



Occupational mismatch and network effects: Evidence from France

Arnaud Herault

► **To cite this version:**

Arnaud Herault. Occupational mismatch and network effects: Evidence from France. Journées de Microéconomie Appliquée, Jun 2019, Casablanca, Morocco. hal-02860040

HAL Id: hal-02860040

<https://hal.univ-angers.fr/hal-02860040>

Submitted on 8 Jun 2020

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.

Occupational mismatch and network effects: Evidence from France

Arnaud Herault*

Abstract

How does the social environment of immigrants influence the probability of being in an occupational mismatch situation? To answer this question, we use the Labor Force Survey (2005-2012) to assess the impact of peers and the neighborhood on the use of referees to find a job on the one hand, and the probability of being in occupational mismatch situation on the other hand. With a probit model, we estimate the probability of using a referee to find a job as well as the probability of being in an occupational mismatch situation for immigrants. Endogeneity is controlled with a recursive bivariate probit model for the use of a referee to find a job and the probability of being in an occupational mismatch situation. The results show that the neighborhood effect has a greater effect than the peer effect on using referees to find a job. Moreover, the role of the referee on the probability of being in an occupational mismatch situation is not homogeneous according to the origins.

Keywords: Occupational mismatch; immigration; labor market; networks; neighborhood

JEL Code: J24; J61; R23

*GRANEM | Université d'Angers, 13 allée François Mitterrand, 49036 ANGERS. Contact: arnaud.herault@univ-angers.fr | (+33)6.82.71.80.11

1 Introduction

The process of integration of immigrants in host countries is determined by various factors relating to the origin of individuals (Europeans VS non-Europeans, Bisin et al. (2011)), their place of residence (Kain (1968), Edin, Fredriksson, and AAslund (2003)), the occupational choice of immigrants and their descendants (Constant and Zimmermann (2003)) and social relations (Munshi (2003), Patel and Vella (2013)). Over the past decades, the labor market performance of immigrants and their descendants has been analyzed to understand what factors influence - or not - the process of integration into the labor market. The interest of the analysis of these determinants is to model immigrants' behaviors and actions on the labor market. From the point of view of public policies, the objective of the analysis is to be able to adapt the determinants in order to maximize the use of the skills available on a territory.

For immigrants, holding a job for which they are qualified allows them to maximize their income. The advantage of having a perfect match between the skills of the individual and the skill-requirement of the job is that society benefits from this full use of the resources available in the territory to maximize revenues.

The analysis of the integration of immigrants into the labor market has been supplemented by the analysis of the networks of relations. Social networks, since the work of Granovetter (1974), have highlighted the importance of the transmission of information on the labor market in the performance of individuals in this market. The social relations of individuals can be considered according to different prisms. Social relations can be characterized from the social environment, the residential environment of individuals (Bisin and Verdier (2001)) or according to the groups belonging to individuals, social networks defined from the concept of homophily (Granovetter (1974)).

The objective of this paper is to analyze to what extent social relations affect the choice of occupation of immigrants. Does the social network allow the immigrant to validate his skills to obtain a job for which his skills are perfectly matched? What effects, between the peer effects and the neighborhood effects, have the most impact on the probability for an immigrant to be overqualified? The contribution of the paper is twofold. On the one hand, we use the French Labor Force Survey to identify the role of the network effect on the occupational mismatch for different origins. On the other hand, the network effect is analyzed from two different points of view: the concentration of immigrants according to their origin in a region and the influence of the neighborhood.

We use data from the French Labor Force Survey for the period 2005-2012. The information available in this database allows us to identify both the origins of immigrants (and their date of

arrival in France) and the characteristics of the neighborhood of individuals.

The next section presents the existing literature on occupational mismatch as well as on network effects. In a third part, we present the database. Then, we present a theoretical model allowing me to highlight the key variables of the analysis. In a fifth section, we present the descriptive statistics. Finally, the following sections present the models used for the regressions and the results.

2 Prior research on the occupational mismatch and the network effects

To study the theoretical and empirical work related to research questions, we will first break down the literature according to two themes: the occupational mismatch and the network effect. The purpose of this decomposition is to present the main results for each of these themes while presenting the results obtained for the immigrants. We conclude with a section regrouping the works putting together both themes. The objective is to identify recent results while presenting the contributions of this article with respect to the existing literature.

2.1 Occupational mismatch

Before studying the analysis done on occupational mismatch in the literature, we first present the definition of this term as well as the different categories that allow us to determine whether or not there is a mismatch. The term mismatch refers to a notion of imbalance. Kain (1968) introduced this term in the literature from working with residential segregation. The mismatch referred to by Kain (1968) is linked on the one hand to the place of residence of the individuals and on the other hand to the location of the jobs. His report is based on the organization of American cities. These are characterized by a central city with a low level of employment and a high concentration of Black Americans while the peripheries have a higher level of employment and a low concentration of Black Americans. These spatial disparities in terms of population diversity and in terms of employment lead to inequalities in the labor market. The further away individuals are from jobs, the more difficult access to employment is. Kain (1968) also highlights the fact that dwellings close to employment areas have a higher cost than dwellings remote from these areas. In general, the analysis on this topic is based on the fact that a characteristic linked to an individual (his place of residence as presented by Kain (1968)) can discriminate against the individual in the labor market. In the literature, this notion of *mismatch* has subsequently referred to education, to the skills of

individuals. Freeman (1976) was the first to present overeducation cases focusing on the US labor market. According to Freeman (1976), overeducation cases can be explained by variations in both the supply side and the demand side. In order to find a job, high-skilled workers had to lower their job search criteria by accepting jobs for which they were overqualified. Compared to the theory of human capital presented by Becker (1964), the contributions of Freeman (1976) highlight that individuals can make errors of anticipation in their investment in education. This results in the non-use of skills available on the labor market because of this bad anticipation of the evolution of the labor market, both on the demand side and the supply of available work. These errors of anticipation may be due to structural changes or difficulties in the transfer of human capital acquired in the country of origin for immigrants.

In fact, there is an imbalance when the level of education, the highest diploma obtained by an individual, does not correspond to the level of education required for a job. More recently, Nordin, Persson, and Rooth (2010) dealt with occupational mismatch in relation to education in the case of Sweden. They define the term mismatch by “*A mismatch may well be caused by a sorting by ability, or a self-reported mismatch might be endogenous and related to the wage, i.e. a self-reported mismatch may be a form of rationalization of a general feeling of disappointment with the wage and/or the workplace*”. This definition puts forward the subjective side of the individual’s feeling of mismatch. This point relating to the measurement of the mismatch is the subject of the last part of this section. The goal is for researchers to be able to measure the mismatch according to a standard, a common basis.

We distinguish three cases:

- Overeducation : this means that the individual has a higher educational level than the required educational level of the job.
- Required education : The educational level of the individual corresponds to the educational level required of the job held.
- Undereducation : the educational level of the individual is lower than the educational level required for the job held.

Overeducation and *undereducation* correspond to an occupational mismatch. The interpretation of the occupational mismatch between undereducation and overeducation cases differs. Indeed, undereducation cases do not correspond to a case where the individual is *disadvantaged* in the

labor market. A company may recruit an individual with an educational level below the required level of education if the latter has skills - other than education - to offer to the firm (past work experience, for example). Sicherman (1991) has shown that individuals in the undereducation situation tend to have higher work experience than individuals in the overeducation situation. This can be explained by the acquisition of new skills during professional experiences. Undereducation rates for immigrants also tend to be higher than for natives (Chiswick and Miller (2010), Aleksynska and Tritah (2013)).

Concerning *overeducation*, the individual holds a job for which he has a higher level of education than the level required for the job. The fact of holding a job for which an individual is overqualified has been explained by various factors in the literature. One of these factors may be related to the origin of the individuals. Diplomas obtained in the country of origin of individuals may not be recognized in the host country (Støren and Wiers-Jenssen (2010), Aleksynska and Tritah (2013)). The probability of occupational mismatch for immigrants who have graduated in their country of origin is higher than immigrants who have graduated in the host country.

This explanation can be supplemented with fluency in the language of the host country. Indeed, an immigrant who does not master the host country's language will find it more difficult to enter the labor market. The language level of the host country for immigrants has a positive effect on the probability of being employed (Chiswick and Miller (1990), Dustmann and Fabbri (2003), Zorlu and Hartog (2018)). In order to limit the negative effects on the relationship between language level and the probability of being in employment, immigrants can mobilize their network in order to increase the probability of being employed (Lewis (2011)). The residential environment makes it easier for immigrants to fit into these markets because of a concentration of people sharing the same language. These results make it possible to highlight the importance of the social environment in the integration of immigrants, whatever their language level, into the labor market.

The consequences of the occupational mismatch on society are not limited to the individual. If there is an occupational mismatch, from an economic point of view, it means that all available resources in the labor force are not used. With regard to education, the fact that there is an occupational mismatch means that the return on investment is not maximized, both for the individual and for society.

Finally, the heart of the empirical analysis on the topic of occupational mismatch lies in its measurement. Two measurement approaches emerge in the literature. The first approach is called *subjective*. This approach is based on the statements of individuals. From a questionnaire, the

occupational mismatch is determined thanks to the perception of the individual on the matching between his educational level and the level of education required for his job. For this, two methods are used: direct self-assessment or indirect self-assessment. Concerning the first method (direct self-assessment), the question about the individual's feeling about the mismatch is asked directly (Verhaest and Omey (2006), Erdogan and Bauer (2009), Verhaest and Omey (2009)). In the second method (indirect self-assessment), the questions relate to the level of education needed to fill the job (Duncan and Hoffman (1981), Verhaest and Omey (2006)). One of the limitations of this approach is that the occupational mismatch is determined from the statements of individuals. However, for the same job, the same educational level, two individuals may have a different feeling of mismatch. The second approach is called *objective*. There are two methods associated with this approach. The first method is to define for each job a level of education required. For this, there is a Standard Occupational Classification System to determine for each job an educational level. As a result, there is an occupational mismatch if an individual has a higher educational level than the educational level defined by the classification system for employment (and conversely) (Baert, Cockx, and Verhaest (2013)). This approach is open to discussion because for each job title there is an associated level of education. However, the missions and tasks may be different despite an identical job title.

