FACULTAD ECONOMÍA Y NEGOCIOS
UNIVERSIDAD DE CHILE
DEPARTAMENTO DE ECONOMÍA

SERIE DE DOCUMENTOS DE TRABAJO

SDT 479

The Return to Private Education: Evidence from School-to-Work Transitions

Autores:
Dante Contreras
Jorge Rodríguez
Sergio Urzúa

Santiago, Enero de 2019

sdt@econ.uchile.cl
econ.uchile.cl/publicaciones

The Return to Private Education: Evidence from School-to-Work Transitions*

Dante Contreras
Universidad de Chile

Jorge Rodríguez
Universidad de Los Andes

Sergio Urzúa
University of Maryland
and NBER

This version: September 25, 2018

Abstract

This paper investigates the labor market returns to high school types. We exploit comprehensive administrative data describing the school-to-work transition for the universe of Chilean students attending tenth grade in 2001. We discuss the role of self-selection into school types, pre-labor market abilities, firm characteristics, and present bounds for the parameters of interest. Attending private high schools has long-lasting effects on earnings. Moreover, the long-term returns to school-level value-added measures and monetary investments in education are larger among private-school students. Our findings provide new insights into the association of school choice and the inertia of income inequality.

*Dante Contreras, University of Chile; email, dcontrer@econ.uchile.cl. Jorge Rodríguez, Universidad de los Andes, Chile; email, jp.rodriguez.osorio@gmail.com; Sergio Urzúa, Department of Economics, University of Maryland; email, urzua@econ.umd.edu. An earlier version of this paper circulated under the title "On the Origins of Inequality in Chile." We are thankful to the seminar participants at LACEA 2013 Meeting, UNU-Wider 2014 Conference, University of Chicago, University of Stockholm, University of Chile, and Pontificia Universidad Católica de Chile. We benefited from comments and suggestions from Derek Neal, Magne Mogstad, Tomás Rau, Loreto Reyes, and Cristián Dagnino. Dante Contreras thanks the support of Centro de Microdatos at the University of Chile through the Millennium Science Initiative sponsored by the Chilean Ministry of Economics, Development and Tourism, Project NS100041. Contreras also thanks financing provided by the Center for Social Conflict and Cohesion Studies (CONICYT/FONDAP/15130009). Jorge Rodríguez and Sergio Urzúa thank the Ministry of Finance of Chile for providing access to administrative data during 2013. This research was supported by the National Institutes of Health under award number NICHD R01HD065436. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

1 Introduction

The issue of unequal access to high quality education represents one of the most important challenges for developed and developing countries. A longstanding body of research has documented how disparities in the quality of educational services can explain the emergence of early test-score gaps (Fryer and Levitt, 2004), heterogeneous returns to education (Card and Krueger, 1992) and imbalances in long-term outcomes (Wachtel, 1975; García et al., 2016). However, potentially bigger issues might arise in settings in which access to high quality education is granted to only those individuals with the financial resources to afford it (Murnane and Reardon, 2017). Under these circumstances, which are commonly prevalent in developing countries, private education might transmit and amplify early inequalities over the life cycle.

This paper examines the long-run effects of attending different school types —public vs. private—in Chile. We exploit unique longitudinal data combining multiple sources of the country's administrative records. We gather information for the universe of Chilean students attending tenth grade in 2001—including individual-level test scores, a comprehensive list of school-level variables and students' socioeconomic characteristics—and detailed data describing their labor market trajectories (up to age 28). By using standard econometric arguments, we define bounds on the average effect of attending a high school type on labor market outcomes more than a decade after graduation. In constructing these bounds, we allow for the possibility that students and families self-select into school types based on observed and unobserved characteristics (e.g., student ability). Importantly, our results take into account firm-specific components.

When it comes to our main findings, we provide new and robust evidence of positive effects of attending private education on labor market outcomes. First, attending a private-fee-paying instead of a public high school in tenth grade boosts average adult earnings by 100-140 dollars a month, equivalent to 15-22% returns. The estimated effect is mostly driven by within-firm variation in earnings, whereas geographical location plays a minor role. Second, students attending private schools benefit more in the long run from improvements in academic achievement and school quality. In particular, the long-term impact of a one-standard-deviation increase in tenth grade math test score is higher for private-fee-paying students (154 dollars per month) compared to public high school students (95 dollars per month). Likewise, the impact of school value-added measures on earnings is 60% higher for those who attend private-fee-paying high schools.[1] Third, the returns to monetary investments in private-fee-paying schools exceed those in other institutions: a one-percent increase in monetary investments on these schools boosts their students' monthly adult earnings by 23 to 34 dollars, an effect five times larger than the corresponding estimate for public schools.

In light of our analysis, we contribute to a large literature examining the impact of attending private schools on students' outcomes. So far, the evidence have not reached a consensus on the relative effectiveness of private versus public schools. Neal (1997), Grogger and Neal (2000), and Altonji et al. (2005a) report positive but modest effects on test scores and relatively large effects on high school graduation for students attending private schools. A set of studies using voucher lotteries in the U.S. find relatively small achievement gains for students who were offered the voucher (Rouse and Barrow, 2009; Epple et al., 2017; Neal, 2018).[2] In Colombia, a voucher program (PACES) produced large effects on achievement and high school graduation (Angrist et al., 2002, 2006).[3] In an influential study, Muralidharan and Sundararaman (2015) disentangle the individual and aggregate effects of a school choice program using a two-stage lottery at the individual and market level—thus, their results are more comparable to ours.[4] They find that attending a private school produces null

[1]This result is in line with evidence that finds a causal impact of teachers' value-added on labor market outcomes (Chetty et al., 2014b).

[2]In a related literature, Angrist et al. (2013b), Dobbie and Fryer (2011), Abdulkadiroglu et al. (2011), Angrist et al. (2012, 2013a), and Fryer (2014) analyze the effects of charter schools (publicly funded institutions run independently by nonprofits or for-profit organizations) on lottery winners' test scores.

[3]The mixed evidence may be explained by the different systems under which the voucher programs was implemented. For example, voucher experiments in the U.S. allowed students to choose from any private school in a district. In Colombia, the voucher included an incentive component.

[4]In voucher programs, the average effects on those who exercise the right to use the voucher not only reflects differences in the productivity of private relative to public schools but also changes in peer composition.

effects on most of the subjects they study (Math, English, and Social Studies) but document that private schools spend less per student, suggesting a productivity advantage of private schools.

Our findings extend these efforts in at least three fronts. First, we look at adult earnings whilst most of the research focuses on short-term outcomes such as standardized test scores, high school completion, or college graduation rates.[5] Focusing just on standardized test scores may be problematic as improvements in test scores may not be caused by an effective boost in the student's human capital.[6] Second, we move beyond specific school types (charters, Catholic schools, etc.) to a broader assessment of school's effectiveness, covering a substantial part of the school spectrum within educational systems (private, voucher and public schools). Third, by exploiting multiple sources of information, we document the extent to which effects of school types on students' labor market outcomes can be determined by pre-labor market ability measures, firm-fixed effects, school value-added measures, and educational monetary investments.

Taken together, our findings imply that private education does have long-term consequences, which in turn might shape income inequality. As our empirical framework allows for general and strategic sorting partners, our findings not only help understanding the inertia of labor market disparities in Chile but also in other developed and developing countries. Thus, and more broadly, our study is also related to the recent empirical efforts to quantify the role schooling systems play in shaping adult earnings disparities (Chetty et al., 2014b).

The rest of the paper is structured as follows. Section 2 briefly describes Chile's education system. Section 3 presents our methodology to construct the bounds on the average effect of high school choices. Section 4 describes the available data sets. Section 5 presents our results and Section 6 concludes.

[5] Dobbie and Fryer (2016) and Sass et al. (2016), in the charter school literature, and Bravo et al. (2010), for the Chilean case, are exceptions.

[6] See Koretz (2002), Jacob and Levitt (2003), and Neal (2013) for a related discussion.

2 The Chilean Education System

Chile has historically exhibited high levels of income inequality and served as a breeding ground for the academic debate on the role of competition in education markets. This section provides a brief overview of the nation's education system and discusses its institutional features motivating this paper.

Three types of schools co-exist in Chile: public schools ran by local governments, private-voucher schools funded by the government and ran by private (for- and non-for profit) entities,[7] and private fee-paying institutions funded and ran by the private sector. To a large extent, this organizational structure had been in place during 20th century, but it was deepened after the wave of large-scale reforms implemented in the 1980s. The reforms decentralized the educational administration by transferring responsibility for public schools from the Ministry of Education to municipalities. They also expanded the voucher system for both public and private schools.[8]

The 1980s reforms led to a sharp redistribution of the educational system, giving a strong push to the private subsidized sector. In fact, the proportion of students at private-voucher schools rose from 15% in 1981 to 56% in 2011. Even though private-voucher and public schools face similar funding program, most of the private subsidized schools charge families with extra tuition (co-payment), while the opposite happens for public schools. This co-payment exited before the 1980s reforms but it was a second wave of changes (1993) that stimulated its use. Overall, the consequences of these reforms are considered a landmark example on the potential effects of a school choice scale-up.[9]

Table 1 presents tenth-grade average math an language test scores across high school

[7]While in 1981 religious institutions ran most private subsidized schools, most of the new post-reform schools were organized as for-profit institutions. By 1988, 84% of new schools were of this kind (Hsieh and Urquiola, 2006).

[8]Prior to the reform, there were already a few private subsidized schools, mainly religious institutions, with subsidies that were 50% of those given to public schools.

[9]See Hsieh and Urquiola (2006) for evidence on the Chilean case. For a review of the evidence on vouchers in other contexts see Rouse and Barrow (2009), Bettinger (2011), and Urquiola (2016).

types.[10] In terms of academic achievement, private-fee paying schools outperform voucher schools, which in turn outperform public schools. For math and language test scores, the table shows that private-fee-paying schools score 0.42 and 0.63 standard deviations higher than voucher schools. At the same time, private-voucher test scores surpass those of public schools by 0.34 and 0.32 standard deviations in math and language.

Figures 1 and 2 compare school types and labor market outcomes. Figure 1 presents kernel-based estimates of the distribution of log earnings across the high school type the individual attended while in the tenth grade. The figure shows that the distribution of earnings for private high schoolers is entirely skewed to the right with respect to the distributions of public and private-voucher students. Figure 2 plots a non-linear regression of academic achievement and earnings by high school types. It shows that conditioning on test scores does not fully remove the earnings gaps between private-fee-paying and the rest of the students. Furthermore, this gap is increasing in test score levels. On the contrary, the voucher-public earnings gap is smaller and does not increase with test scores levels. The figure favors the hypothesis that the earnings gap between private-fee-paying and the rest of the schools is not exclusively explained by differences in schools' production of academic skills.

Since schools operate under different market conditions, any comparison between private- and public-school students must be taken with a grain of salt. On the one hand, public schools are prohibited from selecting students (except in the case where demand exceeds the number of available slots). On the other hand, private schools can choose any student from their pool of applicants and charge tuition, so their student body is naturally skewed towards students from wealthier families.[11] This assertion is confirmed in Table 2, which presents the

[10] The tables and figures from this section use the same sample we use in our main regressions. These gaps are stable over the years. For more details, see Section 4.

[11] Private schools in Chile define their own admission policies, unlike voucher systems in other countries. In the voucher systems of Netherlands, Belgium, or Sweden, the private sector plays a significant role in education, but schools are not allowed to arbitrarily select their students. For example, in Sweden, private schools must operate on the first-come, first-served basis, and cannot select students based on ability, income, or ethnicity. Thus Swedish private schools are consistently found to have similar socioeconomic composition as public schools (Sandström and Bergström, 2005; Bohlmark and Lindahl, 2008)

proportion of students with college-educated mothers and average family income by school types. The average income in private-fee-paying schools is 3.7 and 2.6 times larger than that of public and private-voucher schools, respectively. Moreover, the proportion of students with college-educated mothers in private-fee-paying schools (39%) vastly exceeds the one in private-voucher (16%) and public schools (9%).

Our purpose is to assess if these types high schools have different impacts on future earnings. However, identifying causal effects in this context conveys multiple obstacles. In fact, the available evidence does not allow us to reach a general consensus on the effectiveness of attending a private-voucher over a public school.[12] If families perceive that some schools are better than others in terms students' skills production, and that skills production in a school is a function of baseline ability, then we should expect a nonrandom sorting across high school types. In the following section, we explain how we deal with this type of selection bias.

3 From biases to bounds on causal effects

In Chile, students self-select into school types. This endogenous process challenges the identification of the causal effects of interest. This section describes our identification strategy. It exploits our panel data and it is based on the definition of bounds for the average effect of high school choices on adult earnings.[13]

For simplicity, consider three time periods. In the first period, the agent chooses to attend a high school of type $j \in \{1, ..., J\}$, by taking into account her observed characteristics and inherent ability (A_i). In the second period, academic achievement ($\tilde{A}_i$) is observed. In the

[12] Several studies have analyzed the impact of to these different schooling institutions on students' academic achievement. See Carnoy and McEwan (2000), McEwan (2001), Carnoy and McEwan (2003), Hsieh and Urquiola (2006), Contreras et al. (2010), Auguste and Valenzuela (2006), Elacqua et al. (2011), and Gallego (2013). Unlike these papers, we show long-term impacts of private-voucher schools. Bravo et al. (2010) uses a structural approach to estimate the impact of voucher reform on labor market outcomes.

[13] Altonji and Mansfield (2014) exploit group-level averages of observed characteristics to construct bounds on the variance of the treatment effect of school choice. Even though our focus is on bounding conditional means, our methodology shares some elements with Altonji and Mansfield (2014) in that they use observed characteristics to account for selection based on unobservables.

third period, and after completing formal education, labor market outcomes are realized.

