The version of record of this article has been published and is available in Research Papers in Education (date of publication 18 May 2023)

https://doi.org/10.1080/02671522.2023.2214137

Early dropout, earnings and skills in later life: evidence from Spain

Oscar David Marcenaro-Gutierrez [1]

Luis Alejandro Lopez-Agudo[2]

Abstract

Most education systems have set a minimum age until which students must stay at school. In the case of Spain, students can drop out the same day they reach that age, even without finishing that academic year. In the present research work we intend to analyse the influence of early dropout on later life outcomes for the Spanish population, i.e. literacy and numerical skills and earnings. In order to do this, we employ PIAAC 2012 data and a fuzzy regression discontinuity design, focusing on a transition period between education laws: a first period without minimum dropout age and a second period (law approved in 1970) which obligated students to be at least 14 years-old to dropout. Our results show, for this particular cohort of the Spanish population, that dropping out is substantially negatively associated with literacy and numerical skills in around 1.5 standard deviations, while it did not influence future earnings.

Keywords: dropout; literacy skills; numerical skills; earnings; fuzzy regression discontinuity.

JEL Codes: J01, I20, I21, I28.

Acknowledgements: This work has been partly supported by *Ministerio de Ciencia e Innovación – FEDER – funding* (under Research Project PID2020-119471RB-I00); by the Andalusian Regional Government (SEJ-645) and the *Universidad de Málaga* under Research Project B1-2022_23.

[1] Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: odmarcenaro@uma.es. Tel.: +34 952137003. ORCID: 0000-0003-0939-5064.

[2] Corresponding author. Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: lopezagudo@uma.es. Tel.: +34 952137003. ORCID: 0000-0002-0906-3206.

1. Introduction

Dropout is at the core of the education debate, mainly because a wide body of literature has found that those students who fail in completing at least compulsory education tend to have lower skill levels, less qualified jobs and, then, lower earnings than those who did finish it (Tyler, Murnane, & Willett, 1999; Lacuesta, Puente, & Villanueva, 2012; Akbulut-Yuksel, 2017; Cappellari, Castelnovo, Checchi, & Leonardi, 2017; among others). Dropout is such a key issue that it extends to other aspects of society; for instance, it has been found to increase the likelihood of performing criminal activity (Bäckman, 2017) or having drug disorders (Reingle, Salas-Wright, Connell, Jetelina, Clipper, & Businelle, 2016). Although the precise motives of this dropout are quite difficult to identify, the literature seems to point towards a relevant influence of the access conditions to the labour market (Lacuesta, Puente, & Villanueva, 2012; Serrano, Soler, & Hernández, 2013; Bayón-Calvo, Corrales-Herrero, & Ogando, 2017) and family background (Choi & Calero, 2013; González-Rodríguez, Vieira, & Vidal, 2019; Almås, Cappelen, Salvanes, Sørensen, & Tungodden, 2016), although some other motives have been found to be relevant, such as e.g. low school quality (Hanushek, Lavy, & Hitomi, 2008) or levels of expenditure per pupil, class sizes or pupil-teacher ratios (Mora, Escardíbul, & Espasa, 2010). Nevertheless, in spite of the relevance of this topic and the prolific body of existing literature, it still remains an issue. In respect to this, Tyler and Lofstrom (2009) highlighted that "Although hundreds of dropout-prevention programs exist, from small, discrete programs to whole-school reform models, little hard evidence reveals what does and does not work to decrease the probability of dropping out" (p. 96).

Particularly, in this paper we will analyse this dropout issue for the case of Spain. In this country students can drop out the same day they reach the school dropout age, even without finishing the academic year[3]. Dropout rates have presented a decreasing trend since 2002[4], which were reduced from 30.9% in that year to 17.9% in 2018 (IECA, 2020); nevertheless, these rates were 7.3% higher than those of the European Union in 2018 (INE, 2020). Many researchers explored this issue for Spain and indicated that, in fact, the conditions of the Spanish labour market seem to play a relevant role in dropping out. Following Bayón-Calvo, Corrales-Herrero, and Ogando (2017), who employed data from Eurostat, the Spanish Ministry of Education and the Spanish National Statistics Institute for the period 2001-2011, this decreasing trend in dropout rates may be explained by an increase in public expenditure and a reduction in the accessibility to the less qualified construction labour market by young people. In this sense, Lacuesta, Puente, and Villanueva (2012) used data from the Spanish Labour Force Survey (1992-2009) and indicated that, for those born after 1976 in Spain, a relative increase of 10% in the earnings of less qualified construction workers provoked an increase between 2% to 5% of students with only compulsory education, in detriment of high school graduates. In fact, according to Serrano, Soler, and Hernández (2013), who analysed dropout between 2000 and 2013 in Spain using data from the Spanish Labour Force Survey, the Spanish Salary Structure Survey and the Spanish Employability Survey of the Youth, those who abandoned school mostly did it to start working. They also remarked that those who dropped out their studies had also higher risk of being unemployed for a long period of time. Similarly, Bayón-Calvo (2019) analysed this issue employing data from the Spanish Labour Force Survey for the 2000-2015 period, finding that the characteristics of the Spanish labour market

[3] This minimum school dropout age changed with some education laws, ranging from no minimum age (BOE, 1953) to 14 years (BOE, 1970a) until 16 years (BOE, 1990), being the latter the current minimum age.

[4] This is the first year with available data on dropout rates for Spain.

have played a relevant role in dropout rates (and, particularly, the easy access for young people until the 2008 crisis), together with immigrant status and mother education.

Regarding to the characteristics of those who drop out, the Spanish case has also been widely explored. Following Enguita, Mena, and Riviere (2010), who used data on dropout from the Spanish Ministry of Education for the period 1990-2008, in Spain boys tend to dropout earlier, closer to the minimum dropout age, while girls tend to stay more in the education system. In addition, many authors who analysed the Spanish education system and dropping out risk indicated that family and background characteristics are the most related to dropout (e.g. Choi & Calero, 2013, who used PISA[5] 2009 data, or González-Rodríguez, Vieira, & Vidal, 2019, who conducted their own survey among 134 Spanish teachers in 2017, among others). For instance, Serrano, Soler, and Hernández (2013) studied a sample of 18-24 Spanish individuals between 2000-2013 (using the data from the Spanish Labour Force Survey, the Spanish Salary Structure Survey and the Spanish Employability Survey of the Youth) and highlighted that there are differences in dropping out rates conditioned on the socio-economic background of the parents, as those who dropped out at earlier ages tend to have parents with compulsory education or lower, while those with parents with higher school qualification tend to abandon later. They also found that other factors such as being female or native reduced the risk of dropping out. Other authors such as Mora, Escardíbul, and Espasa (2010) also found for the Spanish 18-24 age group (using a sample of data from the Spanish Ministry of Education for the period 2003-2012) that higher levels of expenditure per pupil, lower class sizes and pupil-teacher ratios were correlated to lower dropout rates.

Unfortunately, all these research works which have analysed the determinants of dropout and the influence of dropping out in Spanish adults' outcomes have been performed using correlational methodologies. In order to identify which factors cause (and not simply which factors correlate with) the changes in earnings and skills, we have to deal with the potential bias of the dropout variable. Specifically, if we want to use the dropout variable as a factor that potentially explains the changes in earnings and skills, we have to "replace" it with a very similar one, because the dropout variable is correlated with other (unobserved) variables which also explain changes in earnings and skills. To achieve this, we employ an advanced methodological (fuzzy regression discontinuity) approach, focusing on a transition period from an education law which did not stablished a minimum dropout age (students could dropout whenever they wanted; BOE, 1953) to another in which the dropping out age was set at a minimum of 14 years (BOE, 1970a); hence, our cohort of interest would be that of those born around 1970. The way this methodology works, in simple words, is by using an instrumental variable (i.e. a variable that works as instrument) which contains similar information to (i.e. it resembles as much as possible) the problematic (dropout) variable and, at the same time, is not correlated with the error term (in which unobserved variables may be included). The law change will be that variable which will replace, in our estimations, the dropout variable, thus acting as instrument for the latter.

Furthermore, the use of "fuzzy" regression discontinuity approach requires to focus the estimations in a window of years (that is why it is "fuzzy" and not "sharp") around that law change, so students before and after the education law change are as similar as possible. Therefore, choosing a proper window of years around this education law (together with a proper instrumental variable) is also an important requirement for our methodology to work and solve the potential issues of the dropout variable.

[5] PISA stands for "Programme for International Student Assessment".

In particular, we want to answer the following question:

Does dropping out from school suppose a reduction on literacy and numeracy skills and/or earnings?

The data employed for this analysis is that from the Programme for the International Assessment of Adult Competencies (PIAAC) 2012 for Spain. The results of this programme (OECD, 2013) provide us with a general view of the relative position of Spain within the 20 OECD participant countries, which is really alarming: Spanish population was ranked in the 19th position in terms of literacy skills and the 20th in numeracy skills. Thus, it seems that the Spanish education system needs special attention in order to improve the competences of its students and, with this, those of the adult population in the future.