The second objective method is to calculate the average level of education obtained for each job. From there, the researcher can associate with each job a level of education. Thereby, if the individual has an educational level higher (or lower) than the one previously defined, there is an occupational mismatch (Verdugo and Verdugo (1989), Bauer (2002)).

In line with the work of Freeman (1976), a part of the literature on occupational mismatch has focused on the polarization of the labor market. The underlying idea is that the structural change in the demand for labor induces workers to obtain jobs for which they are overqualified. The polarization of the labor market results in a simultaneous increase of highly skilled and low-skilled workers. At the same time, the share of workers with an intermediate level of qualification decreases. This change in the distribution of workers has the effect of redistributing workers according to changes in the demand for work, *i.e.* for low or high-skilled jobs¹ These results have been confirmed in the literature (Autor (2010) and Sarkar (2017)): the more the labor market of a country is polarized, the more cases of overeducation. The rapid evolution of labor demand does not allow

¹See David, Katz, and Kearney (2006) and Goos, Manning, and Salomons (2009) for more details on the polarization of the labor market for the US and Europe.

workers to adapt their skills to these structural changes.

It is necessary to distinguish between the jobs according to the level of skills required as well as the notion of “routine job”.

The frequency of overeducation according to the characteristics of the jobs (either skill-requirement or routine work) have implications for the polarization of the labor market. Indeed, the share of the middle class in society decreases due to a simultaneous increase in jobs located at the bottom and the top of the distribution. This is explained by the increase in overeducation cases for low and middle-skilled jobs.

The occupational mismatch has been presented here according to a logic where the determinants of these situations are related either to the macroeconomic characteristics (case of the polarization of the labor market for example), or to anticipation errors in terms of the human capital of individuals.

2.2 Network effects

The network is a central element in understanding the *“transmission of information about job opportunities. [...] They are the basis for the provision of mutual insurance in developing countries, [...] in determining how diseases spread, which products we buy, which languages we speak, how we vote, as well as wheter we become criminals, how much education we obtain, and our likelihood of succeeding professionally”* (Jackson (2010)). This quote makes it possible to highlight all the subjects dealt with in the literature and relating to networks. The network is not static. Due to its constant evolution in its composition, the evolution of the environment, its characteristics are not fixed in time. The evolution of the social network may allow individuals belonging to the network to be more or less integrated into society over time.

The first reference relating to social relations in the scientific literature dates from 1954. It already highlighted the dynamic aspect of the social network. It is stated that *“The third social field [...] is made up of the ties of friendship and acquaintance which everyone growing up in [the] sociey partly inherits and largely builds up for himself.”* (Barnes (1954)). This reference emphasizes on the one hand the dynamic nature of the social relations and on the other hand the influence of these relations on the society in a general. These characteristics of the networks are still analyzed today in research.

The analysis of the networks has been the subject of a very varied literature concerning its applications. Social networks can affect individuals both in terms of their educational performance and in terms of their labor market performance (Munshi (2003), Patel and Vella (2013)), (Calvó-Armengol

and Zenou (2004), Liu et al. (2012)).

Granovetter (1974) laid the groundwork for the role of relations in getting a job. In fact, about a third of individuals find their jobs thanks to a relations (Granovetter (1974), Montgomery (1991), Patel and Vella (2013)). The objective of this literature on social networks is to understand what impact they can have on the performance of individuals in the labor market. The focus of this literature on immigrants is based on the role of networks for this population. Social networks intervene as soon as in the migration decision. The introduction of the family network into the theory of migration choice dates back to the 1980s (Stark and Bloom (1985)). Although this literature is not directly related to the labor market performance of immigrants, it helps to understand the importance of the network for immigrants. The family network offers insurance against climate risks, for example for families living in developing countries. The network makes it possible to diversify this risk and to ensure a guaranteed income for the family remaining in the country of origin.

The information transmitted between the immigrants and the families who stayed in the country of origin allows those who remained in the country of origin to learn about job opportunities in the host countries. Therefore, if there are interesting opportunities, these are transmitted between individuals and allow a village, a city to benefit from a specific information channel. As a result, new immigrants have information enabling them to fit more easily into the host country (Bauer, Epstein, and Gang (2002)).

The network allows individuals to increase the probability of finding a job. The network also enables individuals to perform better employment rate in the job market (Calvo-Armengol and Jackson (2004)). Beyond the network, the neighborhood effect can also impact the probability of individuals finding a job or adopting a similar behavior (Anne and Chareyron (2017), Gibbons, Silva, and Weinhardt (2013)).

The fact that the information is better transmitted between individuals with a common origin is explained - among other things - by a social and cultural proximity, the term used to characterize this behavior is homophily. Homophily can be defined as the fact that individuals with cultural, social, genetic or behavioral proximity have a higher probability of interacting. The socio-demographic distance between individuals is low when there is homophily (McPherson, Smith-Lovin, and Cook (2001)). This particularity of exchanging information between similar individuals is a central point both in the creation of social networks as well as in the explanation of the behaviors and incentives of the individuals composing the networks.

Social networks influence immigrants' occupational choice, both newcomers and those already settled (Munshi (2003), Patel and Vella (2013), Schuetze and Wood (2013)). The literature has shown that the network effect can influence immigrants' choice of occupation both negatively and positively. When the network influences immigrants' choice of occupation without taking into account the skills of these individuals, this can lead to occupational mismatches (Bentolila, Michelacci, and Suarez (2010))

All information transmitted either by the social or residential environment is likely to modify the behavior or actions of individuals on the labor market. Therefore, if there is a network effect, if there is a neighborhood effect, the performance of immigrants can be modified by the relations they have within the host society.

Depending on the origin of the individuals, social networks may be more or less effective. This effectiveness is a function of both the number of peers settled in the host country, the economic integration of these immigrants and the level of education. In the case of an inefficient network, immigrants may experience difficulties in entering the labor market and, more generally, integration difficulties in the host country society.

2.3 Occupational mismatch and network effects

Through the two previous sections, we have shown that the literature on occupational mismatch and network effects has made it possible to understand what factors could positively or negatively influence the individuals performance - and more particularly immigrants - into the labor market. On the one hand, immigrants have individual characteristics that may discriminate them in the labor market, on the other hand, the network, even if it can lead to increased segregation, may in some cases allow individuals to find a job more easily into the job market. The ambiguity of certain effects has led researchers in recent years to look at both the network effect and more specifically the role of the network effect on the occupational mismatch. The measurements of both the network effect as well as the occupational mismatch are complex (often inobservable) and difficult to access in the databases.

The aim of this article is to study to what extent the network effect can positively (or negatively) influence the occupational mismatch of immigrants according to their origin. In the previous two sections, we have shown that immigrants are more likely than natives to be in an occupational mismatch situation and at the same time, network effects help to explain the occupational choices of individuals and more particularly the occupational choices of immigrants. Our article analyzes

the contribution of the network effect to explain occupational mismatch situations for immigrants. The analysis of causality of the network effect on the occupational mismatch has been studied in particular for the cases of Senegalese immigrants (Chort (2017)) and for the case of Australia (Kalfa and Piracha (2018)). The results between these two papers are opposite. Indeed, in the work of Chort (2017), the results show that the network allows immigrants to find a job more in line with their educational level. The results of Kalfa and Piracha (2018) show that ethnic concentration as well as peer relations tend to increase the probabilities of occupational mismatch. Several points distinguish though these two papers:

Results observed in the literature are not homogeneous. In the article of Chort (2017)², the results show that the network allows immigrants to find a job more in line with their educational level. However, the results of Kalfa and Piracha (2018) show that ethnic concentration as well as peer relations tend to increase the probabilities of occupational mismatch. These different results could be explained by some facts. Even if we notice the difference between these results, it may be explained by different factors as the population, the countries and the period studied.

We focus in this paper on the economic integration of immigrants in France.

The period that we study is from 2005 to 2012. To analyze the network effect, we distinguish two measures to characterize it. The first index is the concentration of individuals according to their origin in French regions. The second index focuses on the residential environment³. Moreover, we suppose that the network effect is not homogeneous according to the origins. To test this hypothesis, we distinguish immigrants across origins⁴ or across regions of origin (for instance, Europeans VS non-Europeans). In the literature,

- The countries of destination are different, Australia for Kalfa and Piracha (2018) and France, Italy, Mauritania and Côte d'Ivoire for Chort (2017)
- The populations studied are also different: Senegalese for Chort (2017) and natives and immigrants for Kalfa and Piracha (2018)
- Differences in the determination of the network effect: questions asked directly or based on the concentration of immigrants according to their origin within the regions

²The article analyze the economic integration of Senegalese in France, Italy, Mauritania and Côte d'Ivoire. The data were collected in 2009 and 2010.

³Thanks to the LFS, we can identify the residential characteristics for each individual.

⁴The database allows us to distinguish 9 different origins.

