

Networks and Self-employment of Migrants

Joan Martín-Montaner, Universitat Jaume I and IEI

Guadalupe Serrano, Universitat de València

Francisco Requena, University of Sheffield

PRELIMINARY DRAFT. PLEASE DO NOT QUOTE

ABSTRACT

1. INTRODUCTION

The reasons and determinants behind an individual's decision of becoming self-employed have been analysed very deeply for a long time and have generated a huge literature in the labour economics field. However, only a small share of these studies focused on the particularities of individuals born abroad until recent years. To the best of our knowledge, Borjas (1986) was the first study of the self-employment in the case of migrants. Nowadays, there is an increasing interest in this issue, both from the academic and the political fields. In fact, the potential contribution of migrant entrepreneurs to the host-country's economic growth has drawn the attention of policy makers. For this reason, several OECD countries have introduced specific migration policies to support migrant entrepreneurs, both to those already established and to newcomers.

On average for the OECD countries, migrants experiment rates of self-employment very similar to those of natives. However, there are important differences among individual countries (OECD, 2011). If foreigners show behavioural patterns towards self-employment that deviate from that one of the natives, it is not unlikely that one of the reasons is related to the national composition of the foreign population living in that country and the degree of assimilation they have experienced. There is a branch of literature that deals with these issues by analysing to what extent individuals from the same country (or the same group of reference, language group, for instance) tend to concentrate in the same areas, and whether this concentration contribute to the appearance of networks.

In this paper, we focus on the Spanish case. We exploit the individual information obtained from the Active Population Survey for the period 1999 to 2009 to identify the existence of regional networks of migrants and the effect they could have on the probability to become self-employed. The paper is structured as follows: in section I, we quickly review the existing literature and establish the main hypothesis to be tested. In section II we describe our database and in Section III we present our main results. Some robustness analyses are displayed in Section IV. We finish with some preliminary conclusions in Section V.

2. LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

The determinants of the self-employment decision and entrepreneurship have been on the spotlight of economics and management research for a long time. However, it has not been until the last decade that a significant share of research on this field has focused on the peculiarities of migrants becoming entrepreneurs in the host country. See, among others, Lofstrom (2002, 2011), Hammarstedt (2004) or Constant and Zimmermann (2006).

Empirical research in the sociological field, however, had paid attention to this subject earlier. Light (1972) analysed the case of Chinese and Japanese migrants and postulated the disadvantaged theory: The relatively higher propensity of some immigrant groups to be self-employed reflects more limited options: first, the lack of language skills in the host country, labour skills not transferable to the host country, lack of documents to legally work, discrimination in the labour market. However, it does not explain why ethnic groups, who are considered disadvantaged, have very low self-employment rates. Light (1984) also argued the importance of ethnic and class resources, the former referred to network and financial support through ethnic institutions, both formal and informal. The latter includes human capital, attitude and information transferred among co-ethnics. Under this perspective, there could even exist a “culture” of self-employment transferred from the home country to the host one.

More recently, Fairlie and Meyer (1996) observe differences in self-employment rates across ethnic and racial groups, but they conclude that the disadvantage theory fails to explain them. Neither they are related to the self-employment rates in their countries of origin, rejecting therefore the “culture” argument. However, there is mixed evidence on this issue, as Yuengert (1995) had previously found a positive link. Also Foreman-Peck and Peng (2013) find the strength of entrepreneurial culture is different across ethnics.

Neither of the papers above takes into account the role of networks in the probability of immigrants becoming self-employed. These networks are more likely to appear because of the fact that some immigrants tend to cluster in certain locations (Chiswick and Miller, 2005). These enclaves could have a positive impact on the self-employment outcome if migrants have comparative advantage in the production of goods that are specific to people from a given country or language group, for instance. These advantages could appear because of the existence of economies of scale and/ or cheaper costs hiring people from the same groups of reference. But also self-employment could appear because linguistic concentrations are expected to have a negative effect on the destination language proficiency of immigrants and, therefore, worsen their performance in the labour market as employees. In fact, Lee (2000) had found a positive impact of enclaves only for linguistic groups, but not if they are defined by country of birth.

On the other hand, these ethnic enclaves could hamper self-employment for several reasons: first, previously established co-ethnics could block entry of potential immigrant entrepreneurs; second, some enclaves could be poor areas with low purchasing power (see Clark and Drinkwater, 2000).

