Economics of Education Review 36 (2013) 198–215

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

The Returns to Private Education: Evidence from Mexico

Chiara Binelli [a,b,*], Marta Rubio-Codina [c]

[a] The University of Southampton, Department of Economics, United Kingdom
[b] Rimini Center for Economic Analysis, Italy
[c] Centre for the Evaluation of Development Policies, Institute for Fiscal Studies, United Kingdom

ARTICLE INFO

Article history:
Received 20 August 2012
Received in revised form 8 March 2013
Accepted 18 June 2013

JEL classification:
I20
I24
J31
J24
C36

Keywords:
Private and Public Schools
Returns to Education
Mexico

ABSTRACT

Despite the rapid expansion and increasing importance of private education in developing countries, little is known on the impact of studying in private schools on education and wages. This paper contributes to filling this gap by estimating the returns to private high schools in Mexico. We construct a unique data set that combines labour market outcomes and historical census data, and we exploit changes in the availability and size of public and private high schools across states and over time for identification. We find that attending a private high school does not affect school progression to college nor high school wages but it does positively affect wages conditional on college completion. Results are robust to a number of robustness tests on the validity of the instruments.

1. Introduction

In recent years there has been considerable debate over the privatization of the education sector in low and middle income countries where private education has fast expanded. Several international organizations have been involved in this expansion with the International Financial Corporation (IFC) being the largest multilateral investor: as

argue that the private sector can be used as a means of increasing access to education via efficient supply: attending a private school has been associated with better test score results, increased school attainment and higher wages (e.g. Riddell, 1993). In high income countries, an extensive literature has estimated the effect of attending private schools on education and wages using a number of identification strategies to control for selection bias. Brown and Belfield (2001) review the studies for the US and the

te schools ranging

tween 10% and 23%

ducation and on the sample considered.

On the contrary, the evidence for low and middle income countries remains mainly descriptive. Largely due to data limitations, most studies simply compare differences in mean wages and school attainment between private and public schools' students without controlling for self-selection into the school of choice (e.g. Chudgar & Quin, 2012 for India; Asadullah, 2009 for Bangladesh and

* Corresponding author at: Economics Division, School of Social Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom. Tel.: +44 23 8059 8651; fax: +44 23 8059 3858.
E-mail address: chiara.binelli@gmail.com (C. Binelli).

[1] www.ifc.org.

0272-7757
http://dx.doi.org/10.1016/j.econedurev.2013.06.004

Pakistan; Calónico & Ñopo for Peru; Bedi & Garg, 2000 for Indonesia). A noticeable exception is the paper by Bravo, Mukhopadhyay, and Todd (2008), which uses data from a school voucher program introduced in Chile in 1980 to identify a dynamic structural model of private school attendance and wages. The authors find that the voucher program increased enrolment in private subsidized schools, which subsequently had a positive significant effect on labour force participation and wages.[2]

Thus, despite the important related policy implications, little is known on the relative efficiency of private and public schools in improving educational attainment and wages in low and middle income countries. This paper contributes to filling this gap by estimating the returns to private high schools in Mexico. We construct a unique dataset that combines individual-level data on school choices and wages with state-level school census data since the late 1980s. We first describe the evolution in the availability and size of public and private high schools across states in Mexico during the 1990s, and we show that the public high school sector expanded much faster than the private sector. We then exploit these changes in relative availability and size of public relative to private high schools to identify the effect of studying in a private high school on school attainment and on wages. More specifically, we instrument the choice of attending a private high school with measures of the relative availability and relative size of public high schools in the state and year when the high school choice was made for a sample of workers aged 23–35 in 2008.

We find that attending a private high school does not affect school progression from high school to college nor wages after high school, while it does increase wages upon college completion. These results are robust to a number of validity checks on the strength and exogeneity of the instruments, including controlling for measures of education quality and other omitted factors that may affect both school choice and wages, potential demand-driven changes in high schools' availability, and between-state migration. Even if this evidence is reassuring, we apply the method recently developed by Nevo and Rosen (2008) to further assess the robustness of our findings to weaker identification assumptions. In particular, we explicitly allow the instruments to be correlated with the error term in the outcome equation and estimate bounds of the returns to private high schools. When we relax the assumption that the instruments are exogenous, we find that the returns range between 8% and 58%. This range of estimates is too wide to be informative on the point estimate of the returns. However, the lowest point estimate of 8% is higher than the OLS returns, estimated at 5%, which, alongside the other checks on the validity of the instruments, suggests that there are robust positive wage returns to private high school for those that complete college education.

This paper contributes to the existing literature by establishing a positive causal relationship between attending a private high school and wages in a middle income country. Two main limitations are worth bearing in mind when interpreting the findings. First, we are unable to estimate the exact magnitude of the wage returns because our estimates are likely to partly capture the impact of unobserved factors, such as peer effects and family background variables, which we cannot control for with the available data. Second, our estimates must be thought of as the short-term returns to private high school after college graduation since we work with a sample of young workers. Returns to private high school could reduce in the long-run if, over time, job experience becomes a more important determinant of earnings than the type of school attended.

The remainder of the paper is organized as follows. Section 2 describes the private and public high school sectors in Mexico and their evolution between 1970 and 2000. Section 3 outlines the empirical framework. Section 4 describes the data and presents summary statistics. In Section 5 we present the main results, and in Section 6 we discuss the validity of the instruments. Section 7 concludes.

2. Public and private high schools in Mexico

The Mexican education system is one of the largest in Latin America covering 33.3 million students in 2008, or 31.5% of the country's population (SEP, 2008a, 2008b). There are twelve years of formal education prior to college: six years of primary, three years of secondary, and two or three years of high school. Three main types of high school education are offered: (i) *bachillerato general*, which leads students on an academic track in preparation for college; (ii) *bachillerato tecnológico*, which has a more technical focus and prepares students for either vocational work or higher education to become qualified technicians in specific areas; and (iii) *profesional técnico*, which is a two-year program designed for students seeking a more technical or vocational training that does not allow continuation onto college.

High school education is offered by a mixture of public and private institutions.[3] Enrolment in both public and private high schools is generally open up to capacity and only in the event of excess demand schools administer an entrance exam.[4] However, there are monetary barriers to entrance, especially in the private sector where schools are primarily financed via tuition fees. Public high schools are free of charge and fully funded by the federal, state, or municipal governments. Even then, students are often

[2] Also for Chile, Anand, Mizala, and Repetto (2009) find no difference between the academic achievement of students in the fee-charging private voucher treatment group relative to their counterparts in free private voucher schools, while they find that students in fee-charging private voucher schools have higher scores than students in public schools.

[3] Private schools are legally established after having obtained a license granted by the federal or state government, the *Reconocimiento de Validez Oficial* (RVOE), which guarantees that basic standards are met. Once RVOE has been granted, there is no further regulation as to the quality or type of academic programs offered by private institutions.

[4] An exception are public high schools in the metropolitan area of Mexico City, which recruit students through public competition (http://www.comipems.org.mx).

encouraged to give a voluntary contribution, of varying amount depending on school needs, the poverty level of the area and the administrative authority. In addition, students in both private and public high schools have to cover exam fees, schooling materials, transport and other living costs. Using data from the Mexican National Consumption and Expenditure Survey (*Encuesta Nacional de Ingresos y Gastos de los Hogares* or ENIGH) for 2002, we estimate that the average direct costs of studying (fees, books, exams, etc.) in a public high school represents about 15% of median yearly household income, whereas the average direct cost of attending a private high school amounts to approximately 23%.

A limited number of scholarships help students meet these costs. They are mainly offered in the context of *Oportunidades*, the largest anti-poverty program in the country, which provides cash transfers to poor families conditional on primary, secondary and, starting in 2001, high school attendance. Since 2003, *Jóvenes con Oportunidades*, additionally provides monetary incentives to complete high school.[5] Student loans are also limited in scope and coverage and only finance college education.[6]

In 2008 the enrolment rate at high school was 61%, of which about 19% was in private schools. Of those enroled, 61% were enroled in *bachillerato general* (25% in private schools), 29% in *bachillerato tecnológico* (11% in private schools) and 10% in *profesional técnico* (17% in private schools) (SEP, 2008a, 2008b and own calculations).

2.1. Trends in the provision of high school education

We next document the evolution in the provision of high school education by the public and the private sector using data from the Mexican School Census (*Censo Escolar* or *Estadística 911*), which is collected yearly by the Secretariat of Public Education (*Secretaría de Educación Pública* or SEP). The school census contains the number of schools, teachers, students and classes for each of the 32 Mexican states by education level and by type of school (public/private). We have combined paper records from 1970 to 1989 and online records since 1990[7] to construct a dataset with the number of schools, teachers, students and classes by year and by state in private and public high schools from 1970 to 2000. Figs. 1–3 plot average trends in the number of schools and students by sector over this period.

As shown in Fig. 1, the public sector has expanded faster than the private sector. After a period of sustained growth, the number of public high schools over the 16–18 age population group[8] increased dramatically in the mid-1980s to resume a more moderate growth path thereafter. This was the result of a long term change in educational policies that started in the 1950s with the expansion of primary and secondary education, and accelerated in the 1970s as a result of a government strategy to promote educational attainment beyond compulsory education (Gómez, 1999). Moreover, following recommendations by international organizations

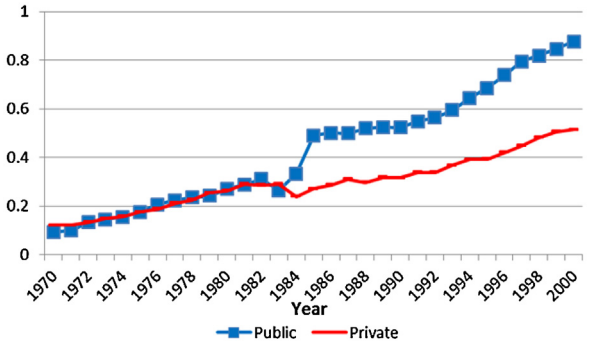

Fig. 1. Per capita number of public and private high schools over the 16–18 age population.
Source: authors' calculations based on data from the Mexican School Census.

