Queen Mary
University of London

Queen Mary University of London
Registry Services
Research Degrees Office
Graduate Centre, Room GC.213
Mile End Campus
London, E1 4NS
arcs.qmul.ac.uk/research-degrees
Email: researchdegrees@qmul.ac.uk
Tel: +44 (0)20 7882 7474

31 May 2024

Name of Student	Gabriel Chaves Bosch
Student Number	200230904
Date of Birth	26 August 1995
Intended Award	Doctor of Philosophy
Programme of Study	Research in Economics
Mode of Study*	Full-time (minimum 24 weeks study per year)
Date of Entry to Programme	22 September 2021
Expected Date of Completion	22 September 2025
Term Time Address	Flat 46, Bath House 25 Dunbridge Street London E2 6JD
Home Address	Carrer Solsones 3 URBANITZACIÓ AIGÜES BONES, BUSTIA 497 Caldes de Malavella 17455 Spain

THIS IS TO CERTIFY that the above-named is a registered student reading for an award of Queen Mary University of London. **The degree certificate will be issued, by Queen Mary University of London, normally within 3 months of the date of the award.**

We issue this certificate, without liability, confirming the home and term-time addresses as provided by the student.
On enrolment all students sign a declaration certifying that the details they provide are correct.

T. J. Mason

Trudy Mason
Director of Registry Services

*For research students full-time study constitutes a minimum programme duration of one academic year which consists of 52 weeks less 30 days annual leave plus bank holidays and Queen Mary closure days. At least 21 hours per week is directed study. This includes students whose Mode of Study is "writing-up – prev FT".

Moving to (Create) Opportunity: The Impact of Immigration on Native Entrepreneurship*

Gabriel Chaves Bosch†

[*Preliminary and Incomplete. Please do not circulate.*]

June 14, 2024

Abstract

This paper examines the impact of immigration on native entrepreneurship, by using rich social security data and a unique immigration episode in Spain. Using variation across local industries and employing a modified shift-share instrumental variable for identification, I find immigration has a positive effect on native entrepreneurship. The effect is driven by the entry of new native entrepreneurs transitioning from wage work, and who tend to have above-median levels of education, previous wages and occupational skill. To understand the drivers of the effect, I propose and calibrate a model of occupational choice and immigration. The models encapsulates two key channels shaping natives' entrepreneurship decisions: (i) the impact of immigration on native wages and (ii) on potential profits for entrepreneurs. The latter channel emerges as the primary driver behind the increase in native entrepreneurship. The increased availability of cheaper immigrant labour decreases the opportunity cost of becoming an entrepreneur by raising potential entrepreneurial profits, particularly for more skilled natives.

*I am extremely grateful to my supervisors François Gerard, Simon Franklin and Marco Manacorda for their continuous guidance and encouragement on this project. For very helpful discussions and suggestions, I would also like to thank Jan David Bakker, Clément Imbert, Parag Mahajan, Tim Lee and Reem Zaiour, as well as seminar participants at Queen Mary University of London, Paris School of Economics Summer School, Junior Migration Seminar, SITES-Naples, RIEF Network, University of Essex Inaugural Applied Economics PhD Conference, and the Sciences Po CEPR Symposium. I thank the Spain's Dirección General de Ordenación de la Seguridad Social for providing data for this project.

†Queen Mary University of London. Email: g.chavesbosch@qmul.ac.uk

Immigration episodes are one of the most common types of labour supply shocks in developed economies. Contrary to the pessimistic and widespread belief that immigrants displace native workers and lower their wages, there is no consensus amongst economists on the employment and wage effect of immigration.[1] This lack of consensus arises because the characteristics of immigrant-induced labour supply shocks and the settings in which they occur lead to varied responses from native workers and firms. For workers, a key adjustment mechanism present in many settings is changes in occupational choice, where natives specialise in production tasks in which they have a comparative advantage relative to immigrants (Peri and Sparber, 2009; Foged and Peri, 2016). Similarly, a growing body of research studies the impact of immigration on firms, documenting changes in firm production structures and increases on firm entry (Dustmann and Glitz, 2015; Mahajan, 2022; Imbert and Ulyssea, 2024). A relatively less explored margin of adjustment that ties together these responses is business creation or destruction among natives. Immigration might affect natives' relative opportunity cost of becoming or ceasing to be entrepreneurs. This aspect is particularly important given the significant contribution of small and young firms to job creation (Haltiwanger et al., 2013).

Entrepreneurship is a relevant margin of adjustment to labour market shocks. While previous literature has documented the effect of labour demand shocks on entrepreneurship[2], less is known about whether and how labour supply shocks affect entrepreneurship. In the case of an immigration-induced labour supply shock, entrepreneurship amongst natives might increase or decrease due to several reasons. First, immigrant entrepreneurs might complement or substitute native entrepreneurs (Fairlie and Meyer, 2003; Duleep et al., 2021). Second, immigrant workers can impact native wages. If the impact is negative, entrepreneurship may serve as a form of self-insurance (Bohnet et al., 2021). Alternatively, if immigration raises native wages, natives may prefer to remain workers than start their own firm. Third, immigration might represent an opportunity to hire relatively cheaper labour.[3] Finally, labour supply shocks have general equilibrium effects operating through immigrant consumption that can affect entrepreneurship decisions of natives. In practice, all these channels can coexist, and thus determining the effect of immigration on native entrepreneurship remains an empirical question.

[1] For a review, see Dustmann et al. (2016) and Edo (2019).

[2] Babina (2019) and Hacamo and Kleiner (2022) document that weak labour demand pushes some individuals into entrepreneurship, as the opportunity cost of entrepreneurship falls during periods of distress.

[3] Even, at the same skill level, immigrants may be willing to accept lower salaries than natives, for instance due to migrations being temporary and/or consuming in the country of origin (Albert and Monras, 2018; Adda et al., 2022), or compensating differentials like human capital accumulation (Duleep et al., 2021).

This paper examines the impact of immigration on native entrepreneurship. I focus on Spain during the period from 1999 to 2008. This setting is ideal for two reasons. First, Spain experienced one of the largest post-war international immigration episodes among OECD countries, in which the share of immigrants amongst working-age individuals expanded from 2% to 14%.[4] This unexpected influx provides key quasi-experimental variation for identification (Fernández-Huertas Moraga et al., 2019). Second, I use high-quality administrative data with unique features. Spanish Social Security data, which captures both wage work and entrepreneurship spells, enable a detailed analysis of labour market transitions over extended periods, unlike rotating labour force survey data. Additionally, the Spanish population registry data provide reliable measures of international migration, including both documented and undocumented immigrants.

To determine the effect of immigration on native entrepreneurship, I leverage spatial variation in exposure to the immigration episode across Spanish local industries.[5] My empirical strategy employs a long-difference specification, comparing differences in native labour market outcomes between local industries with varying exposure to the immigration episode. The lack of immigration inflows before my study period and the use of the long-difference specification address dynamic sources of bias typically present in the migration literature (Jaeger et al., 2019). To minimise selection bias, I use a modified version of the traditional migrant networks shift-share instrumental variable (SSIV), pioneered by Card (2001), which accounts for country-of-origin advantages across industries and existing immigrant networks across provinces, thereby isolating exogenous immigrant inflows across local industries. I provide a battery of checks to ensure that the findings are robust to alternative variables, specifications and sample definitions, and that the SSIV does not suffer from the problems outlined in recent literature (Goldsmith-Pinkham et al., 2020).

To proxy entrepreneurs, I use self-employed individuals.[6] As self-employed in Spain pay pension contributions, they appear in the Social Security data. I follow Levine and Rubinstein (2016) and distinguish between unincorporated and incorporated self-employed. The latter are typically deemed as "high-quality" entrepreneurs.[7] Together with information on individual and previous work characteristics, I can provide an exhaustive characterisation of which type of entrepreneurs are more affected by the immigration episode.

[4]Figure 1 compares the magnitude of the shock to other countries usually studied in the labour market impact of immigration literature.

[5]These are defined using the 50 Spanish provinces and a classification of industries into 5 groups, a total of 250 units.

[6]I provide a detailed explanation of this measure in Appendix A.1.

[7]Levine and Rubinstein (2016) argue incorporated self-employed are closer to the perception of entrepreneurs as successful business owners, who usually engage in activities that demand comparatively strong non-routine cognitive abilities.

In the first part of the paper, I quantify and characterise the positive impact of immigration on native entrepreneurship. I find that the difference between the 25th and 75th percentile of exposure to the immigration episode results in an increase in entrepreneurship by 3% with respect to the baseline average employment in a local industry. To put this result in context, I analyse the impact on other labour market outcomes. Consistently with previous literature (Gonzalez and Ortega, 2011), I find no impact on native employment, which consists of both wage workers and entrepreneurs, and a small but positive impact on native wages.

After establishing the baseline result, I show that most of the increase on native entrepreneurship is primarily driven by incorporated entrepreneurs and high-educated individuals.[8] Incorporated entrepreneurship accounts for over 75% of the increase, while high-educated individuals account for around 83% of the rise, with significant overlaps between these groups. To understand whether this effect is due to higher entry rates into entrepreneurship, lower exit rates, or both, I analyse and decompose the effects into inflows and outflows across different labour market states.

The positive effect of immigration on entrepreneurship is predominantly explained by the entry of new native entrepreneurs. I find that immigration boosts inflows into entrepreneurship while having no effect on outflows. This effect is almost entirely driven by transitions from wage work to entrepreneurship, specifically from individuals who were wage workers in 1999 but became entrepreneurs by 2008. Observing these individuals as wage workers in the baseline year allows me to investigate which types of workers transition to entrepreneurship after the immigration episode. I categorise workers based on their baseline wages and a skill-based occupations classification. The results indicate that the inflow into entrepreneurship is driven by wage workers with above-median wages and those in high-skill occupations.

In the second part of the paper, I interpret the empirical results through a model of occupational choice to explore the potential channels. Since my data does not include information on the business entrepreneurs own, empirically identifying these mechanisms is not possible. To address this issue, I propose a simple model of occupational choice and immigration that considers two channels: (i) immigration affecting natives wages and (ii) immigration changing potential entrepreneurial profits for natives.[9]

[8] I divide individuals between high and low educated depending on whether their education attainment is above or below the median, respectively.

[9] The model abstracts from complementarities between immigrant and native entrepreneurs due to the lack of an increase in immigrant entrepreneurship found in the data and empirical results. The model also abstracts from immigrant increased demand and differing patterns of consumption, as these are netted out in the empirical section in order to focus exclusively on the immigration-induced labour supply impact.

In the model, natives decide whether to start firms, where they produce using other natives and immigrants as inputs of a constant elasticity of substitution (CES) production function, or to become wage workers. This choice depends on the relative value of wages versus potential business profits. Native wages are constant within the two levels of education: high or low. Entrepreneurial ability is drawn from a distribution that differs across education level. Immigrants provide labour inelastically and cannot become entrepreneurs. The main prediction of the model is that the impact of immigration will depend crucially on the complementarity or substitutability between immigrants and natives, and the distribution of entrepreneurial ability of natives. Under a large increase in immigration, depending on the complementarity or substitutability in production between immigrants and natives, native wages may increase or decrease. Immigrant wages decrease as immigrants are perfect substitutes with respect to themselves. Therefore, the impact of immigration on entrepreneurship depends on the opportunity cost of entrepreneurship, i.e. wages when working as an employee, and the potential entrepreneurial profits, which depend crucially on input prices, i.e. wages of both natives and immigrants. With this framework, I examine which of these two endogenous responses dominates.

I calibrate the model by minimising the distance between data moments and moments generated by the model. I use the treatment effects estimated in the reduced-form analysis, alongside with baseline moments from the data, to discipline the model. Hence, in the calibration process, I generate two equilibrium by changing the number of immigrants: one baseline equilibrium and a post immigration episode equilibrium. This mimicks the episode experienced in Spain from 1999 to 2008. From comparing the two equilibria, I obtain the dynamic moments that are matched with treatment effects. While the model is highly streamlined, it manages to capture the entrepreneurship responses documented in the empirical section.

Using the fitted model, I perform two counterfactual exercises. I either fix potential profits or native wages at baseline and then let the economy adjust to the immigration episode under these two settings. When fixing profits, most of the effect of immigration comes from changes in native wages, and vice versa. The counterfactual simulations show that the increase in entrepreneurship is explained by immigration lowering immigrant wages, which in turn raises potential business profits. Therefore, cheaper immigrant labour incentivises the creation of businesses that otherwise would not be profitable.

The insights of this quantitative exercise, along with the relevance of lower labour costs driving the increase in entrepreneurship, are consistent with empirical evidence. The entrepreneurship increase is mostly explained by incorporated and highly educated entrepreneurs, who are more likely to be employers,[10] as well as by the fact that immigrants were mostly absorbed in low-paying and low-skilled occupations.[11]

This paper contributes to two strands of literature. First, and more broadly, this paper builds on a substantial body of work that documents the impact of immigration on labour market outcomes of natives (Edo, 2019). Previous literature has looked at the effect of immigration on wages and employment of natives (Gonzalez and Ortega, 2011; Foged and Peri, 2016; Dustmann et al., 2017; Edo, 2019), as well as margins of adjustment such as internal migration (Piyapromdee, 2020) or changing patterns of occupational choice (Peri and Sparber, 2009; Amuedo-Dorantes and De La Rica, 2011). However, entrepreneurship as an additional labour market state has been largely overlooked and has typically not been considered as a possible margin of adjustment. I contribute to this literature by showing that immigration represents an opportunity for potential native entrepreneurs to hire relatively cheaper labour, and thus entrepreneurship constitutes a relevant margin of adjustment to immigration episodes. Moreover, by omitting this margin, previous literature may be overestimating the negative impact of immigration on employment.