The existence of ethnic enclaves and their relationship with the probability of becoming self-employed allows us to introduce the role of network, as they imply that enclaves do not mean just agglomeration, but also links and interaction. Social networks affect individual behaviour through two important channels: (1) an informational channel (how a person's knowledge depends on the behaviour of others) and (2) a normative channel (to what extent a person's preferences may depend on the behaviour of others)

Therefore, the existence of enclaves may increase the **quantity** of contacts with people belonging to the same groups, but it does not necessarily guarantee quality in the provision of information and knowledge. We need also to measure the "**quality**" of the contacts that characterise the existence of networks. Empirical research has not provided extensive evidence regarding the effect that cohesive social networks may have on the likelihood of migrants to become entrepreneurs. However, the scarce literature on this subject clearly concludes that migrants are more likely to become self-employed if networks operate. To the best of our knowledge, the only papers focusing on this subject are Andersson and Hammarstedt (2012) and Toussaint-Corneau (2012).

3. DATA DESCRIPTION

Our data come from the Spanish Labour Force Survey (EPA, *Encuesta de Población Activa*). This sample survey aimed to collect individual information in relation to labour market issues. The basic unit is the family house. Individuals fill the survey quarterly, providing not only information about personal issues such as their marital status, age and (very important) country of birth and nationality but also their labour status (being unemployed, self-employed or employee), their occupation, sector of activity and educational attainment. In the case of immigrants, it also includes the number of years since their arrival to Spain.

This means that those members of the family who happen to be at home fulfil the survey. In other words, respondents will not necessarily be the same each quarter, although the family house stays in the sample. Thus, our data are not of panel nature as we are not able of following individuals across time.

Data correspond to the years 1999 to 2009. The total number of answers by year is displayed in Table I. It is straightforward to observe that the number of individuals born abroad increases steadily across the years. Not being specifically aimed to generate statistical information on migrant population, the existence of strong immigration flows in Spain forced to adapt the survey to this evidence in two ways: first, by including a higher number of foreign-born respondents; secondly, being increasingly specific with regard to the country of birth of respondents.

Table I. Number of individuals

	Natives	Foreigners						
		Total	Europe	Africa	North America	South America	Asia	Oceania
1999	79790	1991	585	725	26	566	82	7
2000	75139	2099	652	692	23	643	82	7
2001	72777	2370	706	695	15	862	85	7
2002	70905	2901	819	752	28	1223	72	7
2003	72789	3655	1105	902	32	1525	83	8
2004	72977	4213	1252	982	21	1850	99	9
2005	64901	4516	1275	983	24	2111	117	6
2006	67791	5693	1608	1215	31	2709	125	5
2007	69671	6925	2085	1357	30	3274	172	7
2008	70495	7662	2302	1485	27	3633	208	7
2009	69609	7816	2238	1556	27	3745	242	8

The national composition of the foreign-born respondents reflects the relative size of migrant population in Spain. Thus, most of immigrants are originated in European and South American countries, being people from Africa the third group (mostly Moroccan). Although in 1999, the biggest groups correspond to people from African countries, this group does not experience such an intense growth as people from Europe and (mostly) from South America. Notwithstanding this evolution, the share of self-employed over total employment remains highly stable for most of the period, as displayed in Table ***. In the case of foreigners (as a whole), the share is quite close to that of native's in the first years, but it falls very quickly to 11%-12%. As we should expect, heterogeneity across country groups is high: migrants from North America and Asia are more prone to self-employment.

The key point is to what extent this behaviour responds to the ability to form networks. We will explore these issues in the following section.

Table II. Self-employed (share over total employment)

	Natives	Foreigners					
		Total	Europe	Africa	North America	South America	Asia
1999	0,20	0,17	0,17	0,13	0,05	0,20	0,38
2000	0,19	0,18	0,23	0,15	0,22	0,14	0,33
2001	0,19	0,18	0,23	0,15	0,29	0,13	0,31
2002	0,18	0,13	0,15	0,15	0,12	0,10	0,30
2003	0,18	0,13	0,15	0,14	0,30	0,09	0,38
2004	0,18	0,13	0,15	0,14	0,30	0,09	0,36
2005	0,18	0,12	0,12	0,14	0,13	0,09	0,33
2006	0,18	0,12	0,12	0,14	0,41	0,09	0,31
2007	0,18	0,11	0,12	0,13	0,31	0,09	0,31
2008	0,18	0,12	0,14	0,13	0,23	0,09	0,33
2009	0,18	0,11	0,11	0,15	0,16	0,08	0,30