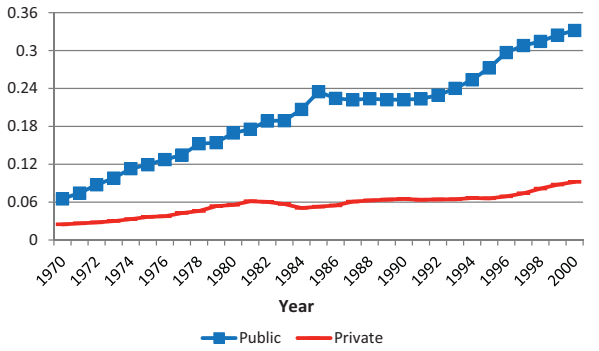

Fig. 2. Enrolment rate in public and private high schools.
Source: authors' calculations based on data from the Mexican School Census.

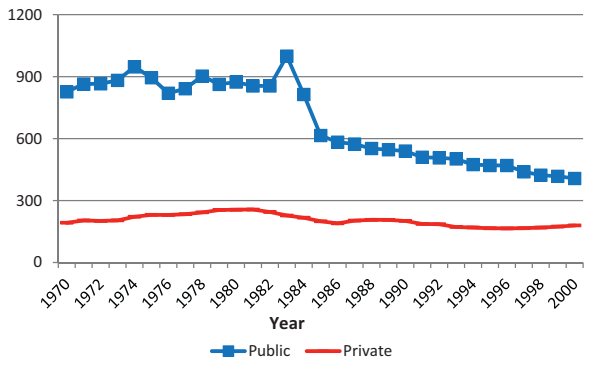

Fig. 3. Number of students per public and private high school.
Source: authors' calculations based on data from the Mexican School Census.

[5] www.oportunidades.gob.mx/informacion_general/main.html. Note that either one of these schemes affected the schooling choices of the individuals in our sample since all individuals started high school before 2001 (See Section 4 for details).

[6] In 2007 only 2% of Mexican students benefited from a student loan. This is a very small proportion even relative to other Latin American countries such as Colombia and Brazil, where, respectively, 9% and 6% of college students have a loan (*Educafin*, 2007).

[7] These are available at http://www.sep.gob.mx.

[8] Population data obtained from the National Population Council (*Consejo Nacional de Población* or CONAPO): http://www.conapo.gob.mx.

(UNESCO, IDB, World Bank, etc.), Mexico shifted public investment in education to those levels offering the highest rates of return—i.e. primary education and technological high schools—as a means to face the economic crisis of the early 1980s. During the 1990s, the public sector continued to expand faster than the private sector, especially after 1993, thus remaining the main high school education provider.

The increase in the number of high schools resulted in increasing enrolment rates (number of entrants over the 16–18 age population group) in both public and private schools (see Fig. 2). From 1970 to the mid 1980s, enrolment rates in public high schools experienced a much larger increase than in private high schools (24% versus 4%). A similar albeit less marked trend is observed during the 1990s, when enrolment rates increased by 10% in public high schools and 3% in private high schools. As a result, by 2000 private high school enrolment, at 9%, was substantially lower than public high school enrolment, at 33%.

Despite the increasing enrolment rates, the number of public high schools increased more than the number of students. Consequently, the size of the public sector (defined as the number of students per public high school) has decreased since the beginning of the 1980s (Fig. 3). In contrast, the number of students per private high school remained almost unchanged from 1970 to 2000.

We expect that these changes in availability and size of private and public high schools affected the decisions of those enroling in high school, particularly since the mid 1980s. As we will discuss in Section 4.2, there is also substantial variation across states. In the next sections we focus on a sample of workers aged 23–35 in 2008, who hence enroled in high school starting in the late 1980s, and exploit the reported variation over time and across states to identify the differential effect of studying in a private high school (relative to a public high school) on wages.

3. Empirical strategy

In order to quantify the wage returns to attending a private relative to a public high school, we specify the following wage equation:

$$w^s_{ija} = \beta_0 + \xi_j + \lambda_a + \delta PvHS_{ija} + X'_{ija}\beta + \varepsilon^s_{ija} \quad (1)$$

where w^s_{ija} is the logarithm of real hourly earnings for individual i aged a with education level $s = \{$high school, college$\}$ and living in state j. $PvHS_{ija}$, our main variable of interest, is an indicator variable that equals one if i studied in a private high school, and zero otherwise. X_{ija} is a matrix of observable individual characteristics that were predetermined at the time when the private/public high school choice was made and thus might have affected this choice, namely gender and a dummy accounting for whether i lives with her parents.[9] ξ_j are dummies for the state of residence in 2008, included to control for any permanent regional differences and labour market trends that could affect earnings, and λ_a are cohort dummies that account for time effects. ε^s_{ija} is the error term.

$\widehat{\delta}$ measures the estimated effect of having studied in a private relative to a public high school on wages. This estimate will be biased if students sort into private and public high schools based on factors that are not accounted for in the regression but that matter for wages. In order to correct for self-selection, we instrument $PvHS_{ija}$ with two measures of the relative availability of public (with respect to private) high schools in the state and in the year when the choice of the type of high school was made. Specifically, we use: (i) the logarithm of the share of public high schools, $SharePb_{ij\bar{a}}$; and (ii) the logarithm of the relative size of public high schools, $SizePb_{ij\bar{a}}$. Both $SharePb_{ij\bar{a}}$ and $SizePb_{ij\bar{a}}$ are measured in the state of residence j and in the year when individual i was $\bar{a} = 14$ years old, which is the age before the median high school entry age in Mexico and thus that when the decision on which high school to attend is made (see Section 4.1 for more details).

We construct *SharePb* by dividing the number of public high schools over the total (private and public) number of high schools in a given state and year. It measures the *share* or proportion of public high schools and hence it is a proxy for the *potential* availability of public relative to private high schools in the year before the individual enters high school. The *effective* relative availability of public high schools will depend both on the number of schools and on the number of vacancies in a school (i.e. on the size of the school). In the absence of data on the actual number of places available in a given school, we use the number of students enroled in a school as a proxy for school capacity. Hence, we construct *SizePb* by dividing the total number of students enroled in public high schools in a given state and year over the total number of students enroled in high schools (private and public) in that state and year, and take this as a measure of the *relative size* of the public high school sector—i.e. the higher the proportion of students in public high schools, the larger the relative size of the public high school sector.

Taken together $SharePb_{ij\bar{a}}$ and $SizePb_{ij\bar{a}}$ capture the *effective availability* of public relative to private high schools in the state and in the year when the decision to attend a private or a public high school is made, and as such should significantly affect this choice. Hence, we use two-stage least squares (2SLS) to estimate the wage Eq. (1) jointly with the following school equation:

$$PvHS_{ija} = \tilde{\beta}_0 + \tilde{\xi}_j + \tilde{\lambda}_a + X'_{ija}\tilde{\beta} + Z'_{ij\bar{a}}\gamma + \omega_{ija} \quad (2)$$

where $Z_{ij\bar{a}} \equiv \{SharePb_{ij\bar{a}}, SizePb_{ij\bar{a}}\}$ are the instruments and ω_{ija} is the error term.

3.1. Identification

By jointly estimating Eqs. (1) and (2), we identify the effect of studying in a private relative to a public high school on wages, $\widehat{\delta}$, through the variation in the relative availability and size of public high schools across age cohorts and states. The identification strategy exploits the

[9] We do not include marital status, head of household nor being employed since these characteristics might have been affected by the type of school attended. Results (available upon request) are, however, robust to the inclusion of these additional variables.

idea that the state and year of start of high school determine how much an individual has been exposed to the increase in the relative availability of public and private high schools. In other words, the choice of studying in a private or in a public high school varies depending on when and where the high school decision was made—i.e. how many years a given age cohort has been exposed to the schooling expansion and the extent to which high school availability increased in the state. The main identification assumption is therefore that in the absence of a relative increase in private high schools' availability, changes in private high schools' enrolment would not have been systematically different in states with high and low relative availability of public high schools.[10]

The validity of our findings depends on the validity of the instruments, which in turn relies on the instruments being strongly correlated with the choice of studying at a private or a public high school. In our setting, this means that both the share and the relative size of public high schools in the state and in the year before high school enrolment have to be important determinants of the individual decision to attend a private/public high school. It seems plausible to expect that if the proportion of public high schools and their relative size increase (decrease) at the state level, students are more (less) likely to enrol in a public (private) high school. One may be concerned, however, that individual schooling decisions are driven by the educational supply at a more local level such as the province or the town of residence. This is unlikely to be the case in Mexico, where within-state migration has become more and more common since 1970, especially towards medium-sized cities (CONAPO, 1999). Further, the first-stage estimates will (and do) provide a direct measure of the strong correlation between the instruments and the high school choice.

In addition, the instruments must be exogenous and satisfy the exclusion restriction. This means that our measures of effective relative availability of public high schools must be uncorrelated with the error term in the school choice Eq. (2) and can only affect wages through their impact on the decision to attend a private or a public high school. In other words, the instruments cannot be correlated with any other variable that may also affect labour market outcomes, such as aggregate growth. In Section 6 we provide direct evidence of the instruments' validity in our context and dismiss a number of potential threats. Further, in order to address any lingering concerns, we impose weaker validity assumptions on the instruments and allow for the exclusion restriction to be violated (Nevo & Rosen, 2008). This allows us to derive bounds for the effect of studying in a private high school on wages, so that we can assess how much the exogeneity assumption drives the results. Finally, the instruments must also have enough variation over time and across states for identification, which is an issue we will return to in Section 4.2.