- Kalfa and Piracha (2018) focus in particular on the dynamics of overeducation and the role of social relations in this dynamic while Chort (2017) is more interested in the influence of the network in the probability of being in an occupational mismatch situation
- The periods studied are very different between the two papers. The role of the business cycle as well as the means of communication can influence the results. The period studied is from 1995 to 2001 for Kalfa and Piracha (2018) and the data were collected between 2009 and 2010 for Chort (2017)

The differences associated with the studied period can more characterize the effects relative to the network effect. Indeed, in period of economic growth or economic recession, the difficulties of access to jobs can be facilitated or not by this channel. Concerning the studied populations, the transferability of diplomas and experiences can be different according to country of origin of the individuals. Thanks to the LFS, origins can be analyzed with a diversity allowing to take into account the specificities of the origins of the individuals. Finally, we are going to characterize occupational mismatches according to the case of overeducation but also of undereducation.

Our paper is in line with the previous studies:

- We use the French Labor Force Survey to study the role of the network effect on the occupational mismatch for 9 different origins. This analysis allowsto determine if the role of the network is homogeneous according to the origin of the individuals or if there is a heterogeneity between the individuals according to their origins (Europeans VS non-Europeans for instance). This assumption is based on the fact that the use of networks may be not homogeneous according to the origins.
- We identify the network effect according to several criteria. The first index used is the concentration of individuals according to their origin in French regions. The second index used refers to the environment close to the individual.

3 Data

To study the network effect on immigrant occupational mismatch, we use the French Labor Force Survey (LFS). This study is conducted by the National Institute of Statistics and Economic Studies

(INSEE). Prior to 2003, the LFS was conducted annually. From 2003, the LFS was conducted every 3 months. Each questioned individual is followed for 18 months (individuals are interviewed 6 times). Each quarter, one sixth of the base is renewed. The questioned individuals are characterized by two elements:

- Individuals are over 15 years old.
- In each surveyed neighborhood, a group of 20 households is interviewed. There is geographical proximity between these individuals.

It should be noted that the condition required for an individual to be interrogated 6 times is that he remains in the urban area in which he was present during the first interrogation.

The database provides information on different topics. In this survey, the information allows us to identify the gender, age, educational level of the individual as well as that of his/her parents (from 2005), the country of origin of the individual and of his parents. From this information, we can therefore identify the origin of individuals according to 9 distinct groups: French, North Africans, Sub-Saharan Africans, North-Europeans, South-Europeans , East-Europeans, Turks, South-East Asians and Others grouping the rest of the world. In addition, we can also identify the arrival date of immigrants.

Finally, individuals are interviewed in the Labor Force Survey based on their place of residence. Within selected urban areas, 20 households are interviewed 6 times over a period of 18 months. It should be noted that if one of the household moves out of the area during this 18-month period, the household is definitely out of the questioning. In the following sections, we will use this household selection feature to analyze the effect of the environment, neighborhood on the occupational mismatch.

It should be noted that we also use the weights associated with each individual in the statistics and regressions. This adjustment is based on census data. The purpose of these weights is to have a representative sample of individuals in the population.

4 Theoretical framework

5 Descriptive statistics

One of the fundamental points of this analysis is based on two elements: the educational attainment and the occupational choice of individuals and more particularly for immigrants. Therefore, during

this section, we first present some key elements, about immigration in France and the educational level of immigrants in France according to their origin. Then, in a second step, we present the elements relating to occupational mismatch.

As presented in the literature review, the mismatch between the educational level of immigrants and the educational level required in a job can be explained by various factors (non-recognition of diploma and experience or low language level of the host country). In the case of France, Figure 2 in the Appendix present the occupational mismatch cases for immigrants and natives over the period 2005-2012. The proportion of individuals in overeducation or undereducation cases is higher for immigrants than natives for both. The aim of this section is therefore to highlight some descriptive facts that make it possible to characterize the education of immigrants, the role of social interactions in the context of occupational mismatch and to highlight the characteristics associated with each origin.

5.1 Composition of the immigration in France

Before studying these different points, we first present some facts on immigration in France. Since the end of the Second World War, France has experienced various waves of migration from countries in southern Europe (Spain, Italy), the Maghreb and more recently sub-Saharan Africa. Although during certain periods the population flows of these regions may have been higher or lower, the proportion of each of these origins has remained relatively constant during the last decades. The share of immigrants in the French population is about 10%. Among this population, the most represented origins are South Europeans and North Africans (table 1).

Table 1 presents the composition of immigration in France according to 8 origins. The results are calculated for the period 2005-2012. The distinction made between newcomers and established immigrants is made from the date of arrival in France. If an immigrant has arrived in the last 5 years, he/she is considered as a newcomer. If an immigrant arrived before that date, we consider that he/she is an established immigrant. Established immigrants are, for all origins, the most represented. However, for some origins, newcomers represent more than a quarter of the immigrants present in France (East-Europeans). The distribution of immigrants according to their date of arrival in France may have an impact on the network effect. If newcomers make up a significant share of immigrants in the host country, immigrants may not rely on a *effective* network because

Table 1: Share of recent and established immigrants - by origins - 2005-2012

| | Share of each origin among the immigrants (%) | Recent immigrants (%) | Established immigrants (%) |
|----------------------|--|--------------------------|-------------------------------|
| North-Europeans | 8.68 | 19.10 | 80.90 |
| South-Europeans | 17.16 | 8.70 | 91.30 |
| East-Europeans | 5.42 | 23.34 | 76.66 |
| North-Africans | 35.16 | 8.92 | 91.08 |
| Sub-Saharan Africans | 15.51 | 12.21 | 87.79 |
| Turks | 3.26 | 11.72 | 88.28 |
| South-East Asians | 3.38 | 2.62 | 97.38 |
| Others | 11.43 | 15.87 | 84.13 |

of the recent arrival of individuals. This fact could be a determinant to explain the heterogeneity of the network effects across origins.

5.2 The educational attainment of immigrants

As seen previously, the occupational mismatch is determined from the level of education of the individual and the educational level required to occupy the job. The objective of this section is to present an inventory of the level of education for immigrants according to their origin.

Table 2 represents the educational level of immigrants according to their origin. The level of education is broken down into 8 different positions (BAC+5, BAC+4, BAC+3, BAC+2, BAC, CAP and BEP, Brevet des collèges, No diploma).

This table shows a certain heterogeneity according to the origin concerning the level of education. Indeed, if we focus initially on individuals with no degree, two groups stand out in particular: the South Europeans and the Turks. For these two groups of immigrants, the percentage of individuals without a low educational attainment exceeds 50%.

Conversely, the proportion of individuals with a level of education higher than or equal to BAC+2 is high for North-Europeans, East-Europeans and Others (more than 30%). South-Europeans and Turks have a low proportion (less than 10%) of individuals with this level of education. The educational distribution is characterized by a stronger polarization at the extremities for immigrants. They are overrepresented among individuals with a low level of education and they are overrepresented among individuals with a high level of education.

Table 3 refers to the origin of diplomas obtained by immigrants. For this sample, we selected only individuals who graduated (individuals with no diploma and Brevet des collèges are excluded from this sample). Thanks to the information available in the database, we kept the date of the last diploma obtained by the individual. we compare this year with the date of arrival in France. From this information, we can determine whether the last diploma obtained by the individual in the country of origin or in the host country⁵ In addition to information related to the origin of the last diploma obtained by individuals, we also calculate the average age of arrival in France of individuals according to the origin of the degree.

⁵The date of the last diploma obtained does not mean that the individual has completed all his/her schooling in France.

Table 2: Educational attainment - Origins

| | No diploma | Brevet des collèges | Educational attainment | | | | | |
|----------------------|------------|---------------------|------------------------|-------|-------|-------|-------|-------|
| | | | CAP/BEP | BAC | BAC+2 | BAC+3 | BAC+4 | BAC+5 |
| French | 0,11 | 23,73 | 27,76 | 19,51 | 12,14 | 8,24 | 3,33 | 5,18 |
| North-Europeans | 0,19 | 16,14 | 15,25 | 21,17 | 10,56 | 17,25 | 7,55 | 11,9 |
| South-Europeans | 0,97 | 59,32 | 21,22 | 8,31 | 3,52 | 2,91 | 1,5 | 2,26 |
| East-Europeans | 0,3 | 20,92 | 12,7 | 26,32 | 7,68 | 12,97 | 6,48 | 12,64 |
| North-Africans | 0,55 | 43,6 | 16,57 | 15,17 | 6,87 | 6,71 | 3,32 | 7,21 |
| Sub-Saharan Africans | 0,13 | 38,35 | 12,23 | 21,28 | 9,24 | 7,46 | 4,11 | 7,19 |
| Turks | 0,32 | 67,1 | 13,12 | 11,57 | 3,38 | 2,25 | 0,46 | 1,81 |
| South-East Asians | 0,27 | 40,92 | 11,84 | 20,24 | 8,21 | 7,81 | 4,37 | 6,35 |
| Others | 0,15 | 29,64 | 8,37 | 20,95 | 7,67 | 14,61 | 5,74 | 12,87 |
| Total | 0,15 | 25,83 | 26,23 | 19,17 | 11,52 | 8,24 | 3,39 | 5,46 |

Table 3: Origin of the diploma and age of immigrants in France - 2005-2012

| | Graduated in their country of origin | | | Graduated in France | | |
|----------------------|--------------------------------------|-------------|-------------------------------|---------------------|-------------|-------------------------------|
| | In % | Average age | Average arrival age in France | In % | Average age | Average arrival age in France |
| North-Europeans | 49.98 | 44 | 33 | 50.02 | 41 | 9 |
| South-Europeans | 17.91 | 41 | 29 | 82.09 | 43 | 7 |
| East-Europeans | 61.99 | 42 | 30 | 38.01 | 35 | 15 |
| North-Africans | 23.60 | 41 | 29 | 76.40 | 44 | 9 |
| Sub-Saharan Africans | 33.03 | 41 | 29 | 66.97 | 39 | 13 |
| Turks | 43.36 | 37 | 25 | 56.64 | 35 | 9 |
| South-East Asians | 13.95 | 47 | 27 | 86.05 | 42 | 10 |
| Others | 45.37 | 43 | 30 | 54.63 | 36 | 13 |

The average age of arrival in France is higher for immigrants who obtained the diploma in their country of origin than for immigrants who obtained their last diploma in the host country. The average age does not exceed 15 years for individuals who have graduated in France for all origins. Regarding the origin of the diploma, the origins can be classified according to 3 groups:

- East-Europeans : the majority of immigrants have graduated in their home country
- North-Europeans, Turks and Others : the proportion of immigrants who have obtained the diploma in the country of origin or in the host country is the same
- South-Europeans, North-Africans, Sub-Saharan Africans and South-East Asians : the majority of immigrants have graduated in the host country

Tables 2 and 3 made it possible to highlight certain facts concerning the educational distribution of immigrants according to the origin, the origin of the last diploma obtained by immigrants and the average age of immigrants at their arrival in France according to the origin of the diploma. Even if this information does not establish a causality between education and occupational mismatch, it still allows to bring some stylized facts on the structure of education and highlight the heterogeneity existing according to the origin of the individuals.

