Received: 2 October 2019 | Revised: 25 February 2020 | Accepted: 23 April 2020

DOI: 10.1111/pirs.12540

FULL ARTICLE

The regional anatomy of youths' educational attainment in Spain: The role of the employment structure in local labour markets

Luis Diaz-Serrano[1,2] | **William Nilsson**[3]

[1]ECO-SOS, Department of Economics, Universitat Rovira i Virgili, Av. de la Universitat 1, Reus, 43204, Spain

[2]ECEMIN, Universidad Antonio de Nebrija, Calle de Sta. Cruz de Marcenado, 27, Madrid, 28015, Spain

[3]Department of Applied Economics, Universitat de les Illes Balears, Cra. De Valldemosa, Km 7.5, Palma de Mallorca, 07122, Spain

Correspondence
Luis Diaz-Serrano, ECO-SOS, Department of Economics, Universitat Rovira i Virgili, Av. de la Universitat 1, Reus 43204, Spain.
Email: luis.diaz@urv.cat

Funding information
Spanish Ministry of Science, Innovation and Universities, Grant/Award Number: RTI2018-094733-B-100; Obra Social "La Caixa", Grant/Award Number: 2014ACUP0130

Abstract

This paper studies the link between the employment structure of local labour markets and the schooling choices of the youth in Spain. We construct a panel of Spanish provinces, and the effect of local labour markets was identified by using the variation in the share of employment by industry and gender across provinces and over time. A model with province fixed-effects and specific-slopes is used, which makes it possible to control for both time constant and time-varying unobserved heterogeneity across provinces. A sizable impact is found for both boys and girls of the industry structure of employment on educational attainment.

KEYWORDS

educational attainment, labour market, skills, Spain, youth employment

JEL CLASSIFICATION

J21; J24

1 | INTRODUCTION

The human capital theory predicts that investment in education is countercyclical; that is, it increases (decreases) with economic downturns (upturns). The incentive to acquire education is likely to be countercyclical for a variety of reasons. First, the expected real wage is procyclical (Solon, Barsky, & Parker, 1994), and thus the income forgone due to the pursuit of educational endeavours is lower during recessions. Most of the existing international empirical literature linking educational attainment with the business cycle tends to confirm this prediction. Rees and Mocan (1997), Betts and McFarland (1995), Light (1996), Card and Lemieux (2001), Dellas and Koubi (2003), and

Johnson (2013) find this countercyclical relationship in the United States, while Rice (1999) and Clark (2011) find it in the United Kingdom, Reiling and Strøm (2015) find it in Norway, Di Pietro (2006) finds it in Italy, and Peraita and Pastor (2000) and Petrongolo and San Segundo (2002) find it in Spain. Two exceptions in the United States to this general result are Johnson (2013) and Warren and Lee (2003). These authors do not find a link between the unemployment rate and enrollment rates in high school or college.

All the studies cited above use the unemployment rate as an indicator of the local labour market conditions. While this measure is a good proxy for economic activity and captures movements in the business cycle, it does not reflect the peculiarities of local labour markets, namely the industry structure. This makes the unemployment rate in itself unable to predict educational attainment accurately in the event of an economic boom or recession. Individuals without post-compulsory formal education who enter the labour market are generally employed in low-skill jobs. This circumstance suggests that, in the event of an economic boom, the demand for low- or high-skilled labour may differ geographically depending on the type of industry driving this boom. In this context, the role played by the industry structure of the employment of local labour markets is crucial. This is an important issue because local labour markets are key in forming individuals' employment prospects in the short and the long run. In this context, in local labour markets dominated by the demand for low-skilled labour, either structural or cyclical, less motivated students may not have incentives to pursue higher levels of formal education when employment opportunities are favourable. For instance, there exist cross-country (Sachs and Warner, 2001; Auty, 2001; Van der Ploeg, 2011) and more recent within-country evidence (Papyrakis and Gerlagh, 2007; James and Aadland, 2011) that indicate lower economic growth in economies based on natural resources by leading to a less skilled labour force.

Two paradigmatic examples of how an economic upturn may have a different impact on youth educational attainment depending on the industry driving the boom are Spain and Ireland. Between 2000 and 2007, Spain experienced the most important economic boom in its recent history. This boom was driven by the construction industry, which is characterized by employing mainly low-skilled men. During that period, this industry employed more than 20% of the male workforce. This boom was responsible for the dramatic decrease in post-compulsory formal education enrolment among male teenagers during that period (Aparicio-Fenoll, 2016). On the contrary, the economic upturn experienced by the Irish economy during the 1990s was driven by the technological sector, which is characterized by employing high-skilled workers. During that period, this industry employed 20% of the Irish labour force. Contrary to what happened in Spain, this boom encouraged the demand for higher education in technological fields (Wickham & Boucher,).

Despite the relationship between the industry structure of local labour markets and educational attainment is an important issue, the literature analysing this link is surprisingly scarce. All these studies find that booms in industries characterized by employing low-skilled labourers tend to decrease the educational attainment of the population. Aparicio-Fenoll (2016) found that, in Spain, the decline in the returns to education driven by the boom in the construction industry decreased the enrolment rate of young men in high school. Black, McKinnish, and Sanders (2005) found that high-school enrolment rates in Kentucky and Pennsylvania declined considerably in the 1970s and increased in the 1980s in coal producing counties relative to counties without coal. Rickman, Wang, and Winters (2017) find significant reductions in high school and college attainment of residents in Montana, North Dakota, and West Virginia due to shale gas and oil extraction booms. Using data on non-metropolitan counties in Arkansas, Louisiana, Oklahoma, and Texas, Weber (2014) obtains results consistent with the studies cited above. He observes that the population in areas associated with greater activity in natural gas extraction is low-skilled in terms of formal education.

Another relevant issue regards the territorial dimension of the phenomena. Disparities in educational achievement are sizeable and persistent both between and within countries. According to Eurostat, in 2018 in Spain, the share of the population without post-compulsory formal education was among the highest (41%) in the EU, similar to Italy, Malta, and Portugal. This figure contrasts sharply with the lowest rates recorded in other EU economies, namely in Lithuania and Czech Republic (12%), and with the average of the EU (25%). Remarkably, the magnitude of these cross-country disparities is, at best, of the same magnitude as within countries. Thus, in Spain in 2018, this figure was nearly 50% or higher in Andalusia, Castilla-La Mancha, and Extremadura, whereas at a distance of a few hundred kilometres and within a similar institutional framework, these figures were 28% and 31% in the Basque Country and Navarra, respectively. Indeed, even within regions (NUTS 2), at a province-level (NUTS 3), a large degree of variation can be observed. According to own computations based on the Labour Force Survey, in Andalusia, the province with the highest share of individuals without post-compulsory formal education was Huelva (56%), while the province with the lowest share was Córdoba (37%), almost 20 percentage points smaller. This pattern can be observed in all the Spanish regions.

This paper analyses the impact of the industry structure of local labour markets on the educational attainment of youth. The industry structure with the distribution of youth employment (16–24 years old) is measured in each industry. To do so, data are constructed for a panel of Spanish provinces (NUTS3) covering the period 2002–2018, and the effect of local labour markets is identified by using the variation in the share of employment by industry and gender across provinces and over time. In contrast with the previous literature, a model with province fixed-effects and province specific-slopes is used, which makes it possible to control not only for time constant but also for time-varying unobserved heterogeneity across provinces. The consideration of region specific-trends makes it possible to control not only for the potential existence of common trends between educational attainment and the explanatory variables, which may lead to spurious correlations but also for time-varying unobserved heterogeneity that may violate the strict exogeneity assumption in conventional fixed-effect models (Heckman & Hotz, 1989).

This study aims to evaluate if short-run shocks in specific industries differing in the skill composition of the employed workforce can affect educational attainment at a young age. The particular interest in the short run is because detection of such effect is potentially more alarming. Low-educated individuals could be a convenient pool of cheap labour for businesses in a low-skilled sector, but the individual and social costs could be long-lasting. It might be argued that individuals who decide against pursuing further education during young age may return to school at more adult ages. However, the present authors agree with Rickman et al. (2017), in that educational investments not undertaken at a young age may be difficult to make up later.[1] This situation may lead to permanently lower education levels that reduce the individual's long-run earning potential.

This paper contributes to the existing literature in three ways. First, it moves one step away from the most conventional studies linking educational attainment with the unemployment rate. Second, the scarce literature linking educational attainment with specific industries only focus on energy extraction and mining. This paper offers a wider view by considering the whole composition of the industry structure of local labour markets. Third, we use an empirical model that considers regions specific-slopes, in addition to the commonly used region fixed-effects. As discussed below, the inclusion of the individual slopes has no dramatic effects on the estimated effects of our explanatory variables of interest, namely industry employment shares. However, it does have on other variables included in the model. For example, individual slopes remove the effect of the business cycle (unemployment rate). The fact that the estimated effects of the employment shares survive to the inclusion of the individual slopes indicates that the estimated effects for these variables are very robust. We think in this way the empirical strategy is strengthened. The paper's econometric estimates indicate that the composition of the industry in local labour markets is crucial in explaining the educational attainment of youth during the sample period, for both boys and girls. This result persists even after controlling for business cycle variables and all types of unobserved heterogeneity.