4. ECONOMETRIC MODEL AND VARIABLES

In this section, we study the impact of the existence of networks on the likeliness of an immigrant to become self-employed. We focus on the probability of self-employment, thus a probit model is specified:

$$Pr(\text{self_employed}_{ijk} = 1) = \beta_0 \text{Network}_{jk} + \sum_h \beta_h X_i + \alpha_j + \gamma_k$$

Network_{jk} is a measure of the level and quality of the network ties of group k in the area j . A set of personal characteristics of individual i such as gender, marital status, age, years of residence in Spain and educational attainment are included in the vector X_i . Two effects are included for immigrant's area of residence (α_j) and region of origin (γ_j).

We use the indicator proposed by Bertrand, Luttmer and Mullainathan (2000) to measure networking. The indicator has two components: (1) the first term measures the concentration of people belonging to the same collective within an area relative to the size of that collective on the total population; (2) the second term approaches the "quality" of the network relations.

$$\text{Network}_{jk} = \left(\text{Density of group } k \text{ in area } j \right)_{jk} \times \left(\text{Knowledge and attitudes of others from group } k \text{ who live in area } k \right)_{jk}$$

$$\text{Network}_{jk} = \left(\frac{N_{jk}/N_k}{N_k/N} \right) \times (\text{self-employed}_k)$$

This indicator combines both cultural and geographical proximity to define networks. The density terms measures contact availability among people from the same group, whereas the second term proxies the attitudes of the group towards self-employment. We must define both (a) the k groups of reference and (b) the j areas where they locate. With regard to the latter, two alternative definitions of area of residence are considered in the analysis, corresponding to the NUTS regional classification: states (*Comunidades Autónomas*, NUTS2) and provinces inside the states (*Provincias*, NUTS3). With regard to the former, we consider different national or cultural groups that could generate the network. We consider immigrants tend to relate with those coming from the same country, sharing the same language. Just as a robustness check, we also define nine great regions associated to continents or sub-continents as the frame for networking. All this information is displayed in Table III.

Table III. Alternative definitions of areas and groups

Areas of residence		
NUTS2	<i>Comunidades Autónomas (ccaa)</i>	17 CC.AA. plus 2 autonomous cities
NUTS3	Provincias (prov):	50 provinces plus 2 autonomous cities
Group of reference		
Individual countries		123 countries + 7 aggrupation of "Other countries"
Language groups	8 groups	English, French, Spanish, Portuguese, Swedish, Russian, Arabian and Other
Great areas		Europe, Africa, North America, Central America and the Caribbean, South America, East Asia, Western Asia, Southeast Asia and Oceania

With regard to the individual characteristics, all of them have been obtained out of the EPA. A brief mention is due for the educational attainment measurement: the survey includes a highly disaggregated classification following UNESCO's International Standard Classification of Education (ISCED, 1997 version). All the items included have been aggregated up to four levels: no education (omitted), primary, secondary and tertiary education.

4. RESULTS

As we explained before, our data are not of panel nature. Thus, we estimate an individual equation for each year in the sample. The number of observations changes across years not only because more foreign-abroad respondents but also because more detailed information regarding the countries of origin is included. We built the network indicator for j =NUTS2 and k =individual countries. Following Bertrand, Luttmer and Mullainathan (2000), we include the contact availability term of the network indicator $\left(CA_{jk} = \frac{N_{jk}/N_k}{N_k/N} \right)$ as explanatory variable to control for omitted characteristics correlated with the density of the group in the area.

First, it is remarkable evidence that, despite the deep changes in the number of observations from the early years on, no great changes are observed in terms of sign or significance of the estimators for most of the indicators. This is mostly true for individual characteristics such as being *female* or the *age* and *years of residence*. Thus, we can observe a strong negative gender effect on self-employment. On the other hand, the probability of self-employment increases with the age of the migrant: the higher experience and higher financial resources could explain

Table IV. Dependent variable: P(self-employed =1)