5.3 Occupational mismatch

5.3.1 Identification of the occupational mismatch

As we noted in the previous sections, the purpose of this analysis is to identify the network effect on the occupational mismatch. One of the key points of the analysis is to identify the cases where there is an occupational mismatch or not. For that, we use the realized matched method (Verdugo and Verdugo (1989), Chiswick and Miller (2010)). This method requires calculating the average level of education for each job. Then, we compare that level of education required for a job to the educational level of individuals. Individuals are categorized according to three categories: overeducated, correctly matched or undereducated.

To apply the realized matched method, we first calculate the average level of education for each job by considering only natives, E_n . We also calculate the standard deviation σ_n for each occupation. In a second stage, we compare the level of education observed for each individual according to his employment with the average level of education observed for the natives. The different possible situations are summarized below:

$$M_{ijo} = \begin{cases} 1 & \text{if } E_{ijo} > \bar{E}_n + \sigma_n \\ 0 & \text{if } \bar{E}_n - \sigma_n \leq E_{ijo} \leq \bar{E}_n + \sigma_n \\ -1 & \text{if } E_{ijo} < \bar{E}_n - \sigma_n. \end{cases}$$

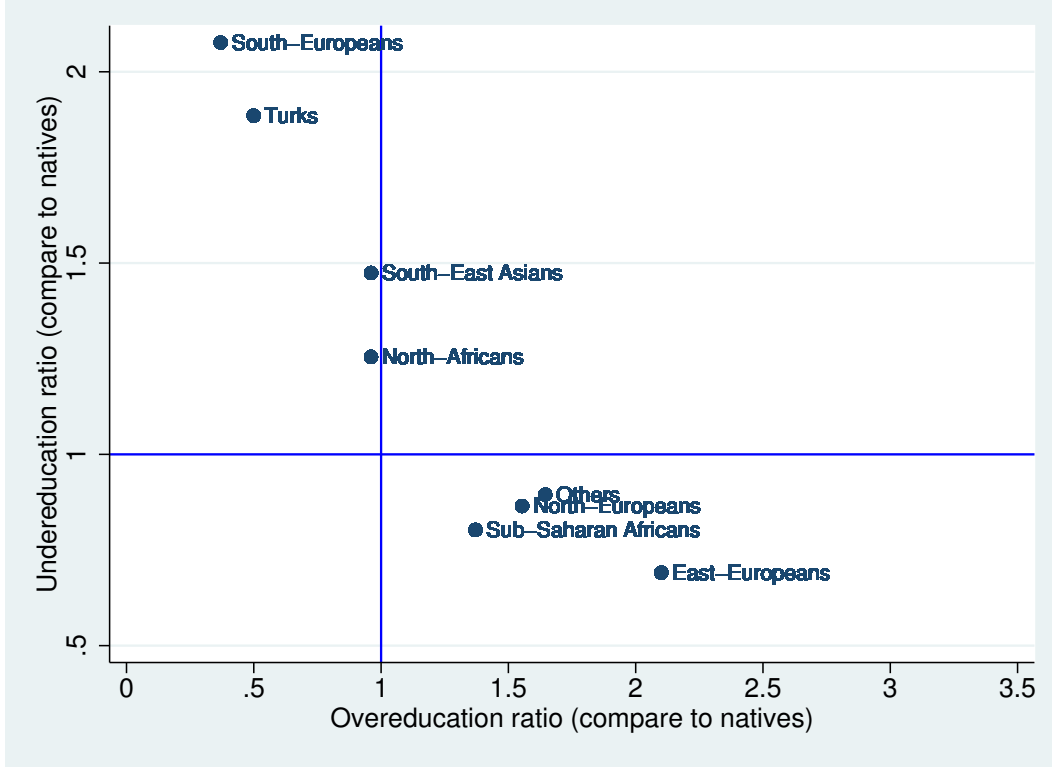
To determine the level of education, we rely on the data available in the LFS. The level of education is detailed thanks to the number of years of study. Based on this information and the age for which education is compulsory in France, we determined the duration of education from the highest degree obtained for each individual as part of their initial training.⁶

From these elements, we can identify individuals according to 3 situations:

- Overeducated if $M_{ijo} = 1$
- Correctly matched if $M_{ijo} = 0$
- Undereducated if $M_{ijo} = -1$

⁶Education is compulsory in France from the age of 6 years. For each diploma (BEP, CAP, BAC ...) and according to a so-called *normal* schooling (ie without repetition and without grade-skipping), the duration of education is calculated from the difference between the theoretical age of individuals when they graduated the highest, minus the minimum age of education, ie 6 years

Figure 1: Relative Mismatch across Origins



5.3.2 Descriptive statistics about the occupational mismatch

Figure 1 is inspired by the work of Aleksynska and Tritah (2013) and is constructed from the educational level of the individuals, the average level of education observed for the natives for each job and the origin of the immigrants. The x-axis represents the ratio between the proportion of immigrants who are overeducated compared to the proportion of natives who are overeducated. The y-axis represents the same ratio (immigrants on natives) for the case of undereducation. Each point represents a different origin. To achieve this result, we rely on this formula:

$$Ratio_{iM} = \frac{\frac{N_{iM}}{N_i}}{\frac{N_{nM}}{N_n}}, \text{ with } M = o, u \quad (1)$$

With N_{iM} corresponding to the number of individuals of origin i being either overeducated if $M = o$ or undereducated if $M = u$. N_i is the number of individuals of origin i . The ratings are identical for the natives, only the index changes with n . The formula 1 is therefore applied to the data in order to obtain the ratio $Ratio_{iM}$ for each case (overeducation or undereducation) and for each origin.

Figure 1 shows that there is heterogeneity in terms of overeducation and undereducation depending on the origin. First, if the rates of overeducation and undereducation had been similar between immigrants and natives, all the points would have had coordinates 1 in abscissa and ordinate. Secondly, if there was a homogeneity between immigrants concerning overeducation or undereducation, all the points would have been grouped together. Few points have similar coordinates for both overeducation and/or undereducation cases.

The fact that Turks and South-Europeans have a high undereducated ratio can be explained with table 2. In fact, in this table, South-Europeans and Turks unqualified represent more than 50% of the individuals.

In contrast, East Europeans have a very high overeducated ratio compared to other origins and natives. Despite a high average level of education for this origin, East Europeans have difficulty in asserting their skills in the labor market. In addition, with table 1, we highlighted the fact that this population has the highest proportion of newcomers. They may therefore find difficulties in entering in the labor market due to non-recognition of their diplomas obtained mostly abroad (table 3) and the social network is inefficient due to a relatively recent arrival date.

Regarding overeducation, six immigrant groups have a higher proportion of overeducated people than natives (South East Asians, North Africans, Sub-Saharan Africans, Others and East-Europeans). Groups with a lower proportion of overeducated individuals than natives are the South-Europeans and the Turks. This result could be anticipated thanks to table 2. Indeed, all groups of immigrants with a higher proportion of overeducated individuals than natives have a higher average level of education or similar to natives. However, for the South Europeans and Turks, the proportion of individuals having a level of education higher than or equal to a BAC + 2 is relatively low (approximately 10% for South-Europeans and 8% for Turks).

About undereducation cases, there are also six immigrant groups with a higher proportion of undereducated people than natives (Turks, South-Europeans, South-Asians, North Africans, Sub-Saharan Africans, Others). North-Europeans and East-Europeans have a lower proportion of undereducated people than natives. These results go in the same direction as the cases of overeducation. Indeed, immigrant groups with a lower level of education are overrepresented in undereducation cases. Three hypotheses can be emitted with this fact: either immigrants have specific skills (excluding education) allowing them to hold a job for which they are, in terms of education, undereducated; immigrants benefit from the support of their network to fit more easily into the labor market; or immigrants may have work experience that allows them to work in jobs for which they do not have

the average level of education of other individuals in the same job. The aim of the next section will be to analyze what is the role of the network as a determinant of overeducation and undereducation cases.

Table 4: Overeducation and undereducation ratios, network effects - 2005-2012

| | Overeducation (%) | Overeducation and network (%) | Undereducation (%) | Undereducation and network (%) |
|----------------------|-------------------|-------------------------------|--------------------|--------------------------------|
| North-Europeans | 28,98 | 11,56 | 8,62 | 16,49 |
| South-Europeans | 6,90 | 13,71 | 20,71 | 29,01 |
| East-Europeans | 39,17 | 12,81 | 6,89 | 24,79 |
| North-Africans | 17,91 | 12,34 | 12,51 | 20,22 |
| Sub-Saharan Africans | 25,55 | 13,41 | 8,00 | 17,89 |
| Turks | 9,32 | 27,45 | 18,80 | 41,54 |
| South-East Asians | 17,91 | 13,83 | 14,70 | 30,06 |
| Others | 30,66 | 15,18 | 8,92 | 27,45 |

Reading : Among the 28.98% of North-Europeans being over-educated, 11.56% of them found their jobs thanks to an acquaintance.