2 | REGIONAL DISPARITIES IN SPAIN

2.1 | Educational attainment

A characteristic feature of the Spanish economy is the high degree of regional heterogeneity in many economic outcomes, such as unemployment (López-Bazo & Motellón, 2013), innovation (López-Bazo & Motellón, 2018), or wage

[1]According to own computations based on the Spanish Labour Force Survey (2002–2018), in Spain, 86% of the individuals complete post-compulsory secondary education before the age of 20, 8% do so during the ages 21 to 24, and only 6% obtain their certificate after the age of 24.

distributions (Motellón, López-Bazo, & El-Attar, 2011). The educational system is not an exception, and school outcomes in Spain are also very heterogeneous across regions. For instance, huge differences in PISA score tests across Spanish regions (NUTS 2) are systematically reported in each wave since testing started in 2000. The southern regions (Andalusia, Extremadura, and Murcia) and the islands (the Balearic and Canary Islands) tend to score similarly as low as some developing countries, for example, Colombia or Ecuador, at the bottom end of the world ranking. On the contrary, other regions, such as Castilla y León, Madrid, and Basque Country, score as high as other European countries leading the world PISA ranking, such as Finland or the Netherlands. Similar patterns might be found regarding educational attainment.[2]

This paper focuses on the educational attainment of the population aged 16–24. This age bracket was chosen because, at this age, most of the individuals decide about their educational choices. In Spain, 16 years old is the age at which compulsory education is completed, and it is also the minimum legal age to work. Post-compulsory non-tertiary education is supposed to be completed at the age of 18 for high school and at the ages of 18–20 for lower and upper vocational studies, respectively. Higher education is supposed to be completed at the ages of 22–24. This paper focuses on those individuals that declare not to be studying and are part of the working population, either employed or unemployed.

Figure 1 shows the evolution of the percentage of the young population (16–24 years old) without post-compulsory formal education during the sample period, split by gender and region (NUTS 2).[3] This figure reveals the high level of regional disparities regarding the level of formal education achieved by the young population. In 2002, the share of the population aged 16–24 without post-compulsory formal education in the southern Spanish regions (Andalusia, Extremadura, Castilla-La Mancha, and Murcia), the islands (the Balearic Islands and Canary Islands), and Valencia was above 65%. This figure contrasts sharply with the educational attainment in some of the northern regions (Navarra, Basque Country, and Cantabria) and Madrid, where the share of youth without post-compulsory formal education was 20 to 30 percentage points lower. Even though the educational attainment of the young population has increased significantly in most of the Spanish regions during the last two decades, and that the gaps across regions have been shortened, regional disparities persist. Thus, in 2018, the share of the young population without post-compulsory formal education in the southern regions, the islands, and Valencia was nearly or above 50%, while in three other regions this figure was below 35%: It was lowest in Basque Country, with only 24%, followed by Cantabria and Navarra, with 32% and 34%, respectively.

The Spanish education system is also characterized by having a wide gender gap in the educational attainment of the population, which is one of the highest in the EU. As with other educational indicators, this gender gap is also quite heterogeneous across regions. During the sample period (2002–2018), the average gap between young men and young women ranges from 20 percentage points in Extremadura to 8 percentage points in Basque Country.

2.2 | Educational attainment and the industry structure of employment

As Figure 1 shows, the persistence of the regional disparities in the educational attainment of the young population across time suggests that such differences are likely to be structural. Figure 1 shows that educational attainment evolves throughout time, with practically identical trends in all regions (NUTS 2) but with almost constant gaps. The Spanish educational system is highly decentralized at both the fiscal (funding and expenditure) and the political (decision making) levels; that is, regions (NUTS 2) practically have full competence to rule their education system. In a

[2]One potential reason for these differences is that the educational system in Spain is highly decentralized. Diaz-Serrano and Meix-Llop (2018) offer an extensive overview of the impact of political and fiscal decentralization on educational achievement.

[3]Despite this econometric analysis being based on data aggregated at the NUTS 3 level, to allow for a clearer interpretation of the figures in this section, NUTS 2 data were preferred. The NUTS 2 classification provides a total of 17 regions, while the NUTS 3 classification provides a total of 50 provinces. Considering either NUTS 2 or NUTS 3 does not change the conclusions raised in this section and still shows the high level of intra-country heterogeneity regarding education and labour market outcomes.

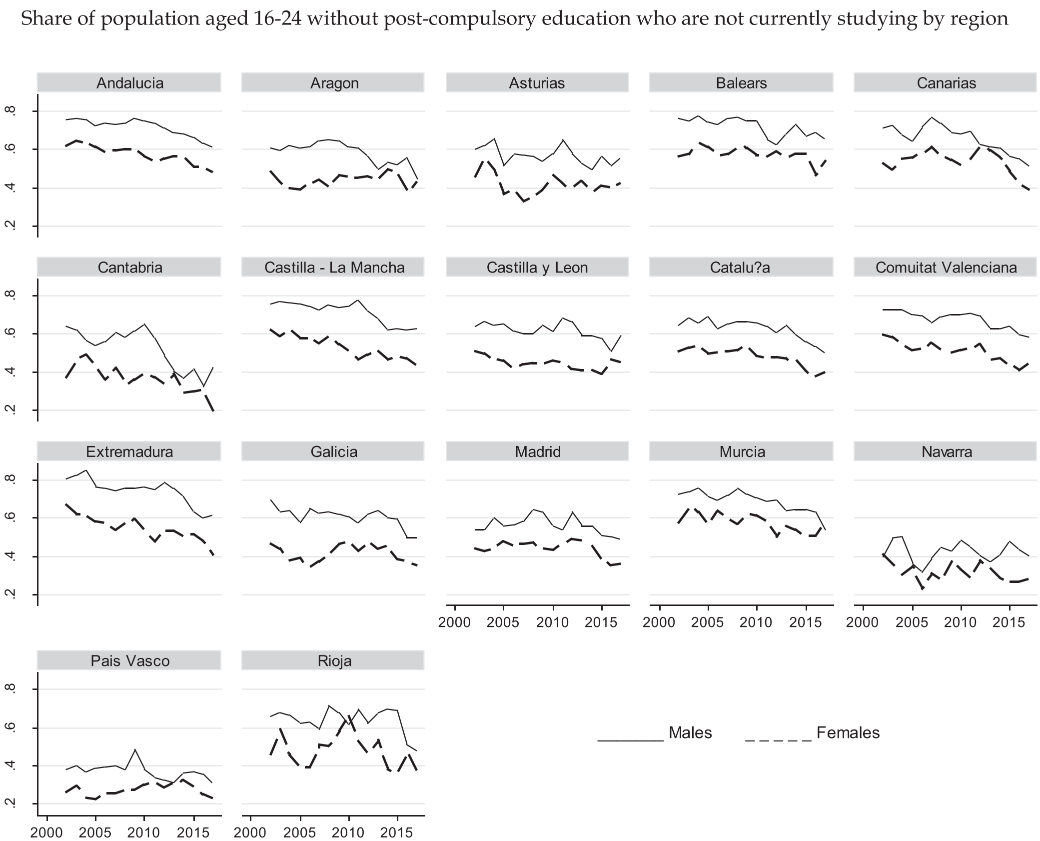

FIGURE 1 Share of population aged 16–24 without post-compulsory education who are not currently studying by region
Source: Own calculations based on the Spanish Labour Force Survey (2002–2018)

decentralized system, if some regional educational authorities perform better than others, they will generate inequalities across regions in school outcomes or increase existing ones. However, data coming from the Spanish Ministry of Education reveal that in 1990, before the educational system was fully decentralized, such differences across regions already existed (Diaz-Serrano & Nilsson, 2018). As was hypothesized in the introduction, and as other studies in the United States have shown, among other potential idiosyncratic factors, disparities in the industry structure of employment in local labour markets are a plausible candidate to explain the huge territorial disparities in educational attainment in Spain.

Figures 2 and 3 depict the relationship between educational attainment (compulsory education or lower) of the young population and employment split by industry. Both figures are gender-specific. At first glance, one can observe a high level of disparities across regions in the industry composition of employment. For boys, in regions such as the Basque Country and Madrid, agriculture represents less than 3% of employment, while in the southern regions (Andalusia, Murcia, Castilla-La Mancha, and Extremadura), this percentage is around 15%. The Balearic and Canary Islands are more intensive in the provision of low-skill services, mostly related to the touristic industry. In these regions, the share of young males employed in low-skill service industries is above 45%, while for young females, these figures are above 65%. Analogously, substantial differences are also observed across regions in the remaining industries.