My results also contribute to the literature on entrepreneurship, which tries to understand why people become entrepreneurs (Poschke, 2013; Levine and Rubinstein, 2016). In particular, labour market shocks may affect occupational choices. Babina (2019) and Hacamo and Kleiner (2022) document that weak labour demand pushes some individuals into entrepreneurship. However, less is know about whether and how labour supply shocks affect entry or exit to entrepreneurship. To the best of my knowledge, there are four papers that investigate the impact of immigration on entrepreneurship[12]. First, Duleep et al. (2021) argue that immigrants might facilitate innovation and entrepreneurship by being willing and able to invest in new skills, with potential positive spillovers on native entrepreneurship. Second, Unel (2022) distinguishes between unincorporated and incorporated entrepreneurs and finds that immigration has a positive effect on the supply of unincorporated entrepreneurs from 2000 to 2018 in the US, without a stance on the mechanism behind this effect. In contrast to these findings, Fairlie and Meyer (2003), using a different empirical approach, argue that immigration has a negative impact on native

[10] In 2008, two thirds of incorporated firms are employers, while only one third of unincorporated firms are employers, (INE, 2008).

[11] I show this for immigrants in the formal sector, but wages in the informal sector, whose size was substantial amongst immigrants in Spain (Bosch and Farré, 2014), are consistently lower (Elias et al., 2022), which means that the actual immigrant wages were, on average, even smaller than I show in this paper.

[12] In addition to these, Ajzenman et al. (2022) relate to this literature by showing that transit migration from refugees across Europe during 2010 to 2016 diminishes native entrepreneurship due to a decrease in risk-taking and confidence on institutions. However, it is difficult to establish to which extent the labour supply shock of transit migration is comparable to the labour supply shock analysed in these papers.

self-employment, and suggest that this is due to increased entrepreneurial competition from immigrants. Finally, Bohnet et al. (2021) analyse the Portuguese *retornados* episode and find a movement of male natives to solo self-employment, which they perceive as low quality entrepreneurship, and thus this movement represents a form of insurance for natives displaced by *retornados* with higher skills. Other papers complement this literature by focusing more generally on firms, finding a positive effect of immigration on firm creation (Olney, 2013; Dustmann and Glitz, 2015; Beerli et al., 2021; Mahajan, 2022)[13]. My findings contribute to this literature by providing a first attempt to characterise in detail the effect of a labour supply shock, in this case driven by immigration, on native entrepreneurship, thus providing a direct link between the impact of immigration on workers' occupational choice and firm creation.

The paper is organised as follows. Section 1 describes the context, data, and descriptive statistics. Section 2 presents the empirical strategy. Section 3 provides the empirical results, and Section 4 discusses the results and provides a formalises a simple model to explain the mechanism behind the results. Section 5 concludes.

1 Context, Data, and Descriptive Statistics

1.1 Context on the Immigration Episode and the Spanish Economy

The immigration episode. Spain experienced a massive immigration inflow from 1999 to 2008. During this period, the number of immigrants increased from less than a million to more than 5 million, over a baseline population of 40 million. The magnitude of this inflow makes it the largest immigration episode in the post-war period in any OECD country with the exception of Israel in the 1990s. In Figure 1, I provide a comparison of the immigration episode with respect to other countries usually studied in the migration literature. This figure shows two striking facts. The first is the magnitude of the episode, expanding the immigrant share of population from slightly above 2 per cent in 1995 to above 13 per cent in 2010. The second fact is that before the late 1990s, Spain was a country with very little immigration experience. Immigration flows into Spain during the 1980s and before were practically zero (Ortega and Peri, 2013). The reasons behind such a sharp increase in immigration during the 1999 to 2008 period are a combination of pull and push factors. The main pull factors were the economic growth of the Spanish economy, the ease of entering Spain, and the labour demand increase in tourism, hospitality, and construction

[13]Note firms are not started exclusively by natives. Indeed, in the case of the US, Mahajan (2022) finds a negative effect of immigration on the count of native-owned establishments and a positive effect on the count of immigrant-owned establishments.

industries, which offered mostly low-skilled and temporary jobs. Since Spain received immigrants from all over the world, the list of country-specific push factors is extensive. Some notable examples are the late 1990s crisis in Latin American countries, tightness in US immigration policy, proximity to Africa, and the European Union expansion. To sum up, the suddenness and magnitude of this episode make it a unique opportunity to understand the impact of immigration on native outcomes.

Immigrants in the labour market. Immigrants during this episode were characterised by working in unskilled and low-paying jobs[14]. Immigrants had larger rates of participation in the informal sector when compared to natives (Bosch and Farré, 2014), and suffered from substantial occupational downgrading (Simon et al., 2014). This created a native-immigrant job disparity, which also entailed a wage disparity. To illustrate this, Figure 2 compares distributions of native and immigrant wages before and after the episode. The distribution of wages changes substantially amongst immigrants at the end of the episode, due to a change of composition from the immigrants who entered during the study period. In Table 1, I compare immigrants and natives aged 20 to 60 in the formal sector by the end of the immigration episode. Consistently with the previous graphical evidence, average wages of immigrants are substantially lower when compared to natives[15], with the difference likely being a lower bound due to the higher participation of immigrants in the informal sector. Taken together, this evidence is consistent with immigrants performing different jobs than natives and thus competing only with natives in low-skill occupations, if at all.

There is a large body of papers documenting the effects of immigrants on the Spanish labour market.[16] When it comes to the effect of immigration on the labour market, the literature agrees on a negligible effect of immigrants on employment and wages of natives from the 1999-2008 immigration episode. Using a spatial correlations approach and by focusing on the 2001 to 2006 period, Gonzalez and Ortega (2011) find no sizeable effect of immigration on wages nor employment of natives. However, Amuedo-Dorantes and De La Rica (2011), by following closely the work by Peri and Sparber (2009) and focusing on the 2000 to 2008 period, show that this negligible employment effect amongst natives masks important relocation towards relatively less manual-intensive occupations. In particular, since immigrants in Spain specialised in relatively more manual tasks, which are

[14]Examples include mostly manual jobs such as construction labourers, waiters, cleaners, caregivers or farm workers.

[15]In a Mincerian regression of wages on sociodemographic characteristics characteristics, I find that being immigrant vis-a-vis being a native has a substantial negative effect on wages, even after controlling for a large set of covariates such as age, tenure, gender, occupational skill (low, medium, high), industry and location.

[16]See De La Rica et al. (2014) for a review.

usually more common in low-skilled occupations, this lead natives to relocate to jobs with a lower content of manual tasks and in which they had a comparative advantage. Additionally, Amuedo-Dorantes and De La Rica (2011) show corporate managers, managers of small enterprises and other professionals are amongst the less manual occupations, and these are more likely than other occupations to be self-employed.

Entrepreneurship amongst natives and macroeconomic context. The period of analysis saw a sharp increase in the number of native entrepreneurs. Table 2 shows that in my sample, composed of natives born between 1954 and 1979, the number of entrepreneurs increased by 86%, compared to a 19% increase in the number of wage workers. This employment growth was fueled by a period of buoyant economic growth, with an average yearly GDP growth of 3.5% during this period. Spain experienced rapid economic growth since the economic and political stabilisation that followed the 1992-1993 crisis, until the country was hit by the Great Recession in 2008.

1.2 Data

I use four sources of data to study the impact of the 1999-2008 immigration episode on native entrepreneurship in Spain: administrative Social Security data on individuals working lives; administrative population registries; labour force survey data; and data from the 1991 population Census. In Appendix A.2 I provide additional information on the data.

A. Muestra Contínua de Vidas Laborales (MCVL). First, I use administrative social security data that includes the working lives for a representative 4% sample of individuals enrolled in the Social Security system. These data include information on working lives of individuals starting in 1966. The data provide detailed daily information on all working spells, including earnings, affiliation type (wage worker, self-employed), incorporation status if self-employed, occupation, industry, as well as socioeconomic variables such as date of birth, gender, education. Importantly, and in contrast with administrative data sets from other countries, a key feature of these data is that they include information on self-employed (see Appendix A.1 for more information on the self-employed definition).

B. Padrón Contínuo. Second, I use administrative data from the population registry for the period from 1999 to 2008. These micro data include information on all people registered as living in a certain town at the beginning of each year. This represents the universe of individuals living in Spain. Regardless of documentation status, immigrants are encouraged to register in order to obtain access to public services such as healthcare. Hence, its universal coverage is key to tracing immigrant stocks across time and provinces.

C. Other data. I complement these sources of data with data from the Encuesta de Población Activa (EPA), the Spanish Labour Force Survey, for the years 1999 to 2008. The survey nature of these data implies that informal workers are also captured, which allows me to calculate the shares of immigrants across different sectors, and to construct population-wide control variables. Finally, since the instrument uses immigrant networks existing well before the immigration episode, I use data from the 1991 Census on the number of immigrants per province and nationality of origin in the late 1980s.

1.3 Sample and Descriptive Statistics

To facilitate the discussion of the descriptive statistics, in this section I first discuss how I construct the main native labour market outcomes sample, the immigration episode measure, and the main dependent variables of interest.

Analysis sample. To construct the main sample of analysis, I first build a panel of yearly observations at the individual level for the period from 1999 to 2008 from the MCVL. To create this panel, I use information on spells for native individuals, born between 1954 and 1979[17], and who were employed at least one year for a minimum of 100 days. I omit workers with missing place or date of birth, or province of residency. Individuals are classified as either wage workers, unincorporated self-employed, incorporated self-employed, or not employed, according to their main source of earnings for each year. Descriptive statistics on this micro data are provided in Table 2. I aggregate this information at the year by province and sector level to obtain labour market outcomes for native workers[18]. I consider the 50 provinces of Spain[19] and an industry classification into 5 groups: agriculture, manufacturing, construction, retail and hospitality, and other services. Therefore, I end up with a sample that contains information of native labour market outcomes across 250 local industries. I refer to this sample as the analysis sample.

Construction of the immigration episode variable. To calculate the magnitude of the immigration episode and construct the migrant networks instrument, I use data from the Padrón Contínuo, the 1991 Census, and the labour force survey, EPA. For each province p and year t, I use the Padrón Contínuo to calculate the number of immigrants and natives aged 20 to 60, which I consider the working age population. I denote this quantity WAP_{pt}. Then, to obtain a proxy for the change in immigration in each local industry, I take

[17] These are the equivalent to the baby-boom generation in Spain, represent the majority of the workforce throughout the period, and since they are highly attached to the labour market during these years, they are the most affected by the immigration episode of the 1999-2008 period.

[18] In the analysis I focus exclusively on the years 1999 and 2008, as I use a long-differences specification.

[19] I exclude the autonomous cities of Ceuta and Melilla, located in Northern Africa.

the number of immigrants aged 20 to 60 in each province and year, also from the Padrón Contínuo, N_{pt}^{F}, and I multiply it by the share of immigrants in province p at year t that work in industry i, ω_{ipt}. This gives me $N_{ipt}^{F} = N_{pt}^{F} \times \omega_{ipt}$, a proxy of the number of immigrants working[20] in province p, industry i and year t. With these quantities, I construct the immigration episode, which is the change in the number of immigrants during the study period in a local industry normalised by the province baseline working age population:

$$\Delta\text{Immigration episode}_{ip} = \frac{N_{ip,2008}^{\text{F}} - N_{ip,1999}^{\text{F}}}{WAP_{p,1999}} \tag{1}$$

This represents the explanatory variable used throughout this paper[21]. I divide by province working age population rather than a proxy of working age population associated with each local industry because some local industries are rather small in the EPA in 1999, leading to unreasonably noisy estimates of the immigration episode. To keep this measure consistent with the dependent variables, I normalise (changes in) native outcomes at the local industry level by province-level baseline native employment. Finally, to relate this measure with previous literature, note that by summing the immigrant episodes across local industries within a province gives the change in the number of working age immigrants over working age population for that province, which is the usual immigrant episode variable used in papers using an spatial correlations approach (Dustmann et al., 2016), such as Sanchis-Guarner (2023) and Ozguzel (2021).

Construction of the dependent variables. The main variable of interest is the change during the period in the number of native entrepreneurs in province p, industry i and year t, E_{ipt}^{N} , normalised by province p baseline native employment, which is the sum of wage workers and entrepreneurs. I calculate this variable using the native labour market outcomes sample as follows:

$$\frac{E_{ip,2008}^{N} - E_{ip,1999}^{N}}{Employed_{p,1999}^{N}} \tag{2}$$

I construct other variables on native labour markets analogously. I keep the normalisation by baseline province employment across different outcomes, except for changes in wages, which I construct as log differences between 1999 and 2008.

Descriptive statistics. In Table 3, I provide descriptive statistics for the main variables used in the analysis, computed using the analysis sample[22]. Across local industries, the

[20] In the unrealistic case that all immigrants were actually working, this would not longer be a proxy.

[21] In Section 2, I explain how I construct an instrumental variable for this quantity, which requires data from the 1991 Census.

[22] Occupational choices change along the life-cycle (Humphries, 2021), with self-employment becoming

number of both native entrepreneurs and wage workers grew, as normalised by baseline province population. The growth in the number of native entrepreneurs accounts for roughly 30% of employment growth, but since the baseline share of native entrepreneurs amongst employed individuals is around 12%, this implies a compositional change, making the share of entrepreneurs across local industries grow by six percentage points, up to 18% by 2008. Within entrepreneurs, the number of both unincorporated and incorporated entrepreneurs grow, but since incorporated entrepreneurs represent a lower share, their growth eventually changes the composition of entrepreneurship, increasing the share of incorporated entrepreneurs across local industries by 4 percentage points, from 34 to 38%. The change in log daily wages (in 1999 euros), obtained as residuals from a regression of wages on quadratic profiles of age and tenure, and occupation and year fixed effects, is negligible and not statistically different from zero. In terms of the immigration episode, the increase in immigrants in the average local industry increases the working age population in the province to which the local industry belongs by around 4.4% with respect to baseline. This growth is substantial, as already noted in Figure 1. In fact, during this period, the working age population increased by 20%, and three-thirds of this increase is accounted by immigrants.

