



# Essays on segregation and identity in France

Pascaline Vincent

## ► To cite this version:

Pascaline Vincent. Essays on segregation and identity in France. Economics and Finance. Université de Rennes, 2015. English. NNT : 2015REN1G022 . tel-01299128

**HAL Id: tel-01299128**

**<https://theses.hal.science/tel-01299128>**

Submitted on 7 Apr 2016

**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.



**THÈSE / UNIVERSITÉ DE RENNES 1**  
*sous le sceau de l'Université Européenne de Bretagne*

pour le grade de  
**DOCTEUR DE L'UNIVERSITÉ DE RENNES 1**

*Mention : Sciences Économiques*

**École doctorale Sciences de l'Homme des Organisations  
et de la Société (SHOS)**

présentée par

**Pascaline Vincent**

préparée à l'unité de recherche CREM (UMR6211)  
Centre de Recherche en Economie et Management  
Faculté de Sciences Économiques

---

**Essays on segregation  
and identity in France**

**Thèse soutenue à Rennes  
le 1er Décembre 2015**

devant le jury composé de :

**Catherine Baumont**

Professeur, Université de Bourgogne  
*Rapporteur*

**Yannick L'Horty**

Professeur, Université Paris-Est Marne-La-Vallée  
*Rapporteur*

**Isabelle Lebon**

Professeur, Université de Caen  
*Examineur*

**Fabien Moizeau**

Professeur, Université de Rennes 1  
*Examineur*

**Benoît Tarroux**

Maître de Conférences, HDR, Université de Rennes 1  
*Directeur*



---

This Ph.D. thesis should not be reported as representing the views of University of Rennes 1. The views expressed are those of the author and do not necessarily reflect those of the University.

---

L'Université de Rennes 1 n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.



*“ C’est, [...] un nom d’une signification plus étendue qu’il faut donner aux autres cités, car chacune d’elles est multiple, et non pas une [...]; elle renferme au moins deux cités ennemies l’une de l’autre : celle des pauvres et celle des riches, et chacune de celles-ci se subdivise en plusieurs autres. ”*

Platon  
La République, 315 av. J.C.



# *Remerciements*

Avant toutes choses je tiens à remercier Benoît Tarroux d'avoir accepté de diriger cette thèse, pour sa disponibilité, son soutien et ses conseils qui ont été nombreux et importants. C'est avec beaucoup de reconnaissance et de fierté que je termine ma thèse sous sa direction.

J'adresse également mes remerciements aux membres du jury : Catherine Baumont et Yannick L'Horty qui ont accepté la tâche fastidieuse qu'est de rapporter une thèse. Merci également à Isabelle Lebon et Fabien Moizeau d'avoir accepté de rejoindre ce jury.

Je remercie aussi Yvon Rocaboy et les membres du CREM pour les nombreuses discussions qui ont pu m'aider à avancer dans mon travail. Beaucoup de personnes seraient à citer, mais je remercie tout particulièrement Isabelle Vialle, pour ses précieuses relectures et Fabien Moizeau pour ses conseils avisés tout au long de la thèse. Je remercie également mes co-auteurs Frédéric Chantreuil, pour ses précieux conseils et Gabin Langevin qui a été présent tout au long de l'aventure.

Un grand merci aux membres du Centre de Sciences Humaines de Delhi qui m'ont aidée et soutenue dans des moments difficiles. Basudeb, Aditi, Devika, Lorraine, Bérénice, Anna, vous étiez tous là. Merci à mes collègues de l'université de Lille 3 qui étaient présents pendant cette dernière année.

Je remercie les chefs d'établissement et enseignants du collège Emile Zola ainsi que du collège des Hautes Ourmes de Rennes. Merci aux collégiens qui ont participé aux sessions ainsi qu'aux membres de l'AFEV sans qui ce projet n'aurait pu voir le jour.

Merci à Ewen et Gauthier avec qui j'ai partagé le meilleur bureau, merci de votre amitié sans faille. Un grand merci à mes collègues doctorants du CREM, Gabin, Henri, Xuan, Clément, Guillaume, Ons, May, Guigui, Nico, Manu, Pauline, Mathieu, Siriki, Gaëtan, Julien et sans oublier Elven.

A mes amis, de Santa Barbara à New Delhi, en passant par Lille, Valladolid ou Poznan. J'aimerais avoir un mot pour chacun d'entre vous. Anne, Guillaume, Isa, Dawa, Elsa, Typhaine, Julien, Anna, Geb, Kereoz, Sarah, sans vous l'aventure n'aurait pas été possible.

Je remercie tout naturellement mes parents, Catherine et Alain, mes sœurs Noémie et Marina pour leur soutien inconditionnel tout au long de mes études, et pour leur présence en tout moment. Je souhaite tout spécialement remercier Samia d'être présente à mes côtés. Enfin je remercie mes neveux Nawel, Abdou-Rahman, Abdou-Ssamad et Abdel-Wahid qui donnent du sens à mon travail.





# Table des matières

<b>Remerciements</b>	<b>vii</b>
<b>Contents</b>	<b>viii</b>
<b>Introduction Générale</b>	<b>1</b>
1 Etat des lieux de la ségrégation en France . . . . .	1
2 Identité . . . . .	3
<b>1 Income segregation in large French cities</b>	<b>7</b>
1 Introduction . . . . .	7
2 Measuring residential segregation by income . . . . .	9
2.1 Segregation under movements . . . . .	11
2.2 Segregation under income transfers . . . . .	13
2.3 Sensitivity to areas definition . . . . .	15
3 Economic segregation in French urban areas . . . . .	16
3.1 Database description . . . . .	17
3.2 Results . . . . .	18
4 Conclusion . . . . .	31
5 Appendix . . . . .	32
5.1 Basic properties . . . . .	33
5.2 Proofs of propositions . . . . .	33
5.2.1 Proof of proposition 1.1 (Normalisation) . . . . .	33
5.2.2 Proof of propositions 1.3 and 1.5 (Segregation under movements) . . . . .	34
5.2.3 Unilateral movements . . . . .	34
5.2.4 Proof of proposition 1.7 (Internal transfer) . . . . .	35
5.2.5 Proof of proposition 1.9 (Sensitivity to areas aggregation) . . . . .	36
<b>2 National Identity and Immigrants' Assimilation in France</b>	<b>37</b>
1 Introduction . . . . .	37
2 Literature review . . . . .	41
3 Data, sample and measures of ethnic identity . . . . .	42
3.1 Data and sample . . . . .	42
3.2 The assimilation index . . . . .	43
4 Treatment and descriptive evidence . . . . .	46
4.1 Treatment . . . . .	47
4.2 Descriptive statistics . . . . .	48

5	Determinants of ethnic identity : whole sample . . . . .	50
6	National identity and assimilation : first and second generations . . . . .	55
6.1	Differences between second and first generations . . . . .	55
6.2	Key differences between national identity and assimilation . . . . .	59
7	Conclusion . . . . .	60
8	Appendix . . . . .	61
<b>3</b>	<b>Segregation and Trust : An experiment in French secondary schools</b>	<b>65</b>
1	Introduction . . . . .	66
2	Litterature review . . . . .	67
2.1	Measuring trust . . . . .	67
2.2	Trust determinants . . . . .	69
3	Design and Background Information . . . . .	70
3.1	Background information . . . . .	70
3.2	Experimental procedure . . . . .	74
4	Results . . . . .	77
4.1	General results . . . . .	77
	Players A . . . . .	77
	Players B . . . . .	79
	Expected and actual returns . . . . .	79
4.2	In-group favoritism ? . . . . .	80
	Players A. . . . .	81
	Players B. . . . .	81
5	concluding remarks . . . . .	82
6	Appendix . . . . .	83
	<b>General Conclusion</b>	<b>97</b>
	 <b>Références</b>	 <b>101</b>
	 <b>List of Figures</b>	 <b>101</b>
	 <b>List of Tables</b>	 <b>103</b>

# Introduction Générale

Les questions d'identité et d'intégration sont au cœur du débat public en France. De nombreuses polémiques ont vu le jour concernant les phénomènes identitaires qui peuvent conduire à la séparation des individus et entraîner des situations de ségrégation. Il n'est donc pas étonnant que les chercheurs en sciences sociales, et notamment les économistes, s'emparent de ces questions. Cette thèse porte sur l'étude des phénomènes de ségrégation urbaine, c'est à dire l'isolement à la fois social et spatial des groupes d'individus au sein des villes. L'objet de cette thèse est à l'intersection de l'étude de la ségrégation urbaine et de la prise en compte de l'identité dans la compréhension des phénomènes et interactions économiques.

## 1 Etat des lieux de la ségrégation en France

Du latin *segregare* qui signifie écarter du troupeau le terme de ségrégation implique l'idée d'une mise à l'écart d'un ou plusieurs individus. Dans le langage courant actuel la ségrégation décrit une "organisation de l'espace en zones à forte homogénéité sociale interne et à fortes disparités sociales entre elles" ([Castells \(1972\)](#)). La ségrégation peut être institutionnalisée ou le résultat de multiples décisions individuelles. Cette division peut se réaliser sur la base de la catégorie sociale, du revenu ou de l'origine (c'est par exemple le cas des migrants). Analyser les phénomènes de ségrégation nécessite d'identifier les groupes sociaux sur lesquels le processus de division s'appuie. Cette question a été étudiée par les sciences sociales. De nombreux travaux se sont attachés à étudier les causes, les conséquences et à mesurer l'ampleur du phénomène de ségrégation. Par exemple, les travaux des sociologues de l'école de sociologie de Chicago, datant du début du XXème siècle, se sont centrées sur les questions de ségrégation ethnique, à savoir la division des populations afro-américaines et blanches dans les grandes villes américaines.

Le phénomène de ségrégation n'est pas un fait nouveau, au contraire, depuis que les villes existent la différenciation socio-spatiale de leurs quartiers semble bien être la norme, comme le souligne la citation de Platon en préambule de cette thèse. Les exemples les

plus connus de ségrégation sont le cas de l'apartheid mise en place en Afrique du sud à partir de 1948 ou bien encore de la ségrégation raciale aux Etats Unis faisant suite à la guerre de sécession. En France, l'exemple de phénomène de ségrégation le plus criant est celui des banlieues où se concentrent les difficultés économiques et sociales. Si la ségrégation a parfois été institutionnalisée, elle résulte aujourd'hui le plus souvent de l'interaction de choix individuels. [Préteceille \(2006\)](#) qui étudie la ségrégation selon les classes sociales montre que les classes sociales les plus privilégiées sont plus ségréguées que les classes sociales populaires. L'existence, réelle ou supposée, d'externalités de voisinage conduit les individus à développer des stratégies de localisation et à choisir soigneusement leur lieu de résidence. L'éducation peut être un moteur des stratégies d'évitement. En effet, les parents recherchent pour leurs enfants l'environnement social et éducatif le plus favorable à leur éducation et leur développement : on recherche les "meilleures" écoles ou celles qui ont la meilleure réputation, de même que des quartiers les plus calmes ou peuplés de familles favorisées. L'enfant pourra alors profiter d'un effet de pair ou d'entraînement positif (complémentarité stratégique en termes économiques). Les catégories sociales favorisées fuient alors les catégories sociales moyennes, elles-mêmes fuyant les familles les plus défavorisées. Ceci a bien été renseigné par la littérature économique. Dans son ouvrage « le ghetto français », [Maurin \(2004\)](#) étudie les phénomènes de ségrégation résidentielle en France. Il montre que les phénomènes de ségrégation ne se limitent pas aux banlieues enclavées mais au contraire correspondent à un phénomène qui touche toutes les classes sociales par des mécanismes de recherche de l'entre-soi, notamment il met en lumière l'existence de ghettos riches où les classes plus aisées de la société se regroupent. Par ailleurs, [Benabou \(1991\)](#) montre que les externalités locales d'éducation, à travers les effets de pairs, poussent les travailleurs qualifiés à ne pas résider à côté de travailleurs non qualifiés. Ces stratégies d'évitement sont renforcées par l'existence de la carte scolaire qui affecte les élèves dans une école selon leur lieu de domiciliation ([Fack & Grenet \(2011\)](#)). D'autre part, [Schelling \(1971\)](#) montre qu'un équilibre ségrégué peut se produire même si individuellement les agents ont des préférences pour la mixité. Basant son analyse sur un échiquier avec des pions de deux couleurs différentes représentant des individus, [Schelling \(1971\)](#) montre encore qu'un équilibre ségrégué se produit lorsque les agents décident de bouger s'ils deviennent minoritaires. Dès lors on atteint une situation de ségrégation alors que les agents n'ont a priori pas de préférence pour la ségrégation.

En France le phénomène de ségrégation est au cœur du débat public depuis plus de vingt ans et l'objectif de mixité sociale est poursuivi à travers l'adoption de politiques de la ville qui ciblent des territoires prioritaires. Cet objectif de mixité s'est notamment traduit par l'adoption de la loi Solidarité et Renouvellement Urbain (SRU) en 2000. Cette

loi impose en outre aux communes importantes de disposer d'au moins 20% de logements sociaux. La mise en place de ces politiques publique implique de disposer d'outils de mesure du phénomène de ségrégation afin d'en caractériser l'ampleur et l'évolution. En France cependant ces études restent rares, elles sont notamment marquées par le biais américain et s'intéressent à la mesure de la ségrégation catégorielle. Certaines études se sont par exemple intéressées à la ségrégation ethnique des immigrés en France ([Verdugo \(2011\)](#), [Safi \(2009\)](#), [Pan Ké Shon \(2010\)](#)), elles montrent que les populations immigrées faisant le plus l'objet d'une ségrégation importante sont les maghrébins, les africains subsahariens, les turcs et les asiatiques. Elles montrent également que sur une période longue (1968-1999) la ségrégation a diminué pour les immigrés.

L'une des questions majeures qui se pose lorsque l'on étudie le phénomène de ségrégation urbaine est donc celle de sa mesure. En outre, la littérature existante sur la ségrégation urbaine propose de nombreux indices mesurant le degré de ségrégation catégorielle mais peu d'études cependant proposent une mesure de la ségrégation sur une variable continue. Le chapitre 1 revient sur la définition de ségrégation et étudie les propriétés générales d'une famille d'indice de ségrégation basée sur la décomposition d'indices d'inégalité. Nous étudions un indice de ségrégation basé sur le revenu le Neighborhood Sorting Index (NSI), introduit par [Jargowsky \(1996\)](#), ainsi que ses interprétations intuitives. Cet indice permet de mesurer le degré de ségrégation économique des individus en fournissant une mesure d'une variable continue. Il est défini comme la racine carrée du ratio de la variance des revenus moyens entre les quartiers sur la variance des revenus des individus dans la ville. L'utilisation du NSI est ensuite illustrée par une application empirique qui mesure et compare le niveau de ségrégation économique en France dans les trente plus grandes aires urbaines françaises entre 2001 et 2009.

## 2 Identité

L'intérêt des économistes pour les questions d'identité est récent et croissant. La notion d'identité peut être difficile à saisir. Elle est multiple : l'identité sociale, l'identité de genre, l'identité ethnique ou l'identité nationale. Il est par conséquent difficile de mesurer cette variable. Dans leur article fondateur, [Akerlof & Kranton \(2000\)](#) tentent de proposer une définition claire de l'identité, ils la définissent comme le sentiment que l'on a de soi et proposent d'intégrer cette variable à la fonction d'utilité.

L'identité peut avoir un effet important sur les choix des individus et peut par conséquent impacter leurs résultats socio-économiques tels que leur niveau de salaire, leur accès à l'emploi etc. L'identité est un concept qu'il est difficile de mesurer, de plus en plus étudiée dans le contexte de la migration on s'intéresse souvent à l'identité ethnique. En migrant, les individus peuvent faire face à un arbitrage entre adopter la culture d'accueil

ou préserver leur héritage culturel. Le résultat de cet arbitrage peut se traduire sous 4 formes d'acculturation définies par [Berry \(1997\)](#) : l'intégration où les individus adoptent la culture d'accueil tout en préservant leur héritage culturel, l'assimilation où les individus adoptent la culture d'accueil en s'éloignant de leur culture d'origine, la séparation où les individus maintiennent leur héritage culturel et n'adoptent pas la culture d'accueil ou la marginalisation où les individus s'écartent à la fois de leur culture d'origine et de la culture d'accueil. Le résultat de cet arbitrage peut avoir des effets importants sur les résultats socio-économiques des individus ([Pendakur & Pendakur \(2005\)](#), [Nekby & Rodin \(2007\)](#), [Manning & Roy \(2010\)](#), [Battu & Zenou \(2010\)](#)). Il convient donc de développer des outils pour mesurer cette distance identitaire, jusqu'ici l'identité ethnique est souvent approchée par la réponse à la question « je me sens [français, marocain, italien...] », qui est également appelée identité nationale. D'autres mesures approchent l'identité à travers la construction d'indices basés sur les caractéristiques identitaires utilisées pour se définir (langue, couleur de peau, origines...) à travers des analyses factorielles. Ces mesures de l'identité ethnique peuvent s'apparenter à la ségrégation dans la mesure où elles capturent une distance au groupe majoritaire.

Dans le chapitre 2 nous proposons une mesure de l'identité ethnique. La notion d'identité ethnique est le plus souvent approchée empiriquement par une variable binaire traduisant le sentiment national (la réponse à la question : je me sens [français/ italien/ Marocain ...]). Ce chapitre propose une méthodologie permettant d'étudier le degré d'assimilation des immigrés en France, ainsi que des descendants d'immigrés. A travers l'analyse de données originales (fournies par la base Trajectoires et Origines) nous construisons un indice qui permet une mesure continue de l'identité ethnique. L'indice que nous construisons est basé sur une analyse factorielle qui permet de prendre en compte le caractère multi-dimensionnel de l'identité en analysant les caractéristiques mentionnées par les individus pour se définir. Nous proposons une comparaison de cet indice avec d'autres mesures de l'identité existant dans la littérature (notamment la variable binaire "je me sens français") et soulignons leur caractère complémentaire. En outre, nous montrons que l'assimilation et le sentiment d'être français ne coïncident pas toujours. Il semble que plus le pays d'origine est distant, à la fois en termes économiques et culturels, plus le sentiment d'être français est fort et plus le degré d'assimilation est faible. Nous étudions ensuite les déterminants de l'identité à travers l'analyse de certaines caractéristiques socio-démographiques. Nous montrons également que l'identité des enfants d'immigrés converge vers celle des natifs, que celle-ci soit approchée par le sentiment national ou le degré d'assimilation.

D'autre part, les phénomènes de ségrégation sont liés à l'existence de groupes sociaux. S'intéressant à l'identité sociale, [Tajfel & Turner \(1979\)](#) développent le paradigme du groupe minimal selon lequel les individus privilégient les membres de leurs groupes,

même si ces groupes sont artificiellement créés. Ces auteurs ont mené une étude auprès d'adolescents où des groupes avaient été créés sur la base de critères arbitraires. Le principal résultat est que les sujets discriminaient positivement les membres de leurs groupes. Le fait que les individus discriminent selon les groupes sociaux peut avoir d'importantes répercussions économiques et sociales, non seulement sur le plan individuel mais également à l'échelle des quartiers d'une ville. C'est pourquoi les questions d'identité représentent une question majeure de politiques publiques. Dans une étude menée dans la ville de Zurich, [Falk & Zehnder \(2007\)](#) ont montré qu'ils existaient des phénomènes de discrimination selon le quartier de résidence des individus. Au sein de cette ville, les participants montraient des niveaux de confiance plus faibles envers les individus résidant dans des quartiers plus ségrégués. Ce résultat implique des conséquences importantes en termes de politiques publiques, notamment les individus peuvent être discriminés dans l'accès à l'emploi simplement parce qu'ils habitent des quartiers différents dans la ville. De plus, cette absence de confiance menace les interactions socio-économiques et peut contribuer à accélérer les processus de ségrégation dans la ville.

Dans le chapitre 3 nous nous intéressons à ces questions et nous étudions les effets de la ségrégation urbaine sur les préférences des individus, notamment sur le degré de confiance aux autres. Ce chapitre vise à étudier les liens existants entre l'environnement social des individus, notamment leur quartier de résidence, et leur degré de coopération et de confiance aux autres. Nous avons menée une étude expérimentale dans plusieurs collèges rennais en proposant à des collégiens de troisième de participer à des jeux de confiance. Nous nous intéressons à des collégiens afin de réduire les problèmes d'endogénéité possibles dans la mesure où les adolescents ne choisissent pas leur lieu de résidence, celui-ci ne dépend donc pas de leurs préférences individuelles. Ces expériences visent à répondre aux questions suivantes : – Comment le contexte social, et notamment le degré de mixité sociale, agit sur le niveau de confiance des individus ? – Le degré de confiance en l'autre est-il influencé par l'identité de son partenaire (en outre, si celui-ci vient du même collège que lui ou non)





# Chapitre 1

## Income segregation in large French cities

In this chapter, we examine the general properties of the family of segregation measures, based on the ratio between the income inequality between spatial areas and the overall inequality, and their intuitive interpretation. The use of two measures of segregation belonging to this family, i.e., the Neighbourhood Sorting Index (NSI) introduced by [Jargowsky \(1996\)](#) and the Gini Segregation Index (GSI), is illustrated measuring and comparing the residential segregation by income in the 30 largest French urban areas from 2000 to 2008.

### 1 Introduction

Sociologists as well as economists have focused more attention in recent years on important effects of segregation, arguing that this residential space crystallizes interactions that influence individual preferences, skills, children's attitude or the choice of school (see, for instance, [Cutler & Glaeser \(1995\)](#), [Cutler et al. \(2008\)](#), [Echenique et al. \(2006\)](#), [Goux & Maurin \(2007\)](#)).

The literature dealing with measures of segregation developed many indicators in order to undertake analyses of *categorical segregation*, that is, the distribution of people across categories (see for instance [Duncan & Duncan \(1955\)](#), [Massey & Denton \(1988\)](#), [Reardon & O'Sullivan \(2004\)](#), [R. M. Hutchens \(1991\)](#), [R. Hutchens \(2001\)](#), [R. Hutchens \(2004\)](#) or [Chakravarty & Silber \(2007\)](#)). The occupational segregation of men and women and the residential segregation of white and black population in cities are the most popular examples of categorical segregation.

Surprisingly, *residential segregation by income* have not been extensively studied by the literature (notable exceptions are [Hardman & Ioannides \(2004\)](#), [Jargowsky \(1996\)](#), [Ioannides & Seslen \(2002\)](#), [Hardman & Ioannides \(2004\)](#), [Davidoff \(2005\)](#) or [Watson \(2009\)](#)). By residential segregation by income, we mean the way by which individuals who can be described by their income are broken down among *local areas* of a city. [Jargowsky \(1996\)](#) proposed a measure of economic segregation, defined as *pure*. This measure, which Jargowsky referred to, is the Neighbourhood Sorting Index (NSI), simply defined as the square root of the ratio of the variance of neighbourhoods' mean incomes over the overall variance. Defining the information theory of segregation, [Jargowsky & Kim \(2005\)](#), stated that all measures of inequality and all measures of segregation are united within a broader and single framework. This framework helps them to define new measures of segregation and to implement well known measures to continuous variables. Furthermore, they derive several properties that should be satisfied by the common measures of inequality and segregation. The inequality ratio property states that there exists a segregation measure for any inequality index, which is based on the ratio of the income inequality measured at the group level and the overall income inequality. The pure measure property states that a measure of segregation must satisfy the inequality ratio property to be valid. The one-to-one relationship between inequality and segregation measures suggests that "a measure of inequality that can be applied to either binary or continuous variables generates a measure of segregation that can be applied to either binary or continuous variables."

The aim of this chapter is to examine theoretically and empirically the family of segregation measures introduced by [Jargowsky & Kim \(2009\)](#), which allows us to compare cities on the basis of the breakdown of individuals who differ in terms of income among a given set of spatial areas. We propose an approach of measuring segregation by income based on *transformations* of the breakdown of a population among local areas. These transformations capture two dimensions of the segregative phenomenon. First, change in segregation may be the result of changes in the "allocation" of individuals among areas while the income distribution among the individuals remains the same. Second, changes in the income distribution (e.g., progressive or regressive transfer) among immobile individuals result in segregation change.

We first study the sensitiveness of the family of segregation measures considered with respect to movements of individual(s). In particular, we argue that a segregation measure should decrease if a rich individual moves from an area - where she is richer than the average individual - to a poorer one *or* if a poor individual moves from an area - where she is poorer than the average individual - to a richer one. Moreover, we consider the impact of a switch of two individuals, which requires that a segregation measure decreases if a

rich individual living in a rich area and a poor individual living in a poor area exchange their location.

The second type of transformations considered leaves the breakdown of individuals unchanged but affects the income distribution among individuals. We study the sensitivity of the segregation measures relatively to an income transfer between two individuals located in the *same* local area. Any income transfer which is *progressive* in the usual sense of Pigou-Dalton contributes to *increase* segregation. The rationale of this requirement lies to the fact that such a transfer increases the homogeneity of the area and hence lowers its social mixing. On the contrary, an income transfer from a poor to a rich individual decreases segregation because it reduces the homogeneity of the local area.

This family of segregation indices is then illustrated measuring and comparing the segregation in the 30 largest French urban areas from 2000 to 2008. More precisely we use the Neighbourhood Sorting Index (NSI) as defined by Jargowsky (1996), and a segregation index based on the Gini index, namely the Gini Segregation Index (GSI). We perform this analysis using households' income and their distribution within the residential space from several scales : *IRIS* and *Grand Quartier*. The main result is that residential segregation by income has decreased over the period considered. Furthermore, we show that the economic segregation in French cities is scale dependent and is not related with income inequality as measured by Gini coefficient. This performed empirical investigation proves that these measures are very appealing for applied analysis : it does not demand prohibitive dataset and it can draw a picture that is somewhat different from the one depicted by inequality measures.

The remainder of the chapter is structured as follows. The next section is devoted to the notion of residential segregation by income. The following section defines formally the family of segregation measures and in particular the NSI as well as the GSI, and explores the properties satisfied by this family of segregation measures. The use of these two measures as part of the measurement of the segregation of large French cities is undertaken in section 4. The final part concludes.