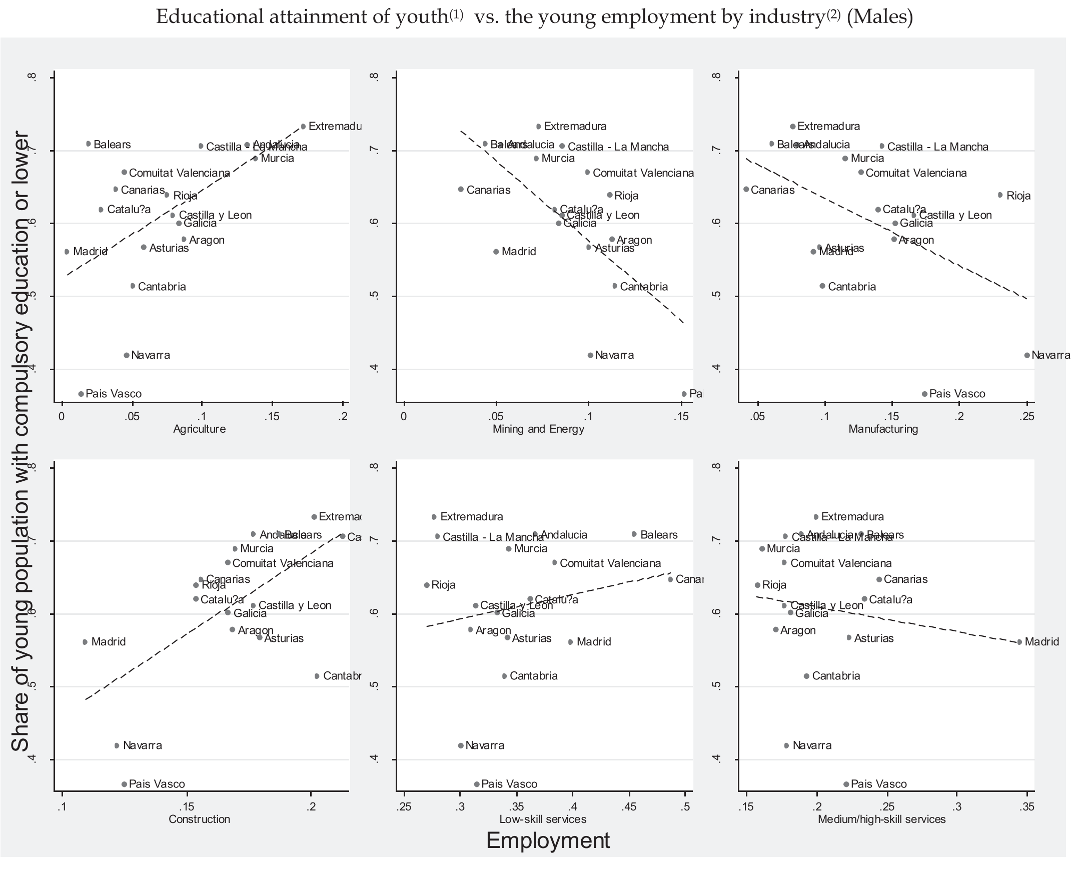

FIGURE 2 Educational attainment of youth (1) vs. the young employment by industry (2) (males)
Source: Own calculations based on the Spanish Labour Force Survey (2002–2018) (1) Share of population aged 16–24 without post-compulsory education who are not currently studying (Y–axis) (2) Share of employment of population aged 16–24 (X–axis)

Both figures also reveal quite sharp associations between the educational attainment of the young population and employment shares in different industries. For boys (Figure 2), there is a strong positive correlation between regional shares of compulsory education and employment in agriculture, construction, and the low-skill services industry. On the contrary, this link is negative in the mining/energy, manufacturing, and high- to medium-skill services industries. Similar associations are also observed between educational attainment and employment by industry for girls (Figure 3).[4]

3 | THE DATA

To carry out the empirical analysis, a panel of data was constructed, aggregated at the province level (NUTS 3) and covering the period 2002–2018. The data contain various policy and economic variables collected from different sources. The *share of the population aged 16–24 without post-compulsory education who are not currently studying* (EDUCOMP) is the outcome variable, and it is based on own calculations from the Spanish Labour Force Survey.

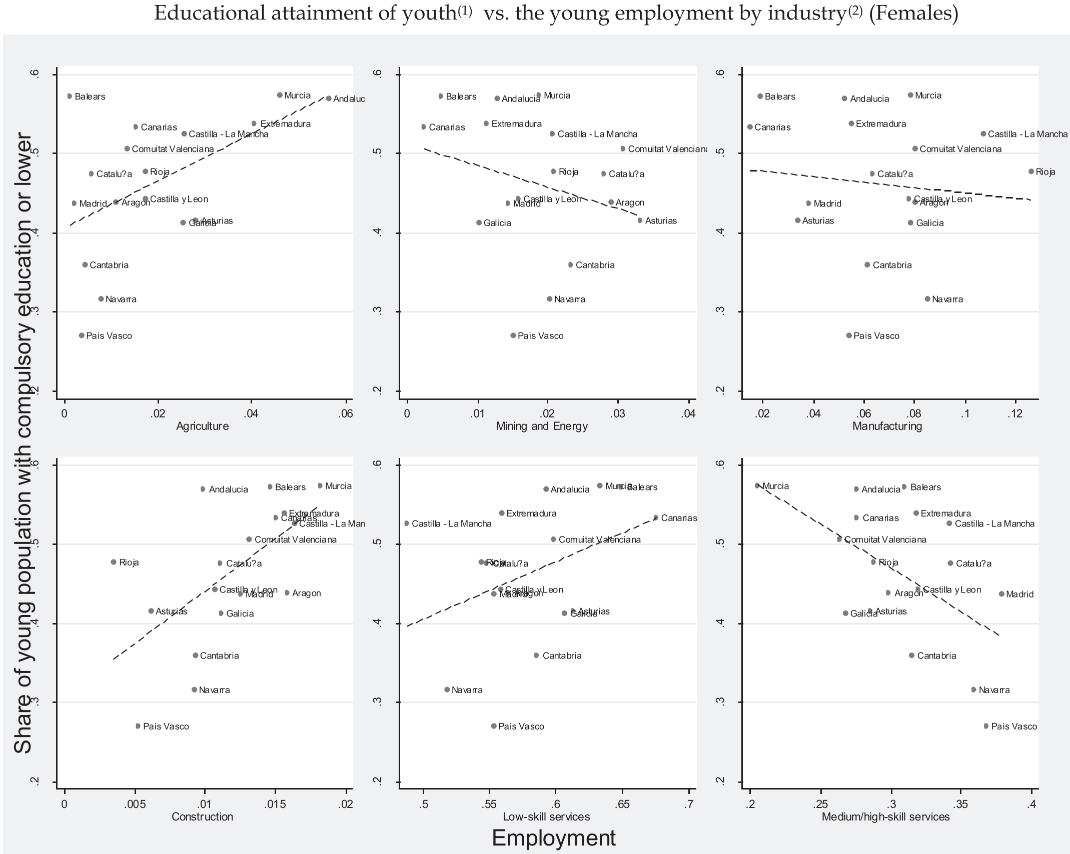

FIGURE 3 Educational attainment of youth (1) vs. the young employment by industry (2) (memales)
Source: Own calculations based on the Spanish Labour Force Survey (2002–2018) (1) Share of population aged 16–24 without post-compulsory education who are not currently studying (Y–axis) (2) Share of employment of population aged 16–24 (X–axis)

Analogously, the key explanatory variables – namely, the shares of youth employment (16–24 years old) in each industry of overall youth employment in each province – are also calculated from the Spanish Labour Force Survey. Since separate estimates are carried out for boys and girls, female and male educational attainment and employment are considered separately. To account for fluctuations in the business cycle and the level of wealth of the regions, the analysis controls for the unemployment rate (UNEMP), also split by gender, and the GDP *per capita* in each province. The unemployment rate is based on own calculations from the Spanish Labour Force Survey, while the provincial GDP *per capita* is taken from the Spanish National Accounts, both provided by the Spanish Statistics Bureau (INE).