Is this labour supply shock associated with native employment growth, and more particularly, growth in native entrepreneurship? In Figure 3 I provide scatterplots of the increase in native wage workers and entrepreneurs with respect to the immigration episode, across local industries. While the correlation of the immigration episode and the increase in native wage workers is positive and statistically significant from zero, the association is graphically weak, potentially driven by larger local industries, and with two clusters of industries experimenting either growth or stagnation, independently of the immigration episode. In contrast, the relationship between native entrepreneurship and the immigration episode is clearly positive.

2 Empirical Strategy

My basic estimation equation regresses native labour market outcomes O_{ip} on the immigration episode across local industries defined by the interaction of province p and indus-

more prevalent after turning 30 years old. The results are robust to recomputing the descriptive statistics using people aged 20 to 60 over the period, and hence they are not driven by the 1954-1979 cohort being older by 2008.

try group i,

$$O_{ip} = \beta \Delta \text{Immigration episode}_{ip} + \gamma' X_{ip,1999} + \gamma_p + \gamma_i + \epsilon_{ip}, \tag{3}$$

where the dependent variables O_{ip} are differences in stocks between 1999 and 2008 in the local industry ip normalised by province p population, as in Equation 2, or changes in log wages. ΔImmigration episode$_{ip}$ is defined as in Equation 1. $X_{ip,1999}$ is a vector of baseline controls at the local industry level which includes the native share of high education, share of males, share of entrepreneurs, as well as the proportion of national employment in industry i accounted by the local industry, as well as the immigrant share. Then, γ_p and γ_i are province and industry fixed effects. Finally, ϵ_{ip} is the random error term. The main parameter of interest is β, which captures the effect of a one-percentage point increase in the immigration episode on native labour market outcomes.

Employing a long-differences specification across local industries has two main advantages. First, by using long-differences, I can take care of dynamic sources of bias, which would be present if I used a stacked regression with multiple shorter time periods and province and time fixed effects, and which are typical in the immigration literature (Jaeger et al., 2019). Also, since my analysis sample follows the same cohort, the impact of compositional changes is minimised. Second, by splitting the sample by provinces, which roughly proxy local labour markets in Spain[23], and industries, I obtain more variation but also the opportunity to control for industry and province fixed effects. Province fixed effects are particularly important, which prevent β from capturing a demand-driven response due to general equilibrium effects of immigration (Mahajan, 2022), such as immigrant consumption, and allow me to identify the effect of the immigrant-induced labour supply shock.

As is well known, OLS estimation of Equation 3 suffers from endogeneity concerns, because unobserved local labour demand shocks might influence both the location decisions of immigrants and the natives' labour market outcomes. This leads to a bias in the estimated $\hat{\beta}$ coefficients. To deal with this issue, I instrument the immigration episode using a modified version of the migrant networks instruments[24]. In its traditional version,

[23]Spanish provinces, designed by Javier de Burgos in 1833, roughly follow a Voronoi diagram, by which the edges of each province are equidistant from the province capital in each side. Since province capitals, and their metropolitan areas, are in most cases the most populated areas of each province, this implies spatial spillovers of shocks across provinces are limited. Thus, provinces serve as a reasonably good approximation to local labour markets.

[24]The traditional migrant networks instrument, pioneered by Altonji and Card (1991) and Card (2001), has also been used in the Spanish setting by Gonzalez and Ortega (2011, 2013), Sanchis-Guarner (2023), Fernández-Huertas Moraga et al. (2019) and Ozguzel (2021).

this instrument consists in predicting the location decisions of immigrants using previous settlement patterns, as immigrants from each nationality are more likely to locate wherever there are more immigrants from the same nationality. I use a similar modification as Mahajan (2022) and leverage the idea that immigrants from certain origins tend to have comparative advantages in certain occupations and industries (Kerr and Mandorff, 2023) in order to allocate immigrants across industries within each province, by distributing them across industries according to their origin baseline distribution across industries.

To construct the instrument, I first consider a predictor Z_{ipt} of the stock of immigrant population aged 20 to 60 in province p, industry i and year t:

$$Z_{ipt} = \sum_{c} \left(\frac{FB_{c,p,1991}}{FB_{c,1991}} \right) \times FB_{ct} \times \omega_{ir(c),1999} \tag{4}$$

The components of this predictor are the following. $FB_{c,p,1991}$ is the number of immigrants from country or region c in province p 1991. $FB_{c,t}$ is the total number of immigrants across Spain from country or region c and year t. So far, these are the traditional elements used in the construction of the migrant networks instrument, and the idea is to distribute the total number of immigrants in the current year t across local labour markets according to their country-specific distribution in year 1991. However, since I use industry variation within each local labour market, I distribute immigrants across industries within each province using the baseline distribution, in year 1999, of immigrants according to their region of origin $r(c)$ at the national level[25]. Since foreign-born population data from 1991 comes from the Census, the subscript c can takes values from a classification of countries with 17 different countries/regions of origin[26], while data on the baseline distribution across industries in 1999 comes from the labour force survey, so $r(c)$ takes values from a coarser classification of 8 different possible regions of origin. Finally, the instrumental variable is constructed as the change in predicted immigration in a local industry divided by the predicted number of working age immigrants plus working age natives in the province, analogously as in Equation 1:

$$\Delta\widehat{\text{Immigration}}\,\text{episode}_{ip} = \frac{Z_{ip,2008} + Z_{ip,1999}}{Z_{p,2008} + WAP^{native}_{p,1999}}$$

The validity of the instrument relies on the exclusion restriction by which both the distribution of immigrants in 1991 and the distribution of immigrants across industries in

[25]By using the baseline distribution at the national level at baseline in order to obtain a share that is exogenous to shocks across local industries.

[26]This classification is constructed using the overlaps in the classification from both the 1991 Census and the Padrón Contínuo.

1999 impact yearly changes in outcomes over the 1999 to 2008 period only through its effect on predicted population changes, conditional on fixed effects and controls. Therefore, the main identification assumption is that local industries with highest exposure to shifts, as distributed by the shares, do not have systematically different potential outcomes than local industries with lower exposure to these shifts, conditional on fixed effects and controls.[27] While this assumption is inherently untestable, in Figure 4 I provide supporting evidence in favour of it, by showing that pre-period outcomes are not affected by the immigration episode in simple reduced form regressions. Pre-period outcomes before the immigration episode are not statistically affected by the instrument. All study period outcomes except the change in native entrepreneurship are also not affected, but this is consistent with existing evidence (see Section 3). Moreover, the coefficients from the regressions of the instrument on the increase in native entrepreneurship before the period and during the study period are statistically different, with the test of the difference in coefficients having a p-value of 0.001. Additionally, the lack of effect on outcomes in the pre-period is to be expected given how sharply immigration increased after 1999 and the virtual lack of immigration in preceding periods. This also mitigates the concerns from Jaeger et al. (2019) that the instrument picks up responses to previous immigration episodes.

Finally, I show the instrument is relevant by exploring the identifying variation in the first stage. In Figure 5 I provide scatter plots of the immigration episode on the instrument, naively in Panel (a) and then netting out covariates and fixed effects in Panel (b). In both cases, there is a clear positive relationship, and even after netting out covariates and fixed effects there is enough residual variation in the instrument in order to identify the reasonably exogenous variation in the immigration episode. In Table 4 I show how the first stage coefficient remains significant after the inclusion of controls and fixed effects. In Column (4), the first-stage F-statistic is 23.11, well above the $F = 10$ cutoff.[28]. Therefore, the instrument satisfies relevance and displays useful identifying variation

3 Empirical Results

This section presents the empirical results, using the identification strategy outlined above.

[27]This implies I embrace identification in terms of the exposure shares, consistent with Goldsmith-Pinkham et al. (2020) and in contrast to Borusyak et al. (2021). However, there is a plethora of independent push factors across different country groups that contribute to the exogeneity of shocks for some countries of origin, which is key to create enough variation in the instrument across local industries.

[28]Throughout the empirical results, I provide the Kleibergeen-Paap rk Wald F statistic, which in the case of one instrument and one endogenous regressor is equivalent to the first-stage F-statistic. The associated 10% maximal IV size critical value is 16.38, and hence the F-statistic still remains above, which provides further confirmation on the relevance of the instrument.

3.1 Native Entrepreneurship, Employment, and Wage Effects

Table 5 provides the results of estimating β from Equation 3 for a set of native employment and wage outcomes. For completeness, in panels A and B, I estimate β using OLS, while in panels C and D, I use the 2SLS estimator consistent with my empirical strategy. To assess the robustness of estimates to controls, in panels A and C, I estimate the model omitting baseline controls. Neither instrumenting nor adding controls makes a substantial difference in estimates, suggesting that there is no strong selection of immigrants into particular local industries. Hereinafter, I use the specification from Panel D.

The main result is that immigration has a positive impact on native entrepreneurship. Column (4) of Table 5 reports the impact of immigration on the change in the number of entrepreneurs. I find a positive effect of the immigration episode on the change in the number of entrepreneurs. In particular, a one percentage point increase in the immigration episode–which corresponds to 33 percent of a standard deviation– increases the growth in the number of entrepreneurs from a given local industry, as normalised by the baseline number of employed workers in the province, by 0.23 percentage points (a 7.8 percent increase from the mean of 2.94 percent). Consistently with previous literature using variation at the province level[29], I do not find an effect of immigration on native employment, which provides further validation of my empirical strategy. In Column (1) I look at overall employment, considering both wage workers and entrepreneurs, while in Column (2) I focus on wage workers. In both cases, the estimates are not statistically significant. Finally, in Column (6) I report the results for wages, where I find a small but positive effect at the 10% significance level.

Is the native entrepreneurship effect large? I calculate that the difference between the 25th and 75th percentile of exposure to the immigration episode results in an additional increase in the number of entrepreneurs of 3% with respect to baseline employment in a local industry in 1999. [30]

Finally, a word of caution regarding the comparability of the employment and wage effects documented in this section. The use of spatial variation at the local industry level would yield similar results to using variation at the province level in the absence of spillovers.

[29] Using the same immigration episode but different identification strategies, both Amuedo-Dorantes and De La Rica (2011) and Gonzalez and Ortega (2011) find no employment effect, while the later find also no wage effect.

[30] This is calculated as $\frac{\hat{\beta}\times(X_{P75}-X_{P25})\times Denom(Y)}{Employed_{1999}}$ where $\hat{\beta}=0.232$ is the estimated regression coefficient, $X_{P75}=0.0627$ and $X_{P25}=0.018$ are the 75 and 25 percentiles of the immigration episode, $Denom(Y)=19476$ is the weighted average of the denominator of Y, at the province level, and $Employed_{1999}=6453$ is the weighted average employment across local industries in 1999.

However, there is some reallocation of workers across industries within a province.[31] Therefore, the employment and wage estimates in this paper are not meant to reply to the question "what is the overall effect of immigration on native wages?" (Dustmann et al., 2016). Instead, the employment and wage effects are instrumental to understand how occupational choice by natives within each local industry reacts to a (immigration-induced) labour supply shock.

3.2 Heterogeneity

The positive effect of immigration on native entrepreneurship raises the question of which type of natives are driving this increase. Immigrants may impact entrepreneurship by impacting the opportunity cost of entrepreneurship, i.e. native wages, or potential profits, as they may impact input prices, i.e. wages of both natives and immigrants themselves. Generally, the impact of immigration on immigrant wages is negative due to the own-price elasticity being negative.[32] However, the impact of immigration on native wages depends crucially on the patterns of complementarity or substitutability. This motivates an analysis by skill, as skill levels may be affected differently due to their differing patterns of substitutability or complementarity with respect to immigrants.

Table 6 provides a decomposition of the main results using education as a measure of skill. Across education levels, there are no effects of immigration on employment or wage worker levels. However, amongst low educated wage workers (those with secondary education or less), there is a positive wage effect. This suggests immigrants are complementary to low educated wage workers. When it comes to entrepreneurship, there is a positive effect on entrepreneurship from all educational levels, although only statistically significant amongst high-educated individuals. While at baseline people with high education represent around half of the sample (see Table 2), they account for more than 83% of the positive effect of immigration on native entrepreneurship.

Previous literature has found that higher education attainment is correlated with high-quality entrepreneurship (Levine and Rubinstein, 2016, 2020). This is consistent with the results being driven by incorporated entrepreneurship–typically used as a proxy for high-quality entrepreneurship–, which accounts for 75% of the increase in native entrepreneurship, as already indicated in Columns (3) to (5) of Table 5. When zooming into education, most of this increase in incorporated entrepreneurship is explained by high educated indi-

[31] Internal reallocation across local areas after an immigation episode also happens in many settings considering only provinces or subnational areas rather than local industries, such as in the US (Monras, 2020).

[32] This is a consequence of assuming perfect substitutability, which is reasonable in this setting. Figure 2 shows average immigrant wages decrease during this period.

viduals, as Column (5) from 6 shows. While the effect is driven predominantly by high educated individuals transitioning to incorporated entrepreneurship, there is also part of the effect explained by high-educated individuals becoming unincorporated entrepreneurs, as well as some low-educated individuals becoming incorporated entrepreneurs.

3.3 Entrepreneurship flows

Comparing cross-sectional quantities misses important dynamic adjustments explaining the effect of immigration on native entrepreneurship. Therefore, the main estimates alone cannot determine whether the effect is driven by increased entry of new entrepreneurs in local industries receiving more immigrants, or alternatively, by lower rates of entrepreneur exit in these industries. In this subsection, I leverage the unique feature of the Spanish administrative data, which records both periods of wage work and entrepreneurship.[33]

The panel dimension of my data allows me to track workers from the time they start paying pension contributions. Most individuals begin contributing to pensions when they start working. By focusing on those born between 1954 and 1979, I ensure that these individuals are between 20 and 55 years old during the study period, making them highly attached to the labour market, thus minimising concerns about selection into the labour force. This enables me to investigate transitions between labour market states over the immigration period and better understand how immigration impacts natives' occupational choices.