## 2 Measuring residential segregation by income

We consider *cities* where each individual is endowed with income and is located in a given area. A city is populated by  $n$  individuals from some finite set  $N = \{1, \dots, i, \dots, n\}$ . Let's assume also that this city is made up of  $m$  geographical areas (or, more generally, of  $m$  subgroups) and populated by  $n^j$  individuals, and define  $M$  the finite set of areas of the city,  $M = \{1, \dots, j, \dots, m\}$ . This city can be defined by two elements : (1) a distribution of income among the population and (2) a partition of the population between the  $m$

geographical areas of the city. Assuming that incomes are drawn from an interval  $D$  of  $\mathbb{R}$ , the overall income distribution of the population can be described by a vector :

$$Y = (y_1, \dots, y_i, \dots, y_n)$$

where  $y_i$  is the income of the individual  $i$ . The breakdown of individuals into a set of areas is described by a partition  $P$  of the population within a set of  $m$  mutually exclusive and exhaustive subsets (areas). We use  $\mathcal{P}$  to represent the set of all possible partitions,  $P = \{N^1, \dots, N^m\}$ , where  $N^j$  is the set of individuals living in area  $j$ . The income distribution within area  $j \in M$  is given by the vector  $Y^j = (y_1^j, \dots, y_i^j, \dots, y_{n^j}^j)$  where  $y_i^j$  is the income of individual  $i$  living in area  $j$ . A city can be thus depicted by a pair  $(P, Y) \in \mathcal{P} \times D^n$ .

By simplicity, we can describe the overall income distribution of the population as :

$$Y = (Y^1, \dots, Y^m) = (y_1^1, \dots, y_{n^1}^1, y_1^2, \dots, y_{n^m}^m)$$

Let  $\mu$  denote the mean income of the city and  $\mu^j$  the mean income of the area  $j$ . The index studied in this chapter is based on the decomposition of an inequality index between within-group and between-group components. Jargowsky & Kim (2009) argue that there exists a segregation measure for any inequality index and this measure is based on the ratio of the income inequality measured at the group level and the overall income inequality (the *inequality ratio property*). For an inequality index  $\mathcal{I} : D^n \rightarrow \mathbb{R}$ , a segregation index, denoted by  $S_{\mathcal{I}} : \mathcal{P} \times D^n \rightarrow \mathbb{R}$ , can be thus defined by :

$$S_{\mathcal{I}}(P, Y) = \frac{\mathcal{I}(Y_{\mu})}{\mathcal{I}(Y)} \quad (1.1)$$

where  $Y_{\mu}$  is the hypothetical income distribution such that each individual receives the mean income of his area, that is,  $Y_{\mu} = (\mu^1 \mathbf{1}^{n^1}, \dots, \mu^m \mathbf{1}^{n^m})$  where  $\mathbf{1}^{n^j}$  is the vector of dimension  $n^j \times 1$  where each element is equal to 1. We may require that the inequality index we select respects some desirable properties : symmetry, replication invariance (i.e., Dalton population principle) and Pigou-Dalton principle of progressive transfers. Subgroup decomposability property<sup>1</sup> is not necessary for segregation index to respect some desirable properties. However the property of the inequality index is useful if we want that  $S_{\mathcal{I}}(P, Y)$  values belong to the interval  $[0, 1]$ .

**Proposition 1.1** (Normalisation). :

- i. If the distribution of mean incomes  $Y_{\mu}$  is of the form  $Y_{\mu} = (c \mathbf{1}^{n^1}, \dots, c \mathbf{1}^{n^m})$  where  $c$  is a scalar, then  $S_{\mathcal{I}}(P, Y) = 0$ .

---

1. Subgroup decomposability imposes that the inequality index can be written as follows :  $\mathcal{I}(Y) = \sum_j f[\mathcal{I}(Y^j)] + \mathcal{I}(Y_{\mu})$ .

- ii. Assume that  $\mathcal{I}$  is additively decomposable by subgroups. If the distribution of mean incomes  $Y_\mu$  is such that  $y_i^p = \mu^p$  for all  $i \in N^p$  and  $\mu^p \neq \mu^q$  for  $p \neq q \in M$ , then  $S_{\mathcal{I}}(P, Y) = 1$ .

## 2.1 Segregation under movements

The starting point of our analysis is the evaluation of segregation when the only change is movement(s) of individual(s) among areas given that there is no change in the income distribution. In other words how can we reduce segregation (or improve social mixing) only by movements of individuals within the city? Consider the following examples.

**Example 1.** Consider the following income distribution :  $Y = (Y^1, Y^2) = (1, 2, 3; 8, 9, 10)$ . The richest individual of area 2 moves to area 1; the new situation is then :  $\tilde{Y} = (\tilde{Y}^1, \tilde{Y}^2) = (1, 2, 3, 10; 8, 9)$ . Our intuition is that  $\tilde{Y}$  is less segregated than  $Y$ .

**Example 2.** Consider now that the richest individual of area 2 moves to area 1 while one individual changes also his location. The new situation is then :  $\hat{Y} = (\hat{Y}^1, \hat{Y}^2) = (1, 2, 10; 3, 8, 9)$ . According to our intuition such movements reduce segregation.

The simplest transformation illustrated by Example 1 is the unilateral movement of one individual from one area to another one.

**Definition 1.2. Movement of one individual.** Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities and  $k$  an individual located in area  $p$ . The city  $(\tilde{P}, Y)$  is obtained from the city  $(P, Y)$  by means of movement of one individual if there exists an area  $q$  such that :

- (a)  $\tilde{N}^p = N^p / \{k\}$
- (b)  $\tilde{N}^q = N^q \cup \{k\}$
- (c)  $Y^\ell = Y^\ell$  for all  $\ell \neq p, q$

We might rewrite conditions (a) and (b) of the previous definition in terms of areas' income distributions. These conditions are equivalent, respectively, to :

- (a')  $\tilde{Y}^p = (y_1^p, \dots, y_{k-1}^p, y_{k+1}^p, \dots, y_{n^p}^p)$
- (b')  $\tilde{Y}^q = (y_1^q, \dots, y_{n^q}^q, y_k^p)$

The following proposition describes the link between unilateral movement and change in segregation as measured by our family of indices :

**Proposition 1.3** (Segregation under unilateral movement). :

Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities such that  $(\tilde{P}, Y)$  is obtained from  $(P, Y)$  by a movement of individual  $k$  from area  $p$  to  $q$ . Assuming that  $\mathcal{I}$  respects the Pigou-Dalton principle of progressive transfer, the following cases can be considered :

- i. If  $y_k^p > \mu^p > \mu^q$  or  $\mu^q > \mu^p > y_k^p$ , then :  $S_{\mathcal{I}}(P, Y) > S_{\mathcal{I}}(\tilde{P}, Y)$ .
- ii. If  $y_k^p > \mu^q > \mu^p$  or  $\mu^p > \mu^q > y_k^p$ , then :  $S_{\mathcal{I}}(P, Y) < S_{\mathcal{I}}(\tilde{P}, Y)$ .

(See appendix for the proof).

The intuition behind this proposition is quite obvious. Any emigration of an individual who is richer than the mean incomes of his initial area and of his new one reduces the gap between mean incomes. Inequality between areas is then lower. We can see by the same token that any emigration of a poor individual from a poor to a rich area reduces also between-areas inequality.

Another transformation is the switch of two individuals, which consists in two simultaneous movements of a rich and a poor individuals.

**Definition 1.4** (Switch of two individuals). :

Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities, within which two individuals  $k$  and  $h$  are located in area  $p$  and  $q$  respectively. The city  $(\tilde{P}, Y)$  is obtained from the city  $(P, Y)$  by means of a switch of two individuals if :

- (a)  $\tilde{N}^p = N_{/\{k\}}^p \cup \{h\}$
- (b)  $\tilde{N}^q = N_{/\{h\}}^q \cup \{k\}$

This transformation can also be interpreted as an exchange of income between two individuals who are located in two different areas and can thus be expressed in terms of areas' income distributions. The two conditions of the definition 1.4 might be substituted by the following conditions :

- (a')  $\tilde{y}_k^p = y_h^q$
- (b')  $\tilde{y}_h^q = y_k^p$
- (c')  $\tilde{y}_i^j = y_i^j$  for  $j = p, q$  and  $i \neq k, h$
- (d')  $\tilde{y}_i^j = y_i^j$  for all  $j \neq p, q$  and  $i = \{1, \dots, n^j\}$

The following proposition describes how a switch affects our segregation measure.

**Proposition 1.5** (Segregation under switch). :

Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities such that  $(\tilde{P}, Y)$  is obtained from  $(P, Y)$  by a switch of two individuals  $k$  and  $h$  between areas  $p$  and  $q$ . Assuming that  $\mathcal{I}$  respects the Pigou-Dalton principle of progressive transfer, the following cases can be considered :

- i. If  $y_k^p < y_h^q$  and  $\mu^p < \mu^q$ , then :  $S_{\mathcal{I}}(P, Y) > S_{\mathcal{I}}(\tilde{P}, Y)$ .
- ii. If  $y_k^p < y_h^q$  and  $\mu^p > \mu^q$ , then :  $S_{\mathcal{I}}(P, Y) < S_{\mathcal{I}}(\tilde{P}, Y)$ .

iii. In other cases, such a switch has no clear implication.

(See appendix for the proof).

A switch of two individuals will decrease our segregation measure if the richest individual is originally located in the richest area and the poorest one in the poorest area. This bilateral movement results in a transfer of mean income from the richest to the poorest area. Since our measure of segregation is based on an inequality index respecting the principle of progressive transfer,  $\mathcal{I}(Y_\mu)$  will decrease and segregation measure  $S_{\mathcal{I}}$  will also decrease. The same type of reasoning can apply when the richest individual is originally located in the poorest area and the poorest one in the richest area.

## 2.2 Segregation under income transfers

We consider now the impact of a transformation of the income distribution on the NSI given that the partition of the population between areas remains the same. In the measurement of income inequality, an elementary transformation is the Pigou-Dalton principle of transfer which states that a transfer of income from a rich individual to a poorer one reduces inequality.

**Example 3.** Consider the following income distribution :  $Y = (Y^1, Y^2) = (10, 15, 21 ; 17, 25, 30)$ .

Transferring income between two individuals in the richest area modifies the income distribution as follows :  $(10, 15, 21 ; 22, 25, 25)$ . While inequality decreases, the richest area becomes more homogeneous in terms of income and the overlapping of areas' income distributions disappears. Indeed, the richest individual of the poorest area is now poorer than the poorest individual of the richest area. Our intuition is that  $\tilde{Y}$  is less segregated than  $Y$ .

**Example 4.** Weigh up now the case of a regressive transfer in the richest area that leads to the following modified income distribution :  $(10, 15, 21 ; 12, 30, 30)$ . The final situation could be evaluated as better in terms of segregation since the social mixing has been increased and the overlapping of the distribution is higher.

A progressive transfer may then be not so relevant in the assessment of economic segregation of a city all the more so as such transfer occurs within the same area. We study the case of an income transfer between two individuals located in the *same* area, namely a *within-area income transfer*.

**Definition 1.6** (Within-area income transfer). :

Let  $(P, Y)$  and  $(P, \tilde{Y})$  two cities and two individuals  $k$  and  $h$  located in area  $p$ . The city  $(P, \tilde{Y})$  is obtained from the city  $(P, Y)$  by means of within-area income transfer if, for any  $\delta \in \mathbb{R}_+$  :



- (a)  $\tilde{y}_h^p = y_h^p + \delta$  and  $\tilde{y}_k^p = y_k^p - \delta$   
 (b)  $\tilde{y}_i^j = y_i^j$  for  $j = p$  and  $i \neq k, h$  and  $\tilde{y}_i^j = y_i^j$  for all  $j \neq p$  and  $i \in N^j$

The question we study now is to ascertain conditions for which a transfer reduces or worsens segregation. A progressive transfer reduces discrepancy of income in a given area *and* in the overall city but inequality between mean incomes is not affected by such an income transfer. We resume in the following proposition the main implications of a within-area income transfer when the initial level of segregation is not zero.

**Proposition 1.7** (Internal transfer). :

Let  $(P, Y)$  and  $(P, \tilde{Y})$  two cities such that  $S_{\mathcal{I}}(P, Y) \neq 0$ . Assuming that  $\mathcal{I}$  respects the Pigou-Dalton principle of progressive transfer, the following cases can be considered :

- i. If  $(P, \tilde{Y})$  is obtained from  $(P, Y)$  by a finite set of within-area progressive income transfers (i.e.,  $\delta > 0$  and  $y_h^p < \tilde{y}_h^p \leq \tilde{y}_k^p < y_k^p$ ), then  $S_{\mathcal{I}}(P, Y) < S_{\mathcal{I}}(P, \tilde{Y})$ .
- ii. If  $(P, \tilde{Y})$  is obtained from  $(P, Y)$  by a finite set of within-area regressive income transfers (i.e.,  $\delta < 0$  and  $\tilde{y}_h^p < y_h^p \leq y_k^p < \tilde{y}_k^p$ ), then  $S_{\mathcal{I}}(P, Y) > S_{\mathcal{I}}(P, \tilde{Y})$ .

(See appendix for the proof).

This proposition says that increasing overall inequality among individuals reduces segregation if increasing overall inequality does not lead to increase inequality among areas mean incomes. That illustrates the conflict, which may exist between segregation by income and income inequality. Another interpretation is that segregation is evaluated in relation to overall inequality. More precisely, segregation is inequality among areas once overall inequality is neutralized. From a “pure” perspective, segregation is a problem of allocation of individuals among areas *given* the income distribution. Such an approach is however unable to compare cities with different levels of inequality (as measured by an appropriate index). To illustrate this point, consider the following cities :  $Y = (1, 2, 3; 3, 4, 5)$  and  $\tilde{Y} = (2, 2, 2; 3, 4, 5)$ . It is not possible to derive  $Y$  from  $\tilde{Y}$  by any movement of individual(s) between the two areas ; the difference between  $Y$  and  $\tilde{Y}$  lies in different income distributions.

Finally it is worth noting that the value of segregation measure  $S_{\mathcal{I}}$  remains the same if the city is initially perfectly mixed ( $S_{\mathcal{I}} = 0$ ). If all individuals have the same income (i.e.,  $\mathcal{I}(Y) = 0$  and thus  $\mathcal{I}(Y_\mu) = 0$ ), it is obvious that we are not able to perform progressive income transfer and a regressive transfer leaves unchanged  $\mathcal{I}(Y_\mu)$ . If all areas have the same mean income (i.e.,  $\mathcal{I}(Y_\mu) = 0$ ), there is no segregation as measured by  $S_{\mathcal{I}}$  and, then, internal transfer leaves segregation unchanged. Consider the following city :  $(1, 1, 10; 4, 4, 4)$  for which  $S_{\mathcal{I}} = 0$  since  $\mathcal{I}(Y_\mu) = 0$ . If we transfer income from the rich individual to one of the two poor ones in the first area, the  $S_{\mathcal{I}}$  value does not increase

since the inequality between areas remains the same. The same observation can be made when we consider a transfer from a poor to the rich in the first area. It means that the change in the  $S_{\mathcal{I}}$  value following a within-area transfer depend on the gap between the mean incomes. We may undoubtedly consider this type of dis-continuity as a theoretical weakness of our segregation measure (1.1). However from empirical perspective, it may not be so problematic since the probability of  $S_{\mathcal{I}} = 0$  is certainly very low.

### 2.3 Sensitivity to areas definition

The previous properties presented above assume spatial partition as given : it may be purely arbitrary or conversely based on geographical or urban meaningful characteristics. We now study how segregation index (1.1) is sensitive to the spatial partition of a given urban area. To illustrate the problem, let consider the following example.

**Example 5** Consider the following income distribution :  $Y = (Y^1, Y^2, Y^3) = (1, 2, 3 ; 5, 6, 7 ; 8, 9, 10)$ . Consider that the first and third areas merge :  $\tilde{Y} = (\tilde{Y}^1, \tilde{Y}^2) = (1, 2, 3, 8, 9, 10 ; 5, 6, 7)$ . It is intuitive to say that there is more social mixing or less segregation in  $\tilde{Y}$  than in  $Y$  since poor people mix with rich people in the new first area.

Jargowsky pointed out that the NSI does not capture the important physical dimensions of segregation. With the next property, we investigate the variability of the NSI with respect to the definition of area unit. Aggregation of areas consists in merging two areas into a unique area.

#### Definition 1.8. Areas aggregation.

Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities such that  $(\tilde{P}, Y)$  is obtained from  $(P, Y)$  by means of an aggregation of areas  $p$  and  $q$  if for  $P = \{N^1, \dots, N^p, \dots, N^q, \dots, N^m\}$  and  $\tilde{P} = \{\tilde{N}^1, \dots, \tilde{N}^{m-1}\}$  :

- (a)  $N^j = \tilde{N}^j \ \forall j < q, j \neq p$  and  $N^j = \tilde{N}^{j-1} \ \forall j > q$
- (b)  $N^p \cup N^q = \tilde{N}^p$

The next proposition considers sensitivity of our segregation measure to aggregation of areas : If merged areas have the same mean income, there is no change in segregation according to the  $S_{\mathcal{I}}$ . However, if one merges two areas with different mean incomes, the aggregation increases income heterogeneity within areas and decreases the variability in mean incomes.

#### Proposition 1.9 (Sensitivity to areas aggregation). :

*Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities such that  $(\tilde{P}, Y)$  obtained from  $(P, Y)$  by means of areas aggregation, then :*

(a)  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(\tilde{P}, Y)$  if  $\mu^p = \mu^q$

(b)  $S_{\mathcal{I}}(P, Y) > S_{\mathcal{I}}(\tilde{P}, Y)$  if  $\mu^p \neq \mu^q$

(See *appendix for the proof*).

In other words, this property states that the  $S_{\mathcal{I}}$  does not decrease when the number of areas increases; because the inequality between mean incomes can not decrease.<sup>2</sup> As [Shorrocks & Wan \(2005\)](#) have showed, the expected value of between-component of any inequality measure increases with the number of areas.

Even though the  $S_{\mathcal{I}}$  fails to account for spatial patterning of areas<sup>3</sup>, this property might help to give a picture of spatial segregation. Indeed, it is sensitive to the definition of the boundaries of areas and assumes that each individual lives near all individuals of his area and far from individuals located in another area (even across the street from one another). Consider two sets of areas' boundaries,  $P$  and  $\tilde{P}$ , such that  $\tilde{P}$  is obtained from  $P$  by a finite set of *adjacent* areas aggregation. In other words, we consider two partitions of individuals among areas,  $P = \{N^1, \dots, N^m\}$  and  $\tilde{P} = \{\tilde{N}^1, \dots, \tilde{N}^\ell\}$ , such that  $\tilde{N}^j$  is the union of two or more *adjacent* (sub)areas  $N^k$  (with  $k \in \{1, \dots, m\}$ ) for all  $j \in \{1, \dots, \ell\}$ . If the segregation index takes the same value whatever the partition  $P$  and  $\tilde{P}$ , it means that areas are adjacent to similar areas (in terms of mean income). On the contrary, the larger difference between the two values taken by the segregation index is, the more areas are heterogeneous. This reasoning holds comparing two cities. Consider two cities with the same value of  $S_{\mathcal{I}}$  for the finest areas division. If the value of  $S_{\mathcal{I}}$  differs according to a different partition, then a city has more homogeneous grouping of sub-areas and is thus more spatially segregated.

### 3 Economic segregation in French urban areas

The residential segregation has been discussed among the social scientists and in the public debate for more than twenty years. By studying changes in social class and income composition of the close neighbourhoods, [Maurin \(2004\)](#) shows that income segregation has been stable over the period from 1991 to 2002. [Préteceille \(2006\)](#) studies the geographical breakdown of social classes in Paris urban area by using the *dissimilarity* index. The author shows that the most privileged social classes are more segregated than the

---

2. One might be interested in considering the case of area division rather than areas aggregation. Obviously, if the division is such that the two new areas have the same mean income, segregation remains the same; because variability in mean incomes is the same before and after the division. But segregation becomes worse if the division of the area allows a sorting of individuals with respect to income.

3. In particular, the *checkerboard problem* and the *modifiable areal unit problem* are ignored by such indices. See for example [White \(1983\)](#), [Reardon & O'Sullivan \(2004\)](#) and [Jargowsky & Kim \(2005\)](#).

*popular* categories. White-collar workers<sup>4</sup> and executive employees<sup>5</sup> tend to be less segregated in 1999 than in 1990. Contrarily, segregation has increased between 1990 and 1999 for the blue-collar workers. Recently, some papers investigate the segregation of immigrants or ethnic segregation (see for instance Verdugo (2011), Safi (2009) or Pan Ké Shon (2010)). Our study complement these empirical investigations by studying how *income* segregation has changed in French urban areas during the 2000s.

Some segregation measures have been used in several papers in order to understand the segregation in the US cities (for instance, Jargowsky (1996), Jargowsky & Kim (2005), Yang & Jargowsky (2006) or Wheeler & La Jeunesse (2006)). Recently, using the census tract level family income data, Watson (2009) shows that income segregation has increased between 1970 and 2000 in 216 US urban areas.<sup>6</sup> Moreover, inequality is found to be positively correlated with the NSI.

### 3.1 Database description

We use the *Revenus Fiscaux Localisés* database, provided by INSEE.<sup>7</sup> This database provides, over a 8 year period (2001-2008) and for each area unit, the mean and the median income, the Gini Inequality coefficient, the quartiles and the deciles. The area unit considered by INSEE is the *IRIS*,<sup>8</sup> defined as an area comprising between 1800 and 5000 inhabitants : *IRIS* are uniform in their habitat type and their borders are based on the large cuts in the urban area, such as main roads, railways, rivers, *etc.* This spatial unit is close to the principle of Tract in the USA. Note that an important characteristic of this database is that INSEE is using exhaustive files providing by the *Direction Générale des Impôts* rather than a sample of inhabitants.

The definition of income adopted for our study is the taxable income, which is established from two different files of the income statement and property tax. INSEE estimates the taxable income for several geographical levels. The *taxable household* is an ordinary household formed by the combination of taxable families listed in the same dwelling. The taxable income is the amount of resources reported by taxpayers on the *income statement*, before any reduction (which is not equivalent to the concept of *disposable income*). Therefore this income variable accounts for wages, unemployment benefices, pensions, capital income and non salaries revenue.

---

4. *Employés* and *professions intermédiaires*.

5. *Cadres* and *professions libérales*.

6. The author uses the *census tract* as the definition of a local area. As such an area is composed by roughly 4,000 people, her results could be compared with ours.

7. Institut National de la Statistique et des Etudes Economiques

8. *IRIS* : Ilots Regroupés pour l'Information Statistique.

The income is expressed in Consumption Unit, which accounts for the size and the structure of household into consideration. Indeed, differences in household structure between areas are sometimes such that the fact of using income per consumption unit offers a different picture of levels and differences in relation to the reasoning per household or per person. This equivalence scale is commonly used by INSEE and Eurostat to study income expressed as *equivalent adult*. For a given household, the first adult counts for one consumption unit, while the remaining persons count for 0.5 consumption unit if they are more than 14 years old, and children (less than 14 years old) count for 0.3 consumption unit.

We calculated the NSI and the GSI for 7 years (2001, 2002, 2004, 2005, 2006, 2007 and 2008) for the 30 largest urban areas<sup>9</sup> and computed it for different geographical levels, testing different scales of spatial areas. We first computed both indices for the *IRIS* level, then for the *Grand Quartier* level, which are less fine partitions of the urban areas. More precisely, a *Grand Quartier* is defined as a grouping of several adjoining *IRIS* inside a city. The size is variable but respects some population norms. A city with 20 000 inhabitants is generally divided into less than 3 *Grand Quartier* and few cities with less than 10 000 inhabitants are identified as a unique *Grand Quartier*.

Table 1.1 presents some general information related to our database, including, for each city, the population expressed in consumption unit, the number of *IRIS* and *Grand Quartier*, the absolute variation of the number of *IRIS* and *Grand Quartier*, and the variation of the population over the period considered (2001-2008). The last column presents the share of population accounted for, dropping the population belonging to *IRIS* for which at least one piece of information of the *Revenus Fiscaux Localisés* database is not available.

## 3.2 Results

In this section we present our empirical results, which exhibit a slightly decrease in segregation over the considered period. We first compare both segregation indices (GSI and NSI) and show that, though correlated, they imply different rankings of the cities over the period. Furthermore, we perform the inter-city comparison based on their ranking and their values and we discuss the nature of the relationship between income segregation and income inequality.

---

9. We use the concept of *unité urbaine* which is defined by INSEE as a set of districts for which at least 2,000 people is living in a continuous developed area.

City	population	Nb of <i>IRIS</i>	$\Delta$ <i>IRIS</i>	Nb of quartier	$\Delta$ <i>quar-tier</i>	$\Delta$ pop (%)	Pop. Included (%)
Paris	6099092.53	3728.00	24	617.14	12	28.54	93.23
Marseille	766363.26	487.00	3	129.71	1	8.61	85.34
Lyon	694787.91	411.43	6	154.43	2	9.04	77.37
Nice	544977.76	301.29	8	52.00	3	18.61	84.35
Lille	534558.39	367.29	6	113.71	0	13.65	85.42
Bordeaux	439039.29	257.14	5	50.14	0	10.72	86.55
Toulouse	369451.76	212.00	2	72.43	0	9.62	70.57
Toulon	305651.26	180.71	1	26.00	1	4.67	82.51
Nantes	298931.66	166.86	3	35.00	1	14.94	83.41
Strasbourg	228236.56	139.71	3	45.86	0	0.30	85.03
Rouen	217186.77	141.29	7	46.14	2	12.56	87.67
Grenoble	202446.64	131.43	15	25.57	4	13.13	75.11
Tours	162520.91	104.86	2	37.14	2	4.15	83.47
Rennes	157381.83	103.14	6	20.00	0	10.08	93.71
Saint-Etienne	146181.40	97.29	2	30.57	1	6.86	78.18
Clermont	142212.33	74.43	1	30.14	0	2.77	85.00
Le Havre	141094.70	95.86	19	28.43	6	45.56	90.46
Montpellier	140584.80	78.14	26	27.43	3	14.05	73.81
Dijon	133689.73	87.86	4	26.29	0	0.26	89.01
Orleans	127882.80	77.86	7	33.00	3	14.12	74.51
Mulhouse	118953.31	74.57	3	24.00	0	1.93	77.91
Reims	116454.09	75.57	3	42.86	1	-0.92	89.19
Brest	114776.97	72.71	18	6.00	5	16.17	88.30
Angers	106392.99	68.71	2	15.00	0	6.94	76.39
Limoges	102001.40	65.14	133	24.86	41	10.89	89.55
Bayonne	100702.46	58.43	11	16.14	4	11.83	78.51
Le Mans	100576.13	67.57	1	26.00	0	5.83	79.93
Caen	98755.80	69.43	15	30.00	5	26.13	82.18
Dunkerque	97341.70	61.14	4	28.71	0	7.79	82.57
Amiens	79778.21	53.86	4	21.00	2	6.80	81.95

TABLE 1.1: The data

*A slightly decrease in segregation.* Table 1.2 presents the evolution of the mean values of both segregation indices and some inequality index. Each city is weighted by the size of its population in 2006 in order to estimate the average segregation *experienced* by people. Whatever the geographical scale, income-based segregation measured through the NSI increases between 2001 and 2004 and then decreases. In 2001, 32% of income inequality (as measured by coefficient of variation) can be explained by the inequality *between IRIS*, while this rate is 30% in 2008. Segregation measured by the GSI is rather stable over the period : indeed, in 2001 and in 2008, 49% of income inequality (as measured by the Gini Inequality index) can be explained by the inequality *between* neighborhoods.

