



Immigrants and entrepreneurship: A road for talent or just the only road?☆

Susana Iranzo *

Universitat Rovira i Virgili and ECO-SOS, Spain

ARTICLE INFO

JEL classification:

F22

J61

J71

Keywords:

Migrant entrepreneurship

P propensity to self-employment

Entrepreneurial quality

Negative sorting

Human capital losses

Information problems

Spain

ABSTRACT

The available evidence suggests that migrant entrepreneurs contribute to a country's growth and innovation, but entrepreneurship might also be chosen by less talented migrants who have limited chances in the labor market. This paper develops a theoretical framework that features specific constraints migrants face to better understand the mechanisms at play in their occupational choices. I test the model predictions using data for Spain right after the migration boom occurred from the mid-1990s to the mid-2000s. I find that human capital losses upon arrival are largely responsible for the low propensities to self-employment of the three largest migrant groups (Romanians, Moroccans and Ecuadorians) observed in the data. Also, I find no evidence of negative sorting into self-employment. Yet, the relatively large self-employment rate of Moroccans, once human capital is properly accounted for, and their low entrepreneurial quality is consistent with this group of migrants being subject to penalties in the labor market associated to information problems that push them into self-employment.

1. Introduction

Migrant entrepreneurship exhibits two facets. On the one hand, it can be a route for economic advancement and assimilation in the receiving country, with talented migrants contributing to the local economies. On the other hand, business ownership and self-employment might be the only path for migrants who stand few chances in the labor market. In effect, much of the documented productivity and innovation gains of migration -see for example Peri et al. (2015), Stuen et al. (2012) and Hunt (2011)- are channeled through entrepreneurship. In a survey of more than 1,500 engineering and technology companies founded between 1995 and 2005 in the US, Wadhwa et al. (2007) document that at least one key founder in 25% of those companies was foreign-born, with the vast majority of them being highly educated with degrees in STEM-related disciplines. Likewise, in a sample of US high-tech firms Hart and Acs (2011) find that 16% of them had a foreign-born owner with skills in science and engineering. Azoulay et al. (2022) find that migrants in the US are more entrepreneurial than natives suggesting that overall migrants contribute more to labor

demand expansion (job creation) than to labor supply expansion (job taking).

At the same time, it is also true that being new in a country, migrants endure specific problems that limit their possibilities in the labor market. Some of their skills are not transferable, and they might suffer additional penalties in the labor market due to information gaps regarding their ability and other attributes. As a result, they are likely candidates to enter 'necessity-driven' entrepreneurship. Cross-country data seem to confirm this conjecture. As the OECD data reproduced in Table 1 show, in many countries self-employment rates are higher among foreign-born than among natives. Furthermore, in countries like the US, migrant business owners are concentrated at the top and the bottom of the educational distribution (Fairlie, 2008). These stylized facts seem to suggest that entrepreneurship is chosen by the most talented migrants (positive sorting) but also by less skilled migrants who would struggle to find opportunities in the labor market (negative sorting).¹

☆ This project received financial support from MICIU /AEI/10.13039/501100011033/ by ERDF "A way of making Europe" project PID2022-137382NB-I00. I thank seminar participants at the Bank of Greece, Universitat d'Alacant, Universidad de Navarra, the IX Meeting on International Economics at Universitat Jaume I and EEA-ESEM 2018, for useful feedback on previous versions of this paper. I also thank co-editor David K. Nagy and two anonymous referees for very helpful comments and suggestions.

* Correspondence to: Department of Economics, Avda. Universitat 1, 43204 Reus, Spain

E-mail address: susana.iranzo@urv.cat.

¹ Negative sorting into entrepreneurship or self-employment is not exclusive to migrants. Evidence of a U-shaped relationship between education and entrepreneurship in the overall population has been documented in studies such as Poschke (2013) for the US, Astebro and Chen (2011) for Korea, Schjerning and LeMaire (2007) for Denmark or Blanchflower (2000) for other OECD countries. None of these papers focus on migrants, though some may include immigrants as part of their sample. As will be explained in Section 3, there are reasons why migrants might choose self-employment that do not necessarily apply to natives.

Table 1
Self-employment rates for natives and foreign-born for some OECD countries.
Source: OECD Settling in 2018. Indicators of Immigrants Integration

Country	Natives	Foreign-born
Australia	14.8	16.8
Austria	8.5	8.7
Belgium	12.2	14.1
Bulgaria	8.8	23.7
Canada	12.8	16.2
Croatia	8.0	12.9
Cyprus	11.7	9.7
Czech Republic	15.3	22.3
Denmark	6.8	7.8
Estonia	8.8	10.0
Finland	10.0	13.1
France	8.8	11.2
Germany	8.9	9.3
Greece	23.8	11.9
Hungary	8.6	10.0
Ireland	11.9	10.3
Italy	21.5	14.3
Latvia	9.2	10.5
Lithuania	8.2	11.1
Luxembourg	8.0	8.9
Malta	12.0	14.6
Netherlands	13.5	14.7
New Zealand	12.0	12.5
Norway	4.9	7.3
Poland	11.5	26.0
Portugal	11.2	13.5
Slovak Republic	14.7	22.2
Slovenia	10.4	9.5
Spain	14.6	15.3
Sweden	7.6	8.2
Switzerland	10.6	8.8
United Kingdom	12.7	16.7
United States	8.1	10.9

From a policy perspective, understanding the mechanisms that govern the occupational choice of migrants is important for at least two reasons. First, many countries have specific business visa programs or facilitate visa procedures for migrants who intend to set up businesses. If there is negative selection into migrant entrepreneurship, easing restrictions for potential entrepreneurs across the board might not be the most sensible immigration policy. Second, if migrants are pushed into self-employment due to problems in the labor market unrelated to their ability or productivity, such as lack of information or statistical discrimination, entrepreneurship would not be a route to assimilation but just a perpetuation of the lack of opportunities. Policies aimed at correcting such information problems would not only be justified on moral grounds but they would also lead to a more efficient allocation of skills across occupational choices.

Previous works have attempted to explain the factors that lead migrants to become entrepreneurs or self-employed, from a primarily empiricist approach. By and large, they have obtained mixed results, and we lack a clear framework to think about the occupational choice of migrants. In this paper, I develop a theoretical model that considers group-specific constraints that migrants face when choosing between entrepreneurship and salaried work. Negative sorting into entrepreneurship can occur due to severe penalties in the labor market, beyond mere human capital losses, associated with information problems.

I test the model predictions using data for Spain, a country that underwent a remarkable migration episode. Its foreign-born population increased by 8 times in about a decade, from the mid-1990s to the mid-2000s. I use survey data from 2006/2007, right after the migration boom and before the global financial crisis, which provides me with an ideal overview of the migration boom. My proxy for entrepreneurship is self-employment and, despite not being the same, I will use both terms interchangeably throughout the paper. A first glance at the

data on Spanish immigration could point to negative selection into self-employment because this occupational choice is relatively high among foreign-born individuals with college education and those with little schooling. However, given the wide heterogeneity in schooling quality across countries, education does not seem the best proxy to compare migrants' skills. Moreover, the aggregation of migrants masks important differences in self-employment across migrants of different origins. Using a proxy for skills based on a common metric for migrants' human capital, I find no evidence of negative sorting into self-employment. Across all the main migrant groups in Spain, self-employment is higher at higher human capital levels. Yet, the relatively large self-employment rate of Moroccans, once human capital is properly accounted for, and their low entrepreneurial quality suggest that this group of migrants is subject to information problems in the labor market that push them to self-employment.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 sets out the theoretical framework and conducts comparative statics that generate testable predictions regarding positive and negative sorting into entrepreneurship, the likelihood and the quality of entrepreneurship. These predictions will allow me to disentangle the main factors driving the choice between self-employment and salaried work across groups of migrants. Section 4 describes the data, including the construction of the human capital measure I use, and the choice for the reference group that will be used as *natives* in the analysis. Section 5 presents the empirical results and discusses some alternative explanations. Finally, Section 6 concludes.

2. Related literature

Migrant entrepreneurship is a relatively understudied issue within the burgeoning literature on migration that has emerged over the last two decades.² Yet, some interesting patterns have been documented. First, self-employment rates and firm creation are found to be higher among migrants than natives in several countries, including the US (Fairlie and Lofstrom, 2015; Kerr and Kerr, 2020; Azoulay et al., 2022), Denmark and Sweden (Andersson and Wadensjö, 2004), the UK (Clark and Drinkwater, 2000) and other OECD countries (OECD, 2018). Second, migrant entrepreneurs are, on average, more educated than their native counterparts (OECD, 2010), and in countries like the US, migrant business owners are concentrated at both the top and bottom tails of the educational distribution (Fairlie, 2008). Third, employment creation by migrant entrepreneurs appears to be lower than that of their native counterparts, and survival rates are also lower (OECD 2010, (Bahar et al., 2023)).³

Most of the existing works on migrant self-employment have attempted to empirically explain the different propensity for self-employment of migrants compared to natives. For instance, early work by Borjas (1986) assessed the role of geographical migrant enclaves in explaining the migrant-native self-employment differential in the US and found enclaves to be important for several Hispanic groups. Yuengert (1995) tested three explanations for the self-employment differential: (i) the home-country self-employment hypothesis, which posits that migrants from countries with relatively high self-employment rates have a preference or more adequate skills for self-employment; (ii) the tax avoidance hypothesis, which argues that migrants located in high-tax states tend to be self-employed in order to avoid taxes; and (iii) the enclave hypothesis. He found support

² Within that literature, the closest strand of papers are those examining the relationship between migrants and innovation activity – for instance, Hunt (2011), Hunt and Gauthier-Loiselle (2010), Kerr and Lincoln (2010), Peri et al. (2015) or Burchardi et al. (2020).

³ Bahar et al. (2023) document that although Venezuelan foreign-owned firms in Colombia are more capitalized than local firms when founded, they also have lower survival rates than Colombian-owned firms.

for the home-country and tax avoidance hypotheses but not for the enclave hypothesis. Using data from the Current Population Survey, [Uwaifo and Belton \(2012\)](#) found support for a modified version of the home-country self-employment hypothesis: migrants in the US coming from developed countries had a higher tendency towards self-employment than those from less developed countries. They explained this result by arguing that developed countries are closer, culturally and institutionally, to the US. [Lofstrom \(2002\)](#) found significant differences in self-employment rates across national migrant groups in the US after controlling for individual characteristics, years since migration and arrival cohort dummies. Furthermore, these differences remained largely unexplained after controlling for enclave effects and measures of success of self-employment in the geographical area.

Similar in spirit are the papers that try to explain the differences in the propensity for self-employment across ethnic groups. Among them, [Fairlie and Meyer \(1996\)](#) document large differences in self-employment rates across 60 ethnic groups that remain almost as large after controlling for individual demographic characteristics. They test different explanations for such differentials (the home-country self-employment hypothesis, the role of self-employed versus wage earnings, as well as several factors of advantage and disadvantage of ethnic groups on self-employment) and obtain mixed results. While each of the explanations finds some support in at least one ethnic group, all of the theories also have counter-examples. Similarly, [Fairlie \(2018\)](#) finds that some individual characteristics (such as wealth, education or age) are important to explain business ownership rates for some ethnic groups but do not play as big a role for others.

Outside the US, [Alden et al. \(2021\)](#) have documented differences between migrant and native self-employed in Sweden that persist over time, particularly with respect to capital income. For the UK, [Clark and Drinkwater \(2000\)](#) found important differences in self-employment rates among ethnic minorities (many of them immigrants) in England and Wales, but their empirical analysis did not reach any conclusive assessment of the relative importance of the factors considered. Likewise, [Clark et al. \(2017\)](#) consider foreign-born and UK natives and find that, after controlling for individual socioeconomic characteristics, significant differences in the propensity for self-employment remain across migrants of different origins. By contrast, [Constant and Zimmermann \(2006\)](#) do not find significant differences in the propensity for self-employment across ethnicities and with natives in Germany once individual characteristics are controlled for.

In sum, the existing literature so far is inconclusive on what explains the different propensities for self-employment across migrant groups, and we lack a clear understanding of how the specific constraints migrants face play out in their decision to choose self-employment over salaried work. I aim to fill this gap by developing a theoretical framework to help us think about the occupational choice of migrants. Building upon [Lucas \(1978\)](#) model of entrepreneurship, I incorporate the following key features affecting migrants' occupational choice in the host economy: (i) heterogeneous entrepreneurship entry costs, reflecting, for example, the difficulties to raise funds or poor understanding of local laws and procedures to start up a business; (ii) the loss of skills upon arrival or imperfect transferability of human capital; and (iii) additional difficulties in the labor market due to information deficiencies. Although natives are also subject to entrepreneurship entry costs and information gaps, as will be argued in Section 3, these problems are likely to be more severe in the case of migrants. The model shows how these factors, along with human capital losses, influence the propensity for entrepreneurship and average entrepreneurial quality of migrant groups. Importantly, the model can also explain negative sorting of migrants into entrepreneurship based on acute information deficiencies in the labor market.

3. The model

Consider [Lucas \(1978\)](#) entrepreneurship model. The economy consists of individuals who have different skill or efficiency levels, denoted z_i , drawn from a distribution function $\gamma(z)$ with support $[\underline{z}, \bar{z}]$, where $0 < \underline{z} < \bar{z} < \infty$, and cumulative distribution function $\Gamma(z)$. Entrepreneurial ability, a , depends positively on general skills z_i , and specific entrepreneurial attributes such as attitude towards risk, ambition, etc. that, although I do not explicitly model here, I will try to control for empirically. It is assumed that $a(0) = 0$, $a' > 0$, and to avoid unnecessary complications, we will also assume $a'' = 0$. Each firm is run by an entrepreneur and firms produce output using workers (n) and capital (K) according to the following production function:

$$a_i \cdot g[f(L, K)] \quad (1)$$

where $L = \sum_{j=1}^n z_j$ is total labor in efficiency units, f is a constant return to scale function and g is a concave transformation so that the production function exhibits diminishing returns to the variable inputs –Lucas' span of control. Notice that more capable entrepreneurs are those who are able to extract more output from the same inputs or, equivalently, those who can generate higher TFP. Given the constant return to scale assumption, the profits of an entrepreneur with ability a_i can be written as

$$\Pi = a_i \cdot g[L \cdot \phi(k)] - wL - rK \quad (2)$$

where $k = \frac{K}{L}$, w is the wage per labor efficiency unit and r is the capital rental.

From the first order conditions for profit maximization, one can obtain the two factor demand functions in terms of entrepreneurial ability, $L(a_i)$ and $K(a_i)$, which, back in the profit function, yield the optimal profits of an entrepreneur with skill z_i :

$$\Pi(z_i) = a(z_i) \cdot g[L(z_i) \cdot \phi(k)] - L(z_i)[w + rk] \quad (3)$$

Using the optimal input conditions and the earlier assumption $a''(z_i) = 0$, it can be shown that $\Pi'(z_i) > 0$ and $\Pi''(z_i) > 0$. That is, the profits of an entrepreneur are a convex function of her skills.

Alternatively, if an individual of skill z_i decides to go to the labor market and work as an employee, she will receive a wage proportional to her productivity

$$W(z_i) = w \cdot z_i \quad (4)$$

I consider entry costs to entrepreneurship that include, among others, regulation and red tape costs of starting up a business.⁴ Following [Guiso and Schivardi \(2011\)](#) and [Klapper et al. \(2006\)](#), I assume these costs, C , to be fixed. Thus, an individual of skill z_i will become an entrepreneur if her net profits are higher than the wage she would get as an employee working for somebody else:

$$\Pi(z_i) - C \geq w \cdot z_i \quad (5)$$

Given that $W(0) = \Pi(0) = 0$, the wage and entrepreneurial profit functions only cross once on the interval $[\underline{z}, \bar{z}]$. Let S be the skill level of the individual indifferent between becoming an entrepreneur or working as an employee, that obtains when expression (5) holds with an equality.⁵ As shown in [Fig. 1](#), individuals with skill level above S become entrepreneurs and those with skill level below S work as employees.