I define flows as the number of people transitioning between different labour market states, which include entrepreneurship, wage work, and non-employment. Additionally, I consider transitions from entrepreneurship across local industries to achieve an exact decomposition of the total entrepreneurship effect. In the regressions, the dependent variable is defined as "the number of people in a given labour market state in 2008 who were in a different labour market state in 1999," normalised by the baseline number of employed people in the province, as described in Equation 2. This approach enables me to decompose the estimates from the previous sections into contributions by different flows.[34]

Table 7 decomposes the impact of immigration on entrepreneurship into inflows and outflows to and from other labour market states. The first column provides the increase in entrepreneurship, which is the same as in Column (3) from Table 5, for reference. Columns (2) to (4) refer to inflows and columns (5) to (7) refer to outflows. Column (3)

[33] The reason behind this is that self-employed individuals in Spain pay public pension contributions. For more information on this characteristic of the data, see Iraizoz-Olaetxea (2022).

[34] For instance, the change in the number of entrepreneurs between 1999 and 2008 can be decomposed into people who were entrepreneurs in 2008 but not in 1999 minus people who were entrepreneurs in 1999 but not in 2008.

shows the main contributor to the increase in entrepreneurship, namely, flows from wage work to entrepreneurship. Comparing inflows to outflows in Panel A shows that most of the effect is driven by inflows, and in particular inflows from wage work. Therefore, the entrepreneurship effect is driven by people who were entrepreneurs in 2008 but wage workers in 1999. However, when zooming across entrepreneur types in Panels B and C, there is a decrease in inflows from non-employment to unincorporated entrepreneurship and a similar sized, but positive, effect on flows from non-employment to incorporated entrepreneurship. Given that the sample focus on people who are potentially attached to the labour market, this is likely an amalgamation of people working informally, studying, temporarily unemployed or doing unpaid work. Therefore, I make no further claims on what is the driver behind this effect and I focus on flows from wage work to entrepreneurship in the rest of this section.

Since most of the inflows into entrepreneurship are driven by transitions from wage work, I can characterise who are these new entrepreneurs by analysing their baseline characteristics when they were wage workers. For this endeavour, I take two defining characteristics that the data include: baseline wages and occupation. Occupation is a 10-category variable reported by employers and used by the Social Security system to classify employees into skill levels[35].

To divide workers by wages, I take quartiles of wages at the industry level at baseline and classify workers according to the quartile in which they belong. Therefore, I classify people who are entrepreneurs in 2008 according to their position in the wage distribution in 1999. In Table 8, I provide the results on flows from wage work to entrepreneurship by quartiles of baseline wages. Overall, 75% of the impact of immigration on native entrepreneurship is accounted by entrepreneurs who were in the top half of the baseline distribution within their industry.

Finally, in Table 9, I show that most of the effect is explained by workers in medium to high skilled occupations. Since wages and occupation skill are positively correlated, the results points toward a similar as for the wage distribution: 68% of the impact of immigration on native entrepreneurship is accounted by entrepreneurs who were in medium and high skill occupations in 1999.

3.4 Robustness

I begin the robustness checks by analysing the sensitivity of the main results in Table 5. I provide a battery of robustness checks in Table B1. First, in Panels A, B and C, I drop

[35] As occupations grow in skill requirement, the minimum and maximum Social Security contribution cutoffs increase. Hence, as occupational skill increases, minimum contributions to Social Security increase.

either Barcelona, Madrid or both. The results remain mostly unchanged, although power decreases and relevance of the first stage as well, although it still is above the $F > 10$ rule of thumb. In Panel D, I drop the agriculture industry as this industry is known for employing large rates of self-employed, almost exclusively natives, which employ typically employ many seasonal immigrant workers informally (Hoggart and Mendoza, 1999). The presence of agricultural sector, in which large numbers of informal immigrants working for low wages enables potential profits of entrepreneurs to raise, contributes directly to the argument proposed in this paper. However, its omission does not change the results. Then, in Panel E I confirm estimates are not driven by outliers in the dependent variable, as dropping the top and bottom 5 percent of observations does not yield different results.

I show that the results are robust to other specifications in which I do not normalise by employment in the province. In Panel F, I normalise the outcome by local industry employment rather than province employment. The estimates get larger and a bit noisier, as some local industries have quite small numbers of employed people in 1999, but the results are maintained. Then, in Panel G I normalise also the independent variable, the immigration episode, by local industry imputed working age population. However, this measure becomes noisier and the first stage of the IV procedure becomes insignificant. Instead, I use OLS and I drop the top 10% provinces in terms of the shock, for which the immigration episode variable becomes unreasonably large. In this case, the results are similar to the OLS results in Panel B of Table 5. Finally, Panel H provides estimates without weighting by baseline population. All results survive qualitatively, although with a smaller magnitude, except for wages where the effect becomes negative and significant but only at the 15% significance level.

The shift-share procedure based on immigrant networks is also robust to the main criticisms in the literature. First, the immigration literature has raised worries about the serial correlation of the instrument confounding shocks at different time periods and potentially driving the results. This is less of a concern in the Spanish setting due to the sharp unexpected increase in immigration. In any case, Table B2 demonstrates that controlling for pre-existing immigration trends from 1996 to 1999[36] and using the multiple instrumentation procedure suggested by Jaeger et al. (2019) does not alter the results. Second, I embrace identification through exogeneity of the shares as in Goldsmith-Pinkham et al. (2020). Previous paper in the Spanish setting have defended the exogeneity of the shares in this setting[37], so I show that the instrument is also robust to using alternatives that try to increase the plausibility of the exogeneity of the shocks. Panel B of Table B3 shows that the

[36]The Padrón Contínuo data is only available from 1996.

[37]For instance Gonzalez and Ortega (2011, 2013); Sanchis-Guarner (2023); Castellanos (2024).

results remain virtually unchanged when using a push-factors instrument as in Sanchis-Guarner (2023). This instrument predicts FB_{ct} from Equation 4 using a "zero-th" stage that regresses FB_{ct} on a plethora of indicators from World Bank data across all countries of origin, and then predicts how many immigrants will move due to push-factors. Panel C shows the results of a leave-one-out (LOO) specification. The LOO instrument subtracts the number of foreign born population from country c in province p at time t, $FB_{c,p,t}$ from $FB_{c,t}$ in Equation 4, thus using only the number of immigrants in other provinces to calculate the shifts. Using a LOO also does not affect the results.

4 Discussion

4.1 Interpretation of the Results

Immigration has a positive effect on native entrepreneurship. This impact can be driven by many factors. I identify four main channels.

First, general equilibrium effects operating through immigrant consumption may increase the demand for native entrepreneurs. For instance, if immigrants demand more houses, more natives may become house builders. However, under the likely condition that immigrants consume products from all local industries, this mechanism will be partially captured by province fixed effects. At the same time, if one industry disproportionately absorbed labour and provided many opportunities for entrepreneurs, industry fixed effects should also capture this. Moreover, immigrants during this period had on average lower wages and were likely to send remittances back home, lessening the potential impact of potential spillovers of immigrant consumption on native entrepreneurship. Hence, I abstract from this channel both in the empirical section and in the model section.

Second, complementarity amongst immigrant entrepreneurs and native entrepreneurs can spur the creation of joint ventures and push natives into entrepreneurship. Although in other settings, like the US, this channel represents a relevant margin increasing native entrepreneurship (Duleep et al., 2021), existing evidence does not support this channel in the case of Spain. In Appendix Table B4 I show that immigration has a limited impact on immigrant entrepreneurship. This finding is consistent with circumstantial evidence showing that immigrants in Spain, while perceiving more business opportunities than natives, are less likely to exploit them (Bolívar-Cruz et al., 2014). This phenomena can be explained by immigrants facing higher legal and institutional problems when starting a business, as well as having larger credit constraints and lower entrepreneurial capital than natives. In consequence, I do not explore this channel.

A third explanation hinges on the increased potential returns from entrepreneurship for natives. If a large labour supply leads to abundant cheap labour, natives may take the opportunity to start their businesses. Since immigrants tend to be perfect substitutes amongst themselves, i.e. the own-price elasticity of factor demand is negative, an increase in immigration lowers immigrant wages. On top of this, immigrants usually work in jobs paying lower wages. There are many reasons behind immigrants' lower salaries: occupational downgrading, informality, lack of country-specific skills, such as language barriers, and inelastic labour supply, to name just a few.

Finally, wages of natives may react to an increase in immigration, leading them to change the relative value of working against entrepreneurship. For instance, if the impact of immigration on native wages is negative, then entrepreneurship becomes more attractive. While this channel is likely to be present in cases where migrants are perfect substitutes to natives and lower native wages, this possibility cannot be ruled out even if migrants and natives are not perfect substitutes[38]. A negative impact on negatives wages arising from lower labour demand may be masked by positive upward forces on native wages coming from general equilibrium effects such as labour supply responses and scale effects, as shown theoretically by Dustmann et al. (2017).

In the next subsection, I propose a simple model of occupational choice and immigration that captures the last two channels listed above. In the model, natives can choose whether to become wage workers or entrepreneurs, depending on the value of each choice. Entrepreneur production follows a CES function with three inputs: high educated native labour, low educated native labour, and immigrants. Hence, the model incorporates the canonical factor demand model used in immigration economics in an occupational choice model, following Lucas (1978). Immigrants impact native occupational choice through their impact on both natives wages and entrepreneurial profits. However, both quantities are determined in equilibrium, making it difficult to quantify how much a change in entrepreneurship is caused by each channel.

To understand how immigration impacts native entrepreneurship, I calibrate the model and perform counterfactual simulations. I first calibrate the model by minimising the distance between data moments and model moments. I combine static moments such as relative wages and entrepreneurship rates at baseline, with dynamic moments such as treatment effects identified in the empirical section. Once the model is calibrated, I perform a counterfactual decomposition by fixing either native wages or potential profits at

[38] Perfect substitutability entails an increase in immigration, keeping the number of natives fixed, lowers the marginal product of native labour, while imperfect substitutability implies an increase. Hence, imperfect substitutability implies that natives and immigrants are complements in employment (Borjas et al., 2008)

baseline. Fixing profits at baseline allows to understand how much of the increase in entrepreneurship is explained by changes in native wages due to increased immigrant supply. Conversely, fixing native wages at baseline allows to isolate the impact on profits operating only through the decrease in immigrant wages stemming from an increase immigrant supply, giving a sense of how much entrepreneurship reacts to the lower opportunity cost of becoming an entrepreneur. While the combination of these two effects does not add up exactly to the overall effect due to the presence of general equilibrium effects, this exercise shows that the latter channel explains the increase in native entrepreneurship.

4.2 A Model of Occupation Choice and Immigration

Set-up. Native individuals can choose whether to be wage workers or entrepreneurs, depending on its value V. If they become wage workers, they obtain a education-specific wage:

$$V^j_{WW} = w^j_N, \quad j \in \{H, L\}$$

If they become entrepreneurs, they employ $h(z)$ high-educated (HE) natives, $\ell(z)$ low-educated (LE) natives, and $i(z)$ immigrants to produce an output $O(z)$ that depends on their entrepreneurial ability z:

$$O(z) = z\left[a(bi(z)^\gamma + \ell(z)^\gamma)^{\frac{\rho}{\gamma}} + h(z)^\rho\right]^{\frac{\alpha}{\rho}} = zQ(z)$$

where a is the relative efficiency of LE workers with respect to HE workers, b is the relative efficiency of immigrants with respect to LE native workers. Then, $\alpha < 1$ is a decreasing returns to scale parameter, as in Lucas (1978). Finally, and most importantly, γ and ρ govern the degree of substitution/complementarity between LE natives and immigrants, and HE native workers and LE workers, respectively, with $\gamma \leq 1$ and $\rho < 1$.[39]

A large literature in immigration economics tries to obtain estimates of γ and ρ.[40] However, existing estimates of γ and ρ are calculated in settings where constant returns to scale (CRS) are assumed, i.e. $\alpha = 1$. Under CRS, HE natives are always complements in employment with respect to immigrants. For LE, it depends on the value of γ: the closer the value to 1, the most likely it is that the marginal product of LE labour decreases when immigration increases, all else fixed.

[39] In this paper I focus on the impact of immigration on wage and employment levels, not relative quantities, where the (inverse of the) elasticity of substitution $\sigma = \frac{1}{1-\gamma}$ plays a main role. When talking about wage and employment levels, both scale effects, due to expanded output from immigration-induced cost reductions, and general equilibrium effects stemming labour supply responses are also key elements to consider (Wagner, 2010).

[40] Ottaviano and Peri (2012) and Manacorda et al. (2012) are two prominent examples.

In the presence of decreasing returns to scale, the values of γ and ρ that make natives and immigrants substitutes or complements in employment are different. To allow complementarity amongst HE natives and immigrants, it is sufficient to have $\alpha < \rho < 1$. For LE natives and immigrants, the degree of substitutability depends on γ, but from the first order condition in Equation 6 it can be seen that whether $\partial w_N^L / \partial i \lesseqgtr 0$ depends on α as well. I don't make any prior assumption on whether natives and immigrants are substitutes or complements, and thus on the values of γ and ρ.

I choose two levels of skill based on education so the model can endogenously capture the positive correlation between entrepreneurial ability and wages suggested by the empirical results. Different skill levels allow for different degrees of complementarity or substitutability across workers of different education levels in the labour market, given by $\rho \neq \gamma$.

The value of being an entrepreneur will be equal to the profit $\pi(z)$. Profit is defined as output minus labour costs. The latter are determined by native and immigrant wages, respectively w_N^j and w_I, taken as given by the entrepreneur:

$$V_{EN}(z) = \pi(z) = zQ(z) - w_I i(z) - w_N^H h(z) - w_N^L \ell(z)$$

The first order conditions for optimal input choice are:

$$w_I = z\alpha \left[a(bi(z)^\gamma + l(z)^\gamma)^{\frac{\rho}{\gamma}} + h(z)^\rho \right]^{\frac{\alpha-\rho}{\rho}} ab(bi(z)^\gamma + l(z)^\gamma)^{\frac{\rho-\gamma}{\gamma}} i(z)^{\gamma-1} \tag{5}$$

$$w_N^L = z\alpha \left[a(bi(z)^\gamma + l(z)^\gamma)^{\frac{\rho}{\gamma}} + h(z)^\rho \right]^{\frac{\alpha-\rho}{\rho}} a(bi(z)^\gamma + l(z)^\gamma)^{\frac{\rho-\gamma}{\gamma}} l(z)^{\gamma-1} \tag{6}$$

$$w_N^H = z\alpha \left[a(bi(z)^\gamma + l(z)^\gamma)^{\frac{\rho}{\gamma}} + h(z)^\rho \right]^{\frac{\alpha-\rho}{\rho}} h(z)^{\rho-1} \tag{7}$$

Wages are determined in equilibrium by labour market clearing and the break-even condition. The first condition condition implies that the labour supply of each factor is equal to its demand by firms. Labour supply by natives is given by LS_N^j, for each education level j. Labour supply for immigrants is given by LS^I. Labour supply for natives is determined endogenously, while labour supply for immigrants is exogenous as they do not start firms.