Note that by jointly estimating (1) and (2) on the sample of workers, we obtain the "overall" return to private high school attendance, which combines the effect of studying in a private high school on college attendance, on college completion, and on wages. To identify each of these effects separately and to understand the mechanisms whereby private high school attendance affects wages, we will separately estimate the effect of studying at a private relative to a public high school on wages by the highest educational achieved—i.e. for those with completed high school education, uncompleted college and completed college. Moreover, we will also estimate the effect of private high school attendance on the probabilities of completing high school, enroling into college and completing college, and on the total number of years of completed education.

4. Data

4.1. Sources

We use two main sources of data for the empirical analysis: the National Survey on Labour and Educational Trajectories (*Encuesta Nacional de Trayectorias Educativas y Labourales* or ENTELEMS) and the Mexican School Census, which we described in Section 2.1

The ENTELEMS survey was collected by the division of SEP in charge of high school education. It was administered to all 15–35 year olds who had completed at least one year of high school education and were living in households included in the third round of the 2008 Mexican National Employment Survey (*Encuesta Nacional de Ocupación y Empleo* or ENOE).[11] If there was more than one household member satisfying these characteristics in any given household, the individual whose birth date was closer to the date of the interview was included. Overall, the ENTELEMS surveyed 34,901 individuals, or about 8.5% of the individuals surveyed in the 2008 ENOE.

The survey contains information on basic individual characteristics (sex, age, marital status, head of household and number of children—although only for women), current and previous employment status, type of employment, hours of work, wages, years of high school and college education completed, state where high school and college were attended, and, crucially for us, whether the individual attended a private or a public high school and/or college.[12] For those still living with their parents, we can use ENOE data to construct parental education and parental employment. In addition, for the sub-sample of

[10] A similar type of age cohort and regional variation identification strategy has been used by Duflo (2001).

[11] The ENOE is Mexico's main employment survey and is collected by the National Statistical Office (INEGI) every three months on a nationally representative sample of 120,260 households. It has a rotating panel structure and it includes detailed information on employment, education and household socio-demographics. The ENOE contains accurate data on wages and has been widely used for studies of the Mexican labour market (e.g. Bosch & Manacorda, 2010).

[12] There is no information on whether an individual has switched between private and public centres during high school. All questions refer to the last high school the individual attended. To the best of our knowledge, the only other available Mexican surveys that include information on the public/private type of high school attended are the Mexican Family Life Survey (MxFLS) and the ENILEMS Survey, although only for a very small number of workers.

Table 1
Descriptive statistics.

	Private high school ($N = 2327$)		Public high school ($N = 6196$)		t-Test
	Mean	SE	Mean	SE	
Age	28.896	0.074	28.487	0.045	4.737
Female = 1	0.560	0.010	0.449	0.006	9.201
Living with parents = 1	0.431	0.010	0.393	0.006	3.267
Employee = 1	0.826	0.008	0.874	0.004	−5.622
Married or in partnership = 1	0.458	0.103	0.518	0.006	−4.918
Head of household = 1	0.304	0.009	0.358	0.006	−4.710

those aged 15–29, the ENTELEMS includes a module on employment trajectories, which we will exploit in Section 6.2. We deflate wages to June 2011 using the most recent National Consumer Price Index data available at the time of analysis.

We merge the ENTELEMS and the Mexican School Census datasets by state of residence and year at age 14, which is the age before the median high school entry age in Mexico and thus the year when the decision on which high school to attend is made. 88% of the individuals in the sample reside in the state where they went to high school and 84% continue to live in the state where they were born. Hence, the choice of the relevant state—birth, high school attendance, residence in 2008—at which to merge the information is trivial, and results (available upon request) are robust to this choice. Regarding the year of entry into high school, the ENTELEMS includes a question on high school entry age. However, this question is only answered consistently—i.e. reports that are not at odds with the age and educational achievement of the respondent in 2008—for about 40% of the sample and hence we deem it unreliable. Instead, we assume that all individuals started high school at age 15, the sample and country median age of high school entry (Domínguez & Pérez Gómez, 1993), and merge the measures of school availability in the year before, at age 14. By doing so, we will obtain a lower bound estimate of the impact of school availability and size on the choice to attend a private or a public high school.[13]

For each wage earner in the sample, the final dataset contains data on wages, education choices, and state-level measures of effective public and private high school availability in the state and year when the choice to attend a private or a public high school was made.[14]

[13] We have investigated whether there are differences in the age at high school entrance by type of school attended amongst those that started high school between the age of 15 and 18. We find that private high school attendees are marginally older (0.8 months) when they start high school. This suggests that, if anything, our estimates underestimate the returns on work experience for those attending private high school since we are de facto assuming that they enroled in high school slightly earlier than they actually did.

[14] We focus on high school education because of data availability but also because private high schools represent a homogenous group to which public high schools can be meaningfully compared. On the contrary, both private and public colleges are characterized by a substantial degree of heterogeneity (there are both very prestigious private universities such as the Autonomous Technological Institute of Mexico, and public universities such as the National Autonomous University of Mexico), which makes the private–public comparison uninformative unless the actual school attended was known.

4.2. Sample and descriptive statistics

Our final sample consists of 8523 workers aged between 23 and 35 in 2008, who were not studying at the time of the survey. This implies that individuals in the sample attended high school between 1988 and 2000. We restrict the sample to workers that are 23 years old and older to maximize the likelihood that they have completed their education, possibly including college. Table 1 presents summary statistics on this sample. On average, those that have attended a private high school (27% of the sample) are older, more likely women, have a higher probability to live with their parents (and consistently, a lower probability of being heads of household), and are less likely to be working as employees (as opposed to being self-employed) and married or in partnership (as opposed to being single).

Fig. 4 shows that the proportion of individuals in the sample that attended a private or a public high school is well distributed across cohorts. We also observe lower enrolments in private high schools amongst younger cohorts: while only 22–24% of the 23–25 year old cohort studied in a private high school, about 30% of the 32–35 year old cohort did. Note that, as shown in Fig. 5, the proportion of individuals in the sample that attended a private or a public high school is less well balanced across states, which may be related to differences between states in the relative availability and size of public relative to private high schools.

Figs. 6–8 show that there is enough variability in the value of the instruments across states and over time, as required for identification. The maps in Figs. 6 and 7 illustrate the variation of our measures of relative availability and size of public high schools (*SharePb* and *SizePb*) across states in 1987 and 1999, which are the years when the oldest and the youngest cohort in the sample were 14 and thus about to start high school. As shown, both the proportion (share) of public high schools and its relative size vary substantially across states. Moreover, not all states follow a similar trend over time: while the relative availability of public high schools increased in some states, it remained almost constant or decreased in some other states.

Fig. 8 shows that these variables also vary over time: the share of the public high school sector increased from a minimum value of 0.61 in 1987 to a maximum value of 0.63 in 1999, whereas the relative size of the public high school sector (proportion of students enroled in public high schools) decreased from around 0.74 in 1987 to 0.70 in the year 1999.

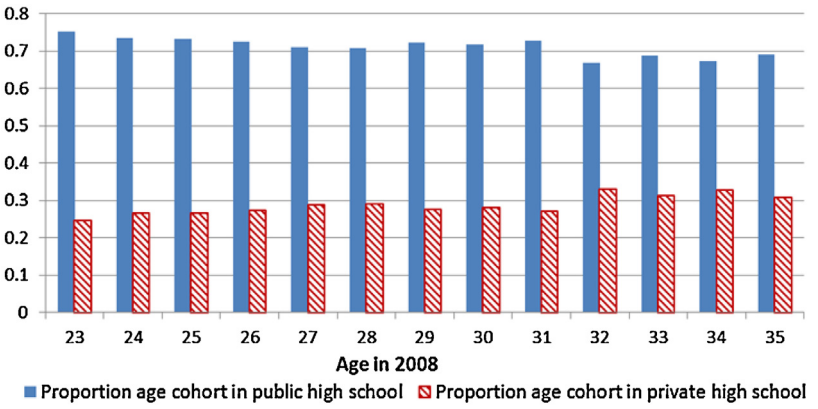

Fig. 4. Proportion of individuals in the sample that attended a public or a private high school by age cohort.
Source: authors' calculations based on the ENTELEMS survey.

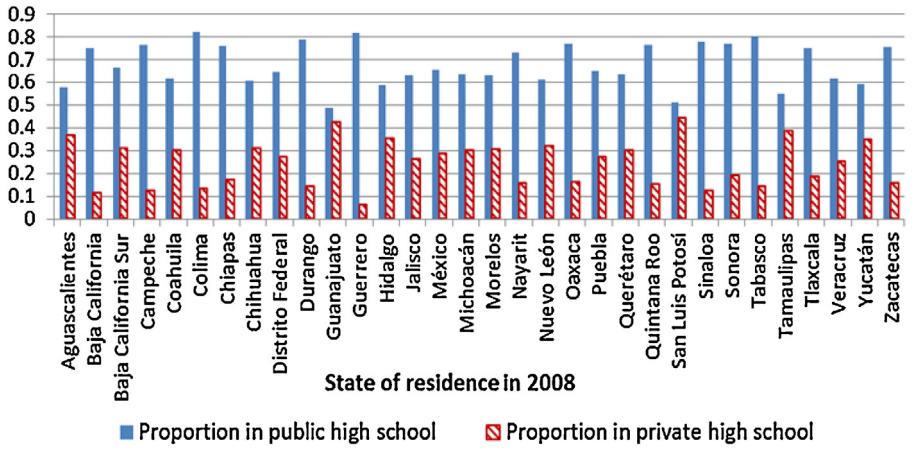

Fig. 5. Proportion of individuals in the sample that attended a public or private high school by state.
Source: authors' calculations based on the ENTELEMS Survey.