Table 4 details the cases of overeducation and undereducation for each origin. In addition to these data, it is also specified the percentage of individuals who have found a job through networks.

Columns 2 and 4 of table 4 specify the percentage of individuals who have found their jobs through social networks and who are either overeducated (column 2) or undereducated (column 4). The role of the network is heterogeneous depending on the origin. The proportion of South-Europeans and East-Europeans having found employment through the network and being overeducated is almost identical. However, the proportion of overeducated individuals is much higher for East-Europeans than South-Europeans. These facts highlight the importance of analyzing the influence of social interactions on occupational mismatch according to the origin of individuals.

6 Method

6.1 Definition of the network

A fundamental point of the analysis refers to the network definition. Before being able to measure the network, we must be able to define the network from the available data.

The first definition of network is based on the job search method. In the database, we can identify individuals who have found a job through a social contact (family, friends). Even if the origin of the job-contact is not available, this information makes it possible to show the direct link between the method of the individual to obtain his job and the role of his network.

The second definition of the network is based on the influence of peers. For each of the French administrative regions, we identify the proportion of immigrants from origin o . The concentration of peers in the region may allow individuals to have more information on the labor market. To determine this regional concentration of immigrants by origin, we have:

$$S_{ogt} = \frac{N_{ogt}}{N_{ot}} \quad (2)$$

where N represents the number of individuals, o corresponds to the origin of the individuals, g corresponds to the region considered and finally t represents the year.

Finally, the last index used to study the network effect is based on the influence of the neighborhood. The database makes it possible to identify individuals living in a restricted area. As a result, we identify the characteristics related to the composition of the neighborhood according to their level of education. The index is based on the average level of education of individuals living in this area.

If high educated immigrants reside in an area where the average level of education is low, social networks may induce the individual towards an occupation for which he is overeducated.

6.2 Empirical strategy

The objectives of this analysis are twofold:

- Highlight the influence of social networks on the probability of being in an occupational mismatch situation.
- Study the influence of the social networks on the occupational mismatch across origins.
- Analyze the effect of the network educational composition on the occupational mismatch.

To study these relations, we first use a simple econometric model: probit. In a second step, we use the bivariate probit model to characterize the relationship between the probability of finding a job thanks to a referee and the probability of being in an occupational mismatch situation. This model allows us to test the endogeneity of the referee's influence.

The dependent variables used in the estimations are:

- The probability to find a job thanks to a referee R_i
- The probability of being in an occupational mismatch situation M_i

The variable relative to the referee, denoted by R_i , takes the value 1 if the individual found his job thanks to a relation. This information is directly available in the database. The value of the variable is equal to 0 otherwise⁷.

Concerning the variable relating to the occupational mismatch, we distinguish different situations allowing us to characterize the different cases relating to the occupational mismatch⁸. In general, if the individual is in an occupational mismatch situation, the value of the variable is 1 and 0 otherwise. However, we distinguish the samples according to the case of overeducation or undereducation. Indeed, if the parameter associated to the variable *referee* has a significant effect in undereducation cases, the network effect plays a positive role in the integration of individuals into

⁷This data does not give access to the referee's information. However, this one makes it possible to highlight the importance of the network in the search for employment.

⁸We compare the average educational level for each job with the educational level of individuals. If the educational level of the individual is higher than the average educational level of the job plus the standard deviation, we consider that the individual is overeducated for the job, so he is in an occupational mismatch situation.

the labor market. However, in the case where the individual is overeducated for his job and the estimated parameter of the variable *referee* is significant, we can assume that the role of referees has a negative effect on the integration of individuals on the labor market. Indeed, if the individual finds a job for which he is overeducated through a referee, the use of the individual's skills is not maximized in the labor market.

First, we estimate the probability of finding a job thanks to a referee. We introduce the explanatory variables relating to the education of the immigrants. The exogenous variables used in our estimates include the origin of the last diploma obtained⁹. In fact, we have distinguished in the database, on the one hand, the individuals who had obtained their highest diploma in the country of origin and, on the other hand, the immigrants who obtained their last diploma in France. If the individual has completed his training or part of his training in the host country, it can be assumed that the immigrant has arrived at a relatively young age. This factor can be influential in the integration of immigrants into the host society (Dustmann and Van Soest (2002), Yao and van Ours (2015)). We also take into account the year of arrival in France of immigrants by distinguishing established immigrants from new immigrants. For this, we have distinguished immigrants who arrived in the last 5 years and other immigrants.

In a second step, our variable of interest is the probability for an immigrant to be in an occupational mismatch situation. As previously stated, we distinguish mismatch situations according to overeducation and undereducation. The purpose of this distinction is to know if the referee has an effect on the probability of being in an occupational mismatch situation and if this effect is significant for the two situations, it makes it possible to know if the relations have an identical effect according to overeducation and undereducation. Indeed, if the estimated parameter associated with the variable *referee* is significant and positive in the case of undereducation, this allows to characterize a positive effect of referees on the integration of immigrants into the labor market. It should be noted that we will make estimates for the immigrant population with and without distinction of origin. This allows us to highlight the heterogeneity of the network effect on the integration of immigrants into the labor market.

The first equation, with referee as dependent variable, is:

$$R_i = \alpha Z_i + u_i \tag{3}$$

Z_i is the vector containing the exogenous factors that can explain the probability of using a referee

⁹The level of education has a positive effect on the integration of immigrants into society (Aoki and Santiago (2015))

to find a job. This vector is - among others - composed of the indexes presented above relating to the residential and social characteristics (peers). As explanatory variables, we include the concentration of immigrants in each regions and for each origins. We include the concentration of immigrants or immigrants according to their origin in each of the administrative regions as well as the index of neighborhood characteristics¹⁰.

The second estimated equation relates to the probability of being in an occupational mismatch situation:

$$M_i = \beta X_i + \delta R_i + \epsilon_i \quad (4)$$

With M_i corresponding to the probability of being in an occupational mismatch situation, X_i is the vector containing the exogenous factors explaining the probability of being in an occupational mismatch situation for immigrants. R_i corresponds to the dependent variable of the equation 3. This variable takes the value 1 if the individual find a job thanks to a referee, $R_i = 0$ otherwise. Finally, we include in this estimate the variable R_i characterizing the use of referees in the job search.

The objective of our analysis is to have an econometric specification allowing us to model the role of the social environment (neighborhood and peers) on the probability of being in an occupational mismatch situation for immigrants.

The second model that we estimate is composed of two binary dependent variables with a system with two equations. The application of bivariate recursive probit assumes that the error terms of the two equations are correlated. The difference between this model and a simple bivariate model is that the dependent variable of the second equation appears in the first equation.

In order to be able to estimate the parameters of the equations, we need to set two specific conditions for both probit models in general and a condition for recursivity of the bivariate probit model:

- Residuals of both equations are normalized to 1.
- At least one explanatory variable of the model having as a dependent variable *Referee* must be absent in the occupational mismatch equation¹¹.

The application of the recursive bivariate probit is characterized by the simultaneous estimation of the equations 3 and 4. We include the dependent variable *Referee* (equation 3) in the equation

¹⁰We include individual and geographical characteristics as gender, age, regions and years.

¹¹Maddala (1983) and Wooldridge (2010)

4. In this framework, with the maximum likelihood method, the two equations are estimated simultaneously.

Our estimation strategy is decomposed in two steps. First, we estimate 3 and 4 independently. The aim of these first estimations is to highlight the determinants of the probability to find a job through network and the probability to be in an occupational mismatch situation for immigrants. Secondly, we consider that the variable *referee* is endogenous. Thanks to a recursive bivariate probit model, we estimate these equations simultaneously (equations 3 and 4).

7 Results

This section is divided into two parts. In the first part, we present the probit models whose dependent variable is the occupational mismatch or the network effect. For each estimate, we also distinguish overeducation cases and undereducation cases. In a second part, we present the biprobit models. The decomposition of the sample is identical to that applied to the probit models.

7.1 Probit estimations

These estimates allow us to implement a simple model to characterize the determinants of either the probability of using the network to find a job or the likelihood of being in an occupational mismatch situation.

7.1.1 Overeducation situations

Network effects

Tables 5 and 6 present the results of the estimates of the probability of finding a job through the network in the case of overeducation and the marginal effects. The results presented are made for the total immigrant population and with a distinction according to origin. We can notice that the influence of the peers is not significant on this probability, neither for the immigrant population without distinction of origin, nor for the immigrants according to their origin.

However, two factors have a significant role in the probability of having recourse to the network to find a job: the year of arrival of the immigrant in the host country and the origin of the diploma. Regarding the origin of the diploma, obtaining a diploma in France decreases the probability of using a referee to find a job. These results are confirmed, however, only for immigrants without distinction of origin and the East-Europeans. The results for East-Europeans, compared to descrip-

tive statistics presented previously, can be explained by a relatively young immigrant population compared to other origins. The use of formal prospecting methods can therefore be preferred to using referees to find a job.

The other significant variable concerns the distinction between new immigrants and immigrants established for more than 5 years in France. Being a newcomer increases the likelihood of using a referee to integrate into the job market. This variable is significant for immigrants without distinction of origin and for North-Europeans. This result confirms the fact that immigrants are accompanied by the network to integrate into the labor market (Piracha, Tani, and Vaira-Lucero (2016)).

Occupational mismatch

Tables 6 and 7 present the results concerning the marginal effects and the probability of being overeducated for immigrants.