Labour market clearing conditions can be written as follows:

$$LS_N^H = \int_{z^{*H}} \mu^H(z)n(z)dz + \int_{z^{*L}} \mu^L(z)n(z)dz$$
$$LS_N^L = \int_{z^{*H}} \mu^H(z)\ell(z)dz + \int_{z^{*L}} \mu^L(z)\ell(z)dz$$
$$LS_I = \int_{z^{*H}} \mu^H(z)i(z)dz + \int_{z^{*L}} \mu^L(z)i(z)dz$$

where ability z for each type of native j follows a pdf $\mu^j(z)$.

Native labour supply is determined by cutoffs z^{j*}. These cutoffs are obtained from the break-even condition for the marginal entrepreneurs of each education level:

$$\pi(z^{H*}, LS_N^H, LS_N^L, LS_I, w_N^H, w_N^L, w_I) = w_N^H(z^{H*}, LS_N^H, LS_N^L, LS_I, w_N^L, w_I) \tag{8}$$
$$\pi(z^{L*}, LS_N^H, LS_N^L, LS_I, w_N^H, w_N^L, w_I) = w_N^L(z^{L*}, LS_N^H, LS_N^L, LS_I, w_N^H, w_I) \tag{9}$$

Therefore, an equilibrium of this economy consists of wage rates w_N^H, w_N^L, and w_I such that taking wages as given, natives choose optimally between employment and entrepreneurship, entrepreneurs demand inputs optimally, and the labour market clears.

Calibration. I estimate and calibrate the model using a two-step minimum distance estimator. In the first step, I set $\alpha = 0.9$, to obtain a profit share of income[41] of 10%. I choose the proportion of LE natives to be half of the native population to match the data. In the baseline period, I choose the number of immigrants such that its share over total population amounts to 2.14%, consistent with the share of immigrants in the working age population in 1999. Additionally, I suppose z follows a log-normal distribution with mean μ_z and variance σ_z^2. In the second step, I estimate the remaining parameters of the model by minimising the distance between data moments and moments simulated by the model.

To make the calibration consistent with the empirical analysis, I use both baseline moments and dynamic moments. For baseline moments, I take baseline shares of entrepreneurship by education level and relative wages, obtained directly from descriptive statistics of the data in 1999, before the immigration episode. Dynamic moments include increases in wages and entrepreneurship by education level, and are calculated using the treatment effects identified in the empirical section. Therefore, in the model calibration I estimate the model both with the share of immigrants before and after the immigration episode, calculate baseline and dynamic moments, and find the parameters that minimise the objective function that combines the distances between data and model moments.

[41] I use the same α as Poschke (2018).

Table 10 provides the results from the calibration. The model does a good job in matching data moments at baseline. When it comes to matching dynamic moments, the model performs reasonably well except for the wage increase of low-educated workers. The model does not match the increase in wages of low-educated workers, thus leading to relative wage increases across education levels to be biased towards high-educated individuals. Since the main channel this paper hinges on is movements of wage workers into entrepreneurship after an immigration episode, which is captured by the model, I deem the calibration satisfactory.

The calibrated parameters are reasonable when interpreted through the lens of previous research. The substitution parameters ρ and γ from the right panel of Table 10 imply complementarity in employment, as an increase in immigration lead to an increase in the marginal product of both LE and HE natives.[42] This is consistent with previous literature documenting that labour demand of natives can increase due to a complementarities in production between natives and immigrants (Peri and Sparber, 2009; Beerli et al., 2021). Then, $1 > a = 0.75 > b = 0.28$ shows that workers in the low-educated and immigrants nest are less productive than high-educated natives, and also that immigrants are less productive than low-educated natives. Finally, the parameters μ_j and σ_j show that the distribution of entrepreneurial ability amongst high-educated has a higher mean but lower variance. Therefore, the model endogenously generates a positive correlation between wages and ability across education levels.

4.3 Counterfactual simulations

In this section, I perform a counterfactual decomposition by fixing either native wages or potential profits at baseline. The decomposition helps to understand which mechanism, namely (i) the impact of immigration on native wages or (ii) on potential profits for entrepreneurs, is driving the increase in entrepreneurship. The idea behind this exercise is to fix either the left-hand side or right-hand side, respectively, of the break-even conditions expressed in Equations 8 and 9. Fixing profits at baseline allows to understand how much of the increase in entrepreneurship is explained by changes in native wages due to increased immigrant supply. Immigration leading to higher native wages increases the opportunity cost of entrepreneurship. Conversely, fixing native wages at baseline allows to isolate the impact on potential profits operating only through a decrease in immigrant

[42]This can be shown numerically by taking the production function of any entrepreneur and increasing immigration while keeping native employment fixed. In fact, with the current estimates of ρ and γ, HE natives' wages increase more than those of LE natives when immigration increases, i.e. $\partial w_N^H/\partial i > \partial w_N^L/\partial i > 0$.

wages. This scenario gives a sense of how much entrepreneurship reacts to the higher potential entrepreneurial profits. While the combination of these two counterfactual simulations does not add up exactly to the overall effect due to the presence of general equilibrium effects, this exercise provides key insights into which mechanism dominates.

Before and after. Table 11 shows the results of the counterfactual simulations. The first column shows the baseline, or before period, where immigration is around 2%. In the after scenario, in the second column, immigration increases to 14%. Native wages increase, with LE natives experiencing a relatively lower increase. Wages of immigrants drop substantially, as their own-price elasticity is negative because of perfect substitutability.

When it comes to profits, they increase on average for both HE and LE entrepreneurs. However, two mechanisms are at play affecting profits: selection and lower costs. A higher share of entrepreneurs has a negative selection effect: new entrants profits are lower, lowering average profits. However, this is compensated by an upward shift in the profit curve due to immigration lowering the cost of production. Therefore, existing entrepreneurs see their profits increase on net.

The increase in entrepreneurship amongst natives despite higher wages suggest that the impact of immigration on potential profits dominates the impact on native wages. However, to provide further confirmation, I now turn to the counterfactual composition that allows me to show that this is indeed the case.

Counterfactual: fix profits at baseline. Column 3 of Table 11 provides the results of this counterfactual. Shutting down potential profits makes natives' labour supply choices depend on wage changes. In terms of labour demand, wages of HE natives increase more than those of LE natives due to the immigration increase. This is because the former are more complementary with respect to immigrants, as the marginal product of HE natives in production increases relatively more, all else fixed. Regarding entrepreneur shares, both groups of natives see their entrepreneurship decrease when compared to the *Before* setting, but more so for HE natives, as the opportunity cost of entrepreneurship increases more due to higher wages. Since profits are unaffected by changes in the wage structure, the average profits of entrepreneurs increase only because of positive selection into entrepreneurship. However, compared to the *After scenario*, native wages are slightly lower because there is a larger pool of workers, thus putting downward pressure in wages.

Counterfactual: fix native wages at baseline. Column 4 of Table 11 provides the results of this counterfactual. When native wages are fixed at baseline, native labour supply choices are driven by changes in potential profits. Immigration has a negative impact on immigrant wages while keeping native wages unchanged, and thus potential profits increase for both HE and LE natives. This shifts the profits curve up for natives. Since the profit distribution for HE natives is more concentrated due to a lower variance, they experience a relatively higher increase in entrepreneurship.

Taken together, the last two counterfactual scenarios suggest that the increase in profits due to immigrants' lower wages is the main channel behind the increase in native entrepreneurship, as the shift in the potential profits curve more than compensates for the increase in wages amongst HE native workers due to the immigration episode. On net, the channel outlined in the second counterfactual dominates. Immigration increases native entrepreneurship, and this increase is driven by HE entrepreneurs.

5 Conclusion

I have provided evidence on the effects of immigration on native entrepreneurship in the context of Spain. Immigration episodes in developed economies have been pervasive in recent history, and the number of international migrants in developed economies has only grown. Still, there is a widespread belief that immigrants might be an economic burden, and particularly, might have negative consequences on labour market outcomes of natives. In this paper in focus on one of the most massive immigration episodes in the postwar era amongst OECD countries, I argue that international immigration might have a limited impact on employment and wages of natives, while fostering native entrepreneurship. More concretely, my main contribution is to show that these immigration episodes can foster the entry to entrepreneurship amongst natives. This is plausibly explained by immigration lowering labour costs, and thus incentivising entrepreneurship amongst individuals with relatively higher entrepreneurial ability and who, in absence of the immigration episode, would not have become entrepreneurs.

However, the results in my paper depend crucially on the type of immigration episode. Immigrants who entered Spain during the analysis period mostly worked in low-paying manual jobs. In other countries, such as the US in the present century, immigrants represent a high share in skilled occupations (Kerr et al., 2015) and have higher rates of entrepreneurship than natives (Kerr and Mandorff, 2023). Consistently, the impact of immigration on the count of native owned business has recently been found to be negative in the US (Mahajan, 2022). Contrary to this, in my setting, as immigrants lower labour costs but do not generally compete with native entrepreneurs, they have a positive impact on native entrepreneurship. Hence, my findings relate more directly to scenarios in which immigrants suffer substantial occupational downgrading or mostly take low-paying jobs, such as refugee episodes[43] or developing to developed countries migration episodes.

[43]For instance, my results are consistent with a positive effect of refugee migration in Turkey on extensive

and intensive margins of firm production (Altindag et al., 2020).

References

Adda, J., C. Dustmann, and J.-S. Görlach (2022): "The Dynamics of Return Migration, Human Capital Accumulation, and Wage Assimilation," The Review of Economic Studies, rdac003. 1

Ajzenman, N., C. G. Aksoy, and S. Guriev (2022): "Exposure to transit migration: Public attitudes and entrepreneurship," Journal of Development Economics, 158, 102899. 5

Albert, C. and J. Monras (2018): "Immigration and Spatial Equilibrium: the Role of Expenditures in the Country of Origin," CEPR Discussion Papers 12842, C.E.P.R. Discussion Papers. 1

Altindag, O., O. Bakis, and S. V. Rozo (2020): "Blessing or burden? Impacts of refugees on businesses and the informal economy," Journal of Development Economics, 146. 28

Altonji, J. and D. Card (1991): "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in Immigration, Trade, and the Labor Market, National Bureau of Economic Research, Inc, 201–234. 12

Amuedo-Dorantes, C. and S. De La Rica (2011): "Complements or substitutes? Task specialization by gender and nativity in Spain," Labour Economics, 18, 697–707. 5, 7, 8, 15

Babina, T. (2019): "Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship," The Review of Financial Studies, 33, 4061–4101. 1, 5

Beerli, A., J. Ruffner, M. Siegenthaler, and G. Peri (2021): "The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland," American Economic Review, 111, 976–1012. 6, 25

Bohnet, L., S. Peralta, and J. P. dos Santos (2021): "Cousins from overseas: the labour market impact of half a million Portuguese repatriates," NOVAFRICA Working Paper Series wp2114, Universidade Nova de Lisboa, Nova School of Business and Economics, NOVAFRICA. 1, 6

Bolívar-Cruz, A., R. M. Batista-Canino, and E. Hormiga (2014): "Differences in the perception and exploitation of entrepreneurial opportunities by immigrants," Journal of Business Venturing Insights, 1-2, 31–36. 20

Borjas, G. J., J. Grogger, and G. H. Hanson (2008): "Imperfect Substitution between Immigrants and Natives: A Reappraisal," NBER Working Papers 13887, National Bureau of Economic Research, Inc. 21

Borusyak, K., P. Hull, and X. Jaravel (2021): "Quasi-Experimental Shift-Share Research Designs," The Review of Economic Studies, 89, 181–213. 14

Bosch, M. and L. Farré (2014): "Immigration and the Informal labor Market," Cuadernos

Economicos del ICE, 87, 185–204. 5, 7, 49

Card, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," Journal of Labor Economics, 19, 22–64. 2, 12

Castellanos, M. A. (2024): "Immigration, Parenthood and Child Penalties," Working Paper. 19

Colaiacovo, I., M. G. Dalton, S. Pekkala Kerr, and W. R. Kerr (2022): "The Transformation of Self Employment," Working Paper 29725, National Bureau of Economic Research. 47

De La Rica, S., A. Glitz, and F. Ortega (2014): "Immigration in Spain: what have we learned from recent evidence?" Cuadernos Economicos del ICE, 87, 9–28. 7

Duleep, H., D. A. Jaeger, and P. McHenry (2021): "On Immigration and Native Entrepreneurship," CReAM Discussion Paper Series 2108, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London. 1, 5, 20

Dustmann, C. and A. Glitz (2015): "How Do Industries and Firms Respond to Changes in Local Labor Supply?" Journal of Labor Economics, 33, 711 – 750. 1, 6

Dustmann, C., U. Schönberg, and J. Stuhler (2016): "The Impact of Immigration: Why Do Studies Reach Such Different Results?" Journal of Economic Perspectives, 30, 31–56. 1, 10, 16

——— (2017): "Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment," The Quarterly Journal of Economics, 132, 435–483. 5, 21

Edo, A. (2019): "The Impact of Immigration on the Labor Market," Journal of Economic Surveys, 33, 922–948. 1, 5

Elias, F., J. Monras, and Vázquez-Grenno (2022): "Understanding the Effects of Legalizing Undocumented Immigrants," Working paper. 5