⁴ Just to get an idea of how pervasive entry costs might be, [Ciccone and Papaioannou \(2007\)](#) found that red tape procedures (in particular, the time necessary to set up a business) is significantly associated with less growth in industries facing expanding global demand and technology shifts. Specifically, their study revealed that industries in countries where it takes 62 days to register a new business saw approximately 0.4% lower growth in the number of establishments compared to countries where registration takes only 24 days.

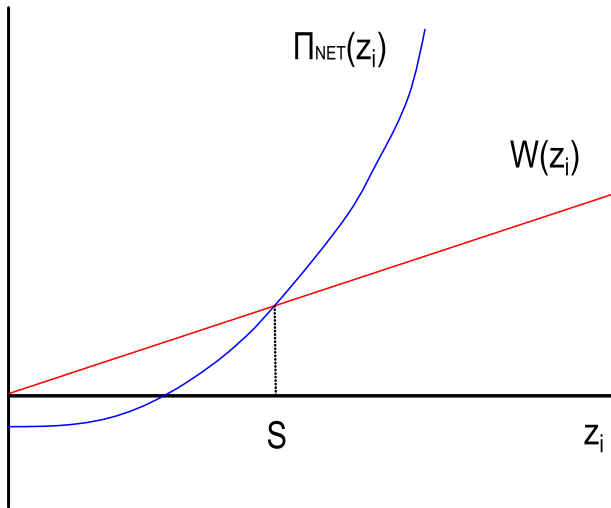


Fig. 1. Model equilibrium.

The share of employees and entrepreneurs in the total working population are respectively:

$$\Gamma(S) = \int_{\bar{z}}^S \gamma(z) dz \quad (6)$$

and

$$1 - \Gamma(S) = 1 - \int_{\bar{z}}^S \gamma(z) dz \quad (7)$$

On the other hand, entrepreneurial average skill is given by

$$E(z|z \geq S) = \frac{\int_S^{\bar{z}} z\gamma(z) dz}{1 - \Gamma(S)} \quad (8)$$

Finally, for the economy to be at equilibrium, w and r must be such that the market-clearing conditions for labor and capital are satisfied

$$1 - \Gamma(S) + \int_S^{\bar{z}} L(z)\gamma(z) dz = 1 \quad (9)$$

$$\int_S^{\bar{z}} k(z)\gamma(z) dz = \frac{\bar{K}}{L}$$

3.1. Immigrants and entrepreneurship

In the case of migrants, expression (5) needs to be modified to account for the specific issues they face when choosing between entrepreneurship and salaried work. First, the entry cost into entrepreneurship might be particularly burdensome to migrants due to lack of information and unfamiliarity with the business culture of the host country. Moreover, in the presence of financial constraints, migrants might find it harder than natives to access credit as they tend to have fewer assets to use as collateral and lack credit histories. Within the Lucas (1978) model, financial constraints are generally incorporated by means of a borrowing constraint in the entrepreneur's profit maximization problem that sets a limit (usually a fraction of their collateral wealth) on the amount the entrepreneur can borrow -see, for example,

⁵ Because abilities other than general skills affect entrepreneurial ability, the net profit function might be higher for some individuals and lower for others and so in practice we observe entrepreneurs along the whole skill distribution. Yet, as long as specific entrepreneurial abilities do not correlate negatively with general skills, condition (5) is still valid to identify the average skill level of the marginal entrepreneur.

Evans and Jovanovic (1989), Cagetti and De Nardi (2006) and Midrigan and Xu (2014). As a result, entry into entrepreneurship is reduced and/or some entrepreneurs are “constrained entrepreneurs”, that is, they invest below optimal levels.⁶ Here, I take a much simpler modeling approach. I consider that, if any, borrowing constraints are embodied in the entrepreneurial entry cost, and I allow for this cost to differ across migrants of different origin.

Second, as is well known, transferability of skills across countries is not perfect. An important skill that is country-specific and plays a key role in the performance of migrants in the host country is language knowledge –Adsera and Pytlikova (2016) and Chiswick and Miller (2015). Formal education and experience are not fully portable across countries either. Many professional credentials (e.g. physicians, lawyers, architects) are not recognized across countries but need to be validated or awarded through the host country's educational system. Likewise, firm training and work experience might not be valued in the same way as in the origin country. There is considerable evidence of downgrading, as reflected in earnings, of migrants compared to natives of similar education and experience levels. For example, Adsera and Chiswick (2007) find that immigrants in Western European countries earn 40% less than natives upon arrival, although the estimated wage gap varies significantly across host and origin regions, and also by gender. Dustmann et al. (2013) comparison of the wages received by migrants in the UK and the estimated wages that similar natives would receive reveal that, on average, migrants earn only 89% of what their native counterparts are paid. I will thus assume that a percentage of a migrant's skills, δ , with $0 < \delta < 1$, is “lost” upon arrival in the receiving country and that this percentage differs by migrant origin.

Beyond human capital losses, migrants might encounter additional difficulties to be hired in the labor market arising from information gaps. It is hard for potential employers to elicit workers' true productivity and other attributes such as job commitment or reliability. This is a pervasive problem in labor markets that applies to all workers, not only migrants, and is often solved by referral-based hiring -see, for example, Ioannides and Loury (2004) for a survey on this. As Munshi (2003) points out, though, we would expect this problem to be stronger in the case of migrants as, by definition, they are newcomers in the labor market and thus more susceptible to the information deficiencies that generate a need for job referrals. The available evidence confirms this. Dustmann et al. (2016) document that migrants in Germany obtain their jobs (particularly their first job upon arrival) through referrals in larger proportions than native workers, and that most referrals come from other migrants from the same country of origin. In the presence of information problems that cannot be resolved by referrals, migrants would need to accept wages below their productivity level in order to be hired, or it might altogether be impossible for them to be employed.⁷ Information gaps might also lead to statistical discrimination, namely, individuals or particular groups being treated according to the average characteristics of the group they belong to, rather than their individual qualities. The average characteristics inferred to the group will be based on employers' knowledge and previous experience with migrants from that particular origin which, in some cases, can be rather limited. Hence, I will allow for the severity of these information gaps to differ by migrant origin and will model them as a wage cut D .

Consequently, for migrants of origin m , the equation that yields the skill level, S_m , of the marginal entrepreneur is

$$\Pi((1 - \delta_m)S_m) - C_m = w \cdot (1 - \delta_m)S_m - D_m \quad (10)$$

⁶ For instance, in their simulated model for the US, Cagetti and De Nardi (2006) obtain that only high entrepreneurial ability individuals with wealth above \$250,000 become entrepreneurs.

⁷ Dustmann et al. (2016) estimates imply that referrals raise wages in a worker's first year at a firm by 2.6% and in subsequent years by 0.4% which gives an idea of the importance of information deficiencies in the labor market and the costs they imply in terms of wage reduction to workers who cannot resort to social connections.

where C_m are entrepreneurship entry costs that group m faces, δ_m is the percentage of human capital losses and D_m is the cost, in terms of a wage cut, of information deficiencies in the labor market.

3.2. Partial effects on the propensity for entrepreneurship and average entrepreneurial skill

Next I perform comparative statistics on expressions (10), (7) and (8) to gauge the impact of the three factors above (entrepreneurial entry costs, human capital losses and information problems in the labor market) on the propensity for entrepreneurship and average entrepreneurship skill.⁸ These partial effects will help me disentangle, in the empirical exercise, the most salient drivers into self-employment for each migrant group.

3.2.1. Effect of entrepreneurship entry costs

In order to isolate the impact of heterogeneous entrepreneurship entry costs, C_m , I assume no human capital losses (*i.e.*, $\delta_m = 0$) and no information problems in the labor market ($D_m = 0$) and differentiate (7) and (10) with respect to C_m . This yields

$$\frac{dS_m}{dC_m} = \frac{1}{\Pi'(S_m) - w} > 0 \quad (11)$$

and

$$\frac{d[1 - \Gamma(S_m)]}{dC_m} = -\gamma(S_m) \frac{1}{\Pi'(S_m) - w} < 0 \quad (12)$$

In words, the skill cutoff of the marginal entrepreneur will be higher and the probability of becoming entrepreneur lower for those groups of migrants with relatively higher startup entrepreneurship costs. Panel A on Fig. 2 offers a graphical illustration of this effect. The intuition is straight-forward. As higher entrepreneurship entry costs reduce the profit function while they do not affect salaried income, it takes a higher skill to make entrepreneurship worthwhile. Consequently, average entrepreneurial skill increases as differentiation of Eq. (8) shows:

$$\frac{dE(z|z \geq S_m)}{dC_m} = \frac{[E(z|z \geq S_m) - S_m]\gamma(S_m)}{1 - \Gamma(S_m)} \frac{dS_m}{dC_m} > 0 \quad (13)$$

3.2.2. Effect of human capital losses

In the absence of migrant-specific entrepreneurship entry costs (*i.e.*, $C_m = C$) and information problems in the labor market ($D_m = 0$), differentiating (7) and (10) with respect to δ_m , we obtain

$$\frac{dS_m}{d\delta_m} = \frac{S_m}{1 - \delta_m} > 0 \quad (14)$$

and

$$\frac{d[1 - \Gamma(S_m)]}{d\delta_m} = -\gamma(S_m) \frac{S_m}{1 - \delta_m} < 0 \quad (15)$$

That is, the skill cutoff of the marginal entrepreneur will be higher and the probability of becoming an entrepreneur lower for those groups of migrants suffering higher human capital losses. This effect is illustrated on panel B of Fig. 2. Notice that as $W[(1 - \delta_m)0] = \Pi[(1 - \delta_m)0] = 0$, the wage and entrepreneur net profit functions still only cross once on the skill range $[\underline{z}, \bar{z}]$ and the effect of δ_m is to tilt both curves causing

⁸ Notice that the analysis is partial equilibrium in that I abstract from any effect of migration on the local wage, w , and interest rate, r . Put it differently, the comparative statics should be understood as the differential impact on the skill level of the marginal entrepreneur at the new, if any, equilibrium wage and interest rate.

them to intersect at a higher skill level. Therefore, if the skill level of the marginal entrepreneur is S_0 for $\delta_m = 0$, the new equilibrium skill threshold for $\delta_m > 0$ is $S_1 > S_0$ such that $W(S_0) = W[(1 - \delta_m)S_1] = \Pi[(1 - \delta_m)S_1] = \Pi(S_0)$.

This result is not qualitatively different from the one obtained previously for higher entrepreneurship entry costs. Unlike the partial effect of entrepreneurship entry costs, though, there is an ambiguous effect of δ_m on average entrepreneurial ability as the derivative of Eq. (8) with respect to human capital losses (δ_m) yields

$$\begin{aligned} \frac{dE[(1 - \delta_m)z|z \geq S_m]}{d\delta_m} &= -E(z|z \geq S_m) + (1 - \delta_m) \frac{[E(z|z \geq S_m) - S_m]\gamma(S_m)}{1 - \Gamma(S_m)} \frac{dS_m}{d\delta_m} \\ &= -E(z|z \geq S_m) + \frac{[E(z|z \geq S_m) - S_m]\gamma(S_m)}{1 - \Gamma(S_m)} S_m \end{aligned} \quad (16)$$

The first term of (16) is the effect of human capital losses on average entrepreneurial skill keeping the skill cutoff S_m fixed. This is negative because skills are reduced by a percentage δ_m . The second term of (16) is the effect of human capital losses on the skill cutoff of the marginal entrepreneur, S_m . This term is positive because $E(z|z \geq S_m) - S_m > 0$, *i.e.*, the skill level of the marginal entrepreneur is lower than the average entrepreneurial skill. Whether the negative or the positive effect in (16) dominates depends on the particular skill distribution function, and so no unambiguous conclusion can be drawn on the effect of higher human capital losses on average entrepreneurial skill.⁹

3.2.3. Effect of penalties in the labor market associated to information deficiencies

Finally, I assume $\delta_m = 0$ and $C_m = C$ and consider $D_m > 0$ for at least some migrant group. The net wage function of that group shifts down on a parallel fashion so that the skill level of the marginal entrepreneur is now lower. Moreover, if D_m is sufficiently high, it is then possible for the wage and the entrepreneurial profit functions to also intersect at a lower skill level resulting in negative sorting into entrepreneurship. Thus we need to distinguish two cases:

- 1 $D_m > 0$ is such that the net wage and the entrepreneurial profits still cross only once.

Differentiating (7) and (10) we obtain

$$\frac{dS_m}{dD_m} = \frac{1}{w - \Pi'(S_m)} < 0 \quad (17)$$

and

$$\frac{d[1 - \Gamma(S_m)]}{dD_m} = -\gamma(S_m) \frac{1}{w - \Pi'(S_m)} > 0 \quad (18)$$

where the inequality in (17) follows from the fact that $\Pi'(S_m) > w$, *i.e.*, the slope of the profit function is larger than that of the wage function at $z = S_m$.

As for the effect on average entrepreneurial skill, we have

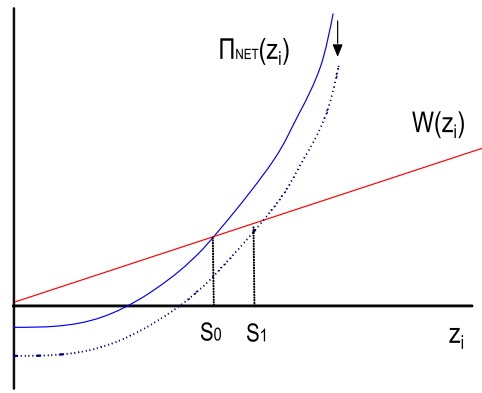
$$\frac{dE(z|z \geq S_m)}{dD_m} = \frac{[E(z|z \geq S_m) - S_m]\gamma(S_m)}{1 - \Gamma(S_m)} \frac{dS_m}{dD_m} < 0 \quad (19)$$

where the inequality follows from the fact that $E(z|z \geq S_m) - S_m > 0$ and $\frac{dS_m}{dD_m} < 0$.

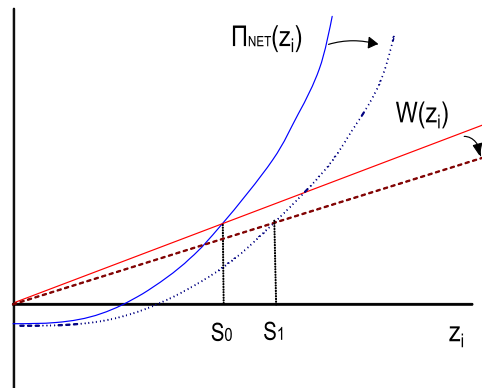
In words, provided the wage and entrepreneurial profit functions only intersect once (positive sorting), the skill cutoff of the marginal entrepreneur will be lower and the probability of becoming an entrepreneur higher for those groups of migrants suffering more severe

⁹ As an example, in the Appendix, I show results for two distribution functions: a simple uniform distribution for which the overall effect on average entrepreneurial skill is negative and a more realistic distribution, a Pareto distribution function, for which there is a null overall effect on average entrepreneurial skill.

Panel A. Effect on S of higher entry costs



Panel B. Effect on S of higher human capital losses



Panel C. Effect on S of higher information problems in the labor market

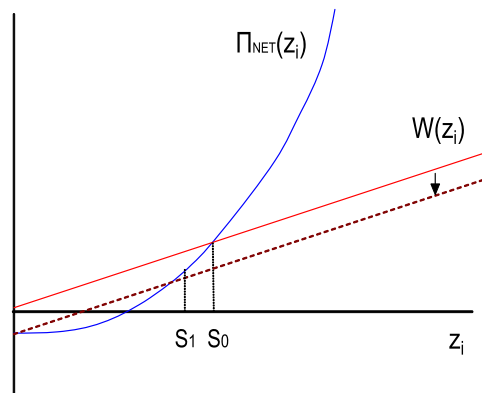


Fig. 2. Comparative statics - Effect on the skill level S of the marginal entrepreneur.

information problems (D_m) in the labor market, whereas average entrepreneurial skill will be lower. This case is illustrated on Panel C of Fig. 2.

