I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15400

Brain Refrain and Human Capital Formation in Spain

Jose Blanco-Álvarez
Christopher Parsons
Sam Tang
Yong Wang

JUNE 2022

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15400

Brain Refrain and Human Capital Formation in Spain

Jose Blanco-Álvarez
University of Santiago de Compostela

Christopher Parsons
University of Western Australia and IZA

Sam Tang
University of Western Australia

Yong Wang
City University of Hong Kong

JUNE 2022

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Brain Refrain and Human Capital Formation in Spain*

We examine how low and high skilled internal emigration causally affect investments in human capital at origin. We provide theoretical and empirical evidence of a disincentive mechanism through which individuals refrain from education should low skilled emigration prove a viable alternative. Our identification strategy leverages administrative records of labor contracts of differing skills signed at migrants' provincial destinations. We document large Brain Gain and Brain Refrain effects. Our results paradoxically demonstrate an improvement in human capital given the trajectory of the Spanish labor market over our sample period. When juxtaposed against provinces' net human capital positions however, most provinces lose.

JEL Classification: F22, 015

Keywords: brain drain, brain gain, brain refrain, internal migration, low skilled migration, high skilled migration

Corresponding author:
Christopher Parsons
Economics (UWA Business School)
University of Western Australia
35 Stirling Highway
Crawley WA 6009
Australia
E-mail: christopher.parsons@uwa.edu.au

* We thank the *Service of Development and Statistical Exploitation of Spanish National Public Service of Employment (SEPE)* for the data they generously provided. We are indebted to Caroline Theoharides for sharing code on the Fourier decomposition for an earlier version. We thank Michel Beine, Caglar Ozden, Ilse Ruyssen, Martin Fernandez Sanchez, Maurice Schiff and Jackie Wahba as well as participants at the 11th OECD-CEPII conference on immigration in OECD countries, December 2021 and the XXIII (Spanish) Conference on International Economics June 2022 for timely and valuable comments.

1 Introduction

> The [Committee] draws attention to the risk brain drain poses to the long-term sustainability of the European project. Sending regions are in a double bind...they need convergence...but are losing their skilled workforce...
>
> *The European Committee of the Regions [COR] (2020)*

The loss of native human capital resulting from emigration, the so-called Brain Drain, constitutes an enduring economic concern.[1] According to this now orthodox perspective,[2] lost skills and the elimination of positive externalities and spillovers associated with the presence of highly talented individuals lower sending nation's (or regions) welfare.[3] Traditionally, these fears have manifest most starkly in relation to developing countries, especially small countries and island states, which often face severe resource constraints.[4] The term 'Brain Drain' was first coined in relation to the emigration of UK scientists to the United States however.[5] Today, Brain Drain once again represents a key concern to policymakers in *developed* countries, given the unbalanced development of *regions*, the trajectories of which may be exacerbated by lower migration costs stemming from the freedom of movement, for example within the EU.

Advocates of the Brain Gain emphasize a (potential) countervailing (incentive) mechanism of emigration. Provided that higher levels of education grant both admission into (foreign) skilled labor markets, while also simultaneously increasing a worker's probability of migrating, additional induced investments in education at origin can more than compensate for lost skills therein yielding welfare gains.[6] If (populous) winners successfully compensate losers the global economy can benefit on net (Beine et al., 2008). Since no such compensation mechanism exists however, the implication is that many countries and regions across the world are hemorrhaging human capital.

Both of these long-established literatures fail to consider the low skilled counterpart to Brain Gain, namely the disincentive mechanism through which individuals *refrain* from investing in human capital should low skilled emigration prove a viable alternative; which we refer to as the *Brain Refrain* effect. This is particularly likely in the case of internal migration for which migration costs are low relative to international migration.

[1]See: Kwok and Leland (1982), McCulloch and Yellen (1977), J. Bhagwati and Rodriguez (1975), Hamada and Bhagwati (1975), and J. N. Bhagwati and Hamada (1974).

[2]Despite first being viewed more positively (Johnson, 1967; Grubel & Scott, 1966). For comprehensive literature reviews see Docquier and Rapoport (2012), Gibson and McKenzie (2011), and Commander et al. (2004).

[3]Through for example: fiscal deficits, stifled investment, an uncoupling of human-capital value chains and weakened public sectors; in turn linked with reduced education provision, hemorrhaged medical services, poorer institutions, ineffectual public spending and lower economic growth.

[4]Croix et al. (2014) and Docquier et al. (2007).

[5]For a discussion about the origin of the term see Godwin et al. (2009).

[6]Stark (2004), Stark and Wang (2002), Beine et al. (2001), Stark et al. (1998), Vidal (1998), Mountford (1997), and Stark et al. (1997).

First, we introduce a simple micro-founded theory that integrates both the disincentive (Brain Refrain) and incentive (Brain Gain) effects to invest in human capital – as induced by both low and high skilled emigration – into a unified framework. As a corollary of our model, there exists a threshold level in equilibrium at which the marginal individual is indifferent between investing in higher education or not, when the expected gains of high and low skilled emigration are equalized. Changes in this threshold level are therefore able to explain, at least in part, the skill composition of origins' populations. Since during our sample period, the tuition fees of tertiary education institutions increased in accordance with the austerity policies implemented at the time; we exploit these price hikes to further corroborate our theoretical model. As expected, higher fees resulted in fewer enrollments.

We subsequently test the theory at the local labor market level, exploiting rich aggregated survey and administrative data in the context of *internal* migration in Spain, a country that has experienced significant emigration away from its regions towards the country's main economic hubs (see Section 2 for details). Our resulting sample yields rich identifying variation in a balanced panel of 52 Spanish provinces across 18 years, 2001-2018. Since we observe provincial wages and emigration rates, we can directly quantify the expected gains of both high and low skilled emigration, as well as domestic skill premia; such that we need not rely upon proxy measures that are likely endogenous (Beine et al., 2008, p. 636). Since *"Ideally, the incentive effect of migration on human capital investment should be identified through the impact of migration prospects on expected returns to education"* we employ as our outcome measure provincial enrollments in upper-secondary education as in Theoharides (2018). This ensures a tighter conceptual fit between our theory and empirics since we are able to study yearly dynamics, as opposed to relying on ex-post long-run aggregates.

We establish causal estimates of both the Brain Gain and Brain Refrain effects by exploiting variation in emigrants' labor contracts disaggregated by skill level, as captured by administrative records at emigrants' provincial destinations. Our Brain Refrain estimates suggest that a one standard deviation increase in the returns to low skilled emigration, equivalent to €636, is commensurate with a 4.97% point reduction in upper-secondary school enrollments. Our Brain Gain estimates imply that a one standard deviation increase in the returns to high skilled emigration, equivalent to €1271, corresponds with a 4.05% point increase in upper-secondary school enrollments.[7] Our estimates are robust to considering alternative: geographies, outcomes and instruments.

Finally, we quantify what our estimated coefficients imply for each Spanish province on net. First, we calculate provincial-level Brain Gain and Brain Refrain effects that result from observed changes in the expected gains to high and low skilled emigration over our sample period. Paradoxically, since the expected gains to low skilled emigration actually fell in all but three provinces during this period, our estimated Brain Refrain effects result in the majority of Spanish provinces actually *gaining* human capital. When juxtaposed against provinces' prevailing human capital positions however, the majority of Spanish provinces nevertheless constitute net human capital losers.

Schiff (2006) considers a nuanced and less positive view of the Brain Gain by further considering externalities and changes in public and private expenditures.[8] In the

[7] Please refer to our summary table A.1 and column 4 in table 2.

[8] There are also well-founded reasons why fears of the Brain Drain might have been exaggerated, see for example Clemens (2007) in specific relation to to the medical brain drain.

Mexican-U.S. context, an international migrant corridor that is notably negatively selected on skills from origin (Kerr et al., 2016), McKenzie and Rapoport (2011) posit a 'disincentive' effect to invest in human capital since undocumented migration is predominant, such that the incentive to invest in human capital can be negative. Those authors argue that migration can affect education through myriad channels including: remittances, the absence of parents and the role of children in carrying out domestic duties and informal work. As such, these authors examine the aggregate effect of all such channels simultaneously.

The main contribution of this paper is in incorporating the dual incentive mechanisms to invest in human capital as catalyzed by *both* high and low skilled emigration into a unified theoretical construct; and in subsequently testing for both mechanisms in a causal empirical framework. Recent literature has rather focused on causally documenting individuals' incentives to invest in human capital along a single skill dimension in response to the international emigration of workers of specific skill types, in various cultural, geographical and historical contexts (Abarcar & Theoharides, 2020; Fernández-Sánchez, 2020; Theoharides, 2018; Shrestha, 2017; Dinkelman & Mariotti, 2016; Batista et al., 2012).

While first conceptualized at the international level, additional investments in human capital in response to variations in the expected to gains to skilled emigration have also been identified in Chinese internal migration data (De Brauw & Giles, 2017; Ha et al., 2016). Uncovering such effects at the internal level is important, since the number of internal migrants is approximately three to four times larger than their international counterparts (King & Skeldon, 2010). Despite such effects having been identified in response to both internal and international migration however, neither parallel literature incorporates both internal and international migration into the same empirical or theoretical framework. While our focus is on internal migration therefore, our rich administrative data nevertheless allow us to also control for international emigration by skill level.

Such distinctions become more blurred in the context of free mobility areas e.g. the European Union. Our paper therefore notably departs from the existing literature in examining the incentives to invest in human capital in response to emigration in a developed country context, not least since, as exemplified by the opening quote, human capital formation endures as a pivotal issue for European policy makers. Our findings therefore have potentially important ramifications for the development of nation states and economic regions alike.

2 The Spanish and European Contexts

Spain is a suitable case to study for two main reasons. First, Spain is characterized by significant economic disparities between regions, which in turn has resulted in significant inter-provincial internal migration; in part driven by the increased demand of non-tradable services (see for example Pritchett (2006)). Secondly, while starting from a low base of human capital, Spain rapidly expanded tertiary education from the 1960s onward, which itself closely paralleled the nation's development (see Figure A.1). This expansion of tertiary education was reflected in increased enrollment rates and newly established universities. The underlying logic was to expand tertiary education and today all Spanish provinces host at least one university (see Figure A.2). The consequences of this rapid expansion in tertiary education are ambiguous. On the one hand, Spain today has a higher proportion of tertiary educated than the EU15 average. On the other hand however, Spain has also witnessed the highest early school drop rate among the same group of countries. This duality of the Spanish education system is captured in Figure A.3.

Figure 2 provides some preliminary evidence as to what might be driving this apparent contradiction. In our setting of internal mobility in Spain, we observe a strong correlation between the availability of low skilled jobs (in construction) and the numbers of early school leavers over the last 20 years (see Figure 2), most notably following the onset of the Global Financial Crisis. The availability of jobs therefore, especially those for which less education is required, appear to be directly correlated with early school leave. Evidence from a special module of the 2016 labor market survey, supports this supposition since 58% stated they simply 'want[ed] to work' in lieu of studying.

Despite the dramatic expansion of tertiary education, Spain's uneven development has resulted in a significant brain drain from less developed to the more developed provinces (see Figure A.5). As a corollary of our opening quote therefore, the less developed regions of Spain, while investing further in higher education have been frustrated by the resulting emigration dynamic, which has lead to the agglomeration of *both* low and high skilled jobs in Spain's economic centers. It is now common for Spanish commentators to refer to '*España Vacía*' (empty Spain) as both low and high skilled Spaniards vacate their provincial origins in search of superior labor market opportunities in the country's economic hubs. Such agglomerations of human capital have been widely documented in recent years (Kerr et al., 2017, 2016; Artuc et al., 2015; Özden et al., 2011) although the focus has been on concentrations of high skilled workers. In this paper, we provide at least a partial explanation for these observed patterns since we capture both low and high skilled migration dynamics in a unified framework.

Human capital mobility is now front and center in the minds of developed country policy makers, not least those based in the EU. Such concerns have traditionally been voiced in regards to international emigration from EU countries. More recently however, internal mobility within EU countries have evoked fears of demographic and skill deficits. At the inaugural session of the seventh term of the European Committee of the Regions convened in February 2020 for example, a declaration on the necessity to address the brain drain at 'every level' was approved. In speaking to these EU wide concerns of depopulated regions and growing disparities and agglomerations of economic activities, we study internal mobility within Spain – while controlling for international emigration by skill – between which many parallels can be drawn in the sense that international emigration with few or no immigration barriers (the case of EU) is comparable to the process of internal migration within national borders.