At the same time, inequality as measured by Gini Inequality index and the coefficient of variation increases over the considered period.

	NSI based on :		GSI based on :		Income inequality	
	<i>IRIS</i>	<i>Grand Quartier</i>	<i>IRIS</i>	<i>Grand Quartier</i>	Gini Inequality coefficient	Coeff. of variation
2001	0.32	0.28	0.49	0.42	0.38	1.08
2002	0.33	0.28	0.48	0.42	0.38	1.06
2004	0.33	0.28	0.49	0.42	0.39	1.09
2005	0.32	0.28	0.49	0.42	0.39	1.13
2006	0.31	0.27	0.50	0.43	0.39	1.21
2007	0.31	0.27	0.50	0.43	0.40	1.24
2008	0.30	0.26	0.49	0.43	0.40	1.26

TABLE 1.2: Income segregation and income inequality (weighted mean)



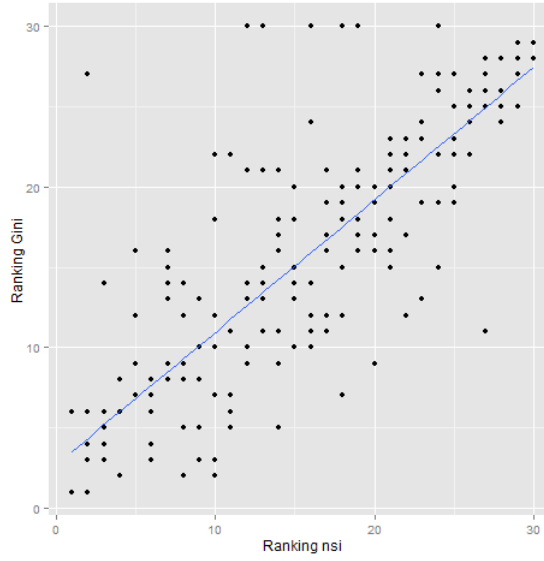


FIGURE 1.1: Ranking correlation between GSI and NSI (*IRIS*scale)

*Comparison between GSI and NSI.* For the 30 cities, table 1.3 shows the individual and neighborhood income inequality measured by the Gini Inequality coefficient as well as the coefficient of variation, and the corresponding segregation indices. Indeed, both indices are built as the ratio of between-neighborhood income inequality over individual income inequality, where inequality is measured through Gini coefficients in the case of GSI and through coefficient of variation in the case of the NSI.

Figures 1.1 and 1.2 illustrate the correlation between the ranks of the two indices for both level (*IRIS* and *Grand Quartier*). Both indices are strongly correlated, indeed, the Pearson correlation coefficient at *IRIS* level is high (0.78) and highly significant (see table 1.4). This correlation is even stronger at the *Grand Quartier* scale, where the Pearson coefficient correlation is equal to 0.8962 (see table 1.5). However, these indices draw different pictures when we make an inter-city comparison over time.

*Inter-city comparisons.* Here we compare the cities throughout the period according to both indices. Figures 1.4 and 1.3 give respectively the values taken by both indices and the ranking of the 30 cities over the period.<sup>10</sup> These figures highlight some differences in the segregation pattern of the cities drawn by both indices. For all the cities, the GSI appears to be higher and more stable over time than the NSI. As me mentioned before, both indices are highly correlated but we observe quite important differences in

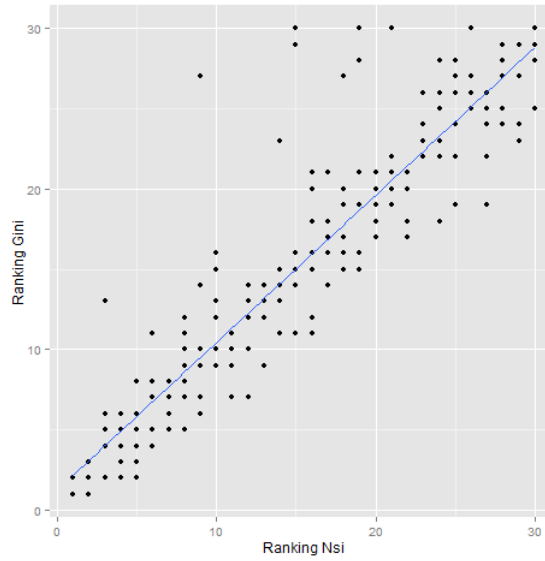
10. As an ordinal measure, the NSI and GSI are particularly useful to classify and rank urban areas. But it cannot quantify the segregation and indicate whether segregation represents an amount  $x$  or  $y$ .

Urban Area	Gini coefficients			Coefficient of variation		
	Income Inequality	Neighborhood Inequality	GSI	Individual Household Income	Neighborhood Mean Income	NSI
Paris	0.42	0.23	0.55	0.0057	0.0018	0.32
Marseille	0.40	0.20	0.50	0.0234	0.0085	0.37
Lyon	0.37	0.17	0.47	0.0285	0.0090	0.32
Nice	0.39	0.15	0.38	0.0400	0.0097	0.24
Lille	0.39	0.21	0.54	0.0326	0.0128	0.39
Bordeaux	0.35	0.14	0.40	0.0441	0.0116	0.26
Toulouse	0.36	0.14	0.40	0.0528	0.0145	0.28
Toulon	0.36	0.14	0.40	0.0513	0.0153	0.30
Nantes	0.34	0.15	0.43	0.0663	0.0182	0.28
Strasbourg	0.39	0.19	0.49	0.0917	0.0305	0.34
Rouen	0.36	0.18	0.50	0.0778	0.0288	0.37
Grenoble	0.35	0.15	0.43	0.0844	0.0273	0.33
Tours	0.34	0.15	0.45	0.0998	0.0325	0.33
Rennes	0.36	0.15	0.41	0.1272	0.0361	0.29
Saint-Etienne	0.35	0.13	0.36	0.0981	0.0250	0.26
Clermont	0.35	0.14	0.39	0.1426	0.0343	0.25
Le Havre	0.37	0.20	0.53	0.1049	0.0423	0.41
Montpellier	0.41	0.17	0.42	0.1308	0.0395	0.31
Dijon	0.34	0.14	0.43	0.1276	0.0391	0.31
Orleans	0.34	0.14	0.42	0.1257	0.0387	0.31
Mulhouse	0.36	0.18	0.50	0.1363	0.0505	0.37
Reims	0.37	0.18	0.48	0.1567	0.0512	0.33
Brest	0.33	0.12	0.38	0.1239	0.0360	0.29
Angers	0.35	0.15	0.43	0.1504	0.0470	0.31
Limoges	0.35	0.13	0.36	0.1655	0.0444	0.27
Bayonne	0.35	0.11	0.31	0.2043	0.0396	0.20
Le Mans	0.34	0.15	0.43	0.1487	0.0484	0.33
Caen	0.36	0.16	0.45	0.1584	0.0524	0.33
Dunkerque	0.36	0.15	0.42	0.1298	0.0443	0.34
Amiens	0.40	0.20	0.50	0.2460	0.0798	0.34

TABLE 1.3: Mean Gini Inequality coefficients and coefficient of variation of individual and neighborhood income distributions, and economic segregation measures

	Gini Segregation	NSI	Gini Inequality
GSI	1.0000		
NSI	0.7780	1.0000	
Gini Inequality	0.5401	0.2294	1.0000

TABLE 1.4: Pearson's correlation coefficients (*IRIS* scale)

FIGURE 1.2: Ranking correlation between GSI and NSI (*Grand quartier scale*)

	Gini Segregation	NSI	Gini Inequality
GSI	1.0000		
NSI	0.8962	1.0000	
Gini Inequality	0.5241	0.3439	1.0000

TABLE 1.5: Pearson's correlation coefficients (*Grand Quartier scale*)

the rankings for some cities. For instance, Paris is the most segregated city according to the Gini Segregation, while regarding the NSI Paris is in the middle of the ranking. On the contrary, Dunkerque appears less segregated regarding the GSI than measured with the NSI.

According to *the GSI*, Bayonne, Limoges and Saint Etienne are the less segregated cities. In 2008, 31% of the overall income inequality in the city, as measured by the Gini Inequality coefficient, is explained by inequality between neighborhoods. While in Paris, the most segregated city, we can explain 56% of the overall inequality by inequality between areas.

According to the NSI, the cities of Bayonne, Nice and Clermont-Ferrand are the less segregated cities throughout the considered period, while the most segregated urban areas are Le Havre and Lille. In 2008, while 16% of the overall inequality as measured by the standard-deviation is explained by inequality between areas in Bayonne, the less segregated city of our sample, we can explain 36% of overall inequality observed in Le Havre by inequality between areas. Thus, the highest NSI is 2.25 times higher than the lowest NSI in 2008. As we see in figure 1.4 it is not surprising to observe that NSI values are not volatile. Few exceptions concern Nantes, Amiens, Montpellier or Strasbourg, for which NSI values follow a more unstable trend. Figure 1.3 gives also

interesting information on the change in ranking of cities. We notice that the ranking according to the NSI is not stable across the period. Bayonne, Nice or Lille are few exceptions of segregation stability. Montpellier, Nantes or Strasbourg observe a strong degree of segregation volatility. Some cities are less and less segregated compared to other cities : such is the case of Lyon and Grenoble. On the contrary, cities like Paris see their rank increasing. However, as we can deduce from the figure 1.4, these changes might be explained rather by close NSI's values, for which a *small* change in the NSI's value can cause important changes in the ranking.

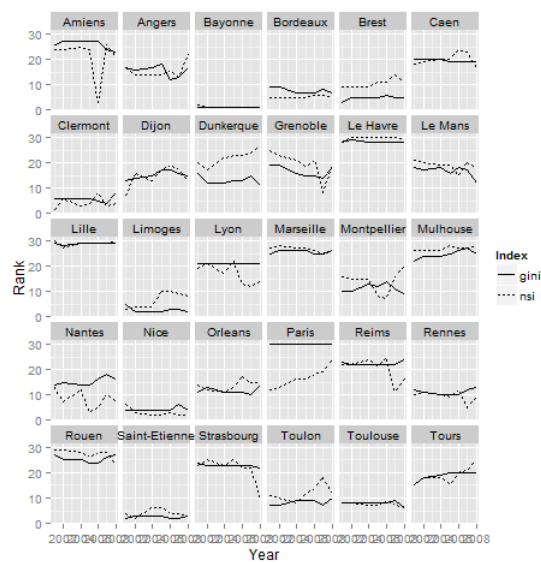
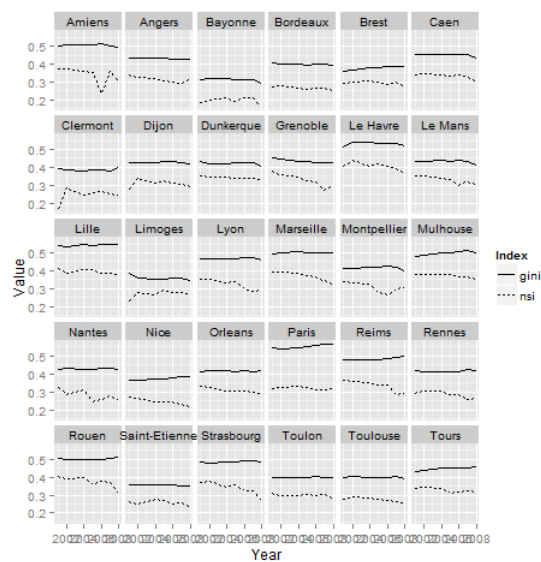
Table 1.6 presents mean value of NSI and GSI for each *large region* in 2008. We divide France into 4 large regions :<sup>11</sup> North and *Bassin Parisien* (9 cities), West (9 cities), East (7 cities) and Mediterranean region (5 cities).<sup>12</sup> Northern cities seem to be more segregated than other cities and western cities seem to be the most mixed of French cities. However we can note that standard-deviation of NSI values is larger within western cities than within other regions.<sup>13</sup> Table 1.6 shows also the decomposition of the variance of the NSI values into within and between regions. The variability of NSI values among French cities at the *IRIS* level is mainly explained by variability within regions than between regions : around 30% of the variance of the *IRIS*-based NSI values and around 42% of the variance of the *IRIS*-based GSI values can be explained by variance between regions. Hence, even if the number of observation per group is small, we are not able to identify some clusters of segregated or mixed cities. We can illustrate this point by the following examples. In the Mediterranean region, Marseille is highly segregated (*IRIS*-based NSI value is equal to 0.37) whereas Nice is one of the most mixed cities (*IRIS*-based NSI value of 0.24).

---

11. Of course such a cutting-up is arbitrary. To check robustness of our results, alternatives were implemented but the qualitative result was not affected.

12. North and *Bassin Parisien* : Paris, Le Havre, Orléans, Reims, Rouen, Tours, Amiens, Lille and Dunkerque. West : Angers, Bayonne, Bordeaux, Limoges, Caen, Le Mans, Brest, Nantes and Rennes. East : Dijon, Grenoble, Lyon, Mulhouse, Saint-Etienne, Clermont-Ferrand and Strasbourg. Mediterranean region : Marseille, Montpellier, Nice, Toulon and Toulouse.

13. However, the lower mean NSI value and the higher standard-deviation can be explained by the fact that Bayonne belongs to this region. If we exclude Bayonne from this region, we would find a mean *IRIS*-based NSI value of 0.28 and a standard-deviation of 0.023 ; a mean *Grand Quartier*-based NSI value of 0.21 and a standard-deviation of 0.047.


FIGURE 1.3: Ranking over time (*IRIS* scale)

FIGURE 1.4: Value over time (*IRIS* scale)

			NSI				Gini Segregation			
			<i>IRIS</i>		<i>Grand Quartier</i>		<i>IRIS</i>		<i>Grand Quartier</i>	
			Mean value	Std- deviation	Mean value	Std- deviation	Mean value	Std- deviation	Mean value	Std- deviation
Bassin	Parisien	and	0.3239	0.0298	0.2954	0.0241	0.492	0.0533	0.4481	0.0523
North										
East			0.2828	0.0408	0.2463	0.0395	0.435	0.0524	0.3781	0.0605
Mediterranean			0.2739	0.0403	0.2148	0.0519	0.4147	0.0486	0.3187	0.0793
West			0.2639	0.0455	0.197	0.0527	0.3916	0.045	0.2872	0.0715
Analysis of variance										
Between-region			0.0180		0.0477		0.0482		0.1283	
Within-region			0.0401		0.0470		0.0648		0.1099	
Total variance			0.0582		0.0947		0.1130		0.2382	

TABLE 1.6: NSI and GSI values per region, 2008

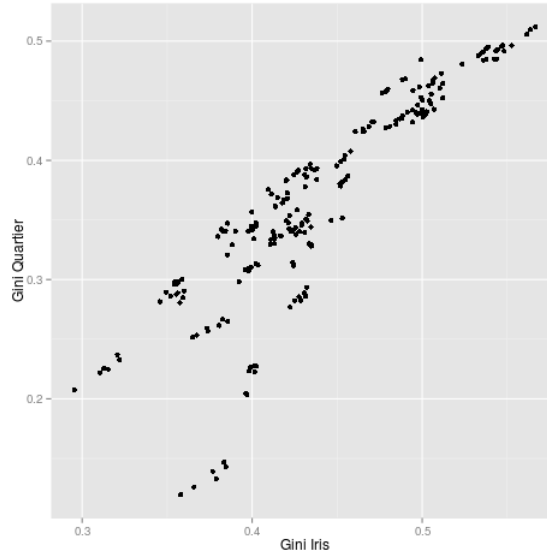


FIGURE 1.5: The value of GSI according to the scale area unit

Figures 1.5 and 1.6 depict, for both indices, the relation between both scales in terms of values which illustrates the proposition 1.9. With respect to this proposition, migrating from *IRIS* scale to *Grand Quartier* scale should not increase the segregation. Indeed, all the points are located under the first bisector. This figure highlights also the strong correlation between segregation index calculated on the basis of *IRIS* and on the basis of *Grand Quartier*. The unweighted coefficient is equal to 0.8859 for the NSI and 0.8835 for the GSI (for all cities and for all years) and are significantly different to zero ( $p$ -value  $< 0.001$ ).

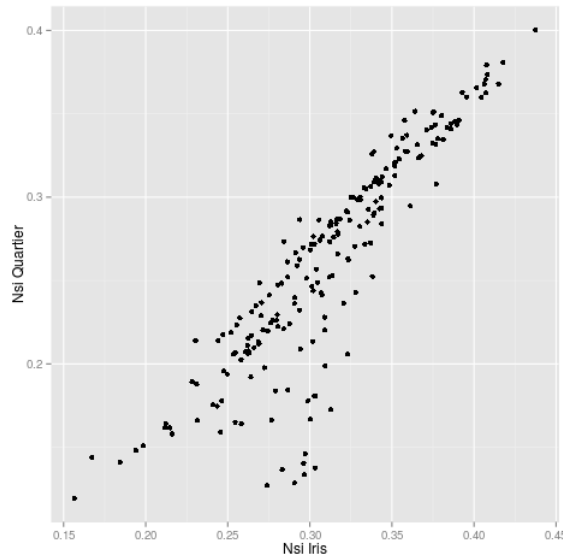


FIGURE 1.6: The value of NSI according to the scale area unit

*Segregation by income and income inequality.* Figures 1.7 and 1.8 represent correlation between inequality (measured by the Coefficient of Variation and the Gini Inequality coefficient<sup>14</sup>) and the segregation indices at the *IRIS* level. Whatever the index, there is no clear relation between inequality (measured with Gini Inequality coefficient<sup>15</sup> and segregation. The coefficient of correlation between the Gini Inequality coefficient and the segregation measured by the NSI is low (0.2294) but statistically different to zero ( $p$ -value  $< 0.001$ ). Similarly, the coefficient of correlation between the Coefficient of Variation and the GSI is low (0.2951) and significant ( $p$ -value  $p$ -value  $< 0.001$ ). The same conclusion holds when we study these correlations based on the *Grand Quartier*. However, given that there are several observations per each city, it is likely that this estimate of the correlation does not make any sense and does not assess properly the link between inequality and segregation. Consequently, we use a simple tobit model with random effects in order to estimate the link between the NSI's rank and the (inequality) Gini's rank and the link between the Gini Segregation's rank and the Coefficient of Variation's rank.<sup>16</sup> Table 1.7 summarizes our results<sup>17</sup>. It appears that the rank with respect to the Inequality Index does not influence the rank with respect to the segregation Index. Interestingly, thanks to a similar tobit model with random effects, we find that the value of NSI is negatively correlated with the value of Gini Inequality coefficient. That is to say, the more a city is equal, the more segregated the city is. However this correlation is significantly positive between the GSI and the Coefficient of Variation (though low, +0.02) The econometric results are summarized in table 1.7.

---

14. As GSI and the Gini Inequality coefficient are intrinsically correlated we focus on the correlation between GSI and the Coefficient of Variation, similarly, we study the correlation between the NSI and the Gini Inequality coefficient

15. Readers interested in properties verified by inequality measures could refer to Chakravarty (1999) or Cowell (2000)

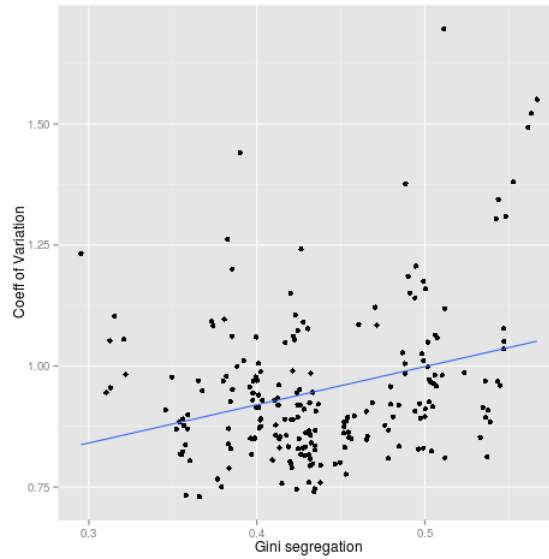
16. The purpose of these simple econometric estimates is not to give a complete overall explanation of the segregation rankings and measures but to give an idea of the correlation between inequality and segregation.

17. Three-stars means that the coefficient is significantly different to zero at 99%; two-stars means that the coefficient is significantly different to zero at 95%;  $W$  is the Chi-2 Wald statistic of test;  $\rho$  is the percent contribution to the total variance of the panel-level variance component.



Dependent variable	Independent variable	Constant	Coefficient	L	W	rho
NSI rank ( <i>IRIS</i> )	Gini Inequality rank	13.16***	+0.15	-604.69	1.70 (p=0.1924)	0.84
NSI rank (Grand Quartier)	Gini Inequality rank	13.10***	+0.15	-548.47	2.59 (p=0.1073)	0.91
GSI rank ( <i>IRIS</i> )	Coeff. of Variation rank	14.89***	+0.04	-433	3.31 (p=0.0689)	0.98
GSI rank (Grand Quartier)	Coeff. of Variation rank	15.03***	+0.03	-379.58	3.50 (p=0.0615)	0.99
NSI value ( <i>IRIS</i> )	Gini Inequality value	0.85***	-1.46***	452.09	29.34 (p=0.0000)	0.91
NSI value (Grand Quartier)	Gini Inequality value	0.76***	-1.38***	469.62	35.52 (p=0.0000)	0.95
GSI value ( <i>IRIS</i> )	Coeff. of Variation value	0.42***	+0.02***	652.28	16.12 (p=0.0001)	0.98
GSI value (Grand Quartier)	Coeff. of Variation value	0.34***	+0.02***	628.48	16.72 (p=0.0000)	0.99

TABLE 1.7: Tobit regressions

FIGURE 1.7: Correlation between GSI and Coefficient of Variation (*IRIS* scale)

To sum up, the results of this simple econometric analysis are : (i) the ranking of cities according to Gini Inequality coefficient is not correlated with the one according to NSI but (ii) The value of GSI is weakly positively correlated with the value of Coefficient of Variation while the value of NSI is negatively correlated with the value of Gini Inequality coefficient. The former result allows us to justify the use of a segregation index like NSI. Indeed, since NSI and Gini Inequality coefficients do not draw the same picture, using NSI provides an interesting and useful information : the more unequal cities are not the more segregated cities. According to the latter result, segregation tends to decrease with an increase in inequality. How can we explain this result ? Even if Gini Inequality coefficient and NSI are not mechanically linked, segregation as measured by NSI is not independent of inequality. Indeed, NSI could be interpreted as the ratio between the Coefficient of Variation between local areas and the Coefficient of Variation for the overall city. Considering that Gini inequality coefficient and coefficient of variation are positively correlated,<sup>18</sup> an increase in income inequality as observed during 2000s (see table 1.2) does logically lead to a decrease in NSI, *ceteris paribus*.

## 4 Conclusion

The goal of this chapter is to lay out a set of properties satisfied by a family of measures of residential segregation by income. The approach studied in this chapter is as follows : once selecting a measure of income inequality which respects some basic properties (such as the Pigou-Dalton principle), a segregation measure is the ratio between

<sup>18</sup>. Note the coefficient of correlation between Gini Inequality coefficient values and Coefficient of Variation values are in average equal to 0.68

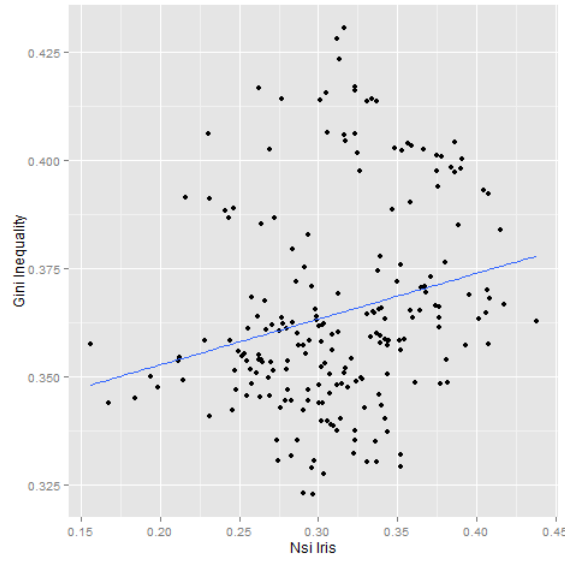


FIGURE 1.8: Correlation between NSI and Coefficient of Variation (*IRIS* scale)

inequality between areas and overall inequality. We relate this approach with elementary transformations capturing that changes in segregation may be the result of changes in allocation of individuals among areas or income transfer between individuals. The basic intuitions underlying these transformations is that segregation increases if income heterogeneity within local areas increases and/or gap between local areas' mean incomes reduces. We show also that such indices can compare cities with different number of individuals and/or different number of areas and/or different overall mean income, and does not depend on whether areas are labelled or whether individuals within areas are named. The most serious drawback of the NSI, as pointed out by Jargowsky (1996), is that it is sensitive to the modifiable areal unit problem and to the physical location of are unit vis-vis one another.

This chapter concludes with an illustration as part of the measurement of the residential segregation by income in the 30 largest French cities. As pointed out in the previous sections, the NSI and the GSI are very appealing for applied analysis. On the first hand, the necessary dataset for their implementation is not prohibitive. On the other hand, these indices can provide some interesting information and usefully complement inequality measures. In the case of the 30 of the largest French cities, we show that, though correlated, the NSI and the GSI draw different rankings of the cities over time. Furthermore, they draw a ranking of cities that is different (and not correlated) from the one according to Gini Inequality index.

## 5 Appendix

## 5.1 Basic properties

**Proposition 1.10** (Symmetry). :

*Symmetry within area.* Let  $(P, Y)$  and  $(P, \tilde{Y})$  two cities such that  $\tilde{Y}^j = D^j Y^j$  ( $j \in M$ ) with  $D^j$  an individual permutation matrix and  $\tilde{Y}^q = Y^q$  for all  $q \neq j$ . Then,  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(P, \tilde{Y})$ .

*Proof.* By definition  $Y_\mu = \tilde{Y}_\mu$ . If  $\mathcal{I}$  respects symmetry,  $\mathcal{I}(Y) = \mathcal{I}(\tilde{Y})$  and then  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(P, \tilde{Y})$ . QED.

*Symmetry between areas.* Let  $(P, Y)$  and  $(P, \tilde{Y})$  two cities such that  $\tilde{Y} = DY$  with  $D$  an areas' permutation matrix. Then,  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(P, \tilde{Y})$ .

*Proof.* If  $\mathcal{I}$  respects symmetry,  $\mathcal{I}(Y_\mu) = \mathcal{I}(\tilde{Y}_\mu)$  and then  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(P, \tilde{Y})$ . QED.