We seek to identify the average effect of attending a school of type j on adult earnings relative to a baseline category, J. Let $\boldsymbol{S}_i$ be a vector containing a set of $J-1$ dummies, $\{S_{i,1}, S_{i,2}, ..., S_{i,J-1}\}$, where $S_{i,j}$ equals 1 if the student i attends high school of type j, and 0 otherwise. Let w_i denote student's adult earnings. We define our parameter of interest as:

$$\kappa_j \equiv E[w_i \mid S_{i,j} = 1, A_i] - E[w_i \mid S_{i,J} = 1, A_i]. \tag{1}$$

In general, κ_j comprises two distinctive effects. First, school type j may rise individual's human capital, possibly increasing labor market productivity and earnings. Second, and conditional on inherit ability, attending a school of type j (relative to J) can have a direct impact on earnings through, for example, networks (Zimmerman, 2017). Individuals choose school j anticipating its impacts (schooling decision is endogenous), breaking down the identification of κ_j from a simple comparison of wages between those who attended j and the baseline alternative J.[14]

One potential strategy leading to the identification of κ_j is matching. To illustrate this method, consider the following linear regression model for wages:

$$w_i = \boldsymbol{S}_i'\boldsymbol{\kappa} + \boldsymbol{X}_i'b + \alpha A_i + \epsilon_i, \tag{2}$$

where $\boldsymbol{\kappa} = [\kappa_1 \ \kappa_2 \ ... \ \kappa_{J-1}]$ is the vector collecting our parameters of interest, $\boldsymbol{X}_i$ is a vector containing a rich set of observed characteristics (e.g., gender, age, family background variables, etc.), A_i represents individual's inherent ability, and ϵ_i is the error term. Under matching, $S_{i,j} \perp\!\!\!\perp \epsilon_i \mid \boldsymbol{X}_i, A_i$ for all $j = 1, ..., J-1$. Thus, ϵ_i does not have information that could explain movements in $S_{i,j}$, after conditioning on individual's characteristics. As suggested in Heckman and Navarro (2004), this strategy can be justified under the presumption that the researcher has access to enough data to mimic the individual's information set

[14] In identifying κ_j, we cannot infer the effect of increasing private competition on the market and academic achievement. See Hsieh and Urquiola (2006) and Urquiola (2016) for a formal treatment of this point.

at the time of the schooling decision. Thus, agents should not foresee that attending high school j instead of J has benefits beyond those anticipated on the basis of $\boldsymbol{X}_i$ and A_i, a prediction that can also be assessed by the econometrician.[15]

In practice, however, researchers usually do not have access to—or have imperfect measures of—A_i, which leads to biased results. In what follows, we use these biases to construct bounds for the parameters of interest.[16]

The upper bound: Omitted relevant variables. Consider a version of model (2) for the estimation of the economic returns to school type where A_i is omitted:

$$w_i = \boldsymbol{S}_i'\boldsymbol{\rho} + \boldsymbol{X}_i'\beta + v_i, \tag{3}$$

where $\boldsymbol{\rho} = [\rho_1 \ \rho_2 \ ... \ \rho_{J-1}]'$ and v_i is the error term. In this regression model, student sorting is a major hurdle preventing the direct identification of κ_j. Families might select school types based on the (short- and long-term) gains of choosing j over the alternatives. In this context, $E[v_i \mid \boldsymbol{S}_i, \boldsymbol{X}_i] \neq 0$ and OLS estimators would not identify the average earnings effect of attending a high school of type j.

Even though self-selection into schools based on A_i generates bias in the estimation of

[15] Appendix A presents a model of school choices which is consistent with stratification of students across public and private schools based on observed and unobserved characteristics (Epple and Romano, 1998). This model justifies our matching assumption and illustrates how our bounds for the average effect of attending private high schools can account for selection on unobservables. Appendix B shows that the probability of attending a private schools for public-school students traces almost all of the unit interval, although not many observations are close to 1. Estimates of our main regressions do not significantly change when we disregard observations outside common support (Table B1 in Appendix B), suggesting lack of support is not a major source of bias (Heckman et al., 1997). Appendix C shows that, in our context, the matching condition is a reasonable assumption and that possible deviations from it are likely to be economically insignificant.

[16] In principle, the method of instrumental variables could also be pursued to identify $\boldsymbol{\kappa}$. However, in settings in which selection into schooling types is driven by unobserved gains, without further assumptions, this approach might not identify the parameter of interest (Heckman et al., 2006b). Moreover, the availability of potential instruments might be limited by the complexity of the schooling choice problem. For example, general equilibrium effects can link preferences, educational production process, costs, prices, and institutional structure to outcomes (Epple and Romano, 1998), turning what might have been initially assumed as an exogenous source of variation affecting schooling choices (valid instrument) into a mediator of different outcomes. Finally, the identification through experimental or quasi-experimental methods, although can provide credible knowledge of "the effects," might limit our understanding of the consequences of the endogenous self-selection of students and families into schools, preventing the recovery of quantities that are useful for policy (Deaton, 2009).

κ_j, we can use this bias in our favor. With a suitable choice for a baseline school J, we can obtain upper biased estimates for every κ_j, $j = 1, ..., J-1$. Let ρ_j^{OLS} be the OLS estimate of ρ_j and $\boldsymbol{\rho}^{OLS}$ the vector collecting them for all j. To see how an OLS estimation on equation (3) generates upper bounds, note that the sign of the omitted relevant variable bias on each of the components of $\boldsymbol{\rho}^{OLS}$ depends on the signs of α and $Cov(S_{ij}^*, A_i)$ for all j, where S_{ij}^* is the residual of a regression of S_{ij} on the set of control variables. On the one hand, we can safely assume that $\alpha \geq 0$ given the literature documenting non-negative effects of different dimensions of ability on earnings (Heckman et al., 2006a). On the other hand, we can infer the signs of $Cov(S_{ij}^*, A_i)$ by examining the empirical association between schooling choices and academic test scores (imperfect proxies for ability). If these conditions for upward biased estimates hold, we can use OLS to generate an upper bound for κ_j for all $j = 1, ..., J-1$ from equation (3).

The lower bound: Proxies for ability and measurement error. With data on individual-level test scores, $\tilde{A}_i$, one could estimate a version of (2) by simply substituting A_i with $\tilde{A}_i$:

$$w_i = \boldsymbol{S}_i' \boldsymbol{p} + \boldsymbol{X}_i' \tilde{b} + \tilde{\alpha} \tilde{A}_i + \tilde{\epsilon}_i. \tag{4}$$

Although this simple approach might seem appealing, including $\tilde{A}_i$ in the wage regression still creates biases in the estimation of causal effects. Nonetheless, we can exploit these biases to obtain lower bounds on κ_j.

There are two sources of biases affecting the previous regression and we use both of them to form a lower bound on κ_j. First, the inclusion of test scores as proxies for ability generates measurement error biases. In the context of a multivariate regression, when only one regressor is measured with error, the estimators of all regression coefficients are potentially biased. The coefficient of the error-ridden regressor is biased towards zero, and the sign of the biases of the other parameters can be estimated consistently. These biases depend on the sign of the elements of the last column of $\Sigma^{-1} = \left([\boldsymbol{S}\boldsymbol{X}\tilde{A}]'[\boldsymbol{S}\boldsymbol{X}\tilde{A}]\right)^{-1}$, which can be consistently estimated in the data. In our empirical application, we obtain the sign of the bias associated to the school choice dummies to ensure that this first condition for estimating a lower bound is met.

The second source of bias emerges by noting that $\tilde{A}_i$, the second period academic achievement measure, is itself an outcome of school choice.[17] If we include $\tilde{A}_i$ in the wage regression, we would capture part of the effect of school type on earnings, as schools might have indirect effects on earnings through their contribution on raising academic achievement in the short term.[18] If academic achievement has a positive impact on earnings, then the estimated parameter associated with S_{ij} is downward biased.

Below we discuss the empirical conditions supporting the estimation of lower and upper bounds on the parameters of interest.

4 The Data

Our empirical analysis exploits various sources of administrative information. Our database is unique in combining administrative records on students' test scores at the tenth grade (2001) with labor market trajectories (2004-2013).

To recover information on school choices, family background, and test scores, we use data from the 2001 Measurement System of Education Quality (SIMCE). Every year, with the goal of measuring the individual attainment on minimum curricula requirements, the Ministry of Education conducts a national standardized exam to all Chilean students at some schooling level. In 2001, the Chilean government surveyed tenth graders (16 year of age on average). We measure academic achievement using SIMCE's math and language test scores. Additionally, the SIMCE database records student characteristics and family background

[17] Along similar lines, Altonji and Mansfield (2014) use averages of students characteristics at the school and neighborhood levels to absorb part of the impact of school and neighborhood choices. This strategy allows them to estimate a lower bound of the across-school component of the variance of student outcomes.

[18] Appendix A works out a model where the researcher has access to test scores after students make their choices. We show that κ_j is not identified in a wage regression where $\tilde{A}_i$ is as a control variable.

variables ($\boldsymbol{X}_i$ in equations (2)-(4)). Here, we include gender, previous attendance to pre-primary education, a dummy variable indicating whether a student has repeated previous schooling levels, and region dummies (out of 13 regions across the country). We also include two sets of variables that trace out parents' answers to "*At home, during a normal week, your pupil studies...*" and "*Does the student has a job outside schools?*"[19] As proxies for family background, we include mother's and father's education, per-capita household income,[20] number of books at home, mother' and father's age, dummies for whether the mother or father (or both) live with the student, guardians' answer to "*how far along do you think you pupil will get in school*," whether guardians attend parents meeting at school, and if the guardians know previous results of SIMCE scores of the school. Finally, in the SIMCE database, we are able to recover information on private and public educational expenditures. We use this data in Section 5.5 to estimate the impact of educational expenditures on earnings.

We use the Unemployment Insurance (UI) administrative database to gather information on students' earnings. We observe monthly gross earnings for 2004-2013, for all of those who have reported at least one formal job contract up from 2002 to 2013. Students who have not had a job up to 2013 are still considered in the final data, where their wages are recorded with a zero.[21] Using the monthly records of earnings, we construct annual average monthly earnings, where we define earnings to be zero for those months with a missing record.[22]

In most of our empirical analysis, we use our last available year of earnings (2013), meaning that, by that year, students are 29 years old on average. According to evidence from the United States and Sweden (Chetty et al., 2014a; Nybom and Stuhler, 2016; Chetty

[19]For the first variable, we construct a group of dummies with the following answers (i) "*every day*", (ii) "*some days*", (iii) "*only if there is an exam coming*", (iv) "*almost never*", (v) "*I don't know.*" For the second variable, there are only two possible answers (yes or no).

[20]Household income is self-reported. The variable has 13 categories corresponding to non-overlapping intervals of monthly family income. We take the middle point of each interval and use the resulting variable in our regressions.

[21]For most of the cases, a student who is not in the UI database should not be in college; on-track high school graduation occurs by 2003, leaving 10 years to complete a post-secondary degree.

[22]For the year 2013, the monthly average considers data up to October.

and Hendren, 2015), individuals' ranking in the income distribution when they are in their early 30s are highly correlated to the ranking in the distribution of lifetime income. This evidence supports our choice of using students' earnings in their late 20s instead of, say, using an average of all earnings reported in our data.

The process to obtain our final sample is as follows. First, we disregard 1,663 students with a missing or duplicate national identifier. Eliminating these observations leaves us with 191,282 observations. Second, we drop students with missing values in some of the covariates included in our regression analysis and also students attending special education (such as schools serving students with disabilities). Our final sample contains information for 111,395 students. Table 3 presents descriptive statistics for all of the variables that we use in our regressions. Table 4 compare statistics on key variables between the original and final sample. The table indicates that the distribution of students across schools remain almost the same, although baseline ability levels are potentially higher in the final sample.

Conditions to attain lower and upper bounds. As shown in Section 3, the conditions for obtaining lower and upper bounds for κ_j depend on the definition of high school types (the baseline J) as well as the empirical relationships between test scores and earnings, and between test scores and school choices.

In order to obtain our upper bound, we define our baseline school J to be the public school category. Given this choice, we must estimate the sign of $Cov(S_j^*, A_i)$, for the private schools dummies, to check if we can have an upper bias on the causal effect on our parameter of interest. One problem in estimating $Cov(S_j^*, A_i)$ is that we do not observe A_i; however, we can approximate it by $\tilde{A}_i$. In our sample, the correlation of the (residualized) private-school dummy with measures of math and language equal 0.11 and 0.08, respectively. For the private-voucher dummy, the correlations of the corresponding residual with math and language equal 0.04 and 0.05, respectively.[23] Under these conditions, the OLS estimation of

[23]Having all-positive covariances means that the baseline school is at the bottom of academic performance ranking. See Appendix A for a formal derivation in the context of a school choice model.

equation (3) delivers our upper bound for κ_j for $j \neq J$.

To obtain lower bounds, we need to analyze the pattern of correlations between our school dummies and our proxies for ability. Appendix D analyzes the case of multiple regressors subject to measurement error in the context of our research question. In this Appendix, we show that—for a range of sensible values of unobserved parameters—the inclusion of test scores will generate negative biases on the relevant regression coefficients. This result justifies our interpretation of p_j^{OLS} (estimated using equation 4) as our estimated lower bound for κ_j for $j \neq J$.

