RETURNS TO EDUCATION AND EDUCATIONAL OUTCOMES: THE CASE OF THE SPANISH HOUSING BOOM

Ainhoa Aparicio-Fenoll*†

Collegio Carlo Alberto

March 10, 2014

Abstract

This paper provides a novel identification strategy to estimate how returns to education affect school enrollment. It also explores the consequences of changes in returns to education on students' performance as measured by grade completion. The identification strategy relies on the fact that the construction sector employs mostly uneducated men and hence the Spanish housing boom significantly decreased the difference in returns to education between men and women. Results show that a 10% increase in the ratio of wages of educated to uneducated individuals leads to a 2% increase in the probability of being enrolled in school and a 0.2% increase in grade completion among 16 to 18 year-olds. These findings suggest that the influence of returns to education on educational outcomes is sizeable and wider than previously thought.

JEL codes: R12, J24, J22, I20, L74

Keywords: Returns to education, school enrollment, grade completion, housing boom, Spain

*Collegio Carlo Alberto. Address: Via Real Collegio 30, Moncalieri (TO) 10024, Italy. Email: ainhoa.aparicio@carloalberto.org

†This paper has benefited from comments by Jerome Adda, Saku Aura, Ghazala Azmat, Cristian Bartolucci, Gordon Dahl, José Ignacio García Pérez, Libertad González, Francis Kramarz, Zoe Kuehn, Steve Machin, Julián Messina, Mario Pagliero, Núria Rodríguez Planas, Kurt Schmidheiny, Chris Taber, Thijs Van Rens, Marian Vidal-Fernández, Till von Wachter and Stèphane Bonhomme.

1 Introduction

When making decisions about whether or not to enroll in school and about how much effort to exert in studying, do individuals take into consideration current returns to education? Becker (1964) formalizes how education decisions depend on expected returns to education. In his model, individuals choose their education by comparing the flow of expected returns under different educational levels. Following this reasoning, young individuals decide to become educated if they infer that the expected sum of their discounted future wages if they acquire education is higher than the expected sum of discounted future wages if they enter the labor market. Therefore, theory predicts that enrollment increases with returns to education. Besides, returns to education could also affect the performance of enrolled individuals. If students that respond to rises in returns to education by enrolling in school are low ability students, student performance may decrease due to peer effects because newly enrolled students could exert a negative influence on the performance of the rest.[1] On the other hand, if returns to education increase, enrolled students could increase their effort in order to increase the probability of completing their studies. In this paper, I measure returns to education by wages of educated workers relative to wages of uneducated workers and I test the impact of returns to education on the propensity to enroll in post-compulsory education and on the probability of student grade completion.

Estimating the impact of returns to education on educational outcomes is challenging. Simple correlations are misleading for two reasons: First, rises in the supply of uneducated workers reduce wages of uneducated relative to educated workers and hence increase returns to education (reverse causality). Second, many unobserved factors could affect both educational outcomes and returns to education. For instance, an improvement in the curricula of high school education could make high school graduation more difficult for students and more valuable to employers, which implies lower educational outcomes and

[1] Evidence about the relevance of peer effects can be found in Zimmerman (2003) and Falk and Ichino (2006), among others.

higher returns to education. In order to address both concerns, I use a novel identification strategy that relies on exogenous variation in returns to education for males with respect to females induced by the Spanish housing boom (1997-2006).

The Spanish housing boom modified returns to education of men because the construction sector typically demands a large share of uneducated male workers. Figure 1 shows the composition of the labour force for the construction sector and for the entire economy during the period 1997-2006 in terms of gender and level of education. Fractions displayed in the graph show that men are overrepresented in construction and that men working in construction are disproportionately uneducated. During the housing boom, 23% of male workers with less than high school education worked in construction. In contrast, only 5% of male workers with at least high school education worked in the construction sector during the same period. Women represent a negligible part of the construction workforce (5% of total construction workers are female) and only 1.42% of female workers worked in the construction sector at that time. The housing boom implied an increase from 7% to 12% in the contribution of construction to total value added from 1997 to 2006. This translated into an increase of 30% in the number of construction workers and an increase of 47% in construction wages, leading to a substantial improvement in labor market prospects for uneducated male individuals.

The validity of my identification strategy is justified by two facts: (i) the housing boom generated substantial variation in construction activity along time and across regions, and (ii) the housing boom is exogenous to differences in educational outcomes between men and women.

The housing boom represented a huge increase in construction activity in many countries as a response to the sharp rise in housing prices, in what the Economist labeled "the biggest bubble in history"[2]. The contribution of the construction sector to total value added increased between 1997 and 2006 by 15% in the United States, 28% in the

[2] "The global housing boom", June 16th, 2005.

United Kingdom and 67% in Spain[3]. These numbers become even more impressive if one considers that from 1997 to 2006 total value added increased by 70% in the United States, 107% in the United Kingdom and 146% in Spain. Figure 2 shows the evolution of the contribution of construction to total value added in Spain. We observe that the share of the construction sector in value added has been continuously increasing during the period 1997-2006. Regarding geographical variation, the housing boom took place in all the 52 Spanish provinces, however with different intensities (see Figure 3). The differences in the magnitude of the housing boom across provinces produce the geographic variation in gender differences in returns to education that allows me to identify how educational outcomes change with returns to education.

The housing boom originated in an increase in the demand of housing as a consequence of excessive lending. Caballero and Krishnamurthy (2009) suggest that excessive lending originates in financial market's rising demand for non-risky assets to be used for securitization of risky assets. According to Rajan (2010), political pressure to make housing affordable for all citizens created excessive lending in the financial sector. This translated into an increase in housing prices and distortion in private-sector incentives that led to excessive investment in construction. Hence, there is consensus in the economic literature that the determinants of the housing boom are unrelated to educational outcomes. However, credit standards are very homogeneous across regions, so what drives the geographical variation in the housing boom? The evidence in Gonzalez and Ortega (2013) suggests that lax credit standards had greater impact on the demand for housing in some regions due to the presence of immigrants. To the extent that male and female students are equally affected by the presence of immigrant students, the proposed identification strategy should be valid.

Given the considerations above, the Spanish housing boom provides exogenous variation in differences in returns to education between genders that can be used to identify the causal impact of returns to education on enrollment in post-compulsory education and on

[3]Data source: National Accounts Main Aggregates Database. United Nations Statistics Division.

grade completion. As a first approximation, I estimate the impact of the housing boom on the difference between male and female educational outcomes. These estimations are useful to assess whether returns to education have an effect on educational outcomes and in this case, whether the sign of the effect is consistent with the predictions of Becker's model. However, it does not provide a sense of the magnitude of the effect. In my baseline estimations I provide a sense of the magnitude of the impact of returns to education on enrollment and on grade completion. I perform IV estimations where the level of construction activity as measured by construction over total value added is the instrument for differences in returns to education between genders.