This research paper is closely related to the works performed in the field of returns to education, which has been widely explored and has consistently shown that higher years of education are translated into better skills and earnings in the future (Harmon, Oosterbeek, & Walker, 2003; Sianesi & Van Reenen, 2003; Hanushek, Schwerdt, Wiederhold, & Woessmann, 2015; Hanushek, Schwerdt, Wiederhold, & Woessmann, 2017; Lee & Wie, 2017; among others). Within this literature, the use of changes in education laws as instruments (like the present study) has also been extended. For instance, Braga, Checchi, and Meschi (2013) explored the influence that education legislation reforms (for the 1930-2000 time period) had on students' years of education using data on 24 European countries, finding that those reforms which were based on inclusive policies (e.g. increasing the length of compulsory education or facilitating the access to tertiary education) favoured the bottom tail of the achievement distribution, while selective reforms (e.g. those related to school quality) favoured both tails. Authors such as Cappellari, Castelnovo, Checchi, and Leonardi (2017) also employed different exposures to educational reforms in 20 countries participating in PIAAC 2012 in order to analyse the influence of years of education on earnings, using a recursive model (i.e. a set of interconnected equations forming a system of equations, in which the explained variable – outcome – of an equation acts as explanatory variable in the next equation) and education laws as instruments for years of education; they found that 1 standard deviation increase in years of education was positively associated with 22% higher earnings. Akbulut-Yuksel (2017) also employed school dropout legislations to analyse their influence on the academic attainment and earnings of a sample of United States' individuals from the 1930-1975 cohorts, finding that school-leaving laws were positively associated with the educational attainment of American students born between 1930 and 1940 but, for younger cohorts, this association was not kept. In terms of earnings, however, these legislations did not influence the labour market success of any of these cohorts. However, our research work focuses only on dropout, i.e. we study whether finishing basic education would influence adults' skills and/or earnings or not for the cohort under analysis. To the best of our knowledge, this is the first time that a research work has got so close to the causal influence of dropout on skills and earnings for Spain by trying to solve the bias that the dropout variable may introduce through the use of the previously described methodology.

The rest of the paper is structured as follows: first, the data and methodology employed are described. Then, the obtained results are presented and discussed in the conclusions section.

2. Data

The data employed in this research paper is that from PIAAC 2012 for Spain. The sample for this country is composed by 6,055 participants aged between 16 and 65 who were born between 1947 and 1996. Participants also answered a cognitive test in which their literacy and numerical skills were assessed[6]. Furthermore, PIAAC also contained a background questionnaire which included variables related to socio-economic background, the development, engagement and maintenance of literacy and numeracy skills and ICTs. In addition, information on the current labour activity of the participants, their income and employment status were also collected. Finally, participants were also asked about their health status, social trust, volunteering and political effectiveness.

Within this information we obtain the relevant variables employed for our analysis. Concretely, as dependent (explained) variables, we employ participants' standardised[7] literacy and numerical competence scores, together with hourly earnings excluding bonuses for wage and salary earners in Purchasing Power Parity (PPP) dollars and hourly earnings including bonuses for wage and salary earners in Purchasing Power Parity (PPP) dollars. Furthermore, some contextual variables were also employed; particularly, dropout, respondents' gender, immigrant status, level of education of the father and the mother, number of books at home (when the respondent was 16 years old), region and, for those working at the time of the survey, current job status, years of working experience in current job, ISCO-08 code of current job and sector of current job. To the extent that our dependent variables are measured in 2012 and the rest of controls that we will use are measured earlier, we avoid future variables explaining past variables, which may bias our results, to the extent that the "future" can never explain the "past".

Finally, the PIAAC data requires the use of methodologies such as weighting, jackknife and balanced repeated replication weights and ten plausible values, which will be applied in the current paper[8].

3. Methodology

In order to get closer to the causal influence of dropping out on literacy and numeracy skills and wages, as previously indicated, we have employed a fuzzy regression discontinuity approach. For this purpose, the sample under analysis is the cohort of people who started primary compulsory education once the education law of 1970 (*Ley General de Educación*, LGE; BOE, 1970a) was completely implemented (i.e., those born in 1969,

[6] Participants also answered a problem-solving skills questionnaire, but not in the case of Spain.

[7] This standardisation has been performed by the use of the mean and standard deviation for Spain in PIAAC 2012 (in literacy skills the mean is 251.8 and standard deviation 49, while for numeracy skills the mean is 245.8 and the standard deviation is 51.3; OECD, 2013).

[8] These adjustments are required to the extent that PIAAC is based on a sample of individuals, so it is not a whole population (census) data. Regarding weights, they are employed to scale the sample up to the size of the population, while jackknife and balanced repeated replication weights are employed to obtain robust standard errors in the estimations. In the case of plausible values, they are obtained due to the design of the cognitive test of PIAAC, in which individuals do not answer all literacy and numeracy questions, but a subset of them. This design was created by the OECD in order to reduce the burden of answering a long list of questions. Then, individuals' scores in literacy and numeracy are obtained by combining the scores of the questions that the person actually answered and the scores they would have obtained if they had answered the questions they were not administered. This is made through a multiple imputation procedure (i.e., the scores of the questions that were not administered are imputed by considering their answers to other questions and their socioeconomic characteristics). Due to the uncertainty of this imputation procedure, instead of obtaining one final score, ten potential scores are drawn, which are named as plausible values. Therefore, each estimation is performed ten times (i.e. once per plausible value) and then averaged in order to get a final estimation. More details can be found in OECD (2016); this procedure is commonly used in international large-scale assessment tests such as PISA.

who were six – the compulsory starting age – in the academic year 1975/76)[9]. This legislation of 1970 stablished a dropout cut-off at the age of 14 years; before, education was not compulsory and students could drop out whenever they wanted[10]. We used a ± 5 year of birth limit from those born in 1969 in order to stablish our sample of analysis (i.e. those born between 1964-1968 and 1969-1973). In fact, choosing this 1969 cut-off is based on another structural change: before the 1970 education law, primary education lasted 5 grades (followed by 6 grades of high school) but, when it was implemented, primary education lasted 8 grades. Hence, those students who were in the first grade of high school before the 1970 law (corresponding to 6th grade of the 1970 new system) were attending secondary education schools and not primary schools so, even if they repeated, they did not change to the 6th grade of the new 1970 system until this new law was completely implemented; this is because that would suppose being changed from a secondary to a primary school again (to start the new 6th grade of primary education).

Another element that this strategy needs is an "instrument" for dropout. As indicated in the Introduction section, in simple words, an instrumental variable is a variable (in this case, the change in education law) that substitutes the dropout variable (to avoid its potential correlation with other – unobserved – variables). This requires the instrument to be highly correlated to the dropout variable. In this case, we will employ an instrument which takes the value "1" for those born between 1969 to 1973 (those who started compulsory education when the new education law of 1970 was completely implemented) and "0" for those born between 1964 to 1968 (those who started when the education law of 1970 was being implemented). Due to the use of such an instrument, the main assumptions and properties that this instrument must fulfil are described in the following, together with the conditions that our data have to accomplish to make our methodology work.

3.1. Testing for exogenous variation

First, the instrument has to be "as good as randomly assigned", i.e., accomplish the assumption that the instrument is not correlated with the equation error term, which happens with LGE for the education legislation and also the school dropout age. However, we have employed a fuzzy instead of a sharp regression discontinuity approach because not all students who were born in 1969 and reached 14 years dropped out because of that law, and not all of those who wanted to drop out did it at that age.

3.2. Testing for discontinuity in covariates

In order to make our fuzzy regression discontinuity design to work we need to compare similar control (those born between 1964-1968) and treatment groups (those born between 1969-1973). Hence, in Table A2 (Appendix A) the main descriptive statistics of the socio-economic background variables of the control and treatment groups have been presented[11] (and also current work characteristics for those who

[9] The implementation calendar of this law was followed and indicated that compulsory education was completely implemented in the academic year 1974/75 (BOE, 1970b), so students who started compulsory education in the 1975/76 academic year (those who were six in 1975, i.e. those born in 1969) were the first ones who could not dropout until the age of 14. Table A1 (Appendix A) shows a summary of the implementation calendar of the 1970 education law.

[10] This previous education law was called "*Ley de Ordenación de la Enseñanza Media*" (BOE, 1953) and was approved in 1953.

[11] These variables have been chosen following the previous literature review, in which authors explored the factors which were found to condition dropout rates in Spain (Enguita, Mena, & Riviere, 2010; Mora, Escardíbul, & Espasa, 2010; Choi & Calero, 2013; Serrano, Soler, & Hernández, 2013; González-Rodríguez, Vieira, & Vidal, 2019).

work), together with a balance test between both groups (using a test of mean differences). We can see that both groups look similar, so that we can believe that our methodology is comparing people who are similar in the most relevant socio-economic background variables[12]; however, we will control for these socio-economic variables in our estimations, in order to avoid potential biases.

