Is Private Education Worth it? Evidence from School-to-Work Transitions*

Dante Contreras
Universidad de Chile

Jorge Rodríguez
Universidad de Los Andes, Chile

Sergio Urzúa
University of Maryland
and NBER

June 15, 2020

Abstract

We estimate the impact of private schools on labor market outcomes. Using rich longitudinal data from Chile, we find that attending private-fee-paying schools boosts average adult earnings by 100-140 dollars a month, equivalent to 15-22% return (relative to public schools). The advantage of private education also emerges at the intensive margin of investments. For private-fee-paying schools, a one percent increase in per-pupil spending leads to at least 23 extra dollars in adult monthly earnings. For public and voucher schools the effect is negligible. These results are robust to different specifications, including regressions that control for pre-labor market ability, school quality measures, individual-level education spending, and firm fixed-effects. Our analysis provides new insights into the association between school choice and income inequality.

Keywords: School choice, returns to education, income inequality.
JEL codes: I24, J24

*Dante Contreras, University of Chile; email, dcontrer@econ.uchile.cl. Jorge Rodríguez, Universidad de los Andes, Chile; email, jrodriguezo@uandes.cl; 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, Pontificia Universidad Católica de Chile, and SECHI. We benefited from comments and suggestions from Derek Neal, Magne Mogstad, Cristian Pop-Eleches, Tomás Rau, Loreto Reyes, and Cristián Dagnino. Dante Contreras thanks financing provided by the Center for Social Conflict and Cohesion Studies (CONICYT/FONDAP/15130009). We thank Pablo Sánchez and Raúl Navarrete for excellent research assistance. 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. Web Appendix: http://econweb.umd.edu/∼urzua/WebAppendixCRU.pdf.

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., 2017). However, potentially bigger issues might arise in settings in which access to high-quality education is granted only to those families with the financial resources to afford it (Murnane and Reardon, 2018). Under these circumstances, which are commonly prevalent in developing countries, private education might transmit early inequalities over the life cycle.

This paper investigates the role education systems play in shaping adult earnings disparities by exploiting novel and comprehensive longitudinal information. We focus on the long-run effects of attending private high schools in Chile. We exploit unique data combining multiple sources of administrative records. In particular, we gather information for the universe of Chilean students attending tenth grade in 2001 and recover their labor market trajectories using matched employer-employee data (up to age 28). In this way, we have the unusual opportunity to examine student-, family-, school- and firm-level information. We link individual data on school choices, family income and education, parental investments, expectations regarding students future performance, and standardized test scores with adult earnings. In addition, the data allows us to control for school-level financial investments—including tuition costs and municipality-level direct transfers to public and voucher schools—and to capture employment sorting across firms during adulthood, assessing new channels through which school choice might affect labor market outcomes in the long run.

By connecting the long-standing literature on ability bias (Griliches, 1977) with the recent evidence on the importance of latent ability (Hansen et al., 2003; Rodriguez et al., 2016), we exploit simple econometric principles to bound the impact of attending a high school type (private, voucher or public) on earnings more than a decade after graduation. The empirical strategy exploits a comprehensive set of control variables together with assumptions regarding different sources of biases to partially identify our treatment effects. More precisely, we take advantage of classic OLS biases from which we obtain upper and lower bounds of the effect of private education on earnings. First, we exploit the standard problem of omitted-variable biases (student ability) to construct an upper bound. In particular, as OLS estimates should overstate the impact of school type when ability is not properly controlled for, we show a bound can be recovered from regressions of earnings on school choice dummies. Second, to obtain a lower bound, we note that test scores, as imperfect proxies for ability, suffer from measurement error problems and/or are endogenous to school choice, two situations well-documented in the literature (Evans and Schwab, 1995; Heckman et al., 2006a). In this case, OLS estimates should underestimate the effect private schools have on earnings. We present multiple robustness checks supporting the assumptions behind this strategy.

Our results suggest positive and sizable earnings returns to private education. We find that attending a private-fee-paying instead of a public high school boosts average earnings by 100-140 dollars a month, equivalent to a 15-22% return. Importantly, by exploiting longitudinal employee-employer data we show the estimated effects are mostly driven by within-firm variation, which is to the best of our knowledge a new result in the literature. The effects remain positive even after controlling for tuition costs and school-level average test scores—which we interpret as a measure of school academic quality.[1] We also report that the returns to financial investments in private-fee-paying schools exceed those in other institutions: a one-percent increase in funding directed to 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. Lastly, we show that the long-term economic returns to voucher and public schools are of similar magnitude. This is consistent with previous findings for Chile (Hsieh and Urquiola, 2006; Bravo et al., 2010), giving grounds for our main interest in the comparison of private and public schools.

Our novel data enables us to contribute to the literature on the returns to school types. To a large extent, previous research has focused on short- or medium-term student outcomes, such as standardized test scores, high school completion, or college graduation rates at most.[2] A set of

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

[2]Dobbie and Fryer (2016), Sass et al. (2016), Jha and Polidano (2015), Binelli and Rubio-Codina (2013) and Bravo et al. (2010) are exceptions. The last two papers in this list are the closest to ours. However, their results are not directly comparable to those reported herein. In particular, unlike the analysis in Binelli and Rubio-Codina (2013) for Mexico, we do not implement an instrumental variable strategy to identify the impact of private schooling on labor market outcomes. Instead, we exploit the potential endogeneity of schooling choices to construct bounds for the parameters of interest. Bravo et al. (2010), on the other hand, estimate a structural model to assess the long-term effects of the Chilean voucher reform implemented in the early 1980s on labor market participation and

influential studies by Neal (1997), Grogger and A. Neal (2000), and Altonji et al. (2005) report positive but modest effects on test scores and relatively large effects on high school graduation for students attending private schools. The body of the literature exploiting voucher lotteries in the U.S. also reports relatively small achievement gains for students who were offered the voucher (Howell et al., 2002; Mayer et al., 2002; Krueger and Zhu, 2004; Wolf et al., 2007, 2010),[3] with one recent voucher program evaluation even showing negative effects (Abdulkadiroglu et al., 2018).[4] Closer to the institutional background of this paper, Angrist et al. (2002) and Angrist et al. (2006) document large effects on achievement and high school graduation of a voucher program in Colombia. Finally, 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.They find that attending a private school produces null 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. Even though many of these papers have produced convincing evidence—shedding lights on the causal effects of schools on students outcomes—focusing on standardized test scores may be problematic as improvements in test scores may not be the only outcome of an effective boost in the student's human capital.[5] Moreover, by moving beyond specific school types (such as voucher, charters, Catholic, among others), we broaden the assessment of the returns to schools to the study of a substantial part of the school spectrum within an educational system. Finally, by exploiting multiple sources of information, we unravel and evaluate the relative importance of different mechanisms. Specifically, we assess the extent to which the effects of school types on students' labor market outcomes can be determined by pre-labor market ability measures, across-firm sorting, school quality measures, and education spending (investments).

This paper also relates to the literature exploring the identification of treatment effects without relying on quasi-experimental variation (Altonji et al., 2005; Altonji and Mansfield, 2018; Oster, 2019). This set of studies analyze the information underlying the variation in observed variables to inform the importance of selection on unobservables and obtain a range of possible values for

earnings. Our analysis takes the reform as given, present reduced-form results, and focuses on the differences on earnings across school types.

[3] In a related literature, Angrist et al. (2016), Dobbie and Fryer (2011), Abdulkadiroglu et al. (2011), Angrist et al. (2012), 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.

[4] For a review of this literature, see Rouse and Barrow (2009), Epple et al. (2017), and Neal (2018).

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

the treatment effects that researchers are interested in. In connection to this literature, our analysis uses OLS biases yielded by the use of observed variables to obtain bounds for the effects of schools. Furthermore, our strategy allows for the possibility that, as students, individuals (and/or families) self-select into school types based on unobserved ability. This is particularly important given the evidence of sorting within the Chilean education system (Contreras et al., 2010). Our simple identification strategy is also connected to the analysis of Altonji and Mansfield (2018), who use group-level observed variables to bound the proportion of the variance in student outcomes explained by schools and neighborhoods.[6]

The rest of the paper is structured as follows. Section 2 briefly describes Chile's education system. Section 3 presents our strategy to identify the average effect of attending private school during high school on adult earnings. Section 4 describes the available data sets. Section 5 presents our main fundings and robustness checks. Section 6 concludes.