How tight is our lower bound? Suppose $A_i = \tilde{A}_i$, thus controlling for test scores would secure the conditional independence between high school choices and the unobserved components of the regressions, and the identification of κ_j for all $j \neq J$. We test the plausibility of this assumption in two exercises. In the first test, we evaluate the influence of selection bias by using a previously omitted baseline characteristic as a dependent variable. In the second exercise, we estimate a selection model and see if we find support for a statistically and/or economically insignificant correlation between the error terms of the selection and earnings equations. We present the results for these tests in Appendix C. The appendix shows that, even though we could still detect a degree of selection-on-unobservables bias, it is unlikely to modify our main conclusions as the resulting correlation coefficients are in the low range.

5 Results

We now turn to the analysis of the effects of high school types on labor market outcomes. We focus on the economic consequences of attending a private-voucher or private-fee-paying high schools relative to the public tuition-free alternatives and then show heterogeneous impacts.

5.1 Average effects of attending private high schools

Baseline regressions. Table 5 presents our estimated lower and upper bounds. These bounds come from the estimated regressions equations (3) and (4). To get a sense of orders of magnitude, the average earnings from students in private-fee-paying schools is 2.7 hundreds of dollars higher than students in public schools. This unconditional difference most likely overestimates the true causal effect of attending a private-fee-paying school: ablest students—coming from relatively wealthier, more educated families—have a higher probability of attending private schools. If we control for exogenous characteristics and family background—that is, our upper bound—, the earnings differential of private-fee-paying schools fall to 140 dollars. If we add test scores to the regression—thereby obtaining a lower bound—, the earnings gap between private-fee-paying schools over public schools drops to 100 dollars per month. Thus, the average effect of attending private-fee-paying instead of public schools is bounded between 100 and 140 dollars a month. This effect is equivalent to a 15-22% increase in earnings. The return to attending private-voucher versus public schools averages 10 to 22 dollars a month, although the lower bound is statistically insignificant.[24] Appendix E shows that employment effects are economically and statistically insignificant, suggesting that the impact of private schools on earnings is explained by a higher wage offer and/or more hours worked for a given hourly wage.

An alternative interpretation of our results emphasizes how schools may affect dimensions of human capital that are beyond academic achievement. The regression estimates from Table 5 show that the impact of school type does not vanish once we control for test scores. The coefficients on high school types represent the impact of high schools choices beyond schools' skills production. This result is consistent with a networking mechanism by which students at elite high schools get an extra reward in the labor market (Zimmerman, 2017).

[24] Auguste and Valenzuela (2006), Hsieh and Urquiola (2006), and Bravo et al. (2010) estimate the impact of private-voucher entry on students students' outcomes. The cross-sectional evidence is not conclusive: Auguste and Valenzuela (2006) find positive effects on test scores while Hsieh and Urquiola (2006) document statistically insignificant effects on test scores and Bravo et al. (2010) no effects on earnings. However, the estimated parameter in the literature—the overall effect from introducing private-voucher—is not directly comparable to ours—the individual effects from attending a private-voucher versus a public high school.

More general, the non-zero coefficient when controlling for test scores is also in line with the hypothesis suggesting that schools produce a set of skills that are rewarded in the labor market but have a low correlation with academic test scores.[25]

Figure 4 shows the upper and lower bound on the effects of private high school attendance on earnings across years. Each line shows the coefficient associated with the high school dummy: the circled line shows the estimated parameter from regressions that control for test scores (the lower bound) and the squared line depicts the estimated coefficients in regressions without test scores (the upper bound). The figures displays the estimated effects of average monthly earnings from 2007 until 2013. We find that, before 2009, the effect of attending a private-fee-paying versus a public high school is close to zero. In fact, the lower bound for that period is generally not statistically significant at the 5% level.[26] After 2009, all the estimated coefficients are statistically significant. Starting the year 2010, the impact of private-fee-paying schools increases at a decreasing rate. A similar pattern exhibits the estimated impact of private-voucher relative to public schools. However, in most of the years, the lower bounds of the average effect of attending a private-voucher relative to a public schools are not significant, so we cannot rule out null impact.

Effects of one year at high school type j on earnings. We follow Dobbie and Fryer (2016) to estimate the following relationship:

$$w_i = \sum_s NS'_{is}\rho_s + \boldsymbol{X}'_i\alpha + v_i, \tag{5}$$

[25]Petek and Pope (2016) find that value-added estimates based on non-test scores measures (such as classroom behavior) predict students' academic performance in high school. These non-test scores value-added measures have a low correlation with the more traditional test-scores value-added estimates.

[26]The lack of effect of private schools in the first years, and the increasing effect in the following periods, may be explained by the fact that some students are still enrolled in post-secondary degrees. This hypothesis is consistent with evidence showing that high-ability students are more likely to enroll in five-year post-secondary degrees, followed by four- and two-year degrees, and finally staying with a high school diploma (Rodriguez et al., 2016). Hence, the economic return to private schools increases with time as more talented students are entering the labor market each year.

where NS_{is} corresponds to the number of years that student i has spent in school of type s up until tenth grade (that is, the maximum value of NS_{is} is 10). ρ_s captures the effect of attending one additional year at school of type s on earnings. We construct NS_{is} with parents' answer to "*how many years have your pupil spent in this school.*"[27] The OLS estimation of Equation (5) gives an upper bound on ρ_s, while adding test scores to this equation yields a lower bound.

Table 6 presents our estimated bounds on the average effect of spending one year in high school type s (private-fee-paying or private-voucher) relative to spending that year in a public school. We estimate that the effect of spending one year in a private-voucher instead of a public is bounded between 4 and 2 dollars (although the lower bound is not significant). The effect of spending a year in a private-fee-paying school is bounded between 18 and 24 dollars a month (both bounds are statistically significant). Taking this estimate into account, if a student spends 10 years (from first to the tenth grade) in a private-fee-paying instead of a public school, it earns an additional 180-240 dollars of monthly earnings.[28]

5.2 The role of firms and geographic sorting

Our matched employee-employer data allows us to analyze two potential mechanisms explaining the effects of high school choices on labor market outcomes: across-firm and spatial sorting. We analyze how the return to high school choices changes when we control for firm and location ("comuna") fixed-effects.

The purpose of this exercise is to assess the importance of the mechanisms by which individuals cant take up the economic returns to high schools. In practice, individuals can exercise the economic returns to schools in—at least—three ways. First, individuals can sort

[27]A student could have switched from voucher school A to voucher school B. In this case, the parent's answer would take into account only the years the student have spent at B. Assuming a classical measurement error model, this source of error should attenuate our estimates. Thus, we are careful in interpreting ρ_s as the impact of spending one year in a particular voucher school.

[28]As said, these estimates are difficult to interpret as the actual effects of spending a year in a school type s. Nonetheless, even if we take into account a potential attenuation bias, these estimates suggest that the baseline regressions from Table 5—which use a dummy variable for tenth-grade attendance at school s—underestimate the long-run impact of spending a large part of the schooling years in a private school.

into those places where the skills associated to the school she went have the largest economic return. Second, individuals can sort into firms in which the skills acquired have the largest payoff. Third, conditional on the chosen firm, the skills associated to the chosen high school have an impact on earnings through affecting marginal productivity within a firm.

For this analysis, we use earnings for the last available month in the UI database (October, 2013) as the dependent variable. We use the last month on record, instead of the last year, to avoid situations in which workers change location or firms within a year. Also, since we would like to add firm fixed-effects to our regressions, our sample considers only individuals with positive records of earnings.[29] For October, 2013, we have 55,858 workers distributed across 21,599 firms and 329 comunas.[30] The average earnings of workers in this sample equals 1,297 dollars with a standard deviation of 966 dollars. Note that this average is higher than the average monthly earnings for 2013, since this last number includes more "zeros," capturing unemployment spells during the year.

Table 7 presents the estimated lower and upper bounds for three types of regressions: without fixed-effects, firm fixed-effects, and comuna fixed-effects. The bounds of the regression with no fixed-effects closely match those of our baseline regression (if we measure bounds as a percentage of the overall average of earnings in each case), suggesting that working with October 2013 earnings records does not affect overall quantitative conclusions. The fixed-effect regressions from Table 7 show a significant role for within-firm sorting and a minor incidence of across-firm sorting and location in explaining the impact of high school choices on earnings. The firm fixed-effects regressions cut down the estimated bounds by almost 30% for the private-fee-paying bounds and by approximately 20% for the private-voucher bounds. We conclude that a significant part of the return to high schools is explained by within-firms workers' sorting: the impact of attending a private-paid school relative to public school is bounded between 144 and 205 dollars (11-15% with respect to the mean) for

[29] Table E1 suggests that most of the effects of schools occurs at the intensive margin.

[30] Because of some missing values in the comuna indicator, we loose 235 observations in the comuna fixed-effect regressions.

individuals working in the same firm. In any case, the between-firm portion of the economic returns to private school is not negligible and coincides with literature suggesting that the between-firm wage variation explains approximately 20% of total wage variation (Abowd et al., 1999; Card et al., 2018) When we add comuna fixed effects, the estimated bounds are reduced by less than 9%.

5.3 Heterogeneous, long-term impacts of pre-labor market test scores

In this section, we contribute to the literature that finds positive effects of early skills on earnings by documenting heterogeneous effects.[31]

Table 8 presents the effect of test scores on earnings. This regression is equivalent to the estimates from Table 5, only this time we present the estimated coefficients associated with math and language test scores. The estimates show statistically significant impacts of test scores on earnings, both for math and language. The estimated earnings return to a one standard deviation increase in math test scores is 111 dollars a month—a 17%-increase relative to the average. The analogous estimate with respect to the language test score is lower: 7.8 dollars a month.

Figure 3 show evidence of nonlinear effects of improving SIMCE scores on earnings. The figure plots the estimated coefficients associated with dummy variables indicating whether the student belongs in different math or language test scores quantiles. The return on earnings of being in the fifth relative to the fourth quintile is higher than the differences between the the fourth and third, the third and second quintiles, and so on. This result emerges both in language and math test scores.

The second column of Table 8 presents heterogeneous effects of the math test score on earnings. We focus on math test score since this variable has larger effects on earnings than the language test score. On average for students in public schools, a one-standard-deviation

[31]See for example Neal and Johnson (1996) and Heckman et al. (2006a)

increase in math rises earnings by 95 dollars a month (15% relative to the average). We find that the interaction of math with the private-voucher dummy is not significant, so the effect of math on earnings for voucher students is statistically equivalent to that of public-school students. We do find a statistically significant coefficient in the interaction of math and private-fee-paying dummy. A one-standard-deviation boost in math test score increases monthly earnings for private-fee-paying students by 153.6 dollars—which corresponds to the sum of the baseline impact (94.8 dollars) with the additional private-fee-paying effect (58.8 dollars). This impact—a return 24% on average—is 61% larger than the effect we find for public- and voucher-school students.[32]

These results suggests that non-academic factors in private-fee-paying school play a role in generating future earnings. Suppose that each school generates earnings following a production function that depends on academic (A) and non-academic skills (NA): $w = f(A, NA)$. Table 8 shows that $\partial f(A, NA)/\partial A$ is larger for private-fee-paying schools than for private-voucher and public schools. Thus, our results are consistent with the presence of complementarities between academic and non-academic skills in producing earnings, a feature that is particularly stronger for private-fee-paying schools.[33]

5.4 School value-added and earnings

This section explores heterogeneous effects of high school type along the school academic value-added dimension. The general regression we estimate in this section takes the following form:

$$w_{i,\bar{t}} = \boldsymbol{S}_i'\beta_1 + \beta_2\Lambda_i + \Lambda_i \times \boldsymbol{S}_i'\beta_3 + \beta_4\tilde{A}_i + \boldsymbol{X}_i'\alpha + \varepsilon_i, \tag{6}$$

[32]These figures are lower bounds as the set of regressions control for individual test scores.

[33]However, we cannot distinguish if the larger $\partial f(A, NA)/\partial A$ is explained because private schools produce more NA or, for given baseline levels of A and NA, there is a higher degree of complementarity between factors.

where $\boldsymbol{S}_i$ represents a vector of high school dummies and $\tilde{A}_i$ denotes individual-level test scores. Λ_i is the school-level average of SIMCE test scores associated with the high school individual i attended.[34] Hence, we measure school value-added using the school-level language and math test score averages. $\boldsymbol{X}_i$ includes individual and family characteristics.

We interpret our results as the labor market return to school academic value-added, using a similar argument to identify the average effect of school types. Controlling for individual-level test scores allow us to deal with endogeneity in the school value-added measures thereby.[35] Nonetheless, including A_i in the regressions absorbs part of the impact of school quality on earnings and introduces attenuation bias in the estimators. Thus, the causal impact of school value-added is bounded between the corresponding coefficients of the pair of regressions with and without test scores as a control variable.

Table 9 presents the results. The first two columns present lower and upper bounds of school-value added on earnings in regressions without interactions. Our estimates show that school value-added have a statistically significant long-term impact (β_2 of equation 6): a one-standard-deviation rise in school's math average increases the student's future earnings by an average of 118-190 dollars a month (a return of 18-30% with respect to the overall average). The same regression shows that there is a negative, although statistically insignificant impact on earnings from rising the school's average language score.

In the school-interaction panel we document value-added impacts on earnings across school types (β_3 of equation 6). The estimates indicate that the return to school value-added varies by high school type. The differences are sizable. For public and private-voucher students, a one standard-deviation increase in school's math average raises future monthly earnings in the 90-161 dollars range. The same rise in math value-added for a student in a private-voucher school augments monthly earnings by 163-236 dollars. Mirroring the results from interactions of school types with individual-level scores (previous section), these results

[34] For the sake of simplicity, we assume in this equation that $\boldsymbol{A}_i$ and Λ_i are scalars. However, our regressions include math and language value-added measures.