**Proposition 1.11** (Principles of population). :

*Within-area replication invariance.* Let  $(P, Y)$  and  $(\tilde{P}, \tilde{Y})$  two cities such that, for all  $j \in M$  and for any  $\alpha \in \mathbb{N}_+$ ,

$$\tilde{Y}^j = (\underbrace{Y^j, Y^j, \dots, Y^j}_{\alpha+1 \text{ times}}) = (\underbrace{y_1^j, \dots, y_1^j}_{\alpha+1 \text{ times}}, \dots, \underbrace{y_{n_j}^j, \dots, y_{n_j}^j}_{\alpha+1 \text{ times}})$$

Then,  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(\tilde{P}, \tilde{Y})$ .

*Area replication invariance.* Let  $(P, Y)$  and  $(\tilde{P}, \tilde{Y})$  two cities such that, each neighbourhood is replicated  $\ell$  times as

$$\tilde{Y} = (\underbrace{Y^1; \dots; Y^1}_{\ell+1 \text{ times}}; \dots; \underbrace{Y^m; \dots; Y^m}_{\ell+1 \text{ times}})$$

Then,  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(\tilde{P}, \tilde{Y})$ .

*Proof.* If  $\mathcal{I}$  respects Dalton principle of population,  $\mathcal{I}(Y) = \mathcal{I}(\tilde{Y})$  and  $\mathcal{I}(Y_\mu) = \mathcal{I}(\tilde{Y}_\mu)$ . Hence  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(\tilde{P}, \tilde{Y})$ . QED.

## 5.2 Proofs of propositions

### 5.2.1 Proof of proposition 1.1 (Normalisation)

The proof of this proposition is straightforward.

If  $y_i^j = \mu^j$  for all  $j \in M$ , then  $V(Y^j) = 0$  for all  $j \in M$  and  $V_w = \sum_{j \in M} \frac{n_j}{n} V(Y^j) = 0$ . Obviously,  $V(Y) = V_b(Y)$  and  $NSI = 1$  if  $\mu^j \neq \mu^{j'}$  for two at least areas  $j$  and  $j'$  such

that  $j \neq j' \in M$ .

If  $\mu^j = \mu^{j'}$  for all  $j$  and  $j'$  such that  $j \neq j' \in M$ ,  $V_b(Y) = 0$ . Then  $NSI = 0$ .

### 5.2.2 Proof of propositions 1.3 and 1.5 (Segregation under movements)

#### 5.2.3 Unilateral movements

Proof of part i. of Proposition 1.3.

By simplicity, we consider a city with two local areas such that  $\mu^1 < \mu^2$ . Assume a movement of a rich individual from area 2 to area 1. Denote his income by  $\gamma$  with  $\gamma > \mu^2$ . As, by definition, inequality of overall income distribution remains unchanged,  $S_{\mathcal{I}}(P, Y) > S_{\mathcal{I}}(\tilde{P}, Y)$  if and only if  $\mathcal{I}(Y_{\mu}) > \mathcal{I}(\tilde{Y}_{\mu})$ .

We can then write the distribution of mean incomes :

$$\begin{aligned} Y_{\mu} &= (\underbrace{\mu^1, \dots, \mu^1}_{n^1 \text{ times}}, \underbrace{\mu^2, \dots, \mu^2}_{n^2 \text{ times}}) \\ \tilde{Y}_{\mu} &= (\underbrace{\tilde{\mu}^1, \dots, \tilde{\mu}^1}_{(n^1+1) \text{ times}}, \underbrace{\tilde{\mu}^2, \dots, \tilde{\mu}^2}_{(n^2-1) \text{ times}}) \end{aligned}$$

where

$$\tilde{\mu}^1 = \frac{n^1}{n^1+1}\mu^1 + \frac{\gamma}{n^1+1} \text{ and } \tilde{\mu}^2 = \frac{n^2}{n^2-1}\mu^2 - \frac{\gamma}{n^2-1}$$

This unilateral movement can be interpreted as a set of Pigou-Dalton rank-preserving progressive transfers of income between an individual originally located in area 2 and another one originally located in area 1 given that each individual is supposed to have the mean income of her area. That is,  $\tilde{Y}_{\mu}$  can be obtained from  $Y_{\mu}$  by a finite set of rank-preserving progressive transfers :

1. Each individual  $i \in N^1 \cap \tilde{N}^1$  receives from each individual  $i' \in N^2 \cap \tilde{N}^2$  a transfer of income for an amount of :

$$\delta_a = \frac{\frac{1}{n^2-1}(\gamma - \mu^2)}{n^1} > 0$$

Such a transfer of income is progressive. There are  $n^1(n^2 - 1)$  transfers of this type.

2. Each individual  $i \in N^1 \cap \tilde{N}^1$  receives from individual  $i' \in N^2 \cap \tilde{N}^1$  a transfer of income for an amount of :

$$\delta_b = \frac{\mu^2 - \frac{n^1}{n^1+1}\mu^1 - \frac{\gamma}{n^1+1}}{n^1} > 0$$

Such a transfer of income is progressive. There are  $n^1$  transfers of this type.

Hence the new distribution of mean incomes is :

$$\tilde{Y}_\mu = (\underbrace{\mu^1 + \delta_a + \delta_b, \dots, \mu^1 + \delta_a + \delta_b}_{n^1 \text{ times}}, \underbrace{\mu^2 - n^1 \delta_b, \mu^2 - \frac{n^1 \delta_a}{n^2 - 1}, \dots, \mu^2 - \frac{n^1 \delta_a}{n^2 - 1}}_{(n^2 - 1) \text{ times}})$$

As  $\mathcal{I}$  is supposed to respect the Pigou-Dalton principle of progressive transfer,  $\mathcal{I}(Y_\mu) > \mathcal{I}(\tilde{Y}_\mu)$ .

The proof is similar for the other parts of the proposition.

### Switch

Proof of part i. of Proposition 1.5.

By simplicity, we consider a city with two local areas such that  $\mu^1 < \mu^2$ . Assume a switch of two individuals  $i$  (originally in area 1) and  $j$  (originally in area 2) such that  $y_j^2 > y_i^1$ . Once again as overall income distribution remains unchanged,  $S_{\mathcal{I}}(P, Y) > S_{\mathcal{I}}(\tilde{P}, Y)$  if and only if  $\mathcal{I}(Y_\mu) > \mathcal{I}(\tilde{Y}_\mu)$ .

Then, new mean incomes are :

$$\tilde{\mu}^1 = \mu^1 + \frac{\gamma}{n^1} \text{ and } \tilde{\mu}^2 = \mu^2 - \frac{\gamma}{n^2} \text{ with } \gamma = y_j^2 - y_i^1 > 0.$$

Hence the distributions of mean incomes are :

$$\begin{aligned} Y_\mu &= (\underbrace{\mu^1, \dots, \mu^1}_{n^1 \text{ times}}, \underbrace{\mu^2, \dots, \mu^2}_{n^2 \text{ times}}) \\ \tilde{Y}_\mu &= (\underbrace{\tilde{\mu}^1, \dots, \tilde{\mu}^1}_{n^1 - 1 \text{ times}}, \underbrace{\tilde{\mu}^2, \tilde{\mu}^2 - \frac{\gamma}{n^2}, \dots, \tilde{\mu}^2 - \frac{\gamma}{n^2}}_{n^2 - 1 \text{ times}}, \tilde{\mu}^1) \\ \tilde{\tilde{Y}}_\mu &= (\underbrace{\mu^1 + \frac{\gamma}{n^1}, \dots, \mu^1 + \frac{\gamma}{n^1}}_{n^1 \text{ times}}, \underbrace{\mu^2 - \frac{\gamma}{n^2}, \dots, \mu^2 - \frac{\gamma}{n^2}}_{n^2 \text{ times}}) \end{aligned}$$

Assuming that  $\mathcal{I}$  respects symmetry, we can obviously see that a switch implies a series of mean income transfers between individuals located in different areas. Each individual transfer has an amount of  $\frac{\gamma}{n^1 n^2}$  and such a transfer is progressive. As  $\mathcal{I}$  is supposed to respect the Pigou-Dalton principle of progressive transfer,  $\mathcal{I}(Y_\mu) > \mathcal{I}(\tilde{Y}_\mu)$ .

The proof is similar for the other parts of the proposition.

#### 5.2.4 Proof of proposition 1.7 (Internal transfer)

Proof of part i. of Proposition 1.7.

Consider a city with  $m$  area and a progressive transfer between two individuals in area  $p \in \{1, \dots, m\}$ . By definition the distribution of mean incomes is not changed :

$Y_\mu = \tilde{Y}_\mu$  while the overall inequality reduces according to  $\mathcal{I}$  that respects the Pigou-Dalton principle :  $\mathcal{I}(Y) > \mathcal{I}(\tilde{Y})$ . If  $Y_\mu$  is not of the form  $Y_\mu = c\mathbf{1}_m$  (where  $m$  is the total number of areas),  $S_{\mathcal{I}}(P, Y) < S_{\mathcal{I}}(P, \tilde{Y})$ . If  $Y_\mu = c\mathbf{1}_m$ , then  $S_{\mathcal{I}}(P, Y) = S_{\mathcal{I}}(P, \tilde{Y}) = 0$ .

The proof is similar for the other parts of the proposition.

### 5.2.5 Proof of proposition 1.9 (Sensitivity to areas aggregation)

Let  $(P, Y)$  and  $(\tilde{P}, Y)$  two cities. By definition,  $(\tilde{P}, Y)$  is obtained from  $(P, Y)$  by means of an aggregation of areas  $p$  and  $q$ , i.e.,  $P = \{N^1, \dots, N^p, \dots, N^q, \dots, N^m\}$  and  $\tilde{P} = \{\tilde{N}^1, \dots, \tilde{N}^{m-1}\}$ . Then, we can write :

- $N^j = \tilde{N}^j$ ,  $\forall j < q, j \neq p$  and  $N^j = \tilde{N}^{j-1}$ ,  $\forall j > q$
- $N^p \cup N^q = \tilde{N}^p$
- $n^p + n^q = \tilde{n}^p$
- $\mu = \tilde{\mu}$

Then,

$$n^p \mu^p + n^q \mu^q = \tilde{n}^p \tilde{\mu}^p \Leftrightarrow (\tilde{\mu}^p)^2 = [\delta \mu^p + (1 - \delta) \mu^q]^2 \quad (1.2)$$

with  $\delta = \frac{n^p}{\tilde{n}^p}$  and  $(1 - \delta) = \frac{n^q}{\tilde{n}^p}$

Hence,

$$\begin{aligned} S(P, Y) \geq S(\tilde{P}, Y) &\Leftrightarrow \sum_{j=1}^m n^j (\mu_j - \mu)^2 \geq \sum_{j=1}^{m-1} \tilde{n}^j (\tilde{\mu}_j - \tilde{\mu})^2 \\ &\Leftrightarrow n^p (\mu^p)^2 + n^q (\mu^q)^2 = \tilde{n}^p (\tilde{\mu}^p)^2 \end{aligned} \quad (1.3)$$

From (1.2) and (1.3), it follows that

- $S(P, Y) \geq S(\tilde{P}, Y) \Leftrightarrow \delta (\mu^p)^2 + (1 - \delta) (\mu^q)^2 \geq [\delta \mu^p + (1 - \delta) \mu^q]^2$
- $\Leftrightarrow (\mu^q - \mu^p)^2 \geq 0$

## Chapitre 2

# National Identity and Immigrants' Assimilation in France

Determination and changes of immigrants' identity resulting from intercultural contacts impact their socio-economic integration. To precisely assess individuals' identity, we propose a continuous index which aims to overcome interpretation troubles faced by usual measures of ethnic identity. Then, we investigate the determinants of immigrants' ethnic identity in France. We compare our composite and continuous index exhibiting individuals' assimilation with a usual measure of ethnic identity – the national identity ("I feel French" dummy). We underline the importance of some sociodemographic characteristics in ethnic identity formation and detail immigrants' assimilation in France. We are thus able to show that cultural assimilation and national identity do not always coincide. It seems that the further the origin (in cultural terms), the higher the national identity, but the lower the assimilation. We also present evidence of second generations' identity convergence to natives' one, either in terms of national identity (almost total commitment) or assimilation.

### 1 Introduction

Under the impact of immigration, modern societies become *culturally plural* (Berry (1997)). People of many cultural backgrounds have to live together and immigrants' assimilation is thus a crucial issue. Two models of integration can be pointed out. The multiculturalism model adopted by Anglo-Saxon countries is based on the coexistence of various cultures. On the other hand, the assimilation model adopted in France compels



that immigrants perfectly and uniquely integrate the culture of the host country (see Bloemraad (2007)).

Ethnic identity of immigrants is an expression of their integration. This issue has been traditionally addressed by sociologists (Berry (1997)) and political scientists (Abdelal (2009)). Economists have recently paid attention to immigrants' identity since it is likely to impact individuals' behaviors and decisions (Akerlof & Kranton (2000)) and therefore to induce consequences on their socioeconomic integration (see for empirical studies, Pendakur & Pendakur (2005), Constant & Zimmermann (2007), Nekby & Rodin (2007), Manning & Roy (2010), Battu & Zenou (2010), Casey & Dustmann (2010)). In most of these works, the empirical measure of ethnic identity is based on the answer to the question "*Do you think of yourself as* [natives' citizenship] *?*" Moreover, although estimation of the impact of ethnic identity on the socioeconomic integration of immigrants raises econometric issues, the literature stresses a significant and negative influence of ethnic identity (Nekby & Rodin (2007), Constant et al. (2009), Casey & Dustmann (2010), Battu & Zenou (2010), Bisin et al. (2011)).

Nevertheless, this measurement of ethnic identity suffers from a number of drawbacks. The first one is that the surveyed individual has to place herself regarding the native's identity – for instance, through questions like "*do you feel french* (or "*British*", or "*German*", etc.)" That can only be answered by a simple yes or no. This discrete measure of identity loses the fact that the immigrants' identity may be more or less close to the natives' one. For instance, valuable information would similarly be lost by assessing individuals' incomes as "*rich*" or "*poor*" rather than a continuous distribution.

A second drawback raised by Lee (2009) underlines the subjective interpretation that each respondent may give to the reference category ("*as French*", "*as british*", "*as German*", etc.). Indeed, this latter may not be equally salient and valid across individuals and contexts. Similarly, Citrin & Sears (2009) point out that while answering to a "*Do you feel French ?*" question like, it is not clear whether the respondent considers the "*civic*" or "*ethnic*" (say, "*cultural*") conception of a nation? In other words, some may consider themselves as civically French (i.e. in a legal point of view) with, however, multiple ethnic identities (in a cultural point of view). National identity questions cannot report these.

A third one is given by Abdelal (2009). It emphasizes the fact that it may be stigmatizing to answer that you do not feel like the natives, particularly when the pollster is a native.

This chapter makes four contributions.

First, we provide a new measure of ethnic identity that circumvents drawbacks of existing measures. Its key feature is that it is based on characteristics that individuals choose

to define themselves. Precisely, we use the *TeO* survey conducted by INED and INSEE in 2008 which is aimed to inform about the life conditions of immigrants and their descendants in France. To assess ethnic identity, we use the following question “*Which of the following features define yourself the most?*” Individuals can choose at most 4 features among the following 14 items : “*my age/generation*”, “*my gender*”, “*my job occupation*”, “*my educational attainment*”, “*my neighborhood or my town*”, “*my health*”, “*my citizenship*”, “*my origins*”, “*my skin color*”, “*my religion*”, “*my hobbies*”, “*my political opinion*”, “*my region of origin*” and “*my family situation*”. We conduct a Multiple Correspondence Analysis (MCA) on these 14 items and we interestingly observe that the first dimension discriminates characteristics chosen by immigrants from those chosen by native respondents. This crucial dimension represents the ethnic or cultural part of respondents' identity. Note that we do not ask for this immigrants-versus-natives opposition. No information regarding immigration position is included in the MCA. We ex-post observe such an opposition and conclude that the ethnic dimension of identity is the most important.

Following this result, we are thus able to build a weighted composite index of these 14 items where the weights are the correlation coefficients between each item and the first dimension. By construction, this index has to capture respondents' ethnic identity. The bigger the index, the closer the self-definition from the “*typical*” native one. The construction of the composite index follows the methodology provided by [Filmer & Pritchett \(2001\)](#) in a rather different context.

Our index therefore circumvents the difficulties faced by the standard measure of ethnic identity. It is continuous and allows obtaining a measure of the cultural distance between immigrants and natives. Furthermore, it overcomes the salience and validity troubles since respondents do not situate themselves regarding a given category but choose characteristics that ex-post locate them in an ethnic identity axis, thanks to the MCA procedure. By doing so, it also follows [Sen \(2007\)](#)'s call for multiple identities consideration, rather than a one-dimensional position. Finally, in a face-to-face interview, respondents are likely to feel less under pressure by defining themselves through multiple characteristics than regarding the dominant group belonging (natives).

Second, considering that the presence and importance of identity questions largely differ among surveys and countries, the assimilation index is not a ready-made tool for further studies. However, its methodology is exportable and interesting results are extracted from MCA. In particular, we find on the one hand that immigrants and, to a lesser extent, descendants of immigrants are likely to define themselves with attributes related to their ethnic group (origins, nationality, religion, skin color). On the other hand, natives choose more individualistic items (job, level of education, hobbies). At

the risk of paraphrasing [Durkheim \(2013\)](#), the ethnic part of immigrants' identity could be seen as a mechanical identity, because self-definition is made by homogeneity with other ethnic group members. On the contrary, natives' one would be as organic identity, since it highlights individuals' socio-economic interdependence with others members (for more details about mechanical versus organic integration in France see in [SCHNAPPER \(2007\)](#)).

Third, we try to clarify the interpretations that can be done thanks to ethnic identity in terms of socioeconomic integration. Among the existing measures in the empirical literature, national identity is by far the most used. We discuss and qualify conclusions implied by this use by running regressions on national identity and on the assimilation index in parallel. The comparison of the results for both measures enables us to deepen the understanding of national identity and, in particular, to explore its meaning and performance regarding assimilation. This leads us to challenge some conclusions of the economic literature. In particular, [Manning & Roy \(2010\)](#) investigate national identity determinants in Great-Britain and find that *"the process of assimilation is faster for those from poorer and less democratic countries"*. With our data we have a similar conclusion with national identity. However, this conclusion is completely reversed when we use our assimilation index. Then, our contribution is simple, national identity quite imperfectly assesses the convergence of immigrants' ethnic identity toward natives' one. Furthermore, it seems that when assimilation of an immigrant is high, the claim for national identity remains less necessary.

Finally, we give a detailed description of the determinants of ethnic identity. Notably, as found in the literature ([Manning & Roy \(2010\)](#), [Battu & Zenou \(2010\)](#), [Casey & Dustmann \(2010\)](#)), time since arrival is crucial. Thus we show that, compared with first generations, second generations claim more often national identity (93.12%, against 63.33%) and assimilate better. Nevertheless, the assimilation index of second generations is much lower than natives' and some determinants as discrimination or religiosity that do no impact first generations are activated for second's.

The following section presents the literature. Section 3 describes more precisely the identity features and assimilation index achievement after a brief presentation of the survey. We also present the potential measures of identity. In section 4, we introduce the exogenous variables and corresponding descriptive statistics. Ethnic identity's determinants for the whole French population (thanks to representative weights) are presented in Section 5. Section 6 focuses on a comparison of first and second generations of migrants and presents related determinants. Section 7 concludes.

## 2 Literature review

This chapter focuses on the measurement of ethnic identity with the aim of improving the understanding of immigrants' integration. It thus belongs to two strands of literature, economics of immigration and economics of identity.

The literature in economics about identity is recent. [Akerlof & Kranton \(2000\)](#), [Austen-Smith & Fryer Jr \(2005\)](#) propose theoretical frameworks – not specifically in immigration contexts – to explain individuals' socio-economic outcomes through identity. Identity refers to the sense of belonging to a given group. More particularly, ethnic identity is activated when ethnic groups have to cohabit together which is an obvious consequence of immigration. While migrating, an individual might face a trade-off between conforming to the host culture or preserve her own cultural heritage such that it could even lead to adopt oppositional identities ([Battu & Zenou \(2010\)](#)). Consequently, ethnic identity may have direct implications on economic outcomes ([Pendakur & Pendakur \(2005\)](#), [Constant & Zimmermann \(2007\)](#), [Nekby & Rodin \(2007\)](#), [Battu & Zenou \(2010\)](#), [Casey & Dustmann \(2010\)](#), [Bisin et al. \(2011\)](#)). Empirically, immigrant's assimilation is often approached by the national identity ([Manning & Roy \(2010\)](#), [Casey & Dustmann \(2010\)](#), [Bisin et al. \(2010\)](#)). None of these studies investigate the case of France which is interesting in two ways. First, France has a specific model of integration which promotes cultural conformity of people (natives and immigrants) toward the “French Republican ideal”. Second, for two centuries, France has experienced a relatively high rate of immigration<sup>1</sup>. Furthermore, the issue of identity measurement is crucial but poorly documented by the literature. This concern is at the core of this chapter.

Besides national identity, a few other empirical works have focused on the measurement of ethnic identity ([Constant & Zimmermann \(2007\)](#), [Nekby & Rodin \(2007\)](#)) using [Berry \(1997\)](#)'s typology. [Constant & Zimmermann \(2007\)](#) have introduced the ethnosizer for commitment with home or host identity measurement. It allows a ranking of individuals regarding this commitment among four states of acculturation : integration (commitment with both host and home culture), assimilation (commitment with host culture only), separation (commitment with host culture), or marginalization (no commitment with either host or home culture). To compute the index they defined *a priori* identity features such as spoken language, religion, societal interaction... They are then able to emphasize different pre-migration – religious cults, origins – and post-migration determinants – age, education, and hours worked. Our methodology is different and less subjective since we do not have postulate which characteristic should be regarded as typically native or typically immigrant.

---

1. See [Noiriel \(2002\)](#).

In 2011, [Bisin et al. \(2011\)](#) focus on ethnic identities of immigrants in Europe. Thanks to factor analysis, they construct a composite index with three variables that are attachment to religion, attachment to traditions and customs, language spoken at home. They present determinants of identity as origins, age, educational level, gender and years since arrival. Then, they analyze the influence of ethnic identity on educational and economic outcomes. However, their index is based on only three variables. The quality and amount of information given by the *TeO* survey as well as the use of Multiple Correspondence Analysis are assets of our index.

Our work provides new information about immigrant's identity in France. Furthermore, we confront usual measures to a new one which exhibits immigrants' assimilation based on individuals' self image. Thus, compared with previous studies, we can show that national identity and assimilation of immigrants, although tied, are not equivalent.

### 3 Data, sample and measures of ethnic identity

#### 3.1 Data and sample

Our work is based on the French *Trajectoires et Origines* (*TeO*, thereafter) cross-sectional survey developed in 2008 which focuses on integration and discrimination issues. It offers a very large set of variables about today's first and second generation of immigrants in France. It has been conducted by two main French statistical institutes, INSEE and INED.

The scope of the survey is the set of people from 18 to 60 years old residing in metropolitan France according to the 2007 French national census. Survey areas have been chosen – countryside and cities – to be representative of the structure of the whole country. It voluntarily over-samples immigrants and descendants of immigrants but representative weights have been calculated to improve analysis. The major difficulty was the construction of the second generation sample due to their statistical invisibility in French censuses. This has been fixed by assembling information about parents' birth country, nationality, and information about respondents' current situation. Only one person in the household has been interviewed but information about other members in the housing is available. Finally, 21761 respondents have been kept in the overall sample, including 8456 immigrants, 8161 descendants, 712 DOM<sup>2</sup> natives, 651 descendants of DOM natives and 3781 respondents whom both parents are born French and called French natives. After our own selection (some necessary answers about respondents' self image were missing), our final sample is composed of 6964 descendants of immigrants, 6868 immigrants and 3545 natives.

---

2. "Départements d'Outre-Mer", that is French Overseas Departments.

The survey can be divided into 3 main themes : domestic and social environment (classical but numerous socio-demographics), access to social life resources (employment, educational, housing outcomes), different dimensions of origins, and cultural belongings (parents' origin, senses of belonging, discriminations, etc.). One of its main appeals is that it includes an entire part about *“person's self image”*.

### 3.2 The assimilation index

Each one of the existing measures has its limits : interpretation dilemmas with national identity, endogeneity issues in identity formation explanation with the ethnosizer, lack of information and improper statistical method with the composite index of Bisin et al. (2011). Here, we propose a methodology aiming to circumvent drawbacks of these measures.

#### *Construction*

Tajfel (1974) explains that individuals identify themselves by features that both translate their belonging to a specific group and distinguish them from other groups. The *TeO* survey offers a large set of variables regarding respondents' identity. One question is particularly interesting in the questionnaire, *“Which of the following features would you choose to define yourself? Please choose at most 4 of them”*. Then, 14 features are proposed : *“my age/generation”*, *“my gender”*, *“my job occupation”*, *“my education”*, *“my neighborhood or town”*, *“my health”*, *“my nationality”*, *“my origins”*, *“my skin color”*<sup>3</sup>, *“my religion”*, *“my hobbies”*, *“my political opinion”*, *“my region of origin”* and *“my family situation”*. We call these previous features identity attributes in the sense that they are chosen by respondents to define themselves. In other words, these attributes are 14 endogenous variables corresponding to 14 specific dimensions of a person's identity. The challenge is to find a proper way to summarize this heterogeneous information about individuals' identity in a unique measure.

In a totally different context, Filmer & Pritchett (2001) create a weighted index of Indian individuals' material possessions (washing machine, clock, etc.) in order to approach long-run wealth. They run a Principal Component Analysis with all individuals possessions as exogenous and assume that the first dimension exhibit respondents' wealth, i.e. their long-run wealth index. Therefore, weights equal correlation coefficients between each possession variable and the first dimension. Bisin et al. (2011) follow this approach with a limited amount of information.

3. One might think that skin-color is particularly discriminating since it would target particular population. Around 36% of sub-Saharan migrants choose this feature against nearly 10% of French. As regard to the relatively limited gap between the previous proportions we do not think that this feature is biased by restrictiveness.

However, PCA in [Filmer & Pritchett \(2001\)](#) and factor analysis in [Bisin et al. \(2011\)](#) are not suitable procedures when exogenous attributes are discrete. That is why we prefer another multivariate procedure, the Multiple Correspondence Analyses (MCA).

MCA helps us to uncover and summarize the ethnic part of individuals' identity that is common to the 14 identity attributes. Indeed, this statistical procedure aims to classify different variables (here the identity attributes) among several axes, the dimensions, such that the first dimension accounts for as much of the variability in the data as possible and, in turn, each succeeding dimension has to get the highest variance as possible under the constraint that it is uncorrelated with the preceding one. Therefore, each dimension is a weighted linear combination of all the variables entering in the statistical procedure where weights are the corresponding MCA's correlation coefficients.

