

Occupational choices and network effects: Evidence from France

Arnaud Herault* Eva Moreno - Galbis [†] Francois Charles Wolff[‡]

Very Preliminary Version. Do not quote without permission. Do not circulate.

Abstract

Immigrants have a tendency to cluster towards cities, industries and occupations where there is a large share of their country peers. While there is a large consensus in the literature concerning the validity of historical settlement patterns across both host countries and/or occupations by geographical origin as an instrument for immigrants' current inflows across occupations, there are few studies outside the US or Canada analyzing the role of social networks as a determinant of immigrants' occupational choices. Using the French Labor Survey for the period 1993-2012, the contribution of this paper is twofold. First, based on a simple theoretical framework we estimate a reduced form equation using an instrumental variable strategy. We find that social networks clearly determine newly arrived immigrants' occupational choices in France. This is consistent with findings on the US or Canada. Second, contrary to most of the literature, our paper also analyzes the changing role of social networks along the economic cycle. Given the structural labor market characteristics of France, our theoretical framework predicts that social networks should play a more important role as a determinant of immigrants' occupational choices during recessions than during expansions. Our econometric estimations confirm this theoretical result.

Keywords: immigrants, network effects, occupational choices

JEL: J61; J24;

1 Introduction

Social interactions and peer effects have proved to be crucial in various aspects of economic activities, including education, crime, smoking, teenage pregnancy, school dropout, etc. One population

*GRANEM-University of Angers

[†]Corresponding author. GREQAM (University of Aix-Marseille), IUF, IRES (Catholic University of Louvain).

Email: Eva.moreno-galbis@univ-amu.fr

[‡]LEMNA-University of Nantes

subgroup of individuals for whom local interactions are particularly relevant is immigrants. As in Patel and Vella (2013) we define as network the fact that individuals that come from the same country and who are located in the same region are likely to have a relatively higher propensity to interact. This type of network membership might influence an individual's propensity to take employment in the same occupation as the other network members. On the basis of this definition, this paper evaluates how networks influence immigrants occupational choices along the economic cycle, *i.e.* expansions vs. recessions.

As already shown in the nineties by Altonji and Card (1991), immigrants and natives tend to work in different industries and jobs. Immigrants are essentially attracted to cities with a large concentration of their country peers. The empirical economic literature on networks in the labor market has focused either on unemployment or wages while largely ignoring occupational choice. However, empirical evidence reveals that around 50% of individuals obtain or hear about jobs through friends and family (see Holzer (1987), Holzer (1988), Montgomery (1991), Granovetter (1995) or Addison and Portugal (2002)). This share should be even more important among immigrants for whom social networks and peer effects have proven to be crucial on economic activities.

The importance of local social networks for location decisions of migrants and their labor market outcomes has been empirically examined among others by McKenzie and Rapoport (2010), Munshi (2003) or Patel and Vella (2013).¹ McKenzie and Rapoport (2010) examine the role of migration networks in determining self-selection patterns of Mexico-US migration. Using survey data from Mexico, they find that in communities with small migration networks, the probability of migration is increasing with education, resulting in positive selection of migrants. In contrast, in communities with large networks, where migration costs are lower, the propensity to migrate decreases with education. Working with Mexican workers in the US, Munshi (2003) finds that the more established members of the community contribute disproportionately to the network and that it is the disadvantaged members (women, elderly and non qualified) who benefit the most. Contrary to McKenzie and Rapoport (2010), he finds that the same individual is more likely to be employed and to hold a preferred nonagricultural job when his network is exogenously larger. Our paper follows closely the study proposed by Patel and Vella (2013). Based on the US Census

¹The resulting spatial segregation at the residential level is studied in Bartel (1989), Musterd (2005), Cutler, Glaeser, and Vigdor (2008) and Semyonov and Glikman (2009). Firm-level segregation of minority groups is analyzed in Carrington and Troske (1998), Hellerstein and Neumark (2003), Hellerstein and Neumark (2008) and Hellerstein, McInerney, and Neumark (2007) for the US, Aslund, Hensvik, and Skans (2009) for Sweden and Dustmann, Glitz, and Schönberg (2011) for Germany.

data 1980, 1990 and 2000, the authors examine the relationship between the occupational choices of recently arrived immigrants and those established immigrants from the same country. Patel and Vella (2013) find that the occupational share of certain ethnic groups grew drastically in particular labor markets over the period from 1980 to 2000. Moreover, the pattern of growth is consistent with the presence of network effects. The data also does not appear to suggest that the allocation observed is the result of sorting on the basis of a comparative advantage. Waldinger (1996) suggests that the rapid movement of white native workers away from certain jobs in New York city enabled ethnic minorities to form niches. Networks then channelled immigrants of specific ethnic backgrounds into specific occupations. Waldinger (1994) finds that immigrants sorted into different occupations within the New York city government according to ethnic background.

While the role of networks as a determinant of occupational choices has been less analyzed in the literature than other topics, there is large consensus among economics regarding the use of historical settlement patterns across both host countries and/or occupations by origin country as an instrument for immigrants' current inflows when trying to solve the endogeneity problem arising when evaluating labor market outcomes of immigrants in the host country or the impact of immigrants on natives' labor market outcomes. Such an instrument has proven to be a strong determinant of contemporaneous inflows in both the single-country case (see Card (2001), Card (2009), Cortes and Tessada (2011) or Patel and Vella (2013) for the US and Gonzalez and Ortega (2011) or Farré, González, and Ortega (2011) for Spain among others) and in the multi-country setting (see Angrist and Kugler (2003), D'Amuri and Peri (2014) or Moreno-Galbis and Tritah (2016)).

A large consensus arises then on the fact that immigrants tend to cluster into labor markets with a higher share of their country peers. Informational networks are likely to be the main responsible of these geographical and occupational location choices. However, our guess is that the importance of these networks as a determinant of immigrants' choices is likely to vary along the economic cycle. The effect is in principle ambiguous. On the one hand, in an economic context characterized by a constrained number available jobs, informational networks within communities should play a relatively more important role (with respect to economic expansions) in determining occupational choices, since the additional information and the reduction of uncertainty provided by networks is likely to become more valuable during economic recessions.² On the other hand, the opposite reasoning may apply and we can argue that, during recessions, due to the scarcity

²For employers these networks may represent a mechanism for screening potential workers.

of jobs, meritocracy will be the main determinant of job access and relative importance of social networks in occupational choices will decrease.

We propose a simple theoretical model providing foundations to the reduced form equation we estimate using French data for the period 1993-2012. More precisely, we examine if the occupational choices of recently arrived immigrants are influenced by the occupational location of their predecessors. Moreover, we study whether the influence of social networks has changed along the economic cycle. Our reference labor market is the individual French region, therefore, we will exploit heterogeneity across time and across regions of newly arrived immigrants occupational choices, distinguishing by their geographical origin. France is likely to be a particular case. Due to its past colonial history France receives immigrants differing in origin from immigrant population in Canada, Australia, the US or other Western countries for which several studies on networks' effect have already been proposed. Immigrants from different origins are likely to behave differently. To our knowledge, we are the firsts to evaluate the role of social networks on immigrants occupational choices in France. Moreover, we propose an original contribution with respect to other papers in the literature by analyzing the influence of these networks along the economic cycle.

To deal with endogeneity issues (*i.e.* newly arrived immigrants may be flowing towards the same regions and occupations as their country peers simply because these regions and occupations are expanding), we instrument the occupational distribution of the established migrants with that of migrants in the same region 10 years ago. While this identifying strategy may be questionable if there are time persistent regional and occupational effects which affect the occupational choice of particular groups, we feel it is appropriate here due to the conditioning variables we include in the specifications we adopt.

Next section presents a theoretical framework providing theoretical foundations to our econometric estimations. Section 3 describes the database. Empirical motivation for our study is provided in Section 4. Section 5 describes the econometric approach and the main estimation results are presented in Section 6. Section 7 concludes.

2 Theoretical framework

We explain here a basic theoretical setup allowing to provide a simplified representation of the labor market. This framework proposes some theoretical foundations to the econometric approach presented by section 5 and it is largely inspired by the works of Hemet (2015) and Zenou (2015).

2.1 Notation

We consider a particular labor market l , which can be a region, a city or an occupation. This labor market is composed by two groups of individuals who are assumed to be identical except from their nativity origin:

- Individuals from ethnic group j
- Individuals from other groups (majority group or other ethnic groups) $-j$

At date t , the population of size N_t is composed of these two groups whose size equal respectively N_{jt} and N_{-jt} . For notational simplicity, we have eliminated subscript l , since our analysis will always be referred to this particular labor market. The share of each group in total population is denoted as n_{jt} and n_{-jt} , where $\frac{N_{jt}}{N_t} + \frac{N_{-jt}}{N_t} = n_{jt} + n_{-jt} = 1$.

Individuals may be employed or unemployed, so that we have:

$$N_{it} = E_{it} + U_{it} \Rightarrow 1 = e_{it} + u_{it} \quad \text{for } i = j, -j$$

where E_{it} stands for the number of employed, U_{it} for the number of unemployed, e_{it} for the employment rate and u_{it} for the unemployment rate.

While we do know that lots of jobs are obtained through networks, lots of jobs are not. We consider here the possibility of an unemployed worker learning of a vacancy directly from an employer as well as via a social interaction. Unemployed workers hear directly about a job at rate λ_d but job offers circulate also through word of mouth communication by means of employed workers (when the employers do not advertise their vacancies publicly but only tell their own employed workers about them). Employed workers hear of job vacancies at the exogenous rate λ_n . Employed workers lose their job at the exogenous rate δ .

The probability of finding a job will actually be an increasing function of the rate at which the individual hears about job offers directly or through networks. While most of the literature (see Calvo-Armengol, Verdier, and Zenou (2007), Hemet (2015) or Zenou (2015)) considers that the rate at which the individual hears about a job offer equals the probability of finding a job, we will assume here that due to frictions in the market there is not a perfect correspondence between the rate at which the individual hears about a vacancy and the probability of finding a job. The probability of finding a job is though an increasing function of the rate at which unemployed individuals hear about a job opportunity directly or through networks, but there is no strict equality between both of them, *i.e.* $p = f(\lambda_d, \lambda_n)$, where $p'(\lambda_d) > 0$ and $p'(\lambda_n) > 0$.