[35] This result follows if individual-level test scores are partly caused by school value-added and individuals sort into school types taking into account school value-added impacts on individual test scores.

indicate that non-academic factors in private schools have a stronger complementarity with school-level academic production.

5.5 Monetary investment in education and adult earnings

In this section, we pool private and public sources of money spent in students' education to assess the impact of monetary educational investment on earnings.[36] We test whether the differences in school's impact on earnings is explained by the fact certain schools have more monetary resources and show how the productivity of one dollar of educational investment varies by high school type.

For this analysis, we exploit data on private and public money at the school level. Our goal is to obtain the total individual monthly monetary expenditures associated to the student's high school attendance. As we mentioned, we have information on tuition and other education-related expenditures from students' families. In particular, the SIMCE database contains tuition and other self-reported private expenditures.[37] In the case of public schools, we include in our regressions direct monetary transfers from local municipalities to public schools.[38] These transfers constitute the main source of funding for public schools.[39] For students in private-voucher schools, we compute the voucher amount that the schools receive for having the student enrolled.[40] Finally, for students attending private-fee-paying (and some of the private-voucher and public schools as well), we use the self-reported monthly tuition.

Table 10 documents average monthly educational expenditures across school types and

[36]We only measure observed monetary resources. We cannot account for other types of investment, such as parental time and others non-pecuniary elements of human capital investment.

[37]The voucher amounts depends on the amount of monthly average copayment (in the case of shared-funding schools) at the school-level. The same formula applies for public and private-voucher schools. To compute the voucher amount, we use administrative records of 2001 monthly average payments for shared-funding schools.

[38]These data come from the National System of Municipalities' Information (http://www.sinim.gov.cl/).

[39]We exclude from our sample 455 students from public schools (0.6% of the sample) with zero educational expenditures.

[40]The formula can be found in Ministerio de Educación (1998). In 2001, each month, a private-voucher school received 46 dollars per student. This number varies whether the school has JEC or if it charges copayment to families.

math test score. Higher educational monetary educational investment is associated with better academic performance measured by SIMCE, a feature that is independent of which school the student attends. The same table shows that students from private-fee-paying schools have the greatest amount of average investment (\$165.4 a month), followed by private-voucher (\$105.3) and public schools (\$82.0). Thus, the numbers show that the Chilean society invest more on the ablest students.[41]

Next, we estimate the effect of educational expenditures on earnings after accounting for selection into schools and other individual characteristics. Define c_{is} as the total monetary investment on education for student i at school s.[42] We estimate the following equations:

$$w_i = \gamma_1 \log(c_{is}) + \log(c_{is})\boldsymbol{S}_i'\gamma_2 + \boldsymbol{X}_i'\alpha + u_i \tag{7}$$

$$w_i = \tilde{\gamma}_1 \log(c_{is}) + \log(c_{is})\boldsymbol{S}_i'\tilde{\gamma}_2 + \boldsymbol{X}_i'\tilde{\alpha} + \tilde{A}_i'\tilde{\gamma}_3 + \tilde{u}_i \tag{8}$$

where γ_1 captures the effect of an 1% increase in educational expenditures on adult earnings. Using the argument from Section 3, the full impact of the equivalent increase in c_i lies between the OLS estimates of γ_1 and $\tilde{\gamma}_1$.

We show the estimated parameters of equations (7) and (8) in Table 11. Overall, we find that the return of educational monetary resources differs by school type. A 1% rise in monthly educational investment increases adult earnings by 23 to 34 dollars a month for private-fee-paying and 3 to 5 dollars a month for private-voucher (see columns 2 and 3). The estimated bounds on the returns to educational monetary resources for public schools imply larger effects than that of private-voucher schools, but they are not statistically significant.

Table 11 shows that the differences in the impact of schools on earnings are not solely explained by the level of school's monetary resources. To illustrate the quantitative implica-

[41]The evidence on charters schools for the United States indicate that part of the success that some charter schools have is due to the fact that these schools put extra educational resources—better teachers, more instructional hours, among other practices that should increase their annual budget. See for example Angrist et al. (2013c) and Fryer (2014).

[42]This component varies by student because sometimes schools charge different tuition to different families. On the other hand, tuition is self-reported by parents, so it may well be capturing measurement error. If this is the case, out estimates would be downward biased in general.

tions of this argument, consider the following hypothetical empirical exercise. Suppose that a policy-maker wishes to close the adult earnings gap between public and private-fee-paying students by injecting more money into public schools.[43] How much would it cost to close this gap? To answer this question, we compute the necessary increase in educational expenditures for the average public-school individual so that her earnings catch up to those of the average private-fee-paying student.[44] However, our results indicate that we cannot rule out (up to standard significance levels) a zero effect of investment on the earnings of public-school students. If the effect is indeed zero, then increasing monetary resources at public schools (everything else constant) cannot reduce future earnings gaps. Even if we ignore the statistically insignificant lower bound, using our estimated coefficients at face value we find that, in order to close the public-private wage gap, the educational expenditures in a public school student would have to be between 2,500 and 5,000 dollars a month—that is, at least 30 times bigger than the actual monthly monetary investment on the average public-school student. Taking the total number of tenth graders in 2001 and considering annual figures, the necessary, lower-bound investment to close the wage gap is equivalent to 3% of the GDP, or nearly the entire public expenditure in education from 2001.[45]

In sum, we find that educational monetary expenditures have heterogeneous effects on future earnings. Our estimates show that the long-term effect of each dollar invested varies by school type.[46] This source of comparative advantage has important implications for educational public policies searching for having long-term impacts on economically disadvantaged students. Policies that seek to equalize future earnings based on just increasing monetary educational expenditures on public high schools are in general ineffective—or at least far

[43]The 2013 average monthly earnings of public-school and private-fee-paying students correspond to 571 and 845 dollars, respectively.

[44]This exercise leaves fixed the exogenous observed and unobserved characteristics of these two hypothetical students, including their school choice. Also, it does not consider changes in school choices as a result of this policy.

[45]When computing this number we are ignoring the fact that the estimated lower bound is not statistically significant.

[46]These results are in line with experimental evidence from India showing that private schools spend less but produce similar test scores on students (Muralidharan and Sundararaman, 2015).

from being cost-effective.

6 Conclusions

This paper explores the impact of attending different high school types on adult earnings. We take advantage of rich and unique longitudinal information combining administrative records on individual pre-labor market test scores, school-age family background characteristics, school-level variables and labor market outcomes. Our empirical strategy allows for self-selection into school types.

Our main findings can be summarized as follows. First, we document that attending a private-fee-paying instead of a public school in tenth grade (age 16) boosts adult earnings by an average of 100-140 dollars a month (15-22% return). On the other hand, attending a private-voucher instead of a public school raises monthly earnings by 10-22 dollars, although the lower bound is statistically not different from zero. In addition, our results suggest that the large return to private institutions is mostly driven by within-firm variation. Adding firm fixed-effects in our regressions cuts down the middle points of the estimated bounds for the effects of private-fee-paying and voucher schools (relative to the public alternative) by only 30% and 20%, respectively; even for individuals working at the same firm, the monthly earnings return to private-fee-paying over public schools equals 11-16%. And relative to the baseline regression, adding geographical location ("comuna") fixed effects reduces the estimated bounds by 9%. Hence, geographical location in the labor market plays a minor role in accounting for the effects of high school on future earnings.

Second, in line with the literature, we find positive effects of pre-labor market skills on earnings (Neal and Johnson, 1996; Heckman et al., 2006a). However, we also document heterogeneous returns by high school type. The estimated effect of a one-standard-deviation increase in math test scores is higher for private-fee-paying students (154 dollars a month) than the effect obtained among those who attended public high schools (95 dollars a month).

Third, we uncover a robust and positive association between school value-added measures and adult earnings. In particular, we find that the estimated impact of school value-added on earnings is 60% higher for those who attend private-fee-paying high schools than for those at private-voucher or public institutions. Furthermore, the private-public earnings gap persists even after controlling for school value-added, suggesting that the labor market returns to private high schools are not entirely explained by academic factors. These results contribute to the recent literature identifying, for example, the causal impact of teachers' value-added on labor market outcomes (Chetty et al., 2014b).

Fourth, we measure individual-level monetary (public and private) educational investment and estimate its effect on students' earnings. We find that students at private-fee-paying schools have the highest levels of per-capita educational investment: 57 and 100% higher than that of voucher and public schools, respectively. Moreover, we show how these monetary investments have heterogeneous effects on adult earnings: a one-percent increase in monetary investment on education boosts monthly earnings by 23 to 34 dollars for private-fee-paying students (after controlling for test scores and firm fixed effects), an effect five times bigger than that of private-voucher and public school students.

All in all, our results suggest that the Chilean schooling system plays an important role in shaping long-term income disparities. This paper represents a first attempt towards understanding the link between school choice and the intergenerational transmission of income inequality in the context of a developing country.

References

Abdulkadiroglu, Atila, Joshua Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak, "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots," *The Quarterly Journal of Economics*, 2011, *126* (2), 699–748.

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, mar 1999, *67* (2), 251–333.

Altonji, Joseph and Richard Mansfield, "Group-Average Observables as Controls for Sorting on Unobservables When Estimating Group Treatment Effects: the Case of School and Neighborhood Effects," 2014. Unpublished manuscript.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151—-184.

_ , _ , and _ , "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151—-184.

Angrist, Joshua D, Parag a Pathak, and Christopher R Walters, "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, 2013, *5* (4), 1–27.

_ , Sarah R Cohodes, Susan M Dynarski, and Christopher R Walters, "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," 2013. NBER Working Papers No. 19275.

Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag a. Pathak, and Christopher R. Walters, "Who Benefits from KIPP?," *Journal of Policy Analysis and Management*, 2012, *31* (4), 837–860.

Angrist, Joshua, Eric Bettinger, and Michael Kremer, "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *American Economic Review*, jun 2006, *96* (3), 847–862.

_ , _ , Erik Bloom, Elizabeth King, and Michael Kremer, "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review*, nov 2002, *92* (5), 1535–1558.

_ , Parag Pathak, and Christopher Walters, "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, 2013, *5* (4), 1–27.

Auguste, Sebastian and Juan Pablo Valenzuela, "Is it just cream skimming? school vouchers in chile," 2006. Fundacion de Investigaciones Economicas Latinoamericanas, Mimeo.

Bettinger, Eric, "Educational Vouchers in International Contexts," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 4, Elsevier, 2011, pp. 551–572.

Bohlmark, Anders and Mikael Lindahl, "Does School Privatization Improve Educational Achievement? Evidence from Sweden's Voucher Reform," 2008. IZA Discussion Paper No. 3691.

Bravo, David, Sankar Mukhopadhyay, and Petra Todd, "Effects of school reform on education and labor market performance: Evidence from Chile's universal voucher system," *Quantitative Economics*, 2010, *1* (1), 47–95.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, jan 2018, *36* (S1), S13–S70.

_ **and Alan B. Krueger**, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, feb 1992, *100* (1), 1–40.

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman, "2001 Lawrence R. Klein Lecture Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice," *International Economic Review*, may 2003, *44* (2), 361–422.

Carnoy, Martin and Patrick J. McEwan, "The Effectiveness and Efficiency of Private Schools in Chile's Voucher System," *Educational Evaluation and Policy Analysis*, 2000, *22* (3), 213–239.

_ **and** _ , "Does privatization improve education? The case of Chile's national voucher plan," in David N. Plank and Gary Sykes, eds., *Choosing Choice: School Choice in International Perspective*, Teachers College Press: New York., 2003.

Chetty, R., N. Hendren, P. Kline, and E. Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, nov 2014, *129* (4), 1553–1623.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," 2015. NBER Working Paper No. 23002.

_ **, John N. Friedman, and Jonah E. Rockoff**, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, September 2014, *104* (9), 2633–2679.

Contreras, Dante, Paulina Sepúlveda, and Sebastián Bustos, "When Schools Are the Ones that Choose: The Effects of Screening in Chile," *Social Science Quarterly*, 2010, *91* (5).

Deaton, A., "Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development," *Proceedings of the British Academy, 2008 Lectures*, 2009, *162*, 123–160.

Dobbie, Will and Roland G. Fryer, "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 2011, *3* (3), 158–187.

_ and _ , "Charter Schools and Labor Market Outcomes," 2016. NBER Working Paper No. 22502.

Elacqua, Gregory, Dante Contreras, Felipe Salazar, and Humberto Santos, "The effectiveness of private school franchises in Chile's national voucher program," *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 2011, *22* (3), 237–263.

Epple, Dennis and Richard E. Romano, "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects," *The American Economic Review*, 1998, *88* (1), 33–62.

_ , _ , and Miguel Urquiola, "School Vouchers: A Survey of the Economics Literature," *Journal of Economic Literature (forthcoming)*, 2017.

Fryer, Roland G. and Steven D. Levitt, "Understanding the Black-White Test Score Gap in the First Two Years of School," *Review of Economics and Statistics*, 2004, *86* (2), 447–464.

Fryer, Ronald G., "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments," *The Quarterly Journal of Economics*, 2014, *129* (3), 1355–1407.

Gallego, Francisco, "When Does Inter-School Competition Matter? Evidence from the Chilean "Voucher" System," *The B.E. Journal of Economic Analysis & Policy*, January 2013, *13* (2), 525–562.

García, Jorge Luis, James J. Heckman, Duncan E. Leaf, and María J. Prados, "The Life-cycle Benefits of an Influential Early Childhood Program," 2016. NBER Working Paper No. 22993.