$2 D_m > 0$ is sufficiently large so that the net wage and entrepreneurial profit functions cross twice.

Using as reference the graph of Fig. 2 Panel C, it is easy to see that if the net wage function shifts down sufficiently, there will then be skill levels S_1 and S_2 , with $S_1 < S_2$, such that $\Pi_{NET}(S_1) = W_{NET}(S_1)$ and $\Pi_{NET}(S_2) = W_{NET}(S_2)$, so that individuals with $z_i > S_2$ will become entrepreneurs (positive sorting) but individuals with $\underline{z} \leq z_i < S_1$

will also be drawn into entrepreneurship (negative sorting).¹⁰ Trivially, in this case the share of entrepreneurs increases and average entrepreneurial skill decreases even further than in the case of positive sorting only.

This simple model has two main virtues. First, it can explain negative sorting of migrants into entrepreneurship as a consequence of severe information problems in the labor market that result in wage penalties beyond those arising from human capital losses. Second, it

¹⁰ A necessary and sufficient condition to obtain negative sorting is $D_m > w \cdot \underline{z} - \Pi(\underline{z}) + C$.

generates testable predictions regarding the likelihood of entrepreneurship and entrepreneurial quality (or average entrepreneurial skill) that help tease out the dominant factors driving the occupational choice of migrants from different origins. I test the model predictions on the dataset of immigrants in Spain that I describe next.

4. The data

I use the *Encuesta Nacional de Inmigrantes 2007 (ENI)*, a rich survey on immigrants conducted by the Spanish National Statistical Institute between November 2006 and February 2007. Spain has been the subject of several studies on migration because it offers a remarkable migration episode for analysis.¹¹ The foreign-born population in Spain increased by 8 times in a little over a decade; from a mere 1.4% of the overall population in 1995 to a share of 11.4% in 2007. Thus, the ENI survey captures the migration boom while it is prior to the global financial crisis that suddenly worsened economic conditions, impacting on migrants decisions as well.

The survey contains information on a large number of variables for about 15,500 immigrants, defined as individuals born abroad, who at the time of the interview were at least 16 years old and had resided in Spain for a year or longer, or had the intention to remain in Spain for at least a year. The survey excluded those immigrants born abroad to Spanish parents and those who were less than 2 years old when they arrived in Spain. I also eliminated observations of migrants over 65 years old.

To avoid the problems derived from grouping migrants into arbitrary geographical regions, I will look at individual sending countries. In particular, I consider the three main countries of origin of migrants in Spain at the time: Romania, Morocco and Ecuador that accounted respectively for 8.2%, 12.5% and 9.4% of all foreign-born in Spain in 2006, and migrants from some European countries that will serve as reference group. These four groups make a reasonably large and balanced (in size) set of countries. All together, they amount to 6,154 observations in the ENI survey which represent almost half (43%) of the total sample. My analysis will focus on migrants who were working the week prior to the survey (about 4,000 observations) with the sample size being slightly reduced in the econometric analysis as the models are estimated using the observations for which data on all variables are available.

Although it is a cross-section, the dataset enables me to use a rich set of controls to circumvent some of the limitations of cross-sectional data. The survey includes standard demographic characteristics such as age, gender and marital status, the year of arrival and country of origin, and vast information on their working and living conditions in Spain, as well as some information on the employment experience in their country of origin. I know whether migrants are salaried workers or self-employed (my proxy for entrepreneurship), the sector where they work as well as the previous occupational status in their country of origin. Earnings in the current job are reported for 79% of the migrants working while for the remaining 21% this information is reported in earnings intervals. For the latter, I use the midpoint of the earnings interval.

I also know the educational level of migrants, and whether this was obtained abroad or in Spain. Importantly, there is also information on how well they speak Spanish (and the specific languages spoken in some of the regions in Spain). Based on this information, I can construct a variable of language knowledge that takes on three values: 0 if it is poor, 1 if it is reported as good and 2 if it is very good or the mother tongue. Other important information that will be used includes the assets migrants have (for instance, whether they own or rent the dwelling they live in and whether they have invested in other assets),

¹¹ See, among others, Amuedo-Dorantes and De la Rica (2007), Peri and Requena (2010) and Bosch et al. (2010).

their legal status and the year they obtained it, and the existence or not of social ties when they arrived in Spain. The legal status is reflected by a dichotomous variable constructed based on the documents the migrant had to work. According to Spanish law at the time foreign-born could legally work in Spain if they held Spanish citizenship, were citizens of EU member states (except for Romanians and Bulgarians for whom at the time of the survey there was no free labor mobility)¹² or citizens from non-EU members of the European Free Trade Association. Migrants also enjoyed a legal working status if they held the following authorization permits: permanent residency, temporary residency, EU residence permit, refugee status or asylum application. All other cases, that is, non-EU citizens who were not in possession of the necessary authorization to work, are considered illegal. Social ties in Spain are captured by a survey question on whether or not the migrant had any contacts (family, friends, acquaintances, an employer or others) upon arrival in Spain. It is only reported for migrants who arrived after 1989. Thus, in order to be able to use this information, for those migrants arriving prior to 1989, I assume this variable takes value 1.

Next I explain the set of countries that will be used as the reference group, and I detail the computation of effective working experience and the human capital measure that will be used in the analysis.

4.1. The migrant reference group

The model in Section 3 features the specific factors at play in the occupational choice of migrants as compared to natives. The ENI surveyed migrants only, and there is no additional dataset for Spanish workers containing all the key variables in this study. Therefore, I need to resort to some other migrants as reference group for the analysis. The goal is to use migrants who are as similar as possible in characteristics and labor market behavior to Spanish natives. The natural pool of countries to search for are European Union (EU) member states. Not only is there free labor mobility among EU countries, but they also have relatively similar educational systems, institutions and labor markets.

I use France, Germany and Portugal, which are among the most important origin countries of EU migrants in Spain, and together make up a group comparable in size to the other sending countries under consideration. Migrants from these three European countries constitute an appropriate reference group for a number of reasons. First, their self-employment rate is 19% which matches up the self-employment rate of Spanish natives in 2006/2007.¹³ Second, most of those migrants have been in Spain for long (almost 75% of them arrived before 1996) and so they are likely to have assimilated into Spanish labor markets and culture. The vast majority of them have good language competencies (78% are proficient in Spanish, with only 9% reporting having a poor level of Spanish). Moreover, I use Spanish Census data for the year 2001 to compare the demographic characteristics and education of the migrants in the reference group to those of native Spanish.¹⁴

As Table 2 shows, the average age of migrants in the reference group is similar to that of Spanish natives (migrants are less than one year older on average, although this difference is statistically significant) and they have the same marriage rate. There are, however, more women among the working population in the reference group. As for

¹² Although Romania and Bulgaria became EU member states on January 1, 2007, they were subject to a transitional period of 7 years during which EU countries could impose labor mobility restrictions. Spain was among the countries that exercised such restrictions. Therefore, during the years covered by the ENI survey, Romanians and Bulgarians, as migrants from outside the EU, needed to be in possession of an authorization permit to legally work in Spain.

¹³ The share of self-employed of native Spanish was 19.1% in 2006 and 18.9% in 2007 (data from Spanish Ministry of Social Security).

¹⁴ Despite some statistically significant differences, the group of EU migrants from France, Germany and Portugal constitute the most similar group to Spanish natives I could obtain.

Table 2
Average characteristics of Spanish natives and the migrant reference group.

	Spanish natives (I)	Migrant reference group (II)	Mean Difference (III)
Age	38.089 (11.428)	38.846 (8.994)	-0.757** (0.328)
Female	0.380 (0.485)	0.469 (0.499)	-0.089*** (0.014)
Married	0.632 (0.482)	0.637 (0.481)	-0.005 (0.014)
No education (%)	0.050 (0.217)	0.055 (0.227)	-0.005 (0.006)
Primary education (%)	0.167 (0.373)	0.158 (0.365)	0.009 (0.011)
Secondary education (%)	0.573 (0.495)	0.509 (0.500)	0.064*** (0.014)
Tertiary education (%)	0.210 (0.408)	0.279 (0.449)	-0.069*** (0.012)
Potential experience	22.234 (12.866)	21.135 (9.127)	1.099*** (0.370)
Observations	764,120	1,211	

Notes: Means for Spanish natives computed using the 2001 Population Census.

Table 3
Average characteristics of migrants by country of origin.

	Reference group (I)	Romanians (II)	Moroccans (III)	Ecuadorians (IV)
Entrepreneurship rate	0.19	0.06	0.11	0.05
Age	38.85	32.39	36.85	34.02
Years since arrival (YSA)	21.49	3.63	13.22	5.48
Post-1996 arrivals	0.26	0.98	0.55	0.96
Language knowledge - poor	0.09	0.29	0.33	0
Language knowledge - good	0.13	0.40	0.28	0
Language knowledge - very good	0.78	0.31	0.39	1
Schooling (years)	10.91	10.86	7.41	10.25
Potential working experience (years)	21.14	14.95	19.97	16.73
House ownership	0.55	0.19	0.35	0.33
Ownership of another house	0.07	0.00	0.03	0.02
Investments in stock exchange	0.10	0.01	0.02	0.01
Illegal status	0	0.32	0.07	0.04
Observations	1,211	932	910	976

the educational level, the share of workers with no formal education and primary education is about the same in both groups, whereas the percentage of workers with tertiary education in the reference group appears to be statistically larger than that of Spanish natives. This might partly be due to the slightly different way in which the ENI and the Census define tertiary education. While in the ENI survey tertiary education includes those with some college and completed college education (which cannot be distinguished), the Census classification only considers complete college education. As a result of the difference in education and average age, potential experience also appears to be slightly different between the two groups.

The comparison of the group of French, German and Portuguese migrants to the other migrants subject of analysis provides further validation for the choice of the reference group. As reported in Table 3, on average, the migrants in the reference group arrived much earlier in

Spain. The vast majority of Romanians and Ecuadorians (98% and 96% respectively) arrived after 1996, while for Moroccans this percentage is 55%. Migrants from the three individual countries tend to have lower skill levels than the reference group. Except for Romanians, whose average is not different from that of the reference group, the number of years of schooling is generally lower; about 30% of Romanians and Moroccans have poor language competencies and less than half of them report being very proficient in Spanish. Working experience is also lower for all three migrant groups. With regard to assets, migrants from all three countries of origin have lower percentages of home ownership than the reference group as well as lower ownership rates of another house or stock exchange investments. As for the legal status, 32% of Romanians could not be legally employed at the time of the survey, while for Moroccans and Ecuadorians this percentage was 7% and 4%, respectively. Altogether, these statistics show that the reference group of French, German and Portuguese migrants are different from the migrant groups of interest while they are close in characteristics to Spanish natives, which is precisely what is needed from the reference group.

4.2. Construction of effective experience and the human capital measure

The information on the working history of migrants includes variables on whether the current job is the first one they have had in Spain, time searching for a job, as well as the number of unemployment spells and length. This allows me to construct effective (rather than potential) working experience as follows. For those migrants working in their first job upon arrival in Spain, effective experience in Spain is the time elapsed since they were hired. When the year of the current contract is not available, I use the time searching for the first job; that is, effective experience is then potential experience in Spain minus time searching for the first job.¹⁵ For those migrants whose current job is not the first one in Spain, two cases can be distinguished: migrants who have not experienced unemployment spells (presumably, they changed from one job to another without being unemployed) and those who have been unemployed and looking for a job once or more times. Effective experience for the former is computed as potential experience in Spain minus the time elapsed until they got their first job. For those suffering unemployment spells, effective experience is computed as potential experience in Spain minus the time of the longest unemployment spell and as many months as times he has been unemployed for at least a month. This is a conservative way to compute effective experience for workers undergoing several unemployment spells; that is, effective experience will be slightly overestimated (or closer to potential experience) if those other unemployment spells were longer than one month. There is also some measurement error in the computation of effective experience due to the fact that a number of relevant variables are reported in intervals (search time, length of longest unemployment spell and number of spells). Yet, this measure of effective experience is still closer to the real one than potential experience.

The information about the working history in the country of origin is not as rich as that in Spain. I only know whether the migrant had worked or not there. Therefore, for those who had worked, effective experience in the country of origin is set equal to potential experience there (as standard, computed as age at migration minus the years of schooling after 6 years old) and zero for those who had not worked.

Effective working experience and schooling will constitute the basis for the measure of human capital used in the analysis but they need to be comparable across countries. In effect, the use of raw schooling and working experience as proxies for skills poses obvious comparability challenges as there is wide heterogeneity in educational curricula and

¹⁵ Searching time, unemployment spells and the number of times unemployed tend to be reported in intervals. I take the midpoint of the interval in each case.

Table 4
Mincerian logwage regressions by country of origin.

	Reference countries (I)	Romanians (II)	Moroccans (III)	Ecuadorians (IV)
Schooling in CO	0.044*** (0.005)	0.006* (0.003)	0.004** (0.002)	0.002 (0.002)
Schooling in Spain	0.044*** (0.005)	0.007 (0.007)	0.013*** (0.004)	-0.000 (0.004)
Effective experience in CO	0.014*** (0.005)	0.005* (0.003)	0.003 (0.004)	0.007*** (0.002)
(Effective experience in CO) ²	-0.000*** (0.000)	-0.000** (0.000)	-0.000 (0.000)	-0.000*** (0.000)
Effective experience in Spain	0.016** (0.006)	0.032*** (0.012)	0.012** (0.006)	0.020* (0.011)
(Effective experience in Spain) ²	-0.000 (0.000)	-0.000 (0.001)	-0.000* (0.000)	-0.001 (0.001)
Language knowledge - Good	-0.004 (0.056)	0.010 (0.022)	0.016 (0.026)	
Language knowledge - Very good	-0.029 (0.054)	0.022 (0.026)	0.060** (0.029)	
Female	-0.171*** (0.025)	-0.214*** (0.030)	-0.181*** (0.030)	-0.136*** (0.023)
Married	0.063** (0.028)	-0.013 (0.021)	0.061** (0.024)	-0.020 (0.016)
Illegal		-0.081*** (0.026)	-0.143*** (0.042)	-0.036 (0.049)
Log(hours worked)	0.626*** (0.059)	0.471*** (0.038)	0.436*** (0.064)	0.429*** (0.038)
R-squared	0.442	0.658	0.486	0.571
Observations	880	828	748	868

Notes: All regressions include sector, region and arrival cohort fixed effects. Robust standard errors in parentheses *** p<0.01, ** p <0.05, * p<0.1.

quality across countries. Moreover, some of the skills used at work (such as the language, corporate culture values and social norms) are country-specific and thus not perfectly transferable to the Spanish labor market. In order to deal with these issues, I use a common metric for migrants' human capital based on the returns that the Spanish labor market places on the schooling and working experience of migrants from each country of origin. That is, I estimate country-specific Mincerian wage regressions where the monthly (log) earnings of salaried migrants are regressed on schooling (distinguishing between schooling in the country of origin and that obtained in Spain), effective experience in the country of origin and in Spain and their quadratic terms, level of language knowledge, a set of individual controls (gender, marital and illegal status and hours worked), as well as sector, Spanish region and arrival cohort fixed effects.¹⁶

Table 4 shows the results of the Mincerian regressions. As observed, for the reference group one additional year of schooling increases wages in Spain by 4.4%,¹⁷ while the return of one additional year of working experience is between 1.4–1.6%. These estimates are in line with those obtained in other studies for native Spanish workers -see, for example, Arrazola and de Hevia (2008)-,¹⁸ which further validates the use of this set of countries as the reference group. By contrast, the

¹⁶ To minimize the effects of outliers, I eliminate the top and bottom 1% of the wage observations for each country. As the countries of origin are quite heterogenous, I estimate separate models for each one, which allows for different returns across countries on schooling and experience as well as on all other covariates. I checked the sensitivity of the results to different ways of computing the human capital measure, including using the coefficients obtained from pooling the countries together and allowing for different coefficients only for schooling and experience.

¹⁷ For this group of migrants, the returns to schooling obtained in Spain and in the country of origin are not statistically different from each other.