The empirical model presents several innovations. Due to its formulation in gender differences, the model is the first to account for unobserved factors at the province level that vary over time such as the supply of schools or policies aimed at improving the quality of education. Moreover, the use of construction activity as an instrument for differences in returns to education between genders eliminates the influence of confounding factors that affect men and women differently such as gender policies. The instrumental variable strategy also removes the distortion induced by reverse causality, i.e., the fact that changes in the relative supply of uneducated men with respect to uneducated women affect the difference between male and female returns to education.

Estimation results indicate that a 10% increase in returns to education as measured by the ratio of wages of educated to uneducated workers increases the probability of enrollment in post-compulsory education by 2% and increases the proportion of grade completers by 0.2% among 16-18 year-olds.

The rest of the paper proceeds as follows. Section 2 reviews the previous literature. Section 3 explains the methodology. Section 4 describes the Spanish educational system and the data used in the analysis. Section 5 presents the empirical results. Section 6 concludes.

2 Literature Review

This paper belongs to the literature on returns to education and young individuals' educational outcomes. The effect of returns to education on educational outcomes is typically analyzed using as proxies for returns to education certain labor market indicators that affect uneducated workers to a greater extent. The unemployment rate is the most often used indicator. Some of the papers that explore the relationship between the unemployment rate and educational outcomes are Duncan (1965), Rumberger (1983) and Dellas and Sakellaris (2003) for the US, Petrongolo and San Segundo (2002) for Spain and Clark (2011) for Britain. Some of these papers find that the correlation between the unemployment rate and the probability of being enrolled in school is positive and others find that there is no correlation. However, one should be careful when interpreting these estimates as causal effects because school dropouts might sign up for unemployment disproportionately more often.

Other studies use arguably more exogenous labor market outcomes as proxies for returns to education. Neumark and Wascher (1995) study the role of minimum wages, which apply predominantly to uneducated workers, in determining school enrollment. Goldin and Katz (1997) make use of the fact that most manufacturing jobs do not require a high school diploma. Thus, the authors interpret an increase in the fraction of employment in manufacturing as a reduction in returns to education. Abramizky and Lavy (2011) interpret a change in wage determination schemes in Israeli Kibbutzims that makes wages depend more on productivity as an increase in returns to education. The last three papers find positive effects of returns to education on the probability of school enrollment. Again, concerns about the potential endogeneity of returns to education apply to these papers. The proportion of uneducated individuals in the population may affect the determination of minimum wages, the level of manufacturing production and the implementation of wage determination schemes.

Two recent papers quantify the causal impact of returns to education as measured by wages on educational outcomes. Lacuesta, Puente and Villanueva (2012) rely on the

observation that "unskilled young men typically work in construction and retail, while skilled young men work in services to industries, health and education" and measure returns to education using collectively bargained wages in those sectors. As a result, a significant part of their identifying variation comes from the housing boom. Their findings indicate that a 10% increase in the ratio of wages of unskilled workers to the wages of medium-skilled workers increases the fraction of males completing at most compulsory schooling between 2 and 5 percentage points.[4] Black, McKinnish and Sanders (2005) study the effect of the Appalachian coal boom on high school enrollment. The authors make use of the fact that the coal boom increased wages only in mining areas. They find that a 10% increase in the earnings of low-skilled workers as a consequence of the coal boom reduces high school enrollment rates by 5-7%.

Black, McKinnish and Sanders (2005) is the closest paper to mine. Different from my data, the wage data available to them is not disaggregated by level of education and hence they need to rely on wage variation at the regional level being caused exclusively by changes in returns to education. Moreover, their data does not allow to control for time-varying factors at the regional level (they explicitly mention that they would have benefited from data disaggregated by gender). My paper contributes to the previous literature by proposing an identification strategy that jointly: (i) allows to abstain from time-varying confounding effects at the regional level by comparing male and female individuals within a region over time and (ii) controls for endogeneity using the exogenous variation in returns to education induced by the housing boom.

To the best of my knowledge, this is the first paper to study the impact of returns to education on a measure of student performance like grade completion. However, grade completion can be considered as a more detailed outcome for graduation (individuals who complete all grades graduate). The recent paper by Lopez-Mayan (2010), Abramitzky and Lavy (2011) and Lacuesta et al (2012) analyze high-school graduation. They find

[4] Unfortunately my results are not fully comparable to those from Lacuesta et al (2012) because the two studies use different measures of returns to education, refer to different populations and analyze different outcomes.

that graduation rates increase in response to rises in expected returns to education.

3 Methodology

3.1 The Impact of the Housing Boom on Enrollment and Grade Completion

As a first approximation, I estimate the probability of school enrollment as a function of the strength of the housing boom:

$$\begin{aligned} Enrollment_{gpt} &= \beta_0 + \beta_1 CA_{pt} \cdot U_g + \beta_2 CA_{pt} + \beta_3 U_g + \beta_4 X_{gpt} \\ &\quad + \beta_5 V_p \cdot U_g + \beta_6 W_t \cdot U_g + \beta_7 Z_{pt} + \varepsilon_{gpt} \end{aligned} \tag{1}$$

where $Enrollment_{gpt}$ is the school enrollment rate of gender g in province p at time t, and CA measures the level of construction activity. The vector X contains average individual characteristics (age and an immigrant indicator) and average family characteristics (high-school graduated father, high-school graduated mother, university graduated father, university graduated mother, absent mother, absent father, absent parents, number of brothers, number of sisters, and number of siblings in the 0-10, 11-15, 16-20, 21-30 and over 30 age groups). I denote gender, province, and time (quarter and year) dummies by U, V, and W, respectively. Z_{pt} stands for province-specific time trends and ε is the residual.[5]

A negative (positive) estimated coefficient β_1 provides evidence in favor (against) Becker's theory that individuals respond to increases in returns to education by staying longer in school. In order to learn about the magnitude of the impact of returns to skill on enrollment in post-compulsory education, I estimate an alternative specification which

[5] Standard errors are clustered at the province level and population weights account for the number of individuals in each gender-province-time cell.

is described in the next section.

I also estimate the grade completion rate as a function of the strength of the housing boom:

$$\begin{aligned} Completers_{cgpt} &= \alpha_0 + \alpha_1 CA_{pt} \cdot U_g + \alpha_2 CA_{pt} + \alpha_3 U_g + \alpha_4 K_{ct} \cdot U_g \\ &\quad + \alpha_5 V_p \cdot U_g + \alpha_7 W_t \cdot U_g + \alpha_8 Z_{pt} + u_{gpt} \end{aligned} \tag{2}$$

where $Completers$ is the proportion of individuals that successfully complete grade c, K are cohort dummies, and the rest of variables are as defined before.