Grogger, Jeff and Derek Neal, "Further Evidence on the Effects of Catholic Secondary Schooling," *Brookings-Wharton Papers on Urban Affairs*, 2000, pp. 151–193.

Heckman, James J. and Salvador Navarro, "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models," *Review of Economics and Statistics*, 2004, *86* (1), 430–57.

_ and _ , "Dynamic discrete choice and dynamic treatment effects," *Journal of Econometrics*, 2007, *136* (2), 341–396.

_ , Hidehiko Ichimura, and Petra E. Todd, "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *The Review of Economic Studies*, 1997, *64* (4), 605–654.

_ , John Eric Humphries, and Gregory Veramendi, "Dynamic treatment effects," *Journal of Econometrics*, 2016, *191* (2), 276–292.

_ , _ , and _ , "Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking," 2016. Forthcoming in *Journal of Political Economy.*

_ , Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 2006, *24* (3), 411–482.

— , **Sergio Urzua, and Edward J. Vytlacil**, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 2006, *88* (3), 389–432.

Hsieh, Chang-Tai and Miguel Urquiola, "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of Public Economics*, 2006, *90* (8-9).

Jacob, Brian A. and Steven D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *The Quarterly Journal of Economics*, 2003, *118* (3), 843–877.

Koretz, Daniel M., "Limitations in the Use of Achievement Tests as Measures of Educators' Productivity," *The Journal of Human Resources*, 2002, *37* (4), 752–777.

McEwan, Patrick J., "The Effectiveness of Public, Catholic, and Non-Religious Private Schools in Chile's Voucher System," *Education Economics*, 2001, *9* (2), 103–128.

Ministerio de Educación, "Decreto con Fuerza de Ley No. 2, de Educacion, de 20.08.98," 1998. Santiago, Chile.

Muralidharan, Karthik and Venkatesh Sundararaman, "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India," *The Quarterly Journal of Economics*, aug 2015, *130* (3), 1011–1066.

Murnane, Richard J. and Sean F. Reardon, "Long-Term Trends in Private School Enrollments by Family Income," 2017. CEPA Working Paper No.17-07.

Neal, Derek, "The Effects of Catholic Secondary Schooling on Educational Achievement," *Journal of Labor Economics*, 1997, *15* (1), 98–123.

— , "The Consequences of Using one Assessment System to Pursue two Objectives," *The Journal of Economic Education*, 2013, *44* (4), 339–352.

— , *Information, Incentives, and Education Policy*, Harvard University Press, Cambridge: MA., 2018.

— **and William Johnson**, "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 1996, *104* (5), 869–895.

Nybom, Martin and Jan Stuhler, "Biases in Standard Measures of Intergenerational Income Dependence," *Journal of Human Resources*, 2016, pp. 0715–7290R.

Petek, Nathan and Nolan G. Pope, "The Multidimensional Impact of Teachers on Students," 2016. Unpublished manuscript, University of Chicago.

Rodriguez, J., S. Urzua, and L. Reyes, "Heterogeneous Economic Returns to Post-Secondary Degrees: Evidence from Chile," *Journal of Human Resources*, 2016, *51* (2), 416–460.

Rouse, Cecilia Elena and Lisa Barrow, "School Vouchers and Student Achievement: Recent Evidence and Remaining Questions," *Annual Review of Economics*, September 2009, *1* (1), 17–42.

Sandström, Mikael F. and Fredrik Bergström, "School vouchers in practice: competition will not hurt you," *Journal of Public Economics*, 2005, *89* (2-3), 351–380.

Sass, Tim R., Ron W. Zimmer, Brian P. Gill, and T. Kevin Booker, "Charter High Schools' Effects on Long-Term Attainment and Earnings," *Journal of Policy Analysis and Management*, jun 2016, *35* (3), 683–706.

Urquiola, Miguel, "Chapter 4 – Competition Among Schools: Traditional Public and Private Schools," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 5, Amsterdam: North Holland: Elsevier, 2016, pp. 209–237.

Wachtel, Paul, "The Effect of School Quality on Achievement, Attainment Levels, and Lifetime Earnings," in "Explorations in Economic Research, Volume 2, number 4," NBER, October 1975, pp. 502–536.

Zimmerman, Seth, "Making Top Managers: The Role of Elite Universities and Elite Peers," 2017. NBER Working Paper No. 22900.

Tables

Table 1: Academic performance by school type

	Language		**Math**	
	Average	**Std. Dev.**	**Average**	**Std. Dev.**
Public	-0.19	0.90	-0.17	0.97
Private-voucher	0.13	0.96	0.17	0.97
Private-fee-paying	0.76	1.12	0.59	1.00
Total	0.09	1.02	0.09	1.01

Notes: We show SIMCE 2001 test scores average (in standard deviations units with respect to the original-sample means) for different school types. The total number of observations is 111,395 students (the sample we use in our main empirical analysis).

Table 2: Student's family socioeconomic background and school choices

Type of school	College (mother)	Family income
A. Source of funding		
Public	11%	384
Private-voucher	20%	556
Private-fee-paying	48%	1561
B. For-profit status		
Nonprofit	21%	669
For-profit	18%	555
C. Co-payment requirements		
Public-nonshared-funding	10%	383
Public-shared-funding	12%	386
Private-voucher-nonshared-funding	15%	506
Private-voucher-shared-funding	21%	569

Notes: The table shows the proportion of college-educated mothers of students and the average monthly family income (2006 US dollars) of students across different types of schools. The sample we use to construct the table is the same as the one we use in our regression analysis.

Table 3: Descriptive statistics of the final sample

Variable	Mean	Std. Dev.	Min	Max
2013 average monthly earnings (hundreds of US$ 2013)	6.51	8.55	0.00	68.80
Language score	0.086	1.011	-4.821	3.056
Math score	0.092	1.023	-4.437	3.266
2001 monthly family income (hundreds of US$ 2013)	6.48	7.79	2.00	38.00
Public school (%)	45.4	49.8		
Private-voucher school (%)	37.7	48.5		
Private-fee-paying school (%)	16.9	37.5		
Male (%)	48.8	50.0		
Mother's education: primary (%)	37.3	48.4		
Mother's education: secondary (%)	30.9	46.2		
Mother's education: secondary vocational (%)	11.4	31.8		
Mother's education: technical institute (undergraduate) (%)	2.9	16.7		
Mother's education: professional institute (undergraduate) (%)	4.8	21.4		
Mother's education: university (undergraduate) (%)	10.5	30.6		
Mother's education: university (graduate) (%)	2.2	14.6		
Father's education: primary (%)	35.8	48.0		
Father's education: secondary (%)	28.8	45.3		
Father's education: secondary vocational (%)	12.5	33.0		
Father's education: technical institute (undergraduate) (%)	3.2	17.6		
Father's education: professional institute (undergraduate) (%)	4.3	20.2		
Father's education: university (undergraduate) (%)	13.2	33.9		
Father's education: university (graduate) (%)	2.2	14.8		
Books at home (<10) (%)	22.4	41.7		
Books at home (10-50) (%)	42.4	49.4		
Books at home (50-100) (%)	19.5	39.6		
Books at home (>100) (%)	15.7	36.4		
Repeated courses=0 (%)	16.5	37.1		
Repeated courses=1 (%)	5.6	23.0		
Repeated courses≥2 (%)	0.0	0.0		
Observations	111,395			

Notes: This table shows descriptive statistics of the variables we use in our regression analysis. Public, Private-voucher and Private-fee-paying are dummy variables that take the value of 1 if the students attends the respective school type and 0 otherwise. Pre-primary variables are dummy variables that equal to 1 if the student has attended a pre-primary school for the correspondent years (one or two) and 0 otherwise. "Only Primary" equals 1 if the student has not attended a pre-primary institution and 0 else. Mother and Father's educations variables are also dummy variables for each level of education. Books variables are dummies indicating the number of books as reported in the 2001 SIMCE.

Table 4: Descriptive statistics by database

Variables	**SIMCE data**	**Valid obs**
Language score (σs)	0.000	0.092
Math score (σs)	0.000	0.086
Public school (%)	47.6	45.4
Private-voucher school (%)	36.5	37.7
Private-fee-paying school (%)	15.9	16.9
Mother's ed: some college (%)	35.5	20.3
Observations	191,282	111,395

Notes: This table presents the average values of key variables associated with different databases. The first column (SIMCE data) corresponds to the original SIMCE 2001 data. The second column (Valid obs) drops observation with missing values in the SIMCE database in at least one of the variables considered in our regressions. This last sample is the one we use in our main regressions. Math and language test scores are expressed in standard deviations with respect to the original sample (first column).

Table 5: The effect of high schools on earnings

Variables	Lower bound	Upper bound
Private-voucher	0.100	0.218**
	(0.095)	(0.109)
Private-fee-paying	0.992***	1.436***
	(0.155)	(0.174)
Exogenous characteristics	Yes	Yes
Family background	Yes	Yes
Test scores	Yes	No
Observations	111,395	111,395
Dependent mean (hundreds of US$)	6.511	6.511

Notes: We show our estimated lower and upper bounds (equations 3 and 4) of the effects of schooling choices on 2013 average monthly earnings (in hundreds of dollars). We obtain the lower bound from a regression that includes math and language test scores. We estimate the upper bound from a regression that does not control for test scores. Both types of equations control for observed family and exogenous, individual characteristics (see Section 4 for details). The baseline category is public school (coefficients on high school dummies represent impacts relative to the baseline). Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

Table 6: Effects of number of years spent in school on earnings

Variables	Lower bound	Upper bound
Private-voucher	0.019	0.044***
	(0.013)	(0.013)
Private-fee-paying	0.181***	0.244***
	(0.019)	(0.019)
Exogenous characteristics	Yes	Yes
Family background	Yes	Yes
Test scores	Yes	No
Observations	110,228	110,228
Dependent mean (hundreds of US$)	6.511	6.511

Notes: We estimate the lower and upper bound of the effect of one year at a private-fee-paying and private-voucher school relative to spending one year in a public school on 2013 average monthly earnings (in hundreds of dollars). We obtain these estimates running:

$$w_i = \sum_s NS'_{is}\rho_s + \boldsymbol{X}'_i\alpha + v_i,$$

where NS_{is} corresponds to the number of years that student i has spent in school s up until 10th grade. This table shows upper and lower bound on $\rho_z - \rho_{\text{public}}$, where $z \in \{\text{private-fee-paying, private-voucher}\}$. To obtain the bounds, we compare regressions with and without math and language test scores. In parenthesis, we show bootstrapped standard errors that are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

Table 7: The effect of high schools on earnings: firm and location fixed-effects

Variables	No F.E.		Firm F.E.		Comuna F.E.	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Private-voucher	0.263**	0.484***	0.223*	0.376**	0.259**	0.467***
	(0.122)	(0.173)	(0.118)	(0.149)	(0.114)	(0.164)
Private-fee-paying	1.987***	2.845***	1.440***	2.045***	1.822***	2.627***
	(0.241)	(0.300)	(0.234)	(0.273)	(0.224)	(0.283)
Exogenous characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Family Background	Yes	Yes	Yes	Yes	Yes	Yes
Test scores	Yes	No	Yes	No	Yes	No
Observations	55,858	55,858	55,858	55,858	55,622	55,622
Dependent mean (hundreds of US$)	12.97	12.97	12.97	12.97	12.97	12.97

Notes: We show our estimated lower and upper bounds (equations 3 and 4) of the effects of schooling choices on earnings from October, 2013 (not including zeros, expressed in hundreds of dollars). We obtain the lower bound from a regression that includes math and language test scores. We estimate the upper bound from a regression that does not control for test scores. Both types of equations control for observed family and exogenous, individual characteristics (see Section 4 for details). We present upper and lower bounds for three types of models: without fixed-effects, firm fixed-effects, and location fixed-effects. The baseline category is public school (coefficients on high school dummies represent impacts relative to the baseline). Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

Table 8: The impact of early test scores on earnings

Variables	Baseline	Interactions
Language	0.078**	0.091**
	(0.039)	(0.038)
Math	1.114***	0.948***
	(0.047)	(0.074)
Math×Private-voucher		0.076
		(0.089)
Math×Private-fee-paying		0.588***
		(0.127)
Exogenous characteristics	Yes	Yes
Family Background	Yes	Yes
Test scores	Yes	Yes
Observations	111,395	111,395
Dependent mean (hundreds of US$)	6.511	6.511

Notes: We show estimates of the effect of math and language test scores on 2013 average monthly earnings (in hundreds of dollars). Test scores are expressed in standard deviation units. Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

Table 9: The impact of school average academic achievement on earnings

Variables	Baseline		School interaction	
	Lower bound	Upper bound	Lower bound	Upper bound
$\overline{\text{Language}}_j$	-0.237	-0.230	-0.017	-0.003
	(0.277)	(0.275)	(0.288)	(0.284)
$\overline{\text{Math}}_j$	1.182***	1.905***	0.897**	1.609***
	(0.273)	(0.267)	(0.351)	(0.344)
$\overline{\text{Math}}_j \times$ Private-voucher			-0.216	-0.217
			(0.233)	(0.233)
$\overline{\text{Math}}_j \times$ Private-fee-paying			0.732**	0.750***
			(0.287)	(0.288)
Exogenous characteristics	Yes	Yes	Yes	Yes
Family Background	Yes	Yes	Yes	Yes
Test scores	Yes	No	Yes	No
Observations	111,395	111,395	111,395	111,395
Dependent mean (hundreds of US$)	6.511	6.511	6.511	6.511

Notes: We show our estimated lower and upper bounds (equations 3 and 4) of the effects of school-level value-added on 2013 average monthly earnings (in hundred of dollars). We obtain the lower bound from a regression that includes math and language test scores. We estimate the upper bound from a regression that does not control for test scores. Both types of equations control for observed family and exogenous, individual characteristics (see Section 4 for details). The baseline category is public school (coefficients on high school dummies represent impacts relative to the baseline). $\overline{\text{Math}}_j$ and $\overline{\text{Language}}_j$ are school-level test scores averages. Test scores are expressed in standard deviation units. Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

Table 10: Total average educational monetary investment by high school type and academic performance

	Math test score				
	< 200	$200-300$	$300-400$	> 400	Total
Public	83.4	80.7	86.8	98.1	82.0
	[37.3]	[30.1]	[29.0]	[34.2]	[31.8]
Voucher	97.0	103.9	116.1	124.2	105.3
	[42.9]	[42.7]	[48.2]	[48.0]	[44.2]
Private-fee-paying	75.9	125.5	223.0	249.0	165.4
	[100.6]	[121.6]	[103.3]	[77.4]	[124.1]
Total	87.3	96.0	147.3	193.9	105.0
	[45.9]	[58.2]	[91.3]	[94.0]	[68.3]

Notes: We show monthly averages of educational monetary investments by types of schools and math test scores ranges. We compute average educational expenditures as the sum of the monthly tuition cost paid by families and other self-reported monthly expenses from the SIMCE data (in dollars). We add to this last number the amount of monthly subsidy for private-voucher and public schools and direct monthly transfers from municipalities to public schools. We also consider the additional monthly subsidies schools with JEC receive.