Any agent can be matched to an individual from his own ethnic group or from the other group, according to the size of each group in the population and with a bias towards his own group. Once a pair of agents is randomly formed, the individuals can exchange information about job opportunities. Employed workers are aware of a job offer at rate λ_n and will transmit this information if they meet unemployed workers. Because this economy features homophily, *i.e.* individuals tend to form links and gather with people similar to themselves along ethnic lines, employed workers systematically transmit information to individuals of their own ethnic group, but they communicate a job offer to someone of the other ethnic group only with a probability $\beta \in (0, 1)$. The network is thus characterized by a bias making it more difficult to get information from someone of a different group.

2.2 Flows

Employment dynamics between t and $t + \Delta t$ in labor market l for group j is given by:

$$\dot{E}_{jt} = p(\lambda_d, \lambda_n, t)U_{jt} - \delta E_{jt} \quad (1)$$

where:

$$p(\lambda_d, \lambda_n, t) = \theta e^{\lambda_d + \lambda_n \frac{E_{jt}}{N_t} + \lambda_n \beta \frac{E_{-jt}}{N_t}}$$

For technical simplicity, an exponential functional form is assumed to define the relationship between the probability of finding a job and the rate at which individuals hear about vacancies. This exponential relation is smoothed by the parameter θ . Note that $\theta < 1$ represents the efficient parameter of job finding after hearing about an employment opportunity. Because $\frac{E_{it}}{N_t} = \frac{E_{it}}{N_t} \frac{N_{it}}{N_{it}} = e_{it}n_{it}$ for $i = j, -j$, the probability of finding a job can be rewritten as:

$$p(\lambda_d, \lambda_n, t) = \theta e^{\lambda_d + \lambda_n e_{jt}n_{jt} + \lambda_n \beta e_{-jt}n_{-jt}}$$

Information on employment opportunities can arrive directly (at rate λ_d) or through employed individuals (networks). If these employed individuals hearing about a job offer belong to the ethnic group j , they will directly transfer the information to the unemployed individual of group j when they meet him. This happens at rate $\lambda_n e_{jt}n_{jt}$. In contrast, if the employed individual who is aware about a job offer belongs to the ethnic group $-j$, he will only transmit the information with probability β . This happens at rate $\lambda_n \beta e_{-jt}n_{-jt}$.

Equation (1) can then be rewritten as follows:

$$\dot{E}_{jt} = U_{jt} \theta e^{\lambda_d + \lambda_n e_{jt}n_{jt} + \lambda_n \beta e_{-jt}n_{-jt}} - \delta E_{jt} \quad (2)$$

Exits from employment in labor market l correspond to the proportion δ of employed individuals whose job is destroyed and entries to employment equal the proportion $p(\lambda_d, \lambda_n, t) = \theta e^{\lambda_d + \lambda_n e_{jt} n_{jt} + \lambda_n \beta e_{-jt} n_{-jt}}$ of unemployed individuals that have managed to find a job.

2.3 Steady State

At the steady state equilibrium, entries to employment should equal exits, so that the number of employed workers remains constant, $\dot{E}_{jt} = 0$. Dividing by N_{jt} equation (2) once $\dot{E}_{jt} = 0$, we find:

$$\delta e_{jt} = u_{jt} \theta e^{\lambda_d + \lambda_n e_{jt} n_{jt} + \lambda_n \beta e_{-jt} n_{-jt}} \quad (3)$$

Taking logs of the previous expression we find:

$$\log e_{jt} = \log \theta u_{jt} - \log \delta + \lambda_d + \lambda_n e_{jt} n_{jt} + \lambda_n \beta e_{-jt} n_{-jt} \quad (4)$$

The employment rate of the ethnic group j in a particular labor market l , which may be a region, a city, an occupation, etc. depends on the share of individuals of the same ethnic group that are employed and have access to job offers in the same labor market l , on the share of individuals of the other groups, $-j$, that are employed in l and have access to job offers and on the structural characteristics of labor market l : job arrival rate, λ_d , job destruction rate, δ and the efficiency of the matching process between unemployed workers in l and job offers, θu_{jt} .

Our econometric analysis estimates a reduced form of equation (4), where employment in occupation l of recently arrived immigrants belonging to group j will depend on network effects as well as on the structural characteristics of the considered labor market (occupation l) and individual characteristics of recently arrived immigrants.

2.4 Out of the steady state

Our econometric approach will also try to identify if the influence of networks as a determinant of occupational choices of recently arrived immigrants evolves along the economic cycle. We will then focus on the dynamics of employment out of the steady state. First, let us compute time variation of the employment rate $e_{jt} = \frac{E_{jt}}{N_{jt}}$:

$$\begin{aligned} e_{jt} \cdot N_{jt} &= E_{jt} \\ \frac{\partial e_{jt}}{\partial t} N_{jt} + \frac{\partial N_{jt}}{\partial t} e_{jt} &= \frac{\partial E_{jt}}{\partial t} \\ e_{jt} N_{jt} + \dot{N}_{jt} e_{jt} &= \dot{E}_{jt} \end{aligned} \quad (5)$$

Replacing the previous expression in equation (1) and dividing by N_{jt} we obtain:

$$\begin{aligned}
\dot{e}_{jt}N_{jt} + \dot{N}_{jt}e_{jt} &= p(\lambda_d, \lambda_n, t)U_{jt} - \delta E_{jt} \\
\dot{e}_{jt} + \frac{\dot{N}_{jt}}{N_{jt}}e_{jt} &= p(\lambda_d, \lambda_n, t)u_{jt} - \delta e_{jt} \\
\dot{e}_{jt} &= p(\lambda_d, \lambda_n, t)u_{jt} - \delta e_{jt} - \gamma_{N_j}e_{jt}
\end{aligned} \tag{6}$$

where γ_{N_j} stands for the growth rate in the number of people from the ethnic group j .

To identify the main drivers of employment dynamics along the economic cycle in France, we combine information from two sources: OECD Statistics and the paper by Hairault, Barbanchon, and Sopraseuth (2015). As shown in Appendix A during the Great Recession, French employment rates remained fairly stable as compared with OECD average or as compared with the progression of the unemployment rate, which followed a continuously increasing path. Moreover, working with the French Labor Flow Survey Hairault, Barbanchon, and Sopraseuth (2015) find that the job separation rate and the job finding rate contributed to unemployment fluctuations during the nineties (50:50 split), while in the last decade the job finding rate was more significant and explained around 65% of the French unemployment fluctuations.

Putting together OECD statistics and the findings of Hairault, Barbanchon, and Sopraseuth (2015), equation (6) reveals that French employment dynamics over the past years is strongly driven by the terms $\gamma_{N_j}e_{jt}$ and $p(\lambda_d, \lambda_n, t)u_{jt}$. Concerning $\gamma_{N_j}e_{jt}$, as revealed by Appendix A, the employment rate has remained fairly constant during the last recession (2009-2014). Actually, if we consider the evolution of the employment rate between 2003 and 2014, we realize that, apart from the period going from 2007 to the end of 2008, the average employment rate in France has always been around 64%. This relative stability is in clear contrast with respect to the unemployment rate, which has followed a continuously increasing path since 2008 and has considerably increase the gap with respect to the average OECD unemployment rate since 2010.

This progression of the employment and unemployment rates suggest that the dynamics of employment in a labor market l during the economic cycle is driven by both entries to employment, *i.e.* $u_{jt}p(\lambda_d, \lambda_n, t)$, and the growth rate in the number of individuals from group j , N_j (for the term $\gamma_{N_j}e_{jt}$), since the term δe_{jt} should not have changed significantly.

It is interesting now to focus on entries to employment, *i.e.* $p(\lambda_d, \lambda_n, t)u_{jt}$. As concluded by Hairault, Barbanchon, and Sopraseuth (2015), job finding rates have been the main driver of unemployment dynamics during the past decade, suggesting that during recession periods they strongly fall and then recover during expansions. As specified previously, in our case, there is not

a direct correspondence between the probability of finding a job and the rate at which individuals hear about job offers. Following the findings of Hairault, Barbanchon, and Sopraseuth (2015), we consider that during recessions there is a decrease in the rate at which firms post job offers and therefore the rate at which individuals hear about these job offers, *i.e.* λ_d and λ_n . At the same time, according to OECD Statistics, e_{jt} and e_{-jt} should have remained fairly constant along the economic cycle. For a given θ , the size of networks, measured by the share of ethnic group j in total population, n_j , should increase their relative importance during recessions since λ_d and λ_n are falling and e_{jt} and e_{-jt} remain almost unchanged. Therefore, during contraction periods, the probability of finding a job *i.e.* $p(\lambda_d, \lambda_n, t) = \theta e^{\lambda_d + \lambda_n e_{jt} n_{jt} + \lambda_n \beta e_{-jt} n_{-jt}}$, will be more influenced by networks (size of n_{jt} and n_{-jt}) than during expansions periods when the e_{it} for $i = j, -j$ is high and so are λ_d and λ_n .

In sum, in a French type economy, characterized by relative stable employment rates along the economic cycle and strongly counter-cyclical job finding rates, the size of the networks should increase its relative importance as a determinant of the probability of finding a job during recessions periods. Our econometric approach will try to test this result.

3 Data

Contrary to many papers dealing with migration issues (see Ortega and Verdugo (2014) or Patel and Vella (2013)), this paper is not based on census data. We use data from the French Labor Force Survey (LFS) for the period going from 1993 to 2012. The LFS was launched in 1950 and established as an annual survey in 1982. Redesigned in 2003, it is now a continuous survey providing quarterly data. Participation is compulsory and it covers private households in mainland France. All individuals in the household older than 15 are surveyed.