3 Theoretical model

In this section we formalize our arguments into a simple theoretical model. At its core is the supposition that individuals consider their expected earnings (from both low and high skilled emigration) before deciding whether to invest or not in education. This fundamental idea, viewing education as an investment, lies at the heart of the very concept of human capital (Heckman, 2000; Becker, 1964; Schultz, 1961). Migrants therefore weight up the costs and benefits of being located at home else abroad, in which case particular skills will be more or less valued (Sjaastad, 1962). The key insight of the Brain Gain literature (see for example: Stark and Wang (2002) and Beine et al. (2001)) is that individuals consider emigration prospects when making their decisions whether to invest or not in human capital.

Our model distinguishes between two types of individuals, high skilled (s) and low skilled (u). Low skilled workers earn a wage W_u in their home province else W_u^* if they successfully migrate. Similarly, high skilled individuals earn a wage W_s while employed at home or W_s^* if they migrate. High skilled wages are assumed to be higher than low skilled wages i.e. $W_s > W_u$, a supposition corroborated by our data.

The cost of education of individual i is given by C_ι. Aside from a fixed amount, this cost depends upon individuals' unobserved ability levels, as well as socioeconomic factors including: parents' status, financial or otherwise, the provision of public facilities, grants and education subsidies. For simplicity, we suppose a uniform distribution in an interval $[C_{min}, C_{max}]$. We suppose that the probabilities of emigration for both low and high skilled individuals are exogenous.

Since we focus on internal migration, these probabilities cannot be interpreted as the result of immigration policies or comparable restrictions. Instead, we interpret these probabilities as individuals' subjective prospects to emigrate as determined by physical, cultural and psychic factors (Sjaastad, 1962). While many individuals search at least initially for employment close to home, job searches nevertheless extend to other destinations about which information will probably be scarcer. This is more likely to be the case should employment prospects be lower at origin. While every individual holds different preferences with regards emigration, we assume that the observed province-to-province emigration stocks are good proxies for these probabilities for an average individual from each provincial origin. We further assume that the probabilities of emigration are higher for more highly skilled workers such that: $P_s > P_u$, which is also supported by our data, please refer to Table A.1.

In summary, before deciding whether to invest or not in education, individuals naturally consider the job opportunities available to them nationally, not least since migration, at least in Spain, is free from a policy perspective. Individuals make this decision in the presence of imperfect information, facing varying distances (physical, cultural and psychological).[9] We subsequently express the expected earnings for low (1) and high skilled (2) individuals as:

$$E(W_u) = (1 - P_u)W_u + P_u W_u^* \tag{1}$$

$$E(W_{\iota,s}) = (1 - P_s)W_s + P_s W_s^* - C_\iota \tag{2}$$

Individuals are risk neutral such that they maximize their life time income. This assumption has been ubiquitously implemented in the literature to date (Beine et al., 2008; Stark & Wang, 2002; Beine et al., 2001; Vidal, 1998; Mountford, 1997). The condition for both types of workers to be indifferent between investing in education or refraining from doing so is therefore:

[9] In 2016, 43.72% of inactive and unemployed young Spaniards (aged 16 to 34 years old) declared their willingness to change their place of residence for the sake of employment (INE, 2017).

$$(1-P_u)W_u + P_u W_u^* = (1-P_s)W_s + P_s W_s^* - C_\iota \quad (3)$$

The marginal worker who remains indifferent between investing (denoted by ι^*) is therefore given by:

$$C^* \equiv C_{\iota^*} = P_s(W_s^* - W_s) - P_u(W_u^* - W_u) + (W_s - W_u) \quad (4)$$

There exists a unique value of $C_{\iota^*} = C^*$ that represents the relevant threshold delineating between investing or not. We suppose that $C_{min} < C^* < C_{max}$.[10] All individuals whose education cost lie below this cut-off will invest and become educated and therefore highly skilled. In contrast, all individuals lying above this threshold will remain low skilled (see Figure 1).

We could further assume that liquidity and credit constrains operate, in which case we would define a positive parameter δ, which represents a minimum consumption level required in the first period, therefore displacing the effective cost to invest in education to $[C_{min+\delta}, C_{max+\delta}]$. Human capital will therefore be lower for any given value of C^* (see Figure 1). In other words, facing the same set of expected gains from skilled and unskilled migration, the resulting human capital level would be lower.

Figure 1 graphically depicts our model. Note that the absolute number of individuals who decide to invest in education (Q_s) is given by the difference $C^* - C_{min}$, while the absolute number of low skilled (Q_u) is given by $C_{max} - C^*$. A rise (decrease) in the probability to emigrate or in foreign low skilled wages will decrease (increase) C^* and *ceteris paribus* result in a lower (higher) number of individuals seeking education. Vice versa, a rise in high skilled wages in foreign regions or a higher premium at home will raise the level of C^* and so increase the number of high skilled individuals.

The aggregate human capital formed in country S prior to migration is:

$$H = (C^* - C_{min})H_s + (C_{max} - C^*)H_u \quad (5)$$

where $C^* - C_{min}$ is the mass of skilled workers and $C_{max} - C^*$ their unskilled counterparts. The average pre-migration human capital in the source country is therefore given by:

$$h = \frac{(C^* - C_{min})}{(C_{max} - C_{min})}H_s + \frac{(C_{max} - C^*)}{(C_{max} - C_{min})}H_u = \lambda_s H_s + \lambda_u H_u \quad (6)$$

where $\lambda_s = \frac{(C^* - C_{min})}{(C_{max} - C_{min})}$ and $\lambda_u = \frac{(C_{max} - C^*}{(C_{max} - C_{min})}$ are the pre-migration proportions of skilled and unskilled workers respectively and $\lambda_s + \lambda_u = 1$. Since $H_s > H_u$, the average human capital formed in the source country in (3) increases with C^*. Consequently, (1) predicts that the pre-migration average human capital level in the source country correlates positively with the expected gain from skilled migration, $P_s(W_u^* - W_s)$ and the skilled wage premium $(W_s - W_u)$, but negatively with the expected gain from unskilled migration, $P_u(W_u^* - W_u)$.

Consequently, our key empirical model based on Equation 6 is:

$$h_{it} = \beta_0 + \beta_1 T_{it,s} + \beta_2 T_{it,u} + \beta_3 S_{it} + \beta_4 X_{it} + \mu_{it} \quad (7)$$

[10] If not, all individuals would become high skilled or remain low skilled, which is unrealistic.

where h_{it} is the pre-migration average human capital in i at time t, $T_{is} = P_{it,s}(W^*_{t,s} - W_{it,s})$ and $T_{i,u} = P_{it,u}(W^*_{t,u} - W_{it,u})$ represent the expected gains from skilled and unskilled migration, respectively, $S_{it} = (W_{it,s} - W_{it,u})$ and X_{it} is a vector of control variables. The model implies that $\beta_1 > 0$ (Brain Gain effect), $\beta_2 < 0$ (Brain Refrain effect) and $\beta_3 > 0$ (positive skill premia).

4 Data

4.1 Education

The dependent variable typically employed in brain gain studies is the average level of human capital in a country or region as enumerated in census data or other comparable sources e.g. Barro and Lee (2013). In our case, we rather leverage educational attainment recorded in the EPA *Encuesta de Población Activa*, (the Spanish Labor Survey), to compute the percentage of those with tertiary education.[11] These data from the EPA survey are representative at the provincial level, while providing other relevant details of the native population (province of birth, province of residence, age and gender).

Our preferred outcome variable following for example Theoharides (2018) and Ha et al. (2016)) is the *current* enrollment rate in upper-secondary education. This choice is governed by our ambition to employ data that allow us to best identify the mechanisms at play i.e. whether students choose to sit exams to enter university else employment, when faced with differing expected gains from low and high skilled emigration. Annual gross flow data are necessarily superior for capturing these dynamics relative to net long-run averages.

We further refine our enrollment measure by considering those sitting the *non-mandatory* "Bachillerato", the main objective of which is to prepare students to undertake the national exam to enter university. For the sake of robustness we also directly employ the gross rate of students taking the national exam to enter university. The only disadvantage of these data is that they are only available at the level of Spanish autonomous region (as opposed to province).

4.2 Emigration

Internal provincial emigration rates by skill level are constructed from emigrant stock data derived from the EPA. As opposed to relying on decennial census data, we rather compute emigrant stocks by skill level using aggregated quarterly data from 2001 to 2018.

The stock of *international* emigrants by province of birth is also employed as a control in our model as recorded in Spain's register of citizens living abroad. Registrations are strongly recommended, not least should consular assistance be required or should individuals want to vote in Spanish elections. Although the Spanish National Institute of Statistics (INE) only publish these data from 2009 onward, we were fortunate to take receipt of a more complete dataset from INE, which comprises Spaniards aged 18-64 by province of birth and country of residence as from the 1st January annually from 2002 until 2018.

[11] We follow this procedure to replicate Beine et al. (2008) in A.3.

Provincial international emigration rates by skill are calculated as the shares of total emigration from individual provinces to each foreign destination. These annual shares are subsequently merged with decennial data from the OECD DIOC and DIOC-E databases for the years 2000 and 2010 to further delineate between high and low skilled emigrants. While these measures serve as proxy measures, these inclusions represent a departure from the existing literature that focuses on either internal or international migration in isolation.

4.3 Wages

In theory, wage differentials represent the primary determinant of emigration (see Grogger and Hanson (2011)). The paucity of comparable international wage data however, explains the omission of accurate wage measures in many brain gain studies. A distinct advantage of working with our internal migration data is the availability of annual provincial wage data. Low and high skilled wages were derived from data obtained from the Spanish Tax Agency (which reports the declared annual gross income from wages by broads groups and province). Following Grogger and Hanson (2011), we estimate the high skilled wage of each province as the wage equal to the 80th percentile and the low skilled wage as the equivalent to the 45th percentile. Note that we use a higher threshold for low skilled wages (Grogger and Hanson used the 20th percentile) since we distinguish between two educational categories: high (those with tertiary education) and low (those without tertiary education), whereas Grogger and Hanson delineate three such groups.

4.4 Other control variables

Unemployment is also computed from EPA data (see Figure 2). In the Spanish context since employment opportunities are relatively scarce, the opportunity cost of education decreases when unemployment rises, which in turn increases enrollment rates and vice-versa.

Population Density data rather derive from the INE *Instituto Nacional de Estadística* (Spanish National Office of Statistics). This is included since more urbanized or densely populated areas likely have superior access to education. Neither can we disregard any potential agglomeration effects, for example the fact that rural schools tend to have lower student:teacher ratios.

Tuition fees data, the price of public universities in Spain, is set by the regional governments, in a price range previously approved by Ministry of Education. We compute these data for the 17 CCAA plus the 2 autonomous cities of Ceuta and Melilla.

5 Empirical Strategy

We estimate the following equation that is based on Equation 7:

$$H_{it} = \beta_0 + \beta_1 \Upsilon_{it,s} + \beta_2 \Upsilon_{it,u} + \beta_3 \Omega_{it} + \beta_4 X_{it} + \mu_{it} \tag{8}$$

Where: H_{it} is the average annual human capital level in province i, at time t, remaining after the emigration of both low and high skilled individuals. $\Upsilon_{it,u}$, $\Upsilon_{it,s}$ are the expected gains of low and high skilled migration given respectively by the expressions

$\Upsilon_{it,u} = P_{it,u}(W^*_{t,u} - W_{it,u})$ and $\Upsilon_{it,s} = P_{it,s}(W^*_{t,s} - W_{it,s})$. The average wage is calculated from the set of possible destinations W^*_t. Ω_{it} is the skill premium that those educated enjoy (the difference between high and low skilled wage) in province i, at time t. X_{it} is a vector of control variables. Including the unemployment rate, the density of population and a proxy that captures the extent of international emigration by province of origin.

As detailed in the theoretical section our model implies that $\beta_1 > 0$ (Brain Gain effect), $\beta_2 < 0$ (Brain Refrain effect) and $\beta_3 > 0$ (positive skill premia).

5.1 Identification strategy

Brain drain/gain studies are plagued by endogeneity concerns. Reverse causality means it remains inconclusive as to whether emigration rates affect investments in human capital or vice-versa. Arguably a greater threat to causal identification however, is the threat of omitted variables that are simultaneously correlated with both human capital and emigration rates.

Following in this tradition, we propose two sets of instruments for our low and high skilled emigration rates, both of which derive from alternative measures of labor demand. Under Spanish law all labor contracts need to be registered before any contractual relationship can legally begin. We obtained these data from SEPE *Servicio Público de Empleo Estatal* ("Public Service of Employment") that comprise all contracts signed between 2001 and 2018, to use as our first measure of labor demand.[12] These data detail the province where workers lived at the time of their signing, in tandem with the province where their future place of work is located. When the two locations differ, we can logically conclude labor emigration occurred.

Formally, we have $\Gamma^j_{it,\sigma}$, where Γ is the number of contracts signed with place of work in province j (destination) by workers with place of residence in province i (origin) in year t. The skill level of the worker is given by σ and takes the values ($\sigma \in [p, s, t]$) that are primary, secondary and tertiary education respectively.[13] We denote the sum of primary and secondary contracts as our measure of low skilled labor.