3.3. Testing for continuity of the density and the education law instrument

In the following we present a graphic in which dropout rates are plotted by year of birth (Figure 1). For the birth dates under analysis a reduction in dropout rates occurs for the population born before the cut-off year (1969), while an increase appears for those born after. This may be the result of a shock produced by this law on those born before the cut-off (who studied during the transition period), from which the dropout rates of younger cohorts were increased again, although with lower maximum rates from that point. The correlation coefficient of year of birth and dropout rates is -0.85 so, as a preliminary approach, it seems that the instrument may be correlated to the dropout variable that it instruments, i.e. it accomplishes the *relevance assumption*[13].

-Insert Figure 1-

3.4. Testing for discontinuity in the outcome variables

As indicated by Feir, Lemieux, and Marmer (2016), in order to make our methodology work we need a discontinuity in the outcomes, i.e. a clear change has to be appreciated before and after the 1969 birthdate cut-off. Thus, this has been checked by plotting the relationship, alternatively, between standardised scores and hourly earnings with year of birth (Figure 2). It can be appreciated that, in fact, the education law instrument (observed through the year of birth) seems to be related to these outcomes and influence the discontinuity in their evolution for the years under analysis (1964-1973). This correlation seems to indicate that our regression discontinuity strategy may work (to the extent that the education law instrument is not correlated to the error term; we will check this when applying the Wooldridge, 1995, endogeneity test in the Results' section). Furthermore, our variable of interest (dropout) is negatively correlated to standardised scores (a correlation of -0.46 for literacy and -0.48 for numeracy skills) and to hourly earnings (-0.19 without and with bonuses) too.

-Insert Figure 2-

There are two additional properties that our instrument needs to accomplish. First, the only causal channel of the instrument on our outcomes (skills or hourly earnings) should be through dropout. As we have previously found in the paragraph above, the instrument may fulfil this restriction, as it seems to be related to the outcomes of interest (as found in Figure 2) and also to the dropout variable, so many potential unobserved variables that may condition the relationship between the outcomes and the instrument may be controlled by both the use of a fuzzy regression discontinuity design and the addition of socio-economic background conditional variables. Another necessary property that our education law instrumental variable needs to accomplish is the

[12] There are variables which show some significant differences between both groups as e.g. years of working experience in the current work (which, in this particular case, may be due to age differences between those born in 1964-1968 and those born in 1969-1973); because of that, we have also controlled by these variables in our estimates.

[13] Nevertheless, a Stock and Yogo (2005) test of weak instruments is needed to further check this, as we will see in the Results' section.

monotonicity property (Fiorini & Stevens, 2014, Barua & Lang, 2016, or Dhuey, Figlio, Karbownik, & Roth, 2019). Following Barua and Lang (2016, p. 348) it can be defined as "while the instrument may have no effect on some individuals, all of those who are affected should be affected unidirectionally"[14]. In the view of Figure 2 we can see that, in fact, those who started primary education once the 1970 law was completely implemented (those born in 1969) actually seemed to perform higher and earn less than those who started during the implementation process or before, so our data may accomplish this property.

3.5. Sharp regression discontinuity with ordinary least squares

First, we can check the results that we may obtain if we consider that a sharp regression discontinuity approach estimated by ordinary least squares would be enough in order to obtain the influence of dropping out on Spanish literacy and numerical skills and hourly earnings for the period under analysis. Using this approach would assume that those born before and after the discontinuity are pretty similar (as we actually found in subsection 3.2) and, thus, so are their unobservable variables. Hence, we stick to our sample of analysis (those born between 1964 and 1973) and estimate two models, one which uses literacy and numerical standardised scores (alternatively) as outcomes and one for logarithmic hourly earnings (excluding and including bonuses, alternatively) as outcomes[15].

Nevertheless, as previously indicated, the influence of the dropout variable in these models would be biased if we tried to estimate them using sharp regression discontinuity, due its potential correlation with unobserved variables. Hence, we move on to our fuzzy regression discontinuity strategy.

3.6. Fuzzy regression discontinuity and two-stage least squares

In order to implement our fuzzy regression discontinuity approach, we need to make use of our education law instrumental variable to the extent that it is estimated using two-stage least squares methodology. This variable will take the value "1" for those born between 1969 and 1973 and "0" for those born between 1964 and 1968[16].

After estimating these models for the same outcomes as in the sharp regression discontinuity methodology, the dropout coefficient may not be influenced by the previously indicated issue.

4. Results

First, as previously described, we run our sharp regression discontinuity models without any kind of control but the dropping out variable, and present the results in Table 1. In this table dropping out seems to reduce both literacy and numerical skills by 0.90 standard deviations (SD) and hourly earnings by 3.4%. However, although these results may seem logical, this sharp regression discontinuity strategy may leave out many unobservable variables (as not all students who were born in 1969 and reached 14 years dropped out because of that law, and not all of those who wanted to drop out did it at that age), so we move on to our fuzzy regression approach.

-Insert Table 1-

[14] Other definition for the monotonicity property by Fiorini and Stevens (2014) is that "for a given change in the value of the instrument, it cannot be that some individuals increase treatment intensity while others decrease treatment intensity" (p. 2).

[15] The equations for these models can be found in Appendix B.

[16] The equations for these models can be found in Appendix B.

The results for this model are presented in Table 2. We can see that, without any additional control (specification I), the negative influence of dropping out becomes even more pronounced when replacing dropping out with the education law (instrument). Particularly, this influence is increased to 1.7 SD for literacy skills and 2 SD for numeracy skills. Nevertheless, for hourly earnings, dropping out does not seem to have an influence. In order to delve more into these results, we included some additional control variables on the first stage of our two-stage least squares approach. Then, in specification II, we find that this dropout influence is reduced to 1.3 and 1.6 SD for both literacy and numeracy skills, respectively, while it seems to remain null for hourly earnings[17].

-Insert Table 2-

In order to check the robustness of our results for hourly earnings due to keeping only the sample of people who work, in Table A3 (Appendix A) specification I we have replicated both specifications I and II from Table 2 not taking logarithms for our hourly earnings variables and imputing a "0" to those who do not earn money (i.e. those unemployed, pupil or student, in retirement or early retirement, permanently disabled, fulfilling domestic tasks or looking after children/family) and our results do not change. The underlying idea of this test is to avoid losing so many observations and the potential selection bias coming from omitting from our sample those people with a current non-working status which may be the result of higher difficulties obtaining a job if they had dropped out.

Alternatively, in Table A3 (Appendix A) specification II, specifications I and II in Table 2 have been replicated using multiple imputation to fill in the earnings of those currently working who have missing information on their earnings and those who were currently unemployed. In order to do this, we have stuck to the cohorts under analysis (1964-1973) to make the imputation, so that we can avoid using workers' profiles which may not belong to the cohorts under analysis. Currently unemployed people have been included to this estimation using their socio-economic characteristics (dropout, respondents' gender, immigrant status, level of education of the father and the mother, number of books at home, autonomous community) and their *last job* characteristics have substituted their current job characteristics (last job status, years of working experience in last job, ISCO-08 code of last job and sector of last job); then, multiple imputation for their earnings and for those currently working but with missing information on their earnings has been run. The results presented in this specification are similar to those in Table 2, so we can conclude that our results for earnings are robust[18] to the sample reduction of using only those with earnings' information and those currently working, so that we can account for potential difficulties of access to the labour market.

5. Discussion and conclusions

In the present research paper we have analysed the influence of dropping out from basic education on current Spanish people skills and earnings, for those born around 1970. Due to the characteristics of dropout on this analysis we have employed a fuzzy regression

[17] In Table 2 we have also included two instrumental variable tests in order to check the reliability of our instrument. The Wooldridge (1995) endogeneity test – whose null hypothesis is that the instrumental variable is not correlated to the error term – is not significant in any of the cases, even at 10%, and the Stock and Yogo (2005) test of weak instruments – whose null hypothesis is that the instrument is weak, i.e. not strongly correlated to the dropout variable – is significant in all the cases at 1%. Hence, we can conclude that our instrument seems to solve the bias of the dropout variable parameter.

[18] However, multiple imputation is not compatible with Wooldridge (1995) and Stock and Yogo (2005) tests, so their results cannot be reported.

discontinuity analysis, using as instrumental variable a change in the education legislation which set that students must be 14 years before dropping out, coming from an education law without minimum dropout age. Our main results indicate that dropping out seems to be negative for Spanish people skills, while earnings do not seem to be influenced by it.

Although our results on skills may seem logical, those related to earnings may seem counterintuitive. However, some authors such as Clark and Martorell (2014) or Akbulut-Yuksel (2017) have also found similar results; particularly, the former indicated that a high school diploma in Texas was just perceived as a piece of paper which did not prove workers' productivity (for those born around the 70s), hence having low signalling properties for employers. Thus, it seems that the Spanish labour market could have had similar characteristics for the cohorts under research (i.e. those born between 1964 and 1973) when they accessed to it. Although we cannot be completely sure when attributing this null result to particular factors, maybe the labour market had facilities for those who dropped out their studies to start, continue or access to jobs with similar limit of earnings' potential to those accessed by people who did not dropout. This may explain, in some way, the low value of the basic education diploma in that period in Spain.