Obviously, MCA is a blind procedure and interpretation of each dimension that it produces is left to researchers. Here is an important assumption stipulating that the first dimension (i.e. the one that captures the highest possible part of the common information) exhibits ethnic identity. Running the MCA – without postulating on what should be immigrants' or natives' characteristics (these information about individuals' origins are not used neither needed to run the MCA) – we obtain the satisfaction of this assumption since immigrants are retrospectively located on negative and natives on positive values of the first dimension. This dimension can be extracted as an indicator thanks to correlation coefficients, this indicator being henceforth our assimilation index.

The index will then have the following form for respondent  $j$  :

$$I_{1j} = w_{11}.X_{1j} + w_{12}.X_{2j} + \dots + w_{1Z}.X_{Zj}$$

with  $I_{1j}$  being the index (equal to the first dimension) for individual  $j$ ,  $w_{1Z}$  the correlation coefficient between the  $z$ th identity attribute and the first component (that is weights) and the value for the  $z$ th attribute. Its minimum is -12.25 and its maximum 7.89 with a mode around 0.41 and a standard deviation equal to 3.47. The higher it is, the closer (resp. the further) the respondent self-image or identity from the natives' typical one (resp. immigrant's). We interpret moves of the index as improvement or worsening of individual's assimilation since it corresponds to Berry's typology and is positively correlated with the assimilation state of the ethnosizer<sup>4</sup>.

---

4. Thanks to our data, we were able to build the ethnosizer developed by Constant et al. We then observed that assimilation was the state of acculturation which was the most correlated with our index. In the interest of clarity of the chapter, we do not develop the construction of our ethnosizer hereafter. However, further details and results about our “*TeO-based*” ethnosizer and its relations with our index are available upon request.



<i>Weights for assimilation index</i>	
<b>Identity</b>	<b>Weights based on</b>
<b>attributes</b>	<b>MCA</b>
Religion	-4.019
Skin color	-3.852
Nationality	-2.547
Origins	-1.831
Region	-0.394
Gender	0.039
Family	0.498
Health state	0.648
Education	0.831
Town/Neighborhood	0.843
Generation	1.580
Political view	1.725
Job occupation	2.207
Hobbies	2.378

TABLE 2.1: Weights for assimilation index

Results for MCA-based weighting procedure are presented in table 2.1. There are 9 identity attributes which increase the assimilation index when chosen : hobbies, job occupation, political view and generation being the four most representative. Conversely, there are 5 attributes that decrease the index with particularly high coefficients for religion, skin color, nationality and origins. By construction, the assimilation index does not represent the entire individual identity but only the ethnic dimension of it.

We insist on the fact that these attributes are not dummies which would indicate whether the respondent is a male or a female, is graduated or not, has a religion or not and so forth, but dummies which take 1 if she believes that these attributes actually define herself, 0 otherwise. To say it even simpler, every respondent has a family position but everyone does not select it to define herself.

### *Index performances*

Reader used to multivariate analysis knows that the percentage of inertia is crucial to determine the importance of each dimension. Here, the percentage of total inertia explained by the first dimension may seem too low (10.53%). However, M. Greenacre (2005) points out that, with MCA, it would be *"futile to expect a good approximation of a matrix of zero and ones in a two dimensional map of points"*. He advises to find other ways to assess the quality of the dimensions and notably to look at the correctness of predictions. To do so, we run 14 probit regressions with each of the fourteen identity items as endogenous and the assimilation index as the unique regressor. We then



<i>Predictions of each attribute by the index only</i>			
<b>Good predictions</b>	<b>Not selected</b>	<b>Selected</b>	<b>Total</b>
Religion	77.18	91.14	84.16
Skin color	73.16	89.90	81.53
Hobby	71.54	80.51	76.03
Job occupation	78.00	69.85	73.93
Nationality	75.44	68.48	71.96
Generation	70.25	66.14	68.20
Origins	49.38	80.19	64.79
Political view	85.98	42.72	64.35
Quartier	58.27	63.27	60.77
Education	58.20	62.45	60.32
Health	82.66	28.38	55.52
Family	56.50	52.32	54.41
Region	68.93	36.54	52.73
Gender	66.14	35.37	50.75
Total	69.40	61.95	65.67

TABLE 2.2: Predictions of each attribute by the index only

check whether or not predicted values of these simple models match actual answers of respondents. 2.2 gives the percentages of correct predictions for each identity attributes. The assimilation index alone predicts almost 70% of no (zero) and 62% of yes (ones). These relatively high percentages suggest that the first dimension (the index) succeed in summarizing the information that is common to all identity attributes.

As main works in this literature, we use national identity (*"I feel French"* dummy). Nevertheless, most papers directly regard national identity as an expression of ethnic identity and even more abusively of immigrants' assimilation.

Our assimilation index, with its combination of multiple attributes, has the advantage to not directly ask people about which group they belong (which is more or less implicitly what national identity and the ethnosizer<sup>5</sup> do) but about which personal characteristics they identify and then to locate them on an endogenous ethnic identity dimension. By doing so, it deepens individuals' identity investigation and enlightens national identity.

## 4 Treatment and descriptive evidence

5. The appendix provides a comparison of the assimilation index and our TeO-based calculation of the ethnosizer.

## 4.1 Treatment

In this section and the following, we only focus on national identity and the assimilation index. We look at their determinants thanks to probit – for national identity – and OLS regressions – for the ipdex. Endogeneity is obviously a main concern, insofar as we do not clearly distinguish how far identity explains individual behaviors from how far individual positions influence identity. That is the reason why we focus, as far as possible, on impact of exogenous demographic characteristics as :

For each group,

- Age ;
- Gender ;
- Education, 1 if higher than certificate of general education ("brevet des colleges"), 0 otherwise ;
- Language spoken by parents during childhood ;
- Parents' origins ;
- Discrimination feeling, 2 types (due to skin color and due to origin) ;
- Proportion of immigrants in living area, 1 when respondent says that at least half of her neighborhood inhabitants are immigrants ;
- Mother education, 1 if higher than certificate of general education, 0 otherwise ;
- Parents' religion ;
- Home country sense of belonging ;
- Framing effect control ;

For first generations only

- Years since arrival in France ;
- French citizenship ;

For second generations only,

- Mixed origin, 1 if respondent has exactly one of her parents who is a French native and 0 otherwise ;

We are aware of the risk of endogeneity of some of the previous variables (education, discrimination, for instance) and we try to minimize it as far as possible. However, we cannot completely avoid it and interpretations must be cautious.

The survey provides representative weights that will be used in every treatment of our study. It also proposes a huge amount of other possible variables and we chose the most relevant ones after trying a lot of them. Context variables (as unemployment rate in the neighborhood, average level of education, etc.) could be used but, surprisingly, none

<i>Means of ethnic identity measures</i>				
	Overall	Natives	1 <sup>st</sup> Generations	2 <sup>nd</sup> Generations
National Identity (“ <i>I feel French</i> ”)	93.75	97.87	63.33	93.12
Assimilation Index	1.596	2.004	-0.671	0.810
N	17377	3545	6868	6964
Deeper descriptive analysis of national identity thanks to French <i>TeO</i> survey can be found in Simon (2012) and Simon and Tiberj (2012).				

TABLE 2.3: Means of ethnic identity measures

of them has a significant impact on either national identity or assimilation index and, more importantly, they would increase endogeneity issues.

#### *Framing effect issue*

A well-known issue raised by this type of question is the framing effect. Indeed, while answering a multiple answers question, respondents are likely to choose the first ones. Anticipating this, the fourteen items were written in two different orders (A or B). They were then randomly proposed to respondents. 50.08% of respondents answered a A-order questionnaire, 49.98% answered a B-order questionnaire. Not surprisingly, the comparison of answers among the two series shows that a framing effect occurs. However, in this work, our first interest is not to find the characteristics that people choose to define themselves but if there are differences in these choices between groups. The two series being randomly determined, the difference between each group choice (native, immigrants, and descendants of immigrants) is not impacted. However, the framing effect also influences MCA coefficients. In order, to address this issue, a dummy corresponding to the items' order proposed to each respondent is integrated in the following regressions.

## 4.2 Descriptive statistics

*TeO* survey allows differentiation of respondents as regard to their personal migratory history (that is their parents' or their own origin). In particular, it is possible to know whether an individual belongs to the French natives group (that we define as respondents with no migratory history for less than two generations), the first generation group (immigrants) or the second generation group (descendants of immigrants). Here we present some descriptive statistics for each of these groups and for the whole sample.

Table 2.3 gives weighted means of our both endogenous measures of ethnic identity. Obviously, weights have important impacts on statistics for the whole sample because

of over-representation of immigrants in the survey. The proportion of respondents which agrees with the sentence "*I feel French*" largely corresponds to majority (93.75%, 97.87% for natives and 93.12% for descendants). One could consider the proportion of first generations which agrees as low (63.33%) but it is quite similar with the value found by Manning & Roy (2010) or Battu & Zenou (2010) for British national identity and much higher than what Casey & Dustmann (2010) obtain for German national identity among immigrants with very similar questions in both cases.

Interpretations for the assimilation index are more difficult and regressions results will be much more interesting. Concerning relations between national identity and the assimilation index, we observe a positive correlation between both measures<sup>6</sup>. By construction of the sample, second generations are much younger than natives and even more than first generations (see the appendix for detailed figures). Such a gap in age is likely to induce differences of identity. This will be controlled as far as possible and notably by using representative weights and performing distinct regressions for first and second generations.

Table 2.4 allows for analyzing the changes of the structure of the French immigration. Indeed, here are presented origins of parents. If we compare second and first generations origins, we then roughly consider two different, although wide, waves of immigration (in average, around the late 60's for second generations' parents, and in the late 80's for first generations). Maghreb immigration has slightly decreased in proportion of overall immigration while the proportion of western European immigration is almost divided by two. Eastern Europe immigration stays at a relative low level when Sub-Saharan's, Asian's and Turkish's have significantly increased. Finally, 41% of second generations' mothers or fathers are French.

Complementary tables of descriptive statistics are available in the Appendix. We notably observe that average time since arrival of first generations is quite high (almost 21 years). The highest differences can be observed in religions. While almost none of the natives' parents are Muslims (less than 1%), this religion was followed by parents of 43% of immigrants and more than 25% of descendants. On the contrary, Christian and atheistic represents almost the totality of natives' parents, against less than a half of immigrants' parents and around two thirds of descendants' parents.

European and Arabic languages are the most spoken foreign languages in France (see appendix ). We distinguish Sahelian languages from other African languages following

---

6. In order to observe the relation between both measures, we run a simple OLS regression with the assimilation index as endogenous and national identity as regressor, estimate = 1.744\*\*\*.

<b>Origin</b>	<i>Weighted shares of origins</i>		
	<b>Whole Sample</b>	<b>1st Generation</b>	<b>2nd Generation</b>
Other Africa	6.66	10.10	3.30
Northern America	0.70	0.81	0.58
South America	1.98	3.11	0.70
Asia	4.98	7.81	2.26
Eastern Europe	7.11	6.27	7.95
Western Europe	40.14	27.72	52.26
North Africa	29.57	31.36	27.91
Middle East	1.82	2.52	1.14
Sahel	3.04	4.00	2.10
Turkey	4.01	6.30	1.80
	100.00	100.00	100.00

TABLE 2.4: Weighted shares of origins

the work of some French researchers (notably [Lagrange \(2013\)](#)) who pointed out a more difficult integration of Sahelian natives due to cultural gaps. 41.28% of first generations immigrants already acquired French citizenship at the time of the survey. The literature regularly points out the importance of the citizenship's acquisition on national identity of immigrants. Furthermore, we observed that 83.4% of immigrants with French citizenship claim national identity against 49.2% for those without. This is going to be verified by further econometric analysis.

## 5 Determinants of ethnic identity : whole sample

Our aim is to investigate the importance of each determinant and wonder to what extent national identity is an informative measure of ethnic identity thanks to its comparison with the assimilation index. Table 2.5 presents the results of a weighted probit on national identity and a weighted OLS regressions on the assimilation index on socioeconomic regressors for the French representative population.

The first remark is that, for this sample, determinants' impacts are quite similar (in terms of signs and significance) for both measures. Though, the assimilation index seems to depend on more numerous determinants than national identity. This is particularly visible for languages<sup>7</sup>. Generally speaking, French taken as control, the use of a foreign language by parents leads to increase respondents' distance from native identity (either assessed by national identity or by the assimilation index). However, one can remark

7. We consider the first language used by both parents to speak with the respondent during her childhood. When the language spoken by mother and father was different (and different from French) we picked the language spoken by the father. Though subjective, this choice does not affect the analyses since this situation represents only 1.6% of our immigrants' sample (first and second generations).

some specificity. European languages only affect national identity. Arab only affects assimilation. We are going to focus deeper on languages in the next section since they can be seen as an assimilation effort of parents for second generations and an asset in terms of human capital for both second and first generations.

Not surprisingly, French origin taken as the reference, respondents' origins have globally negative effects on commitment toward natives' identity. However, situations differ among ethnic groups regarding the selected measure of ethnic identity. While some origins only affect national identity (Northern America, Northern Europe), others exclusively impact the assimilation index (South America, Asia, Sahelian and other Africa). This is an important result of our study, immigrants from culturally close origins (roughly speaking, developed countries) less commit toward national identity but have higher score in the assimilation index. The reverse conclusion holds for culturally far origins.

As expected mixed origin (exactly one parent who is born French) increases both national identity claim and assimilation. Furthermore, it leads to the highest variation among all regressors for national identity and the second highest one for the assimilation index.

Age has significant effects on both measures of ethnic identity but these are surprisingly negative. This can be explained by the fact that first generations, who are the farthest from natives' identity, are older than other groups (natives and second generations). There is no impact of gender.

<i>National identity and assimilation index for the whole population</i>		
	<b>National Identity</b>	<b>Index</b>
<b>Origin</b> (ref. French)		
Northern America	-0.105* (-2.55)	-0.487 (-1.28)
South America	-0.0201 (-1.26)	- 1.458*** (-4.44)
Asia	-0.0109 (-0.86)	- 1.522*** (-5.03)
Africa	-0.00761 (-0.83)	- 1.909*** (-5.44)
Eastern Europe	-0.0285* (-2.12)	-0.658* (-2.21)
Northern Europe	-0.0772*** (-3.98)	-0.0345 (-0.14)
Southern Europe	-0.0250* (-1.99)	-0.511* (-2.41)
North Africa	-0.0195* (-2.23)	- 0.587*** (-3.42)
Middle-East	-0.0150 (-1.09)	0.00901 (0.03)
Sahel	-0.0197 (-1.56)	- 1.814*** (-5.52)
Turkey	-0.0168 (-1.01)	-0.418 (-1.31)
Mixed	0.0193*** (5.15)	0.423* (2.40)
Age	-0.000609** (-3.11)	- 0.0166* (-2.31)
Gender	0.00223 (0.47)	-0.0314 (-0.35)
Diploma	0.0125 (1.74)	0.799*** (7.35)
Discrimination origin	-0.0172** (-3.04)	-0.467* (-2.31)
Discrimination skin color	-0.00724 (-1.32)	- 0.637** (-2.62)
Immigrants share in neighborhood	-0.00390* (-2.49)	-0.102* (-2.52)
Mother diploma	0.000508 (0.09)	0.636*** (5.70)
<b>Citizenship</b> (ref. foreigner)		
French by reintegartion	0.0252*** (7.37)	0.0225 (0.06)
French by acquisition	0.0272*** (9.38)	0.132 (1.12)
French	0.0735*** (4.70)	0.0675 (0.38)
Home country national identity	-0.0255*** (-5.62)	- 0.975*** (-9.74)

*National identity and assimilation index for the whole population*

	<b>National Identity</b>	<b>Index</b>
<b>Mother religion</b> (ref. atheist)		
Christian	-0.00961 (-1.06)	-0.00770 (-0.05)
Muslim	0.00312 (0.32)	-0.597* (-2.11)
Buddhist	-0.0145 (-0.70)	0.640 (1.47)
Other	0.0135 (1.29)	0.289 (0.87)
<b>Father religion</b> (ref. atheist)		
Christian	0.000844 (0.08)	0.222 (1.46)
Muslim	-0.0183 (-1.27)	-0.126 (-0.50)
Buddhist	-0.0192 (-0.88)	-0.291 (-0.73)
Other	-0.0422 (-1.45)	-0.320 (-1.06)
Importance of religion in received education	0.00228 (0.79)	- 0.261*** (-5.02)
<b>Parents' language</b> (ref. French)		
European	-0.0393** (-2.92)	-0.223 (-1.28)
Arab	-0.0157 (-1.85)	- 0.658*** (-3.96)
Berber	-0.0228 (-1.81)	-0.451 (-1.87)
Sahel	-0.0337* (-2.02)	-0.906* (-2.55)
African	-0.0279* (-2.25)	-0.988** (-3.00)
Asian	-0.0309* (-2.23)	-0.704** (-2.69)
Turkish	-0.0807* (-2.50)	-0.649* (-2.03)
Other	-0.0343 (-1.12)	-1.213 (-1.89)
N	17377	17377
adj. R-sq	0.316	0.167



Education may be a source of endogeneity since the causality way with ethnic identity is unclear. Many previous researches use the time of education which does not solve this issue. Thus, we created a dummy with the certificate of general education (CGE) – *Brevet des colleges* – which is the first national diploma that all French pupils have to pass (around 14 or 15 years old, compulsory). Respondents who have a higher diploma are assigned 1 and 0 otherwise. we thus minimize time before diploma and then the potential reverse impact of identity on educational choice. Interestingly, national identity does not depend on diploma whereas the impact of this latter on assimilation index is the highest.

Discrimination feeling is obviously subjective and may imply endogeneity troubles. One must thus be cautious. Yet, it seems that skin color<sup>8</sup>based discrimination does not impact national identity, contrary to origin discrimination. Regarding assimilation index, the coefficient associated with origin discrimination is lower (and less significant) than the skin color's one. In any case, the sign is negative and discrimination increases the distance from natives' identity<sup>9</sup>.

During the survey, respondents had to give their own estimation about the proportion of immigrants in their living area. Although there is no way to control for over- or underestimation, this variable is of first interest and could be seen as a perceived segregation indicator. Estimates are significant and suggest the likelier conservation of ethnic identity in segregated areas.

A strong identification with home national identity implies a lower commitment toward identity of natives. This was expectable for the assimilation index since assimilation is a process that does not theoretically accept the conservation of ethnic identities. In a sense, this goes against the hypothesis of multiple identities claiming that ethnic identities are not substitutes but complements.

Pursuant to previous findings, a positive impact of years since arrival can be exhibited (Manning & Roy (2010) ; Casey & Dustmann (2010) ; Battu & Zenou (2010)) on national identity. However, it does not impact the assimilation index (it does in the next section when we focus on first generations in particular).

Finally, we are able to look at parents' religions and the importance of religion during respondent's childhood. National identity is not significantly driven by parents' religion and the only significant impact on the assimilation index comes from the practice of Islam by the mother. However, robustness of this result is weak since it does not hold in

---

8. Surprisingly we did not find significant crossed effect of African origin and skin color discrimination.

9. Endogeneity could appear if people with strong ethnic identity react stronger to potential discrimination than those close to natives' identity and thus declare it more often to the pollster.

the next section while considering only immigrants. What is actually primordial is not the religion but the importance of it in respondent childhood education.

## **6 National identity and assimilation : first and second generations**

This section investigates the determinants of ethnic identity for first and second generations of immigrants table (2.6) and allows innovative information about their situation in France. We are also able to challenge previous conclusions of the literature. Unfortunately, we do not have panel data and immigration structure has changed between first and second generations of the sample, comparisons in terms of intergenerational integration between these two groups are thus fragile.

### **6.1 Differences between second and first generations**

The determinants of ethnic identity do not have the same impact on first and second generations. Origins well explain identity of first generations (with positive signs for rational identity and negative ones for the index, this result is developed in the next subsection) but do not influence second generations' one. This is an evidence for the process of assimilation since the inertia of home culture decreases among time and generations.

*Comparison of Identities between First and Second Generations*

	<b>National Identity</b>		<b>Assimilation index</b>	
	1st Generations	2nd Generations	1st Generations	2nd Generations
	Marginal effect	Marginal effect	Estimate	Estimate
<b>Origin</b> (ref. other Africa <sup>a</sup> )				
Northern America	-0.313*** (-3.87)	-0.0951 (-1.12)	1.959*** (4.28)	0.802 (1.12)
South America	-0.0912 (-1.45)	-0.00861 (-0.20)	0.864* (2.13)	0.724 (1.24)
Asia	-0.0162 (-0.29)	0.00653 (0.25)	0.174 (0.45)	0.732 (1.50)
Eastern Europe	-0.191*** (-3.34)	0.0184 (1.13)	1.656*** (4.88)	0.274 (0.48)
Northern Europe	-0.291*** (-5.84)	-0.0433 (-1.38)	2.691*** (8.71)	0.759 (1.46)
Southern Europe	-0.196*** (-3.75)	-0.0297 (-1.78)	1.830*** (5.74)	0.549 (1.21)
North Africa	-0.0736 (-1.23)	-0.0103 (-0.60)	1.452*** (3.53)	0.699 (1.51)
Middle-East	-0.0466 (-0.69)	-0.0307 (-0.83)	1.801*** (3.93)	1.486* (2.28)
Sahel	-0.0669 (-1.35)	-0.00932 (-0.41)	0.166 (0.46)	-0.671 (-1.28)
Turkey	-0.0897 (-1.14)	-0.0326 (-0.79)	1.431*** (3.39)	1.068 (1.92)
Mixed	- (-)	0.0172 (1.67)	- (-)	0.472* (2.26)
Age	-0.00268** (-3.00)	0.00109* (1.97)	-0.00118 (-0.21)	0.00196 (0.19)
Gender	0.0927*** (6.29)	-0.00467 (-0.60)	0.386*** (3.98)	-0.0850 (-0.56)
Diploma	-0.000720 (-0.04)	0.0116 (1.26)	0.898*** (8.52)	0.335 (1.65)
Discrimination origin	-0.0893*** (-3.83)	-0.0265* (-2.02)	-0.00278 (-0.02)	-0.119 (-0.70)
Discrimination skin color	-0.0345 (-1.03)	-0.0262 (-1.67)	-0.647* (-2.55)	-1.346*** (-5.29)
Immigrants share in neighborhood	-0.0285*** (-4.79)	-0.00895*** (-3.39)	-0.142*** (-3.68)	-0.0809 (-1.43)
Mother diploma	-0.0281 (-1.21)	-0.00749 (-0.75)	0.723*** (4.75)	0.480** (3.29)
<b>Citizenship</b> (ref. foreigner)				
French by reintegration	0.178*** (4.40)	0.0312* (2.14)	0.418 (1.07)	-1.779* (-2.42)
French by acquisition	0.245*** (16.25)	0.0379*** (5.57)	0.0958 (0.84)	-0.0544 (-0.12)
French	-0.0911 (-0.30)	0.0750* (2.27)	1.731* (2.07)	-0.246 (-2.43)
Home country national identity	-0.153*** (-10.38)	-0.00140 (-0.20)	-0.904*** (-9.08)	-0.763*** (-4.70)
Years since arrival	0.00938*** (11.32)	- (-)	0.0123* (2.28)	- (-)
Framing effect control	0.000202 (0.01)	0.00631 (0.93)	0.649*** (6.76)	0.438** (3.06)

Table 6 (continued)  
*Comparison of Identities between First and Second Generations*

	<b>National Identity</b>		<b>Assimilation index</b>	
	1st	2nd	1st	2nd
	Generations Marginal	Generations Marginal	Generations Estimate	Generations Estimate
Importance of religion in received education	effect -0.0215** (-2.70)	effect -0.00819* (-2.53)	-0.541*** (-11.04)	-0.471*** (-5.06)
<b>Mother religion</b> (ref. atheist)				
Christian	-0.0270 (-0.61)	0.00990 (0.70)	0.221 (0.86)	0.160 (0.55)
Muslim	0.00315 (0.05)	-0.0211 (-0.97)	0.661 (1.60)	-0.347 (-1.05)
Buddhist	-0.0681 (-0.66)	-0.0384 (-0.47)	1.011* (1.96)	-0.468 (-0.95)
Other	-0.0509 (-0.76)	-0.0102 (-0.50)	0.605 (1.42)	0.597 (1.10)
<b>Father religion</b> (ref. atheist)				
Christian	0.0722 (1.91)	-0.00933 (-0.64)	0.0467 (0.21)	0.606** (2.63)
Muslim	0.0456 (0.74)	-0.0118 (-0.65)	-0.632 (-1.58)	-0.122 (-0.47)
Buddhist	-0.0718 (-0.70)	-0.00360 (-0.09)	-0.294 (-0.58)	-0.0161 (-0.04)
Other	0.0391 (0.73)	-0.0272 (-1.20)	-0.102 (-0.28)	-0.421 (-1.35)
<b>Parents' language</b> (ref. French)				
European	-0.100* (-2.52)	-0.0746** (-2.79)	0.0530 (0.24)	-0.393 (-1.65)
Arab	-0.0771 (-1.49)	-0.0202 (-1.49)	-0.231 (-0.59)	-0.866*** (-3.99)
Berber	-0.119 (-1.83)	-0.0372 (-1.31)	-0.637 (-1.35)	-0.236 (-0.58)
Sahel	-0.121 (-1.80)	-0.0270 (-0.92)	-0.466 (-0.98)	0.239 (0.44)
African	-0.117* (-2.38)	-0.0547 (-1.13)	0.0554 (0.18)	-1.786** (-3.09)
Asian	-0.143** (-2.71)	-0.0000113 (-0.00)	-0.128 (-0.37)	0.0240 (0.06)
Turkish	-0.261** (-3.25)	-0.0790 (-1.31)	-0.0790 (-0.18)	-1.371** (-2.92)
Other	-0.171 (-1.46)	- (-)	-0.270 (-0.35)	- (-)

If we now focus on languages spoken by parents during childhood, these differently impact first and second generations. Four groups of languages (European, African, Asian, and Turkish) have a negative impact on first generations' national identity but no one influences the assimilation index. On the opposite, a few second generations' childhood languages impact their assimilation scores (Arab, African and Turkish) whereas only European languages lead to a decrease of their national identity.

Age has different roles on national identity. Considering the first generations (resp. second generations), the elder they are, the lower (resp. the higher) their national identity. Though, it does not drive the assimilation index neither for immigrants nor descendants of immigrants.

French citizenship<sup>10</sup> and its acquisition type give expectable and positive results regarding national identity. However, it seems to have no influence on the assimilation index. French citizenship acquisition by reintegration claim<sup>11</sup> for second generations even has a surprising negative estimate on assimilation.

Gender only affects first generations and males are closer from natives' identity.

A surprising result concerns respondents' diploma. This latter only significantly plays a role in first generations assimilation index formation. This result may come from our restrictive choice of considering only the first national grade that French pupils have to pass as a proxy of respondents' education. By doing so, we aim to reduce the endogeneity bias.

