parts. The first one concerns the valorization of basic skills on the labor market, through chapters 1 and 2, while the second one concerns the production of French skills according to the migratory origins, through chapters 3 and 4.

Chapter 1 is about the returns to basic skills on the labor market, in terms of overeducation/undereducation and in terms of wages. The central hypothesis is to assume that education, through the diploma, is not a good filter for individual capacity. Therefore overeducation (undereducation) for the same level of degree, can result from a low (high) level of basic skills and not necessarily an imbalance between supply and demand for skills. Similarly, for a given level of education, there is some heterogeneity in individual capabilities, so we think that the addition of basic skills in the wage function tends to reduce the return of education and to improve the explanation of the wage formation. In order to avoid a possible selection bias in the estimation of wages equation, our approach has the main advantage to take into account the impact of basic skills in job access. Chapter 2 attempts to have a better understanding of the impact of basic skills on job access, but also on the labor force participation, in a general enrichment context of the work content. The inclusion of these skills should improve the explanation of the situations occupied towards the labor market (employment, unemployment, inactivity), especially for groups already weakened in the job access. More precisely, for a given labor demand, we assume that for every situation corresponds a level of basic skills.

Reading, writing, listening and reading comprehension skills in French language being essential for a successful integration in French society and labor market. The second part of this thesis focuses on the production process of these basic skills in French among immigrants. Indeed, France has a long tradition of welcoming migrants. Their place in French society and their ability to learn its language generates some interest because little research exist about it. Chapter 3 is where we try to determine to what extent the characteristics associated with the migration background hinders the production of French skills, and to what extent the share of inequality in French scores are due to differences in efficiency because of the linguistic distance between the mother tongue and French. Chapter 4, at last, seeks to determine what are the factors that can explain the proficiency level in written French skills of newcomers (considered as producers of their own skills) in working age. The originality of this chapter is the estimation of a stochastic production frontier to explain the scores in written French and the productive performance.

KEY WORDS : Human capital, basic skills of adults, overeducation, private return to basic skills, employability, labor force participation, production skills, immigrants, linguistic distance.