A negative (positive) estimated coefficient α_1 indicates that grade completion increases (decreases) with returns to education. In Section 3.3, I describe an alternative specification that sheds light on the magnitude of the effect.

3.2 Enrollment in Post-Compulsory Education

I estimate the impact of returns to education as measured by relative wages on the probability of enrollment in post-compulsory education. Ideally, I would like to estimate the following model:

$$\begin{aligned} Enrollment_{igpt} &= \gamma_0 + \gamma_1 \frac{W_{\mathrm{E}i\,gpt}}{W_{Uigpt}} + \gamma_2 X_{igpt} + \gamma_3 U_g + \gamma_4 V_p \cdot U_g \\ &\quad + \gamma_5 W_t \cdot U_g + \gamma_6 Z_{pt} + v_{gpt} \end{aligned} \tag{3}$$

where $Enrollment_{igpt}$ is a dummy equal to one if individual i of gender g in province p at time t is enrolled in school and zero otherwise, and $\frac{W_E}{W_U}$ measures returns to education computed as the ratio of the wages individuals would earn if educated with respect to their wages if uneducated. The rest of the variables are as described before. Unfortunately, individual returns to education are unobserved; for educated (uneducated) individuals

the wages they would earn if they were uneducated (educated) are unobserved. Hence, I measure returns to education as the average by region, time and gender group. Given this, one can obtain more reliable standard errors by formulating the model at the gender-province-time level:[6]

$$\overline{Enrollment}_{gpt} = \gamma_0 + \gamma_1 \frac{W_{Egpt}}{W_{Ugpt}} + \gamma_2 \overline{X}_{gpt} + \gamma_3 U_g + \gamma_4 V_p \cdot U_g + \gamma_5 W_t \cdot U_g + \gamma_6 Z_{pt} + \overline{v}_{gpt} \quad (4)$$

where the upper bar denotes the average of the corresponding variable in each gender-province-time cell. Unfortunately, the estimated γ_1 resulting from this model cannot be interpreted as causal effect because returns to education are potentially endogenous. In order to address this, I could restrict the sample to males and estimate the equation above by instrumental variables using construction over total value added as an instrument for male returns to education. Unfortunately, if construction over total value added reflects changes in the availability and/or affordability of housing, it will not be a valid instrument. The reason is that more housing and/or lower housing prices could induce students to drop out of school and start working to buy a house and this violates the exclusion restriction. Moreover, in the regression including only men the coefficient of the province-specific trends that account for different time trends within each province will not be identified. The absence of controls for province trends represents a problem in the presence of unobserved factors like policies implemented by the provincial government. Omitting province trends is also problematic if regions developing more quickly experience stronger housing booms and improvements in the quality of the education system simultaneously. In order to address endogeneity accounting for province-specific time trends, I differentiate equation (4) for males with respect to females for each province and time:

[6] Kloek (1981) and Moulton (1990) explain how the estimates of effects of aggregate variables on microunits improve by aggregating the estimated equation.

$$\overline{Y}_{Mpt}\text{-}\overline{Y}_{Fpt} = \gamma_3 + \gamma_1 \left(\frac{W_{EMpt}}{W_{UMpt}} - \frac{W_{EFpt}}{W_{UFpt}} \right) + \gamma_2 (\overline{X}_{Mpt}\text{-}\overline{X}_{Fpt}) \quad (5)$$
$$+\gamma_4' V_p + \gamma_5' W_t + \overline{v}_{Mpt}\text{-}\overline{v}_{Fpt}$$

where subindexes M and F denote male and female groups, respectively. This formulation of the model eliminates many possible sources of bias in the coefficient γ_1 including all unobserved variables at the province, time and province-time level that are common to males and females. However, the coefficient could still be biased if some of the unobserved variables have different impact on males compared to females (for instance, gender specific policies). Moreover, if the labor market is segmented by gender, issues of reverse causality could still be present because changes in the supply of uneducated men (women) could affect the returns of education of men (women). In order to address these concerns, I estimate the model in equation (5) using the level of construction activity as an instrument for differences in returns to education between genders. In this case, changes in the availability and/or affordability of housing do not cause a problem for the validity of the instrument because they affect both men and women and hence their impact cancels out in the regression in gender differences.[7]

3.3 Grade Completion

I also estimate the impact of returns to education on grade completion. The available data is aggregated by gender, province, year and grade. Ideally, I would like to estimate the following model:

[7] In order to validate the instrument I add total value added to the list of regressors in equation (5). This is useful to make the instrument fulfill the exclusion restriction, i.e., construction activity must have an impact on enrollment only through returns to skill and not through the level of economic activity.

$$
\begin{aligned}
G_{gptc} &= \delta_0 + \delta_1 \frac{W_{Egpt}}{W_{Ugpt}} + \delta_2 U_g + \delta_3 V_p \cdot U_g + \delta_4 W_t \cdot U_g + \delta_5 Z_{pt} \cdot U_g \\
&\quad + \delta_6 G_c \cdot U_g + \delta_7 K_{ct} \cdot U_g + w_{cgpt}
\end{aligned} \tag{6}
$$

where G_{gptc} is the share of students of gender g in province p and year t who complete grade c, G_c are grade dummies, K_{ct} are cohort fixed-effects, and the remaining of variables are as defined above. In the presence of province-time unobserved variables that affect both school performance and returns to education, such as changes in school quality, an OLS estimation of δ_1 could be biased. The coefficient could also be biased if expectations about the level of instruction of future cohorts of workers would affect current returns to education (reverse causality). Again, in order to address endogeneity accounting for province-time variation, I transform the original model by differentiating the model for males with respect to females:

$$
\begin{aligned}
G_{cMpt} - G_{cFpt} &= \delta_2 + \delta_1 \left(\frac{W_{EMpt}}{W_{UMpt}} - \frac{W_{EFpt}}{W_{UFpt}} \right) + \delta_3' V_p + \delta_4' W_t + \delta_5' Z_{pt} \\
&\quad + \delta_6' G_c + \delta_7' K_{ct} \cdot + w_{cMpt} - w_{cFpt}
\end{aligned} \tag{7}
$$

Reverse causality could still be present in the equation in gender differences if the school performance of men with respect to women would modify relative returns to education. Additionally, gender policies at the province level could affect school performance and returns to education simultaneously. Hence, I estimate this equation using construction over total value added as an instrument for gender differences in returns to education.

3.4 Instrument Validity

I use the level of construction activity defined as construction over total value added as an instrument for differences in returns to education between genders. The validity of this

instrument relies on: (i) its relevance to explain changes in returns to education and (ii) the exogeneity of construction activity with respect to differences in educational outcomes between genders.