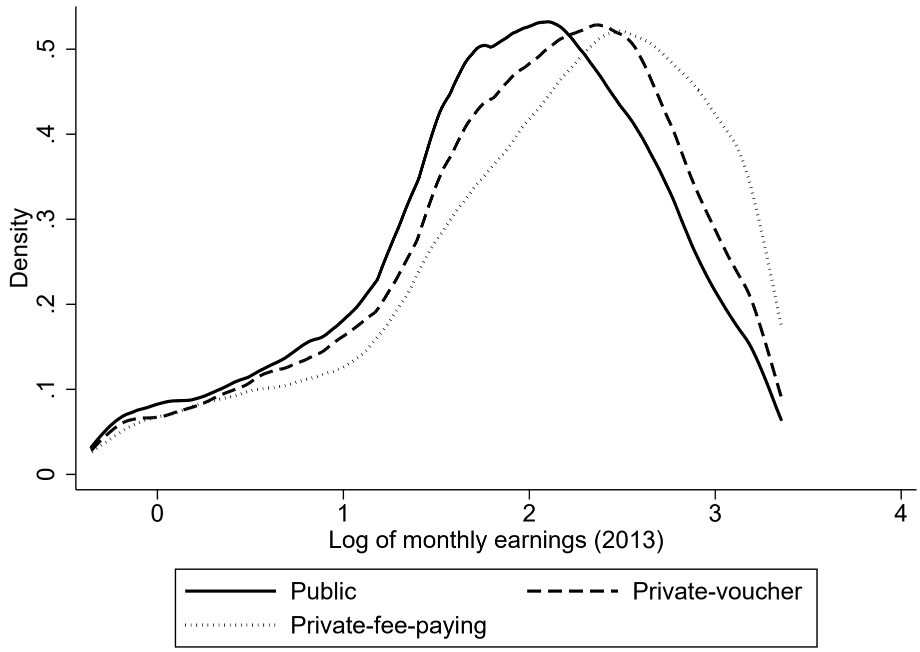
Table 11: The effect of educational monetary investment on earnings

Variables	Lower bound	Upper bound
Log(cost)	0.041	0.093
	(0.083)	(0.088)
Log(cost)*Private-voucher	0.025	0.050**
	(0.022)	(0.025)
Log(cost)*Private-fee-paying	0.227***	0.335***
	(0.035)	(0.040)
Exogenous characteristics	Yes	Yes
Family background	Yes	Yes
Test scores	Yes	No
Observations	107,282	107,282
Dependent mean (hundreds of US$)	6.511	6.511

Notes: We show our estimated lower and upper bounds (equations 3 and 4) of the effects of educational monetary resources on 2013 average monthly earnings (in hundred of dollars). We obtain the lower bound from a regression that includes math and language test scores. We estimate the upper bound from a regression that does not control for test scores. Both types of equations control for observed family and exogenous, individual characteristics (see Section 4 for details). Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

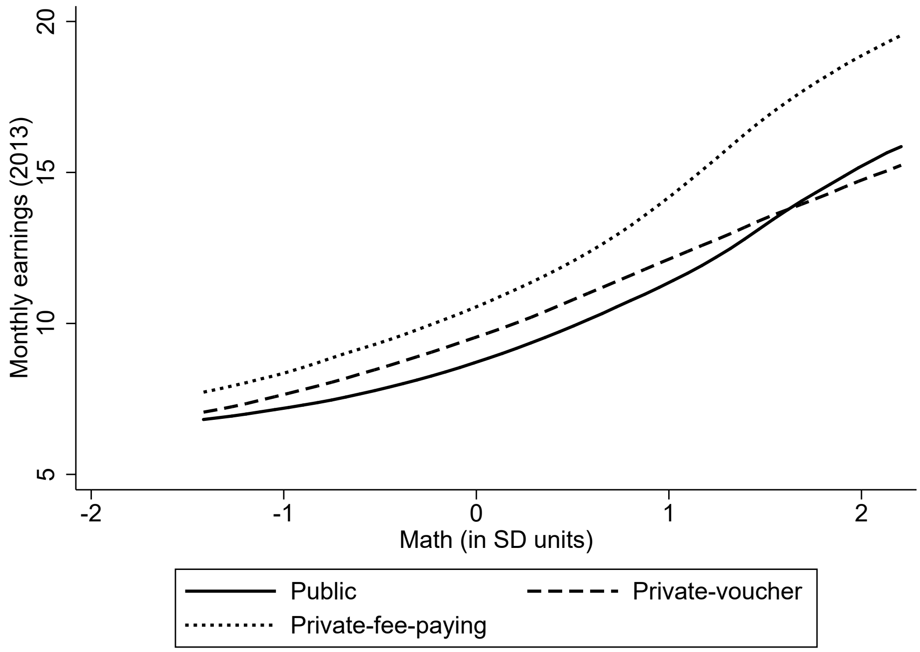
Figures

Figure 1: Log of earnings (2013) distribution and school types



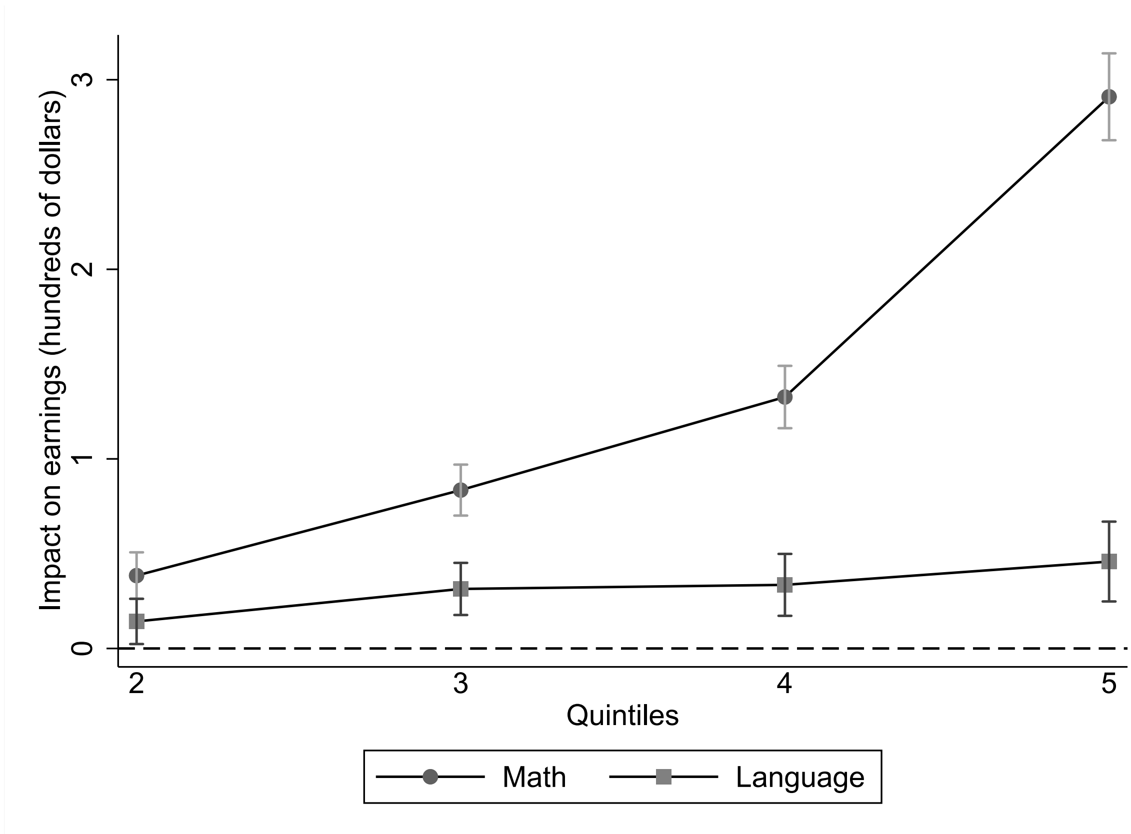
Notes: We show the estimated distribution of the 2013 log of monthly earnings for different school types.

Figure 2: Earnings (2013) and SIMCE test scores



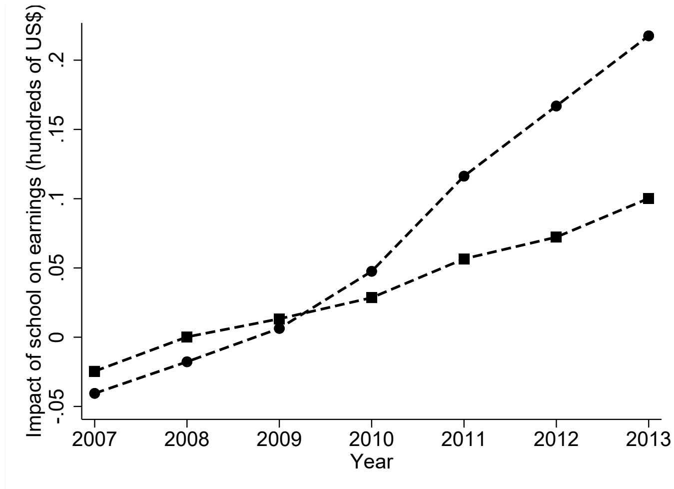
Notes: We show the fitted values of a local polynomial regressions of monthly earnings (2013 average in hundreds of dollars) and math test scores (in standard deviation units).

Figure 3: Labor market returns of academic achievement

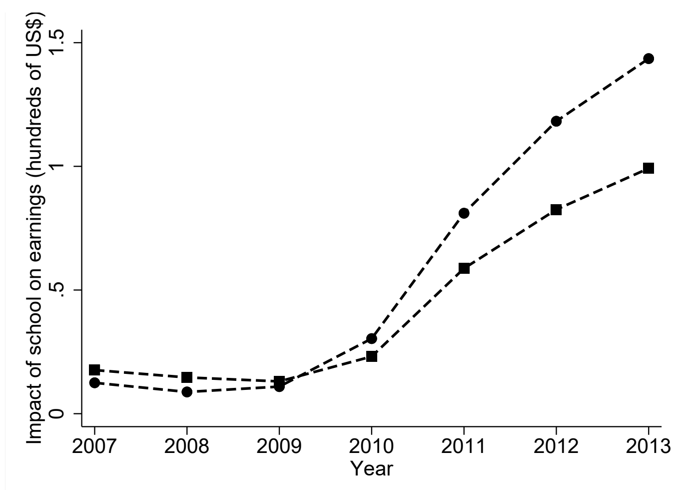


Notes: We show OLS estimates of a regression of 2013 monthly average earnings (in hundred of dollars) on a set of dummy variables indicating test scores quintiles. The baseline category is the first quintile of each test. Each point in the graph represents the effect of scoring in each quintile relative to the baseline. We control for exogenous characteristics, family background, and test scores (see Section 4 for details). Whiskers indicate a 95% confidence interval based on clustered robust standard errors.

Figure 4: Impact of private-voucher and private-fee-paying schools relative to public schools



(a) Private-voucher



(b) Private-fee-paying

Notes: We show OLS estimates of a regression of monthly average earnings (2004-2013, in hundred of dollars) on school choices (observed in 2001). Each dot represents the estimated coefficient associated to private-voucher or private-fee-paying attendance (see equations 3 and 4). The baseline category is public school (coefficients on high school dummies represent impacts relative to the baseline). The squared dots show the coefficient on a regression that controls for individual exogenous characteristics and family background. The circled dots show the coefficient adding test scores to the previous regression. The causal effect of attending private-voucher and private-fee-paying across time is bounded between the pairs of estimates for every year.

Appendices

Appendix A High school choices, test scores, and earnings regressions

In this section, we present a selection model of high school choices, academic achievement, and earnings. We show that our bounds method identifies the average earnings effect of high school attendance if agents behave as in this model. We begin by laying out the main ingredients and conclude with a discussion about the underlying assumptions.

The model. Consider the case where families can choose from a discrete set of school types $\mathcal{J} = \{1, 2, ..., J\}$. A cohort of families choose schools by maximizing their pupil's expected academic outcome from attending $j \in \mathcal{J}$. After families' choices are made, ex-post academic achievement is realized at period $t < \bar{t}$. Then $\bar{t} - t$ years later, wages are observed.