¹⁸ The estimated returns on schooling for the reference group are within the lower range of those obtained by Arrazola and de Hevia (2008) for Spanish workers in a simple Mincerian logwage model with only schooling, experience, experience squared and region fixed effects.

Table 5
Descriptive statistics of the Human Capital measure by country of origin.

Variable	Mean	Standard deviation	Min	Max
Reference group	0.650	0.152	0	1.009
Romanians	0.174	0.074	-0.039	0.570
Moroccans	0.125	0.074	-0.048	0.344
Ecuadorians	0.141	0.050	-0.047	0.500

returns to schooling are lower for the other sending countries (between 0.2–1.3%) which is also consistent with the widespread finding across host countries that the return to education is lower for immigrants than natives -Chiswick (1978) and Dustmann (1993). The return to schooling obtained in the country of origin tends to be about the same to the return to schooling obtained in Spain, except for the group of Moroccans for whom schooling in Spain brings about a considerably higher (and statistically different) return. In the case of Ecuadorians, the returns to schooling are quite imprecisely estimated with the point estimate being even negative in the case of schooling obtained in Spain. As the t- test reveals no statistically difference between the returns to schooling in Ecuador and in Spain, I will just use the point estimate for schooling in Ecuador in both cases. Effective experience in Spain (at least the linear term) is significant for all migrants and, except for the group of Moroccans, so is effective experience in the country of origin. It is also worth noting that compared to the reference group, the overall return to effective experience tends to be lower for all other migrant groups. Hence, the estimated returns to schooling and working experience suggest that migrants from Romania, Morocco and Ecuador suffer from significant human capital losses.

With all the point estimates of the returns to schooling and effective experience, I then compute a measure of human capital for each migrant (whether a salaried worker or self-employed) as follows:

$$\begin{aligned}
 HC \text{ measure} = & \beta_{co}^s * school_{co} + \beta_{sp}^s * school_{sp} + \beta_{co}^e * effexp_{co} + \\
 & + \beta_{co}^{e2} * (effexp_{co})^2 + \beta_{sp}^e * effexp_{sp} \\
 & + \beta_{sp}^{e2} * (effexp_{sp})^2
 \end{aligned} \quad (20)$$

where the β 's coefficients are the point estimates from each country's Mincerian logwage regression.¹⁹

Fig. 3 displays the kernel distribution of the computed human capital measure for the four migrant groups, and Table 5 shows basic descriptive statistics by country of origin. As observed, the computed human capital for the reference group of European migrants is clearly skewed to the right compared to the other three migrant groups. The reason for this is twofold. On the one hand, migrants from the reference group have higher levels of raw schooling and working experience and, as evident in Table 4, their returns to schooling and experience are also higher.

To complete the description of the data, Table 6 shows the sector of occupation of the migrants and the geographical distribution within Spain, while Table 7 reports summary statistics of all the variables used in the empirical analysis. As seen in Table 6, migrants are predominantly concentrated in the construction sector (that underwent an economic boom during the 1990s and 2000s) followed by manufacturing, retailing, hospitality and the domestic workers sector. Geographically, they are concentrated in the regions along the Mediterranean coast where the real state boom was most important (Catalunya, Valencia and Murcia), as well as in Madrid that hosts the capital and was also a fast-growing region during those years.

¹⁹ To check the sensitivity of the results, I also experimented computing human capital using only the returns that were statistically significant and imposing equality of returns across groups whenever hypothesis testing suggested so. All these alternative ways to measure human capital did not qualitatively change the results.

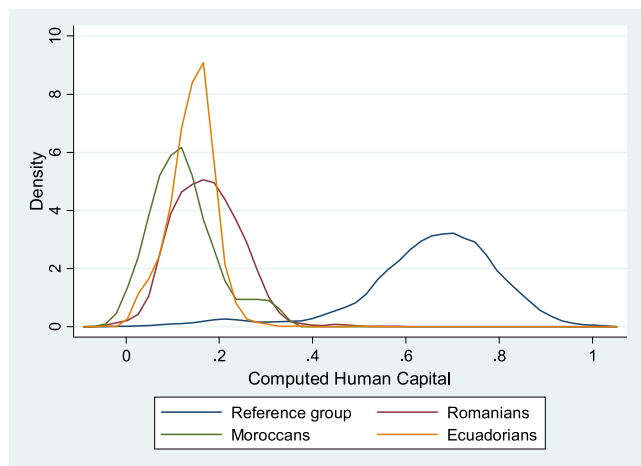


Fig. 3. Kernel density function of Human Capital.

Table 6
Sector and geographical distribution of migrants.

	All migrants (% of total)		All migrants (% of total)
Agriculture, fishing and mining	9.33	Andalucía	7.37
Manufacturing	12.76	Aragón	5.09
Construction	23.21	Baleares	2.31
Retail	10.62	Canarias	5.46
Hospitality	10.30	Canarias	2.11
Transportation and communications	10.30	Cantabria	2.23
Real state	4.84	Castilla y León	4.07
Domestic workers	7.10	Castilla-La Mancha	6.06
Education, health and other personal services	10.42	Catalunya	11.39
Others	8.36	C. Valenciana	8.29
	3.05	Extremadura	2.98
		Galicia	1.84
		Madrid	12.83
		Murcia	11.07
		Navarra	8.24
		País Vasco	2.63
		Rioja	4.34
		Ceuta and Melilla	1.69

5. Testing the model

The following table summarizes the model predictions regarding the way entrepreneurship entry costs, human capital losses and information deficiencies in the labor market affect sorting into entrepreneurship, the likelihood and the quality of entrepreneurship:

	Negative sorting	Likelihood of entrepreneurship	Entrepreneurial quality
Higher entry costs ($C_m > C$)	No	Low	High
Human capital losses ($0 < \delta_m < 1$)	No	Low	Ambiguous
Information problems in the labor market ($D_m > 0$)	Possible	High	Low

I will use these predictions to disentangle the main factors governing the occupational choice of Romanians, Moroccans and Ecuadorians. First, in Section 5.1, I assess whether there is negative sorting into self-employment by looking at self-employment rates along the skill distributions. Then, in Section 5.2, I estimate probit models on the likelihood of self-employment and examine entrepreneurial quality across countries of origin. Finally, additional explanations are discussed in Section 5.3.

Table 7
Descriptive statistics of the variables used in the empirical analysis.

Variable	Mean	Standard Deviation	Min	Max
Self-employed	0.107	0.309	0	1
Reference group	0.301	0.459	0	1
Romanians	0.233	0.423	0	1
Moroccans	0.224	0.417	0	1
Ecuadorians	0.243	0.429	0	1
Total years of schooling	9.960	3.843	0	16
Years of schooling in CO	7.617	5.335	0	16
Years of schooling in Spain	2.343	4.692	0	16
Total years of potential experience	18.324	9.448	0	49
Effective experience in CO	7.811	8.645	0	44
Effective experience in Spain	8.606	8.818	0	47.833
Language knowledge - Poor	0.169	0.375	0	1
Language knowledge - Good	0.194	0.395	0	1
Language knowledge - Very Good	0.637	0.481	0	1
Computed human capital measure	0.300	0.253	-0.048	1.009
Entrepreneurial experience in CO	0.082	0.275	0	1
Log earnings	6.846	0.585	0	9.210
Years since arrival (YSA)	11.594	11.977	0	62
Age	35.688	9.428	16	65
Female	0.434	0.496	0	1
Married	0.634	0.482	0	1
Has children	0.646	0.478	0	1
Illegal	0.100	0.301	0	1
Share of foreign-born in province (%)	12.003	4.670	2.388	21.759
Homeownership	0.366	0.482	0	1
No of other investments	0.108	0.370	0	5
Social ties in Spain	0.872	0.334	0	1
Province network size (% own-group in populat.)	1.553	2.047	0	15.807
Exposed to 2005 regularization program	0.371	0.483	0	1
Documented after 2005	0.227	0.419	0	1

5.1. Positive or negative sorting into self-employment?

Positive sorting into self-employment implies that self-employed migrants are drawn from the top of the skill distribution. However, if self-employment is also chosen by less capable migrants (negative sorting), we should observe relatively high self-employment rates at the bottom of the skill distribution too.

My preferred proxy for skills is the human capital measure computed in Section 4.2. As argued, it is comparable across countries of origin as it is based on a common metric, namely, the value that the Spanish labor market places on the education and working experience of migrants. Thus I will focus on this proxy. Yet, as raw education and working experience are standard proxies for skills used in labor economics, for the sake of completeness, Figs. 9 and 10 in the Appendix also show the self-employment rates along the distributions of education and effective experience. While self-employment tends to be higher among migrants with more years of effective working experience (see Fig. 10), examination of the plots of Fig. 9 leads to mixed conclusions regarding the sorting of migrants into self-employment across educational levels. On the plot for all migrant groups combined, self-employment rates appear to be higher than average at high educational levels (college and some college) but also at low levels (some primary education). However, if we look at individual countries, we only observe disproportionately high self-employment rates at low levels of education for Ecuador.

A clearer pattern of self-employment is observed for human capital. Fig. 4 displays self-employment rates along the distribution (deciles) of the computed measure of human capital for all migrants and for

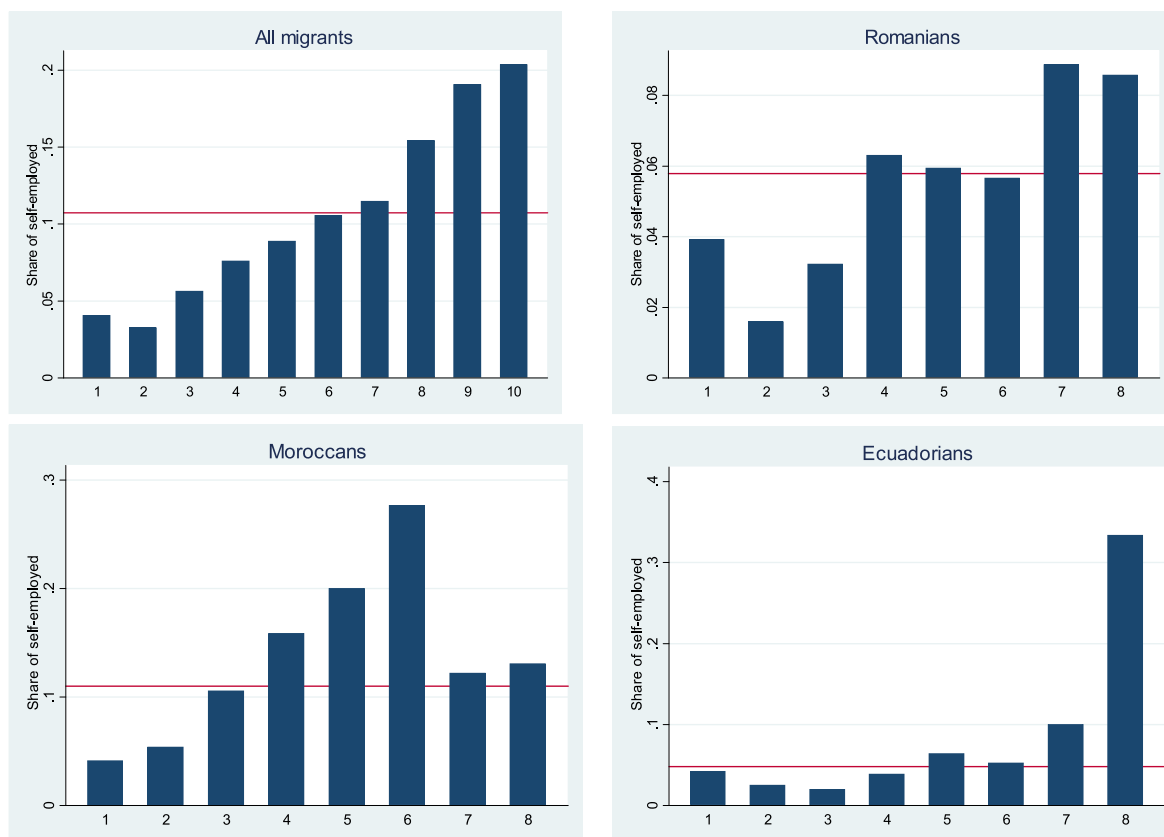


Fig. 4. Self-employment rates by deciles of computed Human Capital.

each one of the countries of interest.²⁰ The red line in the plots shows the average self-employment rate for each group. As observed, there is no evidence of negative sorting into self-employment. For each individual country (and for the whole set of migrants), the highest rates of self-employment are found at the top deciles of human capital. We observe higher than average self-employment rates at the 7th and 8th deciles of human capital (which are the top deciles) for Romanians and Ecuadorians.²¹ Interestingly, the plot for Moroccans features a hump shape with higher than average self-employment rates from the 4th decile onwards, and with the 4th, 5th and 6th deciles displaying the highest rates. To complete the analysis, Fig. 5 shows self-employment rates by another important skill, the level of language knowledge. As seen, this rate is higher than average among migrants with very good language competencies.

As complementary analysis, in the Appendix I take a different methodological approach on the question of sorting into self-employment. As in Fernández-Huertas (2011) and Chiquiar and Hanson (2005), I compare the distributions of education, effective experience and the computed measure of human capital for self-employed and salaried migrants, and test whether they are statistically different from each other. Except for the distributions of education, for which the picture is again blurry,²² the distribution of all other skill proxies of the self-employed tends to be skewed to the right compared to that of salaried

migrants. Moreover, in the case of effective experience, the Smirnov-Kolmogorov test of equality of distributions is rejected at sensible significance levels for all migrants and for Moroccans (it cannot be rejected for Romanians and Ecuadorians), and it is rejected in all cases for the computed measure of human capital. That is, self-employed migrants have statistically higher levels of human capital than salaried ones.

5.2. Entrepreneurial entry costs, human capital losses or information deficiencies in the labor market?

The lack of evidence of negative sorting into self-employment leaves room for any explanation (entrepreneurial entry costs, human capital losses or information deficiencies) on the differential occupational choice of migrants. Hence, I attempt to disentangle the relative importance of each of these factors by examining their predictions regarding the propensity for self-employment and the quality of self-employment.

I start by estimating the following baseline probit model on the likelihood of self-employment

$$\Pr(E_i|X) = \beta_0 + \beta_1 \cdot CO_i + \beta_2 \cdot X_i + \mu_s + v_r + \eta_a + \varepsilon_i \quad (21)$$

where E_i denotes if the migrant is self-employed, CO_i is the set of country-of-origin dummies (Romania, Morocco and Ecuador, with the reference group of European migrants being the omitted category) and constitute the variables of interest. X_i is a vector of explanatory variables for self-employment that includes standard individual controls (age, gender, marital status and whether the migrant has

²⁰ Individual plots for the reference group are not shown because this is not where the interest lies. Migrants in the reference group are like the natives, and I am interested in uncovering the mechanisms at work for the occupational choice of migrants upon arrival in the host country.

²¹ The plots for individual migrant countries display only up to eight human capital deciles because that is the range of human capital for those countries -see Fig. 3.

²² The Smirnov-Kolmogorov test does not rule out the equality of educational distributions of self-employed and salaried workers.