Construction activity is highly correlated with differences in returns to education between genders because: (i) the construction sector employs mostly uneducated men and (ii) the housing boom implied sharp increases in wages of construction workers.[8] This translated into reductions of male returns to education and consequent reductions in the differential in returns to education between males and females. Moreover, differences in the strength of the housing boom across Spanish provinces provide the necessary variation that translates into high correlation between gender differences in returns to education and construction activity.

The second requirement for the IV estimation to be valid is the exogeneity of the housing boom with respect to differences in education between genders. The level of construction activity during the housing boom is arguably exogenous with respect to differences in educational outcomes across genders once general economic conditions, time trends and province characteristics are controlled for.

The housing boom would have been endogenous to gender differences in educational outcomes if a rise in the supply of uneducated male workers and the consequent reduction in wages would have reduced construction labor costs and hence fostered housing investment. However, the housing boom was characterized by significant increases in construction wages which suggests that the sequence of causal events that would have led to endogeneity did not occur. Still, one may be concerned that the strength of the housing boom at the province level may be partially driven by differences in educational outcomes between men and women or unobserved factors correlated with these differences. In order to address this, I regress construction over total value added at time t+1 on differences in educational outcomes between men and women at time t. In these regressions I in-

[8] The average wage of the construction sector was below the national average in 1997 but it exceeded the national average by 2006. This becomes more impressive if one considers that construction workers are mostly uneducated individuals.

clude the same controls as in the equations in sections 3.2 and 3.3 and find that all the p-values associated to the coefficients for differences in educational outcomes exceed 0.45. The coefficients of gender differences in educational outcomes are positive. This indicates that, if anything, rises in male to female educational outcomes are related to increases in construction activity in the following period. This rules out the possibility that increases in the supply of uneducated male workers fostered the housing boom. It also constitutes evidence against the existence of unobserved factors correlated with gender differences in educational outcomes that invalidate the instrument.[9]

3.5 Women as Control Group

The formulation of the empirical specification in gender differences serves two purposes: (i) It is useful to ensure the validity of construction activity as an instrument for returns to education. As discussed above, the inclusion of women as control group eliminates the concern that construction activity may have a direct impact on educational outcomes through changes in availability and/or affordability of housing. (ii) It eliminates a relevant part of the unobservable variation in educational outcomes. The reason is that taking differences of men with respect to women allows to control for changes in educational outcomes at the province level that are common to both genders. Figure 4 provides evidence in favor of the usefulness of females to control for part of the province-time variation in male educational outcomes. In this figure, I compare time trends in enrollment rates for males and females in each of the 52 Spanish provinces prior to the housing boom. Province-specific time trends for males and females present significant co-movement from one year to another. I also provide further evidence on the fact that women account for a significant part of the province-time variation in male returns to education using regression analysis. I run a regression in which male enrollment rates are a function of

[9] One may also be concerned that the housing boom has income effects such that families invest more in relative terms in the education of their daughters as they become richer. In Spain, high-school education is free and the opportunity cost of sending young family members to school is higher for sons so such gender-biased income effects should not exist.

female enrollment rates controlling for province as well as year and quarter dummies. The estimated coefficient is positive and highly significant and the R-squared is high.[10]

4 Data and Descriptive Statistics

4.1 The Spanish Education System

In the Spanish education system schooling is compulsory up to completion of the 10th grade (which cannot be done before age 16) or until the age of 16. Individuals who pass all 10 grades obtain the compulsory school diploma. In contrast, individuals who fail some grade may leave school as soon as they turn 16. However, in this case they do not obtain the compulsory school diploma. The Spanish law contemplates grade repetition in case students do not fulfill the minimum requirements for passing a grade. However, individuals are not allowed to repeat grades more than twice during compulsory education. Students who, after having repeated two grades, fail an additional grade have to leave school and pass an official test provided by the administration to obtain the compulsory school diploma. Only students that obtain the compulsory school diploma by passing all grades are entitled to continue to upper secondary education. Upper secondary education is composed of two grades.[11] Hence, conditional on successfully completing all grades, students are typically 16 years old by the time they finish compulsory education and 18 by the time they finish upper secondary education.

In the academic year 2007-2008 only 71.5% of individuals in the cohort who turned 16 during that year received the compulsory school diploma. In that same year, 44.7% of individuals who turned 18 during that year graduated from high school.

[10] The point estimate is 0.5, the standard error (clustered at the province level) is less than 0.03 and the R-squared reaches 0.74.

[11] Some vocational programs are an exception in that they are composed of only one grade.

4.2 Enrollment in Post-Compulsory Education

In the study of the impact of returns to education on the probability of enrollment in post-compulsory education I combine several databases; I use individual data from the Spanish Labor Force Survey and combine them with wage data from the Social Security Records and with data about construction activity at the province level from the Regional Economic Accounts:

- The Labor Force Survey includes a representative sample of all households living in the Spanish territory. The survey has been collected quarterly since 1976 and it provides information about demographic characteristics, labor market status and education for each member of the surveyed households. Unfortunately, the survey does not contain information about wages which is required to compute returns to education.

- The Social Security Records provide data on wages. I use the 2007 wave that contains information on a random sample including 4% of individuals who were registered as employers, employees or unemployment benefit recipients by the Social Security administration in 2007. The information consists in individual monthly earnings since the moment each person was first registered as employer or employee. I use this data to construct average wages in each gender, education, province and time cell and to compute returns to education.

- The Spanish National Statistics Institute provides Regional Economic Accounts data which are used to compute construction over total value added in each province.

The sample included in the estimation of the probability of enrollment in post-compulsory education is composed by individuals between 16 and 18 years of age. I include individuals in this age range for two reasons: First those are the ages in which individuals who successfully complete all grades are enrolled in upper secondary education. Second, 18 is the maximum age for individuals who repeat grades to obtain the compulsory school diploma.

The included time period spans from 2001 to 2006. I restrict to this time period for two reasons. First, because the data on wages is representative at the gender-province-year-quarter cell only after 2000 due to the fact that the Social Security Records only contain wages of individuals who are still active in 2007. Second, in this period the housing boom was sufficiently strong and arguably exogenous to educational outcomes.

I define individuals enrolled in post-compulsory education as those who have attended school during the four weeks preceding the interview or those who are formally enrolled but are on holidays. I compute returns to education as the ratio of wages of educated to uneducated workers where educated workers have some post-compulsory education and uneducated workers have at most compulsory education.[12]

The Spanish Labor Force Survey only provides information about individuals who are part of the selected households. This implies that the information on parental education is missing for parents who are not part of the household. I have included indicators for absent father, absent mother and both parents absent to account for missing information in the variables for parental education.