Regarding individual characteristics, we notice that age and *be a woman* increase the probability of being overeducated. Two origins have significant estimated parameters associated with these two variables: East Europeans and North Africans.

Regarding the variable *Referee*, we note that the use of referees has a positive role in the integration of immigrants into the labor market. Indeed, using a referee reduces the probability of being in an occupational mismatch situation. None of the origins presents opposite results. Even if the results do not highlight the origin of the referee, this characterizes the integration of immigrants into society in order to integrate into the labor market. It is necessary to specify that this integration is not penalized by a non-transferability of the diploma when the immigrants resort to a referee.

7.1.2 Undereducation situations

Network effects

Tables 8 and 9 present results on the probability of using referees and the marginal effects. The sample used concerns undereducated immigrants and immigrants not in an occupational mismatch situation.

In contrast to the sample of immigrants in over-education, the variables relating to the year of arrival in the host country and the origin of the diploma are not significant in these estimates. These characteristics *individual migratory* do not influence the probability of using a referee to find a job.

However, the average level of education in the neighborhood ¹² decreases the probability of using a referee to find a job.

Occupational mismatch

The marginal effects, related to the probability of being undereducated, are grouped together in table 9. Estimates of the probability of being in an occupational mismatch situation and more particularly to be undereducated are grouped in the table 10.

The proportion of immigrants in undereducation situation is lower than for immigrants in overeducation situation. The situation of undereducation means - according to our methodology - that immigrants have a level of education lower than the average level of education observed in the job minus the standard deviation. However, the reasons for undereducation situations can be varied: the social environment can accompany immigrants to find a job despite the lack of qualifications required for employment; individuals can value professional skills and experiences that we do not account for here because of our methodology. Being undereducated in a job through a referee can be interpreted as being a positive effect for individuals.

Regarding our observations, we observe that the *Referee* variable is not significant. This does not influence the probability of being undereducated. However, the year of arrival in France influences the probability of being undereducated. Indeed, recently arrived immigrants increase the probability of being in undereducation situation for the South-Europeans and the Sub-Saharan Africans.

7.2 Biprobit estimations

Table 11 presents the marginal effects for the recursive bivariate probit model for overeducation situations. The main lesson in this table is the marginal effect of the estimated parameter of the *Referee* on the probability of being overeducated. If we consider all immigrants without distinction of origin, using a referee to find a job reduces the probability of being in an overeducation situation. This remark can also apply to the case of North-Europeans. However, this social effect takes opposite values according to the origin of the immigrants. For North Africans and the Sub-Saharan Africans, the use of referee increases the likelihood of being overeducated. If we take into account the estimation of the variable *Referee*, we observe that the probability of finding a job thanks to a referee is driven by the neighborhood average education level.

¹²The neighborhood is characterized by neighborhoods. In the LFS, household selection is defined from neighborhoods. 20 households belonging to the same neighborhood are followed for 18 months.

We also observe that for all immigrants, whatever their origin, the only effect relative to the social environment concerns the variable *Neighborhood average education level*. The variable for the share of peers in the administrative region is not significant for any origin. Regarding the neighborhood effect, we observe that this effect has an opposite effect depending on the origin of the individuals. For North-Europeans and South-Europeans, an increase in the level of education in the neighborhood reduces the likelihood of using a referee to find a job. For North Africans and the Sub-Saharan Africans, an increase in the level of neighborhood education increases the likelihood of having a referee.

8 Conclusion

The issues relating to the integration of immigrants into the labor market are necessary to understand the demographic shocks observed over the last decades: global increase in the flow of immigrants, aging of the population in the OECD countries... The understanding of the labor market failures of the occupational mismatch is essential in all societies in order to maximize the use of skills available in a society. In addition to these economic issues, the integration of immigrants into the labor market also implies the integration of immigrants into the host society. The aim of this work was to characterize the role of the social environment on occupational mismatch by characterizing neighborhood and peer effects. From the LFS, our estimations were aimed initially at characterizing the occupational mismatch (overeducation and undereducation) and the network effect independently. The results highlight the relatively limited role of the peer effect on the probability of finding a job through a referee. The neighborhood effect, characterized by the level of education of individuals, is more important in the probability of finding a job through a referee.

9 Appendix

Figure 2: Occupational mismatch - Natives VS Immigrants

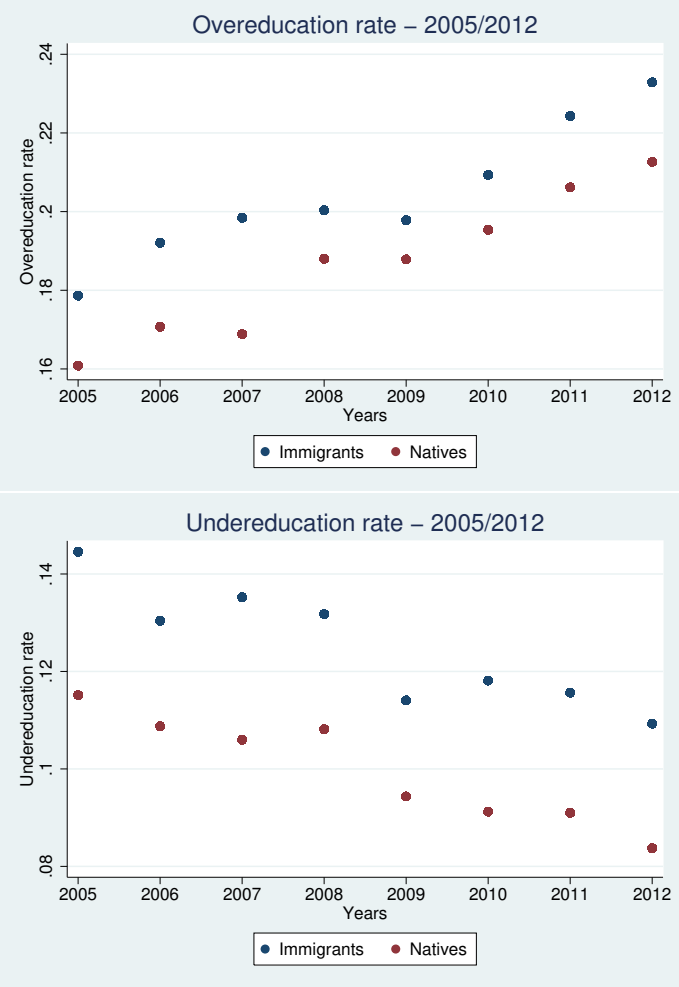


Table 5: Dependent variable: Referee (overeducation)

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|--------------------------------------|----------------------|---------------------|---------------------|----------------------|----------------------|----------------------|--------------------|--------------------|----------------------|
| | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | Turks | South-East Asians | Others |
| Neighborhood average education level | -0.016 (0.011) | -0.018 (0.034) | -0.064** (0.032) | -0.042 (0.040) | -0.005 (0.022) | 0.035 (0.029) | -0.173* (0.102) | -0.003 (0.063) | -0.043 (0.027) |
| Share of peers (region) | 0.188 (0.467) | 2.722 (2.549) | -1.187 (0.982) | 0.583 (0.533) | 1.190 (0.858) | -1.381 (1.083) | -0.513 (1.611) | -1.113* (0.596) | 0.240 (0.270) |
| Arrived in France <5 years | 0.071** (0.035) | 0.252** (0.112) | 0.033 (0.086) | 0.032 (0.198) | 0.033 (0.057) | 0.082 (0.097) | 0.168 (0.254) | -0.079 (0.182) | -0.080 (0.113) |
| French diploma | -0.077** (0.032) | -0.142 (0.100) | -0.187* (0.107) | -0.352*** (0.122) | -0.004 (0.062) | -0.061 (0.078) | -0.286 (0.202) | -0.005 (0.222) | -0.115 (0.086) |
| Constant | -1.081*** (0.295) | -2.017** (0.836) | 0.352 (0.736) | 0.961 (0.780) | -1.401*** (0.520) | -0.811 (0.837) | -0.209 (1.511) | -1.289 (1.643) | -1.689*** (0.569) |
| Observations | 15,715 | 1,995 | 1,876 | 1,129 | 5,373 | 2,555 | 312 | 469 | 1,974 |

Robust standard errors in parentheses

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 6: Probit, marginal effects (overeducation)

| Probit - Marginal effects - Dependent variable: Occupational mismatch (overeducation) | | | | | | | | |
|---|---------------|-----------------|-----------------|----------------|----------------|----------------------|-------------------|---------------|
| Variable | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | South-East Asians | Others |
| Referee | -.03416407** | -.04716396 | -.07967784*** | -.05377409 | -.03262112 | .01297174 | -.08590314 | -.02253394 |
| Arrived in France <5 years | -.09851691*** | -.01267758 | -.02689076 | -.10181605 | -.10425027*** | -.11151062*** | -.09004878 | -.17834726*** |
| French diploma | -.05064081*** | -.03922264 | -.1202552*** | .05265545 | -.07183882*** | .00887962 | -.13508632* | -.01421445 |
| Married | .01928768* | -.01715612 | -.04367308* | -.03013356 | .04825966** | .02200966 | -.03655568 | .05299785 |
| Gender | .07693092*** | .12563821*** | .08462107*** | .17952987*** | .06141224*** | .01372262 | .03576337 | .11618656*** |
| Child | -.03796765*** | -.05753434* | .01771258 | -.06439492 | -.03983954* | -.03042464 | -.11786662* | -.05140295 |
| Probit - Marginal effects - Dependent variable: Referee (overeducation) | | | | | | | | |
| Variable | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | South-East Asians | Others |
| Neighborhood average education level | -.00399075 | -.00412963 | -.01864412* | -.01000371 | -.00119582 | .00762373 | -.00086311 | -.01105868 |
| Share of peers (region) | .04683964 | .6163037 | -.34367289 | .13860509 | .27945354 | -.30156328 | -.28242783 | .06147576 |
| Arrived in France <5 years | .01780413* | .05702235* | .00958512 | .00758581 | .00773257 | .01786082 | -.019978 | -.02054899 |
| French diploma | -.01920773* | -.03221182 | -.05398849 | -.08372196** | -.00086556 | -.01110217 | -.00138136 | -.02953705 |