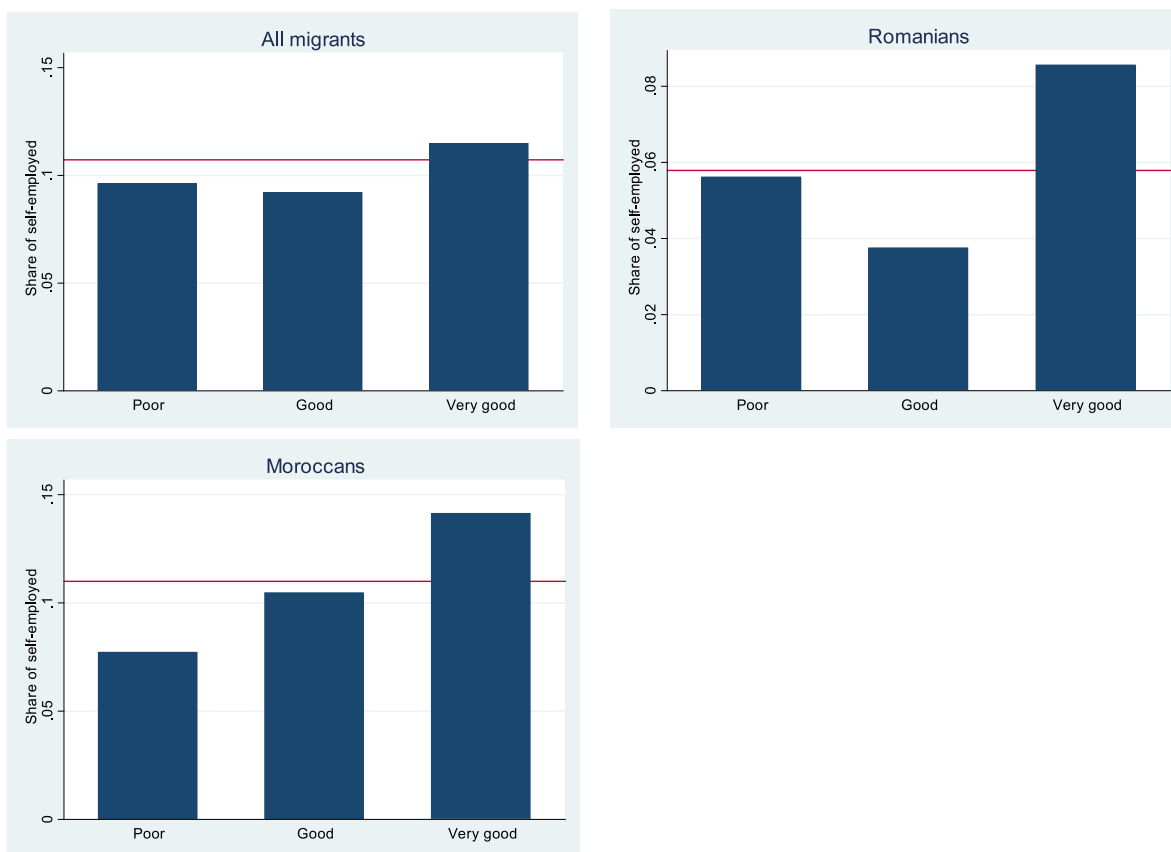


Fig. 5. Self-employment rates by levels of language proficiency.

children), years since arrival in Spain (YSA), whether the migrant had entrepreneurial experience in her country of origin and her legal status in Spain. In the baseline model, the proxies for skills are simply years of schooling, effective experience in the country of origin and in Spain, and quadratic terms of these variables. Additionally, I include the proportion of foreign-born in the province, as well as sector (μ_s), region (ν_r) and arrival cohort fixed effects (η_a).²³

I am interested on the difference in the propensity for self-employment of migrants from each main sending country with respect to migrants in the reference group, and thus the result tables will just report the marginal effects associated to those country-of-origin dummies. Table 14 on the Appendix reports the average marginal effects of all the variables included in the model. As expected, self-employment is significantly correlated to schooling and effective experience, both in Spain and in the country of origin. Migrants who were self-employed in their country of origin are also significantly more likely to be so in Spain. As for the propensity for self-employment across migrant groups, the reported marginal effects indicate that on average Romanians, Moroccans and Ecuadorians are respectively 11, 9 and 12 percentage points significantly less likely to be self-employed than migrants in the reference group. These self-employment differentials are also reported in the baseline model (columns I) of Table 8. What are these differences due to? According to the model, the lower propensity for self-employment is consistent with migrants experiencing higher

entrepreneurship entry costs compared to the reference group, and/or higher human capital losses. I examine the role of each of these factors next.

5.2.1. The role of entry costs and liquidity constraints

The literature has emphasized the importance for entrepreneurship of assets and wealth in the presence of financial constraints –e.g. Hurst and Lusardi (2004), Cagetti and De Nardi (2006) or Midrigan and Xu (2014). In this sense, migrants from Romania, Morocco and Ecuador might face higher entry costs than migrants in the reference group for at least two reasons. First, as Table 3 shows, they have less assets to invest or to use as collateral to access funds. Second, they have been in Spain for less time and thus are less likely to have credit histories, which again makes their access to credit more difficult. To account for these potential disadvantages, on column II of Table 8 Panel A, I add to the baseline probit model in (21) variables correlated with liquidity constraints, and examine by how much is the gap in the propensity for self-employment reduced.

As in other empirical studies - e.g. Constant and Zimmermann (2006) or Disney and Gathergood (2009)- I include a dummy variable for whether the migrant owns her house in Spain. Owning a house, even if the mortgage has not been fully paid back, is not only an indicator of the migrant's wealth but it is also informative of her possibilities to successfully access credit. I interact homeownership with the country-of-origin dummies to allow for a differential role of liquidity constraints across migrant groups. The other variable I include is the number of investments, other than the house of residence, made in Spain. These are investments in the stock exchange market, in land, in another property, in other people's business or lending money to others. While the explanatory power of each one of these investments alone is low

²³ The province is a geographical administrative unit smaller than the region. In Spain there are 18 regions and 52 provinces. As for the arrival cohort fixed effects, as this is a cross-section that cannot support the variable YSA and yearly arrival cohort dummies, I construct 5-year arrival cohorts before 1996 and 2-year arrival cohorts after 1996 when the arrival of migrants intensified.

Table 8

Differences in propensities for self-employment with respect to the reference group by country of origin - Average marginal effects and marginal effects at key representative values.

Panel A. Adding variables correlated with liquidity constraints			
	Baseline model	Number of investments other than house	No. of investm. & at Homeownership = 1
	(I)	(II)	(III)
Romanians (co2)	-0.109*** (0.038)	-0.106*** (0.038)	-0.096*** (0.038)
Moroccans (co3)	-0.090*** (0.019)	-0.088*** (0.019)	-0.071*** (0.025)
Ecuadorians (co4)	-0.120*** (0.031)	-0.130*** (0.023)	-0.116*** (0.031)
Panel B. Accounting for human capital losses			
	Baseline Model	At Language knowledge = Very Good	Human Capital & at Lang knowledge = Very Good
	(I)	(II)	(III)
Romanians (co2)	-0.109*** (0.038)	-0.079** (0.037)	-0.042 (0.035)
Moroccans (co3)	-0.090*** (0.019)	-0.066*** (0.023)	0.045 (0.065)
Ecuadorians (co4)	-0.120*** (0.031)	-0.105*** (0.031)	-0.087** (0.040)
Panel C. Adding variables correlated with information problems			
	Baseline Model	Share of own group & Ties in Spain	Share of own group & at Ties in Spain = 0
	(I)	(II)	(III)
Romanians (co2)	-0.109*** (0.038)	-0.094** (0.045)	-0.097* (0.055)
Moroccans (co3)	-0.090*** (0.019)	-0.082*** (0.021)	-0.030 (0.047)
Ecuadorians (co4)	-0.120*** (0.031)	-0.103*** (0.035)	-0.134*** (0.043)

Notes: The baseline model includes schooling, effective experience in CO and in Spain, and quadratic terms of these variables, YSA, the individual characteristics of age, gender, marital and legal status, whether the migrant has children, whether she had entrepreneurial experience in CO, the share of total foreign-born in the province and sector, region and arrival cohort fixed effects. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

(I only know whether the migrant has made such investment but not by which amount), the number of investments is informative of her relative wealth and, indeed, it is positively correlated with self-employment. The average marginal effects reported on column II of Table 8 Panel A are only slightly smaller, in absolute value, than those of the baseline model (see columns I and II) for the case of Romanians (-0.106 vs -0.109) and Moroccans (-0.088 vs -0.090). However, when we look at the marginal effects for those migrants who own a house (column III on Panel A), we observe that the difference in the probability of self-employment is reduced across all migrant groups (by around 2 and 1 percentage points for Moroccans and Romanians respectively, and by less than 1 percentage point for Ecuadorians). Thus, although liquidity constraints alone cannot explain the lower propensities for self-employment of these migrants (the marginal effects are still statistically different from zero), these results are indicative of some liquidity constraints hampering entry into self-employment.

Entry costs can also be related to information problems regarding self-employment. Migrants might not know how to navigate the requirements and red tape barriers to start a business and they are subject to information problems in the credit market of the host country. I will address this problem later when the role of social ties and networks will be considered.

5.2.2. The role of human capital losses

I account for human capital losses on Panel B of Table 8. One skill that is clearly not portable across countries is knowledge of the

language - Chiswick (1978), Friedberg (2000) and Chiswick and Miller (2015). I add to the baseline model the categorical dummy on language knowledge and interact it with the dummies of origin, which allows me to compute heterogeneous effects of this skill across groups. Column II on Panel B of Table 8 reports the marginal effects evaluated at very good level of language knowledge. The results show that compared to the reference group, Romanians who are proficient in Spanish are 8 percentage points less likely to become self-employed (and this difference is now significant only at the 5% level) while the self-employment gap for Moroccans shrinks to 6.6 percentage points.²⁴ This indicates that part of the gap in self-employment rates is due to poor language knowledge.

On column III of Panel B, I keep language proficiency as a control and, instead of schooling and effective experience, I use the measure of human capital as computed in 4.2. Recall that this measure of human capital is based on the value that the Spanish labor market places on the education and working experience of migrants. Therefore, it overcomes the comparability issues of using schooling obtained in different countries and, more importantly, it accounts for the human capital losses migrants endure upon arrival. Consequently, if higher (than the reference group) human capital losses are responsible for the low propensities for self-employment, the marginal effects associated to the country-of-origin dummies should not be distinguishable from zero when this measure of human capital is used. Column III reports marginal effects at very good language knowledge and at average levels of human capital and all other variables. As seen, the marginal effects for Romanians and Moroccans are now no longer statistically different from zero. That is, when human capital is properly accounted for, the propensity for self-employment of Romanians and Moroccans is not statistically distinguishable from that of the reference group. Another interesting point to notice is that, although not significant, the marginal effect for Moroccans switches sign and is now positive. That is, taken at face value, the estimate indicates that once we account for human capital losses, Moroccans are 4.5 percentage points more likely than the reference group to become self-employed. By contrast, human capital losses cannot completely explain the lower propensity for self-employment of Ecuadorians who are still on average 8.7 percentage points less likely than the reference group to choose self-employment.

In sum, the estimates of Panel B indicate that human capital losses (including poor language knowledge) are behind the lower propensities for self-employment of Romanians and Moroccans, whereas they can only explain part of the self-employment differential of Ecuadorians.

5.2.3. The role of information deficiencies in the labor market

Beyond mere human capital losses, migrants are likely to encounter additional difficulties to be hired in the labor market due to information gaps of at least two sorts. First, it might be difficult for migrants to learn about available jobs. Second, it can be challenging for employers to elicit the true ability and other attributes (e.g. job commitment or reliability) of migrants whose cultural backgrounds they do not know that well, and these information deficiencies are likely to be more severe for migrants from countries culturally more distant to the host country.

To account for the role of information gaps, on Panel C of Table 8 I use the following two variables: (i) the share of other co-nationals in the total population of the province they live, which is a proxy for migrant network effects in the vein of Munshi (2003) or Dustmann et al. (2016),

²⁴ As Ecuadorians are Spanish native speakers, the change in their marginal effect from -0.120 in the baseline model to -0.105 might be surprising. This is actually explained by the lower propensity for self-employment of Spanish proficient migrants in the reference group. Even accounting for the lower self-employment propensity of the reference group at this level of Spanish proficiency, the gap in self-employment of Romanians and Moroccans is reduced by about 1.5 and 1 percentage points, respectively.

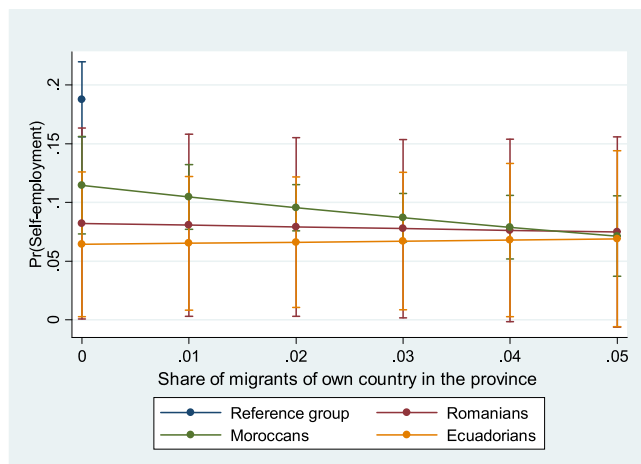


Fig. 6. Predicted propensities for self-employment by size of province network.

and (ii) a binary variable indicating whether the migrant had social ties in Spain (relatives, friends, acquaintances, employer or others) when she arrived.

There is extensive evidence of the importance of networks and connections in labor markets as many workers are hired through informal social contacts (friends and relatives) and referrals - [Calvó-Armengol and Jackson \(2004\)](#), [Ioannides and Loury \(2004\)](#) and [Dustmann et al. \(2016\)](#). However, social networks also help entrepreneurship as they can provide information about market opportunities, buyers and suppliers. Furthermore, they might be key to accessing capital when migrants lack collateral wealth and/or credit histories in the host country. Therefore, it is not clear a priori how the existence of networks and social ties is to correlate with self-employment. That is why I will focus on the opposite situation. What do migrants do when they lack the contacts and the support of a community from their same origin? In the absence of social contacts, migrants need to search for jobs through formal labor markets and this will lead to heterogeneous labor market outcomes across migrant groups depending on how severe information deficiencies are. I expect migrants who lack social contacts and come from countries culturally more distant to Spain to endure more difficulties to be hired and, consequently, to resort more to self-employment.

In Panel C column II, I go back to the baseline model and add the share of migrants from the same country in the province,²⁵ and the dummy variable on social ties in Spain. The average marginal effects reveal an increase in the average propensity for self-employment across all migrant groups (that is, there is less of a gap, in comparison to the reference group, in the likelihood of self-employment). However, average propensities mask quite some heterogeneity by size of migrant networks. [Fig. 6](#) plots the predicted propensities for self-employment for each migrant group by size of the province network, from 0 up to a

²⁵ As the focus is on the role of network effects for the three main groups of migrants, I do not include the share of European migrants (the reference group).

Table 9

Differences in propensities to self-employment accounting for liquidity constraints, human capital losses and information problems - Marginal effects at key representative values.

	Homeownership=1 & Lang knowledge = Very Good (I)	Homeownership=1 & Lang knowledge = Very Good & Ties in Spain = 0 (II)	Homeownership=1 & Lang knowledge = Very Good & Ties in Spain = 1 (III)
Romanians (co2)	-0.034 (0.039)	-0.034 (0.051)	-0.033 (0.040)
Moroccans (co3)	0.069 (0.070)	0.143* (0.084)	0.059 (0.070)
Ecuadorians (co4)	-0.080* (0.046)	-0.117** (0.052)	-0.077 (0.047)

Notes: The model includes homeownership and the number of other investments (to account for liquidity constraints), language knowledge and the computed Human Capital measure (to account for human capital losses), the size of province networks and social ties in Spain (to account for information deficiencies in the labor market). Additionally, it includes YSA, the individual characteristics of age, gender, marital and legal status, whether the migrant has children, whether she had entrepreneurial experience in CO, the share of total foreign-born in the province and sector, region and arrival cohort fixed effects.

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

share of 6%.²⁶ As observed, the propensity of self-employment of Moroccans decreases as the share of other co-nationals in the province gets larger, and is not statistically different (at the 5% significance level) from that of the reference group for a share of 0. For Romanians this pattern is almost flat while for Ecuadorians the predicted propensity for self-employment increases the more Ecuadorians are in the province.

On column III of Panel C, I explore further into the role of social ties by looking at the marginal effects for those migrants with no social ties in Spain. Heterogeneous results across groups of migrants are obtained. While the gap in self-employment (with respect to the reference group) becomes larger for Ecuadorians (13.4 vs 12 percentage points in the baseline model), for Romanians it shrinks by 1 percentage point and for Moroccans the self-employment gap reduces considerably and becomes statistically insignificant.