Estimation method: Probit. *Network defined by country of origin*

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
CA	-2.195*** [0.721]	-3.187*** [1.103]	-1.185** [0.597]	-1.306** [0.560]	-4.056*** [0.897]	-2.603*** [0.736]	-4.957*** [0.829]	-3.154*** [0.697]	-2.785*** [0.465]	-3.614*** [0.483]	-2.870*** [0.467]
Network	12.07*** [2.398]	11.85*** [2.938]	7.289*** [1.606]	5.222*** [1.289]	10.92*** [1.714]	9.297*** [1.687]	14.96*** [1.920]	15.16*** [1.844]	12.57*** [1.266]	16.38*** [1.607]	10.28*** [1.197]
female	-0.224 [0.139]	-0.516*** [0.144]	-0.546*** [0.119]	-0.372*** [0.110]	-0.423*** [0.0978]	-0.571*** [0.0902]	-0.295*** [0.0792]	-0.526*** [0.0729]	-0.415*** [0.0470]	-0.373*** [0.0467]	-0.264*** [0.0478]
married	0.494*** [0.156]	0.250 [0.154]	0.008 [0.122]	0.126 [0.109]	0.140 [0.097]	0.196** [0.094]	0.210** [0.085]	0.124* [0.074]	0.204*** [0.050]	0.287*** [0.048]	0.126** [0.049]
age	0.023*** [0.006]	0.025*** [0.007]	0.029*** [0.006]	0.029*** [0.005]	0.025*** [0.004]	0.024*** [0.004]	0.012*** [0.004]	0.024*** [0.003]	0.020*** [0.002]	0.019*** [0.002]	0.019*** [0.002]
years_res	0.020** [0.009]	0.019** [0.008]	0.026*** [0.007]	0.031*** [0.007]	0.026*** [0.006]	0.030*** [0.006]	0.044*** [0.005]	0.038*** [0.005]	0.012*** [0.002]	0.012*** [0.001]	0.014*** [0.001]
edu_low	0.240 [0.233]	0.003 [0.274]	0.330 [0.308]	-0.171 [0.214]	-0.275 [0.174]	-0.163 [0.203]	0.110 [0.191]	0.012 [0.145]	0.103 [0.111]	0.126 [0.108]	0.069 [0.120]
edu_mid	0.323 [0.240]	0.0184 [0.292]	0.614* [0.314]	0.042 [0.224]	-0.113 [0.185]	0.088 [0.206]	0.324* [0.195]	0.178 [0.146]	0.260** [0.112]	0.200* [0.110]	0.271** [0.121]
edu_high	0.806*** [0.236]	0.522* [0.283]	0.873*** [0.314]	0.062 [0.225]	-0.015 [0.185]	0.149 [0.208]	0.400** [0.196]	0.285* [0.148]	0.329*** [0.113]	0.251** [0.111]	0.399*** [0.121]
Constant	-3.040*** [0.544]	-1.804*** [0.512]	-3.296*** [0.551]	-2.890*** [0.475]	-1.941*** [0.358]	-2.803*** [0.391]	-2.212*** [0.328]	-2.628*** [0.278]	-2.512*** [0.190]	-2.604*** [0.202]	-2.585*** [0.211]
N. observ.	782	841	1,173	1,610	2,298	2,747	3,055	4,060	6,924	7,661	7,796

this outcome. Analogously, more years of residence, which imply a higher knowledge of the host country, also has a positive impact on the self-employed status.

With regard to the outcome for the educational attainment, being primary educated do not seem to have a different impact that not being educated at all (our omitted category), whereas secondary education has a positive impact only in several years (most of them since 2005 on). Tertiary education displays a positive impact in most years. The size of the parameter decreases across time, although the value is higher than that of secondary education in all cases.

Focusing on our variable of interest, the *network* indicator is always positive and significant. Thus, the availability and quality of contacts increase the share of self-employed migrants. At the same time, the contact availability variable also achieves a significant effect in all years, although with the opposite sign. According to this outcome, higher density in a given area discourages people to become self-employed.

However, things are more complex than this because of the fact that the network indicator is constructed relying upon the CA_{jk} variable. Thus, we need to take into account both variables simultaneously in order to observe the whole effect on each individual's probability of self-employment. More specifically, we have that the estimates relating the enclave variable and the probability of being self-employed are:

$$\rho * CA_{jk} + \beta_0 * network_{jk} = \rho * CA_{jk} + \beta_0 (CA_{jk} * \overline{self - employed}_k)$$

$$CA_{jk} (\rho + \beta_0 * \overline{self - employed}_k)$$

Thus, despite the estimate for ρ having a negative sign, the effective effect of the group density will be positive or negative depending on $|\rho|$ being higher or smaller than $\beta_0 * \overline{self - employed}_k$. Focusing on the estimates displayed in Table ***, we can calculate the following threshold values for each year in our sample, defined as those values for which $\rho + \beta_0 * \overline{self - employed}_k = 0$.