Related to this issue, Spanish labour market accessibility also plays a relevant role in determining whether Spanish students drop out or not (Bayón-Calvo, Corrales-Herrero, & Ogando, 2017). Hence, a policy suggestion may be increasing the opportunity cost of dropping out education to access the labour market through, for instance, the need of basic education diplomas or higher studies to access a job, which may be translated into a reduction in dropout (Tumino & Taylor, 2015; Cabus & De Witte, 2016; Adamopoulou & Tanzi, 2017). Furthermore, a highly competitive labour market may demand higher levels of qualification to people, requesting them to invest in education in order to find a job in the future.

Further policy implications can be derived. It seems that, in fact, keeping students in the education system helps in the development of their skills. This may indicate that recent Spanish education laws such as the 1990 education law (BOE, 1990), which increased dropout age until 16 years, may have potential influences on population's skills; however, this would need further checking. Following Hupfeld (2007), resiliency-oriented practices such as the development of an exigent curriculum and high and constant support for students may also help to solve the dropout issue, as well as early detention programs to identify potential dropping out students before they can even think about it. Regarding earnings, dropping out does not seem to be a relevant factor in their determination. However, the most recent access conditions to the Spanish job market should be explored, as the average similarity in earnings for those who dropped and not dropped out for the cohorts under analysis may be a result of previous labour market conditions (those when these cohorts accessed to their jobs), so our results may not be applicable for younger cohorts.

This research has some limitations. First, we are employing a random Spanish sample so, although all OECD estimation procedures have been followed, we have to be cautious in the interpretation of our results. Second, the results for earnings are focused only on people who were actually working at the moment of the survey. Although we intended to deal with this issue, these results should also be dealt with caution. Third, these results are only applicable for the cohorts under analysis and hence may have accessed to the labour market when it had particular conditions; at the present, the Spanish labour market access has changed, so our results cannot be generalised for younger cohorts. Fourth, we focus only on Spain, so our results have internal validity, but not external validity for other countries.

References

Adamopoulou, E., & Tanzi, G. M. (2017). Academic Drop-Out and the Great Recession. *Journal of Human Capital, 11*(1), 35–71. doi: 10.1086/690650

Akbulut-Yuksel, M. (2017). Do legal school leaving rules still affect schooling and earnings? *Social Science Research, 61*, 195–205. doi: 10.1016/j.ssresearch.2016.06.013

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., & Tungodden, B. (2016). What Explains the Gender Gap in College Track Dropout? Experimental and Administrative Evidence. *American Economic Review, 106*(5), 296–302. doi: 10.1257/aer.p20161075

Bäckman, O. (2017). High School Dropout, Resource Attainment, and Criminal Convictions. *Journal of Research in Crime and Delinquency, 54*(5), 715–749. doi: 10.1177/0022427817697441

Barua, R., & Lang, K. (2016). School Entry, Educational Attainment and Quarter of Birth: A Cautionary Tale of a Local Average Treatment Effect. *Journal of Human Capital, 10*(3), 347–376. doi: 10.1086/687599

Bayón-Calvo, S. (2019). Una radiografía del abandono escolar temprano en España: Algunas claves para la política educativa en los inicios del siglo XXI. *Revista Complutense de Educación, 30*(1), 35–53. doi: 10.5209/rced.55465

Bayón-Calvo, S., Corrales-Herrero, H., & Ogando Canabal, O. (2017). Los factores explicativos del abandono temprano de la educación y la formación en las regiones españolas. *Investigaciones Regionales, 37*, 99–117.

Braga, M., Checchi, D., & Meschi, E. (2013). Educational policies in a long-run perspective. Economic Policy, 28(73), 45–100. doi: 10.1111/1468-0327.12002

BOE (1953). *Ley 26 de Febrero de 1953 sobre Ordenación de la Enseñanza Media*. Spain: BOE 58, 23rd February 1953, 1119–1130.

BOE (1970a). *Ley 14/1970, de 4 de agosto, General de Educación y Financiamiento de la Reforma Educativa*. Spain: BOE 187, 6th August 1970, 12525–12546.

BOE (1970b). *Decreto 2459/1970, de 22 de agosto, sobre calendario para aplicación de la reforma educativa*. Spain: BOE 213, 5th September 1970, 14613–14617.

BOE (1990). *Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo*. BOE 238, 4th October 1990, 28927–28942.

Cabus, S. J., & De Witte, K. (2016). Why Do Students Leave Education Early? Theory and Evidence on High School Dropout Rates. *Journal of Forecasting, 35*(8), 690–702. doi: 10.1002/for.2394

Cappellari, L., Castelnovo, P., Checchi, D., & Leonardi, M. (2017). Skilled or Educated? Educational Reforms, Human Capital, and Earnings. *Skill Mismatch in Labor Markets, 45*, 173–197. doi: 10.1108/S0147-912120170000045005

Choi, A., & Calero, J. (2013). Determinants of the risk of school failure in Spain in PISA-2009 and proposals for reform. *Revista de Educación, 362*, 562–593. doi: 10.4438/1988-592X-RE-2013-362-242

Clark, D., & Martorell, P. (2014). The Signaling Value of a High School Diploma. *Journal of Political Economy*, *122*(2), 282–318. doi: 10.1086/675238

Dhuey, E., Figlio, D., Karbownik, K., & Roth, J. (2019). School Starting Age and Cognitive Development. *Journal of Policy Analysis and Management*, *38*(9), 538–578. doi: 10.1002/pam.22135

Enguita, M. F., Mena, L., & Riviere, J. (2010). *Fracaso y abandono escolar en España*. Barcelona: Fundación Bancaria la Caixa.

Feir, D., Lemieux, T., & Marmer, V. (2016). Weak Identification in Fuzzy Regression Discontinuity Designs. *Journal of Business & Economic Statistics, 34*(2), 185–196. doi: 10.1080/07350015.2015.1024836

Fiorini, M., & Stevens, K. (2014). *Assessing the Monoticity Assumption in IV and fuzzy RD designs*. The University of Sydney Economics Working Paper Series, 13, 1–52.

González-Rodríguez, D., Vieira, M. J., & Vidal, J. (2019). La percepción del profesorado de Educación Primaria y Educación Secundaria sobre las variables que influyen en el Abandono Escolar Temprano. *Revista de Investigación Educativa*, *37*(1), 181–200. doi: 10.6018/rie.37.1.343751

Hanushek, E. A., Lavy, V., & Hitomi, K. (2008). Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries. *Journal of Human Capital*, *2*(1), 69–105. doi: 10.1086/529446

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, *73*, 103–130. doi: 10.1016/j.euroecorev.2014.10.006

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2017). Coping with change: International differences in the returns to skills. *Economics Letters*, *153*, 15–19. doi: 10.1016/j.econlet.2017.01.007

Harmon, C., Oosterbeek, H., & Walker, I. (2003). The Returns to Education: Microeconomics. *Journal of Economics Surveys*, *17*(2), 115–156. doi: 10.1111/1467-6419.00191

Hupfeld, K. (2007). A review of the literature: Resiliency skills and dropout prevention. Retrieved from http://www.scholarcentric.com/wp-content/uploads/2014/03/SC_Resiliency_Dropout-Prevention_WP_FNL.pdf

IECA (2020). *Premature dropout rate by sex*. Last time accessed January 2020 from http://www.juntadeandalucia.es/institutodeestadisticaycartografia/indsoc/indicadores/1038.htm

INE (2020). *Abandono temprano de la educación-formación de personas de 18 a 24 años en la UE*. Last time accessed January 2020 from https://www.ine.es/jaxiT3/Tabla.htm?t=12544

Lacuesta, A., Puente, S., & Villanueva, E. (2012). *The schooling response to a sustained increase in low-skill wages: evidence from Spain 1989-2009*. Working Paper No. 1208. Madrid: Banco de España.

Lee, J.-W., & Wie, D. (2017). Returns to Education and Skills in the Labor Market: Evidence from Japan and Korea. *Asian Economic Policy Review*, *12*(1), 139–160. doi: 10.1111/aepr.12169

Mora, T., Escardíbul, J.-O., & Espasa, M. (2010). The effects of regional educational policies on school failure in Spain. *Revista de Economía Aplicada, 54*(18), 79–106. http://www.redalyc.org/articulo.oa?id=96915832004

OECD (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris: OECD Publishing. doi: 10.1787/9789264204256-en

OECD (2016). *Technical Report of the Survey of Adult Skills (PIAAC)*. Paris: OECD Publishing.

Reingle Gonzalez, J. M., Salas-Wright, C. P., Connell, N. M., Jetelina, K. K., Clipper, S. J., & Businelle, M. S. (2016). The long-term effects of school dropout and GED attainment on substance use disorders. *Drug and Alcohol Dependence, 158,* 60–66. doi: 10.1016/j.drugalcdep.2015.11.002

Serrano, L., Soler, A., & Hernández, L. (2013). *El abandono educativo temprano: análisis del caso español*. Madrid: Ministerio de Educación, Cultura y Deporte.