2 The Chilean education system

Historically, Chile has been a breeding ground for the academic debate on the role of competition in education markets, as large-scale, landmark reforms have been implemented in the country. This section provides a brief overview of the nation's education system and discusses its institutional features motivating this paper.

The Chilean education system provides families with three types of schools they can choose from: public, vouchers, and private schools. Public schools are ran by local governments, voucher schools are funded by the government and ran by private (for- and non-for profit) entities,[7] and private fee-paying institutions are funded and ran by the private sector. To a large extent, this organizational structure had already been in place during 20th century, but it was deepened after the wave of 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.

[6]We recognize that, even after conditioning on individual and group-level characteristics, selection into school types on the basis of unobserved variables would limit our chances of identifying meaningful causal effects. However, results reported in Section 5.1 suggest that selection based on unobservables might not be a critical concern in our setting.

[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).

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 voucher schools rose from 15% in 1981 to 56% in 2011. Even though voucher and public schools face a similar funding program, most of the private subsidized schools charge families with extra tuition (co-payment) during the period of analysis, while the opposite happens for public schools. This co-payment was in place before the 1980s reforms but it was a second reform (1993) that stimulated its use. Overall, the consequences of these reforms are considered a landmark example of the potential effects of a school choice scale-up.[8]

As explained in the introduction, the literature analyzing the impact of private education has mainly focused on short- or medium-term outcomes, such as achievement test scores. Table 1 presents this analysis for Chile. It reports average math and language test scores across high school types for the universe of tenth-graders in 2001, the sample studied herein. On average, private-fee paying schools clearly outperform voucher schools, which in turn outperform public schools. For math and language test scores, the table shows that private-fee-paying schools score 0.63 and 0.42 standard deviations higher than voucher schools. At the same time, average test scores of students attending voucher schools surpass those of public schools by 0.34 and 0.32 standard deviations in math and language, respectively.[9]

Figures 1 goes beyond short-term outcomes as it reports earnings distributions in 2013 by school type (reported in 2001) for the same sample of individuals used in Table 1. We observe that the distribution for former private high schoolers is skewed to the right with respect to the distributions of former public and voucher students. The differences in achievement test scores might be one

[8] See Hsieh and Urquiola (2006) for evidence of the effects of the Chilean voucher reform on test scores, repetition rates, and years of schooling. For a review of the evidence on vouchers in international contexts see Rouse and Barrow (2009), Bettinger (2011), and Urquiola (2016).

[9] A potential concern with this analysis, which could also affect our main findings, is the utilization of the school-type reported in tenth grade instead of, for example, the number of years attending private schools up to that grade. Since families are free to re-evaluate their schooling decisions every period, hence focusing on the information reported at a specific grade could be arbitrary. In the Chilean context, however, the enrollment by school type is highly persistence over time. For example, the analysis of administrative records for the period 2004-2006 indicates that out of the universe of eight-graders attending private school in 2004 (last year of primary school), 85.02% enrolled in private high schools in 2006 (tenth-grade). On the other hand, 91.13% of tenth-graders attending private schools in 2004 also reported attending private schools in 2004 (eight grade). This feature also emerges from the analysis of transitions during high school. The analysis of the universe of students attending this level in 2004-2007 indicates that conditional on attending a private school in tenth-grade, the probability of attending private school in twelve-grade is 94.60%. And more than 96% of those attending private schools in the last year of high school, also attended private high school in tenth-grade. Unfortunately, for these cohorts we cannot gather and merge labor market outcomes but they illustrate the stability of enrollment by school type, which is key for our analysis.

important factor behind these patterns. Figure 2 explores this possibility. It plots a non-linear regression of academic achievement and earnings by high school types. It shows that accounting for test scores does not fully remove the earnings gaps between those attending private-fee-paying schools more than a decade before and the rest of the students. Furthermore, the gap is increasing in test score levels, particularly among high-achieving students. Interestingly, 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 alternatives is not exclusively explained by differences in schools' production of academic skills. Moreover, the differences in slopes indicate that an increase in academic skills has larger earnings returns for students at private schools, a hypothesis we formally test later on.

Since schools operate under different market conditions, any comparison across school types 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.[10] This assertion is confirmed in Table 2, which presents the proportion of students with college-educated mothers and average family income by school types. The average income in private-fee-paying schools is 4.1 and 2.8 times larger than that of public and voucher schools, respectively. Moreover, the share of students with college-educated mothers in private-fee-paying schools (48%) exceeds that of voucher (20%) and public schools (11%). Table 3, on the other hand, documents average monthly educational expenditures across school types and math test score. We measure educational expenditures from tuition costs (for private-school students) and the value of the subsidy (for public and public-subsidized school) at the individual level. Students from private-fee-paying schools have the greatest amount of average investment ($165.4 a month), followed by voucher ($105.3) and public schools ($82.0). In addition, regardless of school type, extra funding (investments) for education is associated with better academic performance.

[10]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; Böhlmark and Lindahl, 2015)

Our purpose is to assess if attending these types of high schools have different impacts on adult earnings. However, identifying causal effects in this context conveys multiple obstacles, as the evidence from this section suggests. 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 potential biases in across-school comparisons.

3 From biases to bounds on treatment effects

Conceptually, our empirical strategy is grounded in a simple, yet general, model of school choices and long-term outcomes. For simplicity, consider three time periods. In the first period, the agent chooses to attend a high school of type $j \in \{1, ..., J\}$. At this stage, we do not specify this set, but in our application j can take three values—public, voucher or private—, with J being public. The decision is made based on individual's observed characteristics and unobserved ability, A_i, which forces us to discuss the identification of causal effects under self-selection based on unobserved heterogeneity. We interpret A_i as the latent outcome of early human capital investments, which include both inherent and environmental aspects. In the second period, academic achievement test scores ($\tilde{A}_i$) are observed. In the third period, and after completing formal education, labor market outcomes (w_i) are realized. For the sake of notational simplicity, we have not added time subscripts in what follows.

In general, we seek to identify the average effect of attending a school of type j on adult earnings, w_i, relative to a baseline category, J. This is a standard problem in the program evaluation literature. 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 school of type j, and 0 otherwise. In principle, this schooling choice problem could be solved for every grade. However, given the high persistence characterizing these decisions (see Section 2), we abstract from this dynamic complexity. Moreover, since our goal is to study school-to-work transitions, we focus on the type reported during high school. Thus, if we let w_i denote student's adult earnings, our parameter of interest can be written 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 different effects. First, school type j may rise individual's human capital, possibly increasing labor market productivity and earnings. Second, and conditional on inherited ability, attending a school of type j (relative to J) can have a direct impact on earnings through other channels, such as signaling (Weiss, 1995) or social networks (Zimmerman, 2019). Individuals choose school j anticipating its impacts (schooling decision is endogenous), breaking down the identification of κ_j from a simple comparison of earnings between those who attended j and the baseline alternative J.[11]

Our partial identification analysis starts by assuming a linear structural model for earnings of the form:

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

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 ability, and ϵ_i is the error term. In this equation, $S_{i,j} \perp\!\!\!\perp \epsilon_i \mid \boldsymbol{X}_i, A_i$ for all $j = 1, ..., J-1$. Thus, ϵ_i does not have information that could explain movements in $S_{i,j}$, after conditioning on observed control variables $\boldsymbol{X}_i$ and unobserved ability A_i.

As suggested in Heckman and Navarro (2004), one could estimate this equation under the presumption that the researcher has access to enough data to mimic the individual's information set at the time the schooling decision is made (i.e., a matching assumption). In this case, 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. In practice, however, researchers usually do not have access to, or have imperfect measures of, ability, which leads to biased results. In what follows, we use these biases to construct bounds for the parameter of interest, $\boldsymbol{\kappa}$.[12]

[11] 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.

[12] The Web Appendix 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 conditional independence assumption and illustrates how our bounds for the average effect of attending private high schools can account for selection on unobservables.