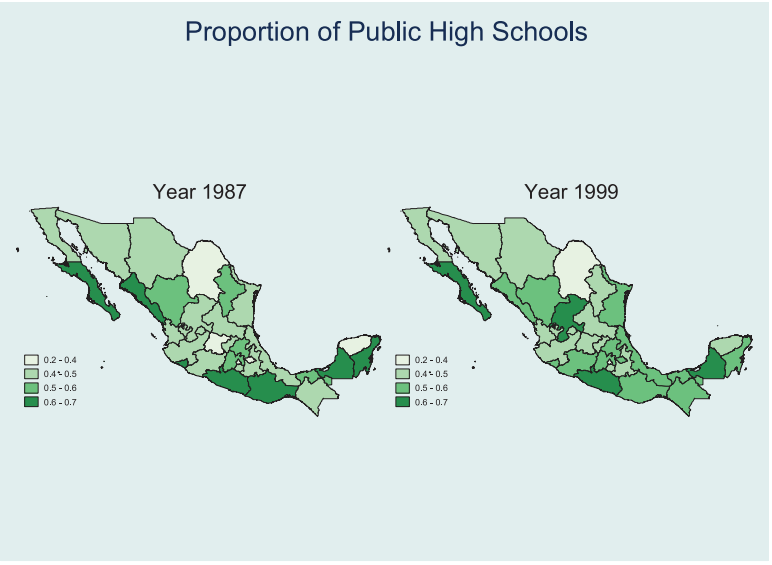

Fig. 6. Proportion of public high schools by state in year 1987 and 1999.
Source: authors' calculations based on data from the Mexican School Census.

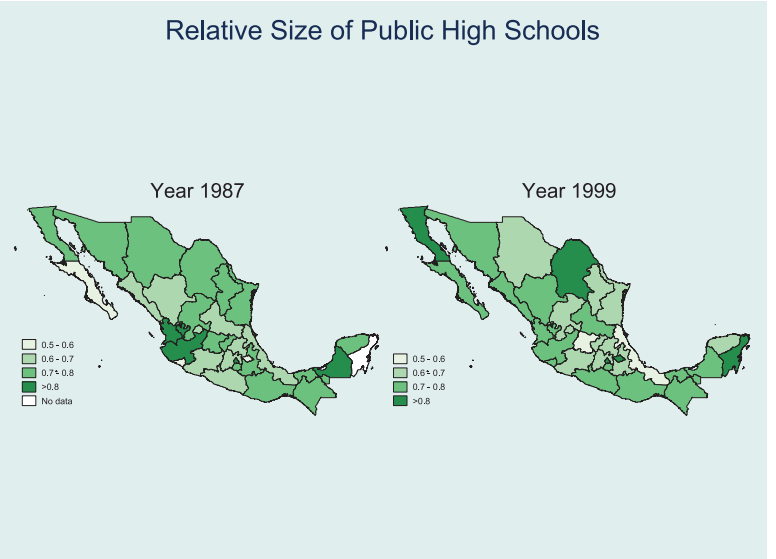

Fig. 7. Relative size of public high schools by state in year 1987 and 1999.
Source: authors' calculations based on data from the Mexican School Census.

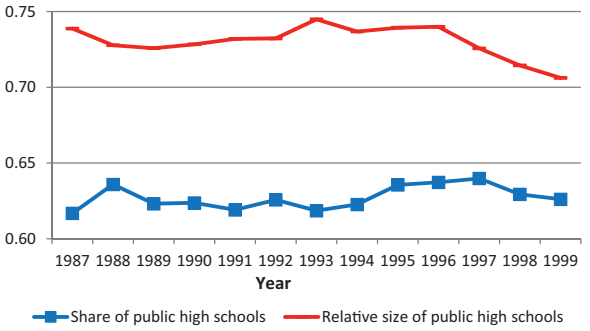

Fig. 8. Share and relative size of public high schools.
Source: authors' calculations based on data from the Mexican School Census.

5. Results

In this section, we first report OLS estimates of the relationship between private high school attendance, wages and school achievement. These estimates do not control for the endogeneity of the school choice and are reported for comparison purposes. We then report 2SLS estimates, alongside the first-stage relationships between school choice and the instruments.

5.1. Preliminary evidence: ordinary least squares estimates

Table 2 presents OLS estimates of the returns to private relative to public high schools. In the first column, we report estimates for the entire sample of analysis. In the following columns we report estimates by the highest level of education achieved: completed high school, uncompleted college and completed college. We find that the estimated returns to studying in a private high school are only statistically significant for the sub-sample of college graduates, and at around 5%. This is suggestive of a wage premium associated to studying at a private high school, for example due to a higher quality of education offered by private schools. Consistently with this indicative evidence,

Table 2
Wage returns to private high school (OLS).

	Log hourly earnings			
	All	HS graduates	Some college	College graduates
Private high school = 1	0.019	0.029	0.017	0.046^{+}
	(0.018)	(0.038)	(0.021)	(0.027)
Other covariates				
Female = 1	−0.107**	−0.146**	−0.067*	−0.062*
	(0.014)	(0.025)	(0.022)	(0.027)
Living with parents = 1	−0.074**	−0.101**	0.068**	−0.083*
	(0.015)	(0.018)	(0.022)	(0.026)
Observations	8523	2807	4217	3154

Notes: State-clustered standard errors in parentheses. Sample of workers aged 23–35 trimmed at bottom and top 0.5% of the earnings distribution. Hourly earnings are expressed in June 2011 prices. All specifications include state and cohort dummies. $^{+}p \leq 0.10$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

Table 3
School attainment (OLS).

	High school completion = 1				
	Unconditional	Conditional	College attendance = 1	College completion = 1	Years completed education
Private high school (HS) = 1	−0.057**	−0.135**	0.068*	−0.041*	0.018
	(0.015)	(0.029)	(0.024)	(0.016)	(0.092)
Other covariates					
Female = 1	0.042**	0.074**	0.007	0.014	0.328**
	(0.006)	(0.029)	(0.014)	(0.011)	(0.061)
Living with parents = 1	0.027**	−0.025	0.112**	0.012	0.567**
	(0.009)	(0.019)	(0.015)	(0.011)	(0.051)
Mean dependent variable	0.824	0.652	0.600	0.911	13.903
Observations	8523	4306	7024	3461	8520

Notes: State-clustered standard errors in parentheses. Sample of workers aged 23–35 trimmed at bottom and top 0.5% of the earnings' distribution. All specifications include state and cohort dummies. $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

individuals in the sample report perceived prestige and quality of the school as the most common reason why they chose to study in a private high school (37%), while this is only the third most common reason why students choose a public high school (20.9%). On the other hand, the most commonly reported reason to study in a public high school is distance to the house (34.4%), while this reason comes in third for private high schools, at 18.9%.[15]

Table 3 presents OLS estimates of private school attendance (relative to public) on school attainment both at the extensive margin—the probability of completing high school, of entering college and of successfully completing college—and at the intensive margin, that is the number of years of completed education. The first column reports the effect of private high school attendance on the (unconditional) probability of completing high school for the entire sample and, the second column reports the same probability conditional on entering the labour force upon high school completion. Perhaps surprisingly, we find that studying in a private high school has a significant negative effect on completing high school and college, while it has a positive effect on enroling into college.

Since OLS estimates do not control for selection into the type of school, they are very likely to be driven by unobserved individual and household characteristics that simultaneously affect school choice, school performance, educational attainment and wages, amongst other factors. The choice of attending a private rather than a public high school is likely to be positively correlated with the error term if private high schools attract better performing students, or if better performing students prefer to study at private high schools because these schools are perceived to provide education of higher quality (as suggested above). Or it could also be the case that it is universities that perceive private high schools as higher quality with prestigious colleges preferentially enroling students that attended a private high school. Such positive correlation could be reduced by negative selection bias if private high schools can only be afforded by rich families who would send their children to study in the best schools disregarding the level of their children's ability. In order to obtain unbiased estimates, we next control for selection.

5.2. Two stage least squares estimates

Table 4 presents 2SLS estimates of Eqs. (1) and (2). The structure of the table is analogous to that of Table 2. We find an overall wage premium of 44% to having attended a private high school for the entire sample of workers—i.e. irrespective of whether they stopped studying at high school or whether they continued onto college (column 1). This premium increases to 53% for those that entered but did not complete college (column 3), and to 60% for college graduates (column 4). Nonetheless, we find no differential wage returns to completing a private relative to a public high school for those that entered the labour market upon high school graduation (column 2).

In the three regressions where we find significant returns (columns 1, 3 and 4) both instruments in the first stage are highly significant and have the expected sign: an increase in the share and relative size of public high schools decreases the probability of attending a private relative to a public high school. The first-stage statistics show that the instruments are strong predictors of the private/public high school choice (probability of the first stage F-test statistic <0.05 and Cragg–Donalds F-statistic >10), and the Sargan test for over-identifying restrictions cannot reject the null that the instruments are uncorrelated with the error term.[16]

Note that these returns are for the sample of workers who report a positive wage. Such inclusion criterion could bias our results if attending a private or a public high school affected the probability of working or the probability of reporting a wage. To dismiss this concern we jointly estimate Eqs. (1) and (2) but replace the wage outcome variable in Eq. (1) with the probability of working (or that of reporting a positive wage). Results (available upon

[15] Interest in the subjects and courses offered by a given high school ranks as the second most important reason to attend both private (25.9%) and public high schools (34.4%).

[16] The correlation coefficient between the log of *SharePb* and the log of *SizePb* is −0.35, that is at most a moderate correlation, which allows to credibly using the Sargan test for over-identifying restrictions.

Table 4
Wage returns to private high school (2SLS).