As mentioned in the previous section, the Spanish education system is very decentralized. Regional governments have a high degree of autonomy regarding educational budgets and political decisions affecting their education system. Since decentralization of the Spanish education system gives autonomy to regional governments (NUTS 2) and

[4]These associations are obtained with the raw data; however, as will be seen in the empirical section, some of these associations reverse after controlling for province fixed-effects and specific slopes, year dummies, and the set of controls.

not to provinces (NUTS 3), these policy variables vary only by region (NUTS 2). Furthermore, this information at the province level is not available. To capture the impact of regional policies on the educational attainment of youth, the analysis also includes regional public spending on education as a percentage of the regional GDP (PUBEXP), the public regional expenditure on private education as a percentage of the total regional public expenditure in education (PRIVEXP), and the share of non-Spanish students attending compulsory education of the overall population of young students in compulsory education (IMSTUD). These three variables are taken from the Spanish Ministry of Education. The variable IMSTUD is included to control for the fact that the immigrant population achieves lower levels of education than its native counterpart. Finally, to capture other dimensions that are not found among the economic and policy variables, the analysis also considers the population density of the province (POPDEN), the share of the incoming young population (INCPOP), and the share of the outgoing young population (OUTPOP) in each province. Incoming population is considered to be those individuals who live in province k in year t but lived in a different province in year $t-1$. Analogously, outgoing population is defined as those individuals who lived in province k in year $t-1$ and live in a different province in year t. Inter-province mobility is considered to account for the fact that young persons may move away from lagging regions towards booming regions, which probably will also have good higher education institutions. This type of mobility will also have an impact on the aggregated level of educational attainment in each province. The expected sign of these two variables will depend on whether young migrants are more or less educated. The variables INCPOP and OUTPOP are based on own calculations from the Spanish Labour Force Survey.

Table 1 shows a description of the variables used in the analysis. The summary statistics are the averaged values for the period 2002–2018 of the aggregated values at the province level (NUTS 3). A first sight, Table 1 reveals a very high difference between the minimum and the maximum values in all variables, which shows the high degree of heterogeneity across Spanish provinces regarding employment and economic and educational indicators. This heterogeneity regarding the employment shares by industry indicates a high degree of specialization in some provinces in specific industries. The concentration of girls employed in services is remarkable: Almost 57% of girls are employed in low-skill services and 30% in medium- to high-skill services. Only 10% of girls are employed in non-service industries. On the contrary, the distribution of employment by industry for boys is less unequal. Compared to girls, the share of boys employed in services is notably smaller, 53%. After services, the construction industry is the one employing the most boys, 17%. Inter-provincial mobility of the young population is very low. On average, the shares of incoming and outgoing young population are 0.7% and 0.6% for boys and 0.8% and 1% for girls, respectively. Finally, it is also worth noting that, compared to boys, girls are much more educated. Almost 63% of boys aged 16–24 do not have post-compulsory education, while this figure is 48% for girls.

4 | EMPIRICAL MODEL

4.1 | A model with province fixed-effects and province specific-slopes

To carry out the empirical analysis, a linear model with province fixed-effects, year dummies, and province specific-trends is used. This model has the interesting feature of controlling not only for time-constant unobserved heterogeneity but also for time-varying unobserved heterogeneity across provinces. The inclusion of province specific-trends can be necessary for two reasons. On the one hand, conventional fixed-effect models may fail because the strict exogeneity assumption is violated due to the existence of time-varying unobserved heterogeneity, which is not captured through the fixed-effects. This problem was recognized by Heckman and Hotz (1989), Polachek and Kim (1994), and Winship and Morgan (1999), among others. On the other hand, the inclusion of province specific-trends makes it possible to control for the existence of common trends between the covariates and the outcome variable that might cause spurious correlations. Although the omission of time-varying unobserved heterogeneity is a common problem in studies using panel data, they seldom account for it. This type of model is especially suitable for

TABLE 1 Description of the variables

Variable	Description	Males				Females			
		Mean	s.d	Min	Max	Mean	s.d.	Min	Max
	Province level (NUTS 3)[a]								
EDUCOMP	Highest level of education people aged 16–24 who are not currently studying (Compulsory or lower)	62.8%	0.100	31.5%	78.1%	47.9%	0.082	21.4%	62.7%
AGRICUL	Share of workers aged 16–24 employed in agriculture	8.5%	0.058	0.4%	24.1%	2.8%	0.034	0.1%	18.1%
MANUFACT	Share of workers aged 16–24 employed in manufacturing	12.9%	0.054	3.6%	25.0%	6.9%	0.031	0.9%	13.3%
ENERGY	Share of workers aged 16–24 employed in mining and energy extraction	8.3%	0.036	2.9%	19.3%	1.7%	0.012	0.2%	6.7%
CONTRUCT	Share of workers aged 16–24 employed in construction	17.3%	0.032	10.0%	26.7%	1.2%	0.005	0.3%	2.6%
LS-SERVICE[b]	Share of workers aged 16–24 employed in low-skill services	33.8%	0.060	23.9%	49.0%	56.9%	0.053	47.4%	68.8%
MH-SERVICE[c]	Share of workers aged 16–24 employed in medium/high-skill services	19.2%	0.050	10.8%	35.1%	30.5%	0.046	20.5%	43.2%
POPDEN	Population density (inhabitants per Km^2)	126	164	9	778	126	164	9	778
OUTPOP	Outgoing the young population (aged 16–24)	0.6%	0.005	0.1%	2.5%	0.8%	0.008	0.1%	4.9%
INCPOP	Incoming the young population (aged 16–24)	0.7%	0.004	0.1%	2.0%	1.0%	0.006	0.3%	3.0%
UNEMP	Unemployment rate (active the population aged 16–65)	13.0%	0.038	7.2%	22.4%	15.2%	0.040	9.0%	23.9%
GDPpc	GDP *per capita*	21,093	4,280	15,530	33,167	21,093	4,280	15,530	33,167
	Region level (NUTS 2)[d]								
PUBEXP	Public expenditure in education as % of the GDP	3.2%	0.006	1.7%	4.8%	3.2%	0.006	1.7%	4.8%
PRIVEXP	Public expend. Private education as % total public expenditure in education	16.6%	0.055	9.2%	29.9%	16.6%	0.055	9.2%	29.9%

(Continues)

TABLE 1 (Continued)

		Males				Females			
Variable	**Description**	**Mean**	**s.d**	**Min**	**Max**	**Mean**	**s.d.**	**Min**	**Max**
IMSTUD	Share foreign students overall the population of students (compulsory education)	7.1%	0.029	2.7%	13.7%	7.1%	0.029	2.7%	13.7%

Notes:

[a]All the employment shares are calculated over population that are not currently studying. All the variables at NUTS 3 level are calculated from the Spanish Labour Force Survey (2002–2018).

[b]Commerce, reparation, transportation and warehouse, hostelry.

[c]Financial, Insurance, real estate, professional and scientific activities, services to firms, public administration and defense, social security, education, health and social services, artistic and leisure activities, home production of goods and services, international organizations, other services.

[d]All the variables at NUTS 2 level are obtained from the Spanish Ministry of Education.

the type of data and period used here, since employment variables, as well as educational achievement, tend to display a drifting behaviour. The econometric model reads as follows:

$$y_{it} = u_i + \delta_i f(t) + X_{it}\beta + \sum_{t=1}^{T-1} \gamma_t d_t + \varepsilon_{it}; \quad t = 1, \ldots, T; \quad i = 1, \ldots, N, \tag{1}$$

where y_{it} is the share of the population aged 16–24 without formal post-compulsory education (EDUCOMP) in region i in year t; u_i are region fixed-effects, which control for the time constant unobserved heterogeneity across provinces; and $f(t)$ is a time trend, 1, 2, …, T, specific for each province picking up time-varying unobserved heterogeneity. $f(t)$ can be either linear or a squared polynomial, and δ_i are the province specific-slopes that indicate how important the trend is for each region. d_t is a set of year dummies. The parameters γ_t associated to the year dummies are year fixed-effects capturing temporal global effects that are not picked up by the covariates and the set of regional effects, either fixed or time-varying, considered in the model.

Equation 1 refers to a random growth model (Heckman & Hotz, 1989), for which the general conditions of estimation can be found in Wooldridge (2002). The estimation is done after a detrending procedure that is more general than the demeaning used to estimate a fixed-effects model. The idea is to purge each variable from the individual specific constant (u_i) and the individual specific-trend ($\delta_i f(t)$). In practice, this is done by regressing each variable on $f(t)$ for each province i: $y_{it} = u_i + \tau_i f(t) + v_{it}$, and obtaining the transformed observation $y^*_{it} = y_{it} - (\hat{u}_i + \hat{\tau}_i f(t))$. This transformation is done with the outcome variable (y_{it}) and the k explanatory variables x_k: $x^*_{kit} = x_{kit} - (\hat{u}_{ki} + \hat{\tau}_{ki} f(t))$. Once the variables have been purged from the province fixed-effects and detrended, the following model is estimated with the transformed variables by using OLS with clustered robust standard errors:

$$y^*_{it} = u_i + \gamma X^*_{it} + \sum_{t=1}^{T-1} \gamma_t d_t + \varepsilon_{it}; \quad t = 1, \ldots, T; \quad i = 1, \ldots, N. \tag{2}$$

In a fixed-effects model, it is necessary to have sufficient variation over time in the variables to identify the coefficients with precision. Analogously, in a model with individual specific-slopes, it is necessary to have sufficient variation around the trend so that the coefficients can be identified. An alternative method to estimate a linear random trend model, though it is not exactly the same, is to calculate the first difference of the data in an initial step and then estimate a fixed-effects model with the variables in differences or a pooled OLS on a second differencing (Wooldridge, 2002). In this sense, the model proposed in Equation 1 is more flexible since it allows for polynomic trends. Equation 2 is estimated using Volker (2015).