The quarterly sample is divided into 13 weeks. From a theoretical point of view, the sampling method consists of a stratification of mainland France into 189 strata (21 French regions x 9 types of urban unit) and a first stage sampling of areas in each stratum (with different probabilities, average sampling rate= 1/600). Areas contain about 20 dwellings and among them only primary residences are surveyed. Each area is surveyed over 6 consecutive quarters. Every quarter, the sample contains 6 sub-samples: 1/6 of the sample is surveyed for the first time, 1/6 is surveyed for the second time, ..., 1/6 is surveyed for the 6th (and last) time. When it was run as an annual survey, every year a third of the sample was renewed meaning that each individual was interviewed

4 times. The collection method has always been a face-to-face interview.³ Therefore, even if we work with data from 1993 to 2012 we do not have a panel of individuals, the database must be rather understood as a sequence of cross-sections.

Topics covered by the LFS concern employment, unemployment, underemployment, hours of work, wages, duration of employment and unemployment (length of service), discouraged workers, industry, occupation, status in employment, education/qualification, and other jobs. The French LFS provides the occupation for each employed individual among a list of 350 possible occupations such as “gardener”, “messenger”, “clerk in banking activities”, or “financial manager”. To ensure a sufficient number of observations by occupation we regroup into 315 occupations.

Farmers, civil servants, the military and clergymen are excluded. All jobs related to these categories are dropped from the sample. Some jobs may have disappeared, while new ones are emerging. The French LFS modified the job classification in 2003 in order to take into account the changes in occupations. We paid attention to having a consistent definition of jobs throughout the 20 years of our sample. There are no new occupations that cannot be included in the pre-2003 classification. We keep a 3 digit classification.

As in Patel and Vella (2013), we distinguish between “new” immigrants and “established” immigrants as we focus on the empirical relationship between the occupational choices of these two groups. Since we are working with data for the period 1993-2012, we consider as “established” immigrants those whose first year of residence is over 5 years before the interview. Concerning “new” immigrants, we first consider in this category immigrants arriving in France over the past 5 years before the interview. In our instrumental variable strategy we will though exclusively focus on immigrants arriving in France after 2001, since we instrument the share of established immigrants in the occupation with the share estimated for 1993-1995 (between 10 and 19 years before the arrival date of the immigrant).

While this dichotomy is determined by the limitations of the data it does not seem unreasonable. That is, determining what distinguishes “established” from “new” is arbitrary and the most important requirement is that one can identify recent arrivals, which is the case since we have the first year of residence in France.⁴ Note though that we are unable to know what was the first job

³Since 2003, a telephone interview has been employed for intermediate surveys (2nd to 5th).

⁴For the period 1991-2002, the first year of residence in France is missing for 44000 immigrants. To reduce the number of missing observations for this variable, we approximate the first year of residence in France by the seniority in the job only when this seniority is below 1991. This concerns 23547 observations. We do not make the approximation when seniority in the position is above 1991 since we do not want to potentially confound “new”

the immigrant occupied unless he has been residing in France for less than two years and he has not changed his occupation within this period, which is very unlikely to be the case.

We distinguish 10 nativity groups depending on their origin and their cultural proximity: (1) French, (2) North African: Algerian, Tunisian, Moroccan, (3) African: Original from other African country, (4) Turkish: Turkey, (5) South-East-Asian: Vietnamese, Cambodian, Laotian, (6) South-European: Italian, Portuguese, Spanish, Greek, (7) Northern European countries: German, Belgian, Dutch, Luxembourg, Irish, Danish, British, Swiss, Austrian, Norwegian, Swedish, (8) Eastern Europe and URSS: Yugoslavian, Polish, URSS, (9) South-America and (10) North-America.

To examine occupational choices by immigrant group in local labor markets it is necessary to have an operable definition of a local labor market. To guarantee a sufficient number of observations we adopt the region as the local labor market. France includes 21 regions. We will exploit then variability across 21 regions, 113 occupations and 10 nativity groups to evaluate the role of networks as a determinant of immigrants occupational choices. All in all, we have 655,524 observations and when focusing in immigrants the sample reduces to 73,783 observations. As the data comprises a large number of occupations and we also distinguish by area of birth, the small size of cells (where cells are defined at the region-occupation-nativity level) are likely to be an issue. We only keep region-occupation-nativity cells having at least 20 individuals.

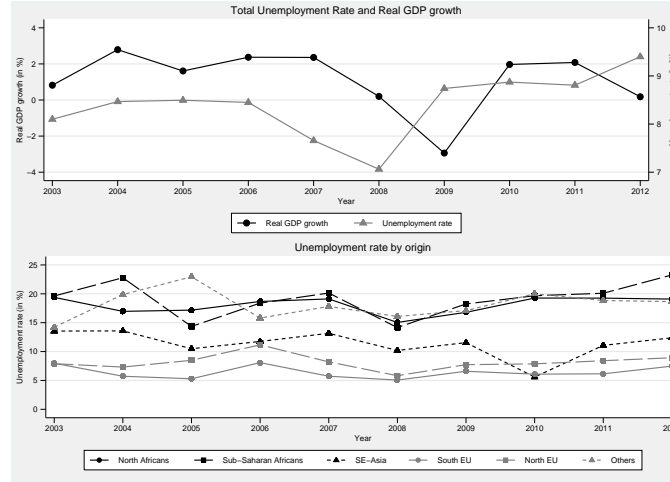
4 Empirical Motivation

This paper seeks to analyze the role of networks as a determinant of newly arrived immigrants' occupational choices as well as the potential changes in the role of networks along the economic cycle. We start providing some descriptive statistics on the French economic context and the impact on immigrants' unemployment rates and origin composition. In a second step, we provide some descriptive illustration on how immigrants occupational choices in France are likely to have been driven by the occupational choices of their peers rather than by a potential comparative advantage.

4.1 The French economic context

Panel (i) of Figure 1 shows the real GDP growth in France as well as the unemployment rate observed during the period 2001-2014. It highlights expansions and recessions undergone by France immigrants and “established” immigrants.

Figure 1: Panel (i): unemployment rate and real GDP growth in France (2001-2012). Panel (ii): unemployment rate by origins in France (2003-2012)

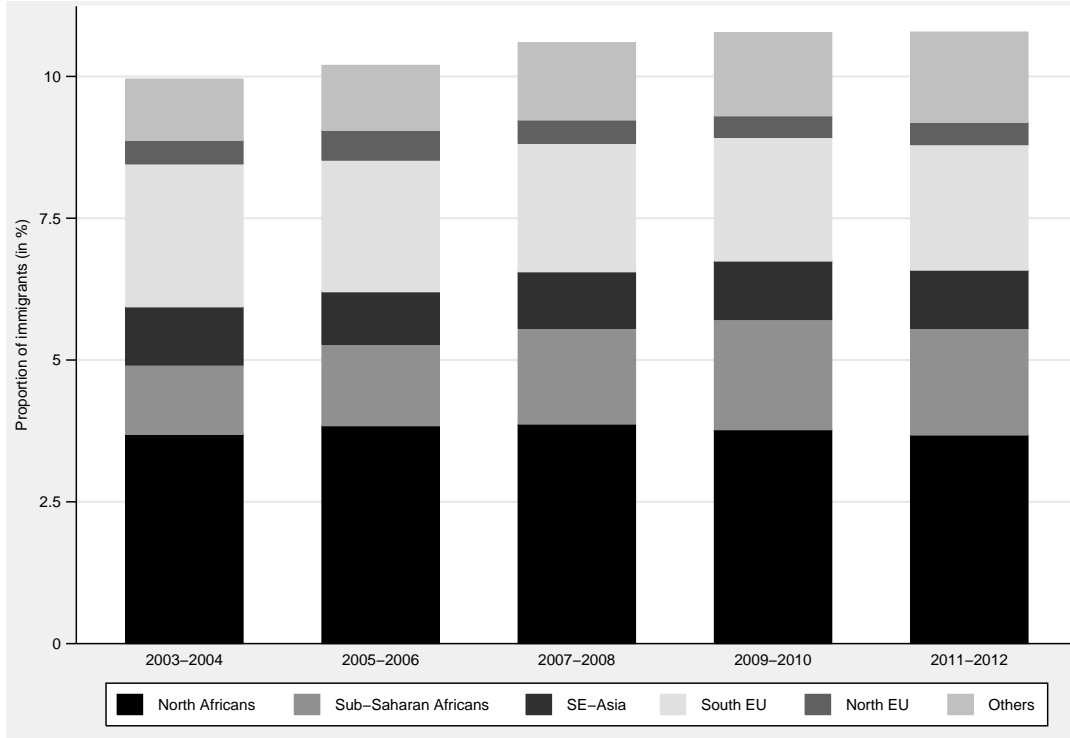


during these years. Unsurprisingly, since the 2008 crisis, unemployment rates in France have experience a very significant increase.

Panel (ii) of Figure 1 decomposes the unemployment rate observed over the same period by country of origin. In general, unemployment rates observed for the immigrant population are higher than the national unemployment rate. However, these rates are not homogeneous across geographical origins. Populations from European countries have lower unemployment rates than populations from countries outside the European Union. The impact of the 2008 crisis on unemployment rates was though very similar whatever the origin. In all cases we observe a rising trend.

While the share of immigrants over total population has remained fairly stable since 1993, its internal composition may have been modified. Figure 2 describes the evolution of migration composition in France since 1993. Because of its colonial past, migration in France is likely to display some peculiarities. Immigrants from North-Africa are the most numerous nativity group among immigrants, representing above 35% of the migrant population. Moreover, their share has remained stable along the considered period. Conversely, the proportion of South-Europeans has followed a continuously decreasing path since 1993, falling from more than 35% of total migrant population to 20%. Apart from Turks and, to a less extent, East-Europeans, whose share has followed an increasing path, the importance of the rest of nativity groups among the immigrant population has remained essentially constant since 1993.