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) gainst of choosing j over the alternatives. If families behave in this manner, $E[v \mid \boldsymbol{S}, \boldsymbol{X}] \neq 0$ and OLS estimators would not identify the average earnings effect of attending a high school of type j. Formally, let S_{ij}^* be the residual of a regression of S_{ij} on the set of control variables and ρ_j^{OLS} be the OLS estimate of ρ_j, and $\boldsymbol{\rho}^{OLS}$ the vector collecting all estimated coefficients. The OLS bias follows the well-known omitted-variable bias formula, where $E[\widehat{\rho}_j \mid S_j^*, \boldsymbol{X}] = \kappa_j + \alpha\delta_j$ with $\delta_j = E\left[\frac{\sum_i (S_{ij}^* - \bar{S}_j^*)A_i}{\sum_i (S_{ij}^* - \bar{S}_j^*)^2} \mid S_j^*, \boldsymbol{X}\right]$, which converges to $\frac{Cov(S_{ij}^*, A_i)}{Var(S_{ij}^*)}$ for $j = 1, ..., J-1$. Thus, $\alpha \geq 0$ and $\delta_j \geq 0$ (for all j) are sufficient conditions for interpreting OLS biases from model (3) as upper bounds.

Despite the fact these two conditions are not directly testable in our setting, in Section 5 we provide suggestive evidence supporting them. In addition, we borrow the results and conclusions from the literature examining the role of unobserved ability to defend them. First, the vast research documenting non-negative effects of unobserved cognitive and non-cognitive ability on labor market outcomes (Heckman et al., 2006a; Borghans et al., 2008; Urzua, 2008), particularly on earnings in Chile (Rodriguez et al., 2016), suggests we can safely assume $\alpha \geq 0$.

When it comes to the second condition, we note it depends upon a positive association between the chances of enrolling in schools other than of type J (public) and A_i, which in our context encompasses the individual's human capital in the first period (high school). We connect this assumption to the extensive literature on the technology of skill formation (Cunha et al., 2005, 2010; Agostinelli and Wiswall, 2020). Under a positive productivity of investments for skills (conditional on $\boldsymbol{X}$), one would obtain increasing latent ability levels as a function of previous human capital investments (Cunha and Heckman, 2007). Based on this channel, if families had preconceived the enrollment in voucher or private schools as an investment decision, one might expect a positive covariance between A_i and S_{ij}^*. The idea of non-public schools being more effective in fostering skill formation would further justify $\delta_j \geq 0$ for $j = 1, ..., J-1$.

The lower bound: Test scores as proxies for ability. With data on individual-level achievement 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 at first glance this specification might be interpreted as a direct approximation to our structural equation, including $\tilde{A}_i$ in the wage regression generates a new potential source of concern. As shown in Hansen et al. (2003) for the U.S., achievement test scores can be conceptualized as constructs affected by ability and schooling at the time of test. In our case, this implies that $\tilde{A}_i$ could be thought as a function, if not the result, of the school choice. To establish the consequences of this phenomenon, let us assume:

$$\tilde{A}_i = A_i + \boldsymbol{S}_i'\gamma + v_i,$$

where, without loss of generality, we normalize the coefficient associated with A_i to one. v_i is the error term. Thus, the schooling coefficients are such that $\boldsymbol{p} = \boldsymbol{\kappa} - \alpha\boldsymbol{\gamma}$. Intuitively, by including $\tilde{A}_i$ in the wage regression we absorb part of the effect of school type on earnings, as schools might have indirect effects on earnings through their contribution on raising academic achievement. If individual's ability has a non-negative impact on earnings ($\alpha \geq 0$) and the elements of $\boldsymbol{\gamma}$ are non-negative (similar logic as for the upper bound but now exploiting the association between school choice and observed test scores), then the OLS estimates of school types in equation (4) should be biased downward, delivering plausible lower bounds.[13,14] Lastly, we rely on the assumption that the error terms in equation (2) are conditionally independent of the set of school dummies. Although

[13] Altonji and Mansfield (2018) 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 that paper in that the authors use observed characteristics to account for selection based on unobservables.

[14] Another potential source of bias is measurement error. If A_i is measured with error, then the OLS estimators of all regression coefficients could be biased. Following standard econometric results, one can show that the coefficient of the error-ridden variable is biased towards zero—preventing the identification of α from $\tilde{\alpha}$—, and the sign of the biases of the other parameters can be estimated consistently. In the context of equation (4), these biases would depend on the signs 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 constructed using the available data. In our empirical application, we confirm that the signs of the relevant biases are such that the measurement error problem reinforces the interpretation of the OLS estimator as a lower bound.

strong, we do not find this assumption too restrictive in our case for two reasons. First, we have at our disposal an unusually large and detailed set of control variables—including achievement test scores. Second, unlike the standard conditional independence assumption—which depends heavily on the set of observed control variables—our partial identification framework does account for selection based on unobserved ability A_i. Section 5 presents a series of empirical robustness checks showing that selection based on unobservables is not likely to contaminate our partial identification empirical analysis.[15]

4 The data

Our empirical analysis exploits multiple sources of administrative information. The database is unique in combining administrative records containing student-level achievement test scores, background characteristics and school-level information at tenth grade (2001) with matched employer-employee data describing individual-level labor market trajectories for the period 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 education authority 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 variables ($\boldsymbol{X}$ in equations (2)-(4)). Here, we include gender, previous attendance to pre-primary education, a dummy variable indicating grade retention, and region dummies (out of 13 regions across the country). We also include two sets

[15] The method of instrumental variables could also be pursued to identify $\boldsymbol{\kappa}$. For instance, Binelli and Rubio-Codina (2013) exploit changes in the availability and size of public and private high schools across states and over time in Mexico to construct instruments, finding positive effects on wages conditional on college completion. 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. Another alternative is a matching estimator (Heckman et al., 1997). Our web appendix presents evidence supporting its key testable assumption (common support), showing that the probability of attending a private schools for public-school students traces almost all of the unit interval. However, conditional on observable characteristics, this method assumes no selection based on latent ability.

of variables that trace out parents' answers to the questions "At home, during a normal week, at what frequency does your child study?" and "Does your child have a job outside schools?"[16] As proxies for family background, we include mother's and father's education, per-capita household income,[17] number of books at home, mother's and father's age, dummies for whether the mother or father (or both) live with the student, parents' answer to question " how far along do you think you pupil will get in school?," whether parents attend school meetings, and if parents are aware of school's SIMCE scores.

In addition, from the SIMCE database we are able to recover information on private and public educational expenditures. We use this data to estimate the impact of educational expenditures on earnings (see Section 5.6). For this analysis, our goal is to obtain the monthly expenditures associated to the student's high school attendance by using self-reported tuition costs and other private expenditures.[18] In the case of public schools, we include in our regressions direct transfers from local municipalities to public schools.[19] These transfers constitute the main source of funding for public schools. For students in voucher schools, we compute the voucher amount that the schools receive for having the student enrolled.[20] Lastly, we complement this information with self-reported monthly tuition costs contained in the SIMCE files.

We use the Unemployment Insurance (UI) administrative database to gather information on students' adult earnings. We observe monthly gross earnings for 2004-2013, for all of those who have reported at least one formal job contract from 2002 up to 2013 (October). Students who have not had a job up to 2013 are still considered in the final data, where their earnings 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.

[16]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).

[17]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.

[18]For students attending voucher schools with co-payment, the monetary value of the subsidy depends on the amount of monthly average copayment. To compute the voucher amount, we use administrative records of 2001 monthly average payments for shared-funding schools.

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

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

[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.

In most of our empirical analysis, we use our last available year of earnings (2013), meaning that, by that year, the average former student has turned 29 years of age. Evidence from the United States and Sweden indicate that individuals' ranking in the income distribution when they are in their early 30s is highly correlated to the ranking in the distribution of lifetime income (Chetty et al., 2014a; Nybom and Stuhler, 2017; Chetty and Hendren, 2018). This evidence supports our choice of using students' earnings in their late 20s instead of 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 missing or duplicate national identifiers. 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 4 presents descriptive statistics for all of the variables that we use in our regressions, whereas Table 5 compares statistics on key variables between the original and final sample. The results indicate that the distribution of students across school types remain almost the same, although baseline ability levels are potentially higher in the final sample.

The main limitation of this setting, despite its comprehensiveness, is the focus on tenth-graders only. Fortunately for us, as explained in Section 2, enrollment in either private or public schools is highly persistence across grades in Chile. This fact alleviates potential concerns arising from the strategic behavior of families revising their decisions dynamically. It also suggests our estimates might capture effects of studying many years in high school of type j. Moreover, in Section 5.2 we control for an imperfect proxy for the years of enrollment in public, voucher of private schools. Our main findings are robust to this addition.