4.2 | Endogeneity

Equation 2 considers that when the employment share increases in a low-skill industry, for example, due to an exogenous boom in that industry, the return to education decreases, and young students find it more attractive to work, instead of acquiring further education. In Equation 2, it is assumed that the employment shares by industry are exogenous. It is, however, possible that school attainment can affect the employment shares in the different industries. For example, an increased school dropout rate would increase the labour supply of low-skilled workers, which could decrease the relative wages in these sectors compared to others. The employment share could increase, due to reduced labour costs, compared to high-skilled sectors, which cannot benefit in the same way from the increased labour supply.

In the previous literature on the relation between the local labour market and school attainment (or dropout), the issue of the endogeneity is rarely considered. Most studies have taken for granted the exogeneity of the local labour market conditions, even in cases when youth unemployment is used as an explanatory variable. There is, of course, a risk of a mechanical relation between the variables as student behaviour, such as school dropout, implies changes in the stock of workers, unemployed, or people outside the labour market. For example, Black et al. (2005) use instrumental variables, but not with the purpose to rule out reverse causality between school dropout rates and the local labour supply.

Accordingly, one cannot rule out that changes in school attainment could determine the employment share in a particular industry in the local labour market. Murtazashvili and Wooldridge (2008) show how random trend models can incorporate instrumental variables, and the model is extended here to incorporate Bartik-like instrumental variables (Goldsmith-Pinkham, Sorkin, & Swift, 2019). The use of Bartik-like instruments is suitable in this application because of the particular interest here in the performance in low-skill sectors as the driving force of educational attainment. This paper considers how a common shock on the country level will provide differential exposure, depending on the industrial structure, which implies different *changes* in educational attainment. The initial employment share is used, and for each year the average growth is calculated from the first period in the employment share based on *all other regions*. Accordingly, a leave-one-out procedure is used to calculate the mean.

To obtain the Bartik-like instrumental variable, the initial employment share is multiplied by the national growth of the variable. Goldsmith-Pinkham et al. (2019) clarify that the key to the identification is the exogeneity of the instrument, *conditional on the observables*, which, in this setting, includes not only location fixed-effects and time fixed-effects but also province specific-trends. The Bartik-like instrument is detrended in the same way as explained previously. With five variables on employment shares for different sectors, an instrumental variable is prepared for each of them, and the estimated model is exactly identified. It is important to remark that the final instruments are highly correlated with their corresponding detrended endogenous explanatory variables, and a full set of year dummies is important for consistency.

5 | ECONOMETRIC RESULTS

5.1 | Specification tests

Tables 2 and 3 report the results of the estimates of Equation 1, with some alternative specifications. All analysis is done separately for boys and girls. Complexity is added sequentially to the model to illustrate initial misspecifications. The first three models (M1 to M3) are estimated with pooled OLS and clustered robust standard errors, where the first model includes only the employment share in the industries (M1). The second model adds control variables (M2), and the third model also includes year fixed-effects (M3). Column M4 reports the results of a fixed-effects model, where only variation over time for the 50 provinces is used to identify the coefficients. Models M5 to M8 add province specific-slopes to the province fixed-effects. Model M5 includes a linear trend, while model M6 considers a quadratic trend. Models M7 and M8 replicate the previous two models but using instrumental variables (IV) to handle the potential endogeneity of the employment shares. The coefficient of determination is included for all models, but it is important to keep in mind that columns M4 to M8 report its within version, that is, once the fixed-effects are purged from the model. The coefficient of determination cannot be used for model selection because any significant change in coefficients would imply invalidating too restrictive models. In all models, standard errors of the estimated coefficients are clustered.[5]

With respect to M1, which includes only employment shares, after sequentially including the set of controls (M2), year dummies (M3), and province fixed-effects (M4), estimated effects associated to industry employment

[5]We also carried out a bootstrap estimation of the standard errors in all models. We obtained that they are very similar to the clustered ones reported in Table 2 and 3. Therefore, we decide to keep the latter in the presentation of the results.

TABLE 2 Estimates of the determinants of the share of the population aged 16–24 without post-compulsory education (males)

	OLS			Fixed-Effects	Specific-slopes		Specific-slopes (IV)	
	M1	M2	M3	M4	M5	M6	M7	M8
AGRICUL	0.550***	0.342***	0.387***	0.212***	0.242***	0.223***	0.213***	0.198***
	(0.0950)	(0.0812)	(0.0777)	(0.0603)	(0.0757)	(0.0797)	(0.0890)	(0.0925)
MANUFACT	−0.158*	0.0989	0.0949	0.112***	0.104*	0.131***	0.0814	0.0989
	(0.0926)	(0.0742)	(0.0691)	(0.0513)	(0.0579)	(0.0613)	(0.0651)	(0.0639)
ENERGY	−0.326***	0.149*	0.110	0.134***	0.263***	0.290***	0.206***	0.235***
	(0.152)	(0.0884)	(0.0917)	(0.0588)	(0.0846)	(0.0898)	(0.0840)	(0.0808)
CONSTRUCT	0.709***	0.650***	0.518***	0.390***	0.258***	0.250***	0.197***	0.196***
	(0.0694)	(0.0698)	(0.0721)	(0.0635)	(0.0630)	(0.0672)	(0.0752)	(0.0790)
LS-SERVICE	0.191***	0.130***	0.160***	0.131***	0.101***	0.119***	0.0682	0.0842
	(0.0704)	(0.0641)	(0.0580)	(0.0473)	(0.0429)	(0.0494)	(0.0551)	(0.0597)
UNEMP		0.509***	0.734***	0.313	0.316	0.229	0.295	0.227
		(0.166)	(0.234)	(0.214)	(0.209)	(0.214)	(0.204)	(0.207)
POPDEN		3.69e-05	5.35e-06	−1.04e-05	2.74e-05	−0.000714	−2.60e-05	−0.000858
		(4.27e-05)	(4.83e-05)	(6.09e-05)	(0.000616)	(0.00124)	(0.000596)	(0.00126)
OUTPOP		0.907	1.002	0.700	0.389	0.801	0.332	0.764
		(0.658)	(0.722)	(0.572)	(0.597)	(0.487)	(0.574)	(0.466)
INCPOP		0.0931	0.364	−0.325	−0.888***	−0.930***	−0.936***	−0.957***
		(0.300)	(0.396)	(0.340)	(0.387)	(0.391)	(0.381)	(0.380)
GDPpc		−1.36e-05***	−1.26e-05***	−8.89e-06***	3.45e-06	4.26e-07	3.72e-06	8.45e-07
		(2.01e-06)	(2.24e-06)	(2.26e-06)	(4.79e-06)	(4.16e-06)	(4.75e-06)	(4.10e-06)
PUBEXP		−0.993	−1.322	−0.193	1.980	2.419	2.003	2.445
		(1.328)	(1.393)	(1.421)	(1.813)	(1.753)	(1.754)	(1.701)
PRIVEXP		−0.359*	−0.330*	−0.415*	0.260	0.437	0.268	0.470
		(0.180)	(0.184)	(0.247)	(0.295)	(0.366)	(0.285)	(0.356)
IMSTUD		1.277***	1.277***	0.721***	0.484	0.880	0.494	0.943
		(0.175)	(0.195)	(0.237)	(0.396)	(0.597)	(0.389)	(0.588)

(Continues)

TABLE 2 (Continued)

	OLS			Fixed-Effects	Specific-slopes		Specific-slopes (IV)	
	M1	M2	M3	M4	M5	M6	M7	M8
Year dummies	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed-effects	No	No	No	Yes	Yes	Yes	Yes	Yes
Province individual slopes	No	No	No	No	Linear	Quadratic	Linear	Quadratic
Specification tests (p-value)					(0.000)[a]		(0.1038) **	(0.6199) **
R-squared (within)	0.417	0.664	0.683	0.4478	0.2248	0.1297	0.223	0.128

Notes: Clustered robust standard errors are included in parentheses;

***$p < 0.01$,**$p < 0.05$,*$p < 0.1$;

[a]H_0: Province specific-slopes are not necessary; [b]H_0: Exogeneity.