Figure 2: Proportion of immigrants from origin o over total population, by periods



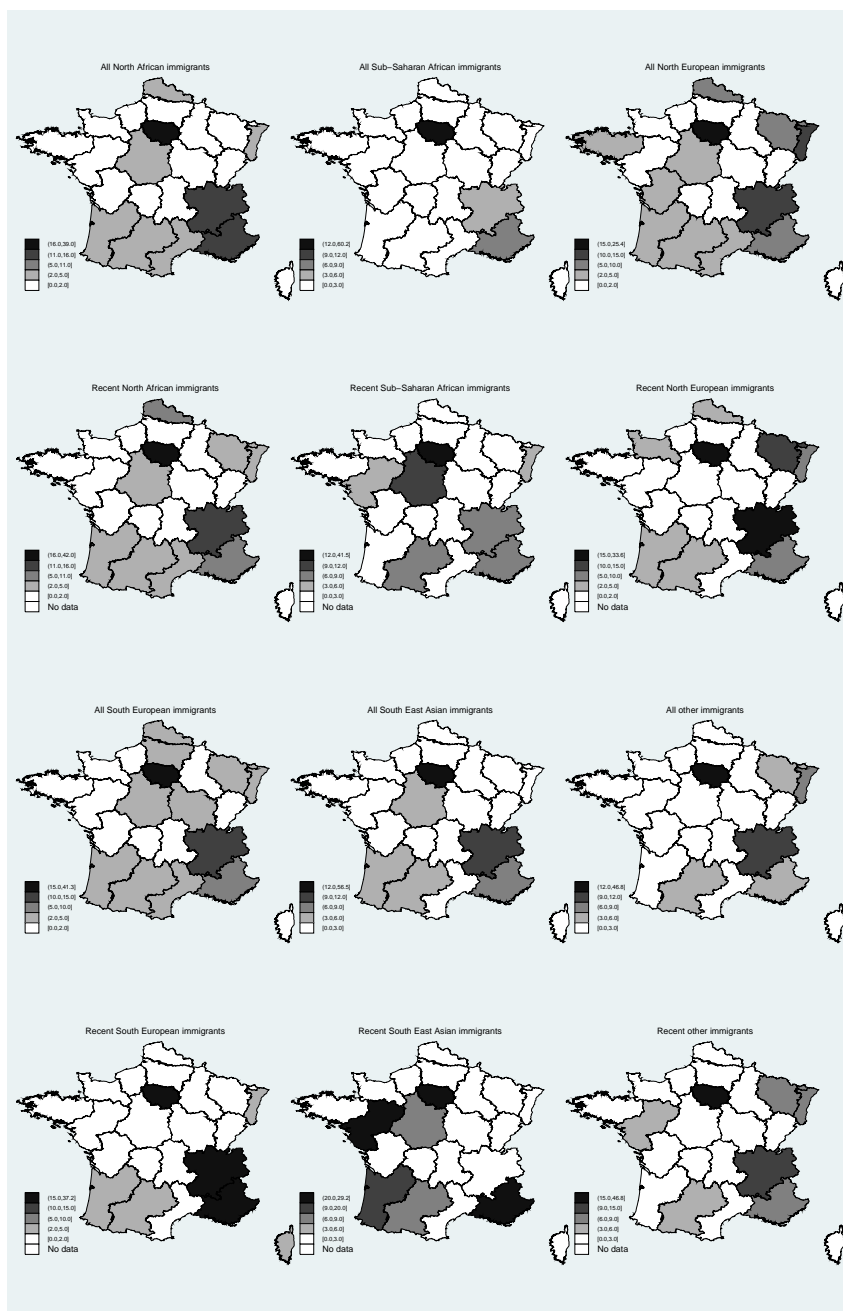
4.2 Immigrants' spatial and occupational allocation

Figure 3 shows the spatial distribution of migrants across the French regions. To ensure a sufficient number of data, we have represent in Figure 3 only individuals from 6 geographical origin: North Africans, Africans, South-Europeans, North-Europeans, South East Asians and "Others" (including Turks, East Europeans, South Americans and North Americans).

Most immigrants are mainly located in Ile de France region. According to the origin there are though "regional preferences". North Africans clearly distribute themselves between Ile de France and southern regions. In contrast, for Africans, Ile de France is by far the most preferred region with 60% of Africans located in the Ile de France. The spatial distribution of South and North Europeans is relatively homogeneous between French regions. However, we can observe a certain geographical preference between regions of residence in France and geographical proximity to the country of origin. Apart from the Ile de France, Southern Europeans are located mainly in French southern regions whereas North-Europeans are located more in French northern regions.

Figure ?? shows the spatial distribution of recently arrived immigrant by geographical origin. Ile de France keeps its attractiveness for new arrivals too. Recently arrived North Africans follow the

Figure 3: Spatial allocation of immigrants and newcomers, by origins, average for the period 1993-2012 (in percentage)



same pattern of spatial distribution across regions as established immigrants from the same origin. Newly arrived Africans respect the spatial distribution of their country peers already established in France, but they seem to be also flowing towards new regions. North and South Europeans follow essentially the same pattern of spatial distribution as their already established peers.

To test the presence of a network effect, we propose an indicator that compares the spatial distribution of immigrants by country of origin in period 1993-1996 with respect to the spatial distribution observed for newly arrived immigrants during the period 2003-2012:

$$NetIndicator = \frac{N_{og0}}{N_{o0}} - \frac{N_{ogT}}{N_{oT}} \quad (7)$$

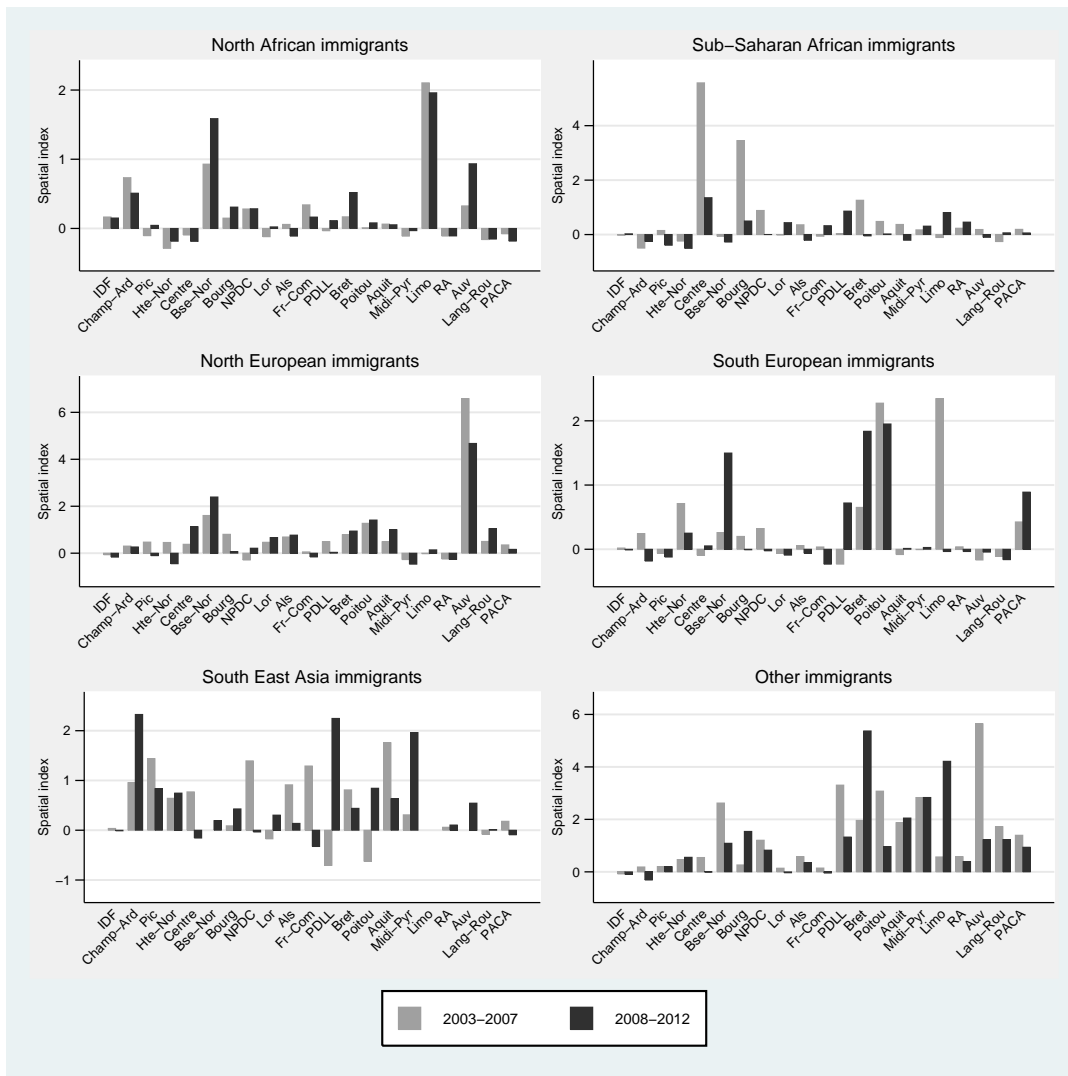
where N_{og0} stands for the number of immigrants from origin o in region g during the period 1993-1996, N_{o0} represents for the total number of immigrants in France from origin o during the period 1993-1996, N_{ogT} corresponds to the number of recently arrived immigrants from origin o in region g during the period 2003-2012 and N_{oT} is the total number of recently arrived immigrants in France from origin o during the period 2003-2012.

If social networks are important in location choices of newly arrived immigrants, we expect our indicator to be close to zero, since the spatial distribution of immigrants for the period 1993-1996 should be a perfect predictor of the newly arrived immigrant distribution for the period 2003-2012. When the indicator is close to 0 the network effect is relevant since the attractiveness of the region remains the same over time for newly arrived immigrants from a given origin. If the indicator is less than 0, there are less newly arrived individuals from the considered origin flowing to the region with respect to what it should be expected if social networks were strongly influencing location choices. Finally, if the indicator is greater than 0, the network effect is reinforced and newcomers from the considered origin are flowing more than it should be expected according to social networks towards a region.

Figure 4 displays our network indicator by region for the four geographical origins being the most numerous ones in France: North Africans, Africans, South Europeans and North Europeans. The comparison of the spatial distribution of new arrivals with respect to the spatial distribution of established immigrants from the same origin between 1993 and 1996 is carried out finally over two periods. This choice is motivated by the periods of expansion and recession observed during 2003-2012: the first period covers the economic expansion of years 2003-2007 and the second period the recession between 2008 and 2012.