For those groups of immigrants showing an average rate of self-employment higher than the threshold values, the contact availability in each NUTS3 area will have a positive impact on the probability of being self-employed. In other words, higher density helps migrants to work for themselves. Not surprisingly, this threshold value changes across time, which imply that this effect could not always be positive or negative for a given country in all cases. In fact, the number of countries for which the effect is mostly positive (more than 50% of the years included in our sample) is relatively small: Italy, the Netherlands, the United Kingdom, the United States, China and India.

Table V. Estimates of threshold values

Year	Estimated parameters	Threshold value for <i>self – employed_k</i>
1999	$\rho = -2,195$ $\beta_0 = 12,07$	0,182
2000	$\rho = -3,187$ $\beta_0 = 11,85$	0,269
2001	$\rho = -1,185$ $\beta_0 = 7,28$	0,163
2002	$\rho = -1,306$ $\beta_0 = 5,22$	0,250
2003	$\rho = -4,056$ $\beta_0 = 10,92$	0,371
2004	$\rho = -2,603$ $\beta_0 = 9,29$	0,280
2005	$\rho = -4,957$ $\beta_0 = 14,96$	0,331
2006	$\rho = -3,154$ $\beta_0 = 15,16$	0,208
2007	$\rho = -2,785$ $\beta_0 = 12,57$	0,222
2008	$\rho = -3,614$ $\beta_0 = 16,38$	0,221
2009	$\rho = -2,870$ $\beta_0 = 10,28$	0,279

4.1 Categories of self-employment

One important issue we have not taken into account yet is the pattern of self-employment followed by the different national groups. The role of networks could not be the same if self-employed rely on their own groups as a market for their activity, or whether self-employed are mostly single workers or hiring other people. The EPA allows us to distinguish two categories of self-employment: (1) Individuals who work in their own business employing other people and (2) individuals who work in their own business without employing other people. Let us label as *entrepreneurs* the former ones and *strict self-employed* the latter ones.

Thus, we can rewrite our econometric model:

$$Pr(\text{entrepreneur}_{ijk} = 1) = \beta_0 Ntwk_ent_{jk} + \sum_h \beta_h X_i + \alpha_j + \gamma_k$$

$$Pr(\text{strict self – employed}_{ijk} = 1) = \beta_0 Ntwk_sself_{jk} + \sum_h \beta_h X_i + \alpha_j + \gamma_k$$

We assume that knowledge and attitudes are different for both types of self-employment, so two different alternatives must be defined for the networking effect:

$$Ntwk_ent_{jk} = \left(\frac{N_{jk}/N_k}{N_k/N} \right) x(\overline{entrepreneur_k})$$

$$Ntwk_sself_{jk} = \left(\frac{N_{jk}/N_k}{N_k/N} \right) x(\overline{strict\ self - employed_k})$$

The results of estimating the model with the new dependent variables are displayed in Table VI. In this case, and for the sake of simplicity, we only show the outcome corresponding to four years, although all the results are available upon request.

Table VI. Dependent variable: strict self-employed or entrepreneurs

Estimation method: Probit. *Network defined by country of origin*

	Entrepreneurs				Self-employed			
	1999	2002	2005	2009	1999	2002	2005	2009
CA	-3.778** [1.514]	-2.575* [1.381]	-3.191*** [1.099]	-1.193** [0.503]	-1.089 [0.715]	-1.067* [0.556]	-4.097*** [0.894]	-3.072*** [0.542]
Network	14.76*** [4.582]	10.24*** [3.089]	17.58*** [3.682]	7.931*** [1.831]	11.07*** [2.961]	6.538*** [1.712]	15.01*** [2.469]	14.14*** [1.709]
female	-0.362* [0.205]	-0.211 [0.167]	-0.293** [0.128]	-0.253*** [0.071]	-0.134 [0.152]	-0.387*** [0.124]	-0.249*** [0.085]	-0.215*** [0.050]
married*	0.553** [0.242]	0.251 [0.184]	0.197 [0.138]	0.122 [0.075]	0.381** [0.171]	0.0605 [0.120]	0.174* [0.092]	0.109** [0.055]
age	0.0004 [0.009]	0.032*** [0.009]	0.015** [0.006]	0.012*** [0.003]	0.027*** [0.007]	0.024*** [0.006]	0.008* [0.004]	0.018*** [0.002]
years_res	0.045*** [0.014]	0.019* [0.011]	0.024*** [0.007]	0.018*** [0.00283]	0.006 [0.010]	0.027*** [0.008]	0.044*** [0.005]	0.009*** [0.002]
edu_low	0.0372 [0.376]	0.439 [0.481]	0.276 [0.293]	0.0567 [0.186]	0.228 [0.251]	-0.291 [0.227]	0.0525 [0.217]	0.0436 [0.132]
edu_med	0.451 [0.368]	0.241 [0.502]	0.101 [0.303]	0.0832 [0.189]	0.188 [0.261]	0.0569 [0.236]	0.368* [0.220]	0.284** [0.132]
edu_high	0.957*** [0.361]	0.515 [0.494]	0.299 [0.301]	0.263 [0.188]	0.475* [0.255]	-0.0313 [0.238]	0.399* [0.221]	0.363*** [0.133]
Constant	-2.906*** [0.817]	-3.327*** [0.699]	-3.171*** [0.591]	-3.044*** [0.333]	-3.117*** [0.589]	-2.674*** [0.490]	-2.130*** [0.354]	-2.594*** [0.232]