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 voucher or private high schools relative to the public tuition-free alternatives and then document the presence of heterogeneous impacts. Before turning to our main results, we present various robustness tests confirming the validity of our identification assumptions.

5.1 Assessing the identifying Assumption

As shown in Section 3, the conditions for obtaining lower and upper bounds for $\boldsymbol{\kappa}$ depend on a series of assumptions. In this section, we assess the likelihood that these conditions are met, in a setting where we estimate the effects of attending private relative public school attendance on earnings (public school is our baseline J). Specifically, we study the empirical relationship between test scores and earnings and between test scores and school choices, and evaluate the plausibility of the conditional independence assumption imposed on equation (2).

Upper bounds. To check the validity of our upper bias for the causal effect on our parameter of interest, we must estimate the sign of $Cov(S_j^*, A_i)$, where S_j^* is a residualized school choice dummy. Of course, a major difficulty in estimating $Cov(S_j^*, A_i)$ is that we do not observe A_i; yet 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 voucher dummy, the correlations of the corresponding residual with math and language equal 0.04 and 0.05, respectively. These positive values support our interpretation of the bias.[22]

In addition, we present a stability analysis, which provides further support of the intuition behind the upper bounds. Appendix A.1 shows that every additional control variable that is included in the regression leads to a fall in the estimated effects of private and voucher school over public schools. Overall, we find unambiguous evidence suggesting that the OLS estimation of equation (3) delivers an upper bound for the coefficients of both private school dummies.

Lower bounds. A necessary condition for having a valid lower bound is the independence of the error term and school dummies in equation (2), conditional on ability A_i and $\boldsymbol{X}_i$. Suppose $A_i = \tilde{A}_i$, thus controlling for test scores would secure the conditional independence between high

[22] As indirect evidence supporting the preconceptions securing $\delta_j \geq 0$ for voucher and private schools, we examine the answers to the question "What aspects of the school did you consider to decide whether to enroll your child in this school?" included in the parental questionnaire of SIMCE 2001 files. Out of a total of 173,359 families, only 26.82% do not list any of the aspects linked to school's academic performance (prestige, school's performance in previous SIMCEs or the college admission test (PAA), quality of its teachers). Among families with children attending public schools the figure is 30.23%, voucher schools 25.39% and private schools 16.93%. Although these proportions come from self-reported information after the enrollment decision was made, they suggest preconceptions aligned with our interpretation of the biases resulting from the OLS estimation of equation (3) as the foundations of upper bounds.

school choices and the unobserved components of the regressions, and the identification of attending effects a private instead of a public school. We test the plausibility of this assumption in series of exercises. In the first test, we evaluate the influence of selection bias by using a previously omitted baseline characteristic as a dependent variable. Appendix A.2 shows that the estimated "effect" of attending private or voucher over public schools on (the previously omitted variable) grade retention is economically insignificant (although statistically significant) when we consider the full set of controls plus test scores.

We also evaluate selection on unobservables by studying how sensitive are the estimated school effects to adding control variables. A lack of sensitivity to some control variables can be related to a lesser influence of selection on unobservables on the estimated coefficients (Altonji et al., 2005; Oster, 2019). Besides showing that the estimated coefficients tend to fall as we add control variables, Appendix A.1 also shows that these estimates hardly move when adding different control variables. Notably, when we include test scores, estimated coefficients once again fall, plausibly because those controls absorb away part of the causal effect we are looking for. However, the R^2 of all of these regressions are in the low range: at most, we estimate an R^2 of 0.08. A relatively low R^2 implies that coefficient stability might not the strongest indicator of lack of selection on unobservables (Altonji et al., 2005; Oster, 2019).

In a third exercise, we estimate a selection model to assess whether we find support for a statistically and/or economically insignificant correlation between the error terms of the selection and earnings equations. Appendix A.3 presents these tests. In most cases, this correlation falls nearly zero when we include the full set of controls. Overall, the three previous exercises suggest that, even though we cannot rule out a certain degree of selection-on-unobservables bias, it is unlikely to modify our assumption leading to the lower bound

A second condition for the validity of our lower bound is having an attenuation bias due to the introduction of noisy measures of ability. To test this assumption, we study the signs of the estimated correlations between the school dummies and the proxies for ability. Appendix A.4 analyzes the case of multiple regressors subject to measurement error in the context of our research question. We show that—for a range of sensible values of unknown parameters—the inclusion of test scores will generate negative biases on the relevant regression coefficients. This result further justifies the validity of our estimated lower bounds for the earnings effects of private school attendance.

5.2 The effects of attending private high schools

Column (1) of Table 6 presents the lower and upper bounds for the impact of high-school types on earnings.[23] These bounds come from the estimated regressions equations (3) and (4). We also present bounds for our partially identified parameters based on Imbens and Manski (2004), which cover the causal effects of interests with probability 95%. We construct this confidence interval using standard errors that are clustered at the school level. Finally, the table reports bootstrapped p-values for the null hypothesis that upper and lower bounds are equal against the alternative that the first exceeds the second one.

To get a sense of the estimated intervals, note that the unconditional average monthly adult earnings of individuals who attended private-fee-paying schools are 270 dollars higher than those reported by individuals who attended public institutions. This difference most likely overestimates the true causal effect of attending a private-fee-paying school as students coming from relatively wealthier, more educated families, have a higher probability of attending private schools. If we control for our wide set of control variables, including various family background characteristics, the earnings differential of private-fee-paying schools falls to 140 dollars. This differential is our upper bound. 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, a range consistent with the wage returns to private schools reported for the US (Brown and Belfield, 2001). The Imbens and Manski (2004) bounds indicate that the partially identified parameter is statistically significant at the 5% level.

From the results in Column (1) of Table 6 we also assess the impact of voucher schools. Previous studies analyzing the issue, but using short-term achievement test scores in Chile, reached mixed

[23] Our framework have the potential to estimating causal effects of some of the control variables. Readers can see full table of estimated coefficients in Appendix B. However, we remark that we have not done any robustness check to see if conditions to identify upper and lower bounds are reasonable in the case of control variables. Additionally, lower and upper bound equations might not structured properly so as to identify causal effects of some of the control variables. To take the example of mother's education, the upper bound equation should not have any variable that is itself an outcome of parental education (such as school choices and family income).

conclusions (Carnoy and McEwan, 2000; McEwan, 2001; Carnoy and McEwan, 2003; Hsieh and Urquiola, 2006; Contreras et al., 2010; Elacqua et al., 2011; Gallego, 2013). Unlike most of the available evidence, we document the long-term effects of these alternatives. Our range of estimates for the returns to attending voucher versus public schools on adult earnings is between 10 and 22 dollars a month. The Imbens-Manski bounds indicate that the partially-identified parameter is different from zero under a 5% level, although the lower bound in this confidence interval is fairly close to zero. These findings are in line with the small point estimates reported in Bravo et al. (2010), from which the authors conclude that the Chilean voucher reform implemented in the early 1980s did not impact overall mean earnings. Since the estimated returns to voucher versus public school attendance is likely to be economically small, in what follows we mainly focus on the contrast between private-fee-paying and public schools.

Interestingly, these findings emphasize how schools may affect dimensions of human capital that are beyond academic achievement. This since the regression estimates from Table 6 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 each alternative 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, 2019). More generally, the non-zero coefficients are 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.[24]

To describe the evolution of these returns over time, Figure 4 reports the upper and lower bounds on the effects of private (panel a) and voucher (panel b) high-school attendance on adult earnings for different years. Each line shows the coefficient associated with the high school dummy, with the line with circle markers corresponds to the estimated parameter from regressions without test scores (the upper bound) while the one with square markers depicts the estimated coefficients in regressions that control for test scores (the lower bound). The shaded area shows Imbens and Manski (2004) confidence intervals. Each panel displays the estimated effects of average monthly earnings from 2010 until 2013. We choose to start this analysis from 2010 as we expect that most

[24] Using data from the Los Angeles Unified School District, 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.

of our sample have completed post-secondary studies by then.