	Log hourly earnings			
	All	HS graduates	Some college	College graduates
Endogenous explanatory variable				
Private high school (HS) = 1	0.436^{+}	0.104	0.532*	0.597*
	(0.268)	(0.501)	(0.263)	(0.271)
Other covariates				
Female = 1	−0.143**	−0.153*	−0.081**	−0.079*
	(0.025)	(0.071)	(0.022)	(0.029)
Living with parents = 1	−0.089**	−0.103*	−0.085*	−0.109**
	(0.019)	(0.033)	(0.027)	(0.032)
First stage				
lg(Share public HS)	−0.199**	−0.120*	−0.297**	−0.299**
	(0.031)	(0.050)	(0.044)	(0.053)
lg(Relative size public HS)	−0.302**	−0.089	−0.376**	−0.412**
	(0.070)	(0.117)	(0.101)	(0.125)
prob > *F*-stat first stage	0.000	0.039	0.000	0.002
Cragg–Donald Wald *F*-stat	22.754	2.948	26.078	19.11
P-Value Sargan test	0.469	0.185	0.934	0.310
Observations	8504	2800	4209	3148

Notes: State-clustered standard errors in parentheses. Sample of workers aged 23–35 trimmed at bottom and top 0.5% of the earnings distribution. Hourly earnings are expressed in June 2011 prices. All specifications include state and cohort dummies. $^{+}p \leq 0.10$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

request) show that having attended a private high school has no significant effect neither on the decision to work nor on the probability of reporting a positive wage, while the instruments remain strong and significant in the first stage.

Table 5 presents 2SLS estimates of the school attainment regressions at the extensive and intensive margins, following the same structure as Table 3.[17] For all regressions, the first stage results show that the instruments are significant and strong determinants of the decision to attend a private or a public high school. However, in contrast to the OLS results, we find that studying in a private high school does not give any advantage in school progression and completion, except for the probability of completing college education, which is negatively affected by having attended a private high school (although the coefficient is only significant at the 10% level).

5.3. Interpretation of the results

Both 2SLS (Table 4) and OLS (Table 2) results show a positive effect of attending a private high school on wages, conditional on college completion. The difference between them suggests that the OLS returns are affected by negative selection bias. This is consistent with the very high costs of private high school attendance and their reputation of being high quality so that while rich families can afford to send their children to the best schools regardless of the child's ability, low income high ability students might not be able to study at private high schools because of binding credit constraints (Card, 1999, chap. 30). Indeed, as discussed in Section 2, the average education cost (over median household income) in Mexico is 8% higher for students in private high schools than in public high schools, and the availability of scholarships is very limited. Consistently, Kaufmann (2009) has identified credit constraints as one of the main reasons behind the low college attendance and completion rates observed in Mexico despite high wage returns.

We find that the 2SLS estimates exceed OLS estimates by around 55%. Are the 2SLS returns over-estimated? An extensive empirical literature has estimated returns to education by using changes in school availability as instruments for educational choices using supply-side variables as a source of identification to estimate demand-side parameters (Card, 2001; Duflo, 2001). As noted by Card (2001), the use of changes in schools availability—such as changes in college proximity—can result in an over-estimation of the returns to school if the correlation between ability and education is different for different individuals, for example if it is higher for individuals that live closer to the schools. Imbens and Angrist (1994) show that in the case of heterogeneous returns to schooling, the IV estimate of the returns to education can be interpreted as a local average treatment effect (LATE) and computed as the weighted average of the marginal returns in the population where the weight of each person is given by the relative size of the increment in schooling due to changes in the instruments. The difference between the LATE and the average return depends on the covariance between the marginal returns and the change in schooling induced by

[17] When estimating the probability of attending or completing a given education level, both the outcome and the endogenous variable are binary, which means that the first stage conditional expectation function is likely to be non-linear. This violates the linearity assumption of the 2SLS estimator (Angrist, 2001), and can be addressed using the Wooldridge (2002, chap. 18) two-steps procedure. When using such procedure, we obtain very similar results, which suggests that the model is not identified off the probit functional form. Hence, we only report the results obtained via the standard 2SLS.

Table 5
School attainment (2SLS).

	High school completion = 1				
	Unconditional	Conditional	College attendance = 1	College completion = 1	Years completed education
Endogenous explanatory variable					
Private high school (HS) = 1	0.021	0.161	−0.009	−0.219^{+}	−0.541
	(0.163)	(0.299)	(0.203)	(0.123)	(1.321)
Other covariates					
Female = 1	0.035*	0.029	0.012	0.018^{+}	0.377**
	(0.016)	(0.048)	(0.017)	(0.011)	(0.119)
Living with parents = 1	0.024*	−0.036	0.116**	0.019*	0.586**
	(0.011)	(0.023)	(0.018)	(0.009)	(0.071)
First stage					
lg(Share public HS)	−0.199**	−0.101*	−0.225**	−0.297**	−0.199**
	(0.031)	(0.041)	(0.034)	(0.050)	(0.031)
lg(Relative size public HS)	−0.302**	−0.266*	−0.233**	−0.459**	−0.302**
	(0.070)	(0.096)	(0.079)	(0.118)	(0.070)
prob > *F*-stat first stage	0.000	0.042	0.000	0.000	0.000
Cragg–Donald Wald *F*-stat	22.754	4.326	24.582	21.847	22.750
P-Value Sargan Test	0.879	0.538	0.486	0.746	0.203
Mean dependent variable	0.824	0.652	0.600	0.911	13.903
Observations	8504	4295	7009	3454	8501

Notes: State-clustered standard errors in parentheses. Sample of workers aged 23–35 trimmed at bottom and top 0.5% of the earnings' distribution. All specifications include state and cohort dummies. $^{+}p \leq 0.10$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

the instruments. An IV procedure based on exogenous variation that affects the educational choices of higher returns individuals more than the educational choices of lower returns individuals will tend to produce an over-estimate of the average return to education. Thus, IV techniques will typically over-estimate the average return if there is heterogeneity in the returns. While we cannot directly test the heterogeneity in returns nor exclude it, in Section 6.6 we will use a technique recently developed by Nevo and Rosen (2008) that relaxes the assumption of instruments' exogeneity and thus allows the instruments to be correlated with unobservables including unobservable individual characteristics that may result in heterogenous returns.

6. Threats to identification

In our setting, the exogeneity condition requires that the instruments are uncorrelated with the error term in the schooling Eq. (2). Similarly, the exclusion restriction requires that changes in the relative availability and size of public high schools affect wages only through their impact on the decision to attend a private or a public high school. As we discuss next, violation of the instruments' validity could occur if there were demand-driven changes in high school provision, endogenous placement bias, omitted state-level factors, and between states migration, amongst others. We also investigate the role of confounding factors, such as quality of education and competition effects, the role of family background and that of following a more academic high school track, which could also violate the validity of the instruments. Finally, we follow Nevo and Rosen (2008) to directly test the robustness of our findings to weaker identification assumptions, namely to explicitly allowing for the instruments to be correlated with the error term in the outcome equation.

6.1. Demand-driven changes in high school provision and endogenous placement bias

Identification is threatened if changes in the relative availability and size of public high schools were demand driven or, in other words, if there was feedback from the individual demand for public and private high school education to the aggregate relative availability and size of high schools in a given state at different points in time.

There are several reasons to partially dismiss this concern. First, as discussed in Section 2.1, while the per capita number of private and public high schools increased (Fig. 1), the number of students per public high school decreased and the number of students per private high school remained stable (Fig. 3). This is consistent with a decrease in the demand for public high school education and a stable demand for private high school education. Further, the inclusion of state and cohort dummies controls for aggregate time-invariant variables—such as the supply of health care facilities in a given state and year—that can be correlated with the individual and, in turn, with the aggregate demand for schooling.

However, the endogeneity problem would not be solved if the individual demand for high school varied over time and across states. Hence, we re-estimate our benchmark specification by including a full set of state-year interactions to control for differential time trends by state. We have divided the 1987–1999 period in four sub-periods: 1987–1989, 1990–1992, 1993–1995, and 1996–1999, and use the 1987–1989 period as the reference group. Returns are even larger after the inclusion of these state-year period dummies (Model 1 in Table 6) and are robust to alternative definitions of the sub-period dummies (results available upon request), while the instruments remain significant and strong in the first stage.

Table 6
Robustness checks.

	Log hourly earnings								
	Benchmark	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Endogenous explanatory variable									
Private high school (HS) = 1	0.597*	0.627*	0.656*	0.693*	0.586*	0.734**	0.643+	0.582*	0.621*
	(0.271)	(0.244)	(0.273)	(0.246)	(0.272)	(0.280)	(0.344)	(0.277)	(0.227)
Other covariates									
Female = 1	−0.079*	−0.081*	−0.081*	−0.093*	−0.080*	−0.079*	−0.083*	−0.081*	−0.091*
	(0.029)	(0.031)	(0.030)	(0.033)	(0.029)	(0.028)	(0.030)	(0.029)	(0.032)
Living with parents = 1	−0.109**	−0.117**	−0.111**	−0.130**	−0.108**	−0.118**	−0.109*	−0.108**	−0.111**
	(0.032)	(0.033)	(0.033)	(0.032)	(0.032)	(0.033)	(0.035)	(0.033)	(0.029)
Yearly GDP growth by state				−0.806					
				(0.766)					
Yearly growth education expenditure by state				−0.000					
				(0.009)					
State of birth same as state of high school = 1						0.061			
						(0.043)			
lg(Number of teachers per class private HS)							−0.407+		
							(0.237)		
lg(Number of teachers per class public HS)							1.738**		
							(0.380)		
lg(Relative size public HS)								−0.163	
								(0.153)	
First stage									
lg(Share public HS)	−0.299**	−0.338**	−0.308**	−0.299**	−0.306**	−0.294**	−0.281**	−0.299**	−0.324**
	(0.053)	(0.076)	(0.076)	(0.083)	(0.075)	(0.076)	(0.075)	(0.075)	(0.079)
lg(Relative size public HS)	−0.412**	−0.369*	−0.407*	−0.306+	−0.419*	−0.452*	−0.346*	−0.412*	−0.392*
	(0.125)	(0.177)	(0.147)	(0.175)	(0.145)	(0.171)	(0.161)	(0.147)	(0.159)
Yearly GDP Growth by state				−0.094					
				(0.558)					
Yearly growth education expenditure by state				0.000					
				(0.004)					
State of birth same as state of high school = 1						−0.069*			
						(0.026)			
lg(Number of teachers per class private HS)							0.367		
							(0.249)		
lg(Number of teachers per class public HS)							−0.486		
							(0.396)		
State dummies × year period dummies		YES							
Cohort dummies × enrolment rates 1980			YES						
Cohort dummies × yearly GDP growth by State					YES				
Dummies State of birth						YES			
Quality of education × relative availability and size public HS in 1980							YES		
Prob > *F*-stat first stage	0.002	0.000	0.001	0.005	0.001	0.002	0.003	0.000	0.001
Cragg–Donald Wald *F*-stat	19.11	20.234	19.388	13.875	19.376	15.800	15.010	38.070	18.957
P-Value Sargan test	0.310	0.832	0.501	0.316	0.316	0.349	0.479		0.163
Observations	3148	3148	3148	2577	3148	3146	3148	3148	2742