Table 1 shows descriptive statistics for the sample included in the estimation for the probability of enrollment in post-compulsory education. The share of individuals who stay in school among 16-18 year-olds is approximately 80%. Average returns to education defined by the ratio of wages of educated to uneducated workers are approximately 2. Construction represents around 11% of total value added. Regarding individual traits, we observe that approximately 4% of surveyed individuals are immigrants, only 2% of surveyed individuals do not live with their mothers, and 21% of surveyed individuals do not live with their fathers.

[12] The classification of educational level in the Social Security Registry is a mixture of occupation and education. I follow the interpretation of the classification in Lacuesta, Puente and Villanueva (2012) to identify the level of education of individuals.

4.3 Grade Completion

In the study of the impact of returns to education on grade completion, I use data provided by the Spanish Ministry of Education. This data includes information about the number of students and the number of grade completers classified by gender, province, year and grade. I then merge this information with wage data from the Social Security records and with data on construction over total value added from the Regional Economic Accounts.

For the model that estimates the effect of returns to education on grade completion I focus on the last grade of compulsory education and the first grade of post-compulsory education. I include those grades because individuals who repeat those courses are between 16 and 18 years of age. The academic years included are 2000-2001 to 2006-2007. This sample covers the same age range and time frame as in the estimation of enrollment rates.

I define grade completers as individuals who move to a higher grade or become high-school graduates in the following period. The dependent variable is then constructed as the proportion of grade completers out of total number of students.

Table 2 shows descriptive statistics for the sample included in the estimation of the impact of returns to education on grade completion. The share of grade completers is 87%. The average values for returns to education and construction activity are statistically indistinguishable from those in the sample for the estimation of the probability of enrollment.

5 Results

5.1 The Impact of the Housing Boom on Enrollment and Grade Completion

Table 3 reports the results of the impact of the housing boom on educational outcomes of men with respect to women. All coefficients of the interaction of construction activity

with a male dummy are negative and highly significant. The magnitude of the estimated coefficients increases significantly when I add individual characteristics to the regression.

The negative sign of the estimated coefficients indicates that the housing boom brought about reductions of the difference between male and female school enrollment and that the housing boom also led to reductions in the difference between male and female grade completion. This indicates that a rise in returns to education causes increases in both school enrollment and grade completion.

5.2 The Impact of the Housing Boom on Returns to Education

Table 4 shows the coefficients from the first stage regressions corresponding to the IV estimations of enrollment and grade completion described in Sections 3.2 and 3.3. The magnitude of the coefficients is consistent between the specifications for enrollment and for grade completion. The standard tests rule out that the instrument is weak in the context of our specifications.

5.3 Enrollment in Post-Compulsory Education

Table 5 shows the OLS and IV results of the estimations of the impact of returns to education on enrollment in post-compulsory education. OLS estimations produce negative estimates while IV estimations result in positive coefficients in all specifications. These positive coefficients are significant at the 5% level.

If one wrongly interprets the OLS estimates as causal effects, one could conclude that results are at odds with Becker's model. The comparison between OLS and IV estimates shows that accounting for endogeneity is crucial to obtain reliable estimates of the causal effect in this framework. The IV coefficients are statistically indistinguishable from each other in all three specifications presented in Table 5. The estimated coefficient in the complete specification is 0.085. This can be interpreted as a 10% increase in returns to

education inducing an increase in enrollment in post-compulsory education of 2%.[13]

Returns to education depend not only on relative wages but also on employment probabilities of educated and uneducated individuals. In order to account for this, I construct an additional measure of returns to education adjusting wages by the unemployment rate. The adjustment consists in multiplying the wage of (un)educated individuals by one minus the unemployment rate of (un)educated individuals. Results, shown in Table 6, are highly invariant with respect to those shown in Table 5. This is not surprising given that unemployment rates were low during the period of study.

5.4 Grade Completion

Table 7 shows OLS and IV results of the estimation of the impact of returns to education on grade completion. OLS estimation provides coefficients that are statistically indistinguishable from zero. Hence, we cannot conclude anything regarding the correlation between returns to education and grade completion. Yet, the IV results show a positive causal impact of returns to education on grade completion. The coefficient of returns to education stays invariant when I add grade and cohort fixed-effects and it is equal to 0.015. This can be interpreted as a 10% increase in returns to education causing an increase in student grade completion of 0.2%.[14]

Again, I have used an alternative measure of returns to education that includes wages adjusted by the unemployment rate. Results are shown in Table 8 and are highly invariant with respect to the ones shown in Table 7.

[13] This percentage is calculated as follows: A 10% increase in returns to education in our sample is equivalent to an increase of 0.2. This number multiplied by the estimated coefficient, 0.085, results in an increase in the propensity to be enrolled in school of 0.17. This divided by the average enrollment rate, 0.797, gives 2.133%.

[14] This percentage is calculated as follows: A 10% increase in returns to education in our sample is equivalent to an increase of 0.011. This number multiplied by the estimated coefficient, 0.015, results in an increase in grade completion of 0.017 percentage points. This divided by average grade completion, 0.873, gives 0.0189%.

6 Conclusion

The extent to which young individuals respond to changes in returns to education has important policy implications for minimum wages, wage subsidies and public employment programs. I address the impact of returns to education on school enrollment and student performance. The empirical strategy has two pillars: (i) the formulation of the model in gender differences and (ii) the use of the housing boom as a source of exogenous variation in differences between male and female returns to education. This empirical strategy provides an estimate of the causal effect of returns to education on educational outcomes controlling for the existence of time-varying factors at the province level. Empirical results show that a 10% increase in returns to education as measured by the ratio of wages of educated to uneducated individuals implies an increase of 2% in the proportion of individuals enrolled in post-compulsory education and a 0.2% increase in grade completion among 16 to 18 year-olds. These findings corroborate the results from recent literature that returns to education have a sizeable effect on enrollment. Yet, the impact of returns to education on grade completion suggests that the consequences of changes in returns to education could extend far beyond their influence on enrollment rates.

The fact that individuals react to the current state of the labor market by changing their education substantially is highly relevant because the state of the labor market at the time of graduation has substantial and time persistent consequences for employment and wages of individuals, as shown in Oreopoulos, von Wachter and Heisz (2013) and Kahn (2010). Hence, my findings show that more individuals benefit from the positive impact of economic booms on lifetime earnings and employment probabilities because more individuals graduate during booms. Similarly, the adverse effect of recessions is mitigated because relatively less individuals graduate at bad times.

References

[1] Abramitzky, R. and V. Lavy (2011) "How Responsive is Investment in Schooling to Changes in Returns? Evidence from an Unusual Pay Reform in Israel's Kibbutzim" NBER Working Paper No. 17093.

[2] Becker, G. (1964). "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education" New York: Columbia University Press.

[3] Black, D., T. McKinnish and S. G. Sanders (2005) "Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust" Industrial and Labor Relations Review 59(1): 3-16.