TABLE 3 Estimates of the determinants of the share of the population aged 16–24 without post-compulsory education (females)

	OLS[a]			Fixed-Effects	Specific-slopes		Specific-slopes (IV)	
	M1	M2	M3	M4	M5	M6	M7	M8
AGRICUL	1.032***	0.707***	0.624***	0.453***	0.287*	0.215	0.294*	0.204
	(0.188)	(0.178)	(0.148)	(0.145)	(0.151)	(0.156)	(0.160)	(0.162)
MANUFACT	0.369***	0.351***	0.232***	0.323***	0.272***	0.296***	0.302***	0.324***
	(0.103)	(0.0909)	(0.0849)	(0.0763)	(0.0599)	(0.0574)	(0.0763)	(0.0774)
ENERGY	0.0469	0.154	−0.00488	0.00243	−0.0426	−0.0108	−0.0536	−0.0138
	(0.222)	(0.194)	(0.183)	(0.133)	(0.130)	(0.115)	(0.120)	(0.106)
CONSTRUCT	1.319***	0.835a	0.425*	0.337c	0.288	0.407**	0.334c	0.441**
	(0.272)	(0.227)	(0.238)	(0.182)	(0.183)	(0.173)	(0.195)	(0.181)
LS-SERVICE	0.288***	0.227***	0.211***	0.211***	0.174***	0.180***	0.178***	0.182***
	(0.0682)	(0.0507)	(0.0458)	(0.0401)	(0.0479)	(0.0521)	(0.0470)	(0.0512)
UNEMP		−0.0386	0.912***	−0.215	−0.484	−0.543*	−0.477	−0.536**
		(0.173)	(0.237)	(0.218)	(0.317)	(0.272)	(0.306)	(0.265)
POPDEN		4.29e-05	−2.03e-05	3.04e-05	0.00195**	0.00167	0.00196***	0.00167
		(4.35e-05)	(3.98e-05)	(3.94e-05)	(0.000729)	(0.00132)	(0.000708)	(0.00128)
OUTPOP		0.0656	0.160	0.444	0.476*	0.539**	0.474	0.540**
		(0.468)	(0.439)	(0.271)	(0.258)	(0.249)	(0.253)	(0.246)
INCPOP		−0.291	0.101	0.167	0.340	0.416	0.345	0.425
		(0.330)	(0.346)	(0.278)	(0.296)	(0.269)	(0.294)	(0.268)
GDPpc		−1.19e-05***	−5.97e-06*	−6.31e-06**	1.83e-07	−5.17e-07	2.28e-07	−4.29e-07
		(3.32e-06)	(3.02e-06)	(3.02e-06)	(3.87e-06)	(5.07e-06)	(3.74e-06)	(4.92e-06)
PUBEXP		0.0562	0.234	0.0603	−1.520	−1.925	−1.502	−1.981
		(1.585)	(1.379)	(1.340)	(2.084)	(2.085)	(2.047)	(2.041)
PRIVEXP		−0.330	−0.387**	−0.612***	−0.357	−0.176	−0.362	−0.181
		(0.197)	(0.178)	(0.207)	(0.473)	(0.552)	(0.461)	(0.537)
IMSTUD		1.253***	1.439***	0.936***	0.353	0.696	0.357	0.694
		0.245)	(0.206)	(0.230)	(0.574)	(0.782)	(0.563)	(0.759)

(Continues)

TABLE 3 (Continued)

	OLS[a]			Fixed-Effects	Specific-slopes		Specific-slopes (IV)	
	M1	M2	M3	M4	M5	M6	M7	M8
Year dummies	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed-effects	No	No	No	Yes	Yes	Yes	Yes	Yes
Province individual slopes	No	No	No	No	Linear	Quadratic	Linear	Quadratic
Endogeneity test (p-value)					(0.000)[a]		(0.8260)[b]	(0.7628)[b]
R-squared (within)	0.198	0.446	0.516	0.2830	0.1360	0.1332	0.1356	0.1328

Notes: Clustered robust standard errors are included in parentheses;
***$p < 0.01$,**$p < 0.05$,*$p < 0.1$;
[a]H_0: Province specific-slopes are not necessary; [b]H_0: Exogeneity.

shares change significantly. Generally, the size of the coefficients tends to decrease gradually in each step. However, for some industries, such as ENERGY and MANUFACT, the sign of the effect is even reversed. When individual slopes are considered, either linear or quadratic, coefficients associated to industry shares also change, but these changes are not very sizeable. In this regard, the consideration of a linear or quadratic trend provides fairly small changes. A version of the Hausman test (artificial regression test) was performed for comparing M4 and M5. The restricted model with only fixed-effects (M4) was rejected; therefore, including province specific-trends is necessary. Regarding the shape of the trend, the use of province-specific quadratic trends (M6 and M8) is considered, since this is more general and avoids misspecification in cases where a linear trend (M5 and M7) is too restrictive.

Models M7 and M8 treat all the employment shares as endogenous explanatory variables. However, formal tests of endogeneity cannot reject exogeneity, neither for boys nor for girls. This circumstance explains why results for models M5 to M8 are so similar. Regarding weak-instrument tests, it was obtained that, for both boys and girls, instruments are strong. For boys, in the model using a linear trend, a correlation of about 0.95 is found for the endogenous explanatory variables and the instruments, except for the construction sector, where it is 0.85. The first stage regression was estimated for the construction sector to evaluate the performance of the instruments. The coefficient of determination for this model is 0.8774, and a $F(5,49)$ statistic of 448.00, clearly rejects that the instrumental variables can be jointly excluded from the model. A $F(1,49)$ statistic for the key instrumental variable is 1962.24, and removing this variable would reduce the coefficient of determination to 0.5393, and, hence, the instrument for this specific sector had a very strong effect. Similar results are found for the rest of the industries and also for the model with quadratic trends. For girls, the correlations of the instrumental variables and the potentially endogenous explanatory variables are also very high. The lowest correlation was found for the manufacturing industry, and results from its first stage regression are reported. The coefficient of determination is 0.8425, with a $F(5,49)$ statistic of 110.71, and a $F(1,49)$ statistic of the instrumental variable for that particular sector is 414.65. The coefficient of determination would be reduced to 0.2096 if this key instrument would be omitted. Again, for girls, it is found that the potentially endogenous employment shares all have their strong instrumental variable.

With such strong instrumental variables, there is a relatively small cost in terms of increased standard errors, and one can, to be on the safe side for avoiding bias, direct the attention to the most complete model. Therefore, for both boys and girls, the interpretation of the results will focus on the IV estimation with quadratic trends (M8), despite not rejecting exogeneity. However, given the similarities of all models with province specific-slopes, conclusions derived from M8 are extensible to models M5 to M7.

Notice that in Tables 2 and 3, a within coefficient of determination (R^2-within) is reported for all models with fixed-effects (M4 to M8), but they are not comparable. The total variance is different for different models because a more general specification will purge more variation from the original variable. Spurious effects can contribute to a higher coefficient of determination for too-simple models, while the same variables can turn out not to contribute at all when that variation is already captured by province specific-slopes. This circumstance explains why, for the models with province specific-slopes (M5 to M8), R^2-within is smaller than for the model with just province fixed-effects (M4), and why, in the models with quadratic trends (M6 and M8), R^2-within is smaller than for the models with linear trends (M5 and M7).

5.2 | The impact of the industry employment shares

This section presents the results regarding the effect of the local markets' industry structure, measured by the employment shares by industry, on the share of the population aged 16–24, who are not currently studying, without formal post-compulsory education (EDUCOMP). Both the employment shares and the dependent variable are measured as proportions. In this setting, the coefficient can directly be interpreted as an elasticity. Since the sum of the shares of employment across industries is 1, one of the industries is left as base category, in this case, employment in medium- to high-skill services (MH-SERVICE); therefore, the estimated elasticities for the industries included in the model should be interpreted as a change with respect to the industry that is left out of the regression. Results are shown in Table 2 for boys and Table 3 for girls. As indicated above, for both boys and girls, results will be based on model M8, which considers province fixed-effects and province specific-slopes estimated with a quadratic trend and instrumental variables.

For boys (Table 2), the coefficients that have turned out to be statistically significant, at 5% level, are the ones associated to employment shares in agriculture (AGRICUL), energy extraction and mining (ENERGY), and construction (CONSTRUCT). The estimated effects for these three industries are quite similar. Statistical tests were carried out for the equality of these coefficients, and it was not possible to reject the null hypothesis that the three coefficients are equal, though the estimated effect for ENERGY is slightly higher. For boys, it is estimated that a one percentage point increase in the employment share in any of these three industries increases the share of the population aged 16–24 without post-compulsory education by around 0.2 percentage points. Only variation beyond a quadratic trend is used to identify the parameters. Accordingly, when the share of workers aged 16–24 employed in these three industries is higher (compared to its potentially quadratic trend), the educational attainment is reduced. This result highlights the importance of the industry mix in local labour markets for educational attainment. The sizeable estimated effects of the employment share in the energy and mining industry are in line with previous evidence from the United States reporting a negative impact of booms in this type of industry on the educational achievement of the population (Black et al., 2005; Rickman et al., 2017; Weber, 2014). The result regarding construction is in line with what Aparicio-Fenoll (2016) observed in Spain. These two industries are characterized by employing mainly low-educated men.