legend: * p<.05; ** p<.01; *** p<.001

Table 7: Dependent variable: Occupational mismatch, overeducation

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|----------------------------|----------------------|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|----------------------|
| | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | Turks | South-East Asians | Others |
| Reference | -0.097*** (0.032) | -0.132 (0.096) | -0.336*** (0.101) | -0.146 (0.112) | -0.095* (0.058) | 0.034 (0.080) | 0.228 (0.199) | -0.255 (0.173) | -0.060 (0.083) |
| Arrived in France <5 years | -0.278*** (0.031) | -0.035 (0.095) | -0.114 (0.094) | -0.276 (0.171) | -0.304*** (0.050) | -0.290*** (0.082) | -0.015 (0.282) | -0.267 (0.167) | -0.472*** (0.097) |
| French diploma | -0.143*** (0.029) | -0.109 (0.085) | -0.508*** (0.111) | 0.143 (0.101) | -0.269*** (0.052) | 0.023 (0.065) | -0.571*** (0.208) | -0.401** (0.198) | -0.038 (0.073) |
| Age | 0.039*** (0.009) | 0.020 (0.025) | -0.039 (0.034) | 0.083*** (0.031) | 0.017*** (0.017) | 0.073*** (0.022) | 0.084 (0.077) | 0.052 (0.066) | 0.020 (0.023) |
| Age ² | -0.001*** (0.000) | -0.000 (0.000) | 0.000 (0.000) | -0.001** (0.000) | -0.001*** (0.000) | -0.001*** (0.000) | -0.001 (0.001) | -0.001 (0.001) | -0.000 (0.000) |
| Married | 0.055** (0.027) | -0.048 (0.076) | -0.184** (0.084) | -0.082 (0.100) | 0.141*** (0.047) | 0.057 (0.062) | -0.284 (0.248) | -0.108 (0.158) | 0.140* (0.074) |
| Gender | 0.217*** (0.024) | 0.350*** (0.068) | 0.357*** (0.081) | 0.487*** (0.086) | 0.179*** (0.042) | 0.036 (0.058) | -0.159 (0.212) | 0.106 (0.130) | 0.308*** (0.063) |
| Child | -0.107*** (0.027) | -0.160** (0.078) | 0.075 (0.093) | -0.175* (0.098) | -0.116** (0.047) | -0.079 (0.062) | -0.226 (0.218) | -0.350** (0.167) | -0.136* (0.071) |
| Constant | -0.981*** (0.195) | -0.744 (0.538) | 0.149 (0.683) | -2.115*** (0.654) | -0.985*** (0.348) | -1.685*** (0.453) | -1.395 (1.450) | -0.590 (1.425) | -0.646 (0.485) |
| Observations | 15,715 | 1,955 | 1,876 | 1,133 | 5,373 | 2,555 | 312 | 483 | 2,011 |

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 8: Dependent variable: Referee (undereducation)

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|--------------------------------------|----------------------|--------------------|-------------------|-------------------|---------------------|----------------------|-------------------|----------------------|---------------------|
| | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | Turks | South-East Asians | Others |
| Neighborhood average education level | -0.023* (0.013) | -0.076* (0.044) | -0.033 (0.036) | -0.054 (0.056) | -0.011 (0.024) | -0.011 (0.039) | -0.090 (0.121) | -0.026 (0.077) | 0.000 (0.035) |
| Share of peers (region) | -0.091 (0.538) | -0.280 (3.396) | -1.333 (1.182) | 1.134 (0.710) | 0.657 (0.915) | -1.971 (1.205) | -0.839 (2.019) | -1.832*** (0.706) | 0.314 (0.354) |
| Arrived in France <5 years | 0.038 (0.040) | 0.210 (0.134) | -0.003 (0.088) | 0.092 (0.262) | -0.005 (0.066) | 0.088 (0.122) | 0.255 (0.290) | 0.050 (0.213) | -0.069 (0.140) |
| French diploma | -0.024 (0.040) | -0.162 (0.124) | -0.127 (0.121) | -0.286 (0.177) | -0.024 (0.076) | 0.015 (0.106) | -0.346 (0.251) | 0.009 (0.285) | -0.025 (0.111) |
| Constant | -1.259*** (0.311) | -0.576 (1.015) | -0.196 (0.814) | -0.075 (1.073) | -1.433** (0.572) | -0.496 (1.025) | -1.306 (1.832) | -3.825** (1.797) | -1.765** (0.716) |
| Observations | 10,546 | 1,315 | 1,633 | 539 | 3,806 | 1,387 | 214 | 302 | 1,096 |

Robust standard errors in parentheses

P<0.01, **p<0.05, *p<0.1

Table 9: Probit, marginal effects (undereducation)

| Probit - Marginal effects - Dependent variable: Occupational mismatch (undereducation) | | | | | | | | |
|--|--------------|-----------------|-----------------|----------------|----------------|----------------------|-------------------|--------------|
| Variable | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | South-East Asians | Others |
| Referee | .00035296 | -.00591703 | .0098043 | -.04622069 | -.003377 | -.01379763 | .11094597** | -.00069169 |
| Arrived in France <5 years | .0160712** | -.00527384 | .02691406* | .04531576 | .01383362 | .02204219* | .0605896 | .01892445 |
| French diploma | -.00055423 | .01601275 | .0073594 | -.04830673 | .0067517 | -.00624149 | (omitted) | -.05173276** |
| Married | .01289362** | .0479554** | .01976683 | .011916 | .0160934* | -.00273933 | .04233959 | -.01954859 |
| Gender | -.00892277* | -.0065231 | -.02757932* | -.01976463 | -.00477543 | -.00521062 | -.02492604 | -.01016043 |
| Child | -.01329114** | -.01489971 | -.01413363 | -.03883822 | -.01265628 | -.01910241* | -.02549014 | .00294335 |
| Probit - Marginal effects - Dependent variable: Referee (undereducation) | | | | | | | | |
| Variable | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | South-East Asians | Others |
| Neighborhood average education level | -.00599585 | -.01773103 | -.0101295 | -.01398383 | -.00258882 | -.00232398 | -.00732467 | .00004154 |
| Share of peers (region) | -.02352407 | -.06503697 | -.40333768 | .29410952 | .1587732 | -.41495609 | -.51416829** | .08107789 |
| Arrived in France <5 years | .00981896 | .04868955 | -.00105211 | .02373831 | -.00115127 | .01848749 | .01415362 | -.01780696 |
| French diploma | -.00605657 | -.03768686 | -.03841658 | -.07412578 | -.00578414 | .00939104 | .00250629 | -.00654068 |

legend: * p<.05; ** p<.01; *** p<.001

Table 10: Dependent variable: Occupational mismatch, undereducation

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|----------------------------|----------------------|---------------------|---------------------|-------------------|----------------------|----------------------|--------------------|----------------------|
| | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans | South-East Asians | Others |
| Reference | 0.004 (0.071) | -0.056 (0.227) | 0.108 (0.136) | -0.429 (0.287) | -0.044 (0.125) | -0.262 (0.234) | 1.069** (0.395) | -0.009 (0.212) |
| Arrived in France <5 years | 0.203*** (0.063) | -0.050 (0.196) | 0.296** (0.135) | 0.421 (0.509) | 0.181* (0.093) | 0.419** (0.188) | 0.584 (0.408) | 0.236 (0.245) |
| French diploma | -0.007 (0.070) | 0.152 (0.180) | 0.081 (0.195) | -0.449 (0.312) | 0.088 (0.129) | -0.119 (0.171) | | -0.646*** (0.209) |
| Age | -0.008 (0.019) | -0.038 (0.043) | -0.004 (0.058) | 0.015 (0.064) | 0.049 (0.037) | 0.015 (0.050) | -0.099 (0.117) | -0.079 (0.055) |
| Age ² | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.001) | -0.000 (0.001) | -0.000 (0.000) | -0.000 (0.001) | 0.001 (0.001) | 0.001 (0.001) |
| Married | 0.163*** (0.056) | 0.455*** (0.136) | 0.217 (0.133) | 0.111 (0.272) | 0.210** (0.095) | -0.052 (0.156) | 0.408 (0.400) | -0.244 (0.179) |
| Gender | -0.113** (0.052) | -0.062 (0.131) | -0.303** (0.120) | -0.184 (0.196) | -0.062 (0.090) | -0.099 (0.151) | -0.240 (0.412) | -0.127 (0.175) |
| Child | -0.108*** (0.056) | -0.141 (0.129) | -0.155 (0.138) | -0.361 (0.269) | -0.165* (0.091) | -0.363** (0.173) | -0.246 (0.387) | 0.037 (0.174) |
| Constant | -1.579*** (0.399) | -1.452 (0.926) | -2.048 (1.277) | -1.289 (1.314) | -3.195*** (0.827) | -2.647** (1.230) | -0.239 (2.335) | 0.056 (1.178) |
| Observations | 10,307 | 1,192 | 1,553 | 412 | 3,729 | 1,365 | 156 | 855 |