We observe that for “small regions” the network effect is accentuated, i.e. positive value of the indicator. These regions (Normandie, Centre, Limousin) are characterized by a small population

Figure 4: Spatial Index, by origins and regions



size. In contrast, for most regions and for all origins the network indicator adopts values that are close to zero, suggesting the existence of a network effect relatively stable along time. Concerning Africans, we also note that, for two regions (Centre and Nord Pas de Calais), the indicator is very high over the period 2003-2007. This shows that newcomers have settled in these regions without benefiting from an existing network effect.⁵ During expansion periods some immigrants seem then be flowing towards some regions where they do not benefit from help from their country peers but they may be taking advantage of the economic opportunities in these regions.

Figures A.1 and B.1 represent in the X-axis each of the 21 regions, and in the Y-axis we find the 113 occupations denoted with the code provided by the Labor Force Survey (the code goes from 110 to 692). For the considered periods, panels provide information on which was the most popular job by education level and origin in each region. If a given origin has a comparative advantage in one occupation, we should observe horizontal lines, since in all regions (X-axis) individuals from the same origin should work in the occupation (Y-axis) for which they have the comparative advantage (which will be identical across regions). Figure A.1 focuses on North-Africans and Figure ?? in South-Europeans.

Figure A.1 reveals that the most popular job for a North-Africans varies across regions and across time, discarding the possibility that immigrants' allocation across local labor markets (regions) is guided by potential comparative advantages linked to the origin. This is the case regardless of the level of education. A similar conclusion can be drawn for South-Europeans (Figure B.1) even if, when considering individuals with an educational level equal or below the baccalaureate there are 14 regions for which South-Europeans have the same most popular occupation.

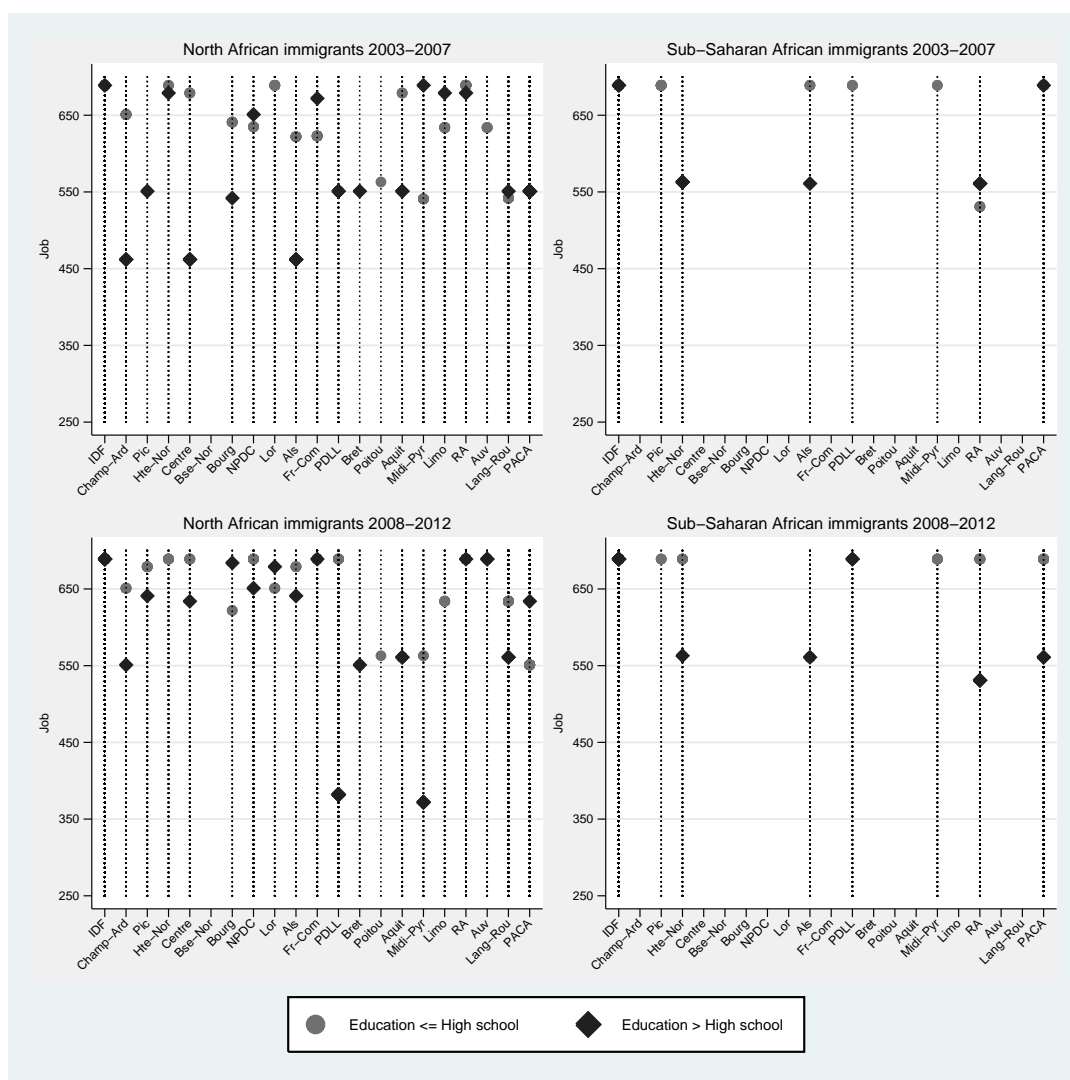
Appendix B proposes the same graphical analysis for African and North-European immigrants. Again, we find that immigrants from identical geographical origin distribute themselves differently across occupations in different regions.

We conclude therefore that a particular geographical origin cannot be associated to a particular comparative advantage in a precise occupation, since immigrants of identical origin specialize in divergent occupations depending on the region considered and regardless of the level of education. Other factors may be influencing immigrants' occupational choices. We focus here on one of these potential factors: social networks.

The main conclusions drawn from this section can be summarized as follows. First, while the

⁵Many African immigrants use Nord Pas de Calais as a stepladder for migrating towards UK, so our numbers may be biased by these immigrants.

Figure 5: Most popular regional occupations and the share in these occupations, for North-Africans and Africans, by education levels. French LFS 2003-2012.



proportion of immigrants over total population in France has not experience dramatic changes since 2001, its internal composition by country of origin has being modified. Second, the geographical distribution across French regions of newly arrived immigrants from a particular origin seems to have followed the spatial pattern of their already established country-peers. This, suggests that social networks influence geographical location choices of recently arrived immigrants in France. What about occupational choices? Our descriptive statistic analysis clearly underlines the fact that immigrants from the same origin and with the same education have divergent popular occupations across regions, suggesting that occupational choices are not driven by a comparative advantage linked to the origin. In the following sections, we seek to gain insights on the potential role of social networks as a determinant of occupational choices and how it varies across the economic cycle.

5 Econometric Analysis

As remarked in Section 3, we distinguish between “new” immigrants and “established” immigrants and estimate whether occupational choices of recently arrived immigrants are driven by the occupational distribution of “established” immigrants. Since we are working with data for the period 1993-2012, we consider as “established” immigrants those whose first year of residence in France is greater than 5 years before the survey. “New” immigrants are those arriving in France during the 5 years preceding the survey. Note though that, because of endogeneity issues, our instrumental variable strategy will exclusively focus on immigrants arriving in France after 2001, since we instrument the share of established immigrants in the occupation with the share estimated for 1993-1995 (between ten and 19 years before the arrival date of the immigrant).

We would like to estimate whether a new arrival chooses the dominant occupation for his immigrant group in the region in which he has arrived. We first define a dummy variable $JobI_{ogjt}$ adopting the unitary value in period t if the individual belonging to the nativity group o in region g is choosing the most popular occupation among individuals from his same nativity group in region g . We then estimate:

$$JobI_{iogjt} = \alpha + \gamma_1 S_{ogjt} + \gamma_2 S_{gjt}^N + \gamma_3 S_{gjt}^O + \eta_{iogjt} \quad (8)$$

where S_{ogjt} denotes occupational share of workers from geographical origin o in region g in occupation j and will measure the importance of the network effect. S_{gjt}^N stands for the share of natives

in occupation j in region g and S_{gjt}^O denotes the proportion of workers employed in occupation j region g .

We control for education, sex, age, age square and marital status. We include regional and origin indicators to control for systematic differences in the distribution of immigrants across regions as well as for systematic differences in the distribution of geographical origins. We also control for aggregate labor demand shocks by introducing year fixed effects, so that we can control for aggregate labor market reforms or other type of exogenous shock. Ideally, we should also introduce fixed effects for the 23 jobs we consider,⁶ so that to take into account the systematic differences in the distribution of immigrants across occupations, however introducing all these dummy variables will hugely increase the number of control variables creating estimation problems due to collinearity. Moreover, to take into account that shocks can be region-specific component we have also introduced the interacted region-year fixed effects.⁷

Because serial correlation within a particular labor market (region-job) is a concern, in all regressions we adjust standard errors for the clustering of observations at the region-occupation level. We have 21 regions and 23 jobs. We also use weighted regressions with weights equal to the immigrant population size in each occupation in the base year period 1993-1995. Using a fixed weight ensures that our results are not affected by changes in the immigrant population size across occupations. Despite our effort to control for unobservable determinants of immigrants' occupational choices, endogeneity biases still remain a concern. This is the case for instance if newly arrived immigrants' regional and occupational choices within a labor market are correlated with changes in unobserved determinants of employment that are also affecting the allocation of established immigrants. It is indeed plausible that a particular region may be going through an economic expansion that attracts both established and new immigrants or, within a region, immigrants would sort into occupations whose demand is growing (and this growing occupation may be different from one region to another). In that case, an endogeneity problem arises since location and occupational choices of new and established immigrants are driven by the same unobserved factors and not by social network effects.

We address this issue with three strategies. First, we control partly for labor-market occupations-

⁶To ensure a sufficient number of observations per job, we regroup the 113 initial occupations into 23 categories.