The impact of private-fee-paying schools (panel a) increases at a decreasing rate over time, with statistically significant results (Imbens-Manski 95% confidence intervals do not cover zero). In contrast, for most of the years, the effects of voucher schools (panel b) remain close to zero. For these estimates, the Imbens-Manksi confidence intervals at the 95% level locate only slightly above zero, so we are cautious about claiming positive earnings effects of voucher over public schools. Nonetheless, our results confirm the positive returns to private education.

School-level characteristics. The rich longitudinal information enables us to assess the importance of two potential mechanisms driving the long-term effects of private-fee-paying high schools: school academic quality and funding (investments). Moreover, by controlling for these school-level characteristics, we test whether the estimated effects hold under a tighter identification strategy. If families have preferences for school-level achievement or investments define school's productivity, then not account for these would not allow us to partially identify causal effects (the lower bound).[25]

We begin by testing whether the fact that private schools have better educational "quality"—by which we mean that the school has a larger conditional SIMCE average—explains away the effect of private education on earnings. Following the logic for obtaining upper bounds, we first exclude school-average SIMCE test scores as a control, as school quality might be correlated with individual-level achievement (this leads to the same upper bounds reported under column (1)). We do control for school-level SIMCE in the lower-bound regression. The results from this exercise are presented in column (2) of Table 6. The lower bound for the estimated impact of private-fee-paying versus public falls from 0.99 to 0.54 dollars after controlling for school average SIMCE. This result is suggestive, although not conclusive, of the fact that school academic quality can explain part of the effects of private schools on earnings. In any case, the statistically significant coefficient does imply that private schools have positive long-term impacts, even after accounting for school academic-level performance.

Second, we test for whether funding differentials across school types can help to explain the impacts of private school on earnings. As in the previous case, we exclude the variable accounting

[25]See Meghir and Rivkin (2011) for a related discussion.

for school funding from the upper-bound equation, and include it in the lower-bound equation. Table 6, column (3), presents the results. The estimated coefficients associated with private-fee-paying dummy remain almost unchanged relative that the bounds reported under column (1). We conclude that educational spending have a lesser role in explaining the effects of private schools on earnings. We revisit these results n our analysis of heterogenous returns (Section 5.6).

The returns to a single year of private education. Given the characteristics of our data, our previous findings do not necessarily capture the effect studying in a private school at 10th grade only, but possibly studying in this type of school for many years before and after. To assess possible effects of an additional year at private schools, we estimate:

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

where NS_{is} corresponds to the number of years that student i attended 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} from the information contained in the SIMCE files. In particular, we use parents' answers to the following question: "For how many years have your pupil attended this school?." 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 7 presents our estimated bounds on the average effect of spending one year in high school type s (private or voucher) relative to spending that year in a public school. We estimate that the effect of spending one year in a voucher instead of a public is bounded between 4 and 2 dollars. However, the Imbens-Manski bound covers 0 as a possible value falling in the 95% range. The effect of spending a year in a private school, on the other hand, is bounded between 18 and 24 dollars a month. This bound implies a statistically significant effect. Taking this estimate at face value, if a student spends 10 years (from first to the tenth grade) at a private instead of a public school, it earns an additional 180-240 dollars of monthly earnings. These estimates are difficult to interpret as the actual effects of spending a year in particular high school type.[26] Nonetheless, it seems that

[26]Despite the evidence of high persistence in enrollment by school type reported in Section 2, one caveat of the present analysis is that a student could have switched from public school A to public school B. In this case, the number-of-years variable 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. For this reason, we are careful in interpreting ρ_s as the impact of spending one year in a particular public school.

the baseline regressions from Table 6—which use a dummy variable for tenth-grade attendance at school s—are not too far from the potential earnings effects of spending from first up to the tenth grade in voucher or private high schools. This evidence is consistent with the high persistence of school choices throughout the schooling years.

5.3 Exploiting employer-employee matches

Individuals can take-up the economic returns to schools in, at least, two ways. First, individuals can be matched to firms where the earnings returns to those skills acquired in school j are relatively large. Second, the returns to school types can vary within firms due to various reasons; for example, studying in a private school increases the probability of having a position at the highest levels of hierarchy in a firm. In what follows, we exploit our matched employee-employer data to assess the quantitative relevance of these channels through which school choice affect earnings.

Only for the analysis of this sub-section, 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 as in our previous results, to avoid situations in which workers change location or firms within a year (linking a student to a firm using annual earnings would make the worker-firm match arbitrary for those who had more than one job in that year). Also, since we add firm fixed-effects to our regressions, our sample considers only individuals reporting at least one job. Thus, we end up with 55,858 workers distributed across 21,599 firms. Average monthly 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 (650 dollars), since the latter figure includes "zeros," capturing unemployment spells for the whole year.

As for the analysis of school-level characteristics, we can explore a potentially more credible identification strategy. If families and students make optimal choices based on unobserved traits unrelated to academic achievement, then our bounds may not contain the causal effects of private schools on earnings. Nonetheless, if these unobserved traits are correlated to earnings returns across firms, then adding firm-fixed effects would eliminate any associated bias. There is evidence showing that an important portion of earnings heterogeneity is actually accounted by between-firm variation (Abowd et al., 1999; Card et al., 2018), thus the importance of controlling for firm-fixed effects in our regressions as a way of picking up any residual unobserved variation beyond academic skills.

Table 8 presents the estimated lower and upper bounds after accounting for firm fixed-effects. The findings suggest a significant role for within-firm sorting and a lower incidence of across-firm sorting in explaining the impact of high-school type on earnings. In particular, the firm fixed-effects regressions cut down the estimated bounds by almost 30% for the private-school bounds and by approximately 20% for the voucher-school bounds, yet the null-effect hypothesis is still rejected at the 95% level. We conclude that a significant proportion of the returns to high-school type 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 individuals working in the same firm. It is worth mentioning that the between-firm variation is not negligible and coincides with literature suggesting it explains approximately 20% of total wage variation (Abowd et al., 1999; Card et al., 2018).[27]

5.4 Achievement test scores, school types and adult earnings

Pre-labor market skills are strong predictors of labor market outcomes.[28] However, to the best of our knowledge, it has not been documented to what extent this association depends on the type of school the individual was enrolled in. In what follows we seek to fill this gap by using SIMCE test scores (tenth grade) as proxies for early skills and taking advantage of our novel database of school choices and earnings.

We first document the estimated average effects of math and language SIMCE test scores on average monthly earnings (again computed on an annual basis), irrespective of school choice (therefore, no reference to bounds is needed). Table 9 presents the effect of test scores on earnings. The results in column (1) mimic those from Table 6, only this time we present the estimated coefficients associated with math and language. 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

[27]Table 8 uses October 2013 earnings as the dependent variable whereas our previous results use annual average earnings. To assess if this change affects our baseline estimates, we compare the regressions without fixed effects from Tables 6 and 8. If we measure the estimated bounds as a percentage of the overall average of earnings in each case (recall that in the fixed-effects estimates the baseline earnings mean is higher), the bounds of the regression with no fixed-effects equal 15 and 22% for both tables, suggesting that working with monthly earnings does not affect our quantitative conclusions.

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

the average. The analogous estimate with respect to the language test score is lower (7.8 dollars a month). Figure 3 complement these findings. It reports nonlinear effects of SIMCE scores on earnings obtained from the regression coefficients associated with dummy variables indicating whether the student belongs in different math or language test scores quantiles. The returns 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 9 presents heterogeneous effects of the math test scores on earnings by high-school types. We focus on math since this variable has larger effects on earnings than language. We estimate the coefficients associated with the interaction of math and school dummies. As the set of regressions control for individual test scores, these estimates are interpreted as lower-bounds. On average, for students enrolled in public schools, a one-standard-deviation increase in math rises earnings by 94.8 dollars a month (15% relative to the average). 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. In this case, a one-standard-deviation boost in math test score increases monthly earnings by 153.6 dollars—which corresponds to the sum of the baseline impact (94.8 dollars) with the additional private-school effect (58.8 dollars). This impact—a 24% earnings return—is 61% larger than the effect reported for public- and voucher-school students. Even though we lack data to directly test this hypothesis, our results are consistent with complementarities between academic and non-academic factors. The fact that the effect of test scores on earnings is larger for private schools could mean that students at these schools foster other traits —produced within the family and/or school—which makes the return to math larger.