Fairlie, R. and B. Meyer (2003): "The Effect of Immigration on Native Self-Employment," Journal of Labor Economics, 21, 619–650. 1, 5

Fernández-Huertas Moraga, J., A. Ferrer-i Carbonell, and A. Saiz (2019): "Immigrant locations and native residential preferences: Emerging ghettos or new communities?" Journal of Urban Economics, 112, 133–151. 2, 12

Foged, M. and G. Peri (2016): "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data," American Economic Journal: Applied Economics, 8, 1–34. 1, 5

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): "Bartik Instruments: What, When, Why, and How," American Economic Review, 110, 2586–2624. 2, 14, 19

Gonzalez, L. and F. Ortega (2011): "How do very open economies adjust to large immigration flows? Evidence from Spanish regions," Labour Economics, 18, 57–70. 3, 5, 7, 12, 15, 19

——— (2013): "Immigration and housing boom: Evidence from Spain," Journal of Regional Science, 53, 37–59. 12, 19, 48

Hacamo, I. and K. Kleiner (2022): "Forced Entrepreneurs," The Journal of Finance, 77, 49–83. 1, 5

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013): "Who Creates Jobs? Small versus Large versus Young," The Review of Economics and Statistics, 95, 347–361. 1

Hoggart, K. and C. Mendoza (1999): "African Immigrant Workers in Spanish Agriculture," Sociologia Ruralis, 39, 538–562. 19

Humphries, J. E. (2021): "The Causes and Consequences of Self-Employment over the Life Cycle," Working paper. 10

Imbert, C. and G. Ulyssea (2024): "Rural Migrants and Urban Informality: Evidence from Brazil," Working Papers DP18160, CEPR. 1

INE (2008): "Explotacion estadistica del directorio central de empresas (DIRCE)," https://www.ine.es/dynt3/inebase/es/index.htm?padre=54&capsel=3920, Last accessed on 2023-01-13. 5

Iraizoz-Olaetxea, A. (2022): "Saving for Retirement through the Public Pension System: Evidence from the Self-Employed in Spain," Working Paper. 17, 48

Jaeger, D. A., J. Ruist, and J. Stuhler (2019): "Shift-Share Instruments and Dynamic Adjustments: The Case of Immigration," Working Paper 24285. 2, 12, 14, 19, 51

Kerr, S. P., W. R. Kerr, and W. F. Lincoln (2015): "Skilled Immigration and the Employment Structures of US Firms," Journal of Labor Economics, 33, S147–S186. 27

Kerr, W. R. and M. Mandorff (2023): "Social Networks, Ethnicity, and Entrepreneurship," Journal of Human Resources, 58, 183–220. 13, 27

Levine, R. and Y. Rubinstein (2016): "Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?*," The Quarterly Journal of Economics, 132, 963–1018. 2, 5, 16

——— (2020): "Selection into entrepreneurship and self-employment," CEP Discussion Papers dp1722, Centre for Economic Performance, LSE. 16

Lucas, R. (1978): "On the Size Distribution of Business Firms," Bell Journal of Economics, 9, 508–523. 21, 22

Mahajan, P. (2022): "Immigration and Busines Dynamics: Evidence from U.S. Firms," Working paper. 1, 6, 12, 13, 27

Manacorda, M., A. Manning, and J. Wadsworth (2012): "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain," Journal of the European Economic Association, 10, 120–151. 22

Monras, J. (2020): "Immigration and Wage Dynamics: Evidence from the Mexican Peso

Crisis," Journal of Political Economy, 128, 3017–3089. 16

OED (2022): "OED Dictionary Online," . 47

Olney, W. W. (2013): "Immigration and Firm Expansion," Journal of Regional Science, 53, 142–157. 6

Ortega, F. and G. Peri (2013): "The effect of income and immigration policies on international migration," Migration Studies, 1, 47–74. 6

Ottaviano, G. I. P. and G. Peri (2012): "RETHINKING THE EFFECT OF IMMIGRATION ON WAGES," Journal of the European Economic Association, 10, 152–197. 22

Ozguzel, C. (2021): "The Cushioning Effect of Immigrant Mobility: Evidence from the Great Recession in Spain," Cesifo working paper no. 9268. 10, 12

Peri, G. and C. Sparber (2009): "Task Specialization, Immigration, and Wages," American Economic Journal: Applied Economics, 1, 135–69. 1, 5, 7, 25

Piyapromdee, S. (2020): "The Impact of Immigration on Wages, Internal Migration, and Welfare," The Review of Economic Studies, 88, 406–453. 5

Poschke, M. (2013): "Who becomes an entrepreneur? Labor market prospects and occupational choice," Journal of Economic Dynamics and Control, 37, 693–710. 5

——— (2018): "The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology," American Economic Journal: Macroeconomics, 10, 1–41. 24

Sanchis-Guarner, R. (2023): "Decomposing the Impact of Immigration on House Prices," Tech. rep. 10, 12, 19, 20

Simon, H., R. Ramos, and E. Sanroma (2014): "Immigrant Occupational Mobility: Longitudinal Evidence from Spain," European Journal of Population, 30, 223–255. 7

Unel, B. (2022): "Effects of Immigration on Native Entrepreneurship in the U.S," Departmental working papers, Department of Economics, Louisiana State University. 5

Wagner, M. (2010): "The Heterogeneous Labor Market Effects of Immigration," CeRP Working Papers 93, Center for Research on Pensions and Welfare Policies, Turin (Italy). 22

Tables

	Natives	Immigrants
Total	651,222	86,562
Share male	0.57	0.61
Average age	39.17	35.52
Average tenure	5.59	1.87
Average daily wage	64.0	43.9
Occupation shares		
Low skill	0.45	0.74
Medium skill	0.35	0.19
High skill	0.20	0.06
Entrepreneurs		
Self-employed	0.17	0.11
Unincorporated	0.12	0.09
Incorporated	0.05	0.02
Industry		
Agriculture	0.02	0.06
Manufacturing	0.16	0.18
Construction	0.11	0.19
Hospitality and retail	0.24	0.25
Other services	0.47	0.32

Table 1: Comparison of labour market outcomes between natives and immigrants

Note: This table provides a comparison of natives and immigrants in the year 2008 using data from the MCVL. The data correspond to individuals aged 20 to 60. Daily wages are calculated in euros and are total yearly earnings divided by the number of days worked, amongst full-time wage workers who were employed all year long. Skills are calculated using occupations, which in the MCVL data correspond to skill levels as viewed by the employer.

	1999	2008
Total	466,925	466,925
Wage Workers	279,874	333,288
Entrepreneurs	38,991	72,506
Unincorporated	27,047	47,349
Incorporated	11,944	25,157
Average age	32.2	40.4
Average tenure	4.2	6.4
Average wage	44.0	49.5
Education		
Low education	57,061	57,061
Medium education	177,630	177,630
High education	227,759	227,759
Occupation (*wage workers*)		
Low skill	97,846	129708
Medium skill	84,302	87,998
High skill	52,526	39,800
Industry (*wage workers*)		
Agriculture	2,276	2,833
Manufacturing	57,396	57,155
Construction	29,743	30,122
Hospitality and retail	67,109	69,041
Other services	123,350	174,137
Industry (*entrepreneurs*)		
Agriculture	4,501	6,096
Manufacturing	4,166	6,564
Construction	5,168	11,565
Hospitality and retail	14,668	25,254
Other services	10,488	23,027

TABLE 2: NATIVE LABOUR MARKET OUTCOMES, ANALYSIS SAMPLE

Note: This table provides information on the analysis sample, splitting by year. All statistics are counts, except for average age, tenure and wage. Age and tenure are expressed in years, while average wage is expressed in daily wages amongst full-time wage workers who were employed during the whole year. Skills are calculated using occupations, which in the MCVL data correspond to skill levels as viewed by the employer. Low education refers to less than secondary, medium to secondary education, and high to more than secondary.

	Mean	Std. Dev.	Max	Min
Main variables				
Change Native Entrepreners (normalised)	0.026	0.013	0.057	-0.000
Change Native Unincorporated Entr. (normalised)	0.016	0.009	0.049	-0.001
Change Native Incorporated Entr. (normalised)	0.010	0.005	0.026	-0.001
Change Native Wage Workers (normalised)	0.057	0.075	0.237	-0.066
Change Log Wages	0.001	0.044	0.674	-0.265
Immigration Shock	0.043	0.030	0.115	-0.000
Supporting variables				
Share Native Entrepreneur over Employed, 1999	0.138	0.117	0.971	0.049
Share Native Entrepreneur over Employed, 2008	0.193	0.128	0.944	0.056
Share Native Incorporated over Entrepreneur, 1999	0.326	0.120	0.827	0.000
Share Native Incorporated over Entrepreneur, 2008	0.369	0.116	0.875	0.016

Table 3: Descriptive statistics of main variables

Note: The table presents descriptive statistics for the main variables used in the analysis, using the analysis sample, that is, using the data aggregated across 250 local industries, from data on natives born between 1954 and 1979, as explained in Section 1.3. The first four variables follow Equation 2 and are the changes in number of native entrepreneurs, unincorporated and incorporated and wage workers between 1999 and 2008 across local industries, normalised by the province native employment (equal to the sum of the province native wage workers and native entrepreneurs) in 1999. The difference in log wages is calculated from 1999 to 2008, and log wages are obtained as residuals from a regression of log daily wages on quadratic age and tenure profiles, and occupation and year fixed effects, using wages only from wage workers. The immigration episode corresponds to Equation 1 and represents the change in a local industry immigrant population over the province baseline working age population. Supporting variables are calculated exclusively for natives. Statistics are weighted according to each local industry baseline native employment.

	Outcome: ΔImmigration Shock$_{ip}$			
	(1)	(2)	(3)	(4)
$\Delta\widehat{\text{Immigration}}$ Shock$_{ip}$	0.603***	0.447***	0.610***	0.472***
	(0.04)	(0.09)	(0.07)	(0.10)
Controls		X		X
Industry/Province FE			X	X
F-statistic	210.61	26.67	67.24	23.11
Observations	250	250	250	250

Table 4: First stage

Note: This table presents first-stage regressions of the immigrant episode on the instrument, as explained in Section 2. The F-statistic corresponds to the F-statistic on the excluded instrument, namely the predicted immigration episode. Observations are weighted by baseline employment in each local industry. Robust standard errors are reported in parenthesis. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

	(1) Δ Employment	(2) Δ Wage Workers	(3) Δ Entrepreneurs	(4) Δ Unincorporated	(5) Δ Incorporated	(6) Δ Wage
Panel A: OLS, no controls						
Δ Immigration Shock	0.269	0.026	0.243***	0.129***	0.114***	0.003
	(0.22)	(0.20)	(0.05)	(0.04)	(0.02)	(0.02)
Panel B: OLS, with controls						
Δ Immigration Shock	0.159	-0.052	0.211***	0.113***	0.098***	0.009
	(0.18)	(0.18)	(0.04)	(0.03)	(0.02)	(0.02)
Panel C: 2SLS, without controls						
Δ Immigration Shock	0.375+	0.062	0.314***	0.145***	0.169***	0.013
	(0.25)	(0.22)	(0.07)	(0.05)	(0.02)	(0.02)
First-stage KP	67.24	67.24	67.24	67.24	67.24	67.24
Panel D: 2SLS, with controls						
Δ Immigration Shock	0.132	-0.100	0.232***	0.051	0.181***	0.058*
	(0.30)	(0.29)	(0.06)	(0.05)	(0.03)	(0.03)
First-stage KP	23.11	23.11	23.11	23.11	23.11	23.11
Baseline workforce share		0.878	0.122	0.085	0.037	
Mean dep. var	9.79	6.84	2.94	1.79	1.15	0.04
Natives per 100 immigrants	4.5	-3.5	8.0	2.0	6.0	
Observations	250	250	250	250	250	250

Table 5: Entrepreneurship, Employment and Wage effects of immigration

Note: This table provides the estimates from estimating β from Equation 3 using either OLS or IV, and adding or not controls. Each column corresponds to an outcome, namely, a difference in the number of native individuals in a given occupational category from 1999 to 2008 in a local industry normalised by baseline province employment for Columns (1) to (5), and to the change in log wages during the same period in Column (6). Additional information on the outcome variables, the immigration episode, or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. The Table provides the mean of each dependent variable in percentage terms (multiplied by 100), as well as baseline shares of each employment category in Columns (2) to (5). I also provide a back of the envelope calculation of the effect of an additional 100 immigrants on the increases (or decreases) of natives in each category. Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

	(1) Δ Employment	(2) Δ Wage Workers	(3) Δ Entrepreneurs	(4) Δ Unincorporated	(5) Δ Incorporated	(6) Δ Wage
Panel A: High education						
Δ Immigration Shock	0.053	-0.142	0.195***	0.064**	0.130***	-0.422
	(0.24)	(0.22)	(0.04)	(0.03)	(0.02)	(0.46)
Panel B: Low education						
Δ Immigration Shock	0.079	0.041	0.038	-0.013	0.051***	1.364***
	(0.15)	(0.15)	(0.04)	(0.03)	(0.02)	(0.51)
Mean dep. var	2.76	1.66	1.10	0.72	0.38	-0.36
First-stage KP	23.11	23.11	23.11	23.11	23.11	23.11
Observations	250	250	250	250	250	250

Table 6: Entrepreneurship, Employment and wage effects of immigration, by education

Note: This table provides the estimates from estimating β from Equation 3 using the migrant networks instrument detailed in Section 2 and controls. Panel A estimates results for high-education individuals, and Panel B for low-education individuals. Each column corresponds to an outcome, namely, a difference in the number of native individuals, by education, in a given occupational category from 1999 to 2008 in a local industry normalised by baseline province employment for Columns (1) to (5), and to the change in log wages for each education level during the same period in Column (6). Additional information on the outcome variables, the immigration episode, or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. The Table provides the mean of each dependent variable in percentage terms (multiplied by 100). Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