N. observ.	569	1,287	2,791	7,796	773	1,610	3,055	7,796
------------	-----	-------	-------	-------	-----	-------	-------	-------

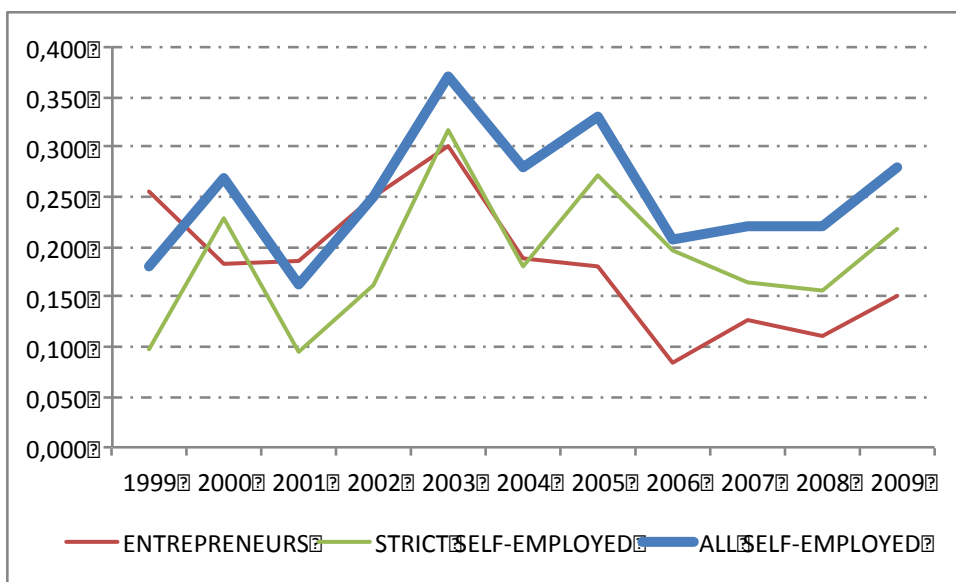
Personal characteristics play a similar role for both types of self-employment, although in this case some indicators are not always significant (the gender indicator, for instance). The educational attainment indicators also follow a pattern of significance close to that achieved before. In the case of the *CA* and *Network* indicators, we still face the same problem of interpretation of the estimates, thus we will replicate the exercise as in the previous section.

Table VII.

Year	Estimated parameters	Threshold value of $\overline{strict\ self - employed}_k$	Estimated parameters	Threshold value of $\overline{entrepreneurs}_k$
1999	$\rho = -1,089$ $\beta_0 = 11,07$	0,098	$\rho = -3,778$ $\beta_0 = 14,76$	0,256
2000	$\rho = -3,570$ $\beta_0 = 15,58$	0,229	$\rho = 2,541$ $\beta_0 = 13,85$	0,183
2001	$\rho = -1,404$ $\beta_0 = 14,81$	0,095	$\rho = 2,509$ $\beta_0 = 13,47$	0,186
2002	$\rho = -1,067$ $\beta_0 = 6,54$	0,163	$\rho = 2,575$ $\beta_0 = 10,24$	0,251
2003	$\rho = -3,734$ $\beta_0 = 11,80$	0,316	$\rho = -2,723$ $\beta_0 = 9,01$	0,302
2004	$\rho = -1,422$ $\beta_0 = 7,89$	0,180	$\rho = -7,222$ $\beta_0 = 38,17$	0,189
2005	$\rho = -4,097$ $\beta_0 = 15,01$	0,273	$\rho = -3,191$ $\beta_0 = 17,58$	0,182
2006	$\rho = -3,862$ $\beta_0 = 19,74$	0,196	$\rho = -1,396$ $\beta_0 = 16,70$	0,084
2007	$\rho = -2,545$ $\beta_0 = 15,46$	0,165	$\rho = -2,264$ $\beta_0 = 17,96$	0,126
2008	$\rho = -2,329$ $\beta_0 = 14,77$	0,158	$\rho = -2,060$ $\beta_0 = 18,66$	0,110
2009	$\rho = -3,072$ $\beta_0 = 14,14$	0,217	$\rho = -1,193$ $\beta_0 = 7,93$	0,150