5.5 School-level academic performance and labor market outcomes

Next we provide additional evidence in line with the hypothesis that high-school types affect the returns to academic achievement. We study the effect of school-level achievement on earnings and explore heterogeneous effects across high-school types. To this end, we consider the following regression models:

$$w_i = \beta_1 \Lambda_i + \Lambda_i \times \boldsymbol{S}_i' \beta_2 + \boldsymbol{X}_i' \alpha + \varepsilon_i, \tag{6}$$

$$w_i = \widetilde{\beta}_1 \Lambda_i + \Lambda_i \times \boldsymbol{S}_i' \widetilde{\beta}_2 + \boldsymbol{X}_i' \widetilde{\alpha} + \widetilde{\beta}_3 \tilde{A}_i + \widetilde{\varepsilon}_i, \tag{7}$$

where, as before, $\boldsymbol{S}_i$ represents a vector of high school dummies, $\tilde{A}_i$ denotes individual-level test scores, and $\boldsymbol{X}_i$ includes individual and family characteristics. Λ_i represents the average high-school level SIMCE score.[29]

The effect of school academic performance is bounded between the estimates of β_1 and $\widetilde{\beta}_1$ in equations (6) and (7), respectively. Intuitively, identification follows the same argument used to identify the average effect of school types. The estimated coefficient on Λ_i, from equation 6, is likely to be upward biased as we do not control for individual ability. By controlling for individual-level test scores, on the other hand, we mitigate the potential consequences of endogeneity in the school performance measures at the price of absorbing part of the impact of school academic achievement on earnings and introducing attenuation bias in the estimands (lower bound). Provided that the assumptions to identify upper and lower bounds hold, we can interpret our findings in this regressions as the earnings effects of average school performance. The interacted effects of school-level academic performance and the school dummies are identified analogously, by comparing the associated coefficients in the two pairs of regressions.

Table 10 presents the estimated effects of school academic quality on student's earnings. Column (1) shows that average academic performance at school-level has a statistically significant long-term impact: a one-standard-deviation increase in school's math average increases student's future earnings by an average of 140-213 dollars a month (a return of 22-33% with respect to the overall average). The same pair of regressions estimate negative effects of increasing school-average language scores. However, in this case, we cannot reject the null hypothesis that lower and upper bounds are equal against the alternative that the upper is larger than the lower bound (p-value$= 0.544$). This last result indicates that we are not able to partially identify the parameter of interest.

Column (2) in Table 10 reports the results allowing for school-interactions. The interacted

[29] For simplicity, we assume $\boldsymbol{A}_i$ and Λ_i are scalars. However, our regressions include both math and language test scores.

partially-identified coefficients are statistically significant according to the 95% Imbens-Manski bounds. The estimates indicate that the effects of school-average math scores is higher for private schools. However, the estimated bounds in the interacted coefficients are tight enough as to not reject the null hypothesis of equality in upper and lower bounds, indicating lack of identifying power. Overall, we must be careful in interpreting these results as causal effects.

5.6 Education spending during high school and gaps in earnings

The results displayed in Table 6 suggest that differences in per-pupil spending across school types cannot explain the advantage of private schools over the alternatives. Can they, however, amplify the edge? To address this question we gather private and public sources of information containing individual-level education spending.[30] We then use our framework to identify the average effect of school types taking into account the intensive margin of investments.

In practice, we estimate the long-term effect of an extra dollar of spending on adult earnings accounting for high school enrollment in private, voucher, and public schools. Formally, if we let c_{is} be total spending in education for student i at school s, we estimate:

$$w_i = \gamma_1 \log(c_{is}) + \log(c_{is})\boldsymbol{S}_i'\gamma_2 + \boldsymbol{X}_i'\alpha + u_i \tag{8}$$

$$w_i = \tilde{\gamma}_1 \log(c_{is}) + \log(c_{is})\boldsymbol{S}_i'\tilde{\gamma}_2 + \boldsymbol{X}_i'\tilde{\alpha} + \tilde{A}_i'\tilde{\gamma}_3 + \tilde{u}_i \tag{9}$$

where γ_1 captures the effect of an 1% increase in educational expenditures on adult earnings. Using the argument from before, the causal impact of c_{is} on earnings is bounded between the OLS estimates of γ_1 and $\tilde{\gamma}_1$.[31]

Table 11 presents the estimated lower and upper bounds from equations (8) and (9). Financial resources at school increases future monthly earnings by 10-30 dollars, on average. However, we find that these returns differ by school type. In fact, most of the earnings average effects of financial resources can be attributed to those students at private-fee-paying schools: a one percent increase

[30] We consider multiple sources of information containing public and private monetary resources allocated to education, including individual-level out-of-pocket education expenses, tuition costs, and public subsidies (vouchers). We do not account for other types of investments, such as parental time and others non-pecuniary investments in human capital. Relative to our baseline sample, we exclude 455 students from public schools (0.6% of the sample) with zero educational expenditures.

[31] c_{is} varies by student as 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.

in monthly educational investment increases monthly adult earnings by 23 to 34 dollars for private-, between 3 to 5 dollars for voucher-, and 4-9 dollars for public-school students (see columns 2 and 3). 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) and with recent evidence suggesting heterogeneous effects of school spending (Jackson, 2018).

The fact that some schools are better than others for every dollar invested in education conveys important public-policy challenges. To illustrate the quantitative implications of this argument, consider the following thought experiment. Suppose that a policy-maker wishes to close the adult earnings gap between public and private-fee-paying students.[32] The evidence from charter schools in the United States suggests that part of their success comes from additional educational resources—better teachers, more instructional hours, among other practices (Angrist et al., 2013; Fryer, 2014). Recent evidence also suggests that school funding has positive effects on student achievement (Jackson, 2018). Hence, it makes sense for policy makers to tackle future earnings inequality by injecting more money into public schools. How much would it cost to close the earnings 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-school student. This exercise leaves fixed the exogenous observed and unobserved characteristics of these two hypothetical students, including their school choice. Using our estimated coefficients (bounds) 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 above the actual monthly funding allocated to 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.

6 Conclusions

This paper explores the impact of attending different high-school types on adult earnings. We take advantage of unique longitudinal information combining administrative records on individual

[32] As a reference, the 2013 average monthly earnings of 2001 public and private-fee-paying students in Chile correspond to 571 and 845 dollars, respectively.

pre-labor market test scores, family background characteristics, school-level variables (including tuition costs and public subsidies), and labor market outcomes. Our empirical strategy allows for self-selection into school types, explores the role of sorting across firms, and identifies heterogeneous effects across individual- and school-level characteristics.

We find that private education in Chile has long-lasting effects. Our results indicate that attending a private-fee-paying high-school boosts average adult earnings by 100-140 dollars a month, equivalent to 15-22% return (relative to public schools). Through different robustness checks, we show these long-term effects are not solely explained by a superior academic quality and/or additional financial resources (OECD, 2010). Second, we find that equivalent increases in individual test scores and funding yield more pronounced increases in earnings for students from private than public institutions. These heterogeneous effects—an indication of relative efficiency differences (Carnoy and McEwan, 2000)—might come from complementarities in the school-specific technology of skill formation (e.g., "soft" and cognitive skills are complements in the production function, and the degree of complementarity is stronger for private-paid schools). Alternatively, they might emerge from a larger concentration of students with better future labor market prospects—from more educated and wealthy families—which, through peer effects, could further boost adult earnings of students with better academic performance. Future research should focus on disentangling these two mechanisms.

Overall, the analysis reveals a major challenge for public policies as it confirms earnings disparities during adulthood are, at least partly, shaped by schooling decisions made early on. In this context, this paper represents a new attempt towards empirically understanding the link between the 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.

_ , Parag A. Pathak, and Christopher R. Walters, "Free to choose: Can school choice reduce student achievement," *American Economic Journal: Applied Economics*, 2018, *10* (1), 175–206.

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

Agostinelli, Francesco and Matthew Wiswall, "Estimating the Technology of Children's Skill Formation," 2020. NBER Working Paper No. 22442.

Altonji, Joseph G. and Richard K. Mansfield, "Estimating Group Effects Using Averages of Observables to Control for Sorting on Unobservables: School and Neighborhood Effects," *American Economic Review*, 2018, *108* (10), 2902–2946.

_ , 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.