		Inflows			Outflows		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Δ Entrepreneurship	Entrepreneurship (Other LI)	Wage Work	Non-Employment	Entrepreneurship (Other LI)	Wage Work	Non-Employment
Panel A: All							
Δ Immigration Shock	0.232***	0.003	0.226***	0.004	-0.003	-0.003	0.005
	(0.06)	(0.00)	(0.04)	(0.04)	(0.00)	(0.01)	(0.00)
Mean dep. var	2.94	0.04	1.73	1.59	0.03	0.29	0.09
Panel B: Unincorporated							
Δ Immigration Shock	0.051	0.002	0.094***	-0.056^{+}	-0.001	-0.002	-0.003
	(0.05)	(0.00)	(0.03)	(0.04)	(0.00)	(0.01)	(0.01)
Mean dep. var	1.79	0.02	1.03	1.05	0.02	0.21	0.07
Panel C: Incorporated							
Δ Immigration Shock	0.181***	-0.001	0.132***	0.059***	0.001	-0.000	0.008***
	(0.03)	(0.00)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)
Mean dep. var	1.15	0.01	0.70	0.54	0.01	0.08	0.02
First-stage KP	23.11	23.11	23.11	23.11	23.11	23.11	23.11
Observations	250	250	250	250	250	250	250

Table 7: Flows to and from Entrepreneurship

Note: This table provides the estimates from estimating β from Equation 3 using the migrant networks instrument detailed in Section 2 and controls. Panel A estimates results for all entrepreneurs, while Panels B and C estimate results for unincorporated and incorporated entrepreneurs, respectively. Each column corresponds to an flow outcome, except Column (1), which corresponds to Column (3) from Table 5. Column (2) are flows from entrepreneurship in other local industries in 1999 to entrepreneurship in the local industry. Column (3) are flows from wage work in 1999 to entrepreneurship in 2008, and Column (4) likewise but from non-employment. Columns (5) to (7) are defined similarly, but as outflows. All flows are normalised by baseline employment in the province. Additional information the immigration episode or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. The Table provides the mean of each dependent variable in percentage terms (multiplied by 100). Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Panel A: All				
Δ Immigration Shock	0.028^{+}	0.020	0.075***	0.103***
	(0.02)	(0.02)	(0.01)	(0.01)
Mean dep. var	0.59	0.44	0.34	0.36
Panel B: Unincorporated				
Δ Immigration Shock	0.012	-0.000	0.035***	0.047***
	(0.01)	(0.01)	(0.01)	(0.01)
Mean dep. var	0.38	0.27	0.19	0.19
Panel C: Incorporated				
Δ Immigration Shock	0.016**	0.020***	0.040***	0.056***
	(0.01)	(0.01)	(0.01)	(0.01)
Mean dep. var	0.21	0.17	0.15	0.17
First-stage KP	23.11	23.11	23.11	23.11
Observations	250	250	250	250

Table 8: Flows from Wage Work to Entrepreneurship by Baseline Wage Quartiles

Note: This table provides the estimates from estimating β from Equation 3 using the migrant networks instrument detailed in Section 2 and controls. Panel A estimates results for all entrepreneurs, while Panels B and C estimate results for unincorporated and incorporated entrepreneurs, respectively. Each column corresponds to the number of people who were wage workers in a given quartile in of the wage distribution in 1999 but entrepreneurs in the local industry in 2008, and each is normalised by baseline employment in the province. Quartiles are calculated from distributions at the industry level. Additional information the immigration episode or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. The Table provides the mean of each dependent variable in percentage terms (multiplied by 100). Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

	(1) LS occ.	(2) MS occ.	(3) HS occ.
Panel A: All			
Δ Immigration Shock	0.067**	0.106***	0.052***
	(0.03)	(0.02)	(0.01)
Mean dep. var	0.92	0.54	0.27
Panel B: Unincorporated			
Δ Immigration Shock	0.017	0.052***	0.025***
	(0.02)	(0.01)	(0.01)
Mean dep. var	0.59	0.30	0.14
Panel C: Incorporated			
Δ Immigration Shock	0.050***	0.054***	0.027***
	(0.02)	(0.01)	(0.01)
Mean dep. var	0.33	0.24	0.13
First-stage KP	23.11	23.11	23.11
Observations	250	250	250

Table 9: Flows from Wage Work to Entrepreneurship by Baseline Occupations

Note: This table provides the estimates from estimating β from Equation 3 using the migrant networks instrument detailed in Section 2 and controls. Panel A estimates results for all entrepreneurs, while Panels B and C estimate results for unincorporated and incorporated entrepreneurs, respectively. Each column corresponds to the number of people who were wage workers in a given occupations in 1999 but entrepreneurs in the local industry in 2008, and each is normalised by baseline employment in the province. Additional information the immigration episode or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. The Table provides the mean of each dependent variable in percentage terms (multiplied by 100). Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Moment	Data	Model
Baseline Entr. Rate HE	0.11	0.07
Baseline Entr. Rate LE	0.14	0.14
Δ Entr. Rate HE	1.12	1.10
Δ Entr. Rate LE	1.03	1.04
Baseline $\frac{w_N^H}{w_N^L}$	1.19	1.19
Baseline $\frac{w_N^L}{w^I}$	1.03	1.03
Δw_N^H	1	1.04
Δw_N^L	1.06	1.01

Parameter	Value
ρ	0.39
γ	0.57
a	0.75
b	0.28
μ_z^H	0.76
μ_z^L	0.63
σ_z^H	0.02
σ_z^L	0.14

Table 10: Moments and parameters from model calibration

Note: The left table provides moments used in the estimation of the model of occupational choice and immigration laid out in Section 4.2. The right table provides the parameters and calibrated values. ρ and γ are the substitution parameters for high-educated and low-educated/immigrant nest, and immigrants and low-educated natives, respectively. Parameters a and b are relative productivity of low-educated nest of the CES production function and relative productivity of immigrants within the low-educated nest in the CES production function, respectively. The parameters μ_j and σ_z^j for $j \in \{L, H\}$ are the mean and standard deviation of the log-normal distribution of entrepreneurial ability for each education level.

	Before	After	After, Fix Profits	After, Fix Wages
High Educated Natives				
Share Entrepreneur	0.073	0.080	0.059	0.090
Average Profits	4.117	4.305	4.840	4.140
Low Educated Natives				
Share Entrepreneur	0.143	0.150	0.142	0.148
Average Profits	7.931	8.134	9.122	8.007
Wages				
Immigrants	2.921	1.253	1.252	1.252
HE Natives	3.627	3.786	3.737	3.627
LE Natives	2.990	3.019	3.005	2.990

Table 11: Counterfactual simulations

Note: The table displays key statistics for HE and LE native entrepreneurs, and wages. Each column shows a different scenario. The first and second columns are the actual before and after scenarios. The third and fourth are the counterfactual simulations in which potential profits or native wages are fixed at baseline, respectively.

Figures

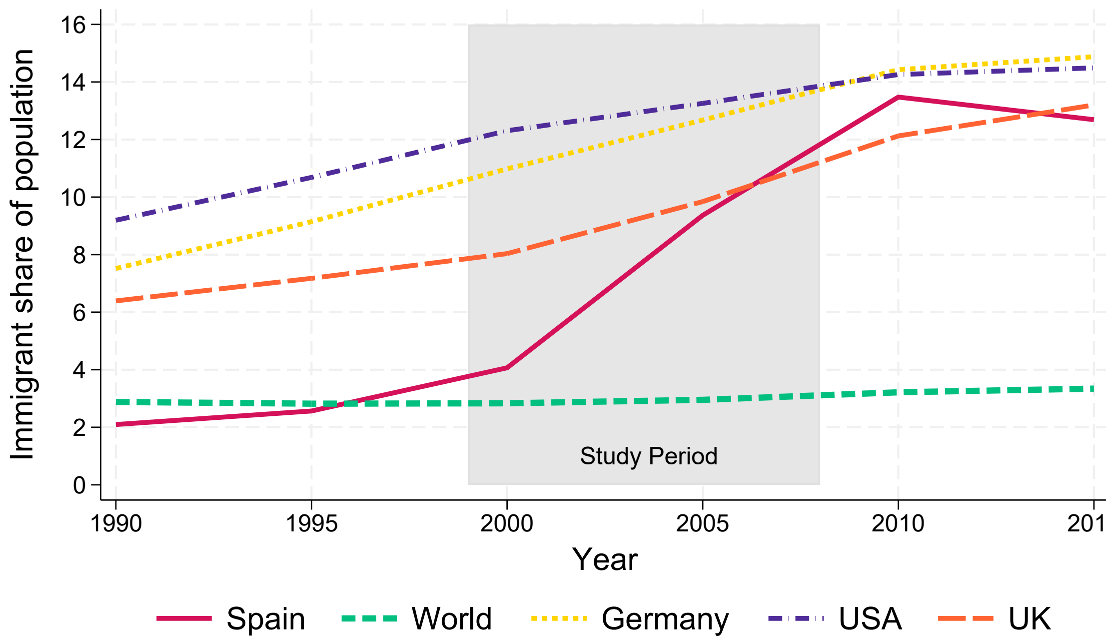
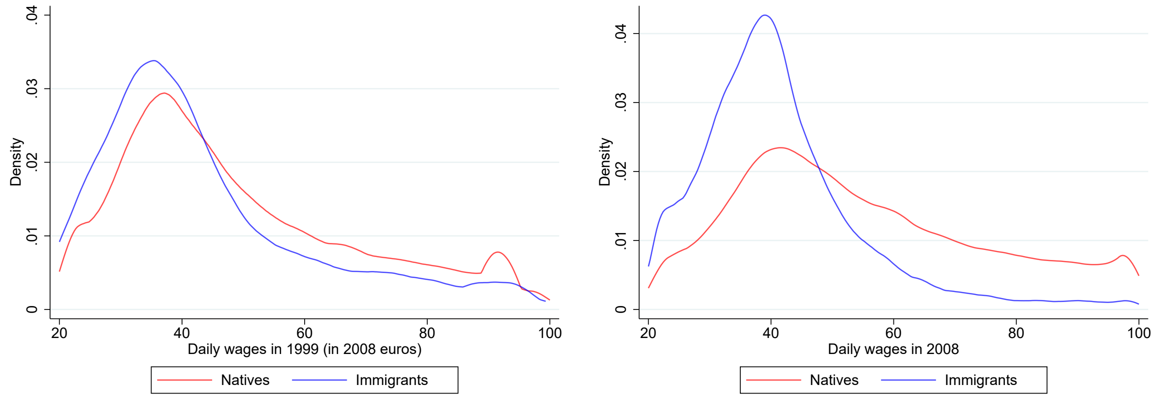


Figure 1: International comparison of immigrant shares

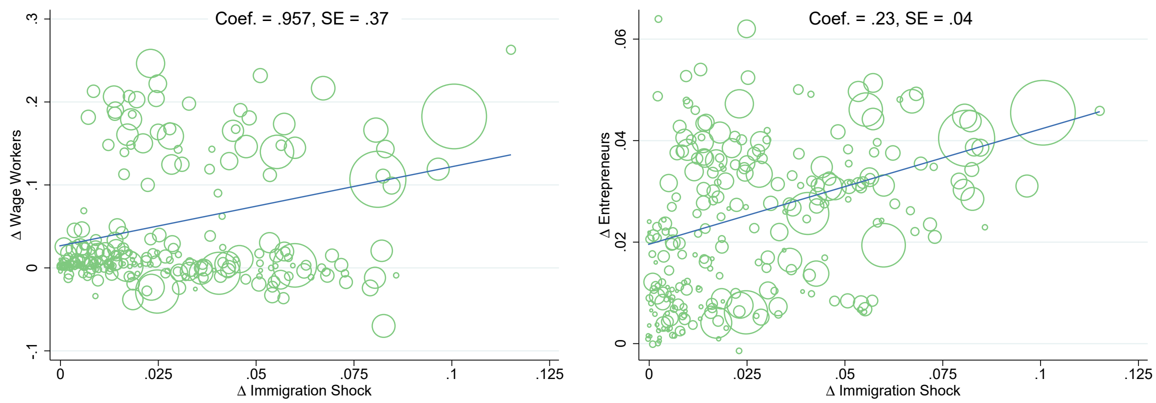
Note: The figure compares the evolution of the immigrant share of population in Spain with that of Germany, United Kingdom, the United States and the world average. Data come from the World Bank.



(a) 1999 (b) 2008

FIGURE 2: WAGE DISTRIBUTIONS IN 1999 AND 2008, FOR NATIVES AND IMMIGRANTS

Note: The figures plot kernel density estimations for wages amongst native males and females, and immigrants, using data for the years 1999 and 2008 from the MCVL. The data correspond to individuals born between 1954 and 1979 for natives, and immigrants aged 20 to 60 for each year. Daily wages are calculated in euros and are total yearly earnings divided by the number of days worked, amongst full-time wage workers who were employed all year long, as in Table 1.



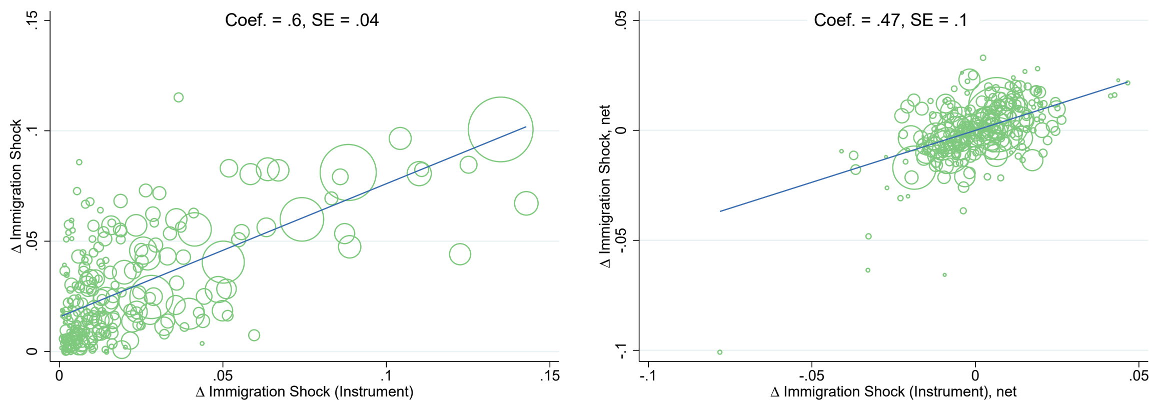
(a) Wage Workers (b) Entrepreneurs

FIGURE 3: GRAPHIC REPRESENTATION OF CORRELATIONS WITH IMMIGRANT EPISODE

Note: The figures plots scatter plots of changes in native wage workers and entrepreneurs, following Equation 2, on the immigration episode, as specified in 1. Each circle represents a local industry, with the circumference proportional to its baseline native employment. The lines of best fit come from simple regressions, also weighted by baseline native employment. In each scatter plot, I also provide the regression coefficient and its associated standard error.