In short, Panel C shows that, in the absence of social ties and networks from the same community, Moroccans tend to resort more to self-employment. This effect is also somehow present, albeit much less marked, for Romanians, and the opposite occurs for Ecuadorians. These results are consistent with Moroccans and Romanians suffering from information problems to be hired in the labor market, and these being much more severe in the case of Moroccans. By contrast, Ecuadorians with no social connections seem to be drawn more into salaried employment.

Finally, in [Table 9](#), I examine how the propensity for self-employment of Romanians, Moroccans and Ecuadorians compares to that of the reference group when all the previous factors combined are accounted for. That is, I estimate a probit model on the likelihood of self-employment including the usual controls and all of the following: homeownership and number of investments (to control for liquidity constraints), language knowledge and the computed measure of human capital (to account for human capital losses) and social ties and province networks (to capture the effect of information deficiencies). Column I of [Table 9](#) reports the marginal effects for migrants who own the house they live in and are proficient in Spanish. The results confirm the earlier findings in [Table 8](#), namely, once human capital is properly accounted for, there is no statistical difference in the

²⁶ After the share of Moroccans in Ceuta and Melilla (which is about 15%), the largest shares of own-country migrants in 2006 were 6% for Romanians in Castellón (in the region of Valencia), 5.3% for Moroccans in Almería (in Andalucía) and 2.5% and 3.7% of Ecuadorians in Madrid and Murcia, respectively.

propensity for self-employment of Romanians and Moroccans compared to the reference group, while Ecuadorians show a smaller, but still significant, self-employment gap. The other two columns of [Table 9](#) report the marginal effects for the same model but for those migrants with no social ties (column II) and with social ties (column III). The marginal effect for Romanians remains practically unchanged whether we look at migrants with or without social ties, which indicates that, overall, social connections are not as important as human capital losses to explain the self-employment gap of this group. Social connections matter, though in opposite directions, for the occupational choice of the other two groups of migrants. Moroccans with no social ties are 14 percentage points more likely that their counterparts in the reference group to choose self-employment, and this difference is statistically significant (see column II). By contrast, for Ecuadorians with no social ties the gap in the propensity for self-employment is about as large as in the baseline model, and this is only reduced and statistically insignificant for migrants with social ties (column III). That is, unlike Moroccans, for Ecuadorians information problems would not be operating to hamper their employability in the labor market but, if anything, to discourage them from becoming self-employed. I will discuss further the explanations for the self-employment gap of Ecuadorians in [Section 5.3](#).

5.2.4. Entrepreneurial quality

The analysis on the propensity for self-employment reveals that human capital losses play a role in the occupational choice of the three main migrant groups in Spain and, particularly for Moroccans, so do information gaps. Focusing on the sub-sample of self-employed migrants, I now complement the analysis by examining the predictions regarding entrepreneurial quality. Recall that the effect of high human capital losses on average entrepreneurial quality was ambiguous, as it depended on the particular distribution of skills. By contrast, the model predicted lower entrepreneurial quality for those migrants experiencing problems to be hired in the labor market associated to information deficiencies.

I use self-employed earnings as a proxy for entrepreneurial quality, and estimate probit models on the probability of earnings belonging to the bottom and the top of the self-employed earnings distribution. Results are reported in [Table 10](#). All regressions include controls for individual characteristics, but not proxies for skills as they are associated to entrepreneurial quality, which is precisely what the dependant variable is to capture. I also allow for sector, Spanish region and arrival cohort fixed effects.²⁷ The focus is on the marginal effect associated to the dummies of origin (the omitted category still being the reference group of migrants from France, Germany and Portugal) and whether they are statistically different from each other. Columns I and II of [Table 10](#) estimate the probability of earnings being at the bottom 25% and 10% of the self-employed earnings distribution respectively, whereas models III and IV estimate the probability of earnings being at the top of the distribution.²⁸

²⁷ In regressions not reported here, I also controlled for hours worked. Although subject to potentially important self-reporting error, it is a way to account for entrepreneurship efficiency. The main qualitative results proved to be robust to this control.

²⁸ It is important to note that the model predictions are based on the assumption that the distributions of skills and entrepreneurial ability are the same across migrant groups. Examination of the distributions of human capital in [Fig. 3](#) rejects this assumption. The distribution of human capital for the reference group is skewed to the right and the support of the distribution also seems to be different from that of the other groups. Thus, we would expect migrants from the reference group to be predominant at the upper end of the self-employed earnings distribution. In any case, though, the interest in this exercise lies on the comparison across the other groups (Romanians, Moroccans and Ecuadorians) which do share the same support for the skill distributions.

Table 10

Probit regressions on the probability of belonging to the bottom and the top of the self-employed earnings distribution - Average marginal effects.

	Earnings <25% (I)	Earnings <10% (II)	Earnings >75% (III)	Earnings >90% (IV)
Romanians (co2)	-0.064 (0.075)	-0.163*** (0.053)	-0.081 (0.105)	-0.162*** (0.051)
Moroccans (co3)	0.179** (0.071)	0.012 (0.065)	-0.295*** (0.046)	-0.174*** (0.031)
Ecuadorians (co4)	0.058 (0.095)	-0.119** (0.058)	-0.124 (0.094)	-0.083 (0.052)
Observations	310	255	308	215
Hypothesis testing (<i>p</i> -value):				
co3 = co2	0.009	0.007	0.011	0.787
co3 = co4	0.235	0.059	0.020	0.023

Notes: All regressions include the individual characteristics of age, gender, marital and legal status, YSA, sector, region and arrival cohort fixed effects.

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

A consistent pattern is observed for Moroccans. Self-employed migrants from Morocco are more likely to have earnings at the lower end of the distribution (see column I in [Table 10](#)). Specifically, Moroccans are 18 percentage points more likely than migrants from the reference group to fall at the bottom quartile of the earnings distribution while no statistically difference is observed in the probability of low earnings of Romanians and Ecuadorians compared to the reference group. When we consider the probability of self-employment earnings falling at the bottom 10% of the self-employed earnings distribution (column II), we see that the marginal effect for Moroccans is not statistically different from that of the reference group, but it is statistically different from that of Romanians and Ecuadorians, which is now negative.²⁹ If instead we look at the top 75% and 90% of the earnings distribution (columns III and IV of [Table 10](#)), self-employed from Morocco are significantly less likely than the reference group to have earnings at this end of the distribution and, by and large, this lower probability is statistically different from that of Ecuadorians and Romanians.³⁰ In sum, controlling for individual characteristics and fixed effects, Moroccan self-employed are consistently more likely to have low earnings and less likely to have high earnings, which is indicative of low entrepreneurial quality.

The results in [Table 10](#) together with those of [Table 9](#) column II, which show a relatively high propensity for self-employment of Moroccans when human capital is properly accounted for and social ties are null, point to Moroccans resorting to self-employment as a way to deal with the additional penalties associated to information problems endured in the labor market. This finding is also consistent with the results obtained in previous studies. In particular, it is in line with the statistical discrimination that [Bosch et al. \(2010\)](#) found Moroccans suffer in the Spanish housing market and the larger employability and occupational gaps that Moroccans experience in the Spanish labor market ([Amuedo-Dorantes and De la Rica, 2007](#)).

As additional evidence, I exploit further the information on arrival cohorts of Moroccans. Compared to the other two main migrant groups, many Moroccans have been in Spain for long. Almost half of them arrived before 1996 while for Romanians and Ecuadorians this percentage is less than 5%. The earliest arrival cohorts were likely to endure the most information problems in the labor market, as they had no social ties and they migrated to a country, Spain, that back then did not host many Moroccans or any other foreigners. [Table 11](#) reports

²⁹ As the bottom 10% corresponds to a very limited number of observations, results are more sensitive to the presence of outliers in the reference group.

³⁰ There is only one instance, when the top 10% of the self-employed earnings distribution is considered, in which we cannot reject the marginal effect for Moroccans being equal to that of Romanians.

Table 11
Pre- and post-1996 arrival cohorts effects for Moroccans- Marginal effects.

	Propensity to self-employment		Prob of self-employment earnings	
	Baseline model (I)	Accounting for HC losses & at Lang Knowledge=Very Good (II)	< 25% (III)	> 75% (IV)
Arrived before 1996	-0.011 (0.027)	0.127* (0.070)	0.150* (0.083)	-0.346*** (0.066)
Arrived after 1996	-0.148*** (0.030)	0.010 (0.074)	0.165 (0.133)	-0.019 (0.122)
Observations	3,961	3,961	311	307

Notes: The baseline model includes schooling, effective experience in CO and in Spain, and quadratic terms of these variables, YSA, the individual characteristics of age, gender, marital and legal status, whether the migrant has children, whether she had entrepreneurial experience in CO, the share of total foreign-born in the province and sector, region and arrival cohort fixed effects.

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

the estimation results on the propensity for self-employment (columns I and II) and entrepreneurial quality (columns III and IV) allowing for a different marginal effect for Moroccans that arrived before and after 1996. As observed in column I, in the baseline model, the post-1996 arrivals have a lower propensity to self-employment than migrants in the reference group while the propensity for the earlier cohorts is not statistically different from that of the reference group. When human capital losses are accounted for (column II), it is the earlier cohorts of Moroccans that display a statistically higher propensity to self-employment. Looking at the probit models of columns III and IV, we see that the Moroccans who arrived before 1996 also appear to have lower entrepreneurial ability. That is, they are statistically more likely to have earnings at the bottom of the self-employed earnings distributions and less likely to have them at the top of the distribution.

5.3. Alternative explanations

Two clear conclusions emerge from the previous analysis: (i) the self-employment differentials of Moroccans and Romanians observed in the raw data are largely explained by human capital losses and (ii) Moroccans are also subject to information deficiencies in the labor market that push them to self-employment. However, the model cannot explain as well the occupational choice of Ecuadorians. After controlling for human capital losses, a significant self-employment gap persists for this group. In this section, I consider additional explanations that could be at play. One is the regularization program of undocumented migrants implemented in 2005, and the other one has to do with the idiosyncratic skills Ecuadorians possess and their sorting across sectors.

In December 2004, the Spanish government announced a regularization program to address the large number of undocumented migrants residing in Spain at the time. Applications for the program could be submitted between February 7 and May 7, 2005. To be eligible to apply, migrants needed to have resided in Spain for over six months (*i.e.*, prior to August 2004), have no criminal record in their home country and in Spain and have a bonafide work contract of at least 6 months. Importantly, self-employed, as well as students, spouses and children could not benefit from the program -Arango and Jachimowicz (2005). As the program favored salaried work over self-employment, the policy could have impacted the occupational choice of migrants who were undocumented at the time of the announcement. Thus, I check whether this regularization program could help explain the self-employment differentials across migrant groups.

There is no information in the ENI survey on whether migrants had applied for the regularization program but I can still construct two variables to identify the migrants who were impacted by the program. The first variable, “Exposure”, accounts for all the migrants

Table 12
Effect of the 2005 regularization program on the propensity to self-employment - Marginal effects.

	Baseline model (I)	Not affected by the program (Exposure = 0) (II)	Not affected by the program (Documented after 2005 = 0) (III)
Romanians (co2)	-0.109*** (0.038)	-0.101** (0.042)	-0.106*** (0.038)
Moroccans (co3)	-0.090*** (0.019)	-0.089*** (0.020)	-0.084*** (0.019)
Ecuadorians (co4)	-0.120*** (0.031)	-0.119*** (0.032)	-0.114** (0.032)

Notes: The baseline model includes schooling, effective experience in CO and in Spain, and quadratic terms of these variables, YSA, the individual characteristics of age, gender, marital and legal status, whether the migrant has children, whether she had entrepreneurial experience in CO, the share of total foreign-born in the province and sector, region and arrival cohort fixed effects.

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

Table 13
Previous entrepreneurial experience in CO of self-employed migrants.

	Share of self-employed with entrepreneurial experience in CO (I)	Difference wrt reference group (II)	Difference wrt reference group controlling for age (III)
Reference group	13.4		
Romanians	18.5	5.1	11.1*
Moroccans	15.0	1.6	2.0
Ecuadorians	31.9	18.5***	22.5***

Notes: *** p<0.01, ** p<0.05, * p<0.1.

who could potentially benefit from the program. That is, migrants who had arrived in Spain before August 2004 and did not have a legal status by February 2005. The other variable, “Documented after 2005”, takes value 1 for those migrants who had arrived in Spain before August 2004 and obtained the legal status after May 2005. I add each one of these variables to the baseline probit model on the likelihood of self-employment to see how much of the self-employment gap can be explained by the regularization program. Table 12 reports the marginal effects for migrants not exposed to the program (column II) and for those who already had a legal status before 2005 (column III). That is, I am focusing on the propensities for self-employment of those migrants whose incentives to become self-employed were not affected by the program. The marginal effects on columns II and III of Table 12 are less than one percentage point lower, in absolute values, to those of the baseline model (column I). In other words, the regularization program of 2005 seems to have played a very minor role, if any, in the occupational choice of migrants in Spain, and certainly does not help explain the self-employment gap of Ecuadorians.

The self-employment gap of Ecuadorians only vanishes once all factors combined are taken into consideration for those migrants with social ties in Spain –see column III in Table 9. This could be due to Ecuadorians facing higher entry costs into self-employment associated to information problems. If this is the case, we should also observe higher entrepreneurial quality for this group. There is some evidence pointing in that direction. First, Fig. 4 suggests very positive sorting of Ecuadorians into self-employment (*i.e.*, the self-employment rate is very high at the highest decile). Similarly, taking the results of Table 10 column IV at face value, the probability of Ecuadorian self-employed obtaining earnings in the top 10% of the distribution is not statistically different from that of migrants in the reference group. Finally, if we believe that entrepreneurial talent is persistent, I use entrepreneurial experience in the country of origin as another proxy for entrepreneurial quality. As Table 13 reports, the share of self-employed migrants with

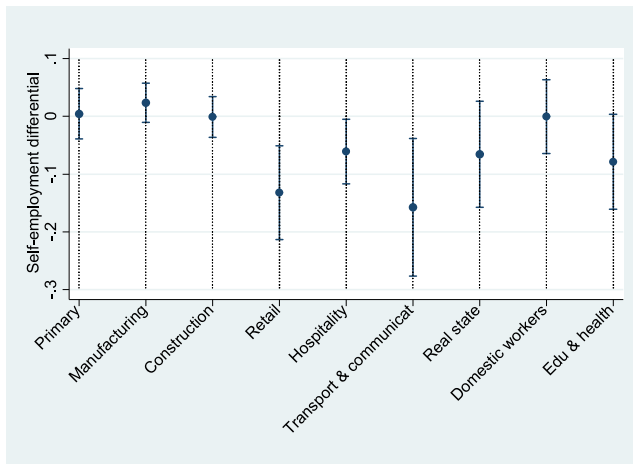


Fig. 7. Self-employment differentials of migrants with language proficiency (wrt those with poor language knowledge) by sector.

Notes: Predicted self-employment differentials obtained from the baseline probit model on self-employment using all the ENI sample of migrants who arrived in Spain after 1996 (i.e., 6,264 observations).

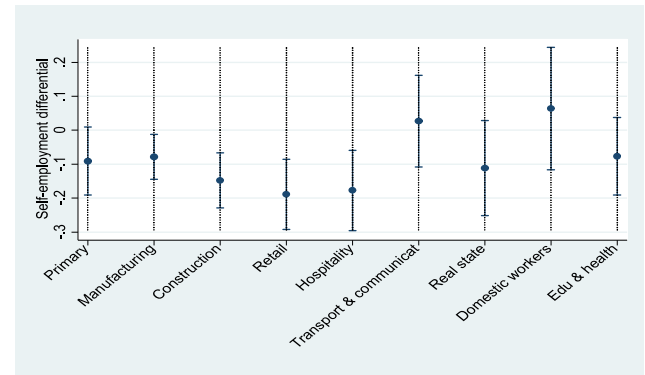
previous entrepreneurial experience is higher for the three main sending countries compared to the reference group. However, this difference is only statistically significant for Ecuadorians (see column II). When age is controlled for (column III), these differences are even larger and again only statistically different at the 5% significance level for Ecuadorians. Taken together these pieces of evidence are consistent with Ecuadorians facing higher entry costs into self-employment.