Angrist, Joshua D., Sarah R. Cohodes, Susan M. Dynarski, Parag A. Pathak, and Christopher R. Walters, "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics*, 2016, *34* (2), 275–318.

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

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.

Binelli, Chiara and Marta Rubio-Codina, "The Returns to Private Education: Evidence from Mexico," *Economics of Education Review*, 2013, *36*, 198 – 215.

Böhlmark, Anders and Mikael Lindahl, "Independent Schools and Long-run Educational Outcomes: Evidence from Sweden's Large-scale Voucher Reform," *Economica*, 2015, *82* (327), 508–551.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 2008, *43* (4).

Bravo, David, Sankar Mukhopadhyay, and Petra E. 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.

Brown, Celia and Clive Belfield, "The Relationship between Private Schooling and Earnings: A Review of the Evidence for the US and the UK," 2001. Occasional Paper N. 27, National Center for the Study of Privatization in Education Teachers College, Columbia University.

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.

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," *The Quarterly Journal of Economics*, 02 2018, *133* (3), 1163–1228.

_ **, 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), 1349–1368.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, May 2007, *97* (2), 31–47.

_ **,** _ **, Lance Lochner, and Dimitriy Masterov**, "Interpreting the Evidence on Life Cycle Skill Formation," *Handbook of the Economics of Education*, 08 2005, *1*.

_ **, James J. Heckman, and Susanne M. Schennach**, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

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*, 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*, 2017, *55* (2), 441–492.

Evans, William and Robert Schwab, "Finishing High School and Starting College: Do Catholic Schools Make a Difference?," *Quarterly Journal of Economics*, 1995, *110*, 941–974.

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," *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," 2017. NBER Working Paper No. 23479.

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 1977, *45*, 1 – 22.

Grogger, Jeff and Derek A. Neal, "Further Evidence on the Effects of Catholic Secondary Schooling," *Brookings-Wharton Papers on Urban Affairs*, 01 2000, *1*, 151–193.

Hansen, Karsten, James Heckman, and Kathleen Mullen, "The Effect of Schooling and Ability on Achievement Test Scores," *Journal of Econometrics*, 06 2003, *121*, 39–98.

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), 30–57.

_ **, 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.

_ **, 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.

Howell, William G., Patrick J. Wolf, David E. Campbell, and Paul E. Peterson, "School vouchers and academic performance: results from three randomized field trials," *Journal of Policy Analysis and Management*, 2002, *21* (2), 191–217.

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), 1477–1503.

Imbens, Guido W and Charles F. Manski, "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 2004, *72* (6), 1845–1857.

Jackson, C. Kirabo, "Does School Spending Matter? The New Literature on an Old Question," 2018. NBER Working Paper No. 25368.

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.

Jha, Nikhil and Cain Polidano, "Long-Run Effects of Catholic Schooling on Wages," *he B.E. Journal of Economic Analysis & Policy*, 2015, (15), 2017–2045.

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.

Krueger, Alan B. and Pei Zhu, "Another Look at the New York City School Voucher Experiment," *American Behavioral Scientist*, jan 2004, *47* (5), 658–698.

Mayer, Daniel P., Paul E. Peterson, David E Myers, Christina Clark Tuttle, and William G. Howell, "School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program," 2002. Mathematica Policy Research Report.

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.

Meghir, Costas and Steven Rivkin, "Econometric Methods for Research in Education," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 3, Elsevier, jan 2011, chapter 1, pp. 1–87.

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," *AERA Open*, 2018, *4* (1), 2332858417751355.

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*, 2017, *52* (3), 800–825.

OECD, *Public and private schools. How management and funding relate to their socio-economic profile*, Paris: OECD, 2010.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

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

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.

Urzua, Sergio, "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices," *Journal of Human Resources*, 2008, *43* (4), 919–971.

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

Weiss, Andrew, "Human Capital vs. Signalling Explanations of Wages," *Journal of Economic Perspectives*, Fall 1995, *9* (4), 133–154.

Wolf, Patrick, Babette Gutmann, Michael Puma, Lou Rizzo, Nada Eissa, and Marsha Silverberg, "Evaluation of the DC Opportunity Scholarship Program: Impacts After One Year," 2007. Executive Summary. US Department of Education, Institute of Education Sciences. Washington, DC: US Government Printing Office.

Wolf, Patrick J., Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, Nada Eissa, Matthew Carr, and Marsha Silverberg, "Evaluation of the DC Opportunity Scholarship Program: Final Report," 2010. NCEE 2010-4018. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute for Education Sciences, US Department of Education.

Zimmerman, Seth D., "Elite Colleges and Upward Mobility to Top Jobs and Top Incomes," *American Economic Review*, January 2019, *109* (1), 1–47.

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
Voucher	0.13	0.96	0.17	0.97
Private	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
Voucher	20%	556
Private	48%	1561
B. For-profit status		
Nonprofit	21%	669
For-profit	18%	555
C. Co-payment requirements		
Public with no copayment	10%	383
Public with copayment	12%	386
Voucher with no copayment	15%	506
Voucher with copayment	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: 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	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 4: 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		
Voucher school (%)	37.7	48.5		
Private 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		
Grade retention: 0 (%)	16.5	37.1		
Grade retentio: 1 (%)	5.6	23.0		
Grade retention: $\geq$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 5: Descriptive statistics by database

Variables	SIMCE original	SIMCE final
Language score (σs)	0.000	0.092
Math score (σs)	0.000	0.086
Public school (%)	47.6	45.4
Voucher school (%)	36.5	37.7
Private 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 6: The effect of school type during high school on adult earnings (in hundreds of US$)

	(1)	(2)	(3)
Private	(0.992,1.436)	(0.543,1.436)	(1.029,1.436)
	[0.816,1.633]	[0.256,1.722]	[0.899,1.636]
p-value: $UB = LB$	< 0.001	< 0.001	< 0.001
Voucher	(0.100,0.218)	(-0.027,0.218)	(0.066,0.218)
	[0.056,0.268]	[-0.131,0.327]	[0.020,0.385]
p-value: $UB = LB$	< 0.001	< 0.001	< 0.001
School-level math score		✓	
School-level funding (tuition cost or subsidy)			✓
Observations	111,395	111,395	107,282
Dependent mean (US$/100)	6.511	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). Imbens-Manski 95% confidence intervals are reported in brackets. We show the p-value of the null hypothesis that the upper and lower bound are equal against the alternative that the upper is larger than the lower bound.

Table 7: Effects of number of years spent in school on earnings (in hundreds of US$)

	(1)
Private	(0.181,0.244)
	[0.150,0.276]
p-value: UB = LB	< 0.001
Voucher	(0.019,0.044)
	[0.000,0.062]
p-value: UB = LB	< 0.001
Observations	110,228
Dependent mean (US$/100)	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 $z \in \{\text{private-fee-paying}, \text{private-voucher}\}$. To obtain the bounds, we compare regressions with and without math and language test scores. Imbens-Manski 95% confidence intervals are reported in brackets. We show the p-value of the null hypothesis that the upper and lower bound are equal against the alternative that the upper is larger than the lower bound.

Table 8: Effect of high schools on earnings and firms fixed effects (in hundreds of US$)

	(1)	(2)
Private	(1.987,2.845)	(1.440,2.045)
	[1.683,3.222]	[1.209,2.315]
p-value: UB = LB	< 0.001	< 0.001
Voucher	(0.263,0.484)	(0.223,0.376)
	[0.195,0.581]	[0.171,0.442]
p-value: UB = LB	< 0.001	< 0.001
Firm F.E.		✓
Observations	111,395	111,395
Dependent mean (US$/100)	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). Imbens-Manski 95% confidence intervals are reported in brackets. We show the p-value of the null hypothesis that the upper and lower bound are equal against the alternative that the upper is larger than the lower bound.

Table 9: The impact of early test scores on earnings (in hundreds of US$)

Variables	Baseline	Interactions
	(1)	(2)
Language	0.078**	0.091**
	(0.039)	(0.038)
Math	1.114***	0.948***
	(0.047)	(0.074)
Math×Voucher		0.076
		(0.089)
Math×Private		0.588***
		(0.127)
Exogenous characteristics	✓	✓
Family Background	✓	✓
Test scores	✓	✓
Observations	111,395	111,395
Dependent mean (US$/100)	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, clustered at the school level, are reported in parenthesis. ***, **, and * indicates a statistically significant coefficient at the 10, 5, and 1% level.