[4] Caballero, R. J. and A. Krishnamurthy (2009) "Global Imbalances and Financial Fragility" American Economic Review Papers and Proceedings, Vol 99, No. 2, May 2009, pp. 584-588.

[5] Clark, D. (2011) "Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrollment in Post-Compulsory Education in England" Economica, Vol. 78 (311), pp. 523-545, 07.

[6] Dellas, H. and P. Sakellaris (2003) "On the Cyclicality of Schooling: Theory and Evidence" Oxford Economic Papers 55: 148-172.

[7] Duncan, B. (1965) "Dropouts and the Unemployed" Journal of Political Economy, 73, 121.

[8] Falk, A. and A. Ichino (2006) "Clean Evidence on Peer Effects" Journal of Labor Economics , Vol. 24, No. 1 (January 2006), pp. 39-57

[9] Goldin, C., and L. Katz (1997) "Why the United States Led in Education: Lessons from Secondary School Expansion" NBER Working Paper No. 6144.

[10] Gonzalez, L. and F. Ortega (2013) "Immigration And Housing Booms: Evidence From Spain," Journal of Regional Science, Wiley Blackwell, vol. 53(1), pages 37-59, 02.

[11] Kahn, L. (2010) "The Long-Term Labor Market Consequences of Graduating College in a Bad Economy" Labour Economics 17(2): 303-316.

[12] Kloek, T, (1981) "OLS Estimation in a Model Where a Microvariable Is Explained by Aggregates and Contemporaneous Disturbances Are Equicorrelated," Econometrica, Econometric Society, vol. 49(1), pages 205-07, January.

[13] Lacuesta, A., S. Puente and E. Villanueva (2012) "The Schooling Response to a Sustained Increase in Low-Skill Wages: Evidence from Spain 1989-2009" Bank of Spain Working Paper No. 1208.

[14] Lopez-Mayan, C. (2010) "Demand for Post-compulsory Education: The Choice Between Academic and Vocational Tracks", Mimeo, CEMFI.

[15] Moulton, B. R., (1990) "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit," The Review of Economics and Statistics, MIT Press, vol. 72(2), pages 334-38, May.

[16] Neumark, D. and W. Wascher (1995) "Minimum Wage Effects on Employment and School Enrollment" Journal of Business and Economics Statistics 13: 199-206.

[17] Oreopoulos, P., A. Heisz and T. von Wachter (2013) "Short- and Long-Term Career Effects of Graduating in a Recession" American Economic Journal: Applied Economics.

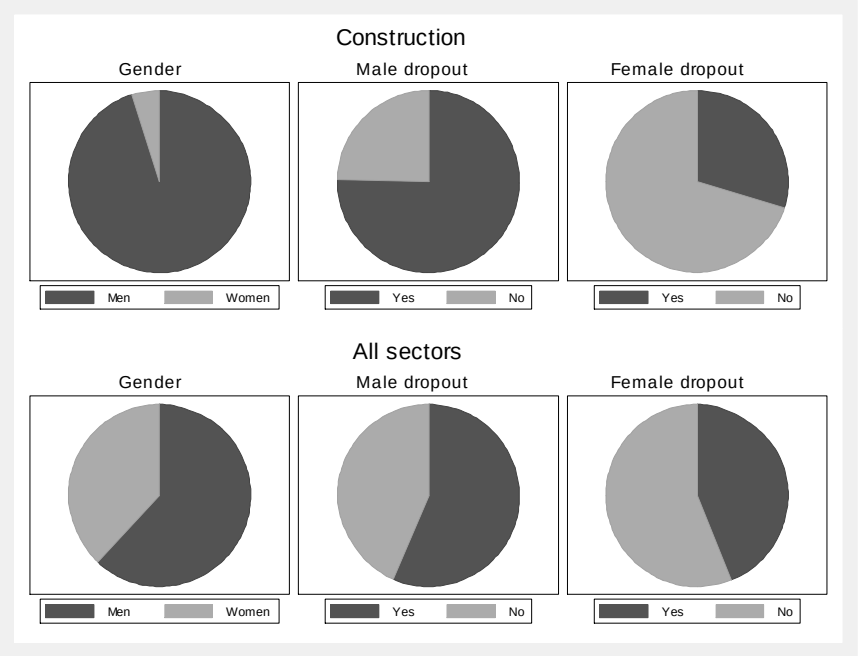
[18] Petrongolo, B., and M. J. San Segundo (2002) "Staying-on at School at 16: The Impact of Labor Market Conditions in Spain" Economics of Education Review, 21 (4), 353-365.

[19] Rajan, R. (2010) "Fault Lines: How Hidden Fractures Still Threaten the World Economy" Princeton: Princeton University Press.

[20] Rumberger, R. (1983) "Dropping out of High School: The Influence of Race, Sex and Family Background" American Educational Research Journal, Vol. 20, No. 2, pp. 199-220.

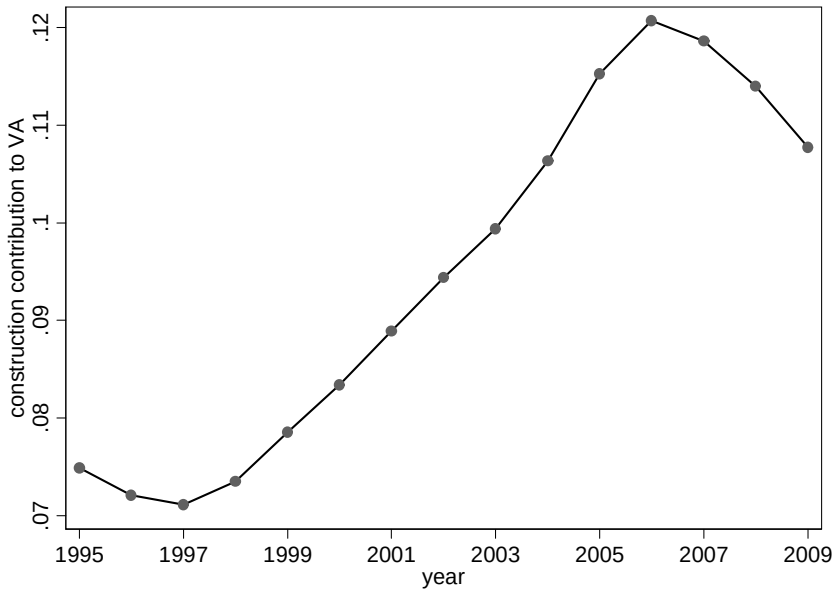
[21] Zimmerman, D. J. (2003) "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment" The Review of Economics and Statistics, February 2003, Vol. 85, No. 1, Pages 9-23