Sianesi, B., & Van Reenen, J. (2003). The Returns to Education: Macroeconomics. *Journal of Economic Surveys, 17*(2), 157–200. doi: 10.1111/1467-6419.00192

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80–108). New York: Cambridge University Press.

Tumino, A., & Taylor, M. P. (2015). *The impact of local labour market conditions on school leaving decisions*. ISER Working Paper Series, No. 2015-14, 1–38.

Tyler, J. H., & Lofstrom, M. (2009). Finishing High School: Alternative Pathways and Dropout Recovery. *The Future of Children, 19*(1), 77–103. doi: 10.1353/foc.0.0019

Tyler, J. H., Murnane, R. J., & Willett, J. B. (1999). *Do the Cognitive Skills of School Dropouts Matter in the Labor Market?* NBER Working Paper No. 7101, 1–21.

Wooldridge, J. M. (1995). Score diagnostics for linear models estimated by two stage least squares. In G. S. Maddala, T. N. Srinivasan, and P. C. B. Phillips (Eds.), *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao* (pp. 66–87). Oxford: Blackwell.

Appendix A

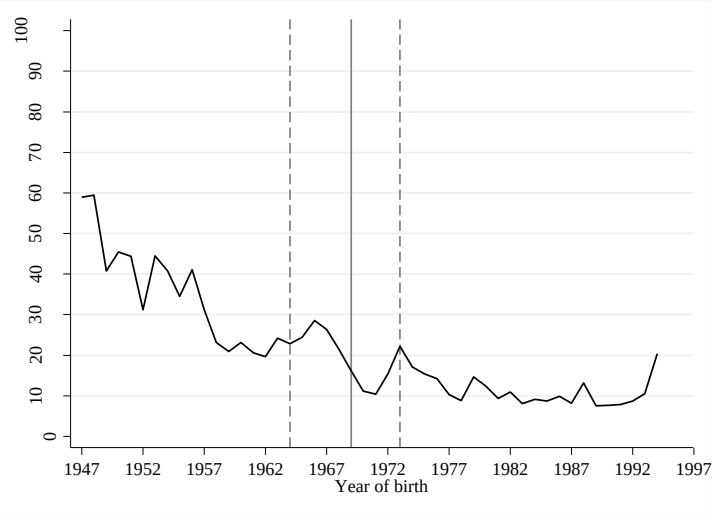
-Insert Table A1-

-Insert Table A2-

-Insert Table A3-

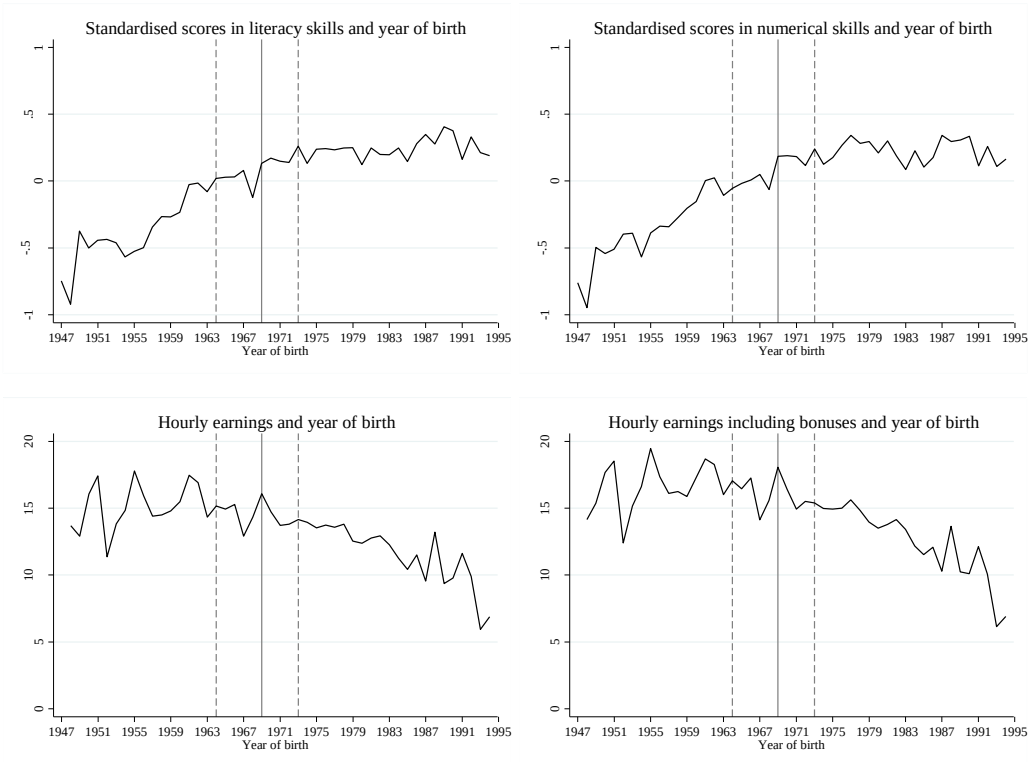
Appendix B

Figure 1. Dropout rates by year of birth



Notes: The procedures indicated in OECD (2016) (weighting and plausible values) have been employed. The bold vertical line represents the cut-off year for the education law instrument and the dotted lines the two ± 5 limits for our regression discontinuity sample.
Source: Authors' own calculations.

Figure 2. Relationship between standardised scores in literacy and numerical skills and hourly earnings with year of birth



Notes: The procedures indicated in OECD (2016) (weighting and plausible values) have been employed. The bold vertical line represents the cut-off year for the education law instrument and the dotted lines the two ± 5 limits for our regression discontinuity sample.
Source: Authors' own calculations.

Table 1. The influence of dropping out on Spanish literacy and numerical skills and earnings

Variables	Literacy skills	Numeracy skills	Logarithmic hourly earnings	Logarithmic hourly earnings + bonuses
The student dropped out (Ref.: did not drop out)	-0.856***	-0.931***	-0.326***	-0.361***
	(0.065)	(0.071)	(0.065)	(0.068)
Constant	0.261***	0.272***	2.616***	2.714***
	(0.029)	(0.028)	(0.023)	(0.024)
Observations	1,415	1,415	722	721

Notes: Standard errors are in parentheses. The procedures indicated in OECD (2016) (weighting and plausible values) have been employed. The sample is that of the people who were born between 1964 and 1973.
Dependent variable: Standardised scores in literacy and numerical skills (standardised using Spanish mean and standard deviation). Hourly earnings and hourly earnings + bonuses are in PPP dollars.
Estimation method: Sharp regression discontinuity (ordinary least squares).
Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.
Source: Authors' own calculations.

Table 2. The influence of dropping out on Spanish literacy and numerical skills and earnings

	Specification I. Without controls				Specification II. With controls			
Variables	Literacy skills	Numeracy skills	Logarithmic hourly earnings	Logarithmic hourly earnings + bonuses	Literacy skills	Numeracy skills	Logarithmic hourly earnings	Logarithmic hourly earnings + bonuses
The student dropped out (Ref.: did not drop out)	-1.665***	-2.016***	-0.133	-0.329	-1.344*	-1.608**	0.090	-0.115
	(0.545)	(0.560)	(0.581)	(0.614)	(0.784)	(0.790)	(0.492)	(0.527)
Female (Ref.: male)	-	-	-	-	-0.139**	-0.310***	-0.185***	-0.192***
					(0.062)	(0.062)	(0.035)	(0.036)
Immigrant status (Ref.: native)								
First-generation immigrant	-	-	-	-	-0.604***	-0.583***	-0.122*	-0.097
					(0.078)	(0.080)	(0.067)	(0.073)
Second-generation immigrant	-	-	-	-	-0.203	-0.100	0.094	0.134
					(0.285)	(0.288)	(0.094)	(0.167)
Level of education of the father (Ref.: ISCED 1, 2, and 3C short)								
ISCED 5 and 6	-	-	-	-	0.262**	0.180*	-0.181**	-0.132*
					(0.103)	(0.096)	(0.081)	(0.070)
ISCED 3 (excluding 3C short) and 4	-	-	-	-	0.007	-0.022	-0.072	-0.092
					(0.091)	(0.092)	(0.073)	(0.077)
Level of education of the mother (Ref.: ISCED 1, 2, and 3C short)								
ISCED 5 and 6	-	-	-	-	-0.129	-0.028	0.114	0.111
					(0.159)	(0.150)	(0.105)	(0.111)
ISCED 3 (excluding 3C short) and 4	-	-	-	-	0.027	0.005	0.070	0.064
					(0.112)	(0.105)	(0.093)	(0.104)
Number of books at home (Ref.: 10 books or less)								
More than 500 books	-	-	-	-	0.510	0.389	0.322**	0.253*
					(0.312)	(0.308)	(0.148)	(0.149)
201 to 500 books	-	-	-	-	0.595*	0.571*	0.455***	0.413***
					(0.313)	(0.311)	(0.126)	(0.132)
101 to 200 books	-	-	-	-	0.418	0.385	0.292**	0.300**
					(0.294)	(0.292)	(0.118)	(0.128)
26 to 100 books	-	-	-	-	0.261	0.267	0.211**	0.207*
					(0.249)	(0.250)	(0.104)	(0.111)
11 to 25 books	-	-	-	-	-0.023	0.000	0.173	0.113
					(0.211)	(0.211)	(0.117)	(0.124)
Current job status (Ref.: Part-time employed (self-employed, employed))								
Full-time employed (self-employed, employed)	-	-	-	-	-	-	-0.037	0.007
							(0.056)	(0.058)
Apprentice, internship	-	-	-	-	-	-	-	-
Other	-	-	-	-	-	-	-0.088	0.109
							(0.175)	(0.299)
Years of working experience in current job	-	-	-	-	-	-	0.021**	0.017
							(0.010)	(0.011)
Squared years of working experience in current job	-	-	-	-	-	-	-0.001**	-0.000
							(0.000)	(0.000)
ISCO-08 code of current job (Ref.: 9 - Elementary Occupations)								
0 - Armed Forces Occupations	-	-	-	-	-	-	0.183	0.228
							(0.186)	(0.189)
1 - Managers	-	-	-	-	-	-	0.727***	0.733***
							(0.139)	(0.141)
2 - Professionals	-	-	-	-	-	-	0.537***	0.502***
							(0.114)	(0.124)
3 - Technicians and Associate Professionals	-	-	-	-	-	-	0.264**	0.240**
							(0.111)	(0.118)