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 11: Biprobit, marginal effects

| Marginal effects - Dependent variable : Occupational mismatch (overeducation) | | | | | | |
|---|---------------|-----------------|-----------------|----------------|----------------|----------------------|
| Variable | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans |
| Referee | -.53063468*** | -.54261888*** | .22882602 | -.21568799 | .49338491*** | .48070367*** |
| Arrived in France <5 years | -.07382163*** | .03006682 | .03045257 | -.09652943 | -.08364029*** | -.09977707*** |
| French diploma | -.06228805*** | -.05175768 | -.080368** | .0364181 | -.05432645*** | .01282024 |
| Age | .00907212** | .00575924 | -.01332596* | .02802132 | .01186036* | .02081441* |
| Age ² | -.00013875*** | -.00009193 | .00014739 | -.0003242 | -.00019959*** | -.00024831* |
| Married | .02125919** | .00096603 | -.0234515 | -.01206015 | .04207958** | .02261212 |
| Gender | .0522444*** | .10705338*** | .04031058* | .16650978* | .04894367*** | .02568293 |
| Child | -.03571671*** | -.03440385 | .02674312 | -.07770352 | -.02526679 | -.02760905 |
| Marginal effects - Dependent variable : Referee | | | | | | |
| Variable | Immigrants | North-Europeans | South-Europeans | East-Europeans | North-Africans | Sub-Saharan Africans |
| Arrived in France <5 years | .0116947 | .04737953 | .01321951 | .00579594 | .01202447 | .01487884 |
| Share of peers (region) | .07981245 | .4420115 | .33426977 | 3.2705071 | .2987521 | .27321178 |
| French diploma | -.01380139 | -.03111849 | -.05772278 | -.08303056 | -.0006418 | -.02053955 |
| Age | .00535049* | -.00005067 | -.00204642 | -.01062939 | -.00106786 | .00275092 |
| Age ² | -.00006183* | -4.077e-06 | .00001884 | .00008653 | .00001748 | -.00001907 |
| Married | .01379902 | .01631399 | .03154021 | .10267411 | -.00694536 | -.01736172 |
| Gender | -.00503177 | .02472925 | -.0090608 | -.03881347 | -.00653522 | -.03085361* |
| Child | -.02040482** | .00518493 | .00167045 | -.09503678 | -.00436911 | .01427053 |
| Neighborhood average education level | -.01993269*** | -.02317642*** | -.02284294* | -.01381388 | .01523274** | .01835391** |

legend: * p<.05; ** p<.01; *** p<.001

References

- Aleksynska, Mariya, and Ahmed Tritah. 2013. "Occupation–education mismatch of immigrant workers in Europe: Context and policies." *Economics of Education Review* 36:229–244.
- Anne, Denis, and Sylvain Chareyron. 2017. "Une analyse spatiale du non-recours aux dispositifs sociaux." *Revue d'économie politique* 127 (2): 227–253.
- Aoki, Yu, and Lualhati Santiago. 2015. "Education, health and fertility of UK immigrants: The role of English language skills."
- Autor, David. 2010. "The polarization of job opportunities in the US labor market: Implications for employment and earnings." *Center for American Progress and The Hamilton Project*, vol. 6.
- Baert, Stijn, Bart Cockx, and Dieter Verhaest. 2013. "Overeducation at the start of the career: stepping stone or trap?" *Labour Economics* 25:123–140.
- Barnes, J. A. 1954. "Class and Committees in a Norwegian Island Parish." *Human Relations* 7 (1): 39–58.
- Bauer, Thomas K. 2002. "Educational mismatch and wages: a panel analysis." *Economics of Education Review* 21 (3): 221–229.
- Bauer, Thomas K., Gil Epstein, and Ira Gang. 2002. "Herd effects or migration networks? The location choice of Mexican immigrants in the US."
- Becker, Gary S. 1964. *Human capital*. New York: Columbia University Press.
- Bentolila, Samuel, Claudio Michelacci, and Javier Suarez. 2010. "Social contacts and occupational choice." *Economica* 77 (305): 20–45.
- Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou. 2011. "Ethnic identity and labour market outcomes of immigrants in Europe." *Economic Policy* 26 (65): 57–92.
- Bisin, Alberto, and Thierry Verdier. 2001. "The economics of cultural transmission and the dynamics of preferences." *Journal of Economic theory* 97 (2): 298–319.
- Calvo-Armengol, Antoni, and Matthew O. Jackson. 2004. "Social networks in determining employment: Patterns, dynamics, and inequality." *American Economic Review* 94 (3): 426–454.
- Calvo-Armengol, Antoni, and Yves Zenou. 2004. "Social networks and crime decisions: The role of social structure in facilitating delinquent behavior." *International Economic Review* 45 (3): 939–958.

- Chiswick, Barry R., and Paul W. Miller. 1990. "Language in the Labor Market: The Immigrant Experience in Canada and the United States."
- . 2010. "Does the choice of reference levels of education matter in the ORU earnings equation?" *Economics of Education Review* 29 (6): 1076–1085.
- Chort, Isabelle. 2017. "Migrant network and immigrants' occupational mismatch." *The Journal of Development Studies* 53 (11): 1806–1821.
- Constant, Amelie, and Klaus Zimmermann. 2003. "Occupational choice across generations."
- David, H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The polarization of the US labor market." *American economic review* 96 (2): 189–194.
- Duncan, Greg J., and Saul D. Hoffman. 1981. "The incidence and wage effects of overeducation." *Economics of Education Review* 1 (1): 75–86.
- Dustmann, Christian, and Francesca Fabbri. 2003. "Language proficiency and labour market performance of immigrants in the UK." *The Economic Journal* 113 (489): 695–717.
- Dustmann, Christian, and Arthur Van Soest. 2002. "Language and the earnings of immigrants." *ILR Review* 55 (3): 473–492.
- Edin, Per-Anders, Peter Fredriksson, and Olof AAslund. 2003. "Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment." *The quarterly journal of economics* 118 (1): 329–357.
- Erdogan, Berrin, and Talya N. Bauer. 2009. "Perceived overqualification and its outcomes: The moderating role of empowerment." *Journal of Applied Psychology* 94 (2): 557.
- Freeman, Richard. 1976. *The Overeducated American*. Academic Press.
- Gibbons, Stephen, Olmo Silva, and Felix Weinhardt. 2013. "Everybody needs good neighbours? Evidence from students' outcomes in England." *The Economic Journal* 123 (571): 831–874.
- Goos, Maarten, Alan Manning, and Anna Salomons. 2009. "Job polarization in Europe." *American economic review* 99 (2): 58–63.
- Granovetter, Mark. 1974. *Finding a job: A study on contacts and careers*. Cambridge, Mass.: Harvard University Press.
- Jackson, Matthew O. 2010. *Social and economic networks*. Princeton university press.

- Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *The Quarterly Journal of Economics* 82 (2): 175–197.
- Kalfa, Eleni, and Matloob Piracha. 2018. "Social networks and the labour market mismatch." *Journal of Population Economics* 31 (3): 877–914.
- Lewis, Ethan G. 2011. "Immigrant-native substitutability: The role of language ability." Technical Report, National Bureau of Economic Research.
- Liu, Xiaodong, Eleonora Patacchini, Yves Zenou, and Lung-Fei Lee. 2012. "Criminal networks: Who is the key player?"
- Maddala, G. S. 1983. "Methods of estimation for models of markets with bounded price variation." *International Economic Review*, pp. 361–378.
- McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a feather: Homophily in social networks." *Annual review of sociology* 27 (1): 415–444.
- Montgomery, James D. 1991. *Social networks and persistent inequality in the labor market*. Center for Urban Affairs and Policy Research Evanston.
- Munshi, Kaivan. 2003. "Networks in the modern economy: Mexican migrants in the US labor market." *The Quarterly Journal of Economics* 118 (2): 549–599.
- Nordin, Martin, Inga Persson, and Dan-Olof Rooth. 2010. "Education–occupation mismatch: Is there an income penalty?" *Economics of education review* 29 (6): 1047–1059.
- Patel, Krishna, and Francis Vella. 2013. "Immigrant networks and their implications for occupational choice and wages." *Review of Economics and Statistics* 95 (4): 1249–1277.
- Piracha, Matloob, Massimiliano Tani, and Matias Vaira-Lucero. 2016. "Social capital and immigrants' labour market performance." *Papers in Regional Science* 95:S107–S126.
- Sarkar, Sudipa. 2017. "Employment polarization and over-education in Germany, Spain, Sweden and UK." *Empirica* 44 (3): 435–463.
- Schuetze, Herbert, and Jonathan Wood. 2013. "Immigrant Networks, Occupational Choice, and Match Quality." Technical Report, Mimeo. University of Victoria.
- Sicherman, Nachum. 1991. "' Overeducation" in the Labor Market." *Journal of labor Economics* 9 (2): 101–122.
- Stark, Oded, and David E. Bloom. 1985. "The New Economics of Labor Migration." *The American Economic Review* 75 (2): 173–178.

- Støren, Liv Anne, and Jannecke Wiers-Jenssen. 2010. "Foreign diploma versus immigrant background: Determinants of labour market success or failure?" *Journal of Studies in International Education* 14 (1): 29–49.
- Verdugo, Richard R., and Naomi Turner Verdugo. 1989. "The impact of surplus schooling on earnings: Some additional findings." *Journal of Human Resources*, pp. 629–643.
- Verhaest, Dieter, and Eddy Omey. 2006. "Discriminating between alternative measures of over-education." *Applied Economics* 38 (18): 2113–2120.
- . 2009. "Objective over-education and worker well-being: A shadow price approach." *Journal of Economic Psychology* 30 (3): 469–481.
- Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. MIT press.
- Yao, Yuxin, and Jan C. van Ours. 2015. "Language skills and labor market performance of immigrants in the Netherlands." *Labour Economics* 34:76–85.
- Zorlu, Aslan, and Joop Hartog. 2018. "The Impact of Language on Socioeconomic Integration of Immigrants."