The production of measured academic skills follows

$$\tilde{A}_i^j = \bar{\pi}_j + \pi_{ij} + \nu_i^j, \tag{9}$$

where ν_i^j represents an i.i.d. mean-zero forecast error, $\bar{\pi}_j$ is a school-level coefficient representing type j's average quality, and π_{ij} is school- and -individual- level term capturing the interaction between individual-level ability endowment and school's value-added. Furthermore, $E(\pi_{ij}) = 0$. Families know their children's ability endowment and school j's value-added, but these are unobserved by the econometrician. The error term ν_i^j is unobserved by both the agent and the econometrician.

Let $S_{i,j}$ be a dummy variable that equals 1 if the observed choice for student i equals j and 0 otherwise. Let J be the baseline school category—that is, all estimated effects of school j are relative to choosing J. Families choose the school type that maximizes household welfare. Welfare depends on academic gains to attending school type j and on a vector of observables, such as family income. S_i^j is determined by:

$$S_{i,j} = \begin{cases} 1 & \text{if } j = \arg\max_{s\in\mathcal{J}}\{V\left(\bar{\pi}_s + \pi_{is}, \boldsymbol{X}_i\right)\}, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

where $V\left(\bar{\pi}_s + \pi_{is}, \boldsymbol{X}_i\right)$ represents the value function associated to choice j. In general, we assume that utility of choice j increases with school quality. Note that utility does not directly depend on student labor market outcomes associated to j, but indirectly though changes in $\bar{\pi}_s + \pi_{is}$. Equations (9) and (10) imply that, for a given $\boldsymbol{X}_i$, families have an ex-ante school-type ranking. The presence of π_{is} in the utility function implies that this ranking needs not to be equal across all families. This decision model is consistent with the general equilibrium analysis of Epple and Romano (1998). Utility might also depend on household income; thus, if we also take into account schools' selection mechanisms, a stratification along the family income dimension emerges.

Expected wages are formed through two main channels. First, academic achievement at period t have an impact of future wages. This is the indirect effect of choosing school j on earnings. Second, for a fixed level of academic outcome at t, there is a further impact of schools on earnings. Let $w_{i,j}^j$ be the the log wage for school $j \in \mathcal{J}$, at period $\bar{t} > t$. Wages are determined by:

$$w_{i,j} = c + \bar{\delta}_j + \delta_{i,j} + \alpha\tilde{A}_{i,j} + \varepsilon_{i,j}.$$

In this last equation, $\bar{\delta}_j$ and δ_{ij} represent the direct impact of schools on earnings, where $E(\delta_{ij}) = 0$. These terms are unobserved by the families and by the econometrician.

Wage regressions. The individual's observed wage corresponds to:

$$w_i = \sum_{j=1}^{J-1} S_i^j w_{i,j} + (1 - \sum_{j=1}^{J-1} S_{i,j})w_{i,J},$$

which implies that:

$$w_i = \widetilde{c} + \sum_{j=1}^{J-1} S_{i,j} \underbrace{\left[(\bar{\delta}_j - \bar{\delta}_J) + \alpha(\bar{\pi}_j - \bar{\pi}_J)\right]}_{\rho_j} + \widetilde{\varepsilon}_i. \tag{11}$$

where $\widetilde{c} = c + \bar{\delta}_J + \bar{\pi}_J$ and

$$\widetilde{\varepsilon}_i \equiv \varepsilon_i^J + \sum_{j=1}^{J-1} S_{i,j}(\varepsilon_i^j - \varepsilon_i^J) + \alpha \sum S_{i,j}(\delta_{ij} - \delta_{iJ}) + \sum S_{i,j}(\pi_{ij} - \pi_{iJ}) + \delta_{iJ} + \varepsilon_{i_J} + \alpha(\pi_{i,J} + \nu_{i,j}).$$

Our goal is to identify ρ_j: the ATE of choosing j relative to J on adult wages. This parameter contains the direct $(\bar{\delta}_j - \bar{\delta}_J)$ and indirect $(\beta(\bar{\pi}_j - \bar{\pi}_J))$ effect of school j relative to the baseline school J.

The OLS estimation on (11) would generate bias as S_{ij} is correlated with ε_i through π_{ij}. Too see how, note that the OLS estimate converges in probability to $E[w_i \mid S_{ij} = 1] - E[w_i \mid S_{iJ} = 1]$. Given how wages are formed, the OLS estimate converges to:

$$\begin{aligned} E[w_i \mid S_{ij} = 1] - E[w_i \mid S_{iJ} = 1] &= \rho_j + E[\widetilde{\varepsilon}_i \mid S_{ij} = 1] - E[\widetilde{\varepsilon}_i \mid S_{iJ} = 1] \\ &= \rho_j + \alpha \left(E[\pi_{ij} \mid S_{ij} = 1] - E[\pi_{iJ} \mid S_{iJ} = 1] \right). \end{aligned}$$

This last equation shows that we can obtain our upper bound following a suitable definition of the baseline school J. The baseline school is such that $\rho_j > 0$ for every j and there is an upward bias in the OLS estimation for every school $j \neq J$. To meet these requirements, J must be such that:

$$J = \arg\min_{k \in \mathcal{J}} \bar{\pi}_k = \min_{k \in \mathcal{J}} E(\pi_{ik} \mid S_{ik} = 1).$$

J is the "worst" school in the following sense. First, J has the lowest average level of skills production: $\bar{\pi}_j \geq \bar{\pi}_J$ for all $j = 1, ..., J-1$ (the first equality). Second, J generates the least amount of academic skills for its students (the second equality), so that $E[\pi_{ij} \mid S_{ij} = 1] - E[\pi_{iJ} \mid S_{iJ} = 1] > 0$ for all j. Under these two conditions, and even if families sort based on unobserved (academic) gains, we can obtain our upper bound for ρ_j, $j = 1, ..., J-1$.

The above analysis holds for observationally equivalent individuals—that is, we can condition every object on a set of observables and result would still go through. However, analysts often introduce a vector of observables $\boldsymbol{X}_i$ additively in the wage regression to account for selection on observables (equation 10). In this case, the construction of the upper bound would still hold. Let S^*_{ij} be the residual of running OLS on $S_{ij} = \boldsymbol{X}'_i\psi + \sum_{k\neq j,J} S'_{ik}\gamma + u_i$. Then the OLS estimate of ρ_j converges in probability to the standard bias formula $\rho_j + Cov(S^*_{ij}, \widetilde{\varepsilon}_i)/Var(S_{ij^*})$; the bias depends on the correlation between the unobservables in the model and the probability of attending j, after controlling for $\boldsymbol{X}_i$. Note further that we can write the unobserved part of equation (11) as $\widetilde{\varepsilon}_i = \beta A_i + \varepsilon^*_i$, where

$$\varepsilon^*_i \equiv \varepsilon^J_i + \sum_{j=1}^{J-1} S_{i,j}(\varepsilon^j_i - \varepsilon^J_i) + \sum S_{i,j}(\delta_{ij} - \delta_{iJ}) + \varepsilon_{i_J} + \delta_{iJ}.$$

Given our behavioral assumptions, the bias formula becomes $\rho_j + \beta Cov(S^*_{ij}, A_i)/Var(S_{ij^*})$. Therefore, we can ensure an upper bias for all ρ_j as long as $Cov(S^*_{ij}, A_i) > 0 \; \forall j = 1, ..., J-1$. The intuition for the result mirrors the univariate case: we need a baseline school that is at the bottom of the school-performance ranking.

Adding a vector of observed variables $\boldsymbol{X}_i$ would help obtaining an upper bound closer to the true ρ_j. Intuitively, controlling for observed characteristics should reduce, but not eliminate selection bias. That is, $E[\pi_{ij} \mid S_{ij} = 1] - E[\pi_{iJ} \mid S_{iJ} = 1] > E[\pi_{ij} \mid S_{ij} = 1, \boldsymbol{X}_i] - E[\pi_{iJ} \mid S_{iJ} = 1, \boldsymbol{X}_i]$.[47]

We obtain our lower bound by eliminating selection bias. If families behave as equation (10) dictates (that is, not using δ_{ij}), we can account for selection bias by controlling for $\tilde{A}_i$.

[47] In our regressions, we show that adding more regressors reduces the coefficient on the private-fee-paying school dummy.

To see this result, note that we can collect all terms associated to $\tilde{A}_i \equiv \sum_{j=1}^{J-1} S_{i,j}\tilde{A}_i^j + (1 - \sum_{j=1}^{J-1} S_{i,j})\tilde{A}_i^J$ and bring them in the "observable" part of the equation captured in A_i:

$$w_i = \tilde{\alpha} + \sum_{j=1}^{J-1} S_{i,j} \underbrace{(\delta_j - \bar{\delta}_J)}_{\tilde{\gamma}_j} + \beta A_i + \varepsilon_i^*. \tag{12}$$

where $\tilde{\alpha} = c + \bar{\delta}_J$ and

In this last regression, the OLS estimate is a consistent estimator (plim$(\tilde{\gamma}_j^{OLS}) = \tilde{\gamma}_j$. Let S_i^{j*} be the residualized choice dummy after controlling for A_i. In the above regression, S_i^{j*} is independent of $\tilde{\varepsilon}_i$, given that δ_{ij} and ε_i^j are i.i.d. Hence, the consistency result follows.

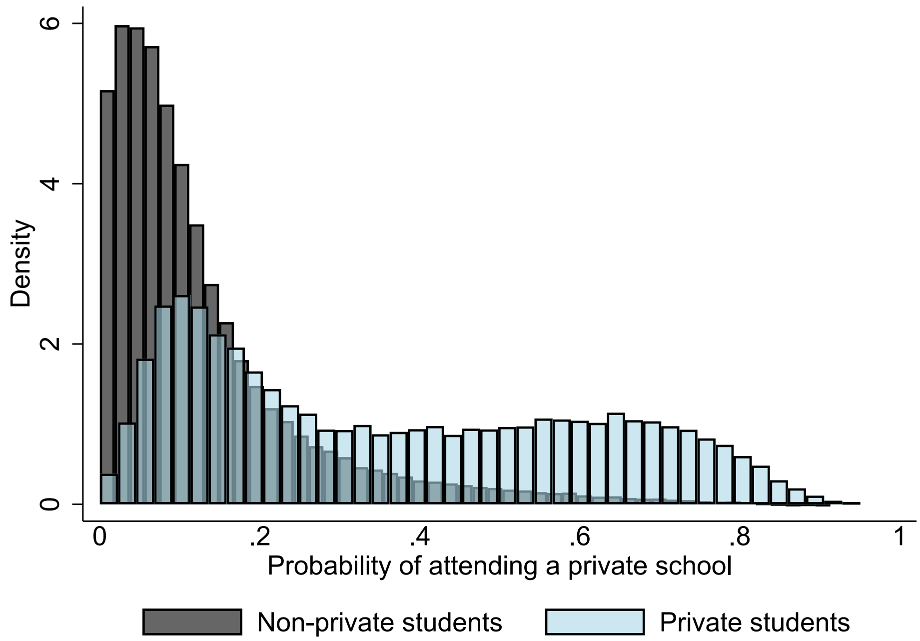
Yet the identified parameter in (12) ($\tilde{\gamma}_j$) is not equal to the original parameter of interest (γ_j). If $\bar{\delta}_{ij} - \bar{\delta}_J \geq 0$, then $\tilde{\gamma}_j \leq \gamma_j$. This result motivates the bounding procedure developed in Section 3.

Discussion. From the point of view of a linear regression, the behavioral assumption contained in equation (10) is a matching statement. It says that $\tilde{A}_i$ accounts for any unobserved factor that influences high school choices. Nonetheless, unlike the usual matching identification argument, our matching condition allows for selection based on unobserved variables. Equivalent matching assumptions have been widely used in (static and dynamic) models of schooling choices and labor market returns, where unobserved skills endowments are obtained from an array of cognitive and noncognitive test scores. In some of these models, the matching assumption is key to identify the joint distributions of counterfactuals. See for example Carneiro et al. (2003), Heckman et al. (2006a), Heckman and Navarro (2007), and Heckman et al. (2016a,b). Furthermore, the matching assumption is consistent with essential heterogeneity: families choices are made based on individual-level comparative advantages (Heckman et al., 2006b).

The important feature of $\tilde{A}_i$ that validates our method is that $\tilde{A}_i$ is itself an outcome of school choice. Because families choose schools based on the individual-level academic ex-ante gain if school type j is chosen ($E(\tilde{A}_i^j) = \pi_{ij}$), controlling for $\tilde{A}_i$ in equation (2) removes any selection bias. The variation contained in $\tilde{A}_i$ accounts for the variance of π_{ij}—the unobserved term by which students sort into schools.

Appendix B Common support check

Figure B1: Distribution of the probability of attending a private high-school, by school type



Notes: We show the estimated probabilities of attending a private-fee-paying school for private-fee-paying students, and public/private-voucher students (grouped as "non-private students.") To construct these probabilities, we estimate a probit model for attending private-fee-paying education (versus any alternative) including as controls the comprehensive set of independent variables discussed in the context of our baseline regressions (including test scores).

Table B1: Baseline regression with observations inside common support

Variables	Lower bound	Upper bound
Private-voucher	0.099	0.217**
	(0.095)	(0.109)
Private-fee-paying	0.992***	1.435***
	(0.155)	(0.174)
Exogenous characteristics	Yes	Yes
Family background	Yes	Yes
Test scores	Yes	No
Observations	111,373	111,373
Dependent mean (hundreds of US$)	6.511	6.511

Notes: This table replicates Table 5 but restricts observations with common support. We disregard individuals who attend private schools such that their predicted probability of attending a private school is greater than the maximum predicted probability for public and private-voucher students.