4 - Clerical Support Workers	-	-	-	-	-	-	0.183	0.167
							(0.115)	(0.125)
5 - Services and Sales Workers	-	-	-	-	-	-	0.004	-0.016
							(0.070)	(0.075)
6 - Skilled Agricultural, Forestry and Fishery Workers	-	-	-	-	-	-	0.060	0.019
							(0.112)	(0.113)
7 - Craft and Related Trades Workers	-	-	-	-	-	-	0.104*	0.099
							(0.056)	(0.061)
8 - Plant and Machine Operators and Assemblers	-	-	-	-	-	-	0.080	0.073
							(0.059)	(0.065)
Sector of current job (Ref.: public sector)								
Private sector	-	-	-	-	-	-	-0.235***	-0.221***
							(0.047)	(0.049)
Non-profit organization	-	-	-	-	-	-	-0.441***	-0.347***
							(0.106)	(0.119)
Autonomous Community (Ref.: Andalusia)								
Aragon	-	-	-	-	-0.159	0.000	0.311***	0.307***
					(0.124)	(0.127)	(0.108)	(0.108)
Asturias	-	-	-	-	0.038	0.027	0.116	0.077
					(0.156)	(0.162)	(0.103)	(0.103)
Balearic Islands	-	-	-	-	-0.151	0.038	0.192	0.230
					(0.209)	(0.229)	(0.181)	(0.187)
Canary Islands	-	-	-	-	-0.151	-0.069	0.006	0.016
					(0.145)	(0.149)	(0.114)	(0.112)
Cantabria	-	-	-	-	0.031	0.054	0.166*	0.170*
					(0.149)	(0.143)	(0.088)	(0.090)
Castile La Mancha	-	-	-	-	-0.201	-0.090	0.013	0.020
					(0.211)	(0.211)	(0.086)	(0.090)
Castile Leon	-	-	-	-	0.104	0.068	0.051	0.019
					(0.138)	(0.140)	(0.092)	(0.104)
Catalonia	-	-	-	-	-0.241*	-0.156	0.099	0.120*
					(0.145)	(0.145)	(0.065)	(0.067)
Ceuta	-	-	-	-	-0.229	-0.194	0.320***	0.364***
					(0.217)	(0.215)	(0.109)	(0.124)
Valencian Community	-	-	-	-	0.186	0.170	0.080	0.079
					(0.151)	(0.150)	(0.062)	(0.063)
Extremadura	-	-	-	-	-0.015	-0.079	-0.127	-0.144*
					(0.148)	(0.152)	(0.077)	(0.084)
Galicia	-	-	-	-	0.096	0.011	-0.028	0.037
					(0.116)	(0.118)	(0.078)	(0.089)
Madrid	-	-	-	-	-0.136	-0.059	0.169***	0.193***
					(0.143)	(0.145)	(0.061)	(0.062)
Melilla	-	-	-	-	-0.429**	-0.441*	0.036	0.032
					(0.209)	(0.239)	(0.121)	(0.118)
Murcia	-	-	-	-	-0.415**	-0.292*	0.080	0.051
					(0.167)	(0.170)	(0.097)	(0.101)
Navarra	-	-	-	-	-0.028	-0.118	0.152**	0.190***
					(0.174)	(0.178)	(0.071)	(0.071)
Basque Country	-	-	-	-	-0.410**	-0.276	0.329***	0.323***
					(0.194)	(0.189)	(0.095)	(0.097)
La Rioja	-	-	-	-	-0.065	-0.117	0.050	0.023
					(0.175)	(0.183)	(0.083)	(0.088)
Constant	0.423***	0.488***	2.591***	2.710***	0.368	0.473	2.245***	2.361***
	(0.113)	(0.115)	(0.080)	(0.084)	(0.464)	(0.464)	(0.283)	(0.303)
Observations	1,415	1,415	722	721	1,415	1,415	722	721

Instrumental variables analysis								
Wooldridge (1995) endogeneity test	2.563	4.873**	0.107	0.003	0.903	1.652	0.009	0.067
Stock and Yogo (2005) test of weak instruments	18.417***	18.417***	6.669***	6.393***	7.878***	7.878***	5.750***	5.523***

Notes: Standard errors are in parentheses. The procedures indicated in OECD (2016) (weighting and plausible values) have been employed. The sample is that of the people who were born between 1964 and 1973. A missing flag has been included in order to prevent missing information on the level of education of the father and the mother, the number of books at home, the current job status, years of working experience in current job, ISCO-08 code of current job and sector of current job. The null hypothesis of the Wooldridge (1995) endogeneity test is that the endogenous variable is now exogenous and the null hypothesis of the Stock and Yogo (2005) test of weak instruments is that the instrument is weak.

Dependent variable: Standardised scores in literacy and numerical skills (standardised using Spanish mean and standard deviation). Hourly earnings and hourly earnings + bonuses are in PPP dollars.

Estimation method: Fuzzy regression discontinuity (regression discontinuity and two-stage least squares). The instrument of "the student dropped out" is the completed transition to an education law without dropout age (for those born before 1969) to another one with a minimum age of 14 years (for those born in 1969 or after). The instrument takes the value "1" for those born after 1969 and "0" for those born before 1969.

Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.

Source: Authors' own calculations.

Table A1. Calendar of implementation for the 1970 education law for primary education

Academic year	Implemented grades	Birth year of those who started the grade in the new education system while it was being implemented (Spanish students start compulsory education at the age of six)
1970/71	1st, 2nd, 3th and 4th grades	1st grade – 1964 2nd grade – 1963 3th grade – 1962 4th grade – 1961
1971/72	5th grade	5th grade – 1961
1972/73	6th grade	6th grade – 1961
1973/74	7th grade	7th grade – 1961
1974/75	8th grade	8th grade – 1961

Source: Authors' own calculations using the implementation calendar of the 1970 education law (BOE, 1970b).