Our instrument relying on these contract data is calculated as the provincial share over the annual number of contracts with inter-provincial migration, by each skill level.[14] This proportion is calculated as:

$$Z_{it,\sigma} = \frac{\sum^j_{j \neq i} \Gamma^j_{it,\sigma}}{\sum_i \sum^j_{j \neq i} \Gamma^j_{it,\sigma}} \tag{9}$$

Our second instrument set is based on the familiar Bartik-style instruments or "migration demand index", following in this context from for example the work of Theoharides (2018). To this end, we employ the bilateral aspects of our emigration dataset in order to calculate initial provincial shares (relative to the national total) in the initial year, which is subsequently shifted by the growth in national migration. This allows us to estimate a "demand index" for emigration from each source province in the following years.

[12] Please note: 1) the shadow economy and informal workers as well as self-employees are not registered in these data and 2) The same worker can sign more than one contract in any given year.

[13] We omit all contracts signed by workers who are not Spanish.

[14] Since we are dealing with internal migration, the total emigration will be equal to total immigration each year.

This migrant demand index is constructed with the following equation:

$$D_{it,\sigma} = \sum_{j} \sum_{i \neq j}^{i} M_{it,\sigma}^{j} \times \frac{M_{i0,\sigma}^{j}}{\sum_{i \neq j}^{i} M_{i0,\sigma}^{j}} \tag{10}$$

Where: $D_{it,\sigma}$ is the predicted number of emigrants M with skill level $\sigma$from province of birth i in year t. This number is estimated for each year t summing up the stock of predicted immigrants with skill level σ from i in all destination provinces j. By each year t and destination province j, we estimate the number of immigrants from i with skill level σ multiplying the total number of immigrants of that skill level from all origins ($\sum_{i \neq j}^{i} M_{it,\sigma}^{j}$), by the proportion that the immigrants from i represent out of total immigration in baseline year 0 ($M_{i0,\sigma}^{j} / \sum_{i \neq j}^{i} M_{i0,\sigma}^{j}$). In our case, we calculate this for both low (U) and high skilled (S) individuals using 2001 as our baseline year. We subsequently use these predicted emigration stocks to calculate our predicted low and high emigration rates. We confirm that our migration index is an excellent predictor of actual emigration rates, as depicted in figure A.6.

6 Results

Despite having migration and wage data for the period 1999 to 2019, data for our preferred instrument (labor contract data) are only available from 2001 to 2018. For the sake of standardization, we employ this reduced sample. Full sample results are presented in the annex (see table A.2).

6.1 Benchmark Results

Table 1 presents our benchmark results of estimating Equation 7, first with OLS in Columns 1-3 and then with 2SLS first using our instruments as described in Equations 9 and 10. Across all estimations we employ the annual enrollment rates in upper-secondary education as our dependent variable.

While controlling for any incentive effects from international migration on investments in human capital, our baseline OLS results lend support for the existence of both a Brain Gain (resulting from the incentives from high skill migration) as well as a Brain Refrain effect (resulting from the incentives from low skill migration). The signs and statistical significance of our estimates are stable across specifications variously including province, year and province and year fixed effects.

Column 4 in Table 1 presents the results from our preferred estimation, the first stages of which are presented in Table A.5. In comparison to the OLS estimates, our IV results pertaining to both the Brain Gain and Brain Refrain are significantly larger. In other words, our OLS estimates are downward biased. Overall we find that an increase in the expected gains from low skilled migration dwarf those from a comparable increase in high skilled migration. In other words, at least in our case, that of internal mobility in Spain, we find that our estimates of Brain Refrain dominate those of Brain Gain, which in exacerbates concerns of the Brain Drain.

To facilitate the interpretation of our results, we report our results in standardized beta coefficients. In our preferred specification – our IV regressions that implement labor contract data as instruments – a one standard deviation increase in the expected gains from low skilled migration (equivalent to 636€) causes a decrease of about twice the standard deviation in upper-secondary enrollment rates equivalent to 6.58% points. This apparently large effect is offset by the brain gain effect, for which we estimate a one standard deviation increase in the expected gains to high skilled emigration (equivalent to 1300€) raises enrollment rates by 5.26 %.

6.2 Changes in C^* (equilibrium threshold)

As a response to the 2008 economic and financial crisis, the Spanish government liberalized the price of university tuition fee, which subsequently resulted in price hikes (see figure A.7).[15] In this sub-section, we therefore employ university tuition fees as a covariate in our regressions, in order to further corroborate our underlying theoretical model, since increases in tuition fees should directly impact the C^* term in our theoretical model (see figure 1).

Our results when including tuition fees (at the CCAA level) are provided in table 3. Confirming our theoretical priors, the higher university tuition is, the lower the corresponding enrollments in upper-secondary education.

6.3 Robustness

For the sake of robustness we compute enrollment and migration rates, wages and other control variables at the CCAA level, as opposed to the provincial level, so as to examine if our effects endure at differing levels of geography. At this geographical level we are then further able to employ an alternative dependent variable, namely the "gross rate of the population passing the university entrance exam".

Our results are presented in table 3. Column 1 presents our OLS estimates based on the inclusion of region and time fixed effects. Our preferred IV estimates employing our labor contract data as instruments are presented in column 2, while our shift-share instrument results are provided in column 3. Columns 4 to 6, present the corresponding estimates when employing our alternative dependent variable.

These results are consistent with our benchmark regression when instead employing data on 52 Spanish provinces. Our preferred specification estimates imply that a €517 increase in the expected gains from low skilled migration causes a 1.9% fall in upper-secondary school enrollments.

6.4 Provinces' Net Human Capital Position

In this section we pose the question: Given the wage dynamics observed over our sample period, what were the net provincial human capital outcomes implied by our analysis? Panel A of Figure A.8 shows, for each of Spain's provinces, the implied *net* impact on upper-secondary investments in human capital following the change in *actual* expected gains to low and high skilled emigration over our sample period. Paradoxically, this Figure shows that the majority of provinces witnessed an *improvement* (38/52) in their human capital position. This is largely due to the fact that over our sample period the expected gains to low skilled migration actually *fell* in all but three provinces.

[15] As figure A.1 shows, the majority of students attend to public universities. The tuition fees of these universities are fixed annually by the regional governments, after consultation with the Ministry of Education. The price hikes followed the implementation of Spain's austerity measures.

Panel B of Figure A.8 goes yet further, juxtaposing the results presented in Panel A of Figure A.8, against provinces' net human capital position in 2018. To this end, we first take our estimates of the net impact of provincial low and high skilled emigration on on upper-secondary school enrollments. We subsequently apply provincial pass rates in order to calculate the implied changes in provincial human capital levels that result from low and high skilled emigration. We then calculate, as of 2018, the net human capital position of each province; to which we apply our implied changes in human capital resulting from our two countervailing mechanisms. Despite the majority of provinces actually gaining human capital as a consequence of Brain Gain and Brain Refrain, the majority of provinces are still nevertheless significantly worse off on net.

7 Conclusion

In this paper, we first provide a unified theoretical framework which encompasses negative and well as positive incentives to invest in human capital as a result of prospects for low and high skilled emigration. We subsequently test the theoretical predictions of our model using rich administrative and survey data. To address endogeneity concerns, we employ two sets of instrumental variables that derive from alternative measures of labor demand. Our preferred instrument leverages administrative data on labor contracts signed by all Spanish workers over our study period. Our baseline results serve to highlight the existence of both Brain Gain and Brain Refrain effects in the case of internal mobility, although the latter effects dominate the former in our specific case of Spain. Given that the majority of Spanish provinces witnessed a decrease in the expected gains to low skilled migration over our sample period however, we subsequently show that most provinces benefited from our newly identified Brain Refrain channel. When our results are juxtaposed against the backdrop of existing Spanish internal human capital mobility however, the majority of provinces nevertheless witnessed significant decreases in their human capital levels, reinforcing once traditional fears of the Brain Drain. Our results serve to highlight this issue in the case of *developed* country regions, which in turn has implications for areas of freedom of movement, not least the EU.

References

Abarcar, P., & Theoharides, C. (2020). Medical worker migration and origin-country human capital: Evidence from us visa policy. *The Review of Economics and Statistics*, 1–46.

Artuc, E., Docquier, F., Özden, Ç., & Parsons, C. (2015). A global assessment of human capital mobility: The role of non-oecd destinations. *World Development*, *65*, 6–26.

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, *104*, 184–198. https://doi.org/10.1016/j.jdeveco.2012.10.001

Batista, C., Lacuesta, A., & Vicente, P. C. (2012). Testing the 'brain gain' hypothesis: Micro evidence from Cape Verde. *Journal of Development Economics*, *97*, 32–45. https://doi.org/10.1016/j.jdeveco.2011.01.005

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education* (1st ed.). Columbia University Press.

Beine, M., Docquier, F., & Rapoport, H. (2001). Brain drain and economic growth: Theory and evidence. *Journal of Development Economics*, *64*(1), 275–289. https://doi.org/10.1016/S0304-3878(00)00133-4

Beine, M., Docquier, F., & Rapoport, H. (2008). Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *The Economic Journal*, *118*(528), 631–652. https://doi.org/10.1111/j.1468-0297.2008.02135.x

Bhagwati, J. N., & Hamada, K. (1974). The brain drain, international integration of markets for professionals and unemployment: A theoretical analysis. *Journal of Development Economics*, *1*(1), 19–42. https://doi.org///doi.org/10.1016/0304-3878(74)90020-0"

Bhagwati, J., & Rodriguez, C. (1975). Welfare-theoretical analyses of the brain drain. *Journal of development Economics*, *2*(3), 195–221.

Clemens, M. (2007). Do visas kill? Health effects of African health profesional emigration. *Center for Global Development Working Paper*, *114*. https://www.cgdev.org/publication/do-visas-kill-health-effects-african-health-professional-emigration-working-paper-114

Commander, S., Kangasniemi, M., & Winters, L. (2004). The Brain Drain: Curse or Boon? A Survey of the Literature. In R. E. Baldwin & L. A. Winters (Eds.), *Challenges to Globalization: Analyzing the Economics* (pp. 235–278). University of Chicago Press. https://www.nber.org/chapters/c9540

COR. (2020, February 12). *Brain drain in the EU: Addressing the challenge at all levels*. https://cor.europa.eu/en/our-work/Pages/OpinionTimeline.aspx?opId=CDR-4645-2019

Croix, D., Docquier, F., Schiff, M., et al. (2014). Brain drain and economic performance in small island developing states. *Population Economics*, 123–144.

De Brauw, A., & Giles, J. (2017). Migrant opportunity and the educational attainment of youth in rural china. *Journal of Human Resources*, *52*(1), 272–311.

Dinkelman, T., & Mariotti, M. (2016). The long-run effects of labor migration on human capital formation in communities of origin. *American Economic Journal: Applied Economics*, *8*(4), 1–35.

Docquier, F., Lohest, O., & Marfouk, A. (2007). Brain drain in developing countries. *The World Bank Economic Review*, *21*(2), 193–218.

Docquier, F., & Rapoport, H. (2012). Globalization, Brain Drain, and Development. *Journal of Economic Literature*, *50*(3), 681–730. https://doi.org/10.1257/jel.50.3.681

Eugenia Núñez, C. (2005). Educación. *Estadísticas históricas de España. Siglos XIX - XX.* (2nd ed., pp. 155–244). Fundación BBVA.

Fernández-Sánchez, M. (2020). Mass Emigration and Human Capital over a Century: Evidence from the Galician Diaspora. *Job Market Paper*. https://www.dropbox.com/s/4fr551b872l10g8/ch1.jmp_mass_emigration_and_human_capital.pdf?dl=0

Gibson, J., & McKenzie, D. (2011). Eight questions about brain drain. *Journal of Economic Perspectives*, *25*(3), 107–28.

Godwin, M., Gregory, J., & Balmer, B. (2009). The Anatomy of the Brain Drain Debate, 1950–1970s: Witness Seminar. *Contemporary British History*, *23*(1), 35–60. https://doi.org/10.1080/13619460801990088

Grogger, J., & Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Symposium on Globalization and Brain Drain*, *95*(1), 42–57. https://doi.org/10.1016/j.jdeveco.2010.06.003

Grubel, H. G., & Scott, A. D. (1966). The International Flow of Human Capital. *The American Economic Review*, *56*(1/2), 268–274.

Ha, W., Yi, J., & Zhang, J. (2016). Brain drain, brain gain, and economic growth in China. *China Economic Review*, *38*, 322–337. https://doi.org/10.1016/j.chieco.2015.02.005

Hamada, K., & Bhagwati, J. (1975). Domestic distortions, imperfect information and the brain drain. *Journal of Development Economics*, *2*(3), 265–279. https://doi.org/10.1016/0304-3878(75)90005-X

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, *54*(1), 3–56. https://doi.org/10.1006/reec.1999.0225

INE. (2017, June 9). *Encuesta de Población Activa (EPA). Módulo sobre los jóvenes en el mercado laboral. Año 2016.* https://www.ine.es/prensa/epa_2016_m.pdf

Johnson, H. G. (1967). Some Economic Aspects of Brain Drain. *The Pakistan Development Review*, *7*(3), 379–411.

Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2016). Global Talent Flows. *Journal of Economic Perspectives*, *30*(4), 83–106. https://doi.org/10.1257/jep.30.4.83

Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2017). High-Skilled Migration and Agglomeration. *Annual Review of Economics*, *9*(1), 201–234. https://doi.org/10.1146/annurev-economics-063016-103705

King, R., & Skeldon, R. (2010). 'Mind the Gap!' Integrating Approaches to Internal and International Migration. *Journal of Ethnic and Migration Studies*, *36*(10), 1619–1646.

Kwok, V., & Leland, H. (1982). An economic model of the brain drain. *The American Economic Review*, *72*, 91–100.

McCulloch, R., & Yellen, J. L. (1977). Factor Mobility, Regional Development, and the Distribution of Income. *Journal of Political Economy*, *85*(1), 79–96. https://doi.org/10.1086/260546

McKenzie, D., & Rapoport, H. (2011). Can migration reduce educational attainment? Evidence from Mexico. *Journal of Population Economics*, *24*(4), 1331–1358. Retrieved February 3, 2020, from www.jstor.org/stable/41488354

Mountford, A. (1997). Can a brain drain be good for growth in the source economy? *Journal of Development Economics*, *53*(2), 287–303. https://doi.org///doi.org/10.1016/S0304-3878(97)00021-7

Özden, Ç., Parsons, C. R., Schiff, M., & Walmsley, T. L. (2011). Where on earth is everybody? the evolution of global bilateral migration 1960–2000. *The World Bank Economic Review*, *25*(1), 12–56.

Pritchett, L. (2006). *Let their people come: Breaking the gridlock on international labor mobility*. Center for Global Development; Distributed by Brookings Institution Press.

Schiff, M. (2006). Brain Gain: Claims about its size and Impact on welfare and growth are greatly exaggerated. In Ç. Özden & M. Schiff (Eds.), *International migration, remittances, and the Brain Drain* (pp. 201–225). The World Bank & Palgrave Macmillan. http://hdl.handle.net/10986/6929

Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review*, *51*(1), 1–17.

Shrestha, S. A. (2017). No Man Left Behind: Effects of Emigration Prospects on Educational and Labour Outcomes of Non-migrants. *The Economic Journal*, *127*(600), 495–521. https://doi.org/10.1111/ecoj.12306

Sjaastad, L. A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy*, *70*(5-2), 80–93.

Stark, O., & Wang, Y. (2002). Inducing human capital formation: Migration as a substitute for subsidies. *Journal of Public Economics*, *86*(1), 29–46. https://doi.org///doi.org/10.1016/S0047-2727(01)00104-9

Stark, O., Helmenstein, C., & Prskawetz, A. (1997). A brain gain with a brain drain. *Economics Letters*, *55*(2), 227–234. https://doi.org///doi.org/10.1016/S0165-1765(97)00085-2

Stark, O., Helmenstein, C., & Prskawetz, A. (1998). Human capital depletion, human capital formation, and migration: A blessing or a "curse"? *Economics Letters*, *60*(3), 363–367.

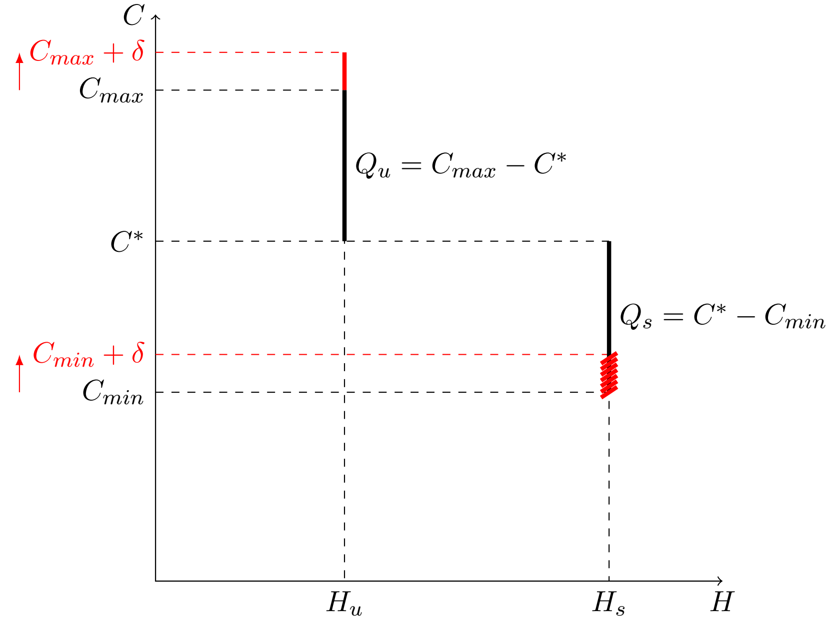
Stark, O. (2004). Rethinking the Brain Drain. *World Development*, *32*(1), 15–22. https://doi.org///doi.org/10.1016/j.worlddev.2003.06.013

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D. W. Andrews (Ed.), *Identification and inference for econometric models* (pp. 80–108). Cambridge University Press. http://www.economics.harvard.edu/faculty/stock/files/TestingWeakInstr_Stock%5C%2BYogo.pdf

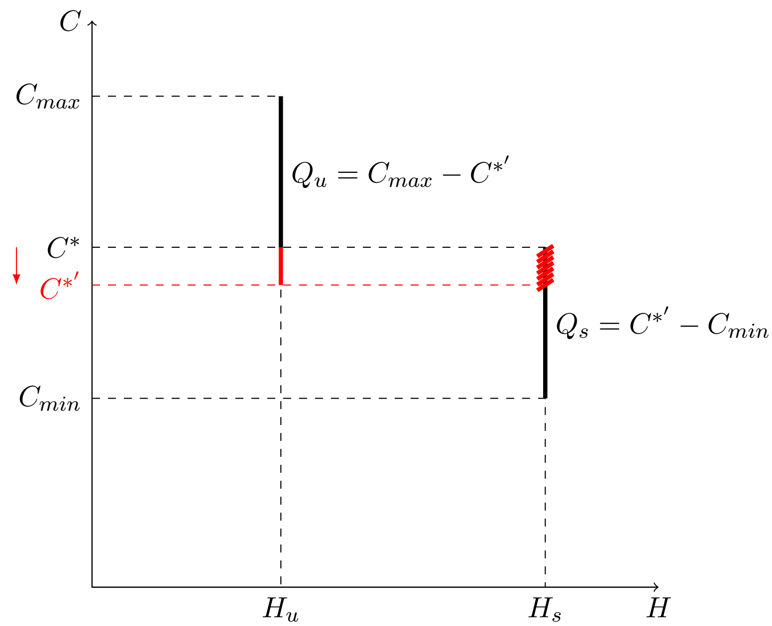
Theoharides, C. (2018). Manila to Malaysia, Quezon to Qatar: International migration and its effects on origin-country human capital. *Journal of Human Resources*, *53*(4), 1022–1049. https://doi.org/10.3368/jhr.53.4.0216.7714R1

Vidal, J.-P. (1998). The Effect of Emigration on Human Capital Formation. *Journal of Population Economics*, *11*(4), 589–600.

List of Figures

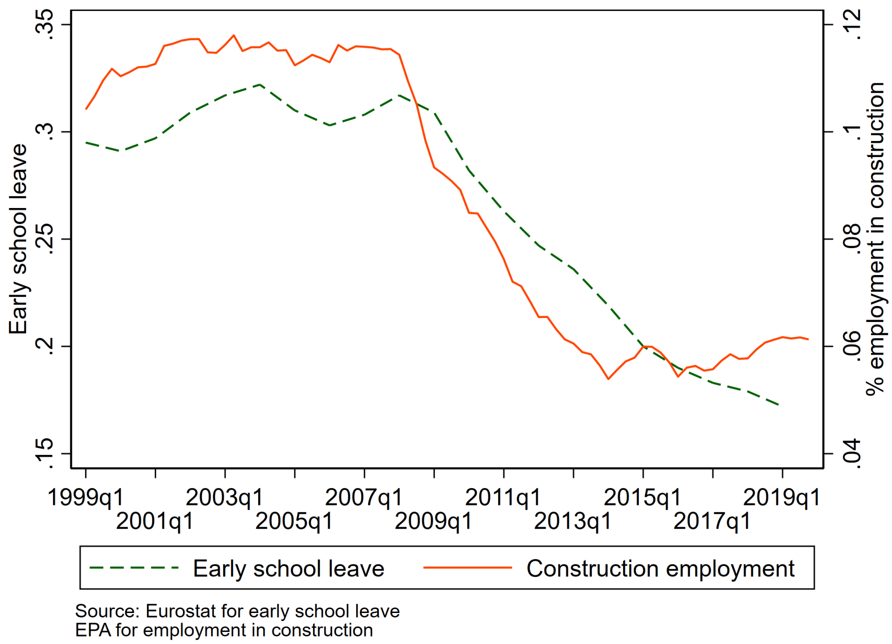


(a) A liquidity and credit constraint (δ) will decrease the number of high skilled

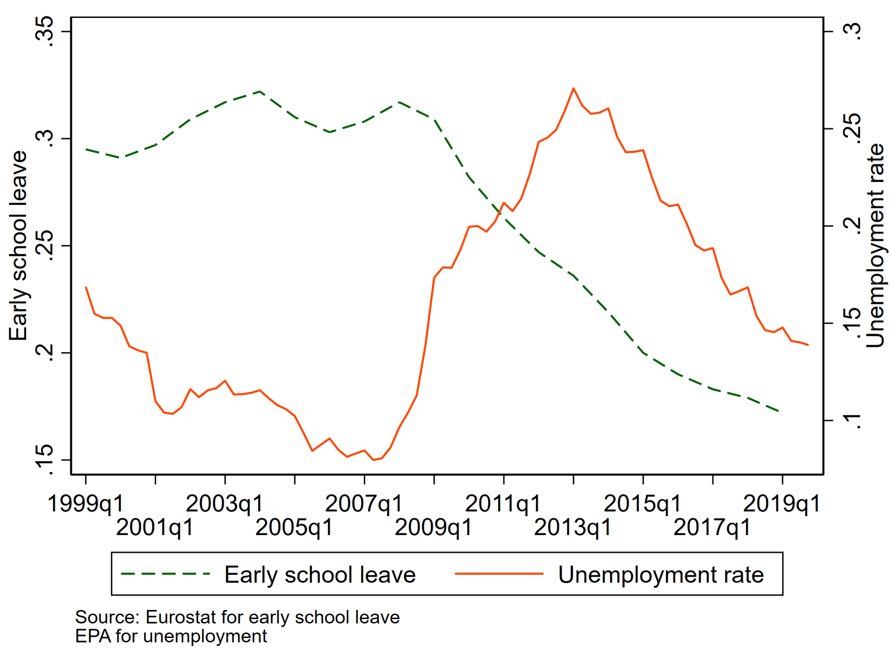


(b) After an increase in P_u, W_u or W_u^* or a decrease in P_s, W_s or W_s^*, the number of low skilled (Q_u) increase while high skilled decrease (Q_s)

Figure 1: Human capital formation in our model, see eq. 4



(a) Relationship between early school leave and employment in construction



(b) Relationship between early school leave and unemployment rate

Figure 2: Relationship between early school leave and scarcity of jobs

List of Tables

Table 1: Benchmark regressions

	enrolment rate in upper-secondary				
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV
Exp. gains from hs mig. (€)	0.045	0.592***	0.192**	1.639***	0.940***
	[0.43]	[4.60]	[2.24]	[3.30]	[5.84]
Exp. gains from ls mig. (€)	-0.292***	-0.382***	-0.170*	-2.037***	-0.448***
	[2.78]	[3.39]	[1.79]	[4.56]	[4.78]
Premium skill (€)	-0.012	0.553***	0.248***	0.940***	0.505***
	[0.16]	[5.94]	[3.09]	[5.00]	[6.92]
Pop. density (hab/km2)	-0.816***	-0.467***	-0.737***	-0.935***	-0.716***
	[3.22]	[4.59]	[4.29]	[2.87]	[4.21]
Unemployment rate (%)	0.544***	0.187	0.224***	0.329***	0.215***
	[12.70]	[1.39]	[3.34]	[4.56]	[5.24]
International hs mig. rate (%)	-0.451***	-0.509***	0.001	0.023	0.012
	[4.98]	[3.29]	[0.01]	[0.30]	[0.24]
International ls mig. rate (%)	0.791***	0.740***	-0.107	-0.012	-0.157*
	[8.52]	[3.32]	[0.84]	[0.09]	[1.86]
Region fe	Prov		Prov	Prov	Prov
Time fe		Year	Year	Year	Year
R^2	0.821	0.627	0.908	0.785	0.890
F	.	74.19	.	79.74	135.53
Underidentification				0.000	0.000
SW weak test 1st inst.				29.99	122.69
SW weak test 2nd inst.				23.73	269.07
KP Wald-F weak test				12.84	61.22
Observations	936	936	936	936	936