One interesting point is that this threshold value shows a pattern across years that can be observed in the Figure I. The pattern shows how finding a positive impact of the agglomeration of natives from the same country in the probability of becoming self-employed required those population being increasingly prone to self-employment until 2003. In that year, more than one third of the population should be self-employed on average. In other words, those collectives where the average rate of self-employment is low will find that too much concentration in a given area will be counterproductive. From 2003 on, the threshold value presents a downward trend.

Figure I. Threshold values



4.2 Robustness checks

We should also check to which extent they are sensible to changes in the area or group of reference. In Table VIII a new set of estimates is displayed, depending on the group of reference being people born in the same country (as in Table III), in the same continent or sharing the same language. All three cases are analysed for the area of residence being NUTS3 (as in Table III) or NUTS2. For the sake of simplicity, we just present the results for our variables of interest, as the outcome for the other explanatory variables hardly changes across alternative estimates. Besides, only four years are displayed, as the pattern of estimates across years is quite similar to those previously shown. All the omitted results are available upon request.

Some patterns appear in our results. First, the sign of the estimates for both parameters remain regardless the level of aggregation used to construct the indicator, as the network indicator displays a positive sign whereas the enclave effect is negative. Secondly, the size of the estimated parameters for both indicators is smaller when the area of reference is NUTS2, whatever the definition of the group (same country of origin, region or language group). This outcome suggests that smaller areas help to create stronger links among individuals. Third, the size of the estimated parameters increases with the size of the group of reference (from country of birth to language group and to great regions/continents). This result also holds in all the years for which we have data. We use the Akaike criteria to test which regression fits

better our data. As we should expect, the better model is the one closer to the “neighbourhood” notion: the highest value is achieved for the NUTS3 by countries of origin estimation. Thus, we will keep the outcome in Table III as our preferred one.

Table VIII Dependent variable: P(self-employed =1)

Estimation method: Probit

		1999		2002		2005		2009	
		Parameter	Akaike IC	Parameter	Akaike IC	Parameter	Akaike IC	Parameter	Akaike IC
Variables calculated by country									
NUTS3	enclave	-2.195***	578.37	-1.306**	826.90	-4.957***	1387.42	-2.870***	3976.76
		[0.721]		[0.560]		[0.829]		[0.467]	
NUTS3	network	12.07***		5.222***		14.96***		10.28***	
		[2.398]		[1.289]		[1.920]		[1.197]	
NUTS2	enclave	-1.434**	584.39	-0.676	831.33	-2.862***	1399.37	-1.100***	4006.80
		[0.567]		[0.465]		[0.524]		[0.274]	
NUTS2	network	9.912***		4.227***		10.63***		7.013***	
		[2.022]		[1.152]		[1.441]		[0.948]	
Variables calculated by region									
NUTS3	enclave	-4.634*	594.31	-2.843	846.25	-8.398***	1431.40	-9.486***	3997.30
		[2.693]		[2.441]		[2.239]		[1.376]	
NUTS3	network	20.20***		13.45**		38.84***		43.27***	
		[5.053]		[6.146]		[6.654]		[5.021]	
NUTS2	enclave	1.280	601.02	-2.773	846.37	-3.429***	1441.55	-3.740***	4241.35
		[2.115]		[1.761]		[1.262]		[0.725]	
NUTS2	network	11.54***		12.10**		34.83***		33.67***	
		[4.142]		[5.300]		[7.057]		[4.650]	
Variables calculated by language group									
NUTS3	enclave	-2.182***	589.84	-1.418**	831.98	-4.850***	1402.14	-2.924***	4010.02
		[0.806]		[0.694]		[1.055]		[0.674]	
NUTS3	network	11.25***		4.836***		15.23***		9.808***	
		[2.594]		[1.380]		[2.297]		[1.433]	
NUTS2	enclave	-1.801***	592.27	-0.859	834.82	-3.110***	1404.71	-0.943***	4027.61
		[0.685]		[0.582]		[0.664]		[0.334]	
NUTS2	network	9.855***		3.985***		11.55***		6.432***	
		[2.248]		[1.211]		[1.650]		[1.037]	