⁷It would have been interesting to introduce region-by-occupation fixed effects to control for the sorting of immigrants into labor markets whose structural determinants of employment are better but due to the large number of occupations considered here, strong collinearity problems arise. Similarly when introducing region-by-origin fixed effects together with the individual fixed effects and region-by-year fixed effects convergence problems arise.

specific productivity shocks using a demand shift index. If an occupation is concentrated in an industry whose employment has grown more than average over the period, we expect the labor demand for this occupation to have grown more than average and, at the same time, to have drawn more newly immigrants within that occupation. To control for this possibility, in our estimated equation, we introduce an occupation-specific labor demand shift index driven by the sectoral composition of natives' occupational employment at the national level : $(Demand\ shift^S)_{jt}$. We focus on natives because we believe their occupational choices are more exogenous with respect to choices of newly arrived immigrants than taking changes in the sectoral composition of established immigrants. Actually, the objective of the paper is to show that occupational choices of newly arrived immigrants are driven by occupational choices of established immigrants, so if we want to control for unobserved shocks that may be affecting both occupational choices, it seems more appropriate to focus on occupational changes of the majority group (natives) since they are likely to influence the occupational allocation of all immigrants (established and newly arrived). Thus, we achieve identification using deviations from occupation-specific trends determined by the initial sectoral composition of occupations. In the spirit of Katz and Murphy (1992) or Katz and Blanchard (1992), this labor demand shift index is constructed as follows:

$$(Demand\ shift^S)_{jt} = \sum_k \nu_{jk1993} L_{kt} \quad (9)$$

where L_{kt} is aggregate natives' employment at the two-digit industry level k at date t , and $\nu_{jk1993} = \frac{E_{jk1993}}{\sum_k E_{jk1993}}$ is the share of native workers in occupation j employed in industry k in 1993, excluding immigrants. We interpret the demand shift index as the predicted employment trend for native workers in an occupation given the distribution of that occupation across sectors. The implicit assumption being that the allocation of established immigrants across occupations is likely to be influenced by occupational shifts promoted by sectoral changes. We purge then from this sectoral effect and analyze the choice of newly arrived immigrants in deviation from this occupational trend resulting sectoral composition changes.

Second, we control for the fact that regions may be asymmetrically affected by shocks and this may promote an increased demand for certain types of occupations. For example, if a particular regional government decides to promote cultural activities, or R&D activities, or the automobile sector, we can expect that the labor demand for occupations linked to the cultural, R&D or the automobile sectors will increase more than average and will thus draw more workers (newly and established immigrants) within that occupation and that region. To control for this unobserved

determinants influencing the occupational choices of both newly and established immigrants, we introduce an occupation-specific labor demand shift index driven by the regional composition of natives' occupational employment:

$$(Demand\ shift^R)_{jt} = \sum_g \nu_{jg1993} L_{gt} \quad (10)$$

where L_{gt} is aggregate natives' employment at region g at date t , and $\nu_{jg1993} = \frac{E_{jg1993}}{\sum_g E_{jg1993}}$ is the share of native workers in occupation j employed in region g in 1993 over the total number of workers in occupation j , excluding immigrants. Again, we focus exclusively on natives. We interpret this index as the predicted trend in native employment in an occupation given the distribution of that occupation across regions and given the changes in regional employment. The implicit assumption being that the allocation of established immigrants across occupations is likely to have followed natives' occupational shifts promoted by regional changes.

The estimation of our coefficient of interest, γ_1 , will not be biased by the correlation between newly arrived immigrants' inflows into a particular labor market (region-occupation) and the better employment prospects due to the labor-market demand shocks driven by changes in the sectoral or regional composition.

Since region and occupation employment may change due to biased economic and technological shocks, our shift-share variable may not be enough. Our third approach to deal with endogeneity biases uses an instrumental variable strategy. This requires a variable correlated with occupational choices of established immigrants in a particular labor market but uncorrelated with unobserved factors driving occupational choices of newly arrived immigrants. We employ the strategy originally developed by Altonji and Card (1991) in a multi-region-occupation setting and use historical settlement patterns across both regions and occupations by geographical origin as an instrument for current occupational distribution of established immigrants. Because of informational networks, immigrants tend to cluster into labor markets with a higher share of their country peers.⁸ Such an instrument has proven to be a strong determinant of contemporaneous distribution of immigrants in the single-country case (see Card (2001), Card (2009), Cortes and Tessada (2011) or Patel and

⁸The importance of social networks for the location decisions of migrants coming from Mexico to the US and their labor market outcomes has been examined by Munshi (2003). Working with 1980, 1990 and 2000 US Census data, Patel and Vella (2013) find that the occupational share of certain ethnic groups grew drastically in particular labor markets over the period from 1980 to 2000. Moreover, the pattern of growth is consistent with the presence of network effects. The data also does not appear to suggest that the allocation observed is the result of sorting on the basis of a comparative advantage.

Vella (2013) for the US and Gonzalez and Ortega (2011) or Farré, González, and Ortega (2011) for Spain among others) and in a multi-country setting (see Angrist and Kugler (2003) and D’Amuri and Peri (2014)).

We focus on immigrants arriving in France after 2001, and we instrument the share of established immigrants in occupation j and region g using the past settlements of their peers. To reinforce the exogeneity of our instrument, we define it in two steps:

- First, we construct the predicted spatial distribution across regions of recently arrived immigrants, using the regional distribution of their peers in period 1993-1996:

$$S_{ogt} = \frac{S_{og0}}{S_{o0}} \cdot F_{ot} \quad (11)$$

where F_{ot} stands for the number of newly arrived immigrants of origin o in France in period t , for $t = 2001, \dots, 2012$, and $\frac{S_{og0}}{S_{o0}}$ represents the share of immigrants from origin o that were living in region g in period 1993 – 1996. We assume then that newly arrived immigrants during the period 2001-2012 are distributed across regions in the same way as their peers in 1993-1996. Economic shocks that could affect the immigrant distribution across regions in 1993-1996 are not likely to be persistent during 2001-2012, so the spatial distribution of new entrants during the period 2001-2012, S_{ogt} , is exogenous to the contemporaneous economic shocks.

- Second, we allocate the predicted stock of new entrants by region and origin, S_{ogt} , across occupations, according to the past allocation of their country peers:

$$S_{ogjt} = \frac{S_{ogj0}}{S_{og0}} \cdot S_{ogt} \quad (12)$$

where $\frac{S_{ogj0}}{S_{og0}}$ represents the share of immigrants from origin o that were living in region g and working in occupation j in period 1993 – 1996. The predicted number of recently arrived immigrants per region, S_{ogt} , is allocated across occupations within the region following the same pattern as their peers in 1993-1996. Again this allocation pattern is likely to be exogenous from contemporaneous economic conditions.

This instrument is highly correlated with the endogenous explanatory variable and has highly significant F-statistics in the corresponding first stage regressions. As the same objections may arise with respect to S_{gjt}^N and S_{gjt}^O we use the same instrumenting strategy for these two variables. As claimed in Patel and Vella (2013) these types of "lagged" instruments is frequently seen as controversial. This is not because of a "weak instrument" concern but rather because they are

seen as invalid. For example, one might argue that there are unobservables, or omitted variables, which are correlated with the regional or occupational distribution of migrants from a specific origin in a particular region in year t which are correlated with those in year $t - \tau$. Our identifying assumption is not that there are no such factors, but that they do not remain after we condition on the variables in our model. This assumption is crucial and can be justified by several reasons:

- First, this type of unobserved skill problem might arise when immigrants are locating in certain regions on the basis of their skills. However, more than half of the residence permits delivered by France correspond to immigrants obtain entry on the basis of family reunification, which means that their initial location decision is exogenous to occupational choices.
- Second, a problem of this type might arise if there were some specific unobserved skill which:
 1. was commonly found in immigrants from a certain country. If this was the case, we should have found that immigrants from a given origin should always be specialized in the same occupation across regions. As shown in section 4 this is not the case, the most popular job by geographical origin changes across regions. Moreover, we are also instrumenting regional allocation.
 2. was in high demand in certain regions of the country. Note though, that we are controlling by the occupational demand shift promoted by sectoral and regional composition effects. Therefore, we are analyzing location choices in deviation from this trend that would arise due to sectoral and regional changes.
- Finally the argument above requires that there are country of origin and region interaction effects which persist over time. Given that we are already controlling for the interaction between yearly shocks (20 years) and regions (21 regions) introducing additional interaction effects between the 10 geographical origins and the 21 regions would require estimating a huge number of parameters and this seems unreasonable. Accordingly, we attempt to somewhat capture the interaction between region-origin effects by including individual origin and regional dummies.

6 Results

In this section, we will use two estimation strategies to highlight the network effect. As a first step, we will use as a dependent variable the proportion of newcomers in the main occupation of

the social network for each of the regions. This cell-based approach allows us to focus on the role of the social network in the choice of occupation of newcomers not at the individual level, but at an aggregate level. The form of our estimation strategy is as follows:

$$S_{iogjt} = \alpha + \gamma_1 S_{ogjt} + \gamma_2 S_{gjt}^N + \gamma_3 S_{gjt}^O + \eta_{iogjt} \quad (13)$$

The variable S_{iogjt} is defined by region, origin, temporality of the migration (only the newcomers are retained here) and by occupation for the social network. The variable represents the proportion of newcomers of origin, region, for each occupation. The interest of this approach is to highlight the role of the social network in the occupational choice of newcomers for each occupation. Indeed, the influence of the social network in the choice of occupation may not be limited only to the main jobs but may also have an influence on the minority occupations of the network. As a result, this approach will allow us to capture all the information related to the influence of social networks on the labor market, especially for newcomers.

In a second step, we will use the individual database to construct our dependent variable described above. This variable is equal to 1 if the newcomer is in the main occupation of the social network in a given region; otherwise the variable is equal to 0. To do this, we have determined for each year, region and origin the main occupation of the social network.

Table 1 presents the estimation results when considering a cell approach, where our cells are defined at the region-occupation-origin level. We present estimates for periods 2001-2012, 2001-2007 and 2008-2012. This decomposition allows us to highlight the network effect along the economic cycle on the proportion of recent immigrants in the main occupation of the social network. The dependent variable is the proportion of newcomers in the most popular occupation of the social network in the region. Columns (4) - (9) introduce the labor demand shift indicators. We have also introduced the interacted fixed effect region-year. The last three columns in Table 1 present estimates using the instrumental variables method described above.