On the other hand, given Ecuadorians' idiosyncratic skills, namely Spanish language proficiency and relative cultural proximity, it is also natural to ask whether they are pushed out of self-employment or rather pulled into salaried work. In effect, although language proficiency is controlled for in Tables 8 and 9, this might not be sufficient to fully account for the skills Ecuadorians possess that give them an advantage, in comparison with the other main migrant groups, in the Spanish labor market. I further investigate the reasons why Ecuadorians might be disproportionately drawn into salaried work by conducting the following sectoral analysis.

I use all the migrants in the ENI survey to estimate the marginal effect on self-employment of language knowledge across sectors.³¹ This way I can identify the sectors where language proficiency is predominantly associated to salaried work, and then I compare them to the sectors where the self-employment gap of Ecuadorians is the largest. Fig. 7 shows that, compared to migrants with poor language skills, for those proficient in Spanish, salaried work tends to be prominent (equivalently, the self-employment gap is the largest) in Transportation and Communications, Retail and Hospitality. On the other hand, Fig. 8 shows the self-employment gap of Ecuadorians (with respect to the reference group) by sector, as predicted by the estimated probit models on self-employment. According to the baseline model (top panel of Fig. 8), the self-employment differentials are the largest, and significant at the 5% level, in Retail, Hospitality, Construction and Manufacturing. Once human capital losses are accounted for (bottom panel of Fig. 8) the self-employment gap diminishes and persists in a significant way only in Retail, Construction and Hospitality. That is, two (out of

³¹ I do not restrict the analysis to the four migrant groups considered thus far but instead use all migrants to prevent the results on language proficiency from being driven solely by Ecuadorians.

According to the baseline model



Accounting for human capital losses

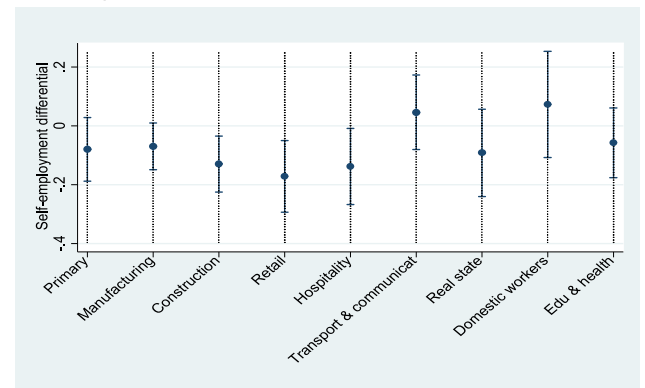


Fig. 8. Predicted self-employment differential of Ecuadorians (wrt the reference group) by sector.

three) of the sectors where Ecuadorians display a low propensity for self-employment (Retail and Hospitality) are sectors where migrants with language proficiency predominantly choose salaried employment. Thus, based on this analysis Ecuadorians also seem to be pulled into salaried work due to their advantage with regard to language skills.

6. Conclusions

This paper develops a framework to better understand the underlying factors driving the choice of migrants between entrepreneurship and salaried work. The model shows that negative sorting into entrepreneurship is possible when migrants are subject to penalties in the labor market, beyond mere human capital losses, associated with severe information problems. It also generates testable predictions regarding the likelihood and the average quality of entrepreneurship, which allows us to disentangle the relative importance of heterogeneous entrepreneurship entry costs, human capital losses, and information deficiencies in the labor market.

I test the model predictions using a rich survey of immigrants in Spain for the years 2006–2007. Spain constitutes an interesting case study because it received an unprecedented migration flow during a short span of time from the mid-1990s to the mid-late 2000s. Thus, the data used here encompasses the migration boom and is prior to the global financial crisis of 2008. To avoid the problems derived from aggregating migrants into arbitrary groupings, I focus on individual countries, in particular, on the three main origin countries of migrants in Spain at the time (Romania, Morocco, and Ecuador) and a reference group of European countries. Entrepreneurship is proxied with self-employment, which is what is observed in the data.

The results are easily summarized. First, there is no evidence of negative sorting into self-employment. By and large, across the three main migrant groups, self-employment rates are higher at the top, not at the bottom, of the human capital distribution. Second, I find that human capital losses alone can explain the gap in the propensity for self-employment of Romanians and Moroccans, compared to the reference group, observed in the raw data. For Ecuadorians, after controlling for human capital losses, a significant, albeit smaller, self-employment gap remains unexplained. Additional analysis for this group suggests that while information gaps could discourage them from becoming self-employed, the idiosyncratic skills Ecuadorians possess (namely, language proficiency and possibly cultural proximity) might have also favored their employability in the Spanish labor market. Third, the results also show that, once human capital is properly accounted for, the predicted propensity for self-employment among Moroccans with no social ties and networks in Spain is significantly larger than that of migrants in the reference group. Furthermore, these migrants appear to have consistently lower entrepreneurial quality. These findings are consistent with Moroccans suffering from information problems in the labor market that push them into self-employment. Public policies aimed at addressing these information gaps would then be necessary to ensure a more efficient allocation of skills across occupational choices.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

A.1. Effect of human capital losses δ on average entrepreneurial skill for some examples of distributions

Consider the uniform distribution with density function $\gamma(z) = \frac{1}{\bar{z}-z}$ and cumulative density function $\Gamma(z) = \frac{z-z}{\bar{z}-z}$ on the interval $[z, \bar{z}]$. The average entrepreneurial skill when the skill of the marginal entrepreneur is S_0 is $E(z|S_0) = \frac{z+S_0}{2}$.

As shown in Section 3.2.2, for a given human capital loss $\delta_m > 0$, the skill of the marginal entrepreneur increases to S_1 with $S_1 = \frac{S_0}{1-\delta_m}$. The average entrepreneurial skill is then $E((1-\delta_m)z|S_1) = (1-\delta_m)\frac{z+S_1}{2}$. Notice that $E((1-\delta_m)z|S_1) - E(z|S_0) = \frac{-\delta_m \bar{z}}{2} < 0$. That is, the average entrepreneurial skill decreases.

Now consider a Pareto distribution with density function $\gamma(z) = \frac{\alpha z^\alpha}{z^{\alpha+1}}$ and cumulative density function $\Gamma(z) = 1 - \left(\frac{z}{\bar{z}}\right)^\alpha$ on the interval $[z, \infty)$. The average entrepreneurial skill when the skill of the marginal entrepreneur is S_0 is $E(z|S_0) = \frac{\alpha}{\alpha-1}S_0$ while the average entrepreneurial skill when the skill of the marginal entrepreneur is S_1 and skills are downgraded by a proportion $\delta_m > 0$ is $E((1-\delta_m)z|S_1) = (1-\delta_m)\frac{\alpha}{\alpha-1}S_1$. Given that $S_1 = \frac{S_0}{1-\delta_m}$, $E((1-\delta_m)z|S_1) = \frac{\alpha}{\alpha-1}S_0 = E(z|S_0)$. That is, the average entrepreneurial skill does not change.

A.2. Additional figures and tables

See Figs. 9–13 and Table 14

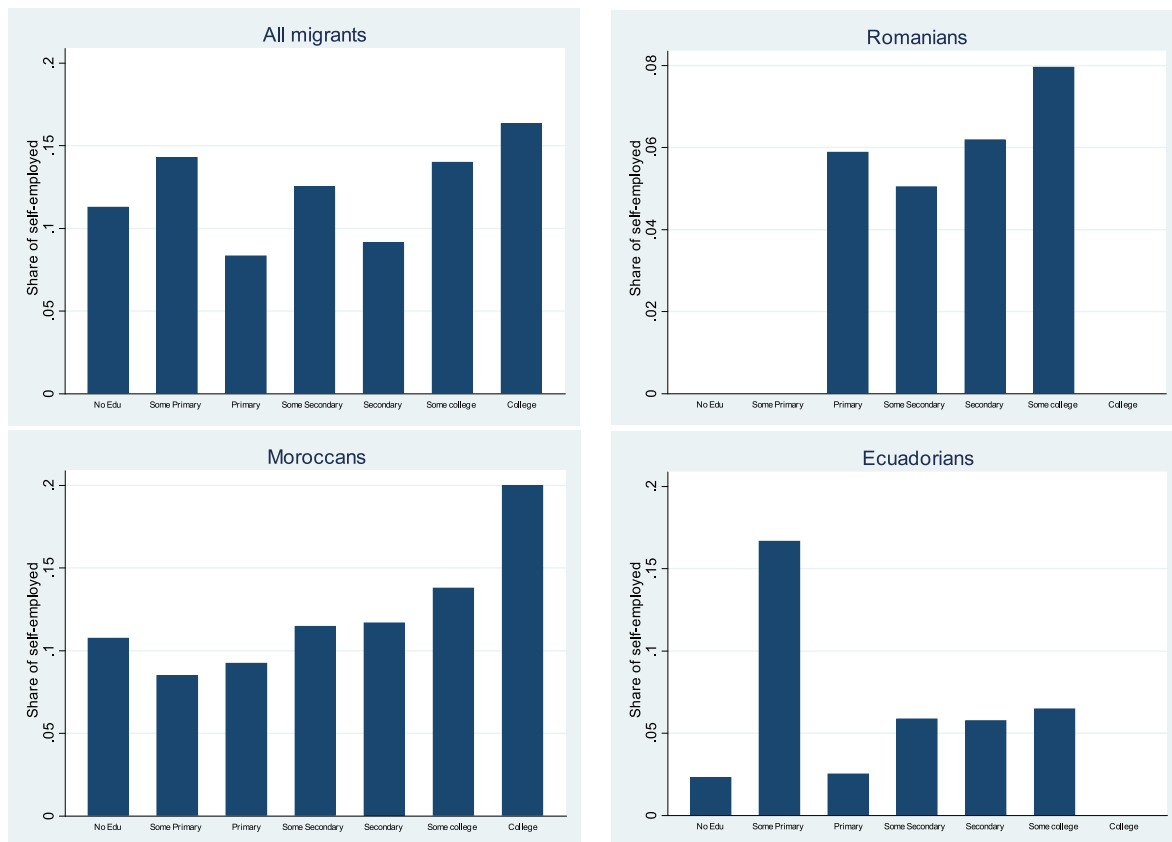


Fig. 9. Self-employment rates by educational level.

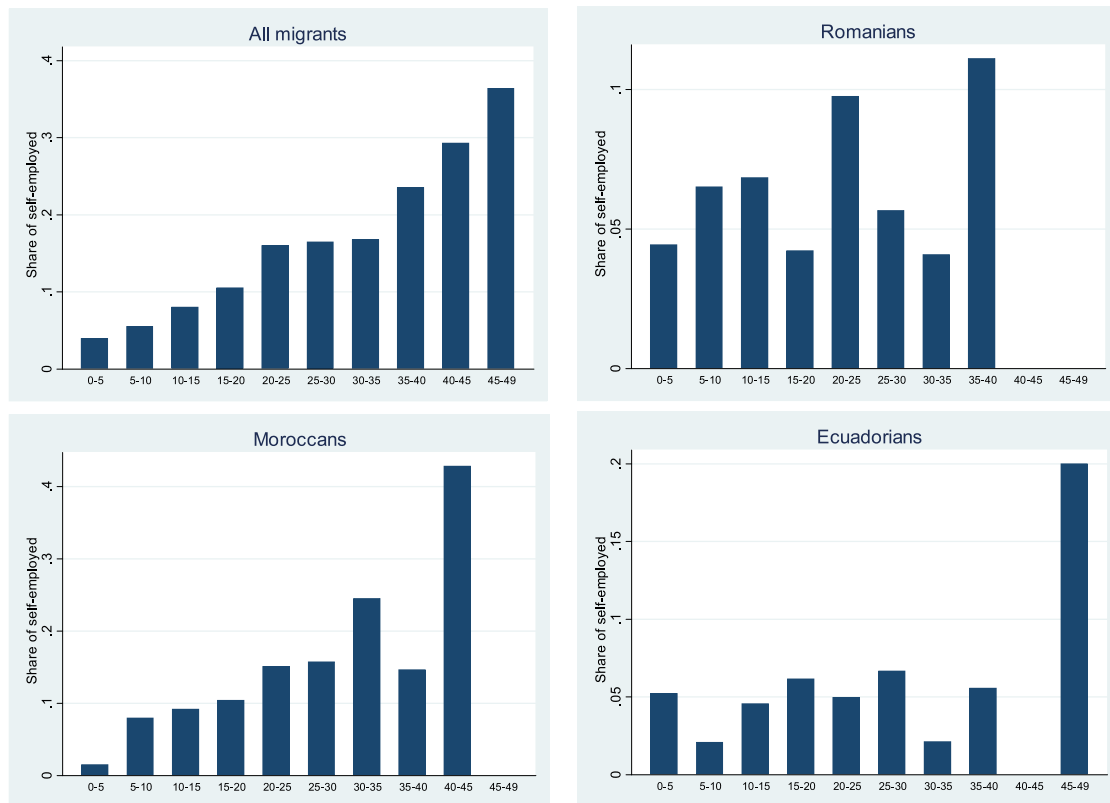


Fig. 10. Self-employment rates by years of effective experience.

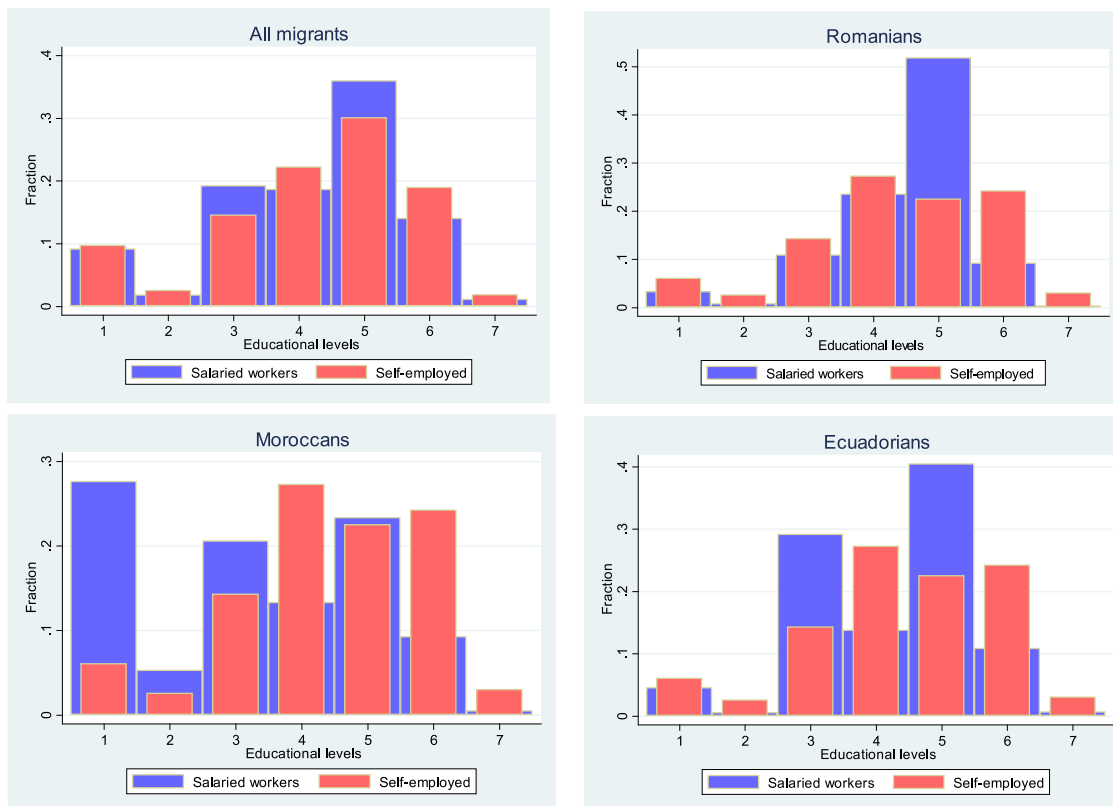


Fig. 11. Distributions of education of self-employed and salaried workers.