5. CONCLUSIONS

We confirm some stylised facts regarding the probability of self-employment (of migrants). We observe that, first it is higher with older people and if individuals are married and, second, the number of years residing in the host country helps to increase the probability, the same as the educational attainment.

We find the existence of both an enclave effect and a network effect. Both effects are stronger as more homogeneous are the groups of reference, and the smaller the area of residence. Besides, higher population density in an area only reduce the probability of being self-employed, except in those cases in which migrants from the same group are themselves prone to self-employment. Finally, the “quality” of the link established among the individuals in the same group help to increase the probability of being self-employed. These results display a high degree of stability across years.

REFERENCES

- Belso-Martínez, J.A., (2005). Equilibrium entrepreneurship rate, economic development and growth. Evidence from Spanish regions. *Entrepreneurship & Regional Development*, 17 (2), 145-161.
- Bender, K.A. and K. Roche, (2013). Educational mismatch and self-employment. *Economics of Education Review*, 34, 85-95.
- Borjas, G.J., (1986). The Self-Employment Experience of Immigrants. *The Journal of Human Resources*, 21 (4), 485-506.
- Constant, A. and K.F. Zimmermann, (2006). The Making of Entrepreneurs in Germany: Are Native Men and Immigrant Alike? *Small Business Economics*, 26, 279-300.
- Evans, M.D.R. (1989). Immigrants Entrepreneurship: Effects of Ethnic Market Size and Isolated Labor Pool. *American Sociological Review*, 54 (6), 950-962.
- Fairlie, R.W. and B.D. Meyer (1996). Ethnic and Racial Self-Employment Differences and Possible Explanations. *The Journal of Human Resources*, 31 (4), 757-793.
- Faria, J.R., J.C. Cuestas and L.A. Gil-Alana (2009). Unemployment and entrepreneurship: A cyclical relation? *Economics Letters*, 105, 318-320.
- Foreman-Peck, J. and P. Zhou (2013). The strength and persistence of entrepreneurial cultures. *Journal of Evolutionary Economics*, 23, 163-187.
- Hammarstedt, M. (2004). Self-Employment Among Immigrants in Sweden – An Analysis of Intragroup Differences. *Small Business Economics*, 23, 115-126.
- Lardiés, R. (1999). Migration and Tourism Entrepreneurship: North European immigrants in Cataluña and Languedoc. *International Journal of Population Geography*, 5, 477-491.
- Levie, J. (2007). Immigration, In-migration, Ethnicity and Entrepreneurship in the United Kingdom. *Small Business Economics*, 28, 143-169.
- Light, I. (1984). Immigrant and ethnic enterprise in North America. *Ethnic and Racial Studies*, 7(2), 195-216.

- Liñán, F., D. Urbano and M. Guerrero (2011). Regional variations in entrepreneurial cognitions: Start-up intentions of university students in Spain. *Entrepreneurship & Regional Development: An International Journal*, 23, 187-215.
- Lofstrom, M. (2002). Labor market assimilation and the self-employment decision of immigrant entrepreneurs. *Journal of Population Economics*, 15 (1), 83-114.
- Lofstrom, M. (2011). Low-skilled immigrant entrepreneurship. *Review of Economics of the Household*, 9, 25-44.
- Portes, A. and J. Sensenbrenner (1993). Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *American Journal of Sociology*, 98 (6), 1320-1350.
- Portes, A., L.E. Guarnizo and W.J. Halier (2002). Transnational Entrepreneurs: An Alternative Form of Immigrant Economic Adaptation. *American Sociological Review*, 67 (2), 278-298.
- Toussaint-Comeau, M. (2012). Ethnic Social Networks and Self-Employment of Immigrant Men in the US, *Eastern Economic Journal*, **38**; 74-98.
- Yuengert, A. M (1995). Testing Hypothesis of Immigrant Self-Employment. *The Journal of Human Resources*, 30 (1), 194-204.