Parents' religion does not drive ethnic identities. However, its importance in respondent's childhood is highly significant for each generation and negatively related to each identity measure.

Finally, many determinants have expectable and similar coefficients. A mixed origin and a well-educated mother are assets to commit toward natives' identity, while living in a segregated area (immigrants in neighborhood), a strong identification with home country (I feel [home country]), and a discrimination experience are not.

Years since arrival for immigrants have a quite high impact on national identity since the average migrant that spent 20 years on the French territory would face a 19% increase of her probability to feel French. Not surprisingly, ethnic identities commitment is a time-costly process. Nevertheless, time since arrival - even though significant - plays a much less important role on assimilation index compared with other determinant.

---

10. One could wonder why a second generation respondent can be foreigner as regard to the French "jus solis" for citizenship. The reason is that a second generation individual, when she attains legal majority, is given the opportunity to choose to keep either home and host nationalities (usual choice) or only one (scarcer case).

11. This procedure refers to people who have had their French citizenship lost and want it back.

## 6.2 Key differences between national identity and assimilation

In section 5, signs and significances generally coincide among ethnic identity measures. This is not the case if we focus on first and second generations only. Many determinants impact one but not the other. Some even have opposite signs.

On the one hand, some determinants affect national identity but not the assimilation index. This is the case of age, origin related discrimination, acquired French citizenship and European languages spoken by parents. On the other hand, skin color related discrimination, mother diploma and the control for the framing effect are significant determinants of assimilation but not of national identity.

More strikingly, origins of respondents significantly lead to different signs of coefficients. To consider only one measure in order to give conclusions on ethnic identity formation is thus likely to imply partially wrong understandings. To highlight this risk, let us focus on origins of first generations. The reference origin, other (i.e. than North and Sahel) Africa, is the same for all models. However, while almost all groups of origins are less likely to claim national identity than the control, they assimilate more French typical self-definition (assimilation index). In particular, western European immigrants have a probability to claim national identity almost 30% lower than Sub-Saharan African ones. This was not expected but consistent with Manning and Roy (2010) findings. Using only British national identity (answer to "What do you consider your national identity to be?"), they indeed conclude that immigrants from poorer and less democratic countries assimilate better. Yet, if we consider origins' estimates for the assimilation index, Western Europe coefficient of first generations is the highest positive one, that is to say that western European immigrants assimilate better than Sub-Saharan. Our conclusion is thus opposite to Manning & Roy (2010) since, here, assimilation of poorer and less democratic countries' natives appears to be lower.

We do not see these remarks as evidence for the failure of one measure or the other, but as the justification of their combination in such a study. They do not measure exactly the same process but, taken separately, they are not sufficient to give proper information on identity formation and, further, on the influence of ethnic identity on socioeconomic outcomes. Their combination can even be helpful to understand their working. For instance, we believe that if assimilation of an immigrant is high, the claim for national identity is made less necessary to belong to the society. The national identity claim can be seen less as a proxy for French identity than a way to fulfill an already existing assimilation gap.

## 7 Conclusion

Ethnic identity has become a key element in studies focusing on immigration. Issues and changes resulting from intercultural contacts are indeed a major concern. This chapter aims to investigate the lack of interest given to ethnic identity measurement in previous studies and to use these understandings to tell more about immigrants' assimilation in France.

In order to complement usual measures, we build an assimilation index that takes into account the distance to the host culture while having identified its cultural characteristics through a statistical method, without making prior assumptions on these characteristics. We thus assess and investigate ethnic identity through two different channels : a dummy for national identity and a continuous variable for assimilation. We claim that this measure brings new information about migrants' assimilation in France, and is more suitable for an analysis of the effects of assimilation on migrants' economics outcomes as it captures an actual distance in identity.

The MCA analysis enables to clearly distinguish different sets of variables that distinguish immigrants' from French natives' self-image. According to this analysis, 5 typical immigrant features increase the distance with the French identity when chosen that are religion, origins, skin color and nationality. On the opposite, the three most representative variables for the natives are hobbies, job and generation.

What this chapter shows is that both measures, national identity and assimilation index, bring different and complementary information about migrants' identity. However, national identity cannot be substituted to assimilation measures since it does not sufficiently and finely approach ethnic identity. We know that it can even lead to wrong interpretations. Indeed, individuals possibly identify themselves like French natives do but reject the national identity, whereas others who do not assimilate French stereotypes claim the national identity. This is even more striking when considering origins of immigrants.

National identity captures an assimilation will or wish, the assimilation index captures, as far as possible, realized assimilation.

Our results qualify Manning & Roy (2010)'s conclusion since we find that assimilation of immigrants from poorest countries is lower, even though they are more likely to "feel French". This justifies the plea of political science for combination of measurements (Sylvan & Metskas (2009), Abdelal (2009)) and, notably, for continuous variables (Lee (2009)).

Thus, we claim that finer measures like the assimilation index would be more suitable to analyze the impact of ethnic identity (as a distance to natives' identity) on socioeconomic outcomes.

Identity theory is particularly suitable for the investigation of the French model of assimilation. Indeed, immigrants in France are expected to economically and socially integrate but they also had to assimilate French values and culture which mainly differs from multiculturalism. Our results – notably regarding time since arrival and the comparison of assimilation levels between first and second generations – suggest that assimilation is a long process and that determinants of ethnic identity formation cannot be well understood by using national identity as a unique proxy.

As regard to our results about ethnic identity's determinants, some facts can be exhibited. As said before, immigration structure has changed in six decades, thus today's and yesterday's immigrants should not be directly compared. However, inertia of origins decreases for second generations' assimilation. Descendants' national identity rate is very close from natives and their assimilation index is much higher than first generations' one.

## 8 Appendix

### A.1. Descriptive statistics



Table A.1

*Weighted means and proportions*

	French	First	Second
	population	Generation	Generation
Age	39.21	40.68	35.21
Gender [male]	49.16	47.81	51.03
Diploma : >EGC	75.59	60.04	74.14
Discrimination : origin	4.11	16.06	12.9
Discrimination : skin color	1.94	7.67	5.69
Migrants in neighborhood [>50%]	11.05	27.29	21.96
<b>Mother religion</b>			
None	24.49	6.95	15.53
Christian	72.00	42.16	53.69
Muslim	0.48	43.06	24.71
Buddhist	0.06	3.37	1.08
Other	2.98	4.47	4.99
<b>Father religion</b>			
None	31.03	9.76	21.15
Christian	63.8	39.17	44.26
Muslim	0.55	42.61	27.19
Buddhist	0.00	2.92	1.00
Other	4.61	5.53	6.40
<b>Religion role in received education</b>			
Not important	37.82	14.65	26.80
Slightly important	33.86	21.60	28.35
Rather important	16.21	24.17	22.46
Very Important	12.12	39.57	22.38
Mother's diploma	17.74	16.32	14.46
French citizenship	93.61	41.28	97.79
I feel [home country]	9.25	55.91	26.85
Community size	5.96	5.61	6.30
Years since arrival		20.71	

Table A2

*Languages : weighted proportions*

	French population	First Generation	Second Generation
French	82.04	4.8	48.68
Arabic	5.08	27.32	17.57
Berber	1.01	6.38	2.63
Sahelian	0.4	2.61	1.08
Other	0.89	7.2	0.9
Africa			
European	8.19	34.02	24.92
Asian	2.39	17.66	4.22



## Chapitre 3

# Segregation and Trust : An experiment in French secondary schools

We design an experiment to study the effect of teenagers' social environment on their trust : in particular, our aim is to test the relation between social mixing and trust behavior. To study this question, we used an extra-laboratory experiment based on a trust game. Subjects are teenagers of two secondary schools located in Rennes (France). These secondary schools are both located in different neighborhoods within a city, contrasting in their degree of social mixing. Thanks to a within-subjects design, it is possible to study whether teenagers condition their trust choice on the school belonging of their partner and how school identity affects their trust and trustworthiness.

## 1 Introduction

Without trust, economic and social relations are hindered if not impossible. Indeed, many relations rely on implicit contracts where trust is therefore a key element. In many real world examples, trust appear to be essential to create or maintain a relationship. An often quoted example is for instance when taking a taxi abroad where only a certain degree of trust between the taxi driver and the user enables this transaction. Moreover, trust is part of social capital held by an individual and might have consequences on economic outcomes, indeed, low levels of trust are associated with poorer economic performance ([Knack & Keefer \(1997\)](#) ; [Zak & Knack \(2001\)](#)). A growing litterature has focused on the effects of social capital, this is therefore not surprising that economists have investigated the topic and accumulated a large body of evidence on the economic consequences of trust.

In this chapter we are interested in investigating the potential links existing between trust and segregation. On one hand, as a pillar of economic and social life, trust enables to facilitate and maintain relationships between individuals. On the contrary, the absence of trust and trustworthiness can lead to enforce segregated environment, where individuals from different groups or with different characteristics do not interact. Therefore, segregation could be a consequence of the absence of trust. One cause of segregation has been discussed by [Schelling \(1971\)](#) who showed that even a small preference for one's neighbors to be of the same color could lead to complete segregation.

On the other hand, segregated environment may impact the degree of trust and trustworthiness of individuals. This aspect lead us to focus on segregation and social environment as determinants of trust. Actually, little is known about trust determinants and whether people experience different degree of trust and trustworthiness regarding their socio-economic background. We investigate the potential relations between trust and the social environment of teenagers. We focus on teenagers as they do not choose their residential location which avoids some endogeneity issues where trust could condition the localisation choice of individuals. Indeed, in the case of adults, social preferences, and among them trust, may impact the residential location choice. Furthermore, we are interested in the following questions : can we observe in-group favoritism in trust, and if so, do people from different socioeconomic groups discriminate similarly ?

We address these questions with the help of a field experiment conducted among 100 teenagers in two middle schools in Rennes, France. The experimental game is a variant of the so-called trust or investment game. In this sequential two-player game first movers (players A) send a transfer to second movers (players B), which is tripled by the experimenters. Second movers then decide how much to return. The amounts sent

and returned inform us about the levels of trust and trustworthiness, respectively. Studying trust discrimination requires a social environment that defines distinct groups. In our study these groups are defined by two secondary schools located in two different neighborhoods in Rennes. These neighborhoods are different in terms of socio-economic segregation, indeed, one of the secondary school is located in a highly segregated area while the other one is located in a mixed neighborhood.

Our results can be summarized as follows. First we find that players A shows relatively high levels of trust and transfer around 60% of their initial endowment to players B. We find that players A in the segregated school significantly trust less than participants in the mixed school. We do not find evidence for in-group favoritism for players A however we find that players B in the segregated school tend to be less egalitarian and discriminate in both ways, whether they favor a participant from their own school or from another school.

The remainder of the chapter is organized as follows. The first part provides a literature review on measuring trust and trust determinants. The second part provides the design of the experiment and gives some background information. Section 3 discusses the results and section 4 concludes.

## 2 Literature review

### 2.1 Measuring trust

In this chapter we retain the generally used definition of trust introduced by Coleman & Coleman (1994) according to whom “an individual trusts if he or she voluntarily places resources at the disposal of another party without any legal commitment from the latter, but with the expectation that the act of trust will pay off”. In order to measure trust economists have developed various tools. Traditionally, the measurement of trust relies on attitudinal questions. Previous research are based on a question in the General Social Survey (GSS) which is the following : “*Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people*”. The latest ranking based on the question in the GSS reveals that 74.2% of people in Norway mostly trust others while they are 3.8% in Trinidad and Tobago. In France they are 18.8% while the average is 26.1%. In studies based on this type of question Arrow (1972) and Fukuyama (1995) conclude that the level of trust in a society strongly predicts its economic success. Similarly, Knack & Keefer (1997) show that an increase of one standard deviation in country-level trust predicts an increase in economic growth of more than one-half of a standard deviation. According to Knack & Keefer (1997) “if you take a broad enough definition of trust, then it would explain basically all the difference between the per

capita income of the United States and Somalia”. In their survey [Algan & Cahuc \(2013\)](#) highlight the dynamics of trust and document the role of climate, weight of history and income on degree of trust. The latest OECD report ([OCDE \(2008\)](#)) highlights that richer countries trust more, indeed the share of people expressing high levels of trust is positively correlated to the median equivalised household income. Moreover, trust is higher when income inequality is lower, there exists a negative correlation between levels of trust and Gini coefficient.

However [Glaeser et al. \(2004\)](#) have pointed out the drawbacks of such a method to measure trust. Indeed, responses are difficult to interpret and the variation in the answers may come from various sources as the different understanding of what it means to be trusted or the different interpretation of who comprises “most people”. Moreover variation in the answers can also come from the fact that people are not willing to answer truthfully to such a question or have different ability to elicit trustworthy behavior from other people.

An other approach to measure trust has thus been developed and is based on experimental games. Introduced as a laboratory experiment by [Berg et al. \(1995\)](#) the trust game consists in a game where two players are anonymously paired. The first mover decides how much on his initial endowment of 10 euros he wants to pass to a second player, the receiver. The second player receives this amount tripled by the experimentalist and decides how much of this transfer he wants to send back to the sender. The sender thus receives : his endowment - his transfer + the back transfer sent by the receiver. The receiver earns : the initial endowment + the transfer multiplied by the experimenter - the back transfer. Therefore, decisions from the sender constitute a measure of trust as it measures the amount someone is inclined to send expecting his partner being reciprocal. Decisions from the second mover constitute a measure of trustworthiness as it measures the degree of reciprocity of this player. Many versions of this game exist as it has been replicated frequently and is a popular choice among trust researchers. Several authors have analysed the relationship between survey-answer and decisions in the trust game and found mixed results. In their study, [Glaeser et al. \(2000\)](#) found that questions about trusting attitudes do not predict trusting behavior but appear to predict trustworthiness while [Fehr et al. \(2003\)](#) find the opposite. [Holm & Danielson \(2005\)](#) find a positive correlation between survey-answers and behavior in trust game in Sweden but not in Tanzania. These results are difficult to compare as the experiments are quite different in these studies.

## 2.2 Trust determinants

Among trust researchers some have focused on trust determinants and little is known about what can influence trust. Using data from U.S. localities, [Alesina & Ferrara \(1999\)](#) estimate that the strongest factors that reduce trust are a recent history of traumatic experiences, belonging to a discriminated group (an ethnic minority or to a lesser extent being a woman), being economically unsuccessful in terms of income and education and living in a racially mixed community and/or in one with a high degree of income disparity.

The trust game introduced by [Berg et al. \(1995\)](#) has been conducted many times in many countries to test the determinants of trust. Here we focus on studies that have tested the effect of social distance on trust. Using an inter-country design where students are paired whether with someone from their university or with someone from another university few studies have shown little or no evidence for social distance effect ([Netzer & Sutter \(2009\)](#); [Willinger et al. \(2003\)](#)) in the sense that the country of residence of their partner does not affect trust behavior of the senders and the recipients. Similarly [Fershtman & Gneezy \(2001\)](#) have implemented a trust game with students from two distinct ethnic groups in Israel : Ashkenazic Jews and Eastern Jews. In this version of the game students are paired whether with someone from their own ethnic group or with someone from a distinct ethnic group. They find an ethnicity effect rather than a social distance effect where Ashkenazic players receive on average a higher amount. Few other studies have focused on the effect of belonging to a common group on experimental trust ([Karlan \(2005\)](#); [Etang et al. \(2011\)](#); [Cadsby et al. \(2008\)](#)) where this belonging to a group can be associated with social distance effect. The weakness in these studies is that group membership could be a treatment effect (joining the group leads people to trust each other) or a selection effect (people who trust each other join the group). To overtake this caveat [Buchan et al. \(2006\)](#) artificially create random groups, this method follows the Minimal Group Paradigm (MGP) approach commonly used in psychology ([Tajfel \(1970\)](#)) which suggests that subjects are more likely to trust people from their own group even if the groups are artificially created. In their study [Buchan et al. \(2006\)](#) set up an experiment where students from different countries can be matched only with a player from their own country after having spent few minutes chatting with their potential partner. When playing the trust game, subjects are told if they are playing with someone from their discussion group (in-group matching) or from a different group (out-group). The effect of this artificial social distance varies across countries, Americans send (and return) more money to in-group members while Chinese students send more to out-group members. A handful of studies have proposed to measure trust through a trust game in a field experiment, [Harbaugh et al. \(2002\)](#) conducted a field experiment in classroom in



order to evaluate trust and trustworthiness across age and examine its development over time. They propose a version of the [Berg et al. \(1995\)](#) trust game for children of different age where individuals received an initial endowment of 4 tokens. [Falk & Zehnder \(2007\)](#) proposed a city-wide experiment in Zurich in order to examine in-group favoritism and evaluate the determinants of trust. They allow senders to base their trusting decision on the neighborhood residency of the receiver. Similarly, [Etang et al. \(2011\)](#) realize a field experiment in Cameroon where they allow the sender to base his decision on the village belonging of the receiver. Both studies highlight in-group favoritism and show that the localization of the receiver matters while choosing the amount sent. Comparably to these studies in our experiment the senders can discriminate regarding the school belonging of their partner.

### 3 Design and Background Information

In this section we give detailed information regarding the city and the secondary schools where the experiment was settled. We then describe the experiment.

#### 3.1 Background information

The experiment took place in two secondary schools located in two different areas in Rennes, France. Collège Emile Zola (school **M** thereafter) is located in the city center while the collège des Hautes Ourmes (school **S**) is located in the south of the city, in a poorer area. Both secondary schools are located in neighborhoods contrasting in their degree of social mixing. School **M** is located in neighborhood 1 (see figure [3.1](#)), while school **S** is located in a more segregated area (neighborhood 11 on the map [3.1](#)) **S** is actually located in a urban renewal zone (*Zone Urbaine Sensible*) characterized as a priority zone for urban public policy. Furthermore, school **S** is classified as a priority education zone (*Zone d'Education Prioritaire*) defined as a zone with higher public dotations and higher autonomy in order to face socio-economic difficulties.

We also present statistics about socio-economic context of each neighborhood. Table [3.1](#) provides the population, the mean and median incomes, inequality indices and a segregation index for each neighborhoods. Concerning segregation index, we use the NSI as introduced by [Jargowsky \(1996\)](#). As used in chapter 1, the NSI is defined as the square root of the ratio of the income variance between IRIS composing a neighborhood over the income variance within the neighborhood. The NSI belongs to the interval  $[0,1]$ , where 1 states for a complete degree of segregation. School **S** is located in the more segregated area in the city, in this neighborhood the median income is two times lower to the median income of the neighborhood where school **M** is located. Moreover, the

neighborhood where school **S** is located is characterized by a strong level of inequality, indeed the interdecile ratio is equal to 16.64.



FIGURE 3.1: Map of Rennes

Neighborhood	Population (UC)	Mean income	Median income	Gini	Interdecile	Income segregation
<b>1</b>	<b>12222,8</b>	<b>27542,09</b>	<b>23977,2</b>	<b>0,4</b>	<b>7,78</b>	<b>0,1</b>
2	15080,4	31007,54	25383,11	0,41	6,46	0,2
3	10008,5	23285,61	21447,92	0,35	5,83	0,12
4	3152,1	23113,52	22620	0,35	5,6	0,02
5	12687,2	18803,65	17656,33	0,38	8,33	0,4
6	7428,7	23511,07	19985,8	0,39	9,34	0,17
7	12732,5	21858,76	20990,71	0,32	5,64	0,15
8	11687,6	23615,64	22322,75	0,35	5,91	0,09
9	8983,8	22548,3	20224,5	0,38	7,33	0,16
10	9663,9	15292,37	14949,06	0,39	17,84	0,27
<b>11</b>	<b>11562</b>	<b>14492,87</b>	<b>13371,56</b>	<b>0,37</b>	<b>16,64</b>	<b>0,43</b>
12	9847,5	17624,45	16530,5	0,33	7,71	0,33

TABLE 3.1: Neighborhoods

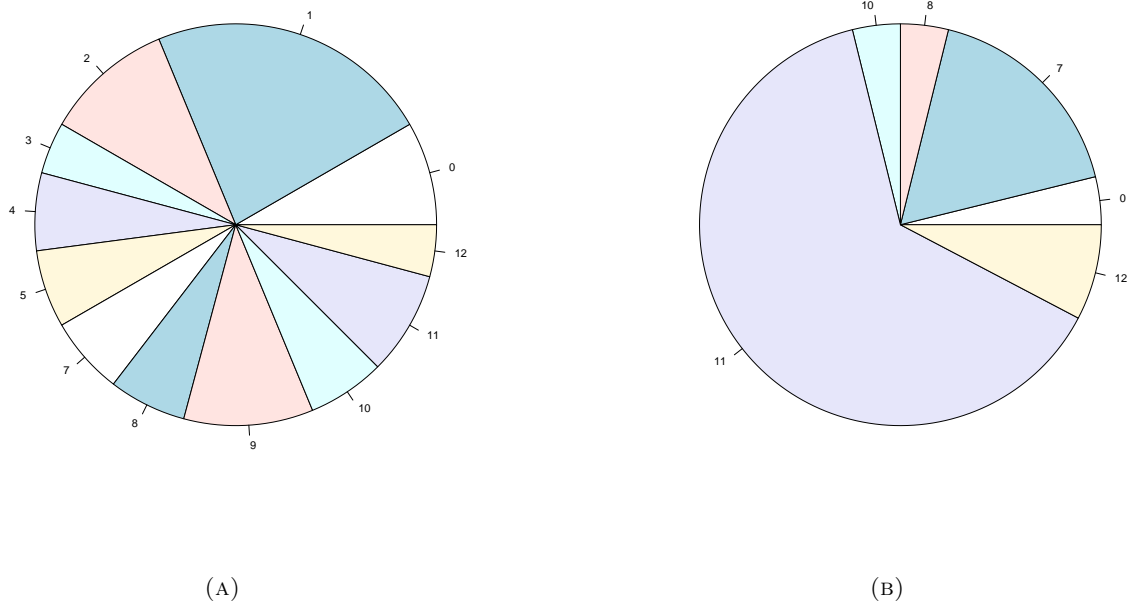


FIGURE 3.2: Neighborhood residency of teenagers (%)  
 (a) stands for teenagers from school M, (b) for teenagers from school S

School mapping is quite different from neighborhood borders, indeed teenagers can come from different areas located in different neighborhoods and inside an area (*IRIS*) teenagers can be associated with different schools. Table 3.3 gives detail information on the areas (*IRIS*) where both schools are located precisely. The median income and the population are calculated in the area where the school are located. The *Iris* considered is thus the heart of the school area, the other adjacent *Iris* are ignored. Here again, the area where school **S** is located is characterized by strong inequalities (the interdecile ratio is equal to 14 while it is equal to 8 in the area where school **M** is located) and by a median income two times lower than in the area where school **M** is located.

Moreover, schools themselves experience different degree of social mixing. Table 3.2 gives some information regarding their social mixity. School **M** is located in Rennes' city center, in this school, 14% of pupils are considered disadvantaged (Merle (2012)) while they are 48% in school **S**.

In a post-experiment questionnaire we asked teenagers their neighborhood residency (see below for more information about this post-experiment questionnaire). Figure 3.2 shows repartition of teenagers in the different neighborhoods. We can see that neighborhood residency is more diversified in school **M** where teenagers live in various neighborhoods, only 23% live in the city center while 63% of teenagers from school **S** live in neighborhood 11 where **S** is located (see also column "same neighborhood" in table 3.2 ).

	Age	Female	Same neighborhood	Disadvantaged <sup>a</sup>	Players
<b>M</b>	14,9	37,50%	23,00%	14,00%	48
<b>S</b>	15,3	36,50%	63,00%	48,10%	52

TABLE 3.2: School comparison

a. Based on parents' socio-professional group (Merle, 2012)

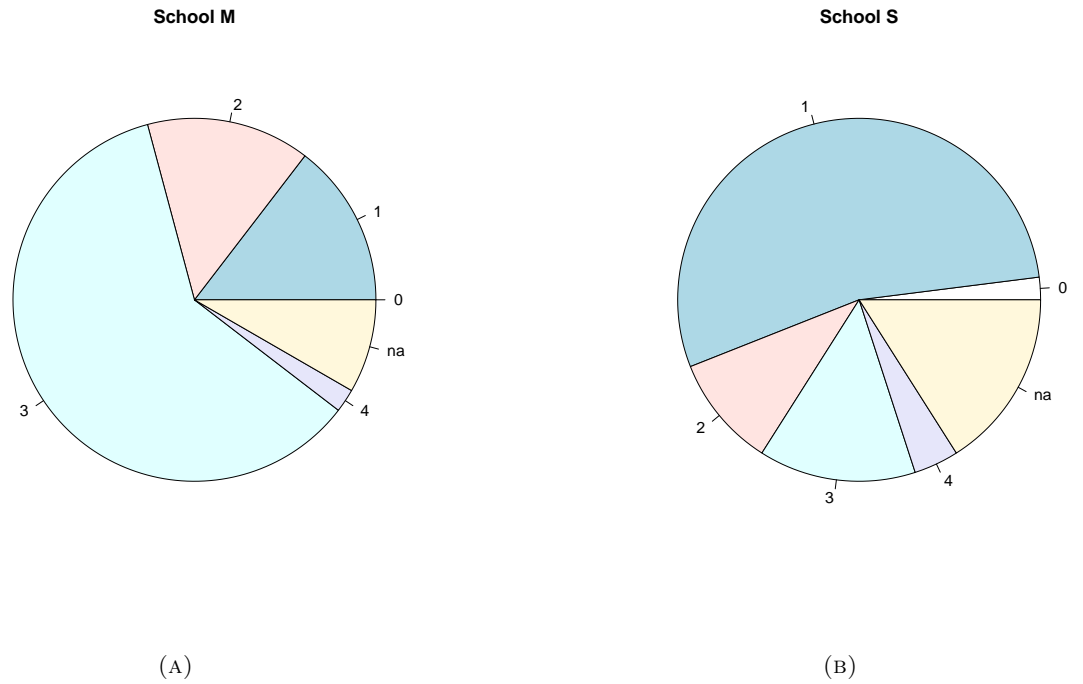


FIGURE 3.3: Father's Socio-Professional Category

0 stands for unemployed, 1 for blue-collar workers, 2 for employees, 3 for executives and 4 for retired.

IRIS	Population (UC)	Median income	Interdecile	Gini
School M	1294,3	21991	8,2	0,4168
School S	1596,3	10115	14,0	0,4138

TABLE 3.3: School' areas comparison

Figure 3.3 shows the different Socio-Professional Category of participants' fathers, 0 stands for unemployed, 1 for blue-collar workers, 2 for employees, 3 for executives and 4 for retired. We can see that most fathers in school **S** are blue-collar workers while in school **M** most fathers occupy an job of executive or assimilated.