Table 3 shows the corresponding results for girls. For this group, it was also found that employment shares by industry are important in explaining the educational attainment of young girls. However, with respect to boys, the industries for which statistically significant effects are estimated are different. For girls, the employment shares that report significant effects are those for manufacturing (MANUFACT), low-skill services (LS-SERVICE), and, again, construction (CONSTRUC). For girls, a one percentage point increase in the manufacturing industry increases the share of girls aged 16–24 without post-compulsory formal education by about 0.32 percentage points. Estimated elasticity for low-skill services is smaller in magnitude but also sizeable, 0.18 percentage points. Interestingly, the highest elasticity, 0.44, is estimated for the construction industry, although the standard error of the parameter is also fairly high. Estimated elasticities for manufacturing and low-skill services are significant at 1% level, while the one for construction is significant at 5% level.

A few restricted models were attempted as an exercise to see how important the employment shares are. The most general model, with quadratic trends and instrumental variables (M8), is the benchmark. The coefficient of determination for the model is 0.1280 for boys. The difference in the coefficient of determination compared to this benchmark indicates how much *unique* variance the employment share in the industry explains, in addition to what other variables already have explained. If the employment share in the construction sector is omitted from the model, the coefficient of determination is 0.1017. The second largest reduction is found for the energy sector, with a coefficient of determination of 0.1044. If all employment shares are dropped, the coefficient of determination is reduced to 0.0874. This is a quite sizeable reduction (1/3). Performing the same exercise for girls implies an initial coefficient of determination of 0.1328, which is reduced by almost half (0.0722) if all the employment shares are omitted from the model. The most important reduction in the coefficient of determination is found in the employment share in low-skill services is removed. The coefficient of determination is 0.0931 for that model. The second largest reduction is found when manufacturing is removed. In that case, the coefficient of determination is 0.1026. All these results taken together indicate that, for both boys and girls, industry employment shares explain an important part of the variation of educational attainment. For boys, the industry that contributes most to explain the variation in educational attainment is construction, while for girls, the most important one is low-skill services.

5.3 | The effect of the remaining controls

In addition to the main explanatory variables of interest, namely the employment share by industry, the model also includes some other controls that otherwise might potentially bias the coefficients associated to the industry employment shares. However, it is worth noting that after including province fixed-effects and province specific-slopes, the risk of omitted variable bias is fairly small. One crucial variable is the unemployment rate (UNEMP) since this variable is picking up the movements in the business cycle. This variable is observed to have no impact on the educational attainment of boys, while for girls, the unemployment rate is negative and statistically significant. In other words, educational attainment is countercyclical for girls, as human capital theory predicts, but acyclical for boys. These results are consistent with the findings of Johnson (2013) in the United States. The signs of the coefficients for boys and girls in these models are also in line with the findings for Spain of Petrongolo and San Segundo (2002), though the statistical significance is reversed.

In these models, a series of tests were also carried out to assess the impact of the unemployment rate in the model. Although it was not possible to reject the null hypothesis about the exogeneity of the unemployment rate, models were estimated where the unemployment rate was treated as endogenous. In these models, the coefficients and their standard errors estimated with the IV method were fairly similar to the ones obtained without instrumenting the unemployment rate. All the models were also estimated with and without the unemployment rate, and in both cases, the coefficients associated to the industry employment shares were observed to be practically identical. This result suggests that these estimated coefficients are quite robust.

It is important to remark that while the coefficients associated to the industry employment shares are not very sensitive to the inclusion of the province specific-slopes, the same cannot be said regarding the remaining of the control variables. For both boys and girls, in models M2 to M4, GDP *per capita* was negative and statistically significant, whereas after including the province specific-slopes, not only does this variable lose statistical significance, but also the sign of the coefficient is reversed. This result suggests that without the province specific-slopes, the estimated relationship between educational attainment and the GDP *per capita* was in some way spurious and driven by a common trend rather than a causal effect. Similar results are found regarding the coefficient associated to the variable IMSTUD, for which statistical significance also vanishes after the province specific-slopes are included in the model. In the case of the unemployment rate, the sign is the same in all models (M1 to M8), but statistical significance vanishes after including province fixed-effects. However, after including the province specific-slopes through a quadratic trend, for girls, the coefficient associated to the unemployment rate becomes significant only at 10% level. After including the province specific-trends, the population of outgoing youth (OUTPOP) becomes negative and statistically significant for boys. This result indicates that the incoming population of boys improves the educational attainment of the province to which they move because they are more educated than the boys already residing in that province. Analogously, after including the province specific-slopes, the coefficient associated to the outgoing young population becomes positive and statistically significant for girls. That is, female inter-province migration increases the share of lower-educated girls in the provinces from which the girls are moving. The results regarding the migration variables for both boys and girls indicate that the young movers are more educated than the young stayers.

6 | CONCLUSIONS

This paper estimates the impact of the employment structure of local labour markets on the educational attainment of the population aged 16–24. With this aim, a panel data set covering the period 2002–2018 was constructed, containing information on Spanish provinces (NUTS 3). The data contain accurate regional measures of employment by industry and gender. A linear model with province fixed-effects and province specific-slopes is used, which makes it possible to control not only for time-constant unobserved heterogeneity across provinces, which can be considered as structural but also for the time-varying unobserved heterogeneity that is not controlled for with the covariates. This empirical strategy makes it possible to overcome the potential problems occurring if the strict exogeneity assumption is violated, which may arise if only province fixed-effects are controlled for. In this regard, it is important to remark that the inclusion of the province specific-slopes has a limited impact on the coefficients associated to most of the key variables, namely industry employment shares. The exception is the construction sector for boys, where an important reduction of the magnitude is found. The impact of including province specific-slopes on the coefficients associated to most of the remaining controls is quite strong. For some of these variables, after including the province specific-slopes, not only do the coefficients lose or gain statistical significance, but, in some cases, the sign is even reversed.

The empirical results are robust. After a battery of specification tests, it is concluded that the best model is the one considering province fixed-effects and province specific-slopes. The analysis found an unequivocal causal impact of the industry structure of the employment of local labour markets on the educational attainment of youth. Estimated coefficients are robust not only to the inclusion/exclusion of the unemployment rate, which captures movements in the business cycle, but also to the inclusion of year dummies, province fixed-effects, and province specific-slopes. In all alternative specifications, the coefficients associated to the employment shares keep their statistical significance, though coefficients change their magnitude, but not dramatically. The estimates generated by these empirical models suggest a crucial role of the industry employment structure in local labour markets. They estimate quite sizeable effects for some industries, though these effects differ between boys and girls. For boys, in local labour markets dominated by agriculture, construction, and mining and energy extraction, lack of educational achievement (individuals without post-compulsory education) is substantially higher. For girls, the manufacturing industry and low-skill services (mainly tourism) are the industries responsible for lower academic achievement. It is also found that, for girls, movements in employment in the construction industry reduce educational achievement.

As mentioned above, specification tests emphasize the necessity of using the model with province specific-slopes versus a model with only province fixed-effects. Although the impact of including province specific-trends on the estimated parameters associated to the industry employment shares was not as important as one might expect, their consideration has had a strong impact on the remaining controls. For some variables, considering province specific-slopes not only results in a gain or loss of statistical significance but, in some cases, causes a reversal in the direction of the effect.

The fact that important policy variables, such as the public expenditure in education, do not exert any impact on educational attainment, jointly with the finding that the industry structure of local labour markets is one of the most important factors affecting the educational achievement of the young population, suggests that policies aimed at increasing educational attainment should pay more attention to local labour markets rather than focusing on specific aspects of the educational system. The results suggest that, for a large number of school leavers, non-skilled labour is more attractive than the classroom, which in turn indicates that increasing the demand for skills in the economy may allow an important share of young workers to pursue further levels of education beyond compulsory education.

We acknowledge that one limitation of our study, as in the existing previous literature, is that we do not estimate the potential existence of the spillover effects of a boom in a specific industry. For instance, because of labour market spillovers, the movements in the employment of a specific industry may have an impact on employment in other industries. Analogously, females could be indirectly affected by a boom in a male-dominated industry. Despite this is an interesting exercise, it is not part of the objectives of our study. One exception is Weber (2014), who estimates the spillover effects of the boom in the gas extraction industry on the mining and manufacturing industries.

As these empirical analyses have shown, in Spain, the industry structure of local labour markets is quite heterogeneous. Indeed, within regions (NUTS 2), there exist sharp differences among provinces (NUTS 3), as big as among regions. In this context, nationwide policies designed by the central government are expected to be less effective than regional policies designed by sub-national (regional/local) governments. This paper claims that to improve the educational attainment of the Spanish young population, it is urgent to achieve a change of the industry structure in local labour markets through regional policies aimed at increasing the demand for skilled labour. This is a difficult challenge that will provide results only in the long term; therefore, Spanish politicians, generally more interested in policies providing results in the short run that allow them to be re-elected, and under the pressure of important business lobbies, have not seriously tried yet to tackle this issue, although this has been a recurrent debate in Spain during the past two decades.