Notes: State-clustered standard errors in parentheses. Sample of workers aged 23–35 trimmed at bottom and top 0.5% of the earnings distribution. $^{+}p \leq 0.10$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$. Model 1: interaction terms between State and year period (1990–1992, 1993–1995, 1996–1999) dummies. Reference group is period 1987–1989. Model 2: interaction terms between cohort dummies and enrolment rates in 1980 before high school expansion. Model 3: yearly growth real education expenditure and yearly GDP growth by State for each year between 1993 and 2000 as additional controls. Model 4: interaction terms between cohort dummies and yearly GDP growth by state. Model 5: state of birth dummies and indicator variable state of birth and state where high school was attended coincide. Model 6: control for measures of education quality and for their interaction with relative availability and size of public high schools in 1980. Model 7: relative proportion private high schools is the only instrument. Model 8: sample of those with *bachillerato general* and *tecnológico*.

The relative increase in the number of public high schools would also be endogenous if the government policy to expand public education had a compensatory nature, that is if the government decided to increase school availability more in states that had fewer high schools, lower enrolment or educational attainment rates, lower income levels, etc. On the one hand, the lack of an identifiable pattern in the variation of the instruments across states and over time plotted in Figs. 6–8 suggests that endogenous placement bias is not a serious concern. On the other hand, the 1992 Compensatory Programs (PARE and PAREIB), which were introduced as part of the 1992 *Ley de la Modernización de la Educación Básica y Normal*, explicitly aimed at closing education inequalities at the primary and secondary school level (Martinez-Rizo, 2001). As such, the concern that the expansion of high school also has a compensatory nature is valid.

We use GDP data and the locality marginality index collected by the National Population Council (*Consejo Nacional de Población* or CONAPO), aggregated at the state level, to classify localities by socio-economic status and investigate the extent to which the high school expansion is correlated with state-level socio-economic variables.[19] Plots of the expansion of public and private high schools by year and type of state ('low SES', 'high SES') (available upon request) show that from 1994 onwards (that is after the 1992 reform) there has been a faster increase in the number of public high schools in low SES and GDP states, and a larger increase in the number of private high schools in high SES and GDP states, which confirms the concern for endogenous placement bias. We address this concern in four distinct ways. First, we re-estimate Eqs. (1) and (2) by including as additional controls interaction terms between cohort dummies and enrolment rates in 1980, that is before the booming expansion of the mid-1980s (Section 2.1). We report results in Model 2 in Table 6. As shown, the returns to college conditional on private high school attendance remain positive and significant, at 66%. The first stage estimates confirm the ability of the instruments to strongly predict the choice of high school.

Second, we include GDP growth and growth in education expenditures in the state as additional regressors (Model 3 Table 6). These variables may be correlated with changes in the availability of high schools and have an independent effect on wages (for example through an increase in job vacancies). In addition, in Model 4 in Table 6, we allow GDP growth to have differential effects by cohort by adding a full set of interaction terms between cohort dummies and GDP growth by state.[20] As shown, results are robust to the inclusion of these variables and the instruments remain strong determinants of the choice of attending a private high school.

Third, we use census data for the year 1980, 1990, 1995, and 2000 to compute the proportion of households with electricity, piped water, and a drainage system by state and year, and we add these variables as additional controls.[21] The omission of infrastructure and services related variables could also bias results since they may simultaneously affect the demand for education and wages. Results (available upon request) remain unchanged after the inclusion of these variables.

Finally, we assign to each individual the value of the CONAPO marginality index in the locality of (current) residence closest in time to the year when the individual started high school, and add this variable as an additional regressor (assuming that the individual still resides in the locality where (s)he attended high school). Results (available upon request) confirm our main findings. In particular, as expected, the marginality index is negatively correlated with the choice of attending a private high school. In addition, we find that it is positively correlated with wages, which is consistent with higher marginality being associated with higher wage inequality. In the first stage the instruments remain very strong and statistically significant, and the returns to private high school are confirmed to be at around 0.60 and statistically significant at the 5% level.

[19] This index is computed every five years using locality socioeconomic indicators collected as part of the Population Census and Conteo (smaller census).

[20] Since GDP data by state are only available between 1993 and 2000, we have interpolated the yearly growth rate of per capita GDP by state for the previous years.

[21] Note that we linearly interpolate the data for the years in between.

6.2. Between-states migration

It is also plausible that high school entrants migrated across states in search for better schools. This type of migration would bias the results since the estimated impact of the schooling availability measures on the choice of school would reflect in part individual self-selection. As discussed in Section 4.1, about 84% of the individuals in the sample attended high school in the state where they were born and 88% live in the state where they went to high school. Hence, in our sample the proportion of those that moved states in search of better educational opportunities at high school is at most 16%, which is likely to be insufficient to bias the estimation results.

Most importantly, the proportion of those that moved between the year of birth and the year of high school attendance does not seem to be driven by a search for more and/or better schools. Indeed, if we compute the proportion of the sample that attended high school in a state other than the state of birth by age cohort, we find that overtime changes in this proportion are not correlated with either changes in the relative availability or in the relative size of public high schools, or with changes in standard measures of education quality (student–teacher ratio and student–class ratio for example). In addition, if we jointly estimate Eqs. (1) and (2) but we replace the wage variable in Eq. (1) with an indicator variable that equals one if the state of birth is the same as the state where high school was attended we find that having attended a private high school does not have any significant effect (coefficient 0.017, se 0.297). As usual, the instruments remain strong and significant in the first stage. These results are available upon request.

As a further robustness check, we re-estimate the model by controlling for state of birth dummies and for an indicator variable that is equal to one if the state of birth and the state of high school attendance coincide. As shown in Model 5 in Table 6, both instruments remain highly significant in the first stage and returns to college and private high school are estimated at a significant 73%.

6.3. Quality of education and competition effects

A further concern on the validity of the instruments relates to education quality. In addition to changes in availability, private and public high schools have followed different trends in terms of education "quality". Figs. 9 and 10 report the evolution of the number of students per teacher and the number of students per class between 1987 and 1999 for private and public high schools, which we have constructed using data from the School Census. Over time, the number of students per teacher is consistently higher in public high schools, which suggests

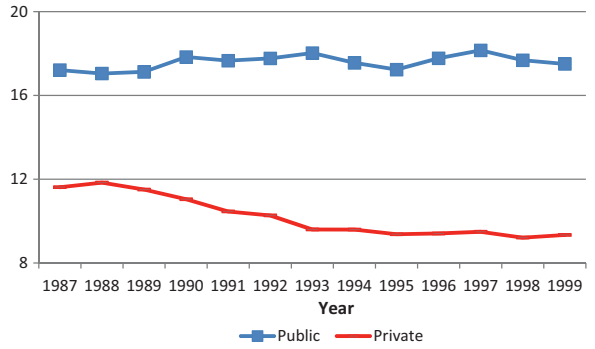

Fig. 9. Students per teacher in public and private high schools.
Source: authors' calculations based on data from the Mexican School Census.

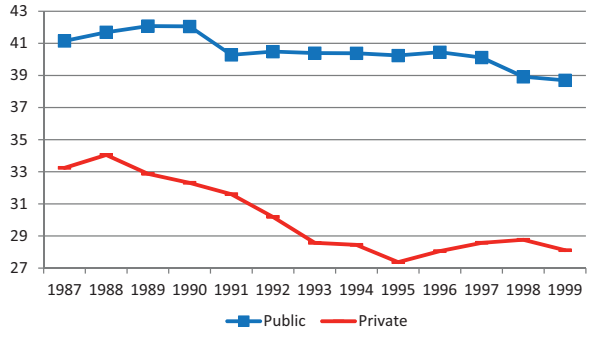

Fig. 10. Students per class in public and private high schools.
Source: authors' calculations based on data from the Mexican School Census.

that these schools offer lower quality of schooling as teachers have to share their time and attention with a larger number of students (Fig. 9). Moreover, this "quality" gap increased during the 1990s with respect to the 1980s: the number of students per teacher in public high schools by 1999 was double the number of students per teacher in private high schools, on average. Fig. 10 provides additional evidence of a quality differential between private and public high schools: despite the reduction in the number of students per class being common to both types of schools, it has been faster in private high schools, which suggests a trend towards higher quality in these schools. On average, in 1999, there were around ten more students per class in public than in private high schools.[22]

The finding that private high schools are of higher quality is consistent with evidence in the literature showing that in Mexico students in private high schools perform better in standardized assessments than students in public high schools (e.g. Gamboa & Waltenberg, 2011; Somers, McEwan, & Willms, 2004; The World Bank, 2005) as well as recent evidence from the Mexican National Assessment of Academic Achievement in School Centers (*Evaluación Nacional del Logro Académico en Centros Escolares* or ENLACE).[23]

The increase in the availability of high schools documented in Fig. 1 could have affected quality of education differentially across states depending on the initial availability of schools in a given state. To assess the extent to which education quality explains the differential return, we include the logarithm of the number of teachers per class in public schools and the logarithm of the number of teachers per class in private schools in our benchmark specification. In order to account for the possibility that education quality changed differentially in states with initial differential availability levels, we also add the interaction between these variables and the relative availability and size of public high school in 1980 (that is before the high schools' expansion). Estimation results reported in Model 6 in Table 6 show that the college premium to having attended a private high school is at around 64%, and that both measures of education quality matter for wages, while they do not have a significant effect on the decision of which type of school to attend.