Figure 1: Composition of the workforce



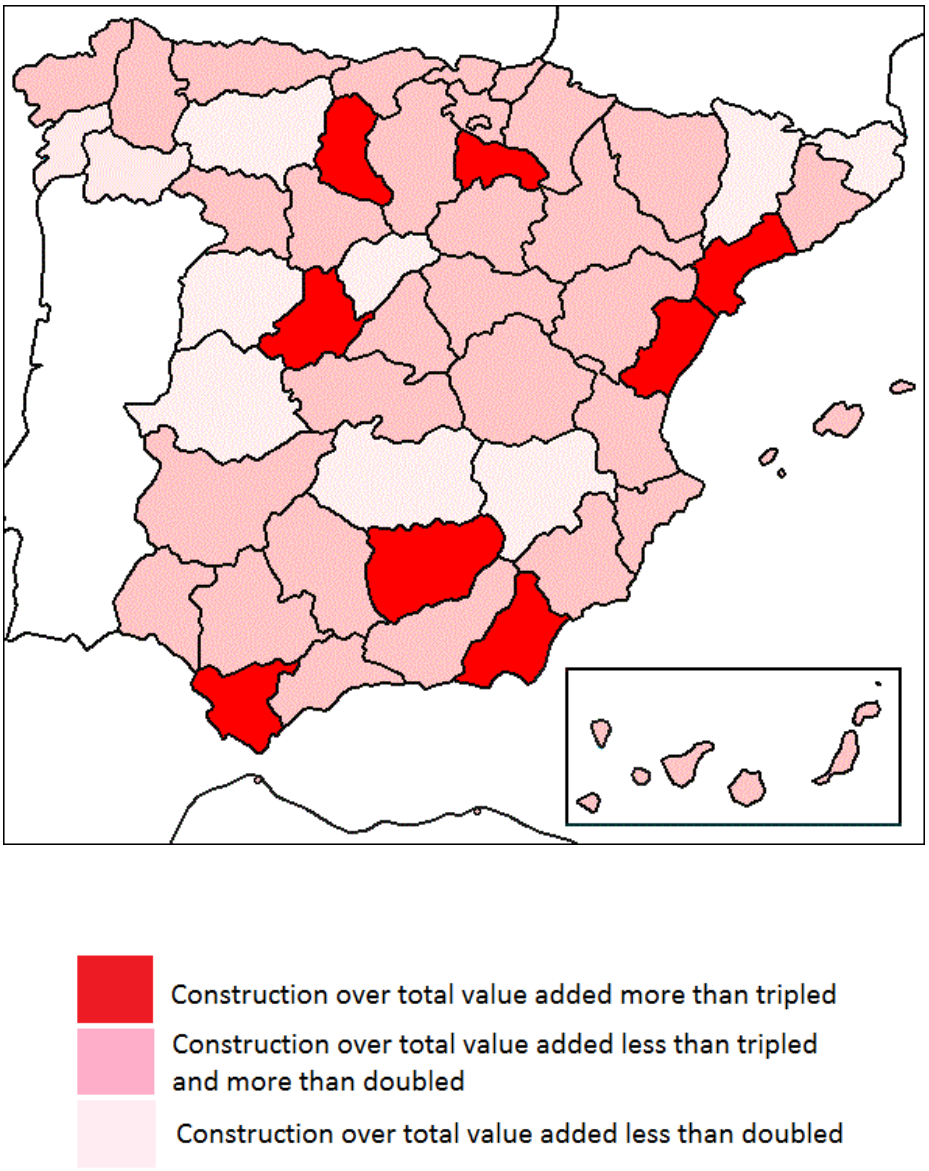
The Gender pie charts show the gender composition of the workforce. The Male (Female) pie charts show the ratio of dropouts over total male (female) workers. I consider as dropouts individuals that have only the compulsory school diploma or no diploma. Data source: Spanish Labor Force Survey for the period 1997-2006

Figure 2: The housing boom over time in Spain



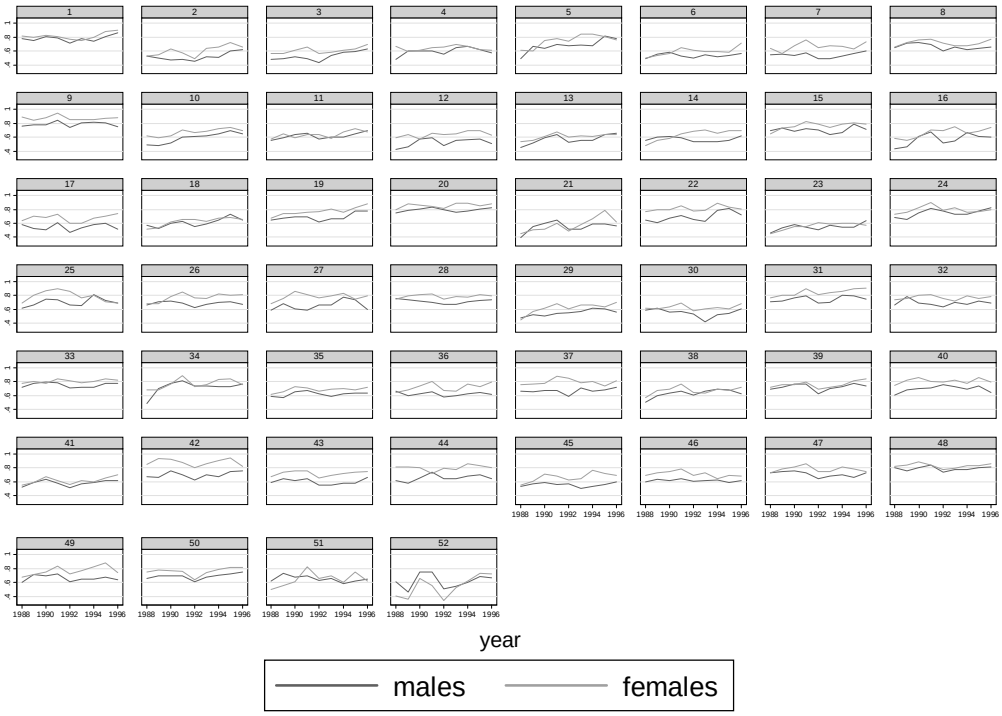
Data source: Spanish Regional Accounts provided by the Spanish National Statistics Institute and available at http://www.ine.es

Figure 3: The housing boom across regions in Spain (1997-2006)



Data source: Spanish Regional Accounts provided by the Spanish National Statistics Institute and available at http://www.ine.es

Figure 4: Enrollment rates by gender in each province



Data Source: Spanish Labor Force Survey for the period 1988-1996.