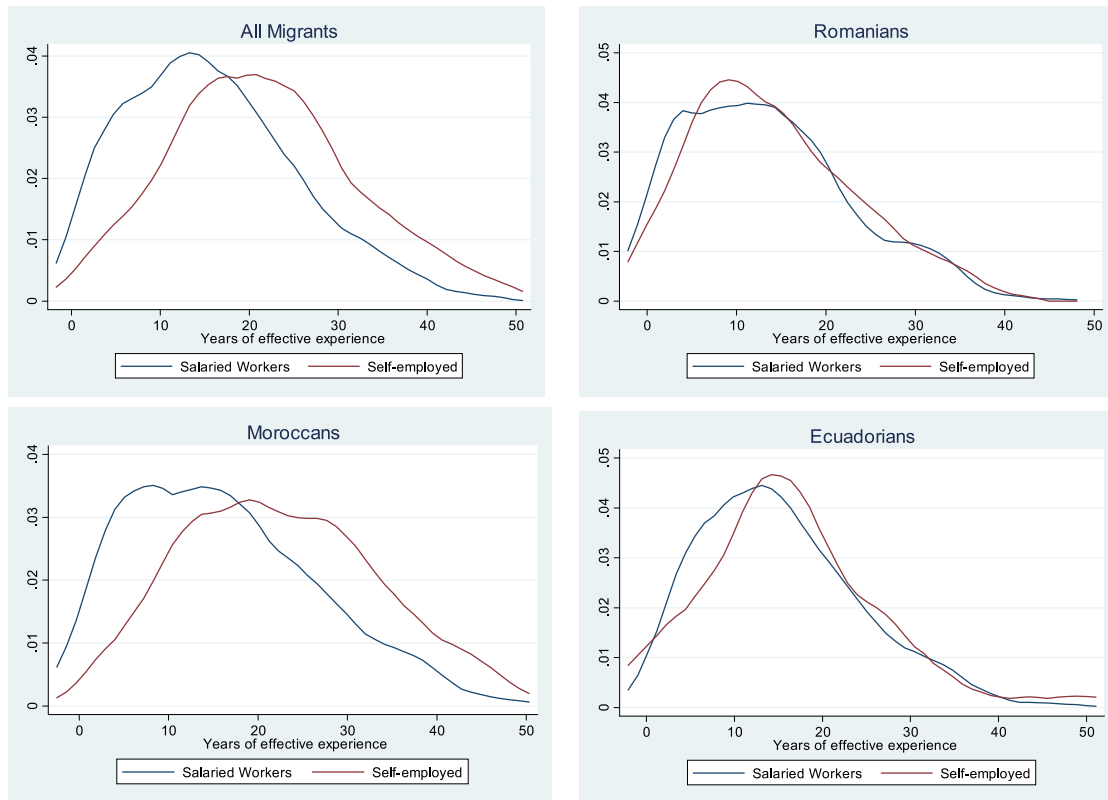


Fig. 12. Distributions of effective experience of self-employed and salaried workers.

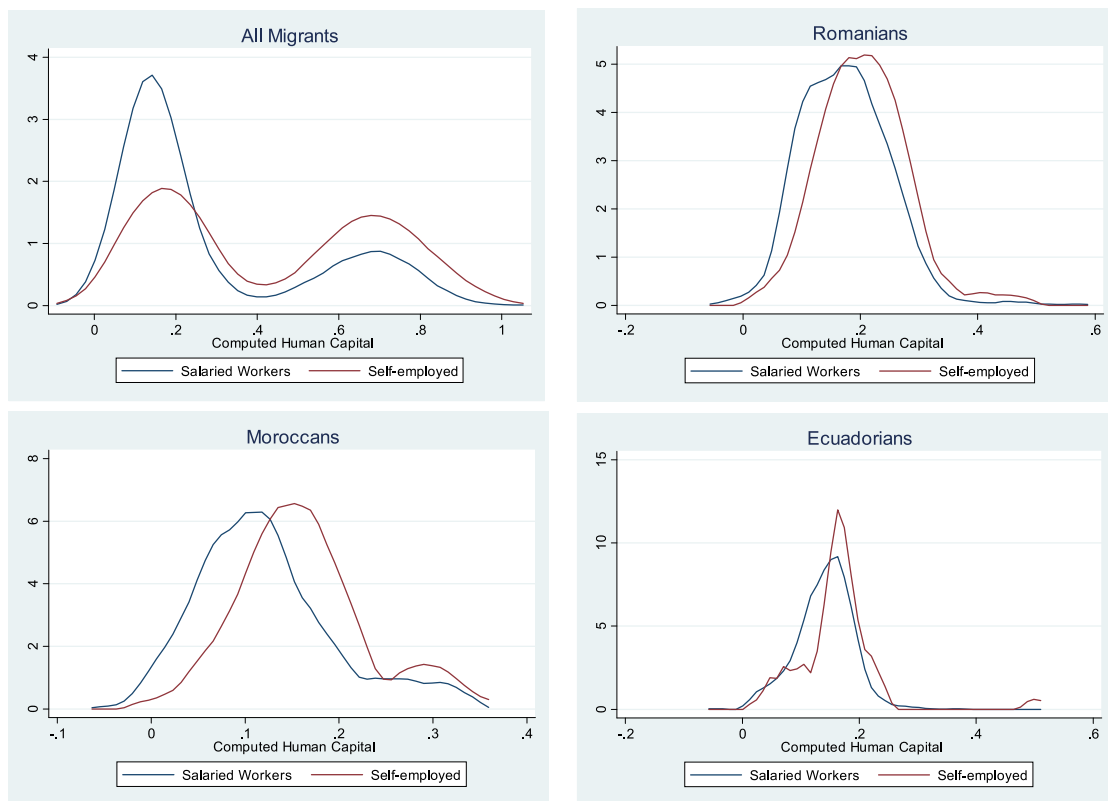


Fig. 13. Distributions of the computed measure of Human Capital of self-employed and salaried workers.

Table 14
Probit model on the propensity to self-employment - Average marginal effects.

	Baseline model
Schooling	0.006*** (0.002)
Effective experience in Spain	0.013*** (0.002)
Effective experience in CO	0.003* (0.002)
Dummy - Romanians (co2)	-0.109*** (0.038)
Dummy- Moroccans (co3)	-0.090*** (0.019)
Dummy- Ecuadorians (co4)	-0.120*** (0.031)
Years since arrival (YSA)	-0.004 (0.004)
Self-employed in CO	0.104*** (0.015)
Age	-0.003 (0.002)
Female	-0.032*** (0.011)
Married	0.013 (0.012)
Has children	0.009 (0.012)
Illegal status	0.043* (0.026)
Share of foreign-born in province	0.041 (0.182)
Pseudo R ²	0.191
Observations	3,961

Notes: Regression includes sector, region and arrival cohort fixed effects
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Data availability

Data will be made available on request.

References

- Adsera, A., Chiswick, B., 2007. Are there gender and country of origin differences in immigrant labor market outcomes across European destinations? *J. Popul. Econ.* 20 (3), 495–526.
- Adsera, A., Pytlikova, M., 2016. Language and migration. In: Ginsburgh, V., Weber, S. (Eds.), *The Palgrave Handbook of Economics and Language*. Palgrave Macmillan, London.
- Alden, L., Bastani, S., Hammarstedt, M., Miao, C., 2021. Immigrant-native differences in long-term self-employment. *Small Bus. Econ.*
- Amuedo-Dorantes, C., De la Rica, S., 2007. Labor market assimilation of recent immigrants in Spain. *Br. J. Ind. Relations* 45 (2), 257–284.
- Andersson, P., Wadensjö, E., 2004. Self-Employed Immigrants in Denmark and Sweden: A Way to Economic Self-Reliance? IZA DP No. 1130.
- Arango, J., Jachimowicz, M., 2005. Regularizing Immigrants in Spain: A New Approach. Migration Policy Institute, mimeo.
- Arrazola, M., de Hevia, J., 2008. Three measures of returns to education: an illustration for the case of Spain. *Econ. Educ. Rev.* 27, 266–275.
- Astebro, T., Chen, P., 2011. Stars and Misfits: Self-employment and labor market frictions. *Manag. Sci.* 57 (11), 1999–2017.
- Azoulay, P., Jones, B.F., Kim, J.D., Miranda, J., 2022. Immigration and entrepreneurship in the United States. *Am. Econ. Rev. Insights* 4 (1), 71–88.
- Bahar, D., Cowgill, B., Guzman, J., 2023. Refugee entrepreneurship: The case of venezuelans in Colombia. *AEA Pap. Proc.* 113, 352–356.
- Blanchflower, D.G., 2000. Self-employment in OECD countries. *Labor Econ.* 7, 471–505.
- Borjas, G.J., 1986. The self-employment experience of immigrants. *J. Hum. Resour.* 21, 485–506.
- Bosch, M., Carnero, A., Farré, L., 2010. Information and discrimination in the rental housing market: Evidence from a field experiment. *Reg. Sci. Urban Econ.* 40, 11–19.
- Burchardi, K.B., Chaney, T., Hassan, T.A., Tarquinio, L., Terry, S.J., 2020. Immigration, Innovation, and Growth. NBER Working Paper No. 27075.
- Cagetti, M., De Nardi, M., 2006. Entrepreneurship, frictions, and wealth. *J. Political Econ.* 114 (5), 835–870.
- Calvo-Armengol, A., Jackson, M., 2004. The effects of social networks on employment and inequality. *Am. Econ. Rev.* 94 (3), 426–454.
- Chiquiar, D., Hanson, G.H., 2005. International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *J. Political Econ.* 113 (2), 239–281.
- Chiswick, B., 1978. The effect of Americanization on the earnings of foreign-born men. *J. Political Econ.* 86 (5), 897–921.
- Chiswick, B.R., Miller, P.W., 2015. International migration and the economics of language. In: *Handbook of the Economics of International Migration*. North-Holland, pp. 211–269.
- Cicccone, A., Papaioannou, E., 2007. Red tape and delayed entry. *J. Eur. Econ. Assoc.* 5 (2/3), 444–458.
- Clark, K., Drinkwater, S., 2000. Pushed out or pulled in? Self-employment among ethnic minorities in England and Wales. *Labor Econ.* 7, 603–628.
- Clark, K., Drinkwater, S., Robinson, C., 2017. Self-employment amongst migrant groups: new evidence from England and Wales. *Small Bus. Econ.* 48, 1047–1069.
- Constant, A., Zimmermann, K., 2006. The making of entrepreneurs in Germany: Are native men and immigrants alike? *Small Bus. Econ.* 26, 279–300.
- Disney, R., Gathergood, J., 2009. Housing wealth, liquidity constraints and self-employment. *Labor Econ.* 16, 79–88.
- Dustmann, C., 1993. Earnings adjustment of temporary migrants. *J. Popul. Econ.* 6 (2), 153–168.
- Dustmann, C., Frattini, T., Preston, I.P., 2013. The effect of immigration along the distribution of wages. *Rev. Econ. Stud.* 80 (1), 145–173.
- Dustmann, C., Glitz, A., Schönberg, U., Brücker, H., 2016. Referral-based job search networks. *Rev. Econ. Stud.* 83 (2), 514–546.
- Evans, D.S., Jovanovic, B., 1989. An estimated model of entrepreneurial choice under liquidity constraints. *J. Political Econ.* 97 (4), 808–827.
- Fairlie, R.W., 2008. Estimating the Contribution of Immigrant Business Owners to the U.S. Economy, Small Business Administration Research Paper.
- Fairlie, R.W., 2018. Racial inequality in business ownership and income. *Oxf. Rev. Econ. Policy* 34 (4), 597–614.
- Fairlie, R.W., Lofstrom, M., 2015. Immigration and Entrepreneurship. CESifo Working Paper no. 5298.
- Fairlie, R.W., Meyer, B.D., 1996. Ethnic and racial self-employment differences and possible explanations. *J. Hum. Resour.* 31 (4), 757–793.
- Fernández-Huertas, J., 2011. New evidence on emigrant selection. *Rev. Econ. Stat.* 93 (1), 72–96.
- Friedberg, R., 2000. You can't take it with you? Immigrant assimilation and the portability of human capital. *J. Labor Econ.* 18 (2), 221–251.
- Guiso, L., Schivardi, F., 2011. What determines entrepreneurial clusters? *J. Eur. Econ. Assoc.* 9 (1), 61–86.
- Hart, D.M., Acs, Z.J., 2011. High-tech immigrant entrepreneurship in the United States. *Econ. Dev. Q.* 25 (2), 116–129.
- Hunt, J., 2011. Which immigrants are most innovative and entrepreneurial? Distinctions by entry visa. *J. Labor Econ.* 29 (3), 417–457.
- Hunt, J., Gauthier-Loiselle, M., 2010. How much does immigration boost innovation? *Am. Econ. J. Macroecon.* 2 (2), 31–56.
- Hurst, E., Lusardi, A., 2004. Liquidity constraints, household wealth and entrepreneurship. *J. Political Econ.* 112 (2), 319–347.
- Ioannides, Y.M., Loury, L.D., 2004. Job information networks, neighborhood effects, and inequality. *J. Econ. Lit.* 42 (4), 1056–1093.
- Kerr, S., Kerr, W., 2020. Immigrant entrepreneurship in america: Evidence from the survey of business owners 2007 & 2012. *Res. Policy* 49 (3), 103918.
- Kerr, W., Lincoln, W., 2010. The supply side of innovation: H-1B visa reforms and U.S. ethnic invention. *J. Labor Econ.* 28 (3), 473–508.
- Klapper, L., Laevena, L., Rajan, R., 2006. Entry regulation as a barrier to entrepreneurship. *J. Financ. Econ.* 82, 591–629.
- Lofstrom, M., 2002. Labor market assimilation and the self-employment decision of immigrant entrepreneurs. *J. Popul. Econ.* 15 (1), 83–114.
- Lucas, R.E., 1978. On the size distribution of business firms. *Bell J. Econ.* 9 (2), 508–523.
- Midrigan, V., Xu, D.Y., 2014. Finance and misallocation: Evidence from plant-level data. *Am. Econ. Rev.* 104 (2), 422–458.
- Munshi, K., 2003. Networks in the modern economy: Mexican migrants in the U.S. labor market. *Q. J. Econ.* 118 (2), 549–599.
- OECD, 2010. *Entrepreneurship and Migrants*. OECD Publishing, Paris.
- OECD/EU, 2018. *Settling in 2018: Indicators of Immigrant Integration*. OECD Publishing, Paris/European Union, Brussels.
- Peri, G., Requena, F., 2010. The trade creation effect of immigrants: evidence from the remarkable case of Spain. *Can. J. Econ.* 43 (4), 1433–1459.
- Peri, G., Shih, K., Sparber, C., 2015. STEM workers, H-1B visas, and productivity in US cities. *J. Labor Econ.* 33 (S1), S225–S255.
- Poschke, M., 2013. Who becomes an entrepreneur? Labor market prospects and occupational choice. *J. Econ. Dyn. Control.* 37, 693–710.

- Schjerning, B., LeMaire, D., 2007. Earnings, Uncertainty, and the Self-Employment Choice. Center for Economics and Business Research Discussion Paper 2007-04.
- Stuen, E., Mobarak, A.M., Maskus, K.E., 2012. Skilled immigration and innovation: Evidence from enrolment fluctuations in US doctoral programmes. *Econ. J.* 112 (565), 1143–1176.
- Uwaifo, R., Belton, W., 2012. Coming to America: Does having a developed home country matter for self-employment in the United States? *Am. Econ. Review: Pap. Proc.* 102 (3), 538–542.
- Wadhwa, V., et al., 2007. Education, Entrepreneurship and Immigration: America's New Immigrant Entrepreneurs. Ewing Marion Kauffman Foundation.
- Yuengert, A.M., 1995. Testing hypothesis of immigrant self-employment. *J. Hum. Resour.* 30 (1), 194–204.