## 3.2 Experimental procedure

To investigate the determinants of trust and trustworthiness we use a variant of the trust game introduced by Berg et al. (1995). The general structure of the trust game is

as follows. There are two players, player A and player B, and each of them has an initial endowment, equal to 4 experimental units in our experiment. In a first stage, player A can send between 0 and 4 units to player B. More precisely, she has to choose an integer between 0 and 4, i.e., 0, 1, 2, 3 or 4 points. This transfer is tripled by the experimenter and sent to the player B. In a second stage, the player B can send back between 0 and  $3 \times$  the transfer of A to B. The final payoff of the two players are then :

- for A :  $4 - \text{transfer to player B} + \text{Back transfer from player B}$
- for B :  $4 + 3 \times \text{transfer from player A} - \text{back transfer to player A}$

Therefore, the amount sent by the sender indicates the “trust in others” as defined by Coleman & Coleman (1994) and the amount sent back by the receiver stands for the degree of trustworthiness. See figure 3.4

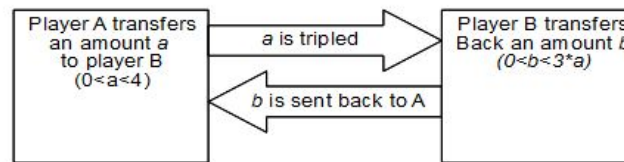


FIGURE 3.4: Sequence of the game

In our experiments, following the strategy method, players can condition their decisions on the school their partner belongs to. That is, each player A makes two decisions : a transfer to a player B in the same school (within-school decision) and a transfer to a player B in a different school (between-school decision).<sup>1 2</sup>. First movers made their decisions on two different sheets. The first sheet stated the following question : **“Suppose the other person belongs to your school<sup>3</sup> ... . How much of your 4 points do you want to send him ?”**. In the second sheet first movers indicated how much they give if the second mover belongs to another middle school. **“Suppose the other person belongs to another school ... . How much of your 4 points do you want to send him ?”**

First movers were also asked to state their beliefs about the second movers’ repayment decision for both cases. Thus first movers made 2 investment decisions and indicated 2

1. The school name and location were not specified

2. Decisions sheets for players A and B can be found in the appendix

3. In that case, they would play with a student from their school belonging to another classroom

beliefs, playing within and between school. Knowing first movers' beliefs allows us to calculate expected back transfers. Similarly, each player B has to decide a back transfer to a player A in the same school and a back transfer to a player in a different school. In order to elicit the willingness to reciprocate of player B, we used once again the strategy method. Then player B has to decide a back transfer for each of the 4 possible transfer amount that the player A can choose ( 1, 2, 3, 4). They had thus two decisions sheets to fill in : **“Suppose player A is from your school, how many points would you like to send him back if he sends you 1 point .... 2 points ... 3 points ... 4 points”** respectively if player A is from another school.

Using the strategy method has several advantages. First it enables us to collect more data on individuals decisions. It also enables us to study the tendency of teenagers to favor someone from their own group. Introduced by [Selten \(1965\)](#) this strategy method has been widely discussed. Besides, one can wonder if this method is not biased as it could drive people to behave more equally in order to not show discriminatory behaviors. [Brandts & Charness \(2011\)](#) compare the strategy method and the direct answers method in a meta analysis. Based on 29 experiments both methods give similar results : no difference exist in 16 over the 29 experiments, significant differences are seen in four experiments and mixed results are found in the 9th other experiments. Most authors insist on the fact that the strategy method is particularly effective in the case of “simple” games. According to [Fong et al. \(2007\)](#) or [Solnick \(2007\)](#) there is no difference between both methods in the trust game. However, [Casari & Cason \(2009\)](#) compare these two methods in a simplified version of the investment game and find no difference for players A but players B tend to return a lesser amount in the strategy method. Finally, here we are interesting in comparing the decisions taken by the participants for a given method.

Our subjects are teenagers in 10th grade in two secondary schools in Rennes. This is important in order to study the determinants of trust and trustworthiness as our subject pool experiment more variance in their socio-economic background than the typical students sample in laboratory. The experiment was conducted in classrooms with pen and papers. In total 100 teenagers took part in the experiment (49 first movers and 51 second movers). The experiment was set up as follows : The instructions of the game and 2 different envelopes containing the two different sheets were disposed on the tables before the experiment. The envelopes were randomly disposed in two different order, first playing with a teenager from their school and then with a teenager from a different school and conversaly. The instructions were read out loud and we asked for understanding questions. Then participants could ask for questions individually answered. Participants opened the first envelope and filled in the first sheet. Then they sealed the envelope and filled in the second sheet. Decisions were thus taken sequentially with no possibility of changing the answers. The envelopes were collected and participants

filled in the additional questionnaire. This questionnaire asked about personal attributes like gender, age, parents' occupation. In addition we elicited information about residential background of students which included questions about neighborhood' residency, duration of residency in the neighborhood, friends' residency. Participants then randomly selected whether they would play with someone from their school or another school by selecting at random a piece of paper. After the experiment we randomly formed pairs for the payments. After a month participants received their payments in the form of gifts vouchers in a sealed envelope with information regarding the school belonging of their partner and the amount received. One experimental unit equals 60 cents, the average payoff is 13.88 euros. The calendar of the session is summarized in figure 3.7

## 4 Results

In this section we detail our results. We first provide general results and then investigate whether we can observe in-group favoritism or discrimination.

### 4.1 General results

**Players A** We first study the results for players A. Our results can be summarized in the following way.

*Result 1.* Regarding players A

- On average first movers reveal a relatively high willingness to trust and invest about 60 percent of their endowment.
- Players A from the non segregated school (**M**) trust more than players from the segregated school (**S**). Indeed, the amount of the transfer is significantly smaller for players A from school **S**, regardless the school belonging of player B.
- Players A exhibit same levels of trust whichever the school belonging of player B.

Figure 3.5 shows the distribution of transfers from first movers in both schools. Table 3.4 indicates the percentage of endowment transferred by players A in both schools. Result 1 is consistent with previous results in the literature where the average transfer is around 60%. Table 3.4 indicates the mean percent of endowment sent by first movers when playing with a player from their school (within-school transfer) and when playing with a player from another school (between-school transfer). Figure 3.5 represents the repartition of these amounts. The amount sent on average by players A from school **M** is higher than the amount sent by players A from school **S**. The figure also shows that the distribution of the transfers is relatively concentrated in the bottom (around 2) for the subjects from school **S**. By contrast, the amounts sent by players A from school **M** are more spread around the upward of the distribution (3 or 4). Wilcoxon tests confirm



the robustness of this result when player B being in the same school ( $p = 0.0472$ ) or, in a lesser extent, player B being in a different school ( $p = 0.1068$ ). Interestingly, players A do not seem (on average) to discriminate regarding the school belonging of player B. 3.5 seems however indicating that players A from school **S** trust more a player B from a different school rather than from the same school. Wilcoxon rank test enable us to conclude that on average there is no discrimination regarding the school belonging of player B :  $p = 0.7041$  for the whole sample,  $p = 0.3173$  for players from school **M** and  $p = 0.8217$  for players from school **S**.

	Within-school	Between-school
All players A	60,70%	59,70%
Players A-School <b>M</b>	67,70%	64,50%
Players A-School <b>S</b>	54,00%	55,00%

TABLE 3.4: Transfers players A (%)

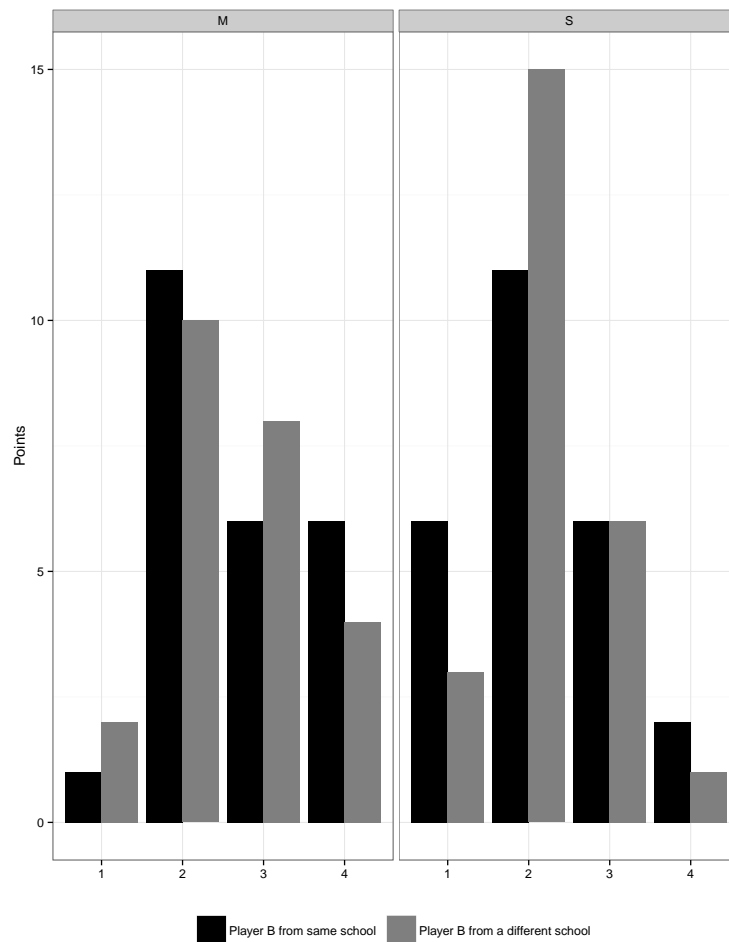


FIGURE 3.5: Distribution transfers A

**Players B** We now turn to the choices made by players B.

*Result 2.* Regarding players B :

- There is no significant difference between players from the non segregated school (**M**) and from the segregated school (**S**) ;
- Similarly, the school belonging of player A does not influence the rate of back transfer from player B.

Table 3.5 shows the average back transfer rates of players B. The first remark is that players belonging to school **M** seem to behave similarly whatever the school belonging of player A. A Wilcoxon test enables to accept the hypothesis which states that the rate of return are identical whatever the school belonging of player A ( $p = 0.9655$ ). By contrast, players from school **S** seem to transfer back a higher amount to players A located in a different school, though this difference is not statistically significant ( $p = 0.5880$ , Wilcoxon test)

Finally, it seems that on average there is not a different behavior between players B from school **M** and from school **S**. In contrast, the repartition of return rates seems to be more concentrated for players from school **M** (around 1/3 and 2/3) than for players from school **S**. However Wilcoxon statistical test indicate that there is no significant difference between the two samples :  $p = 0,9623$  when player A belongs to the same school and  $p = 0,3168$  when he is in a different school.

	Same school	Different school
All players	47.31	50.65
School <b>M</b>	47.40	47.11
School <b>S</b>	47.24	53.81

TABLE 3.5: Back transfer from player B ( % of the transfer)

Figure 3.6 shows the average percentage of back transfers from player B according to the number of points sent by player A. As we used a contingent response method, we have information about the distribution of back-transfers from player B regarding the amount potentially sent by player A.

### Expected and actual returns

While filling the different sheets, players A had to make predictions about the willingness to reciprocate of players B indicating the number of points they think player B will return. On average players A make good predictions as the expected and actual returns are not significantly different. Table 3.6 gives the average percentage of expected and actual returns of player B.

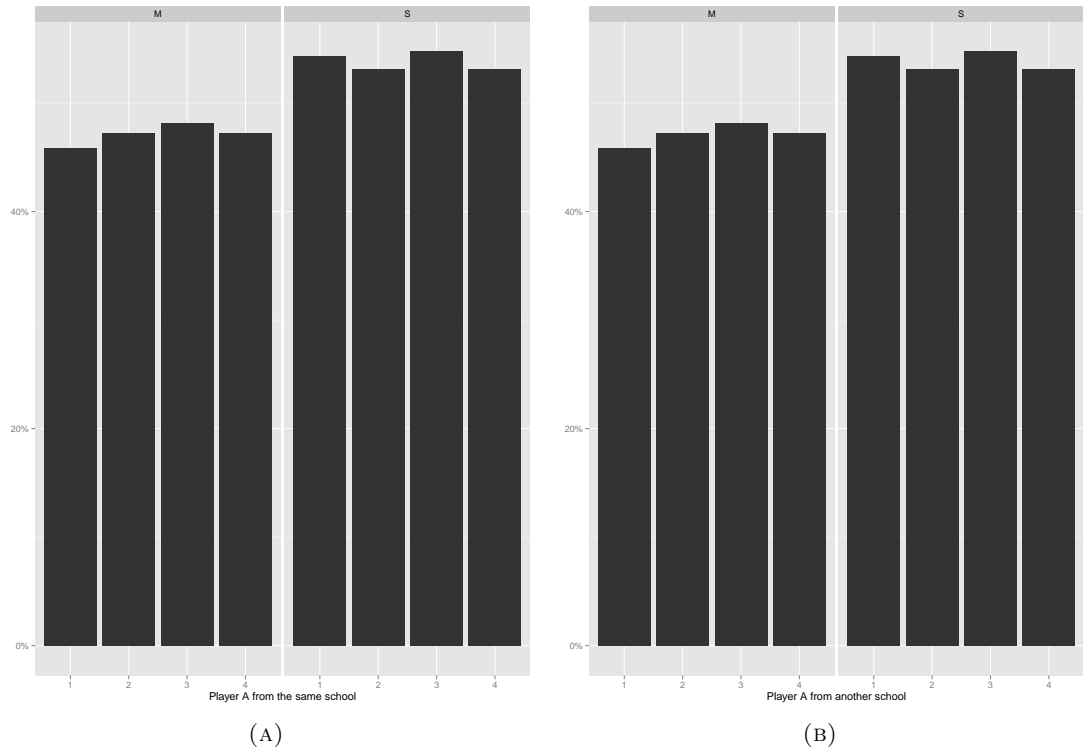


FIGURE 3.6: Distribution back transfers players B

	Within-school		Between-school	
	Expected return	Actual return	Expected return	Actual return
Players A	32,24%	28,71%	32,26%	30,91%
Players A-School M	31,33%	28,31%	30,12%	29,80%
Players A-School S	33,11%	29,07%	34,32%	31,89%

TABLE 3.6: Expected and actual returns

## 4.2 In-group favoritism ?

We study now whether the players behave differently regarding the school belonging of their partner. We distinguish 3 types of attitudes :

- intra-group favoritism : player A always sends more to player B if B is in the same school than in a different school. Respectively player B always send more to a player A from the same school.
- egalitarian behavior : Player A always send the same amount to a player B from the same school and from a different school. Similarly for player B.
- intra-group discrimination : Player A always sends more to a player B from a different school than from the same school. Respectively player B always favor a player A from a different school.

Table 3.7 and 3.8 show the repartition of players regarding these 3 attitudes. Our results can be summarized the following way :

*Result 3.* Players from the non segregated school (**M**) adopt more frequently an egalitarian attitude than players from the segregated school (**S**). These latest discriminate players from a different school as much as players from their school.

**Players A.** A majority of players from school **M** (62,5%) adopt an egalitarian attitude, that is to say they send as many points to a player B in the same school as a player B in a different school. A quarter of them discriminate in favor of their own group while 12.5% discriminate in favor of the other group. Tests of proportion enables to say that a significant higher proportion of players from school **M** adopt an egalitarian behavior than a discriminant behavior (in favor or against their own group) :  $p = 0,0088$  (equality versus intra-group favoritism) and  $p = 0,0003$  (equality versus inter-group favoritism)

Concerning players from school **S**, they uniformly distribute between the 3 attitudes. A  $\chi^2$  test of *goodness-of-fit* indicates that the observed distribution is not significantly different from a uniform distribution ( $p = 0,9608$ ).

	School <b>M</b>	School <b>S</b>
Player A		
Intra-group favoritism	0.25	0.32
Egalitarianism	0.625	0.32
Intra-group discrimination	0.125	0.36

TABLE 3.7: Favoritism players A

	School <b>M</b>	School <b>S</b>
Player B		
Average number of decision/indiv. :		
Intra-group favoritism	0.92	1.19
Egalitarianism	2.04	1.44
Intra-group discrimination	1.04	1.37

TABLE 3.8: Favoritism Players B

**Players B.** Each player B had to take 8 decisions according to the (hypothetic) transfer of player A and regarding the school belonging of the latest. We compare the amount sent back to a player A from the same school to the amount sent to a player A from a different school, given the transfer from player A. We can thus classify the behavior of player B in 4 different cases (each one characterised by the amount of the transfer from player A). Table 3.8 shows the average number of situations in which players B adopt an attitude of intra-group favoritism (that is to say discrimination against the other group), inter-group favoritism or egalitarianism.

Among the players from school **M**, they adopt an egalitarian behavior in average in 2.04 cases among 4. The number of discriminatory decisions against the other group or the own group is 1. Wilcoxon' rank test indicate that there exist significant differences between the number of egalitarian decisions and discriminatory decisions : respectively  $p = 0,0466$  (equality versus intra-group favoritism) and  $p = 0,0766$  (equality versus inter-group favoritism). In contrast there is no significant difference between the two types of discriminatory decisions ( $p = 0,6314$ , Wilcoxon test). Decisions made by players from school **S** tend to be distributed uniformly between the three different behaviors. According to Wilcoxon rank sum test, there is no significant difference :  $p = 0,4375$  (equality versus intra-group favoritism),  $p = 0,7956$  (equality versus inter-group favoritism) and  $p = 0,7882$  (intra-group versus inter-group favoritism).

As each player had two decisions sheets we can analyze whether they discriminate or not regarding the school belonging of their potential partner.

	In-group favoritism	In-group discrimination	Equality
School M	23,00%	25,00%	52,00%
School S	37,00%	35,00%	28,00%

TABLE 3.9: Favoritism according to the payment decisions

Table 3.10 indicates the percentage of in-group favoritism, discrimination or egalitarianism of players B according to the potential transfer of player A (1,2, 3 or 4 points). When a player A potentially send 1 point, 18.52% of players B from school **S** favor their own group while they are 25% in school **M** to send more to a partner from their own school. Some decisions are significantly different depending on the school belonging of player B, we observe significantly more in-group favoritism of players B from the segregated school (**S**) when the potential transfer from player A is 3 (we run a Student test where  $t=0.03$ ). Indeed 44.44% of players B from the segregated school do send a higher back transfer to a partner from their own school while they are 16.67% in school **M** to favor their own group. Similarly players B from school **S** are significantly less egalitarian ( $t=0.007$ ) when players A have potentially sent 3 points. 18.52% of players B from school **S** choose to send the same amount to a player A from his school or from another school having sent 3 points, while they are 54.17% in school **M**.

## 5 concluding remarks

In this chapter we study the relations between social environment and trust. We conduct an extra-laboratory experiment with 100 teenagers in two secondary schools located

in two different neighborhoods in Rennes. We find that players A are characterized by high levels of trust, they send around 60% of their endowment to players B. This result is consistent with previous studies. Players A from the segregated school appear significantly trust-less and send lesser amounts than the participants from the mixed school. Thus the social environment seems to have an effect on trust behavior. Players A from both school do not discriminate and their transfers are not significantly different according to the school belonging of their partner. We find that players A make good predictions about the back transfer from players B as the expected and actual returns are not significantly different.

We then study the existence of in-group favoritism and find no evidence for players A, however we find that players B from the segregated school discriminate more. This discrimination goes in both ways as players B tend to favor equally in-group partners as out-groups partners and appear significantly less equalitarian. In-group favoritism or discrimination can have important economic and social consequences as many socio-economic relations involve some element of trust. At the level of neighborhoods in a city trust discrimination may feed segregation and lead to the spatial separation of groups within the city.

## 6 Appendix

## Joueur A

Tu vas participer à un jeu qui fait partie d'un projet de recherche conduit à l'université de Rennes 1.

Dans ce jeu tu vas pouvoir gagner des points en jouant avec un autre collégien qui peut venir de ton collège ou d'un autre collège participant à ce projet de recherche.

Pendant le jeu, tu devras prendre des décisions. Tes décisions et celles des autres participants sont anonymes. Nous ne communiquerons à personne tes réponses.

Tes décisions te rapporteront des points qui seront par la suite convertis en bons d'achats (Fnac, intersport ...). Ces bons d'achat te seront donnés lorsque nous aurons fait jouer tous les collèges participants.

Pendant toute la durée du jeu, tu ne dois pas communiquer avec les autres joueurs.

Si tu as des questions, n'hésites pas à lever la main et nous viendrons te répondre individuellement.

### Description du jeu :

Dans ce jeu, il y a deux joueurs : joueur A et joueur B.

Le jeu se déroule de la façon suivante :

- Chaque joueur, A et B, dispose de 4 points au départ.
- Le joueur A peut décider d'envoyer tout ou une partie de ses points au joueur B. Il peut donc envoyer 0, 1, 2, 3 ou 4 points au joueur B.
- Le nombre de points envoyés par le joueur A est triplé puis donné au joueur B. Par exemple, si le joueur A a envoyé 2 points au joueur B, le joueur B reçoit  $2 \times 3 = 6$  points.
- Le joueur B décide ensuite de renvoyer tout ou une partie des points qu'il a reçus au joueur A. Il gagne donc 4 points de départ +  $3 \times$  les points reçus – les points qu'il renvoie au joueur A.

Dans ce jeu, tu vas jouer le rôle du **joueur A** : nous allons te demander le nombre de points que tu souhaites envoyer à un autre collégien, qui jouera lui le rôle du joueur B. Ce collégien peut être soit dans ton collège, soit dans un autre collège participant.

Les choix que tu as faits seront ensuite utilisés pour calculer la valeur des bons d'achat que tu vas recevoir. Nous t'expliquerons plus tard comment cette valeur sera calculée.

Quelques exemples :

Lorsque tu envoies un point, le joueur B reçoit 3 points.

Le joueur B peut donc choisir de te renvoyer 0, 1, 2 ou 3 points.

Si le joueur B choisit de te renvoyer 2 points :

Tu gagnes : les 4 points de départ – 1 point que tu as envoyé + les 2 points que tu reçois du joueur B = 5 points.

Le joueur B gagne : les 4 points de départ + 3 x 1 point que tu lui as envoyé – 2 points qu'il t'envoie = 5 points.

*Question de compréhension :*

Passons maintenant une question pour voir si tu as bien compris :

Si tu envoies 3 points, combien de points le joueur B reçoit-il dans un premier temps?  
\_\_\_\_\_ points.

Il décide de te renvoyer 5 points.

Combien de points gagnes-tu ? \_\_\_\_\_ points

Combien de points le joueur B gagne-t-il **au total** ? \_\_\_\_\_ points

Si tu as des questions, n'hésites pas à lever la main, nous viendrons te répondre individuellement.



### Calcul des gains :

Nous allons maintenant t'expliquer comment tes choix vont déterminer les bons d'achat que tu vas recevoir.

En participant à ce jeu tu gagnes d'ores et déjà un bon d'achat de 10 euros valable dans différents magasins.

Tes choix détermineront la valeur du bon d'achat que tu vas recevoir. Cette valeur sera déterminée de la manière suivante :

- 1) Après avoir fait tes choix, tu vas tirer au sort si tu joues avec un collégien de ton propre collège ou d'un autre collège.
- 2) Tes choix seront communiqués à l'autre collégien pris au hasard qui lui jouera le rôle du joueur B.
- 3) Ta décision et la sienne détermineront la valeur de ton bon d'achat ainsi que la valeur du sien.

Chaque point gagné te rapporte 60 centimes d'euros supplémentaires. Par exemple, si le jeu te permet de gagner 5 points, tu gagnes un bon d'achat d'une valeur de 13 euros : 10 euros + 5 points x 60 centimes.

Au maximum, tu peux gagner un bon d'achat d'une valeur de 17,20 euros.

Le bon d'achat te sera remis lorsque nous aurons réalisé les sessions dans les autres collèges .

Nous allons maintenant pouvoir commencer le jeu :

Nous te rappelons que tu joues le rôle du joueur A.

Tu as devant toi 2 enveloppes :

Dans l'une tu vas décider du nombre de points à envoyer à un collégien **de ton collègue**.

Dans l'autre, tu vas décider du nombre de points que tu souhaites envoyer à un collégien **d'un autre collège.**

**Tu peux désormais ouvrir la première enveloppe, celle qui se trouve sur le dessus. Puis tu peux remplir la feuille qui se trouve à l'intérieur.**

Lorsque c'est fait, referme l'enveloppe, tu peux ensuite ouvrir la **seconde enveloppe et remplir la feuille qui se trouve à l'intérieur.**

Tu peux maintenant tirer au sort un papier qui t'indiquera si la valeur du bon d'achat sera déterminée par le jeu avec un collégien de ton collège ou d'un autre collège.

Le jeu est maintenant terminé, nous te remercions pour ta participation, aurais-tu des remarques, critiques, suggestions par rapport à ce jeu ?

[illegible]

**MERCI !**

## Questionnaire collégien

N° d'anonymat : \_\_\_\_\_

### Instructions :

Dans ce questionnaire on te pose des questions sur toi, ta famille, tes ami(e)s, ce que tu aimes faire. Ce n'est pas un test, il n'y a pas de bonne ou mauvaise réponse. Tes réponses seront anonymes, ton nom et ton prénom ne seront pas communiqués. Prends le temps de répondre à chaque question et d'indiquer ce que tu ressens vraiment, en cas de doute tu peux nous demander de l'aide.

-----

### Moi et ma famille

Je suis

- ☐ Un garçon  
☐ Une fille

Année de naissance : .....

Lieu de naissance : .....

Combien as-tu de frères et sœurs ?

..... Frère(s) et ..... sœur(s)

Combien de tes frères et sœurs vivent avec toi à la maison ? .....

Quel est le métier de tes parents ou tuteurs légaux?

Père : .....

Mère : .....

### Moi et le collège

Depuis combien de temps es-tu inscrit dans ce collège ? .....

Avant d'entrer au collège, dans quelle école primaire étais-tu ?

Nom de l'école : ..... Ville de l'école : .....

Aimerais-tu changer de collège ? ☐ Oui ☐ Non

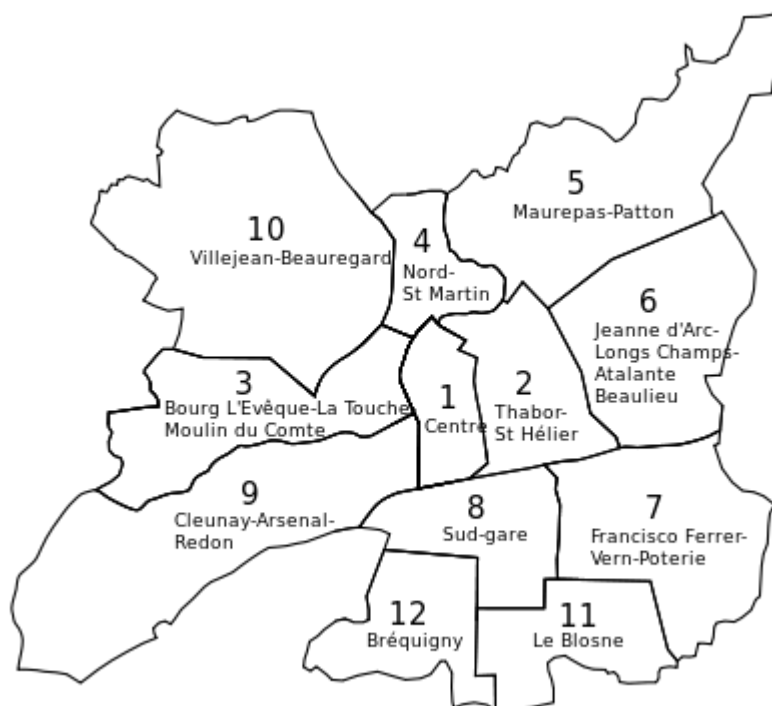
Si tu devais changer de collège, parmi les collèges suivants, dans lequel souhaiterais-tu aller ?