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors clustered by province in OLS regressions, robust errors in IV. Expected gains from hs and ls migration instrumented in 4 and 5. Instruments in 4: provincial shares over total contracts with mobility signed by primary and by tertiary educated workers (see eq. 9). Instruments in 5: shift-share instruments (see eq. 10). Stock-Yogo (2005) weak ID test critical values for 10% maximal IV size: 7.03

Table 2: Benchmark regressions with addition of tuition fees

	enrolment rate in upper-secondary				
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV
Exp. gains from hs mig. (€)	0.045	0.573***	0.150*	1.263***	0.797***
	[0.44]	[4.40]	[1.98]	[2.67]	[5.30]
Exp. gains from ls mig. (€)	-0.269**	-0.359***	-0.150*	-1.548***	-0.399***
	[2.39]	[3.34]	[1.71]	[3.68]	[4.58]
Premium skill (€)	-0.010	0.595***	0.186**	0.724***	0.416***
	[0.13]	[6.89]	[2.58]	[3.84]	[6.02]
Pop. density (hab/km2)	-0.796***	-0.465***	-0.823***	-0.947***	-0.795***
	[3.15]	[4.65]	[4.38]	[3.56]	[4.79]
Unemployment rate (%)	0.525***	0.087	0.186***	0.273***	0.184***
	[9.34]	[0.69]	[2.94]	[4.42]	[4.69]
International hs mig. rate (%)	-0.434***	-0.459***	0.042	0.048	0.046
	[4.44]	[3.00]	[0.48]	[0.76]	[0.95]
International ls mig. rate (%)	0.747***	0.672***	-0.042	0.009	-0.093
	[5.71]	[2.98]	[0.32]	[0.08]	[1.14]
Uni tuition fee (€/credit)	0.031	-0.206**	-0.174**	-0.130***	-0.150***
	[0.45]	[2.65]	[2.60]	[3.95]	[6.27]
Region fe	Prov		Prov	Prov	Prov
Time fe		Year	Year	Year	Year
R^2	0.822	0.645	0.914	0.844	0.901
F	.	60.02	.	99.73	139.90
Underidentification				0.000	0.000
SW weak test 1st inst.				23.11	116.69
SW weak test 2nd inst.				25.88	256.73
KP Wald-F weak test				11.64	58.18
Observations	936	936	936	936	936

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors clustered by province in OLS regressions, robust errors in IV. Expected gains from hs and ls migration instrumented in 4 and 5. Instruments in 4: provincial shares over total contracts with mobility signed by primary and by tertiary educated workers (see eq. 9). Instruments in 5: shift-share instruments (see eq. 10). Stock-Yogo (2005) weak ID test critical values for 10% maximal IV size: 7.03

Table 3: Benchmark regressions by Autonomous Communities (CCAA)

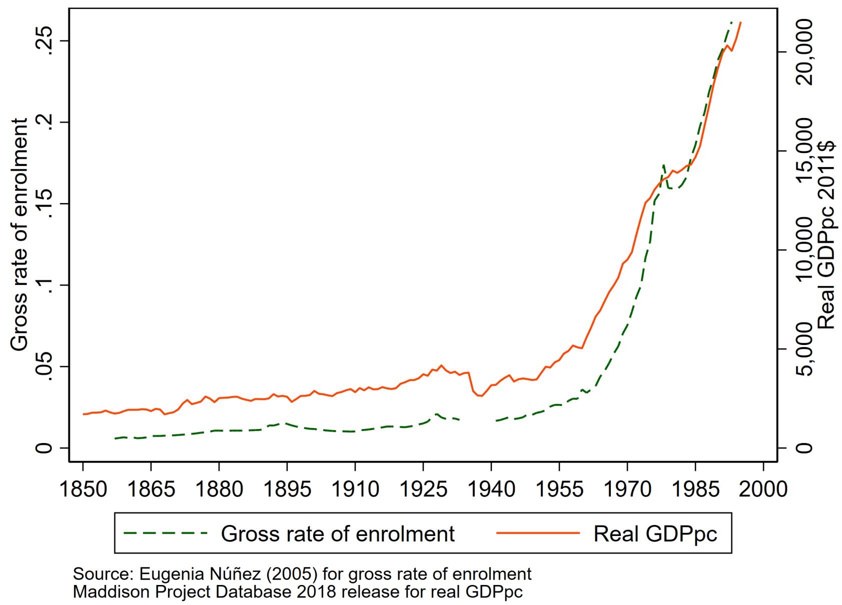
	enrolment rate upper-secondary			gross rate access university		
	(1) OLS	(2) IV	(3) IV	(4) OLS	(5) IV	(6) IV
Exp. gains from hs mig. (€)	0.255*	0.513*	0.584***	0.283	0.675	0.849***
	[1.99]	[1.85]	[3.34]	[1.44]	[1.32]	[3.61]
Exp. gains from ls mig. (€)	-0.521**	-0.932***	-0.630***	-0.565**	-3.057***	-1.062***
	[2.27]	[2.70]	[3.92]	[2.77]	[3.56]	[4.43]
Premium skill (€)	0.220*	0.301***	0.319***	0.195	0.335*	0.367***
	[1.97]	[2.81]	[3.72]	[1.15]	[1.75]	[3.57]
Pop. density (hab/km2)	0.905	0.850***	0.834***	0.013	-0.066	-0.109
	[1.60]	[3.55]	[3.43]	[0.03]	[0.14]	[0.36]
Unemployment rate (%)	0.182*	0.216***	0.188***	0.093	0.312**	0.132*
	[1.81]	[3.32]	[3.19]	[0.58]	[2.38]	[1.90]
International hs mig. rate (%)	0.203	0.273	0.227	-0.074	0.331	0.017
	[1.13]	[1.56]	[1.50]	[0.27]	[0.85]	[0.07]
International ls mig. rate (%)	-0.315	-0.373**	-0.347**	0.020	-0.259	-0.066
	[1.74]	[2.15]	[2.30]	[0.07]	[0.69]	[0.27]
Uni tuition fee (€/credit)	-0.092	-0.097***	-0.091***	-0.034	-0.077	-0.039
	[0.79]	[2.92]	[2.61]	[0.44]	[1.30]	[1.16]
Region fe	CCAA	CCAA	CCAA	CCAA	CCAA	CCAA
Time fe	Year	Year	Year	Year	Year	Year
R^2	0.950	0.947	0.947	0.928	0.831	0.922
F	.	148.03	147.37	.	55.54	104.38
Underidentification		0.008	0.000		0.003	0.000
SW weak test 1st inst.		35.62	70.86		37.31	70.86
SW weak test 2nd inst.		31.14	49.77		16.72	49.77
KP Wald-F weak test		12.46	33.78		4.65	33.78
Observations	306	306	306	306	306	306

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

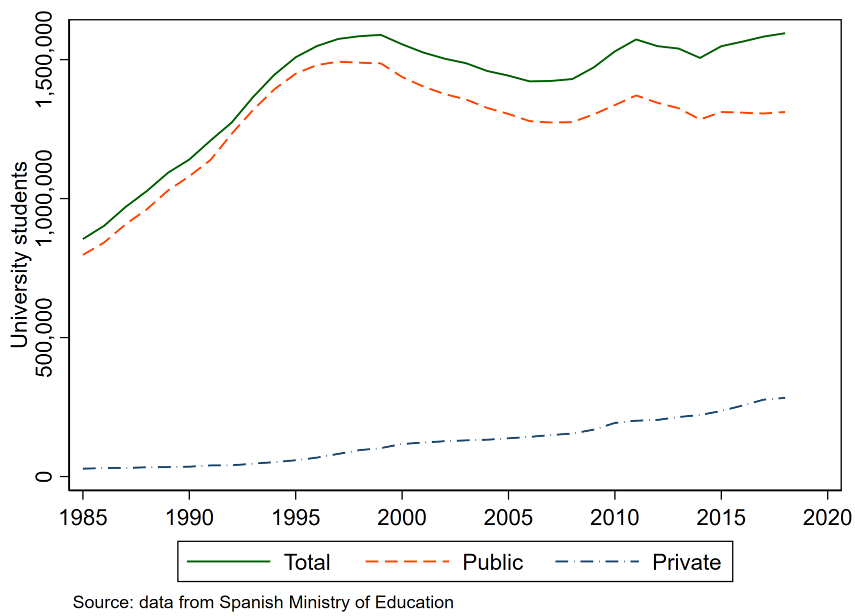
Note: Standardized beta coefficients displayed. Robust errors clustered by CCAA in OLS regressions, robust errors in IV. Expected gains from hs and ls migration instrumented in 2-3 and 5-6. Instruments in 2 and 5: provincial shares over total contracts with mobility signed by primary and by tertiary educated workers (see eq. 9). Instruments in 3 and 6: shift-share instruments (see eq. 10). Stock-Yogo (2005) weak ID test critical values for 10% maximal IV size: 7.03

A Appendix

A.1 Extended results

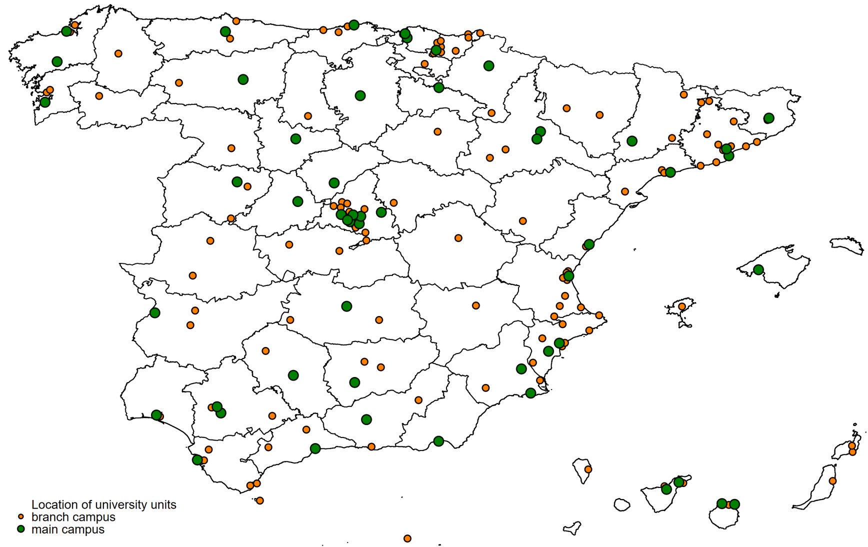


(a) University enrollment and economic development 1850-1993



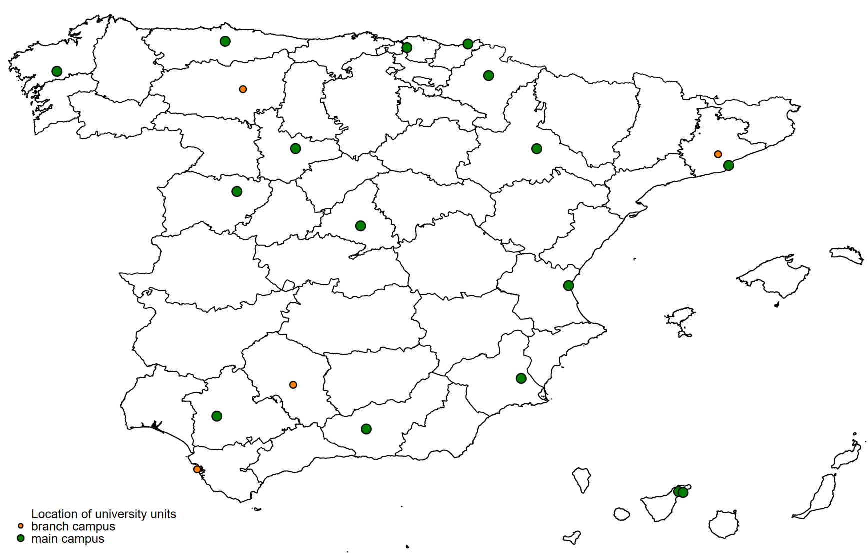
(b) University students 1985-2018 by type of university

Figure A.1: Evolution of university studies in Spain 1850-2018



Source: data from the Ministry of Education and GeoHack.
Each dot represents a municipality with at least one university unit (i.e., a faculty teaching classes).