Quality of education could also have been affected by competition effects between public and private high schools. For example, the increased presence of private high schools in a state may have exerted pressure on public high schools to offer better education in order to retain students. Competition from private schools has been found to raise educational attainment and wages (Hoxby, 1994 for the US; Angrist, Bettinger, Bloom, King, & Kremer, 2002 for Colombia) as well as teachers' salaries (Hensvik, 2010, for Sweden). However, other papers find that private schools are 'cream skimming' drawing the best schools out of the public sector, rather than competing on the basis of value added (Hsieh & Urquiola, 2003, for Chile). Overall, the effect of competition from private high schools on student achievement and wages seems to vary depending on the data set, the outcome, and the competition variable (Jepsen, 1999).

We assess the role of competition effects by allowing the relative size of private high schools to affect wages directly. This implies that, in this new specification, we use the relative proportion of private high schools as the only instrument. Model 7 in Table 6 presents the results. Relative availability remains strong and significant in the first stage and relative size does not significantly affect wages. We thus find no evidence of competition effects from private to public high schools. Importantly, when using relative availability as the only instrument, the impact of studying at a private high school on college wages remains strong and significant and close in

[22] We obtain very similar results if we use alternative measures of education quality such as the number of teachers per school and the number of teachers per class, which are commonly used in the literature as measures of the effective supply of teachers (e.g. Black and Smith, 2006; Card & Krueger, 1996).

[23] ENLACE data can be downloaded from http://www.enlace.sep.gob.mx/. In contrast with the evidence from the raw data, using ENLACE 2008 data De Hoyos, Espino, and García (2011) find no effect of private high school on test scores once a vast array of individual and parental characteristics is controlled for.

Table 7
Jobs After Graduation

	Private high school			Public high school				*T*-stat
	N	Mean	SD	*N*	Mean	SD	dif	
First job after college								
Weekly hours	288	41.500	12.560	816	41.654	14.234	−0.154	−0.265
Monthly wage	286	6075.927	3241.247	813	5686.935	3494.991	388.992	2.004
Self-employed = 1	288	0.052	0.223	816	0.044	0.205	0.008	0.555
Professional A = 1	288	0.431	0.496	816	0.434	0.496	−0.003	−0.119
Professional B = 1	288	0.201	0.402	816	0.223	0.417	−0.022	−1.473
Second job after college								
Weekly hours	124	40.645	15.493	325	42.289	13.237	−1.644	−1.838
Monthly wage	124	7194.193	5037.703	323	6314.749	4080.520	879.444	3.251

magnitude to the effect estimated in the benchmark model.

6.4. Family background and better jobs after graduation

Family background factors such as parental education and income could be one of the main reasons why private high schools are chosen over public ones. First, given the high costs of private schools and the limited availability of funding for private schooling (see Section 2), financial costs are the main barrier to enter private high schools and only high income families might be able to afford private education. Second, families with a preference for private high schools could be better connected socially and professionally, which could in turn facilitate finding better and higher paying jobs for their children. Indeed, network effects seem to be important: when looking at the reasons to study at high school, being influenced by friends and neighbours accounts for double the answers as a reason to attend a private rather than a public high school (6.4% for public high schools and 15.5% for private high schools).

Unfortunately, the available data do not allow to directly control for family background. As noted in Section 4.1, the ENTELEMS survey only reports information on parental level of education and work status for the sub-sample of young that were living with their parents at the time of the survey, which represents less than half of our sample and it is a highly selected sub-sample.[24] However, the available data allow exploring the effect of attending a private or a public high school on the type of jobs after graduation, which is arguably one of the important channels whereby family factors can affect labour outcomes, and hence assess the extent to which the omission of family background variables is likely to threaten identification.

The ENTELEMS survey contains data on the employment trajectories after college graduation for the sub-sample of those aged between 23 and 29 in 2008. We exploit these data to shed light on the role of private high schools in facilitating entrance into the labour market and accessing better jobs by comparing the type of occupations in the first and second jobs after college graduation for private and public high school graduates. Because the data do not include the year of college graduation, we use the year of high school graduation plus five years (which is the average length of a college degree in Mexico) as a proxy. Due to the young age of this sub-sample of workers, only a handful of individuals report having been employed on a third job after college. We therefore only consider the first two jobs after college graduation.

Table 7 compares weekly hours and monthly wages for the first and second job after college graduation for graduates that have attended a private or a public high school. For the first job after college graduation we additionally know whether the individual is salaried or self-employed, and whether he or she holds a professional position. We consider two types of professional jobs: "type A" that includes white collar-type jobs such as teachers, doctors, lawyers, architects, designers, and consultants, executives and other high-rank firm employees, and "type B" that includes liberal (arts) jobs. The final column reports the t-statistic of the difference in means.

Consistently with our main findings, average wages in both the first and in the second job after college graduation are significantly higher for graduates that attended a private high school and increasingly higher in the second relative to the first job after graduation (the wage gap increases from 389 to 879 real Mexican pesos per month). Nonetheless, we do not find any significant difference in the prevalence of self-employment and professional jobs performed by college graduates that have attended a private rather than a public high school, which suggests that both private and public high school graduates have access to similar types of jobs, at least soon after entering the job market, and conditional on college graduation. Further, and consistently with this finding, unreported results, (available upon request), show that our benchmark 2SLS results in Table 4 remain unchanged if we add the

[24] By comparing mean characteristics, individuals still living with their parents are younger, less likely to be married, and more likely to enrol in a private rather than in a public high school. The Mexican Family Life Survey (MxFLS) reports information on parental education level and work status for all individuals in the sample and can be used to construct information on family networks (Angelucci, De Giorgi, Rangel, & Rasul, 2010). However, it does so only for a very small number of workers, which does not allow to credibly estimate the wage returns to private schooling using IV (e.g. in the MxFLS 2005 there is information on 445 workers aged between 23 and 35 that have completed college; out of these, 75 attended a private high school and 370 attended a public high school).

type of job (salaried or self-employed) as an additional control in our model.

6.5. High school track

Finally, we assess the robustness of our results to redefining the estimation sample by conditioning on the type of high school attended. Unreported plots (available upon request) using SEP data on the number of private and public high schools by track and year between 1990 and 2000 show that the expansion of public high schools was mainly due to an expansion of *bachillerato general* and *técnico*, which, as discussed in Section (2), are the two types of high school that allow access to college education.

If attending a private high school has a positive impact on wages, and especially on the wages of those that have completed college, we would expect to estimate a stronger impact of studying at a private high school on the restricted sample of workers that followed a more academic (i.e. college-oriented) track at high school. We exploit the information in the ENTELEMS on the academic versus technical type of high school attended by dropping from the estimation sample all individuals that studied in a technical high school (*professional técnico*). Results in Model 8 in Table 6 confirm our expectations: returns to college if a private high school has been completed are estimated at around 62% on this sub-sample, which is around 2% higher than the average returns estimated for the entire sample in our benchmark specification. The instruments remain strong and significant in the first stage.

Further, in our data the proportion of private high schools' students is at around 34% both for *bachillerato general* and for *profesional técnico*, and at around 8% for *bachillerato tecnológico*, which indicates that *bachilerato general* students are over-represented and *bachilerato técnico* are under-represented with respect to the national averages in 2008 (see Section 2). If we run a robustness check using the sample of *bachillerato general* only, we find that in the first stage the instruments remain very strong and statistically significant, and the returns to private high school continue to be at around 0.61 and statistically significant at the 5% level.

6.6. Relaxing the exogeneity assumption: bounding the return

While the evidence so far is reassuring, doubts often remain as to whether the instruments satisfy validity since the exogeneity condition of no correlation with the error term is not directly testable and could be violated if returns are heterogenous (as discussed in Section 5.3). Also there are some variables—such as family background and school peer effects—that are potentially important to explain wages but which we cannot directly control for in the analysis given the available data. In this section we use the method developed by Nevo and Rosen (2008) to allow for correlation between the instruments and the error term, and derive bounds for the wage returns to private high school. As in other recent applications of this approach (e.g. Fitzsimons & Malde, 2010; Reinhold & Woutersen, 2009), we investigate the extent to which relaxing the exogeneity assumption affects our findings.

Following Nevo and Rosen (2008), we assume that:

1 The correlation between the endogenous variable and the error term, and between the instruments and the error term have the same sign. All the potential threats to identification that we discussed in Section 6 suggest that the potential correlation between our instruments and the error term in the wage equation is likely to be positive. Also, as discussed at the end of Section 5.1, the most likely correlation between the endogenous variable and the error term is positive. Finally, the first stage estimates showed that both instruments have a negative effect on the endogenous variable. If assumption 1 holds and, as in our case, the covariance between the endogenous variable and the instruments is negative, Nevo and Rosen (2008) show that we have a two-sided bound given by $B = [\beta^{OLS}, \beta^{IV}]$.
2 The correlation between the instruments and the error term is less strong in absolute terms than the correlation between the endogenous variable and the error term. This assumption weakens the validity assumption of zero correlation between the instruments and the error term, thus allowing to obtain tighter bounds. We believe that it is reasonable to expect the relative availability and size of public high schools in the state and year before high school enrolment to be less correlated with the error term in the wage equation than it is the individual choice of the type of high school to attend.