Table 1: Descriptive statistics. Enrollment in post-compulsory education

Variable	Mean	Std. Dev.	Min.	Max.
enrollment	0.765	0.424	0	1
wage educated	5325.425	1358.714	2064.371	19356.123
wage uneducated	2985.921	1108.4	168.996	11673.409
construction activity	0.112	0.024	0.059	0.189
total value added	21.964	31.04	1315870	155613526
age	16.89	0.786	16	18
immigrant	0.04	0.195	0	1
father high-school grad	0.159	0.366	0	1
mother high-school grad	0.196	0.397	0	1
father university grad	0.096	0.295	0	1
mother university grad	0.115	0.319	0	1
absent father	0.208	0.406	0	1
absent mother	0.02	0.139	0	1
absent parents	0.006	0.075	0	1
brothers	0.385	0.626	0	5
sisters	0.338	0.582	0	6
siblings aged 0 to 10	0	0	0	0
siblings aged 11 to 15	0	0	0	0
siblings aged 16 to 20	0.299	0.515	0	5
siblings aged 21 to 30	0.41	0.677	0	7
siblings aged more than 31	0.014	0.13	0	3
province	25.63	14.199	1	50
year	2003.37	1.688	2001	2006
quarter	2.486	1.117	1	4

The number of observations is 117,412. All data are from the Spanish Labor Survey, except for wages (from Social Security Records), and value added and construction activity (from Regional Accounts). The sample includes individuals between 16 and 18 years of age in the period 2001-2006. Value added is expressed in millions of euros.

Table 2: Descriptive statistics. Grade completion

Variable	Mean	Std. Dev.	Min.	Max.
completers	0.873	0.067	0	0.341
wage educated	5298.904	1560.337	2064.371	19356.123
wage uneducated	2995.272	1343.39	168.996	16972.988
construction activity	0.113	0.026	0.053	0.202
total value added	14.818	22.828	1.279	167.362
grade	4.5	0.5	4	5
male	0.5	0.5	0	1
province	25.5	14.435	1	50
year	2003.5	2.292	2000	2007

The number of observations is 1,600. The data on grade completion is from the Spanish Ministry of Education. Wages are from Social Security Records. Value added and construction activity are from Regional Economic Accounts. The information is aggregated at the gender, province, year and grade level. The sample includes the last grade of compulsory education and the first grade of high school in the academic years from 2000-2001 to 2006-2007. Value added is expressed in millions of euros.

Table 3: The impact of the housing boom on educational outcomes

	Enrollment			Completers	
	(1)	(2)	(3)	(4)	(5)
construction activity by male	-0.594	-1.212	-1.116	-0.331	-0.331
	(0.128)***	(0.498)**	(0.473)**	(0.093)***	(0.094)***
construction activity	0.478	0.831	0.769	-0.962	-0.962
	(0.257)*	(0.336)**	(0.317)**	(0.149)***	(0.15)***
male	-0.032	0.048	0.052	-0.014	-0.04
	(0.015)**	(0.046)	(0.044)	(0.011)	(0.009)***
Individual characteristics	No	Yes	Yes	No	No
Interactions w/gender	No	No	Yes	No	No
Cohort fixed-effects	No	No	No	No	Yes
No. of observations	117,412	117,412	117,412	1,662	1,662
R^2	0.037	0.04	0.127	0.319	0.357

Enrollment equals one if the individual is enrolled in school and zero otherwise. Completers is the fraction of grade completers in each grade-province-year cell. Construction activity is measured by construction over total value added. Standard errors are clustered by province. The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%.

Table 4: The impact of the housing boom on returns to education

	enrolled			completers	
Dep var: Returns	(1)	(2)	(3)	(4)	(5)
Construction activity	-14.198	-15.09	-15.887	-13.841	-13.841
	(3.99)***	(3.969)***	(3.996)***	(2.724)***	(2.724)***
Individual characteristics	No	Yes	Yes	No	No
Interactions w/ gender	No	No	Yes	No	Yes
No. of observations	1,112	1,112	1,112	738	738

All variables are constructed as the difference between the average value of the variable for males and the average value of the variable for females. The dependent variable is returns to education which is defined as the ratio of wages of educated to uneducated workers. Construction activity is measured by construction over total value added. The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%.

Table 5: The impact of returns to education on enrollment in post-compulsory education

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Returns to education	-0.007	-0.006	-0.006	0.098	0.107	0.085
	(0.004)*	(0.004)*	(0.004)*	(0.049)**	(0.045)**	(0.039)**
Individual characteristics	No	Yes	Yes	No	Yes	Yes
Interactions w/gender	No	No	Yes	No	No	Yes
No. of observations	1,112	1,112	1,112	1,112	1,112	1,112

All variables are constructed as the difference between the average value of the variable for males and the average value of the variable for females. The dependent variable is the proportion of individuals enrolled in school. Returns to education are defined as the ratio of wages of educated to uneducated workers. In the IV specification, returns to education are instrumented using construction over total value added. The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%.

Table 6: The impact of UR-adjusted returns to education on enrollment in post-compulsory education

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
UR-adjusted returns	-0.007	-0.007	-0.006	0.108	0.116	0.091
	(0.004)*	(0.004)*	(0.004)*	(0.055)**	(0.05)**	(0.043)**
Individual characteristics	No	Yes	Yes	No	Yes	Yes
Interactions w/gender	No	No	Yes	No	No	Yes
No. of observations	1,112	1,112	1,112	1,112	1,112	1,112

All variables are constructed as the difference between the average value of the variable for males and the average value of the variable for females. The dependent variable is the proportion of individuals enrolled in school. Returns to education are defined as the ratio of wages of educated to uneducated workers where wages are adjusted by the unemployment rate. In the IV specification, returns to education are instrumented using construction over total value added. The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%.

Table 7: The impact of returns to education on grade completion

	OLS		IV	
	(1)	(2)	(3)	(4)
Returns to education	0.0004	0.0004	0.015	0.015
	(0.001)	(0.001)	(0.008)**	(0.006)**
Interactions w/ gender	No	Yes	No	Yes
No. of observations	738	738	738	738

All variables are constructed as the difference between the average value for males and the average value for females. The dependent variable is the proportion of grade completers out of total students the year following the survey. Returns to education are defined as the ratio of wages of educated to uneducated workers. In the IV specification, returns to education are instrumented using construction over total value added. The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%.

Table 8: The impact of UR-adjusted returns to education on grade completion

	OLS		IV	
	(1)	(2)	(3)	(4)
UR-adjusted returns	0.0005	0.0005	0.015	0.015
	(0.001)	(0.001)	(0.008)**	(0.006)**
Interactions w/ gender	No	Yes	No	Yes
No. of observations	738	738	738	738

All variables are constructed as the difference between the value for males and the value for females. The dependent variable is the proportion of grade completers out of total students the year following the survey. Returns to education are defined as the ratio of wages of educated to uneducated workers where wages are adjusted by the unemployment rate. In the IV specification, returns to education are instrumented using construction over total value added. The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%.