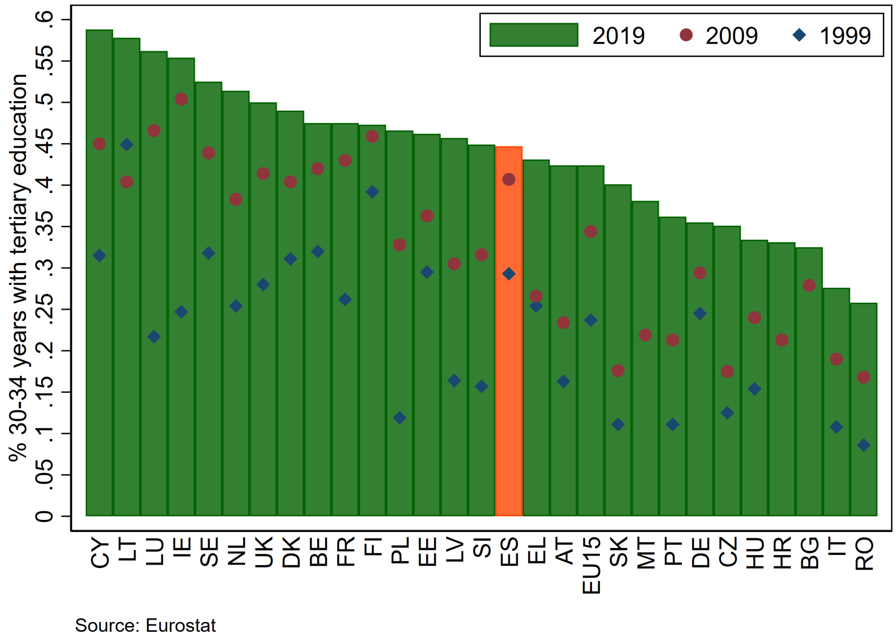
(a) Location of Spanish universities, 2019



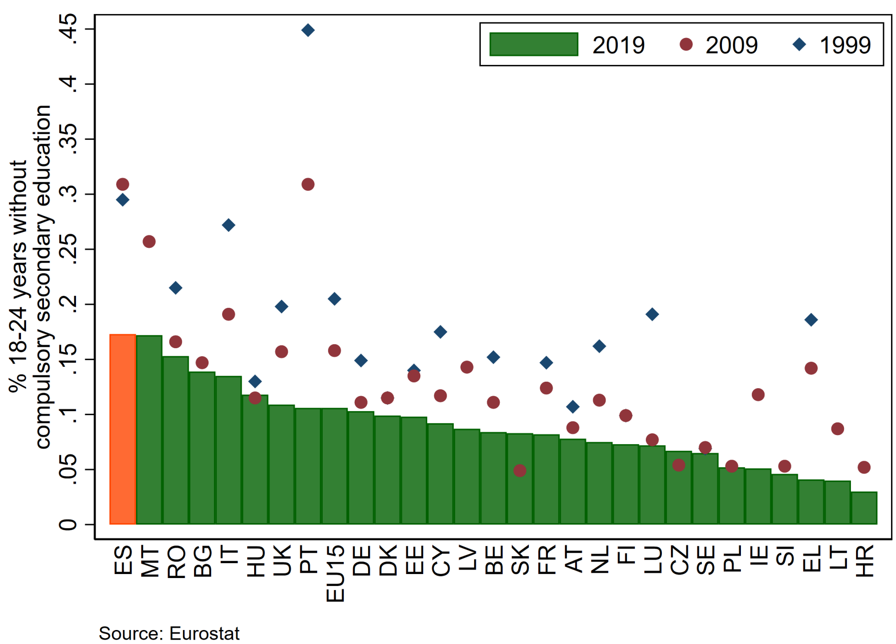
Source: data from the Ministry of Education and GeoHack.
Each dot represents a municipality with at least one university unit (i.e., a faculty teaching classes).

(b) Location of Spanish universities, 1963

Figure A.2: Expansion of the university network in Spain

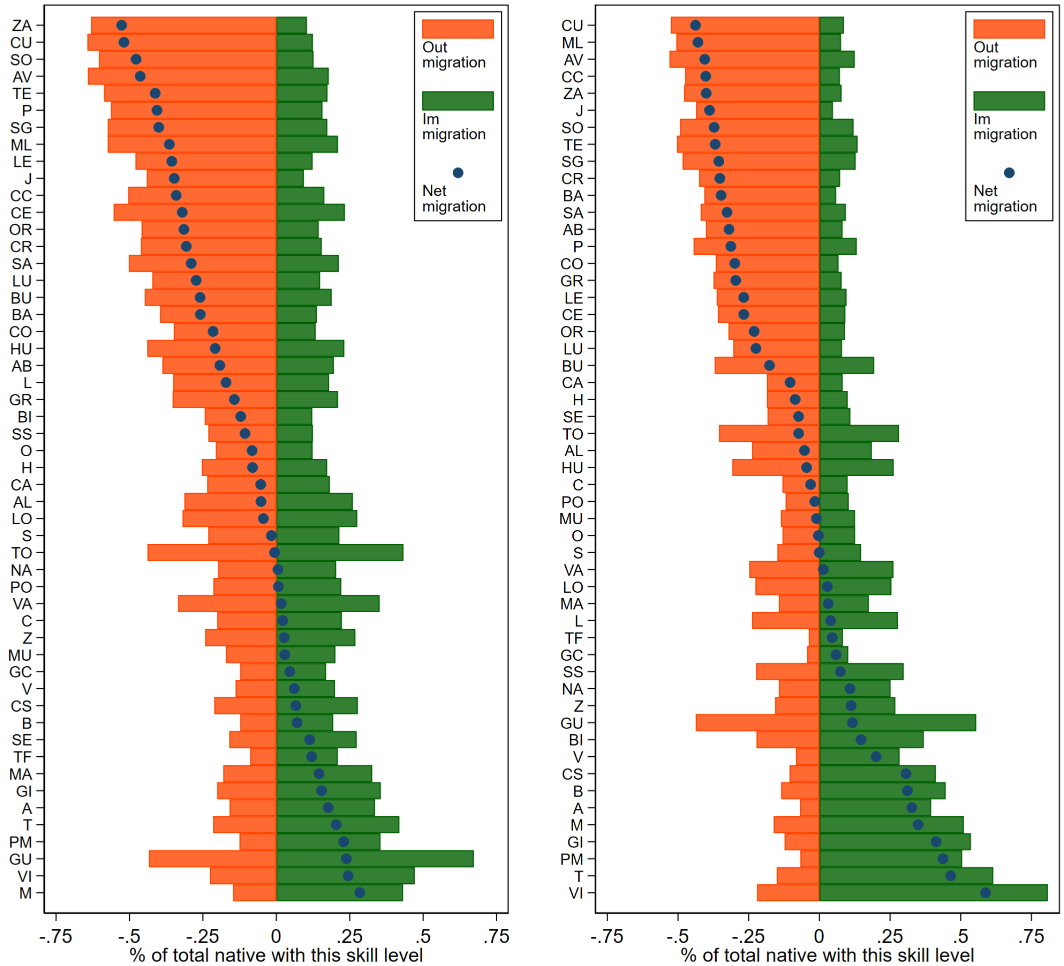


(a) Proportion of people with tertiary education 1999-2009-2019



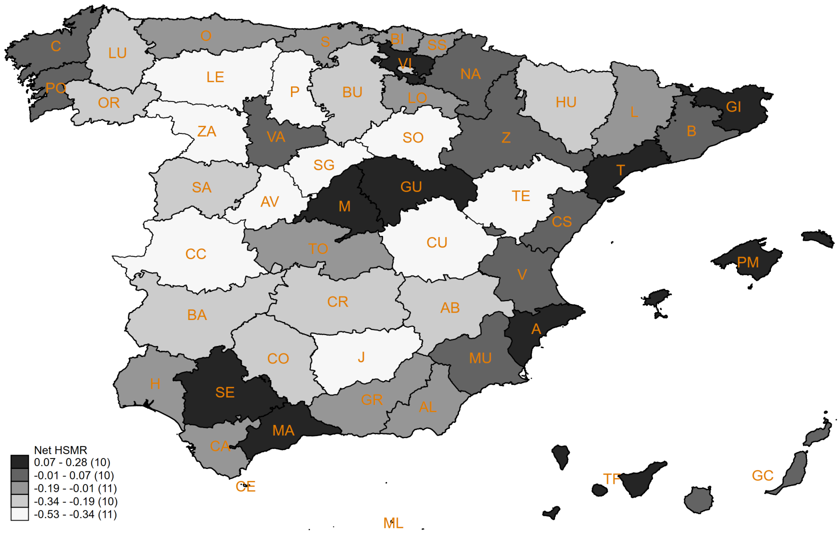
(b) Early school leave 1999-2009-2019

Figure A.3: The duality of Spanish education system

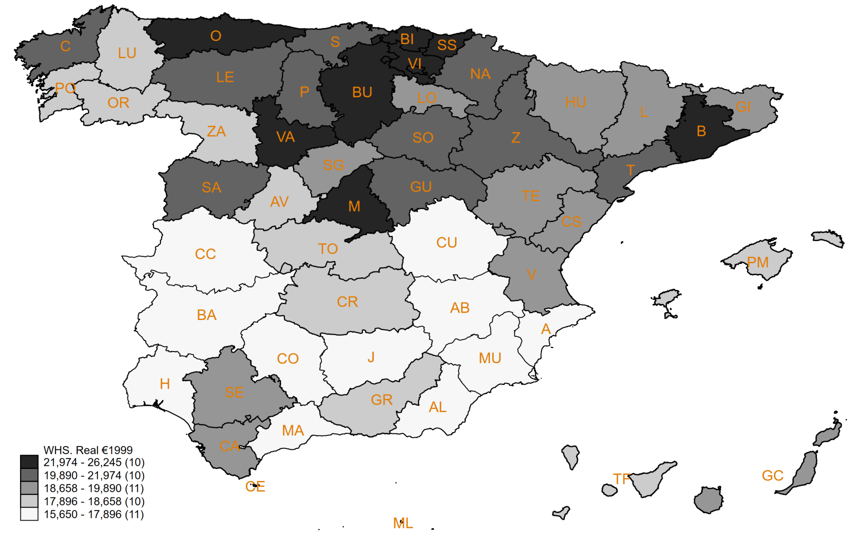


(a) High skilled migration (b) Low skilled migration

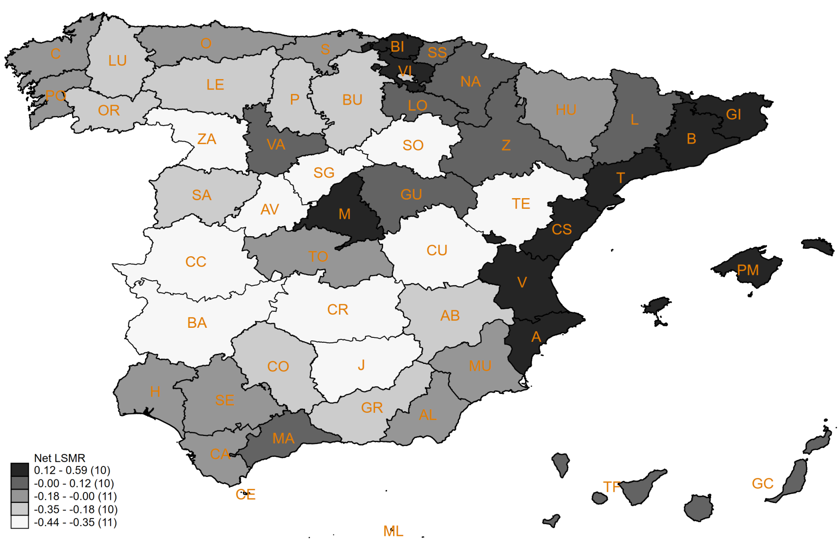
Figure A.4: Average emigration rates by province, 1999-2019



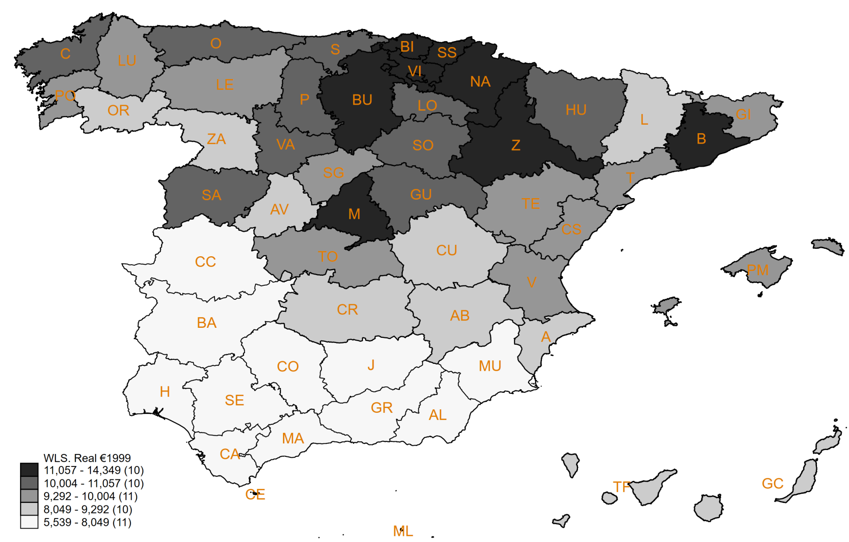
(a) Average net high skilled migration rate, 1999-2019