Table A2. Descriptive statistics and balance analysis

		Sample born between 1964-1968			Sample born between 1969-1973		
		Obs.	Mean	S.d.	Obs.	Mean	S.d.
Competences	**Reading**	706	252.21[S]	48.87	710	260.05[S]	46.47
	Mathematics	706	245.14[S]	51.77	710	255.10[S]	47.42
Earnings	**Hourly earnings**	346	14.90	9.64	377	15.35	10.85
	Hourly earnings + bonuses	345	16.50	11.24	377	17.27	13.45
The student dropped out	**Yes**	706	0.25[S]	0.43	709	0.15[S]	0.36
	No	706	0.75[S]	0.43	709	0.85[S]	0.36
Sex	**Male**	720	0.49	0.50	718	0.52	0.50
	Female	720	0.51	0.50	718	0.48	0.50
Immigrant status	**Native**	706	0.85	0.36	710	0.84	0.37
	First-generation immigrant	706	0.13	0.34	710	0.15	0.36
	Second-generation immigrant	706	0.02	0.12	710	0.01	0.11
Level of education of the father	**ISCED 5 and 6**	689	0.08	0.28	691	0.10	0.30
	ISCED 3 (excluding 3C short) and 4	689	0.10	0.30	691	0.10	0.30
	ISCED 1, 2, and 3C short	689	0.82	0.39	691	0.80	0.40
Level of education of the mother	**ISCED 5 and 6**	692	0.03	0.17	697	0.05	0.21
	ISCED 3 (excluding 3C short) and 4	692	0.04	0.19	697	0.06	0.24
	ISCED 1, 2, and 3C short	692	0.93[S]	0.25	697	0.89[S]	0.31
Number of books at home	**More than 500 books**	698	0.05	0.21	702	0.05	0.21
	201 to 500 books	698	0.07	0.25	702	0.07	0.28
	101 to 200 books	698	0.10[S]	0.30	702	0.14[S]	0.35
	26 to 100 books	698	0.33	0.47	702	0.37	0.48
	11 to 25 books	698	0.22	0.41	702	0.24	0.42
	10 books or less	698	0.23[S]	0.42	702	0.13[S]	0.33
Current job status	**Full-time employed (self-employed, employed)**	495	0.85	0.35	507	0.83	0.38
	Part-time employed (self-employed, employed)	495	0.13	0.32	507	0.14	0.35
	Apprentice, internship	495	0.00	0.00	507	0.00	0.00
	Other	495	0.02	0.13	507	0.03	0.16
Years of working experience in current job		464	12.82[S]	9.11	488	10.28[S]	7.36
ISCO-08 code of current job	**0 - Armed Forces Occupations**	465	0.01	0.08	488	0.00	0.04
	1 - Managers	465	0.09	0.29	488	0.10	0.29
	2 - Professionals	465	0.18	0.38	488	0.16	0.37
	3 - Technicians and Associate Professionals	465	0.10	0.30	488	0.08	0.28
	4 - Clerical Support Workers	465	0.14	0.35	488	0.17	0.38
	5 - Services and Sales Workers	465	0.16	0.37	488	0.15	0.35
	6 - Skilled Agricultural, Forestry and Fishery Workers	465	0.02	0.15	488	0.02	0.13
	7 - Craft and Related Trades Workers	465	0.08[S]	0.27	488	0.12[S]	0.33
	8 - Plant and Machine Operators and Assemblers	465	0.07	0.25	488	0.07	0.25
	9 - Elementary Occupations	465	0.15	0.36	488	0.13	0.33
Sector of current job	**Public sector**	467	0.25	0.43	488	0.22	0.41
	Private sector	467	0.73	0.44	488	0.77	0.42
	Non-profit organization	467	0.02	0.13	488	0.01	0.09
Autonomous Community	**Andalusia**	720	0.18	0.39	718	0.18	0.38
	Aragon	720	0.03	0.16	718	0.02	0.15
	Asturias	720	0.03	0.16	718	0.03	0.15
	Balearic Islands	720	0.02	0.15	718	0.03	0.17
	Canary Islands	720	0.04	0.20	718	0.05	0.21
	Cantabria	720	0.01	0.11	718	0.01	0.12
	Castile La Mancha	720	0.04	0.20	718	0.04	0.21
	Castile Leon	720	0.04	0.20	718	0.05	0.23
	Catalonia	720	0.18	0.38	718	0.15	0.36
	Ceuta	720	0.00	0.05	718	0.00	0.04
	Valencian Community	720	0.11	0.31	718	0.12	0.32
	Extremadura	720	0.02	0.15	718	0.02	0.13
	Galicia	720	0.06	0.23	718	0.05	0.22
	Madrid	720	0.14	0.34	718	0.15	0.35
	Melilla	720	0.00	0.04	718	0.00	0.03
	Murcia	720	0.03	0.18	718	0.03	0.18
	Navarra	720	0.02	0.13	718	0.01	0.11
	Basque Country	720	0.04	0.19	718	0.05	0.21
	La Rioja	720	0.01	0.09	718	0.01	0.09

Notes: The procedures indicated in OECD (2016) (weighting and plausible values) have been employed. "Obs." stands for "Observations" and "S.d." for "Standard Deviation". The "S" superscript indicates that there are significant differences (at 5% or less) between population born between 1964-1968 and population born between 1969-1973. Hourly earnings and hourly earnings + bonuses are in PPP dollars.
Source: Authors' own calculations.

Table A3. The influence of dropping out on Spanish earnings, including those who do not have earnings or using multiple imputation for missing earnings and currently unemployed people

	Specification I. Including those who do not have earnings				Specification II. Multiple imputation of those who were currently unemployed			
	Without controls		With controls		Without controls		With controls	
Variables	Hourly earnings	Hourly earnings + bonuses	Hourly earnings	Hourly earnings + bonuses	Logarithmic Hourly earnings	Logarithmic Hourly earnings + bonuses	Logarithmic Hourly earnings	Logarithmic Hourly earnings + bonuses
The student dropped out (Ref.: did not drop out)	-8.360	-11.007	-6.055	-8.524	0.186	0.112	0.118	-0.069
	(6.323)	(7.310)	(7.169)	(8.386)	(0.981)	(0.933)	(0.573)	(0.561)
Female (Ref.: male)	-	-	-1.483**	-1.671**	-	-	-0.181***	-0.196***
			(0.658)	(0.791)			(0.044)	(0.045)
Immigrant status (Ref.: native)								
First-generation immigrant	-	-	-1.433	-1.650	-	-	-0.118	-0.109
			(0.884)	(1.103)			(0.077)	(0.081)
Second-generation immigrant	-	-	1.947	4.021	-	-	0.012	0.076
			(1.508)	(2.922)			(0.169)	(0.190)
Level of education of the father (Ref.: ISCED 1, 2, and 3C short)								
ISCED 5 and 6	-	-	1.058	1.742	-	-	-0.152**	-0.118*
			(1.377)	(1.620)			(0.070)	(0.062)
ISCED 3 (excluding 3C short) and 4	-	-	0.642	0.621	-	-	-0.036	-0.063
			(1.003)	(1.250)			(0.061)	(0.063)
Level of education of the mother (Ref.: ISCED 1, 2, and 3C short)								
ISCED 5 and 6	-	-	1.050	2.439	-	-	0.103	0.104
			(2.109)	(2.834)			(0.114)	(0.123)
ISCED 3 (excluding 3C short) and 4	-	-	-1.078	-0.578	-	-	0.069	0.064
			(1.612)	(2.316)			(0.086)	(0.089)
Number of books at home (Ref.: 10 books or less)								
More than 500 books	-	-	-4.081	-6.051*	-	-	0.291	0.244
			(2.578)	(3.139)			(0.199)	(0.188)
201 to 500 books	-	-	1.980	1.744	-	-	0.463***	0.423***
			(2.320)	(2.728)			(0.159)	(0.151)
101 to 200 books	-	-	-1.794	-2.174	-	-	0.294*	0.299**
			(2.126)	(2.517)			(0.153)	(0.149)
26 to 100 books	-	-	-1.119	-1.416	-	-	0.210	0.204
			(1.696)	(2.038)			(0.140)	(0.132)
11 to 25 books	-	-	-1.068	-1.735	-	-	0.174	0.123
			(1.622)	(1.917)			(0.133)	(0.124)
Current job status (Ref.: In specification I: Fulfilling domestic tasks or looking after children/family; In specification II: Part-time employed)								
Full-time employed (self-employed, employed)	-	-	2.379*	2.251	-	-	-0.022	0.015
			(1.298)	(1.537)			(0.059)	(0.063)
Part-time employed (self-employed, employed)	-	-	3.045*	2.402	-	-	-	-
			(1.636)	(1.847)				
Unemployed	-	-	0.041	0.038	-	-	-	-
			(0.564)	(0.695)				
Pupil, student	-	-	-3.409	-4.050	-	-	-	-
			(4.488)	(5.260)				
Apprentice, internship	-	-	-	-	-	-	-	-
In retirement or early retirement	-	-	0.494	0.962	-	-	-	-