Table 1: Occupational choices of recently arrived immigrants - Aggregate level

Estimation Method	OLS						IV	
Period	2001-2012	2001-2007	2008-2012	2001-2012	2001-2007	2008-2012	2001-2012	2008-2012
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of immigrants from <i>o</i> in the occupation	0.266*** (0.0477)	0.275*** (0.0443)	0.242*** (0.0949)	0.266*** (0.0477)	0.275*** (0.0443)	0.242*** (0.0949)	-24.81** (9.759)	-407.2*** (87.06)
Share of natives in the occupation	11.38*** (2.820)	9.150*** (2.828)	10.41*** (3.699)	11.38*** (2.820)	9.150*** (2.828)	10.41*** (3.699)	614.7*** (181.4)	-3,917*** (985.0)
Share of workers in the occupation	-12.17*** (2.862)	-9.454*** (2.817)	-10.90*** (3.689)	-12.17*** (2.862)	-9.454*** (2.817)	-10.90*** (3.689)	-488.7*** (167.8)	5,932*** (1,415)
Demand Shift	NO	NO	NO	YES	YES	YES	YES	YES
Origin fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Region fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Region *Year fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	3,212	4,371	1,563	3,212	4,371	1,563	1,946	2,715
R-Squared	0.410	0.375	0.497	0.410	0.375	0.497		899

Note: units of observation are region-occupation, gj , in each year t . The dependent variable is the share of new immigrants arrived in region g in the most popular occupation between 2001 and 2012 (for columns (1), (4) and (7)), or between 2001 and 2007 (for columns (2), (5) and (8)) or between 2008 and 2012 (for columns (3), (6) and (9)) among their already established country peers in the region. The explanatory variable of interest is the share of established immigrants from origin o in region g occupation j among the whole set of immigrants from o in region g . The labor demand shift is based on sectoral and regional distribution of occupations with the weights of sectors based on national sectoral employment. Standard errors clustered at the region-occupation level are reported in parentheses. Regressions are weighted by the total number of immigrants in a region-occupation cell in 1993. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1 presents estimates by the OLS method and the instrumental variables method for the overall period, as well as for periods of economic growth (2001-2007) and economic recession (2008-2012). The results obtained by the OLS method show that the parameter representing the network effect is positive with or without demand shifts. The results obtained with the OLS method show that the parameter related to the network effect is between 0.242 and 0.275.

This means that a 1 percentage point increase in the proportion of immigrants in the main occupation increases the probability that a newcomer will have a job of about 0.25 percentage point. These results are similar for the period considered. Indeed, the results obtained for the parameter of the network effect are identical for each of these periods. As a result, the network effect is identical both in periods of growth and in periods of economic recession on the proportion of new arrivals in the occupation. However, the results obtained with the instrumental variables method are different. The parameter related to the network effect is negative. This highlights the problems of endogeneity with our model. The OLS method for this estimation is inconsistent.

We used the individuals database to make the estimates for the table 2. The dependent variable is set to 1 if the new entrant is in the main occupation of the social network for an origin, region and year; otherwise, the value of the dependent variable is equal to 0. As before, we have made a temporal decomposition of our estimates in order to highlight cycles of economic growth and recession.

Table 2: Occupational choices of recently arrived immigrants - Individual level

Estimation Method	OLS															IV	
	2001-2012	2001-2007	2008-2012	2001-2012	2001-2007	2008-2012	2001-2012	2001-2007	2008-2012	2001-2012	2001-2007	2008-2012	2001-2012	2001-2007	2008-2012	2001-2012	2008-2012
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)		
Share of immigrants from <i>o</i> in the occupation	1.815*** (0.0644)	1.633*** (0.0480)	2.120*** (0.134)	1.938*** (0.0728)	1.860*** (0.0589)	2.189*** (0.136)	1.800*** (0.0636)	1.621*** (0.0473)	2.097*** (0.134)	1.915*** (0.0722)	1.830*** (0.0580)	2.162*** (0.137)	4.359*** (0.297)	3.676*** (0.222)	4.382*** (0.383)		
Share of natives in the occupation	-10.48*** (1.281)	-10.000*** (1.008)	-11.27*** (1.902)	-10.91*** (1.347)	-10.63*** (1.074)	-11.66*** (2.024)	-13.18*** (1.429)	-13.29*** (1.182)	-12.93*** (2.167)	-14.04*** (1.520)	-14.02*** (1.214)	-13.96*** (2.319)	-7.215*** (2.155)	-7.983*** (1.630)	-9.113*** (2.881)		
Share of workers in the occupation	12.53*** (1.446)	11.75*** (1.220)	13.90*** (2.207)	13.03*** (1.518)	12.75*** (1.228)	14.34*** (2.352)	14.60*** (1.517)	14.25*** (1.282)	15.10*** (2.343)	15.47*** (1.609)	15.38*** (1.313)	16.09*** (2.522)	8.198*** (2.461)	8.760*** (1.848)	10.89*** (3.347)		
Demand Shift	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Origin fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Region fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Region *Year fixed effects	NO	NO	NO	YES	YES	YES	NO	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,662	3,686	1,298	2,662	3,686	1,298	2,662	3,686	1,298	2,662	3,686	1,298	2,648	3,671	1,291		
R-Squared	0.461	0.437	0.469	0.527	0.528	0.525	0.465	0.445	0.470	0.530	0.534	0.528	0.108	0.266	0.290		

Note: units of observation are region-occupation, gj , in each year t . The dependent variable adopts the unitary value if the immigrant arrived in region g between 2001 and 2012 (for columns (1), (4), (7), (10) and (13)), or between 2001 and 2007 (for columns (2), (5), (8), (11) and (14)) or between 2008 and 2012 (for columns (3), (6), (9), (10), (12) and (15)) has chosen the most popular occupation

among his already established country peers in the region. The explanatory variable of interest is the share of established immigrants from origin o in region g occupation j among the whole set of immigrants from o in region g . The labor demand shift is based on sectoral and regional distribution of occupations with the weights of sectors based on national sectoral employment. Standard errors clustered at the region-occupation level are reported in parentheses. Regressions are weighted by the total number of immigrants in a region-occupation cell in 1993. Statistical significance: *** $p < 0.01$,

** $p < 0.05$, * $p < 0.1$

The results obtained with the OLS method for the parameter related to the network effect is between 1.621 and 2.189. This means that a 1 percentage point increase in the proportion of immigrants in the main occupation for a region and a given origin, the probability that a new entrant will occupy the main job of the social network increases by 2% . Although the results are relatively similar for the overall period, the period of growth and the period of economic recession, we can show differences between these parameters. Indeed, for each estimation group (with or without demand shifts and / or with or without fixed effects), the parameter related to the network effect is systematically higher in a period of economic recession than in the period of economic recession. This confirms our initial assumption in relation to the business cycle. Indeed, in a period of economic recession, given the scarcity of job offers available, newcomers rely more on the social network in order to enter the labor market. In times of economic growth, the network effect is weaker because of the availability of more jobs available in the labor market.

The second coefficient estimated in our relation concerns the influence of the natives in the occupation considered. We note that the probability that a migrant occupies the main job of his social network decreases when we have an increase in the share of natives in this occupation. These results are close to our theoretical model. We have highlighted that information concerning an employment opportunity is initially transmitted to people of the same origin. The probability of transmission of this job opportunity decreases when individuals are not of the same origin. These elements are linked to the notion of homophily. Our estimations confirm the homophily shown in the theoretical model.

Finally, the third coefficient estimated is related to the variable concerning the proportion of workers in a given occupation in a region. The effect of this variable is positive on the probability that a new entrant will choose the main occupation of the social network. This variable, capturing the relative importance of the occupation in the region, has an expected result on the occupational choice of the new arrivals.

The instrumental variables method gives similar results. The parameter related to the network effect is positive (between 3.676 and 4.382). The influence of the economic cycle is also verified with this method. The estimate of the variable S_{ogjt} is higher in periods of economic recession (2008-2012) than in periods of economic growth (2001-2007). The signs of the parameters and the estimated coefficients related to the proportion of natives in the main occupation and the proportion of workers in this occupation in the region are identical with the OLS method. We may interpret that both the OLS and IV estimates are consistent

Concerning the table 3, we also used the individuals database to make our estimates. The Probit model was used here to make the estimates. In the last three columns, we control our estimates with demand shifts. A time division similar to previous tables related to periods of economic growth and recession has been achieved.

Table 3: Occupational choices of recently arrived immigrants.

Estimation Method		PROBIT					
Period		2001-2012	2001-2007	2008-2012	2001-2012	2001-2012	2001-2012
VARIABLES		(1)	(2)	(3)	(4)	(5)	(6)
Share of immigrants from o in the occupation		16.02*** (0.893)	12.69*** (0.586)	19.63*** (1.711)	16.01*** (0.922)	12.50*** (0.582)	19.60*** (1.744)
Share of natives in the occupation		-38.50*** (7.323)	-34.97*** (4.966)	-50.99*** (13.56)	-40.90*** (8.604)	-46.89*** (6.188)	-52.90*** (17.08)
Share of workers in the occupation		44.98*** (8.408)	40.71*** (5.644)	61.81*** (16.23)	47.10*** (9.268)	51.42*** (6.627)	63.53*** (19.04)
Demand Shift		NO	NO	NO	YES	YES	YES
Origin fixed effects		YES	YES	YES	YES	YES	YES
Region fixed effects		YES	YES	YES	YES	YES	YES
Time fixed effects		YES	YES	YES	YES	YES	YES
Observations		2,662	3,686	1,298	2,662	3,686	1,298

Note: units of observation are region-occupation, gj , in each year t . The dependent variable adopts the unitary value if the immigrant arrived in region g between 2001 and 2012 has chosen the most popular occupation among his already established country peers in the region. The explanatory variable of interest is the share of established immigrants from origin o in region g occupation j among the whole set of immigrants from o in region g . The labor demand shift is based on sectoral and regional distribution of occupations with the weights of sectors based on national sectoral employment.