Appendix C How tight is our lower bound?

Suppose $A_i = \tilde{A}_i$, thus controlling for test scores would secure the conditional independence between high school choices and the unobserved components of the regressions, and the identification of κ_j for all $j \neq J$. In this section, we test the plausibility of this assumption in two exercises. In the first test, we evaluate the influence of selection bias by using a previously omitted baseline characteristic as a dependent variable. In the second exercise, we estimate a selection model and see if we find support for a statistically and/or economically insignificant correlation between the error terms of the selection and earnings equations.

Table C1 performs the first selection-on-unobservable test. We use a dummy for previous failed grades as the dependent variable.[48] This variable equals 1 if the students repeated at least one grade before the 10th grade and 0 otherwise. In the sample, 22% of students have repeated at least one grade.[49] If the regressions that include test scores meet the conditional independence assumption, then we should not find any impact of high school types on this omitted baseline variable. Column (1) of Table C1 shows that students at private-fee-paying schools are 13 percentage points less likely than students at private-voucher and public schools to repeat a grade. Relative to students at public schools, private-voucher are 7 percentage points more likely to have repeated a grade. When we add observed characteristics, the estimated coefficients associated to school dummies are cut down by half.W hen we add test scores into the formula, even though coefficients are still significant, we find that they are largely shrunk relative to the unconditional difference from column (1); the private-voucher coefficient falls 75% to -0.017 and the private-fee-paying coefficient drops 99% to -0.018.

In a second exercise, we estimate a selection model and test whether the correlation between the errors terms in the selection and outcomes equations is statistically significant

[48] In Chile, if a student under-performs (according the a standard GPA cutoff) then she must re-take all courses the following year.

[49] Arguably, this variable should be a good proxy for baseline latent academic achievement, which supports the interpretability of this falsification exercise as a measure of how well test scores deal with a selection-on-unobservables bias.

after controlling for test scores.[50] To keep the analysis simple, consider a model where students must choose between two high school types: private-fee-paying and others. Let S_i equals 1 if the student attends a private-fee-paying school and 0 otherwise. Formally,

$$
\begin{aligned}
Y_i &= \boldsymbol{X}_i'\beta + S_i\delta + u_i \\
S_i^* &= \boldsymbol{X}_i'\beta + \boldsymbol{Z}_i'\gamma + \nu_i \\
S_i &= \mathbb{1}[S_i^* > 0],
\end{aligned}
$$

where $\mathbb{1}[\cdot]$ is an indicator function such that $\mathbb{1}[B] = 1$ if B is true, and $\mathbb{1}[B] = 0$ otherwise; and u_i and ν_i are jointly distributed normal, with correlation coefficient ρ. If ρ is different from zero, then the estimated coefficient associated to the private-fee-paying dummy is biased because of selection on unobservables. From the perspective of our methodology, this bias means that there are not sufficient control variables X_i to eliminate selection on unobservables.

Table C2 presents the estimated correlation coefficients (ρ) and a chi-square statistic for the null hypothesis of $\rho = 0$. Because identifying this selection model requires a large support for $\boldsymbol{Z}_i$, we estimate it for various sets of instruments. This strategy allows us to assess the role of unobserved heterogeneity. Going from the first to the last column, we estimate the selection model using different sets of control variables. The first column includes only a constant, column number 2 adds observed individual characteristics, and column 3 adds test scores as control variables ($\boldsymbol{A}_i$ from equation 2).

The selection models presented in Table C2 show that controlling for test scores, although does not eliminate selection-on-unobservables bias completely, it considerably reduce it. In general, when we add control variables, the estimates of ρ do tend to zero. In those specifications that started with a correlation coefficients that were statistically significant, the estimated ρ (in absolute value) are reduced by around 40% (from 10 to 6%) in three of those

[50] Altonji et al. (2005b) follow a similar approach to test selection on unobservables on regressions that estimate the impact of catholic schools.

equations. In one of the models, the estimated ρ ends up being not significant after including test scores.[51]

[51] Only models 1, 4, 5, and 7 exhibit statistically significant correlation coefficients. Models 2, 3, and 6—which use the average math test and information on educational expenditures by comuna—appear to capture a margin that does not suffer from selection bias, so we will not consider them as informative about the capacity of our test scores to diminish selection on unobservables.

Table C1: Testing selection based on unobservables: regression of omitted baseline variable on high school dummies

Variables	Unconditional	Upper bound	Lower bound
Private-voucher	-0.069***	-0.031***	-0.017***
	(0.010)	(0.007)	(0.006)
Private-fee-paying	-0.128***	-0.061***	-0.018**
	(0.010)	(0.009)	(0.007)
Exogenous characteristics	No	Yes	Yes
Family Background	No	Yes	Yes
Test scores	No	No	Yes
Observations	111,395	111,395	111,395
Dependent mean	0.221	0.221	0.221

Notes: The table shows the impact on high school types on the probability of repeating a grade. This variable equals 1 if the student repeated at least one grade before the 10th grade and 0 otherwise. The first column shows the unconditional differences of the likelihood of failing a grade across school types while columns 2 and 3 the lower and upper bounds of the effects of high schools on grade repetition. The baseline category is public school. Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, ***$p < 0.05$*, *$p < 0.1$).

Table C2: Testing selection based on unobservables: estimating correlation coefficients of error terms in a selection model

Models: instruments	Unconditional	Upper bound	Lower bound
Model 1: Z_1	-0.15***	-0.10***	-0.08
	[258.3]	[96.2]	[61.5]
Model 2: Z_2	0.02	-0.01	-0.01
	[1.6]	[0.2]	[0.8]
Model3: Z_3	0.02	-0.01	-0.01
	[1.3]	[0.5]	[1.0]
Model 4: Z_1, Z_2	-0.11***	-0.07***	-0.06***
	[94.1]	[39.4]	[27.1]
Model 5: Z_1, Z_3	-0.10***	-0.08***	-0.06***
	[81.4]	[44.7]	[30.0]
Model 6: Z_2, Z_3	0.02	-0.01	-0.01
	[1.1]	[0.2]	[0.7]
Model 7: Z_2, Z_2, Z_3	-0.10***	-0.08***	-0.06***
	[81.6]	[42.3]	[28.5]

Notes: The table shows the correlation coefficient between the errors terms in the high school choice and earnings equations. We estimate:

$$Y_i = X_i'\beta + S_i\delta + u_i$$
$$S_i = X_i'\beta + Z_i'\gamma + \nu_i$$
$$S_i = \begin{cases} 1 & \text{if } S_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

In rows we show the estimates across models using different instruments (Z). Model 1 uses the share of private-fee-paying schools by comuna (Z_1). Model 2 uses the average math test score of private-fee-paying by comuna (Z_2). Model 3 uses the average educational expenditures for students enrolled in private-fee-paying schools by comuna (Z_2). Model 4 uses Z_1 and Z_2. Model 5 uses Z_1 and Z_3. Model 6 uses Z_2 and Z_3. Model 7 uses Z_1, Z_2, and Z_3. In columns, we show the estimates of ρ from models 1-7 across different sets of control variables (X). Column (1) includes just a constant. Column (2) includes includes all of the observed control variables discussed in Section 4. Column (3) adds test scores. We present a chi-squared statistic of the null hypothesis of a zero correlation (in square brackets). ***, **, and * indicate rejection of the null hypothesis at the 1, 5, and 10% level.

Appendix D Multiple test scores subject to measurement error

Consider the regression model $Y = \tilde{X}\beta + \varepsilon$ where $\tilde{X}$ is a $n \times k$ matrix. We observe instead $X = \tilde{X} + V$ where V is a $n \times k$ matrix of measurement errors with variance-covariance matrix Ω. Our goal is to infer the sign of the biases. The general formula for the inconsistency of $\hat{\beta}_{OLS}$ is given by:

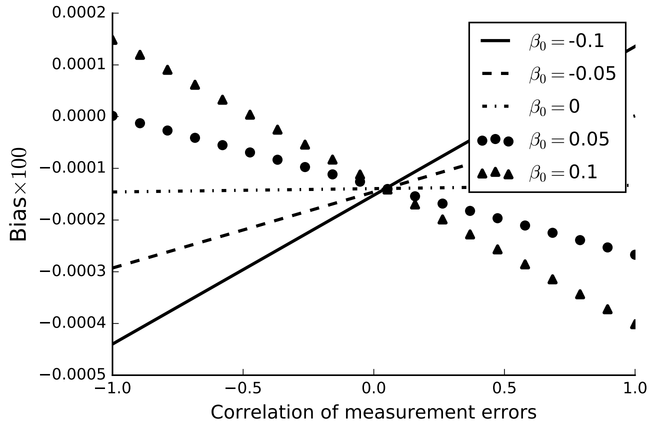
$$\operatorname*{plim}_{n \to \infty} \hat{\beta}_{OLS} = \beta - \Sigma_x^{-1} \Omega \beta,$$

where Σ_x is the variance-covariance matrix of X. Therefore, asymptotically, all of the regression coefficients are biased. The signs of the biases depend on three objects: Σ_x^{-1}, Ω, and β. In our application, only two variables in the matrix X suffer from measurement error problems (math and language test scores).

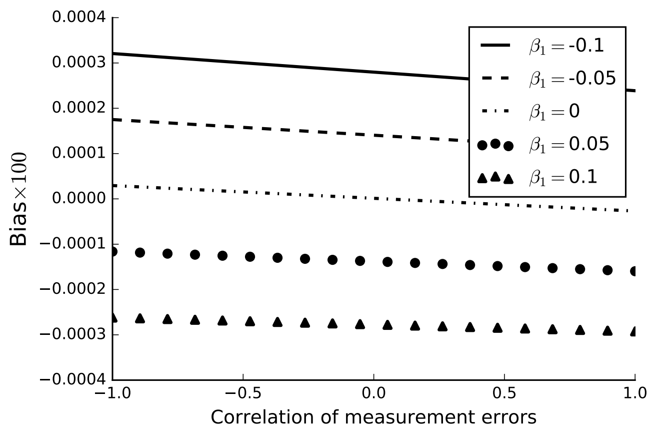
To analyze the consequences of measurement error in our regression models, we use our baseline specification and simulate the bias on the coefficient associated with private-fee-paying school dummy under different parameterizations. Figure D1 presents our results. The top panel depicts the bias as a function of the correlation between the two sources of measurement error and five possible values for β_0 (the coefficient associated with language skills). The bottom panel plots the relationship of the bias as a function of five possible values for β_1 (coefficient on math skills) and the same correlation, assuming $\beta_0 = 0.04$ (estimated value from our baseline regression). In both cases, to compute the bias we take the observed values of Σ_x^{-1}, set the rest of the parameters to their estimated values, and assume that the variances of measurement errors are equal to 0.99. Since by construction test scores have variances close to 1, assuming a variance of 1 for the measurement error puts possible bias values in an upper bound in absolute values.

For positive values of β_0 and β_1 (which is the most plausible scenario), the bias is likely to be negative. The bias is positive if $\beta_0 > 0$ (top panel) only when the correlation of

Figure D1: Bias of the private-school dummy coefficient



(a)



(b)

Notes: The figure shows the bias of the parameter associated with the private-school dummy sas a function of the correlation of measurement errors and for possible values of β_0 (panel a) and β_1 (panel b).

measurement errors is negative, an unlikely case given the nature of our variables.

Appendix E Employment regressions

Table E1 displays employment regressions. The goal of this analysis is to gauge the importance of movements into employment in our baseline regressions. The table maintains the format of Table 5 but uses 2013 employment as the dependent variable instead. This variable equals 1 if 2013 average monthly earnings is positive and 0 otherwise (that is, employment is zero if the individual did not work for the entire year). The unconditional difference in the employment rates between private-voucher and public schools is close to zero and statistically insignificant. The employment rate of private-fee-paying students is 5 percentage points lower than that of public-school students. The estimated regressions that do not control for test score yield school coefficients that are lower than the equivalent estimates of the model with test scores. In these regressions, the sign of the coefficient associated to test scores are negative, which means that these estimates most likely do not meet the conditions to obtain well-defined lower and upper bounds. At such, we cannot conclude much from these regression. However, the estimated lower and upper bounds are not statistically different from zero, both for the private-voucher and private-fee-paying cases. Moreover, the coefficients are tightly estimated around zero, suggesting that effects on employment should not be economically significant.

Table E1: The effects of high schools on employment probability (in percentage points)

Variables	Lower bound	Upper bound
Private-voucher	0.468	0.380
	(0.501)	(0.501)
Private-fee-paying	0.660	0.398
	(0.754)	(0.742)
Exogenous characteristics	Yes	Yes
Family Background	Yes	Yes
Test scores	Yes	No
Observations	111,395	111,395
Dependent mean (in %)	64.63	64.63

Notes: We show our estimated lower and upper bounds (equations 3 and 4) of the effects of schooling choices on 2013 employment (expressed in percentage points). We obtain the lower bound from a regression that includes math and language test scores. We estimate the upper bound from a regression that does not control for test scores. Both types of equations control for observed family and exogenous, individual characteristics (see Section 4 for details). OLS estimates of a regression of 2013 employment on school choices. Employment equals 1 if average monthly earnings is positive and 0 otherwise. Across columns, we show the coefficients associated to high school dummies adding different control variables: exogenous characteristics, family background, and test scores (see Section 4 for details). The baseline category is public school (coefficients on high school dummies represent impacts relative to the baseline). Test scores are expressed in standard deviation units. Standard errors (in parenthesis) are clustered at the school level (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).