(b) Average high skilled wage, 1999-2019



(c) Average net low skilled migration rate, 1999-2019



(d) Average low skilled wage, 1999-2019

Figure A.5: Net migration and wage by province, 1999-2019

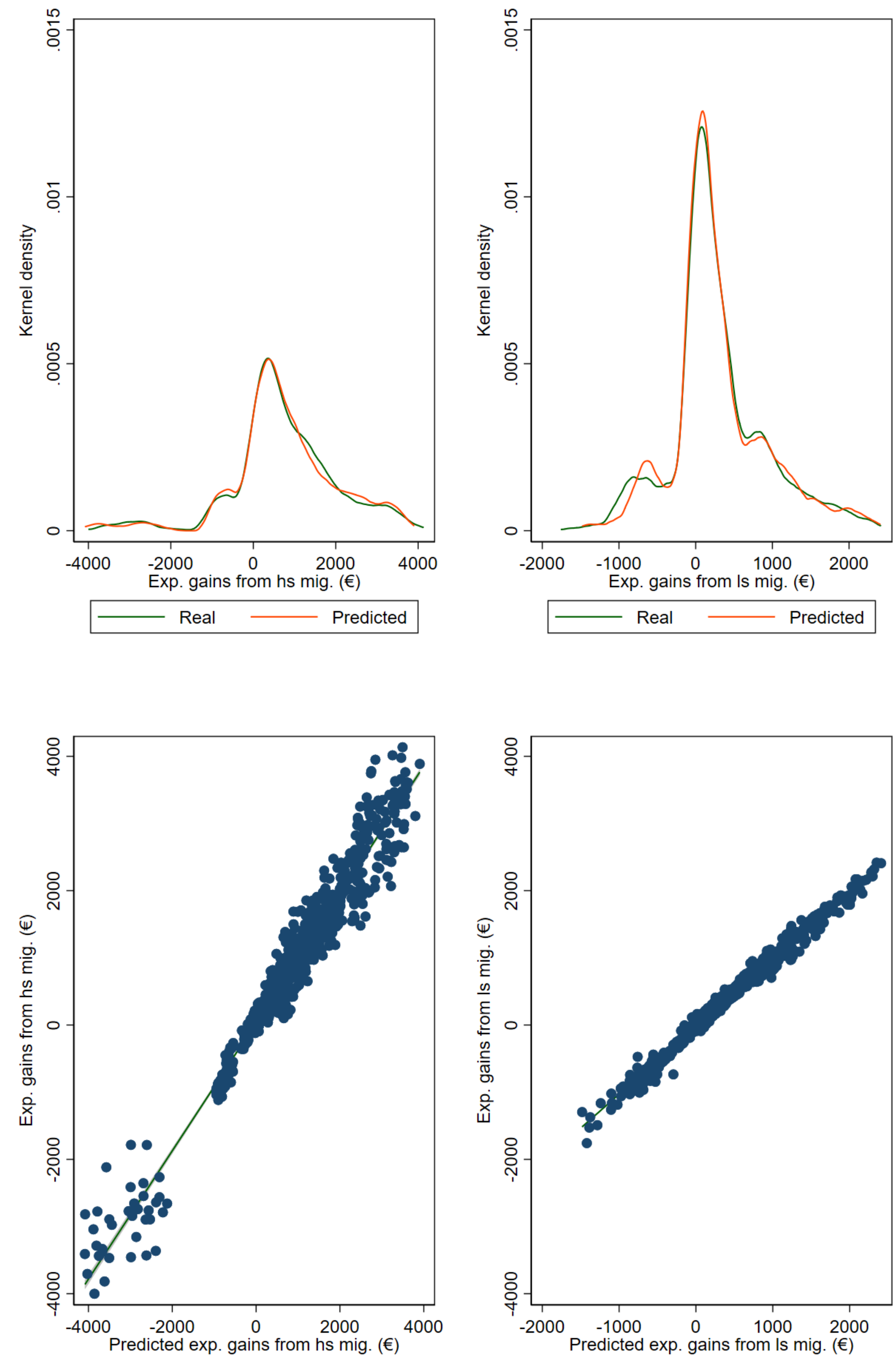


Figure A.6: Graphic overview of the precision of Bartik instruments

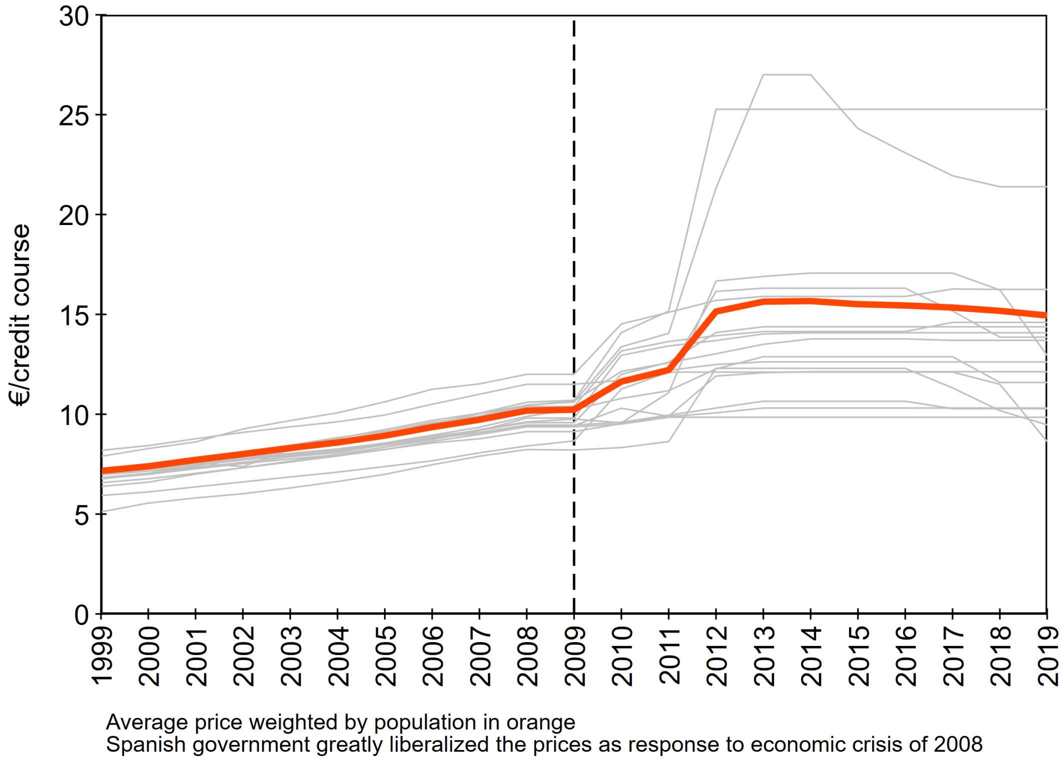
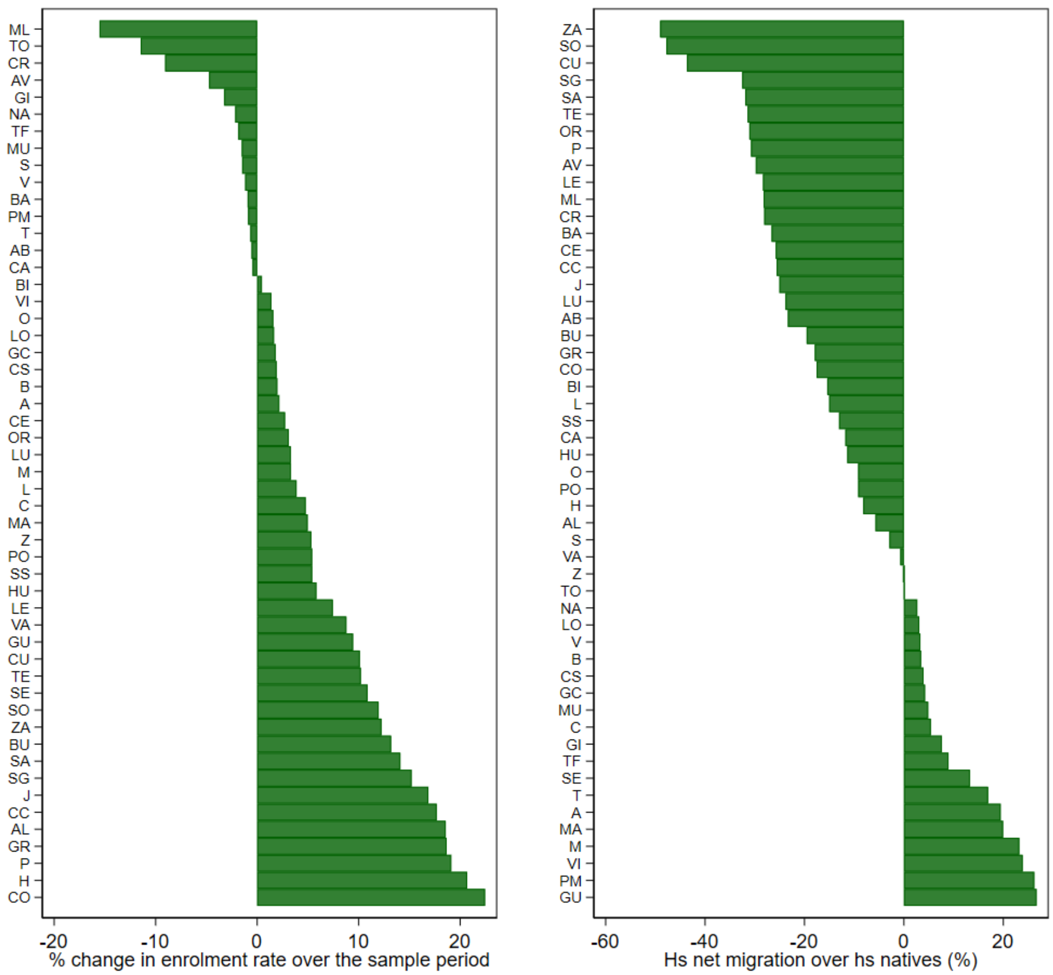


Figure A.7: Increase in tuition fees in Spanish universities by CCAA, 1999-2019



(a) Brain Refrain vs. Brain Gain (b) Net Human Capital Position 2018

Figure A.8: Estimated provincial effects

Table A.1: Summary stats of the variables in our study

	Mean	SD	Min	Max	n
by PROV					
Enrolment upper-secondary ed. (%)	25.09	3.21	17.27	33.75	936
Exp. gains from hs mig. (€)	759.75	1,271.22	-4,000.80	4,135.11	936
Exp. gains from ls mig. (€)	289.46	636.47	-1,758.40	2,417.74	936
Pred. exp. gains from hs mig. (€)	757.56	1,300.41	-4,084.88	3,905.71	936
Pred. exp. gains from ls mig. (€)	302.39	624.11	-1,481.08	2,409.77	936
Premium skill (€)	10,291.13	1,473.91	7,190.23	15,942.03	936
High skilled mig. rate (%)	34.06	16.57	7.11	72.97	936
Low skilled mig. rate (%)	27.51	15.13	2.37	64.95	936
High skilled mig. index (%)	35.07	17.90	8.29	75.94	936
Low skilled mig. index (%)	28.14	16.53	2.90	61.49	936
Unemployment rate (%)	14.84	7.84	2.95	42.66	936
International hs mig. rate (%)	2.85	2.63	0.12	23.44	936
International ls mig. rate (%)	3.38	2.17	0.85	16.26	936
Pop. density (hab/sq km.)	313.43	993.01	8.48	7,017.39	936
by CCAA					
attendance to uni entrance exam (%)	43.86	7.92	27.80	67.10	306
Enrolment upper-secondary ed. (%)	25.32	3.02	19.47	32.60	306
Exp. gains from hs mig. (€)	471.28	810.62	-1,038.78	3,493.47	306
Exp. gains from ls mig. (€)	148.67	517.03	-878.85	2,102.16	306
Pred. exp. gains from hs mig. (€)	555.05	877.44	-951.51	3,493.47	306
Pred. exp. gains from ls mig. (€)	171.66	504.41	-761.48	2,102.16	306
Premium skill (€)	10,370.93	1,321.57	7,873.58	14,714.04	306
High skilled mig. rate (%)	21.36	11.09	5.62	48.87	306
Low skilled mig. rate (%)	17.79	11.23	1.87	48.05	306
High skilled mig. index (%)	22.37	12.74	5.61	49.62	306
Low skilled mig. index (%)	18.41	12.78	1.53	48.40	306
Unemployment rate (%)	13.86	6.97	4.17	35.82	306
International hs mig. rate (%)	0.85	0.54	0.25	3.59	306
International ls mig. rate (%)	2.18	1.15	0.85	8.01	306
Pop. density (hab/sq km.)	133.37	145.50	18.08	674.81	306

Note: All monetary variables deflated with a price index with base year 1999.