Figure 4: Conditional exogeneity of the instrument, reduced form

Note: The figure shows the reduced form impact of the instrument, defined in Section 2, on preperiod (red) and study period (blue) outcomes. The plot provides the coefficient (dots) and 95% confidence intervals. The specification is similar to Equation 3, but the explanatory variable is the instrument, instead of the realised immigration episode. Reduced-form regressions are weighted by baseline province employment (either 1990 or 1999). None of the preperiod and study period coefficients are statistically different, with the exception of the entrepreneurs (p-value equal to 0.001). Preperiod statistics are calculated for the 1954-1970 cohort, as later cohorts were likely not participating in the labour market in 1990. Wages are obtained as residuals from an individual level regression on age and tenure quadratic profiles by gender, occupation and year fixed effects.



(a) Immigration Episode and Instrument

(b) Immigration Episode and Instrument, Netting Out Covariates

Figure 5: Graphic representation of the first stage

Note: The plots show scatterplots of the immigration episode on the instrument across local industries. Lines of best fit, with their coefficients and standard errors on top. The size of each circle corresponds to the baseline size of each local industry. Plot (a) provides the naive correlation. Plot (b) nets out industry and province fixed effects as well as controls from each variable. Therefore, these correspond to the coefficients in columns (1) and (4) from Table 4, respectively.

Appendix A Data and Definitions

A.1 Definition of Entrepreneurship

My definition of entrepreneurship follows strictly the dictionary definition, by which an entrepreneur is ”a person who makes money by starting or running businesses, especially when this involves taking financial risks” (OED, 2022). The self-employed individuals I identify in my data are an strict subset of people in this definition: they start and/or run businesses, and take a financial risk[44]. However, they do not necessarily overlap with innovators and successful business owners, which are what some other papers refer to as entrepreneurs.

In the MCVL, self-employed are those individuals who pay pension contributions under the self-employment scheme[45]. These are individuals who perform an economic activity for profit. For incorporated businesses, the requirement to pay pension contributions under the self-employment scheme is to have effective control of the business. In the Spanish system, an individual is attributed effective control if:

- At least half of the business capital is owned by people in the same household or family members up to second-degree relatives.
- At least a third of the business capital is under the individual ownership.
- At least a quarter of the business capital is under the individual ownership, and the individual has managerial duties.

Therefore, the self-employed category in the MCVL data captures most business owners, but one must note that in large firms, where ownership structure is usually more complex, the main owner or founder might not necessarily appear as self-employed in the data. This is less of a problem given the Spanish context, where most firms are small. For instance, by 2008, 95% of Spanish firms had less than 10 employees.

Finally, it is worth noting that the composition of self-employment during the 1999-2008 period is arguably different to that of succeeding periods, such as self-employed in the post Great Recession period[46]. This is because of two reasons. First, economic conditions during the 1999-2008 made the opportunity cost of becoming self-employed higher, as labour demand was much higher than in preceding and subsequent periods. Second, the recent rise in the gig economy and the false self-employed phenomena has increased the number of self-employed who are de facto employees. These may systematically differ

[44]Unincorporated self-employed respond to debt and liabilities accrued by their business with all their personal assets, while incorporated self-employed respond only with their business capital.

[45]Consult here for more information (in Spanish).

[46]A transformation of self-employment has been documented in the US economy by Colaiacovo et al. (2022).

from other types of self-employed individuals.

A.2 Additional information on the data

Muestra Contínua de Vidas Laborales. The MCVL is a 4% non-stratified random sample of individuals who interacted with Social Security each year, which includes information on their working histories. These are obtained from administrative records that match Social Security records with the Padrón Contínuo. The sample was first drawn in 2004, and each subsequent year some individuals leave the sample (due to not interacting with Social Security in that year). Therefore, new observations are added in order to maintain representativeness. In this project, I use the 2013 version because it is the first that includes information on whether self-employed are incorporated or not. Therefore, I use the MCVL retrospectively, as some other recent studies Iraizoz-Olaetxea (2022), which potentially looses some representativeness. This is less of a problem going back only to 1999 and focusing on the Spanish baby-boom cohort, as people born between 1954 and 1979 were likely to participate by then, but also in 2013, in the labour market[47].

Padrón Contínuo. This administrative data represent a yearly snapshot of people residing in each of the more than 8000 municipalities in Spain. This data are collected by each municipality and updated each year. The public access data contain information on province, age, nationality and place of birth. Individuals are strongly encouraged to register in a municipality, as it offers access to free public healthcare and schooling, it is the main proof of residence in the country, a main requirement to apply for legalisation, and undocumented immigrants can register as they face no threat of prosecution (Gonzalez and Ortega, 2013). Hence, the data are one of the best proxies possible administrative proxies of immigrant population, although it may miss temporary-workers or those who do not register. In any case, this data are used for the sampling of the labour force survey, which I describe below.

Encuesta de Población Activa. The Encuesta de Población Activa, or EPA, is the Spanish labour force survey, conducted quarterly on a representative sample of around 65,000 households (180,000 individuals). I use the EPA in 1999 and 2008 to calculate the shares of immigrants working in each local industry, and in 1999 exclusively to calculate baseline controls, namely the native share of high education (more than secondary education), the share of males and the share of entrepreneurs in each local industry, as well as the overall immigrant distribution across industries in order to construct the instrument. I use the EPA to calculate these controls due to the possibly more representative coverage and

[47] Amongst this group, 92% of the individuals present in the MCVL version of 2013 were present in the MCVL version of 2008.

inclusion of informal workers[48].

[48] By 2000, Bosch and Farré (2014) estimate that around 2.5% of workers were working informally, although this is possibly a lower bound.

Appendix B Additional Tables and Figures

	(1) Δ Employment	(2) Δ Wage Workers	(3) Δ Entrepreneurs	(4) Δ Unincorporated	(5) Δ Incorporated	(6) Δ Wage
Panel A: Dropping Barcelona						
Δ Immigration Shock	-0.115	-0.311	0.196*	0.011	0.185***	0.944
	(0.55)	(0.52)	(0.10)	(0.08)	(0.06)	(0.74)
First-stage KP	13.23	13.23	13.23	13.23	13.23	13.23
Observations	245	245	245	245	245	245
Panel B: Dropping Madrid						
Δ Immigration Shock	-0.566+	-0.754**	0.188***	0.019	0.169***	0.691
	(0.37)	(0.36)	(0.07)	(0.06)	(0.04)	(0.48)
First-stage KP	14.95	14.95	14.95	14.95	14.95	14.95
Observations	245	245	245	245	245	245
Panel C: Dropping Madrid and Barcelona						
Δ Immigration Shock	-0.322	-0.473	0.151+	-0.022	0.172***	0.798
	(0.59)	(0.56)	(0.10)	(0.08)	(0.06)	(0.79)
First-stage KP	10.16	10.16	10.16	10.16	10.16	10.16
Observations	240	240	240	240	240	240
Panel D: Dropping agriculture industry						
Δ Immigration Shock	0.035	-0.177	0.212***	0.019	0.193***	0.915**
	(0.33)	(0.33)	(0.07)	(0.06)	(0.04)	(0.37)
First-stage KP	18.39	18.39	18.39	18.39	18.39	18.39
Observations	200	200	200	200	200	200
Panel E: Droping bottom 5 and top 5 percentile						
Δ Immigration Shock	-0.227	0.775	0.251***	0.058	0.159***	0.797***
	(0.84)	(0.85)	(0.06)	(0.04)	(0.03)	(0.31)
First-stage KP	3.15	4.44	19.71	20.94	13.63	22.13
Observations	226	225	226	226	225	225
Panel F: Using local-industry denominator for outcome						
Δ Immigration Shock	1.224	0.815	0.409+	0.011	0.398***	0.583*
	(0.93)	(0.87)	(0.28)	(0.23)	(0.12)	(0.35)
First-stage KP	23.11	23.11	23.11	23.11	23.11	23.11
Observations	250	250	250	250	250	250
Panel G: Using local industry denominators, OLS						
Δ Immigration Shock	0.077**	0.037	0.040***	0.025**	0.015**	-0.010
	(0.04)	(0.03)	(0.01)	(0.01)	(0.01)	(0.06)
Observations	225	225	225	225	225	225
Panel H: No population weights						
Δ Immigration Shock	-0.169	-0.344	0.175***	0.086**	0.089***	-1.545+
	(0.31)	(0.29)	(0.05)	(0.04)	(0.02)	(0.98)
First-stage KP	40.90	40.90	40.90	40.90	40.90	40.90
Observations	250	250	250	250	250	250

Table B1: Robustness of the main specification

Note: This table provides the estimates from estimating β from Equation 3 using IV and controls, but providing a robustness check in each Panel. Each robustness check is explained in Section 3.4. Each column corresponds to an outcome, namely, a difference in the number of immigrant individuals in a given occupational category from 1999 to 2008 in a local industry normalised by baseline province employment for Columns (1) to (5), and to the change in log wages during the same period in Column (6). Additional information on the outcome variables, the immigration episode, or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. . Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

	(1) Δ Employment	(2) Δ Wage Workers	(3) Δ Entrepreneurs	(4) Δ Unincorporated	(5) Δ Incorporated	(6) Δ Wage
Δ Immigration Shock, 1999-2008	0.122	-0.253	0.375***	0.162**	0.213***	0.040
	(0.42)	(0.39)	(0.10)	(0.07)	(0.04)	(0.44)
Δ Immigration Shock, 1996-1999	0.111	1.656	-1.545+	-1.204*	-0.341	5.873
	(4.38)	(4.16)	(1.00)	(0.67)	(0.52)	(4.94)
First-stage KP	7.75	7.75	7.75	7.75	7.75	7.75
Observations	250	250	250	250	250	250

Table B2: Robustness of the SSIV to multiple instrumentation

Note: This table provides the estimates from estimating β from Equation 3 using the multiple instrumentation procedure suggested by Jaeger et al. (2019), i.e. controlling for lag immigration increases and instrumenting for both predicted current and lagged immigration increases. Each column corresponds to an outcome, namely, a difference in the number of immigrant individuals in a given occupational category from 1999 to 2008 in a local industry normalised by baseline province employment for Columns (1) to (5), and to the change in log wages during the same period in Column (6). Additional information on the outcome variables, the immigration episode, or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

	(1) Δ Employment	(2) Δ Wage Workers	(3) Δ Entrepreneurs	(4) Δ Unincorporated	(5) Δ Incorporated	(6) Δ Wage
Panel A: Baseline instrument						
Δ Immigration Shock	0.132	-0.100	0.232***	0.051	0.181***	0.583*
	(0.30)	(0.29)	(0.06)	(0.05)	(0.03)	(0.35)
First-stage KP	23.11	23.11	23.11	23.11	23.11	23.11
Panel B: Push-factors instrument						
Δ Immigration Shock	0.272	-0.011	0.283***	0.084+	0.199***	0.616+
	(0.33)	(0.32)	(0.07)	(0.05)	(0.04)	(0.42)
First-stage KP	17.10	17.10	17.10	17.10	17.10	17.10
Panel C: Leave one out instrument						
Δ Immigration Shock	0.036	-0.189	0.225***	0.045	0.180***	0.581
	(0.33)	(0.32)	(0.07)	(0.05)	(0.03)	(0.40)
First-stage KP	15.41	15.41	15.41	15.41	15.41	15.41
Observations	250.000	250.000	250.000	250.000	250.000	250.000

Table B3: Robustness of the SSIV to leave-one-out and push-factors specifications

Note: This table provides the estimates from estimating β from Equation 3 using different IV procedures. Panel A provides the baseline results, while Panels B and C use either a push-factors instrument or a leave-one-out, respectively. Each column corresponds to an outcome, namely, a difference in the number of immigrant individuals in a given occupational category from 1999 to 2008 in a local industry normalised by baseline province employment for Columns (1) to (5), and to the change in log wages during the same period in Column (6). Additional information on the outcome variables, the immigration episode, or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

	(1) Δ Employment	(2) Δ Wage Workers	(3) Δ Entrepreneurs	(4) Δ Unincorporated	(5) Δ Incorporated	(6) Δ Wage
Δ Immigration Shock	0.540***	0.493***	0.047^{+}	0.037	0.011^{+}	-0.837
	(0.10)	(0.08)	(0.03)	(0.03)	(0.01)	(2.12)
Mean dep. var	2.90	2.48	0.42	0.34	0.08	-4.35
First-stage KP	10.40	10.40	10.40	10.40	10.40	10.34
Observations	188	188	188	188	188	187

Table B4: Employment and wage effects of immigration on immigrants

Note: This table provides the estimates from estimating β from Equation 3 using IV and controls. Each column corresponds to an outcome, namely, a difference in the number of immigrant individuals in a given occupational category from 1999 to 2008 in a local industry normalised by baseline province employment for Columns (1) to (5), and to the change in log wages during the same period in Column (6). Additional information on the outcome variables, the immigration episode, or the specification is in Sections 1.3 and 2. Robust standard errors are provided in parenthesis. Observations are weighted by baseline local industry employment. The Table provides the mean of each dependent variable in percentage terms (multiplied by 100), as well as baseline shares of each employment category in Columns (2) to (6). Finally, I provide first-stage Kleibergen-Paap rk Wald F statistics and the total number of observations used in the estimation. Significance levels: $^{+}p < 0.15, ^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.