FUNDING INFORMATION

Obra Social "La Caixa", Grant/Award Number: 2014ACUP0130; Spanish Ministry of Science, Innovation and Universities: , Grant/Award Number: RTI2018-094733-B-100

ORCID

Luis Diaz-Serrano https://orcid.org/0000-0001-9479-5091

REFERENCES

Aparicio-Fenoll, A. (2016). Returns to education and educational outcomes: the case of the Spanish housing boom. *Journal of Human Capital*, *10*, 235–265. https://doi.org/10.1086/686154

Auty R. M. (2001). The political economy of resource-driven growth. *European Economic Review*, *45*, 839–846. http://dx.doi.org/10.1016/s0014-2921(01)00126-x.

Betts, J. R., & McFarland, L. L. (1995). Safe port in a storm: The impact of labor market conditions on community college enrollments. *Journal of Human Resources*, *30*, 741–765. https://doi.org/10.2307/146230

Black, D., McKinnish, T., & Sanders, S. G. (2005). Tight labor markets and the demand for education: evidence from the coal boom and bust. *Industrial and Labor Relations Review*, *59*, 3–16. https://doi.org/10.1177/001979390505900101

Card D., Lemieux T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, *116*, 705–746. http://dx.doi.org/10.1162/00335530151144140.

Clark, D. (2011). Do recessions keep students in school? the impact of youth unemployment on enrolment in post-compulsory education in england. *Economica*, *78*, 523–545. https://doi.org/10.1111/j.1468-0335.2009.00824.x

Dellas, H., & Koubi, V. (2003). Business cycle and schooling. *European Journal of Political Economy*, *19*, 843–859. https://doi.org/10.1016/S0176-2680(03)00039-9

Di Pietro, G. (2006). Regional labour market conditions and university dropout rates: Evidence from Italy. *Regional Studies*, *40*, 617–630. https://doi.org/10.1080/00343400600868770

Diaz-Serrano, L., & Meix-Llop, E. (2018). Decentralization and the quality of public services: cross-country evidence from educational data. *Environment and Planning C: Politics and Space*, *37*, 1296–1316.

Diaz-Serrano, L., & Nilsson, W. (2018). School dropouts and the local labor markets: the role of the skills structure of the employment. Department of Economics Working Paper No. 28-2018. Universitat Rovira i Virgili, Reus, Spain.

Goldsmith-Pinkham, P., Sorkin, I., & Swift H. (2019). Bartik instruments: what when why and how. NBER Working Paper Series No. 24408.

Heckman, J., & Hotz, V. J. (1989). Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training. *Journal of the American Statistical Association*, *84*, 862–874. https://doi.org/10.1080/01621459.1989.10478848

James A., Aadland D. (2011). The curse of natural resources: An empirical investigation of U.S. counties. *Resource and Energy Economics*, *33*, 440–453. http://dx.doi.org/10.1016/j.reseneeco.2010.05.006

Johnson, M. T. (2013). The impact of business cycle fluctuations on graduate school enrollment. *Economics of Education Review*, *34*, 122–134. https://doi.org/10.1016/j.econedurev.2013.02.002

Murtazashvili, I., & Wooldridge, J. M. (2008). Fixed-effects instrumental variables estimations in correlated random coefficient panel data models. *Journal of Econometrics*, *142*, 539–552. https://doi.org/10.1016/j.jeconom.2007.09.001

Light, A. (1996). Hazard model estimate of the decision to reenroll in school. *Labour Economics*, *2*, 381–406.

López-Bazo, E., & Motellón, E. (2013). The regional distribution of unemployment: What do micro-data tell us? *Papers in Regional Science*, *92*, 383–405. https://doi.org/10.1111/j.1435-5957.2012.00456.x

López-Bazo, E., & Motellón, E. (2018). Innovation heterogeneous firms and the region: evidence from Spain. *Regional Studies*, *52*, 673–687.

Motellón, E., López-Bazo, E., & El-Attar, M. (2011). Regional heterogeneity in wage distributions. evidence from spain. *Journal of Regional Science*, *51*, 558–584. https://doi.org/10.1111/j.1467-9787.2011.00714.x

Papyrakis E., Gerlagh R. (2007). Resource abundance and economic growth in the United States. *European Economic Review*, *51*, 1011–1039. http://dx.doi.org/10.1016/j.euroecorev.2006.04.001

Peraita, C., & Pastor, M. (2000). The primary school dropout in Spain: the inuence of family background and labor market conditions. *Education Economics*, *8*, 157–168. https://doi.org/10.1080/096452900410721

Petrongolo, B., & San Segundo, M. J. (2002). Staying-on at school at 16: The impact of labor market conditions in Spain. *Economics of Education Review*, *21*, 353–365. https://doi.org/10.1016/S0272-7757(01)00019-X

Polachek, S., & Kim, M. K. (1994). Panel estimates of the gender earnings gap: Individual specific intercept and individual specific-slope models. *Journal of Econometrics*, *61*, 23–42. https://doi.org/10.1016/0304-4076(94)90075-2

Rees, D. I., & Mocan, H. N. (1997). Labor market conditions and the high school dropout rate: Evidence from New York State. *Economics of Education Review*, *16*, 103–109. https://doi.org/10.1016/S0272-7757(96)00037-4

Reiling, R., & Strøm, B. (2015). Upper secondary school completion and the business cycle. *The Scandinavian Journal of Economics*, *117*, 195–219. https://doi.org/10.1111/sjoe.12088

Rice, P. (1999). The impact of local labour markets on investment in further education: Evidence from the England and Wales youth cohort studies. *Journal of Population Economics*, *12*, 287–313. https://doi.org/10.1007/s001480050100

Rickman, D., Wang, H., & Winters, J. (2017). Is shale development drilling holes in the human capital pipeline? *Energy Economics*, *62*, 283–290. https://doi.org/10.1016/j.eneco.2016.12.013

Sachs J. D., Warner A. M. (2001). The curse of natural resources. *European Economic Review*, *45*, 827–838. http://dx.doi.org/10.1016/s0014-2921(01)00125-8

Solon, G., Barsky, R., & Parker, J. (1994). Measuring the cyclicality of real wages: how important is the composition bias? *Quarterly Journal of Economics*, *109*, 1–26. https://doi.org/10.2307/2118426

Van der Ploeg F. (2011). Natural resources: Curse or Blessing? *Journal of Economic Literature*, *49*, 366–420. http://dx.doi.org/10.1257/jel.49.2.366

Volker, L. (2015). XTFEIS: Stata module to estimate linear fixed-effects model with individual-specific-slopes (FEIS), Statistical Software Components S458045 Boston College Department of Economics. revised 21 April 2019.

Warren, J. R., & Lee, J. C. (2003). The impact of adolescent employment on high school dropout: Differences by individual and labor-market characteristics. *Social Science Research*, *32*, 98–128. https://doi.org/10.1016/S0049-089X(02)00021-2

Weber, J. G. (2014). A decade of natural gas development: the making of a resource curse? *Resource and Energy Economics*, *37*, 168–183. https://doi.org/10.1016/j.reseneeco.2013.11.013

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, *25*, 659–706. https://doi.org/10.1146/annurev.soc.25.1.659

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

How to cite this article: Diaz-Serrano L, Nilsson W. The regional anatomy of youths' educational attainment in Spain: The role of the employment structure in local labour markets. *Pap Reg Sci*. 2020;99:1487–1508. https://doi.org/10.1111/pirs.12540

DOI: 10.1111/pirs.12540

Resumen. Este artículo estudia el vínculo entre la estructura de empleo de los mercados laborales locales y las opciones de escolarización de los jóvenes en España. Se construyó un panel de provincias españolas, y se identificó el efecto de los mercados laborales locales por medio de la variación de la proporción de empleo por industria y género, tanto entre provincias como a lo largo del tiempo. Se utilizó un modelo con efectos fijos y pendientes específicas por provincia, lo que permitió controlar dos tipos de heterogeneidad no observada entre provincias: la constante y la que varía con el tiempo. Se observó un impacto considerable de la estructura del empleo del sector industrial en los logros educativos, tanto para los niños como para las niñas.

抄録: 本稿では、スペインにおける地域の労働市場の雇用構造と若者の学校教育の選択との関連を検討する。スペインの州のパネルデータを構築し、すべての州における産業及び性別ごとの就業率の経時的な変動を用いて地域の労働市場の影響を特定する。州の固定効果及び州別の直線の傾きのモデルを用いることにより、時点間で変化しない/変化する、観察されない州間の異質性を調整することが可能になる。男子と女子のいずれにも、産業別の雇用構造の学歴に対するかなりの影響が認められる。