			(1.845)	(2.447)				
Permanently disabled	-	-	1.586	2.374	-	-	-	-
			(2.518)	(2.979)				
Other	-	-	0.608	2.819	-	-	0.066	0.240
			(1.445)	(2.737)			(0.225)	(0.257)
Years of working experience in current job	-	-	0.134	0.061	-	-	0.021***	0.018**
			(0.153)	(0.201)			(0.008)	(0.008)
Squared years of working experience in current job	-	-	-0.007	-0.005	-	-	-0.001**	-0.001**
			(0.005)	(0.007)			(0.000)	(0.000)
ISCO-08 code of current job (Ref.: 9 - Elementary Occupations)								
0 - Armed Forces Occupations	-	-	3.105	3.647	-	-	0.201	0.226
			(2.846)	(3.314)			(0.199)	(0.192)
1 - Managers	-	-	-0.182	0.150	-	-	0.710***	0.702***
			(2.202)	(2.541)			(0.144)	(0.139)
2 - Professionals	-	-	2.786	2.965	-	-	0.548***	0.513***
			(2.308)	(2.602)			(0.142)	(0.142)
3 - Technicians and Associate Professionals	-	-	0.380	0.197	-	-	0.266*	0.234*
			(2.512)	(2.704)			(0.143)	(0.141)
4 - Clerical Support Workers	-	-	-0.693	-0.557	-	-	0.184	0.169
			(2.047)	(2.294)			(0.129)	(0.132)
5 - Services and Sales Workers	-	-	-1.985	-2.167	-	-	0.013	-0.005
			(1.273)	(1.373)			(0.068)	(0.068)
6 - Skilled Agricultural, Forestry and Fishery Workers	-	-	-3.936***	-4.212***	-	-	0.022	-0.010
			(1.439)	(1.623)			(0.160)	(0.168)
7 - Craft and Related Trades Workers	-	-	-1.447	-1.380	-	-	0.132**	0.127*
			(1.174)	(1.322)			(0.063)	(0.069)
8 - Plant and Machine Operators and Assemblers	-	-	-0.167	-0.175	-	-	0.099	0.085
			(1.265)	(1.426)			(0.076)	(0.085)
Sector of current job (Ref.: public sector)								
Private sector			-6.249***	-6.668***			-0.236***	-0.215***
			(1.276)	(1.399)			(0.051)	(0.055)
Non-profit organization			-5.906***	-5.094**			-0.468***	-0.380***
			(1.776)	(2.329)			(0.136)	(0.141)
Autonomous Community (Ref.: Andalusia)								
Aragon	-	-	3.289	3.428	-	-	0.276**	0.290**
			(2.557)	(2.614)			(0.119)	(0.116)
Asturias	-	-	0.784	0.333	-	-	0.108	0.078
			(1.757)	(1.874)			(0.107)	(0.106)
Balearic Islands	-	-	3.576	3.557	-	-	0.174	0.199
			(4.206)	(4.301)			(0.150)	(0.147)
Canary Islands	-	-	-1.144	-1.344	-	-	-0.019	-0.011
			(1.109)	(1.259)			(0.132)	(0.130)
Cantabria	-	-	2.509*	2.571*	-	-	0.160	0.161
			(1.312)	(1.504)			(0.100)	(0.108)
Castile La Mancha	-	-	-0.698	-1.123	-	-	0.014	-0.014
			(1.845)	(2.129)			(0.102)	(0.103)
Castile Leon	-	-	-0.252	-0.753	-	-	0.042	0.020
			(1.186)	(1.431)			(0.095)	(0.094)
Catalonia	-	-	1.351	1.532	-	-	0.091	0.114
			(1.057)	(1.225)			(0.073)	(0.074)
Ceuta	-	-	3.408*	4.057*	-	-	0.347***	0.395***
			(1.750)	(2.129)			(0.119)	(0.119)
Valencian Community	-	-	0.299	0.065	-	-	0.076	0.078

			(1.134)	(1.300)			(0.067)	(0.068)
Extremadura	-	-	-2.286*	-2.990**	-	-	-0.119	-0.095
			(1.193)	(1.371)			(0.105)	(0.109)
Galicia	-	-	-0.906	-0.481	-	-	-0.028	0.038
			(0.857)	(1.070)			(0.118)	(0.126)
Madrid	-	-	2.405**	3.351**	-	-	0.165**	0.195***
			(1.113)	(1.337)			(0.064)	(0.065)
Melilla	-	-	-0.202	-0.564	-	-	-0.002	-0.041
			(1.938)	(2.079)			(0.209)	(0.211)
Murcia	-	-	0.592	0.380	-	-	0.079	0.070
			(1.276)	(1.477)			(0.089)	(0.085)
Navarra	-	-	-1.346	-1.541	-	-	0.166	0.187*
			(1.528)	(1.772)			(0.113)	(0.114)
Basque Country	-	-	1.968	1.999	-	-	0.322***	0.317***
			(1.783)	(2.037)			(0.100)	(0.102)
La Rioja	-	-	-1.489	-1.999	-	-	0.077	0.038
			(1.319)	(1.558)			(0.125)	(0.125)
Constant	9.601***	11.036***	15.535***	18.355***	2.608***	2.706***	2.219***	2.331***
	(1.350)	(1.566)	(5.663)	(6.394)	(0.166)	(0.158)	(0.342)	(0.330)
Observations	1,415	1,415	1,415	1,415	1,181	1,181	1,181	1,181
Instrumental variables analysis								
Wooldridge (1995) endogeneity test	0.330	0.568	0.708	1.037	X	X	X	X
Stock and Yogo (2005) test of weak instruments	18.417***	18.417***	10.415***	10.415***	X	X	X	X

Notes: Standard errors are in parentheses. The procedures indicated in OECD (2016) (weighting and plausible values) have been employed. The sample is that of the people who were born between 1964 and 1973. A missing flag has been included in order to prevent missing information on the level of education of the father and the mother, the number of books at home, the current job status, years of working experience in current job, ISCO-08 code of current job and sector of current job. The null hypothesis of the Wooldridge (1995) endogeneity test is that the endogenous variable is now exogenous and the null hypothesis of the Stock and Yogo (2005) test of weak instruments is that the instrument is weak (these tests are not available for specification II). Multiple imputation in specification II for unemployed people has been performed (10 imputations) using their socio-economic characteristics (dropout, respondents' gender, immigrant status, level of education of the father and the mother, number of books at home, autonomous community) and their last job characteristics (last job status, years of working experience in last job, ISCO-08 code of last job and sector of last job). For people with missing earnings, socio-economic characteristics and current job characteristics have been employed. The cohorts employed for this multiple imputation are those under analysis (1964-1973).

Dependent variable: Specification I: Hourly earnings and hourly earnings + bonuses are in PPP dollars. Those who do not have earnings have a "0". Specification II: Logarithmic hourly earnings and logarithmic hourly earnings + bonuses are in PPP dollars.

Estimation method: Fuzzy regression discontinuity (regression discontinuity and two-stage least squares). The instrument of "the student dropped out" is the completed transition to an education law without dropout age (for those born before 1969) to another one with a minimum age of 14 years (for those born in 1969 or after). The instrument takes the value "1" for those born after 1969 and "0" for those born before 1969.

Coefficient: ***Significant at 1%, ** significant at 5%, * significant at 10%.

Source: Authors' own calculations.

Appendix B

B.1. Sharp regression discontinuity with ordinary least squares

First, we can check the results that we may obtain if we considered that a sharp regression discontinuity approach estimated by ordinary least squares would be enough in order to obtain the influence of dropping out on Spanish literacy and numerical skills and hourly earnings for the period under analysis. Using this approach would assume that those born before and after the discontinuity are pretty similar (as we actually found in subsection 3.2) and, thus, so are their unobservable variables. Hence, we stick to our sample of analysis (those born between 1964 and 1973) and estimate the following regression model:

$$s_i = \alpha + \beta DO_i + \gamma X_i + \varepsilon_i \quad (1)$$

where i represents each individual. s_i are literacy and numerical standardised scores (alternatively); DO_i is a dummy variable which takes the value "1" if the individual dropped out his/her studies before finishing ISCED 2 level and "0" otherwise; X_i are individual's background characteristics; ε_i is the idiosyncratic error term.

For hourly earnings, the sample under analysis is that of people who are currently working, and the specification would be:

$$\ln(y_i) = \alpha + \beta DO_i + \gamma X_i + \delta W_i + \varepsilon_i \quad (2)$$

where $\ln(y_i)$ are logarithmic hourly earnings (excluding and including bonuses, alternatively) and W_i are current job characteristics.

Nevertheless, the β coefficient of interest would be biased if we tried to estimate this model using sharp regression discontinuity, to the extent that there are potential unobservable variables in ε_i which may bias this coefficient. Hence, we move on to our fuzzy regression discontinuity strategy.

B.2. Fuzzy regression discontinuity and two-stage least squares

In order to implement our fuzzy regression discontinuity approach we need to make use of our education law instrumental variable, to the extent that it is estimated using two-stage least squares methodology. This variable is going to be denoted as Z_i and will take the value "1" for those born between 1969 and 1973 and "0" for those born between 1964 and 1968. Then, we estimate models (1) and (2) using two-stage least squares. The *first stage* would be, alternatively for standardised scores (3) and hourly earnings (4), defined as:

$$DO_i = \pi_0 + \pi_1 Z_i + \pi_2 X_i + \vartheta_i \quad (3)$$

$$DO_i = \pi_0 + \pi_1 Z_i + \pi_2 X_i + \pi_3 W_i + \vartheta_i \quad (4)$$

where ϑ_i is the idiosyncratic error term. From each one of these equations we obtain a prediction of the dropping out variable ($\widehat{DO}_i$) and introduce it in the original model in (1) and (2), respectively, specifying the *reduced form* as follows:

$$s_i = \alpha + \beta \widehat{DO}_i + \gamma X_i + \varepsilon_i \quad (5)$$

$$\ln(y_i) = \alpha + \beta \widehat{DO}_i + \gamma X_i + \delta W_i + \varepsilon_i \quad (6)$$

where β would be, in this case, the influence of dropout on the outcomes under analysis. The fact that our fuzzy regression discontinuity approach works depends on the identification of a proper instrument (as it was previously argued) and also that our control variables (X_i and W_i), combined with the regression discontinuity, are able to net out any

potential unobserved variable from the influence of the instrument (through the channel of the predicted dropout, $\widehat{DO}_i$) on the outcome variables.