A.2 Analysis with extended sample

Here we provide the benchmark results with an extended sample period from 1999 to 2019. All variables are available except the international migration rate (2001-2019) and the instruments constructed with labor contract data (2001-2018).

Table A.2: Benchmark regressions with extended sample period

	enrolment rate in upper-secondary				
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV
Exp. gains from hs mig. (1k€)	0.000	0.015***	0.004*	0.041***	0.012***
	[0.11]	[4.43]	[1.74]	[3.30]	[3.70]
Exp. gains from ls mig. (1k€)	-0.015***	-0.020***	-0.011**	-0.103***	-0.015***
	[3.37]	[3.46]	[2.57]	[4.56]	[3.87]
Premium skill (1k€)	-0.002	0.012***	0.006***	0.020***	0.008***
	[1.30]	[6.17]	[3.50]	[5.00]	[6.67]
Pop. density (100hab/km2)	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	[3.21]	[4.44]	[3.98]	[2.87]	[3.50]
Unemployment rate (%)	0.236***	0.073	0.069***	0.135***	0.068***
	[13.27]	[1.37]	[2.93]	[4.56]	[4.25]
International hs mig. rate (%)	-0.412***	-0.561***	-0.053	0.027	-0.044
	[4.81]	[3.56]	[0.66]	[0.30]	[0.83]
International ls mig. rate (%)	0.889***	1.030***	-0.145	-0.018	-0.196*
	[6.99]	[3.43]	[0.96]	[0.09]	[1.83]
Region fe	Prov		Prov	Prov	Prov
Time fe		Year	Year	Year	Year
R^2	0.788	0.599	0.893	0.785	0.889
F	.	80.813	.	79.744	150.806
Underidentification				0.000	0.000
SW weak test 1st inst.				29.99	201.57
SW weak test 2nd inst.				23.73	451.97
KP Wald-F weak test				12.843	99.690
Observations	1092	1092	1092	936	1092

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Robust errors clustered by province in OLS regressions, robust errors in IV. Expected gains from hs and ls migration instrumented in 4 and 5. Instruments in 4: provincial shares over total contracts with mobility signed by primary and by tertiary educated workers (see eq. 9). Instruments in 5: shift-share instruments (see eq. 10) with baseline 1999. Stock-Yogo (2005) weak ID test critical values for 10% maximal IV size: 7.03

A.3 Replication of Beine et al (2008)

As an additional exercise, we replicate the canonical results of the model of Beine et al. (2008) but rather in our context of internal mobility in Spain. To that end we estimate the following regression:

$$\Delta ln(H_{t_1/t_0}) = \beta_0 + \beta_1 ln(H_{t_0}) + \beta_2 ln(P_{it_0,s}) + \beta_3 ln(X_{it_0}) \tag{11}$$

The gross growth of ex-ante human capital is therefore regressed on the initial levels of the explanatory variables, using only the probability to emigrate for high skilled as our proxy to detect the brain gain effect.

Column 1 of Table A.3 presents our OLS results from conducting this exercise. Similarly to Beine et al. (2008) we find evidence of convergence from initial levels of human capital, a correlation in favor of a Brain Gain effect operating.[16]

Table A.3: Replication of Beine et al (2008)

	log increase in % hs nat. pop. 2001-2018		
	(1) OLS	(2) IV	(3) IV
log % hs in the nat. pop at 2001	-0.2843***	-0.2827***	-0.2708***
	[7.47]	[7.75]	[7.01]
log hs mig. rate at 2001	0.0481**	0.0509**	0.0713***
	[2.19]	[2.37]	[2.94]
Population density at 2001	-0.0001**	-0.0001***	-0.0001***
	[2.65]	[2.76]	[2.86]
Constant	0.1812**	0.1869***	0.2303***
	[2.65]	[2.85]	[3.10]
R^2	0.580	0.580	0.573
F	29.31	29.33	34.24
Underidentification		0.000	0.000
Overidentification		0.673	0.045
KP Wald-F weak test		491.58	67.74
Observations	52	52	52

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Robust errors corrected for heteroskedasticity. Hs mig. rate instrumented in 2 and 3, Instruments in 2: log of stock of emigrants and log of population, Instruments in 3: provincial share over the total contracts to tertiary educated workers which implied migration (see eq. 9) and log of population. Stock-Yogo (2005) weak ID test critical values for 10% maximal IV size: 19.93

We address endogeneity concerns in Columns 2 and 3 of Table A.3. Column 2 shows the results from implementing the same instruments as Beine et al. (2008), namely: the stocks of emigrants and the population for each province. Column 3 rather presents the

[16] Contrary to Beine et al. (2008) however, population density is significant in all our specifications. Note that we didn't include a dummy variable for Spanish provinces (as they did with Sub-Saharan Africa).

results when we instead implement our instruments for low and high skilled emigration rates as detailed in Equation 9 when we rather employ provincial shares of contracts signed at emigrants' destinations. The first stages of these IV regressions care presented in Table A.4. Our results are comparable in magnitude to those of Beine et al. (2008)) and we confirm the validity of our instruments as they did.

A.4 First Stage Regressions

Table A.4: First stage regressions in table A.3

	Column 2 in table 1	Column 3 in table 1
	(1)	(2)
	Exp. gains hs mig. 2001	Exp. gains hs mig. 2001
log Total emigrant stock	0.461***	
	[16.12]	
Prov. share tert. ed. contr.		23.176***
		[4.41]
log Total population	-0.613***	-0.807***
	[29.84]	[8.43]
log % hs in the nat. pop at 2001	0.045	-0.617***
	[0.62]	[4.22]
Population density at 2001	-0.000	-0.000***
	[0.36]	[3.95]
F	491.58	67.74
Observations	52	52

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors corrected for heteroskedasticity.

Table A.5: First stage regressions in table 1

	Eq. (5) in table 2		Eq. (6) in table 2	
	(1) Exp. gains hs mig.	(2) Exp. gains ls mig.	(3) Exp. gains hs mig.	(4) Exp. gains ls mig.
Prov. share prim. ed. contr.	0.020**	0.067***		
	[2.08]	[4.36]		
Prov. share tert. ed. contr.	-0.208***	-0.021		
	[5.15]	[0.56]		
Bartik exp. gains hs mig.			0.958***	0.059
			[11.66]	[1.38]
Bartik exp. gains ls mig.			0.019	0.965***
			[0.33]	[31.02]
Premium skill (€)	-0.282***	0.158***	-0.060**	-0.006
	[8.97]	[4.74]	[1.98]	[0.34]
Unemployment rate (%)	0.060**	0.094***	0.003	0.011
	[2.11]	[3.21]	[0.12]	[0.66]
International hs mig. rate (%)	0.008	0.017	-0.033	-0.010
	[0.23]	[0.67]	[1.07]	[0.71]
International ls mig. rate (%)	0.114**	0.124***	0.051	-0.024
	[2.48]	[2.76]	[1.14]	[0.94]
Pop. density (hab/km2)	-0.100	-0.184	-0.035	0.209***
	[0.86]	[1.46]	[0.27]	[2.63]
Region fe	Prov	Prov	Prov	Prov
Time fe	Year	Year	Year	Year
F	13.89	9.54	90.29	698.28
Observations	936	936	936	936

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors corrected for heteroskedasticity.

Table A.6: First stage regressions in table 2

	Eq. (5) in table 2		Eq. (6) in table 2	
	(1) Exp. gains hs mig.	(2) Exp. gains ls mig.	(3) Exp. gains hs mig.	(4) Exp. gains ls mig.
Prov. share prim. ed. contr.	0.026***	0.068***		
	[2.60]	[4.45]		
Prov. share tert. ed. contr.	-0.192***	-0.017		
	[4.51]	[0.44]		
Bartik exp. gains hs mig.			0.953***	0.066
			[11.38]	[1.54]
Bartik exp. gains ls mig.			0.021	0.963***
			[0.36]	[31.06]
Premium skill (€)	-0.289***	0.156***	-0.063**	-0.001
	[9.18]	[4.61]	[2.01]	[0.08]
Unemployment rate (%)	0.052*	0.092***	0.002	0.013
	[1.84]	[3.15]	[0.07]	[0.76]
International hs mig. rate (%)	0.015	0.019	-0.032	-0.012
	[0.45]	[0.73]	[1.01]	[0.86]
International ls mig. rate (%)	0.123***	0.126***	0.053	-0.027
	[2.63]	[2.81]	[1.17]	[1.03]
Pop. density (hab/km2)	-0.115	-0.188	-0.037	0.212***
	[0.99]	[1.49]	[0.29]	[2.67]
Uni tuition fee (€/credit)	-0.030**	-0.007	-0.005	0.008
	[2.45]	[0.66]	[0.54]	[1.50]
Region fe	Prov	Prov	Prov	Prov
Time fe	Year	Year	Year	Year
F	12.42	9.97	85.58	699.98
Observations	936	936	936	936

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors corrected for heteroskedasticity.

Table A.7: First stage regressions in table 3

	Eq. (5) in table 2		Eq. (6) in table 2	
	(1) Exp. gains hs mig.	(2) Exp. gains ls mig.	(3) Exp. gains hs mig.	(4) Exp. gains ls mig.
Prov. share prim. ed. contr.	0.081	0.093***		
	[1.65]	[3.67]		
Prov. share tert. ed. contr.	-0.320***	0.019		
	[5.24]	[0.33]		
Bartik exp. gains hs mig.			1.095***	0.184*
			[9.81]	[1.85]
Bartik exp. gains ls mig.			0.207***	0.771***
			[3.17]	[8.62]
Premium skill (€)	-0.254***	0.029	-0.024	0.007
	[6.97]	[0.82]	[0.62]	[0.19]
Unemployment rate (%)	0.071*	0.105***	0.013	0.089***
	[1.75]	[3.01]	[0.46]	[5.32]
International hs mig. rate (%)	0.048	0.185*	0.025	0.035
	[0.39]	[1.75]	[0.30]	[0.71]
International ls mig. rate (%)	0.034	-0.134	-0.087	-0.081*
	[0.27]	[1.24]	[1.00]	[1.70]
Pop. density (hab/km2)	0.132	-0.234	0.001	-0.168***
	[0.78]	[1.45]	[0.01]	[2.69]
Region fe	CCAA	CCAA	CCAA	CCAA
Time fe	Year	Year	Year	Year
F	17.38	7.64	60.70	222.79
Observations	306	306	306	306

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors corrected for heteroskedasticity.

Table A.8: First stage regressions in table A.2

	Eq. (5) in table 2		Eq. (6) in table 2	
	(1) Exp. gains hs mig.	(2) Exp. gains ls mig.	(3) Exp. gains hs mig.	(4) Exp. gains ls mig.
Prov. share prim. ed. contr.	0.020**	0.067***		
	[2.08]	[4.36]		
Prov. share tert. ed. contr.	-0.208***	-0.021		
	[5.15]	[0.56]		
pEGHSM				
Bartik exp. gains hs mig.			1.035***	-0.035
			[14.24]	[0.96]
pEGLSM				
Bartik exp. gains ls mig.			0.047	1.006***
			[0.99]	[38.34]
Premium skill (1k€)	-0.282***	0.158***	-0.043*	-0.053***
	[8.97]	[4.74]	[1.70]	[3.09]
Unemployment rate (%)	0.060**	0.094***	-0.010	0.017
	[2.11]	[3.21]	[0.51]	[1.11]
International hs mig. rate (%)	0.008	0.017	-0.035*	-0.001
	[0.23]	[0.67]	[1.68]	[0.08]
International ls mig. rate (%)	0.114**	0.124***	0.043	-0.048**
	[2.48]	[2.76]	[1.40]	[2.18]
Pop. density (100hab/km2)	-0.100	-0.184	0.150	0.151**
	[0.86]	[1.46]	[1.45]	[2.22]
Region fe	Prov	Prov	Prov	Prov
Time fe	Year	Year	Year	Year
F	13.89	9.54	161.90	916.57
Weak				
Observations	936	936	1092	1092

* p < 0.1, ** p < 0.05, *** p < 0.01. Absolute t-statistics in brackets.

Note: Standardized beta coefficients displayed. Robust errors corrected for heteroskedasticity.