If, as in our case, assumption 1 and 2 are satisfied, Nevo and Rosen (2008) show that we obtain a two-sided bound given by:

$$B^* = [\beta^{IV^*}, \beta^{IV}]$$

where β^{IV^*} is a 2SLS estimator in which the instruments are re-defined as $SharePb^* = (\sigma_{SharePb} PvHS - \sigma_{PvHS} SharePb)$ and $SizePb^* = (\sigma_{SizePb} PvHS - \sigma_{PvHS} SizePb)$ where σ_{PvHS}, $\sigma_{SharePb}$, and σ_{SizePb} are, respectively, the standard deviation of the endogenous variable and of each instrument. When, as in our case, there are multiple instruments, the upper and the lower bound are identified, respectively, by the lowest β^{IV} and by the highest β^{IV^*} obtained by running separate regressions that use each instrument in turn.

We estimate the bounds of the effect of attending a private high school on wages for the main sample of college graduates for the benchmark specification in Table 6, and we find that $B = [0.077, 0.582]$. Therefore, when we relax the exogeneity assumption the wage returns to private high school are still positive and higher than the wage returns estimated with OLS. Unfortunately, the bounds are too wide to be informative on the point estimate of the returns, but, taken together with the evidence presented in the previous sections, suggest that there are substantial and robust wage returns to private high school for those that complete college education.

7. Conclusions

Assessing the relative efficiency of private and public schools is important for a number of reasons. First, the private sector can be used to expand educational provision under conditions of increasing demand for schooling and limited funding for social development. Second, private schools are often regarded as more efficient than public schools to the extent that families are willing to pay high tuition fees because of the greater choice on offer, which satisfies particular educational preferences (e.g. single-sex, religious schools or different language alternatives), or because private schools are regarded as an easy way of getting a degree in exchange of cash. A number of large-scale education reforms have been proposed encouraging public schools to mimic the technologies of private schools and promoting access to private schools via voucher and education subsidies. A leading example is the nationwide school voucher program implemented in Chile in 1980. A more recent example is the program of subsidies to private schools introduced by the Mexican government in February 2011.[25] Hence, it is important to investigate whether and via which mechanisms introducing a choice between public and private schools would help improving the overall achievement of students (Goldhaber, 1996).

This paper quantifies the relative efficiency of private and public schools by measuring the impact of private high schools on educational attainment and wages in Mexico. We exploit the significant increase in the relative availability and size of public high schools in the country in the 1990s across states and over time as exogenous shifters that affect individual choices to enrol in private and public high schools. We find that attending a private high school does not affect progression from high school to college but does increase wages conditional on college completion. Moreover, the comparison between IV and OLS results suggests that OLS estimates suffer from negative selection bias, which is consistent with the very high costs of private high schools.

The contribution of this paper is to establish a positive causal relationship between attending a private high school and wages in a middle income country. The main limitation of the empirical analysis is that the available data do not allow drawing a conclusive answer on the exact magnitude of the wage returns. Therefore, some notes of caution are necessary when drawing policy implications.

First, while the results are robust to a number of checks of the validity of the instruments, we are not able to account for potentially important unobserved factors that could explain part of the effect of private high schools on wages such as peer effects and family background. Second, with the available data we can estimate the short-term returns to private high schools, which might change in the long-run. Third, enrolment in private high schools is expensive and those currently attending these schools are likely to have access to relatively low-cost financing. For others, the cost of private schooling may be prohibitive. Thus, while private schools can improve access to education and do so efficiently (e.g. Bravo et al., 2008), the route to increased access and equity relates to the implementation of programs that address the equity concerns, such as scholarships targeted at students that cannot afford the costs of private schooling, and programs to increase the quality of education of public schools.

Overall, our findings motivate an effort to collect more detailed data on schools, students, their peers and families in order to ascertain the exact ways in which studying at a private high school affects educational achievement and labour market outcomes, and thus inform the design of effective educational policies.

[25] http://www.eluniversal.com.mx/editoriales/51699.html.

Acknowledgements

We thank two anonymous referees and seminar participants at the University of Mississippi, the University of Oxford, the IFS, the 2010 RCEA conference, the 2011 RES conference, and the 2011 LACEA meeting for useful comments. We are grateful to Armando Correa at CONAPO for sharing the population data; and to Rafael de Hoyos, Juan Manuel Espino and Fernando Ulises Gónzalez at SEP for sharing the ENTELEMS data and the Census 911 books, and for providing institutional knowledge throughout the project. Delphine Ghekiere and James Wisson provided excellent research assistance to prepare the final dataset. A previous version of the paper circulated under the title "The Market Returns to Private High Schools: Evidence from Mexico". The usual disclaimer applies.

References

Anand, P., Mizala, A., & Repetto, A. (2009). Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile. *Economics of Education Review, 8*(3), 370–381.

Angelucci, M., De Giorgi, G., Rangel, M. A., & Rasul, I. (2010). Family networks and school enrolment: Evidence from a randomized social experiment. *Journal of Public Economics, 94*, 197–221.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business and Economic Statistics, 19*(1), 2–25.

Angrist, J., Bettinger, E., Bloom, E., King, E., & Kremer, M. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review, 92*(5), 1535–1558.

Asadullah, M. N. (2009). Returns to private and public education in Bangladesh and Pakistan: A comparative analysis. *Journal of Asian Economics, 20*(1), 77–86.

Black, D. A., & Smith, J. A. (2006). Estimating the returns to college quality with multiple proxies for quality. *Journal of Labor Economics, 24*(3), 701–728.

Bedi, A., & Garg, A. (2000). The effectiveness of private versus public schools: The case of Indonesia. *Journal of Development Economics, 61*(2), 463–494.

Bosch, M., & Manacorda, M. (2010). Minimum wages and earnings inequality in Urban Mexico. *American Economic Journal: Applied Economics, 2*(4), 128–149.

Brown, C., & Belfield, C. (2001). The relationship between private schooling and earnings: A review of the evidence for the US and the UK. *Teachers College Occasional Paper No. 27, Columbia University*.

Calónico, S., & Ñopo, H. (2007). Where did you go to school? Private–public differences in schooling trajectories and their role on earnings. *Well-Being and Social Policy, 3*(1), 25–46.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics, edition 1, volume 3, chapter 30* (pp. 1801–1863). Elsevier.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127–1160.

Card, D., & Krueger, A. (1996). Labor Market Effects of School Quality: Theory and Evidence. *NBER Working Paper Series 5450*.

Chudgar, A., & Quin, E. (2012). Relationship between private schooling and achievement: Results from rural and urban India. *Economics of Education Review, 31*(4), 76–390.

CONAPO (1999). 25 años de cambio de la migración interna en México. Consejo Nacional de Población. México D.F.

De Hoyos, R., Espino, J. M., & García, V. (2011). Determinantes del Logro Escolar en México: Primeros Resultados Utilizando la Prueba ENLACE Media Superior. http://ideas.repec.org/a/elt/journl/v79y2012i316p783-811.html.

Domínguez, Á., & Pérez Gómez, M. H. (1993) El Bachillerato: Su evolución e influencia en la demanda de carreras científicas en el nivel de licenciatura. Perfiles Educativos, 62, Universidad Nacional Autónoma de México, México, D.F.

Duflo, E. (2001). Schooling and Labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review, 91*(4), 795–813.

Fitzsimons, E., & Malde, B. (2010). Empirically Probing the Quantity–Quality Model. *IFS Working Paper 10/20*.

Gamboa, L. F., & Waltenberg, F. D. (2011). Inequality of opportunity in educational achievement in Latin America: Evidence from PISA 2006–2009. *ECINEQ WP* (pp. 2011–2206).

Goldhaber, D. (1996). Public and private high schools: Is school choice an answer to the productivity problem? *Economics of Education Review, 15*(2), 93–109.

Gómez, R. (1999). The modernisation of higher education in Mexico. *Higher Education Policy, 12*(1), 53–67.

Hensvik, L. (2010). Competition, wages and teacher sorting: For lessons learned from a voucher reform. *Institute for Labor Market Policy Evaluation Working Paper 2010* (pp. 8–).

Hoxby, C. M. (1994). Do Private Schools Provide Competition for Public Schools?. *NBER Working Paper No. 4978*.

Hsieh, C. T., & Urquiola, M. (2003). When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program. *NBER Working Paper No. 10008*.

Imbens, G., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica, 62*(2), 467–476.

Kaufmann, M. K. (2009). Understanding the income gradient in college attendance in Mexico: The role of heterogeneity in expected returns. http://ideas.repec.org/p/igi/igierp/362.html.

Jepsen, C. (1999). The effects of private school competition on student achievement. *Northwestern University Working Paper.*

Martinez-Rizo, F. (2001). Las políticas educativas mexicanas antes y después de 2001. *Revista Ibero-Americana de Educación, 27*, 35–56.

Mukhopadhyay, S., Todd, P. T., & Bravo, D. (2008). School voucher effects on education and labor market outcomes: Evidence from a Chile. *2008 Meeting Papers 1069, Society for Economic Dynamics.*

Nevo, A., & Rosen, A. (2008). Identification with imperfect instruments. *CEMMAP Working Paper CWP16/08*.

Riddell, A. R. (1993). The evidence on public/private educational trade-offs in developing countries. *International Journal of Educational Development, 13*(4), 373–386.

SEP (2008). Sistema Educativo de los Estados Unidos Mexicanos. Prinicipales Cifras. Ciclo escolar 2007–2008. México.

SEP (2008). Reforma Integral de la Educación Media Superior. Anexo Único. México.

Reinhold, S., & Woutersen, T. (2009). *Endogeneity and imperfect instruments: Estimating bounds for the effect of early childbearing on high school completion.* http://econ.ucsb.edu/∼doug/245a/Papers/Imperfect%20Instruments.pdf.

Somers, M. A., McEwan, P. J., & Willms, J. D. (2004). How effective are private schools in Latin America? *Comparative Education Review, 48*(1), 48–69.

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

The World Bank (2005). In H. de Wit, I. C. Jaramillo, J. Gacel-Ávila, & J. Knight (Eds.), *Higher Education in Latin America*. The International Dimension.