	Oui	Non	Indifférent
Echange (quartier centre)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rosa Parks (Villejean)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
La Biquenais (Blosne)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zola (Centre)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Y a-t-il un autre collège où tu souhaiterais aller ? .....

Moi et mon quartier

Voici la carte des quartiers rennais,



- ☐ Entoure en rouge le quartier où tu habites
- ☐ Fais une croix dans le ou les quartiers où habitent tes amis (y compris le tien)
- ☐ Souligne le ou les quartiers où tu n'es jamais allé

Si tu habites à Rennes, dans quel quartier habites-tu ? .....

Depuis combien de temps y habites-tu ? .....

Si tu habites dans une autre commune, quel est le nom de cette commune ? .....

Depuis combien de temps y habites-tu ? .....

Aimerais-tu déménager ?

- ☐ Oui
- ☐ Non

Si tu devais déménager, indiques si tu aimerais habiter dans les quartiers suivants :

	Oui	Non	Indifférent
Villejean Beauregard	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nord Saint Martin	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maurepas Patton	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jeanne D'arc- Longs-champs- Beaulieu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Thabor Saint Hélier	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Centre	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Moulin du Comte Bourg Levesque, La touche	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cleunay- Arsenal- Redon	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sud Gare	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Brequigny	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Le Blosne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Francisco Ferrer- La Poterie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Moi et mes amis

Parmi tes amis, combien y a-t-il de garçons que tu considères comme tes meilleurs amis ?  
.....

Et combien y a-t-il de filles que tu considères comme tes meilleures amies ? .....

Parmi tes meilleurs amis ou amies ,

- Combien sont dans ton collège ? .....
- Combien sont dans un autre collège ? .....
- Combien habitent dans ton quartier ? .....
- Combien habitent dans un autre quartier rennais ? .....
- Combien habitent hors de Rennes ? .....

Ce questionnaire est maintenant terminé, nous te remercions d'y avoir répondu. Aurais-tu des critiques, des suggestions ? Certains problèmes qui te paraissent importants auraient-ils dû être évoqués ?

.....

.....

.....

.....

.....

.....

.....

**MERCI !**

---

	Intra-group favoritism		Egalitarianism		intra-group discrimination	
	<b>M</b>	<b>S</b>	<b>M</b>	<b>S</b>	<b>M</b>	<b>S</b>
1	25	18.52	58.33	51.85	16.67	29.63
2	25	25.93	50	37.04	25	37.04
3	16.67	44.44	54.17	18.52	29.17	37.04
4	25	29.63	41.67	37.04	33.33	33.33

---

TABLE 3.10: Players B' favoritism (%)



Si le joueur B vient **de ton collège**,  
combien de points souhaites-tu lui  
envoyer ?

☐0 ☐1 ☐2 ☐3 ☐4

Il va donc recevoir .....Points

Combien de points penses tu qu'il va  
te renvoyer ?

.....Points

Si le joueur B vient **d'un autre collège**,  
combien de points souhaites-tu lui  
envoyer ?

☐0 ☐1 ☐2 ☐3 ☐4

Il va donc recevoir .....Points

Combien de points penses tu qu'il va  
te renvoyer ?

.....Points

Le joueur A vient **de ton collège**, combien de points souhaites tu lui renvoyer :

S'il t'a envoyé **1 point** :

☐0☐1☐2☐3

S'il t'a envoyé **2 points** :

☐0☐1☐2☐3☐4☐5☐6

S'il t'a envoyé **3 points** :

☐0☐1☐2☐3☐4☐5☐6☐7☐8☐9

S'il t'a envoyé **4 points** :

☐0☐1☐2☐3☐4☐5☐6☐7☐8☐9☐10☐11

☐12

Le joueur A vient **d'un autre collège**, combien de points souhaites tu lui renvoyer :

S'il t'a envoyé **1 point** :

☐0☐1☐2☐3

S'il t'a envoyé **2 points** :

☐0☐1☐2☐3☐4☐5☐6

S'il t'a envoyé **3 points** :

☐0☐1☐2☐3☐4☐5☐6☐7☐8☐9

S'il t'a envoyé **4 points** :

☐0☐1☐2☐3☐4☐5☐6☐7☐8☐9☐10☐11

☐12

Sequence of experiments

May :                    **School M**  
                              Players A  
                              **School S**  
                              Players A  
June :                   **School M**  
                              Players B  
                              **School S**  
                              Players B

July : Payments of earnings with gift vouchers.

# General Conclusion

L'objet de cette thèse est à l'intersection de l'étude de la ségrégation urbaine et de la prise en compte de l'identité dans la compréhension des phénomènes et interactions économiques. Rédigée sous la forme de trois essais qui peuvent se lire de façon indépendante, elle propose dans un premier chapitre une analyse de la ségrégation résidentielle en France. Nous avons proposé une étude d'une famille d'indices de ségrégation basée sur la décomposition d'indices d'inégalité. Revenant sur la définition de la ségrégation, nous proposons de mesurer la ségrégation basée sur une variable continue, le revenu. Ce type d'études est rare en France où la ségrégation est le plus souvent mesurée selon des catégories (CSP, origine ethnique). L'importance de la ségrégation et l'objectif de mixité sociale dans le débat public rendent pourtant nécessaire ce type d'étude afin d'évaluer les politiques publiques et mesurer l'ampleur du phénomène. Nous contribuons au débat en proposant une mesure de la ségrégation résidentielle en France dans les 30 plus grandes aires urbaines sur la période 2001-2008, cette mesure permet notamment de voir que le phénomène a peu augmenté voir diminué, et de comparer les villes entre elles.

Nous proposons ensuite dans un second chapitre une étude de l'identité ethnique des individus en France. Les questions d'identité sont récentes en économie mais de plus en plus développées, définies comme le sentiment que l'on a de soi par [Akerlof & Kranton \(2000\)](#), cette variable peut être intégrée dans la fonction d'utilité des individus. Au cœur du débat politique, la notion d'identité est importante notamment car elle a des effets sur les préférences des individus et impactent donc leurs choix et par conséquent leurs résultats socio-économiques. Il semble donc important de pouvoir mesurer cette variable notamment en vue de la mise en place de politiques publiques. L'objet du chapitre 2 est double, nous proposons tout d'abord un apport méthodologique sur la mesure de l'identité ethnique. Elle intéresse les économistes notamment dans les études sur l'intégration des migrants, qui en migrant peuvent faire face à un arbitrage entre adopter la culture d'accueil ou conserver leur héritage culturel. Le résultat de cet arbitrage a des conséquences sur leurs résultats socio-économiques. Le plus souvent approchée par l'identité nationale- la réponse à la question « je me sens [français, marocain, italien...]-

l'identité ethnique est difficile à mesurer. Certaines études proposent de mesurer l'identité via une analyse factorielle sur des critères identitaires définis a priori, c'est le cas notamment de l'ethnosizer de [Constant et al. \(2009\)](#). Similairement, nous proposons une mesure de l'identité ethnique via un indice construit grâce à une analyse factorielle sur des critères identitaires définis a posteriori, résultant de cette analyse factorielle même. Cet indice construit à partir de l'analyse des réponses à la question « parmi ces 14 caractéristiques, lesquelles vous définissent le plus » obtenues dans la base de données originale Trajectoires et Origines (TeO) permet de capturer une certaine distance identitaire entre les migrants et les natifs. Cette distance identitaire est associée à la notion d'assimilation, c'est à dire la proximité à la culture d'accueil majoritaire. Nous comparons cet indice d'assimilation à la mesure plus courante d'identité nationale. Nous montrons que l'assimilation et le sentiment d'être français ne coïncident pas toujours. Il semble que plus le pays d'origine est distant, à la fois en termes économiques et culturels, plus le sentiment d'être français est fort et plus le degré d'assimilation est faible. Dans un second temps nous nous intéressons à la formation de l'identité en étudiant les déterminants de l'assimilation. Nous proposons d'étudier l'impact de certaines caractéristiques socio-démographiques sur le degré d'assimilation des individus. Nous montrons notamment que l'identité des enfants d'immigrés converge vers celle des natifs, que celle ci soit approchée par le sentiment national ou le degré d'assimilation. Enfin dans un troisième chapitre nous mêlons l'étude de l'identité et de la ségrégation en nous intéressant aux effets de la ségrégation sur les préférences individuelles et notamment le degré de confiance aux autres. L'absence de confiance peut avoir des effets importants sur les interactions économiques et par là même avoir des conséquences sur les résultats socio-économiques des individus. En outre, les individus peuvent discriminer leurs semblables selon leur identité sociale, c'est à dire le groupe auquel ils appartiennent. [Tajfel \(1970\)](#) a d'ailleurs montré dans son paradigme du groupe minimal que de la discrimination entre groupes pouvait avoir lieu même pour des groupes créés sur des critères artificiels comme une couleur de ralliement. Des groupes sociaux peuvent être basés sur différents critères comme l'origine ethnique, le lieu de résidence, ou bien encore, comme c'est le cas dans notre étude, sur le collège d'appartenance. Nous avons mené une étude expérimentale dans des collèges rennais afin d'étudier l'existence de favoritisme intra-groupe ainsi qu'évaluer les effets de la ségrégation sur la confiance aux autres. Nous nous intéressons à des collégiens afin de réduire les effets d'endogénéité sur les choix de localisation des individus. L'expérience de terrain a été réalisée au sein de deux collèges qui se différencient par leur localisation géographique, étant situés dans des quartiers différents en termes de ségrégation. Nous montrons que les collégiens du collège ségrégué font globalement moins confiance que les collégiens dans le collège mixte. Nous ne trouvons pas d'évidence claire pour l'existence de favoritisme intra-groupe, nous observons cependant que les participants du collège ségrégué se comportent de manière

---

moins égalitaire favorisant cependant autant les collégiens d'un collège différent que du même collège.



# Table des figures

1.1	Ranking correlation between GSI and NSI ( <i>IRIS</i> scale)	22
1.2	Ranking correlation between GSI and NSI ( <i>Grand quartier scale</i> )	24
1.3	Ranking over time ( <i>IRIS</i> scale)	26
1.4	Value over time ( <i>IRIS</i> scale)	26
1.5	The value of GSI according to the scale area unit	28
1.6	The value of NSI according to the scale area unit	28
1.7	Correlation between GSI and Coefficient of Variation ( <i>IRIS</i> scale)	31
1.8	Correlation between NSI and Coefficient of Variation ( <i>IRIS</i> scale)	32
3.1	Map of Rennes	71
3.2	Neighborhood residency of teenagers (%) (a) stands for teenagers from school M, (b) for teenagers from school S	73
3.3	Father's Socio-Professional Category 0 stands for unemployed, 1 for blue-collar workers, 2 for employees, 3 for executives and 4 for retired.	74
3.4	Sequence of the game	75
3.5	Distribution transfers A	78
3.6	Distribution back transfers players B	80
3.7	Sequence of experiments	96





# Liste des tableaux

1.1	The data . . . . .	19
1.2	Income segregation and income inequality (weighted mean) . . . . .	21
1.3	Mean Gini Inequality coefficients and coefficient of variation of individual and neighborhood income distributions, and economic segregation measures	23
1.4	Pearson's correlation coefficients ( <i>IRIS</i> scale) . . . . .	23
1.5	Pearson's correlation coefficients ( <i>Grand Quartier</i> scale) . . . . .	24
1.6	NSI and GSI values per region, 2008 . . . . .	27
1.7	Tobit regressions . . . . .	30
2.1	Weights for assimilation index . . . . .	45
2.2	Predictions of each attribute by the index only . . . . .	46
2.3	Means of ethnic identity measures . . . . .	48
2.4	Weighted shares of origins . . . . .	50
2.5	National identity and assimilation index for the whole population . . . . .	52
2.6	Comparison of Identities between First and Second Generations . . . . .	56
3.1	Neighborhoods . . . . .	72
3.2	School comparison . . . . .	74
3.3	School' areas comparison . . . . .	74
3.4	Transfers players A (%) . . . . .	78
3.5	Back transfer from player B ( % of the transfer) . . . . .	79
3.6	Expected and actual returns . . . . .	80
3.7	Favoritism players A . . . . .	81
3.8	Favoritism Players B . . . . .	81
3.9	Favoritism according to the payment decisions . . . . .	82
3.10	Players B' favoritism (%) . . . . .	93



# Références

- Abdelal, R. (2009). *Measuring identity : A guide for social scientists*. Cambridge University Press.
- Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *Quarterly journal of Economics*, 715–753.
- Alesina, A., & Ferrara, E. L. (1999). *Participation in heterogeneous communities* (Rapport technique). National bureau of economic research.
- Algan, Y., & Cahuc, P. (2013). Trust, growth and well-being : New evidence and policy implications.
- Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, 343–362.
- Austen-Smith, D., & Fryer Jr, R. G. (2005). An economic analysis of” acting white”. *The Quarterly Journal of Economics*, 551–583.
- Battu, H., & Zenou, Y. (2010). Oppositional identities and employment for ethnic minorities : Evidence from england\*. *The Economic Journal*, 120(542), F52–F71.
- Benabou, R. (1991). *Workings of a city : location, education, and production*. National Bureau of Economic Research Cambridge, Mass., USA.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1), 122–142.
- Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied psychology*, 46(1), 5–34.
- Bisin, A., Patacchini, E., Verdier, T., & Zenou, Y. (2010). *Bend it like beckham : Ethnic identity and integration* (Rapport technique). National Bureau of Economic Research.
- Bisin, A., Patacchini, E., Verdier, T., & Zenou, Y. (2011). Ethnic identity and labour market outcomes of immigrants in europe. *Economic Policy*, 26(65), 57–92.

- Bloemraad, I. (2007). Unity in diversity? bridging models of multiculturalism and immigrant integration. *Du Bois Review*, 4(2), 317–336.
- Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method : a first survey of experimental comparisons. *Experimental Economics*, 14(3), 375–398.
- Buchan, N. R., Johnson, E. J., & Croson, R. T. (2006). Let's get personal : An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior & Organization*, 60(3), 373–398.
- Cadsby, C. B., Song, F., & Bi, Y. (2008). Trust, reciprocity and social distance in china : an experimental investigation. *University of Guelph*.
- Casari, M., & Cason, T. N. (2009). The strategy method lowers measured trustworthy behavior. *Economics Letters*, 103(3), 157–159.
- Casey, T., & Dustmann, C. (2010). Immigrants' identity, economic outcomes and the transmission of identity across generations. *The Economic Journal*, 120(542), F31–F51.
- Castells, M. (1972). La question urbaine.
- Chakravarty, S. R., & Silber, J. (2007). A generalized index of employment segregation. *Mathematical Social Sciences*, 53(2), 185–195.
- Citrin, J., & Sears, D. O. (2009). Balancing national and ethnic identities the psychology of pluribus unum. *Measuring identity : A guide for social scientists*, 145.
- Coleman, J. S., & Coleman, J. S. (1994). *Foundations of social theory*. Harvard university press.
- Constant, A. F., Gataullina, L., & Zimmermann, K. F. (2009). Ethnosizing immigrants. *Journal of Economic Behavior & Organization*, 69(3), 274–287.
- Constant, A. F., & Zimmermann, K. F. (2007). Measuring ethnic identity and its impact on economic behaviour. *Vol.*
- Cutler, D. M., & Glaeser, E. L. (1995). *Are ghettos good or bad ?* (Rapport technique). National Bureau of Economic Research.
- Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (2008). When are ghettos bad ? lessons from immigrant segregation in the united states. *Journal of Urban Economics*, 63(3), 759–774.

- Davidoff, T. (2005). Income sorting : Measurement and decomposition. *Journal of Urban Economics*, 58(2), 289–303.
- Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American sociological review*, 210–217.
- Durkheim, E. (2013). *De la division du travail social*. Presses Électroniques de France.
- Echenique, F., Fryer, R. G., & Kaufman, A. (2006). Is school segregation good or bad ? *The American economic review*, 265–269.
- Etang, A., Fielding, D., & Knowles, S. (2011). Does trust extend beyond the village ? experimental trust and social distance in cameroon. *Experimental Economics*, 14(1), 15–35.
- Fack, G., & Grenet, J. (2011). Que peut-on attendre de la réforme de la sectorisation en france ? quelques enseignements des politiques de choix scolaire. *Revue d'économie politique*, 120(5), 709–737.
- Falk, A., & Zehnder, C. (2007). Discrimination and in-group favoritism in a citywide trust experiment.
- Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2003). A nationwide laboratory examining trust and trustworthiness by integrating behavioural experiments into representative surveys.
- Fershtman, C., & Gneezy, U. (2001). Discrimination in a segmented society : An experimental approach. *Quarterly Journal of Economics*, 351–377.
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears : An application to educational enrollments in states of india\*. *Demography*, 38(1), 115–132.
- Fong, Y.-f., Huang, C.-Y., & Offerman, T. (2007). Guilt driven reciprocity in a psychological signaling game.
- Fukuyama, F. (1995). Trust : the social virtues and the creation of prosperity. *World and I*, 10, 264–268.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, 811–846.
- Glaeser, E. L., La Porta, R., Lopez-de Silanes, F., & Shleifer, A. (2004). Do institutions cause growth ? *Journal of economic Growth*, 9(3), 271–303.

- Goux, D., & Maurin, E. (2007). Close neighbours matter : Neighbourhood effects on early performance at school\*. *The Economic Journal*, 117(523), 1193–1215.
- Greenacre, M. (2005). From correspondence analysis to multiple and joint correspondence analysis. *Available at SSRN 847664*.
- Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk attitudes of children and adults : Choices over small and large probability gains and losses. *Experimental Economics*, 5(1), 53–84.
- Hardman, A., & Ioannides, Y. M. (2004). Neighbors' income distribution : economic segregation and mixing in us urban neighborhoods. *Journal of Housing Economics*, 13(4), 368–382.
- Holm, H. J., & Danielson, A. (2005). Tropic trust versus nordic trust : Experimental evidence from tanzania and sweden\*. *The Economic Journal*, 115(503), 505–532.
- Hutchens, R. (2001). Numerical measures of segregation : desirable properties and their implications. *Mathematical social sciences*, 42(1), 13–29.
- Hutchens, R. (2004). One measure of segregation\*. *International Economic Review*, 45(2), 555–578.
- Hutchens, R. M. (1991). Segregation curves, lorenz curves, and inequality in the distribution of people across occupations. *Mathematical Social Sciences*, 21(1), 31–51.
- Ioannides, Y. M., & Seslen, T. N. (2002). Neighborhood wealth distributions. *Economics Letters*, 76(3), 357–367.
- Jargowsky, P. A. (1996). Take the money and run : economic segregation in us metropolitan areas. *American sociological review*, 984–998.
- Jargowsky, P. A., & Kim, J. (2005). A measure of spatial segregation : The generalized neighborhood sorting index. *University of Texas, Dallas*.
- Jargowsky, P. A., & Kim, J. (2009). The information theory of segregation : Uniting segregation and inequality in a common framework. *Occupational and residential segregation*, 17.
- Karlan, D. S. (2005). Using experimental economics to measure social capital and predict financial decisions. *American Economic Review*, 1688–1699.
- Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly journal of economics*, 1251–1288.
- Lagrange, H. (2013). *Le déni des cultures*. Seuil.

- Lee, T. (2009). Between social theory and social science practice : Toward a new approach to the survey measurement of “race”. In R. Abdelal, Y. M. Herrera, A. I. Johnston, & R. McDermott (Eds.), *Measuring identity* (pp. 113–144). Cambridge University Press. Consulté sur <http://dx.doi.org/10.1017/CB09780511810909.005> (Cambridge Books Online)
- Manning, A., & Roy, S. (2010). Culture clash or culture club ? national identity in britain\*. *The Economic Journal*, 120(542), F72–F100.
- Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social forces*, 67(2), 281–315.
- Maurin, E. (2004). Le ghetto français. *Enquête sur le séparatisme social*.
- Merle, P. (2012). Carte scolaire et ségrégation sociale des établissements. une analyse monographique des collèges rennais. *Espaces et sociétés*, 151(3), 103–121.
- Nekby, L., & Rodin, M. (2007). Acculturation identity and labor market outcomes.
- Netzer, R. J., & Sutter, M. (2009). *Intercultural trust. an experiment in austria and japan* (Rapport technique). Working Papers in Economics and Statistics.
- Noiriel, G. (2002). *Atlas de l’immigration en france : exclusion, intégration*–. Editions Autrement.
- OCDE. (2008). *Society at a glance 2009 : Oecd social indicators*. OECD Publishing.
- Pan Ké Shon, J.-L. (2010). The ambivalent nature of ethnic segregation in france’s disadvantaged neighbourhoods. *Urban Studies*, 47(8), 1603–1623.
- Pendakur, K., & Pendakur, R. (2005). *Ethnic identity and the labour market*. Vancouver Centre of Excellence.
- Préteceille, E. (2006). La ségrégation sociale a-t-elle augmenté ? *Sociétés contemporaines*(2), 69–93.
- Reardon, S. F., & O’Sullivan, D. (2004). Measures of spatial segregation. *Sociological methodology*, 34(1), 121–162.
- Safi, M. (2009). La dimension spatiale de l’intégration : évolution de la ségrégation des populations immigrées en france entre 1968 et 1999.
- Schelling, T. C. (1971). Dynamic models of segregation†. *Journal of mathematical sociology*, 1(2), 143–186.
- SCHNAPPER, D. (2007). Qu’est-ce que l’intégration ?



- Selten, R. (1965). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperimentes..
- Sen, A. (2007). *Identity and violence : The illusion of destiny*. Penguin Books India.
- Shorrocks, A., & Wan, G. (2005). Spatial decomposition of inequality. *Journal of Economic Geography*, 5(1), 59–81.
- Solnick, S. J. (2007). Cash and alternate methods of accounting in an experimental game. *Journal of Economic Behavior & Organization*, 62(2), 316–321.
- Sylvan, D. A., & Metskas, A. K. (2009). Trade-offs in measuring identities : A comparison of five approached. *Measuring Identity : A Guide for Social Science*, 72–112.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5), 96–102.
- Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information/sur les sciences sociales*.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations*, 33(47), 74.
- Verdugo, G. (2011). Public housing and residential segregation of immigrants in france, 1968-1999.
- Watson, T. (2009). Inequality and the measurement of residential segregation by income in american neighborhoods. *Review of Income and Wealth*, 55(3), 820–844.
- Wheeler, C. H., & La Jeunesse, E. A. (2006). Neighborhood income inequality. *FRB of St. Louis Working Paper No.*
- White, M. J. (1983). The measurement of spatial segregation. *American journal of sociology*, 1008–1018.
- Willinger, M., Keser, C., Lohmann, C., & Usunier, J.-C. (2003). A comparison of trust and reciprocity between france and germany : experimental investigation based on the investment game. *Journal of Economic Psychology*, 24(4), 447–466.
- Yang, R., & Jargowsky, P. A. (2006). Suburban development and economic segregation in the 1990s. *Journal of Urban Affairs*, 28(3), 253–273.
- Zak, P. J., & Knack, S. (2001). Trust and growth. *The economic journal*, 111(470), 295–321.



VU :  
**Le Directeur de Thèse**  
(Nom et Prénom)

VU :  
**Le Responsable de l'École Doctorale**

**VU pour autorisation de soutenance**

**Rennes, le**

**Le Président de l'Université de Rennes 1**

**Guy CATHELINEAU**

**VU après soutenance pour autorisation de publication :**

**Le Président de Jury,**  
(Nom et Prénom)



## **Essais sur la ségrégation et l'identité en France**

Cette thèse porte sur l'étude des phénomènes de ségrégation urbaine, c'est à dire l'isolement à la fois social et spatial de certains groupes d'individus au sein des villes. Également, afin de caractériser les groupes d'individus et mesurer leur degré d'intégration sociale, cette thèse s'intéresse à la définition et à la mesure de l'identité, définie comme le sentiment que l'on a de soi. Dans un premier chapitre nous nous intéressons à une mesure de la ségrégation urbaine, en proposant une analyse théorique et empirique d'un indice de ségrégation économique sur une variable continue (les revenus) en France. Nous nous intéressons ensuite dans un second chapitre à l'intégration sociale des immigrants et enfants d'immigrants en France où l'identité ethnique des individus est une expression de cette intégration. Ce chapitre introduit un indice qui permet une mesure continue de l'identité ethnique et permet de capturer une certaine distance identitaire. Nous comparons cet indice à d'autres mesures usuelles de la littérature et étudions les déterminants socio-économiques de l'identité ethnique. Enfin, dans un troisième chapitre, nous nous intéressons aux effets de la ségrégation urbaine sur le degré de coopération et de confiance des individus. Il s'agit dans ce chapitre d'étudier l'impact que peut avoir le quartier de résidence d'un individu sur sa confiance aux autres. Pour cela, nous réalisons une étude expérimentale dans différents quartiers rennais auprès de collégiens de 3ème.

Mots clés : Ségrégation ; Revenu ; Identité ; Identité Ethnique ; Assimilation ; Immigrants ; Ville.

---

## **Essays on segregation and identity in France**

This thesis focuses on the study of urban segregation, ie the isolation both social and spatial of certain groups of individuals within cities. Also, to characterize groups of individuals and measure their degree of social integration, this thesis focuses on the definition and measurement of identity, defined as a person's sense of self. In the first chapter we focus on a measure of urban segregation, proposing a theoretical and empirical analysis of an index of Economic segregation on a continuous variable (the income) in France. We then are interested in a second chapter by the social integration of immigrants and children of immigrants in France where the ethnic identity of individuals is an expression of this integration. This chapter introduces an index that allows continuous measurement of ethnic identity and captures a certain distance in identity. We compare this index with other usual measures in the literature and study the socio-economic determinants of ethnic identity. Finally, in the third chapter, we focus on the effects of urban segregation in the degree of cooperation and trust of individuals. We study the potential impact of the neighborhood residency of a person on his trust. For this, we perform an experimental study in different schools in Rennes among teenagers.

Keywords : Segregation ; Income ; Identity ; Ethnic Identity ; Assimilation ; Migrants ; City.