Standard errors clustered at the region-occupation level are reported in parentheses. Regressions are weighted by the total number of immigrants in a region-occupation cell in 1993. Statistical significance: * * * $p < 0.01$, * * $p < 0.05$, * $p < 0.1$

The results confirm those obtained previously with the OLS method and the instrumental variables method. The increase in the proportion of immigrants in the main occupation in one region and in a given year will increase the probability that a new entrant will occupy the main job of the social network. The effect is the same with the proportion of workers in this occupation. An increase in workers in the main occupation of the social network will increase the probability that a new entrant will occupy that main job. However, this probability will decrease if there is an increase in the share of natives in the main occupation of the social network.

Estimates obtained on the basis of individual data confirm the predictions of our theoretical model. The positive influence of the social network as a determinant of recently arrived immigrants occupational choices is consistent with previous findings in the USA and Canada. Moreover, the social network has a stronger influence in the period of economic recession on the choice of occupation of the new arrivals than in period of economic growth.

7 Conclusion

One population subgroup of individuals for whom local interactions are particularly relevant is immigrants. While there is a large consensus in the literature concerning the validity of historical settlement patterns across both host countries and/or occupations by geographical origin as an instrument for immigrants' current inflows across occupations, there are few studies outside the US or Canada analyzing the role of social networks as a determinant of immigrants' occupational choices. Moreover, to our knowledge, the changing role along the economic cycle of social networks as a determinant of immigrants' choices has traditionally been absent from the literature.

The contribution of this paper is twofold. First, contrary to the majority of studies, this paper considers an European country, France. Because of its colonial past, the geographical origin of immigrants arriving in France clearly differs from that of immigrants arriving in the US, Canada, Australia or the UK. France is then a peculiar case. Second, as shown by the simple theoretical framework proposed in this paper, French labor market structural characteristics imply that social networks should play a more prominent role as occupational choice driver during recessions than during expansions. Our econometric analysis confirms this intuition. In rigid labor market economies, as France, unemployment dynamics is driven by job finding rates, rather than by destruction rates. Because social networks influence job finding rates but not destruction rates, their relative importance is improved during recessions.

Our analysis faces though some important limits that should be underlined. First, our definition of geographical origin is too large. It would be better to work by country, but we do not have this information. Moreover, since we are using survey data and not census data, a more detailed definition of the geographical origin (*i.e.* by country) will lead to ridiculous region-occupation-nativity size cells. Second, for the period 1991-2002, the first year of residence in France is missed for a huge amount of observations. We partly mitigate the problem in two ways. On the one hand, we proxy the first year of residence in France with the seniority in the job only when this seniority is below 1991. On the other hand, when using our instrument variable strategy to correct for endogeneity bias, we focus on the period 2002-2012.

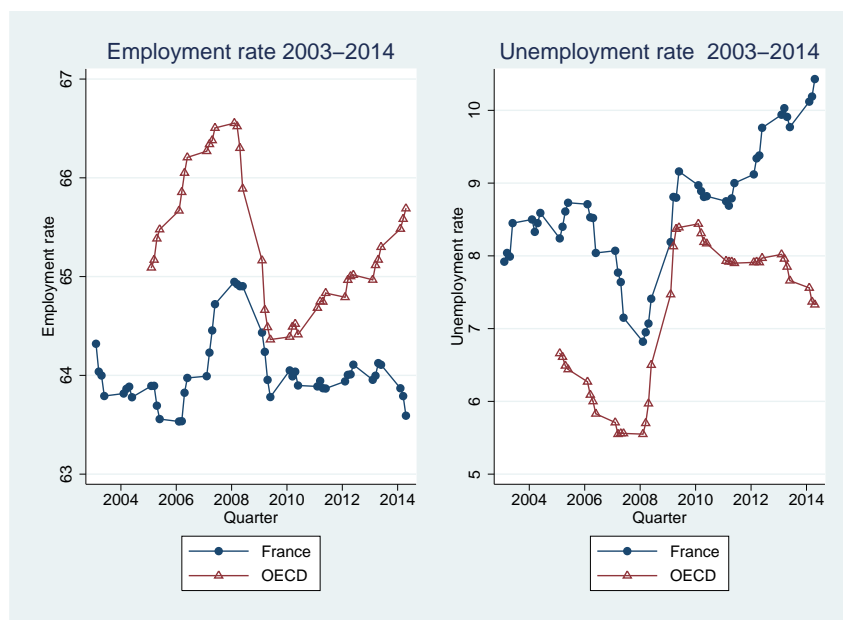
References

- Addison, J., and P. Portugal. 2002. "Job search methods and outcomes." *Oxford Economic Papers* 54 (3): 505–533.
- Altonji, J., and D. Card. 1991. *Immigration, Trade and Labor*. Edited by J. Abowd and R. Freeman. Chicago: University of Chicago Press.
- Angrist, J., and A. Kugler. 2003. "Productive or counter-productive: labor market institutions and the effect of immigration on EU natives." *Economic Journal* 113 (488): 302–331.
- Aslund, O., L. Hensvik, and O. Nordstrom Skans. 2009. "Seeking Similarity: How Immigrants and Natives Manage at the Labor Market." *IFAU Working Paper*, vol. 2009-24.
- Bartel, A.P. 1989. "Where do the new United States immigrants live ?" *Journal of Labour Economics* 7:371–391.
- Calvo-Armengol, A., T. Verdier, and Y. Zenou. 2007. "Strong and weak ties in employment and crime." *Journal of Public Economics* 91:203–233.
- Card, D. 2001. "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, vol. 19 (January).
- . 2009. "Immigration and Inequality." *American Economic Review* 99 (2): 1–21.
- Carrington, W.J., and K. Troske. 1998. "Interfirm Racial Segregation and the Black/White Wage Gap." *Journal of Labor Economics* 16 (2): 231–260.
- Cortes, P., and J. Tessada. 2011. "Low-Skilled Immigration and the Labor Supply of Highly Skilled Women." *American Economic Journal: Applied Economics* 3 (3): 88–123 (July).

- Cutler, D.M., E.L. Glaeser, and J.L. Vigdor. 2008. "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics* 90 (3): 478–497.
- D’Amuri, F., and G. Peri. 2014. "Immigration, Jobs and Labor Market Institutions: Evidence from Europe." *Journal of European Economic Association* 12 (2): 432–464.
- Dustmann, C., A. Glitz, and U. Schönberg. 2011. "Referral-based Job Search Networks." *CReAM WP*, vol. 14-11.
- Farré, L., L. González, and F. Ortega. 2011. "Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women." *The B.E. Journal of Economic Analysis & Policy* 11, no. 34.
- Gonzalez, L., and F. Ortega. 2011. "How Do Very Open Economies Absorb Large Immigration Flows? Recent Evidence from Spanish Regions." *Labour Economics* 18:57–70.
- Granovetter, M.S. 1995. *Getting a Job: a Study of Contacts and Careers*. 2nd. MA. Harvard University Press.
- Hairault, J.O, T. Le Barbanchon, and T. Sopraseuth. 2015. "The Cyclicalities of the Separation and Job Finding Rates in France." *European Economic Review* 76 (May): 60–84.
- Hellerstein, J., M. McInerney, and D. Neumark. 2007. *Changes in Workplace Segregation in the United States: Evidence from the 1990 and 2000 Decennial Employer-Employee Datasets*. The Analysis of Firms and Employees: Quantitative and Qualitative Approaches. Edited by Kathryn Shaw Fredrik Andersson Stefan Bender, Julia Lane and Till von Wachter. University of Chicago Press.
- Hellerstein, J., and D. Neumark. 2003. "Ethnicity, Language, and Workplace Segregation: Evidence from a New Matched Employer-Employee Data Set." *Annales d’Economie et Statistique* 71-72:19–78.
- . 2008. "Workplace Segregation in the United States: Race, Ethnicity, and Skill." *Review of Economics and Statistics* 90 (3): 459–477.
- Hemet, C. 2015. "Ethnic Networks and the Informal Labor Market." *Mimeo*.
- Holzer. 1988. "Search method use by unemployed youth." *Journal of Labor Economics* 6 (1): 1–20.
- Holzer, H. 1987. "Informal job search and black youth unemployment." *American Economic Review* 77 (3): 446–452.

- Katz, L., and O. Blanchard. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 1:1–75.
- Katz, L., and K. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107 ((February): 35–78.
- McKenzie, D., and H. Rapoport. 2010. "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks." *The Review of Economics and Statistics* 92 (4): 811–821.
- Montgomery, J. 1991. "Social networks and labor-market outcomes: Toward an economic analysis." *American Economic Review* 81 (5): 1408–18.
- Moreno-Galbis, E., and A. Tritah. 2016. "The Effects of Immigration in Frictional Labor Markets: Theory and Empirical Evidence from EU Countries." *European Economic Review*, p. forthcoming.
- Munshi, K. 2003. "Networks In The Modern Economy: Mexican Migrants In The U.S. Labor Market." *The Quarterly Journal of Economics* 118 (2): 549–599 (May).
- Musterd, S. 2005. "Social and Ethnic Segregation in Europe: Levels, Causes, and Effects." *Journal of Urban Affairs* 27 (3): 331–348.
- Ortega, J., and G. Verdugo. 2014. "Immigration and the Occupational Choice of Natives: A Factor Proportions Approach." *Labour Economics* 29:14–27.
- Patel, K., and F. Vella. 2013. "Immigrant Networks and Their Implications for Occupational Choice and Wages." *Review of Economics and Statistics* 95:1249–1277.
- Semyonov, M., and A. Glikman. 2009. "Ethnic Residential Segregation, Social Contacts, and Anti-Minority Attitudes in European Societies." *European Sociological Review* 25 (6): 693–708.
- Waldinger, R. 1994. "The Making of an Immigrant Niche." *International Migration Review* 28 (1): 3–30.
- . 1996. *Still the Promised City?* Boston University Press.
- Zenou, Y. 2015. "A Dynamic Model of Weak and Strong Ties in the Labor Market." *Journal of Labor Economics* 33 (4): 891–932.

Figure A.1: Employment and unemployment rates: France vs. OCDE average. 2003-2014. OECD Statistics



A Appendix: drivers of employment dynamics during the economic cycle

B Appendix Motivation

Figure B.1: Most popular regional occupations and the share in these occupations, for North-Europeans and South-Europeans, by education levels. French LFS 2003-2012.