Table 10: The impact of school average academic achievement on earnings (in hundreds of US$)

	(1)	(2)
$\overline{\text{Language}}_j$	(-0.360,-0.363)	(-0.004,-0.001)
	[-0.379,-0.344]	[-0.021,0.015]
p-value: $UB = LB$	0.544	0.457
$\overline{\text{Math}}_j$	(1.404,2.126)	(0.904,1.614)
	[1.084,2.439]	[0.585,1.924]
p-value: $UB = LB$	< 0.001	< 0.001
$\overline{\text{Math}}_j \times$ Voucher		(-0.233,-0.229)
		[-0.245,-0.217]
p-value: $UB = LB$		0.300
$\overline{\text{Math}}_j \times$ Private		(0.810,0.821)
		[0.799,0.832]
p-value: $UB = LB$		0.110
Observations	111,395	111,395
Dependent mean (US$/100)	6.511	6.511

Notes: We show our estimated lower and upper bounds (equations 3 and 4) of the effects of school-level academic performance 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. Imbens-Manski 95% confidence intervals are reported in brackets. We show the p-value of the null hypothesis that the upper and lower bound are equal against the alternative that the upper is larger than the lower bound.

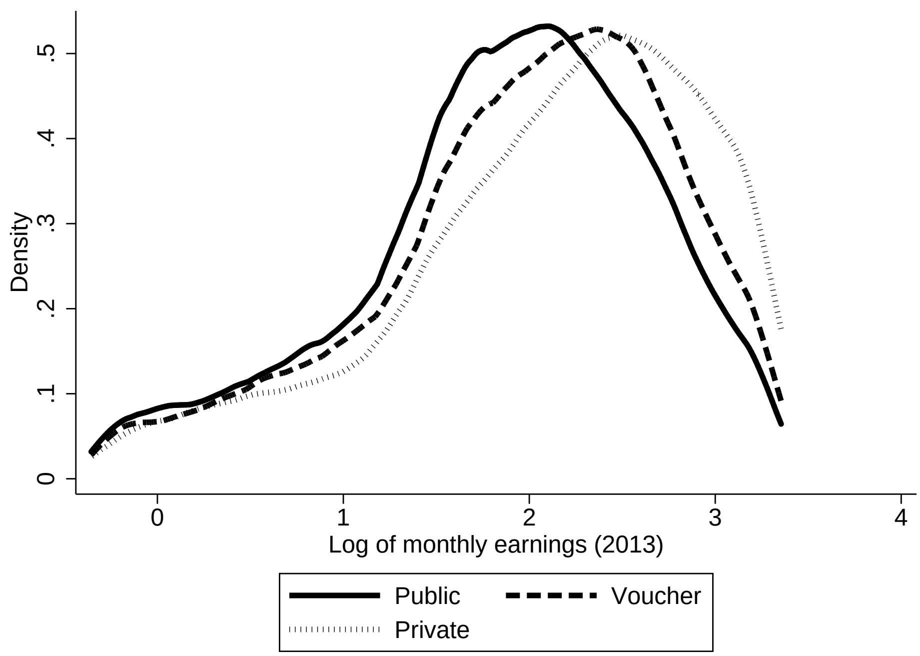
Table 11: The effect of educational monetary investment on earnings (in hundreds of US$)

	(1)	(2)
Log(spending)	(0.137,0.253)	(0.041,0.093)
	[0.092,0.304]	[0.022,0.113]
p-value: UB = LB	< 0.001	< 0.001
Log(spending)×Voucher		(0.025,0.050)
		[0.015,0.061]
p-value: UB = LB		< 0.001
Log(spending)×Private		(0.227,0.335)
		[0.185,0.382]
p-value: UB = LB		< 0.001
Observations	107,282	107,282
Dependent mean (US$/100)	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). Imbens-Manski 95% confidence intervals are reported in brackets. We show the *p*-value of the null hypothesis that the upper and lower bound are equal against the alternative that the upper is larger than the lower bound.

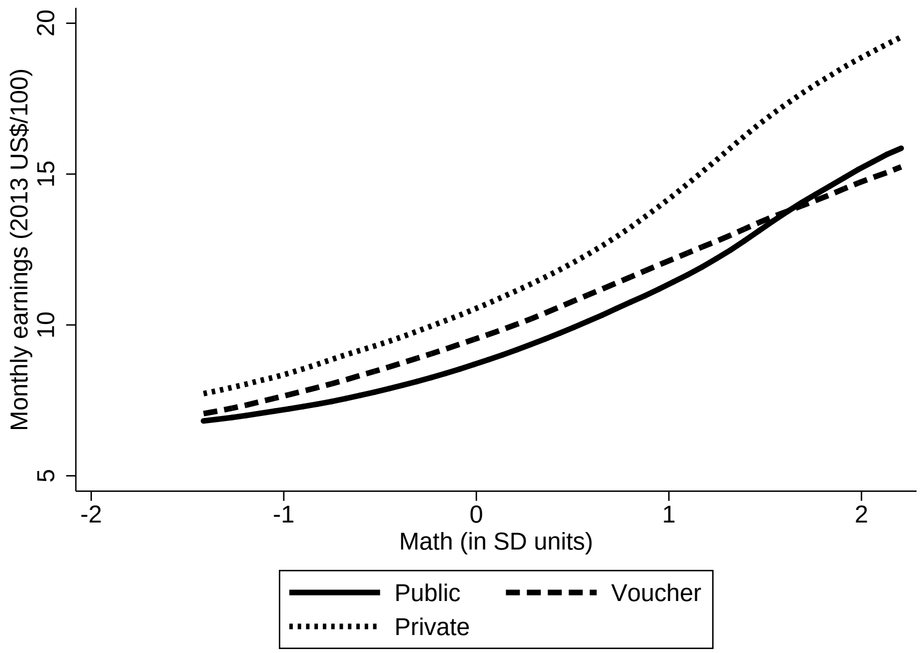
Figures

Figure 1: The distribution of (log) adult earnings in 2013 by hight-school type attended in 2001



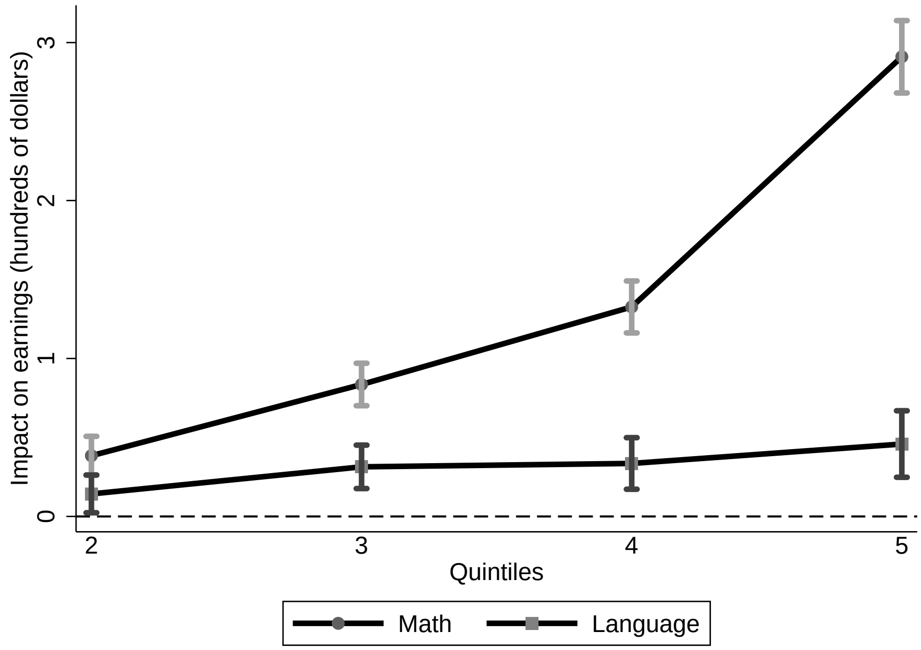
Notes: We show the estimated distribution of the 2013 log of monthly earnings for different school types.

Figure 2: Non-parametric regression of adult earnings (2013) and math test scores by high-school type (2001)



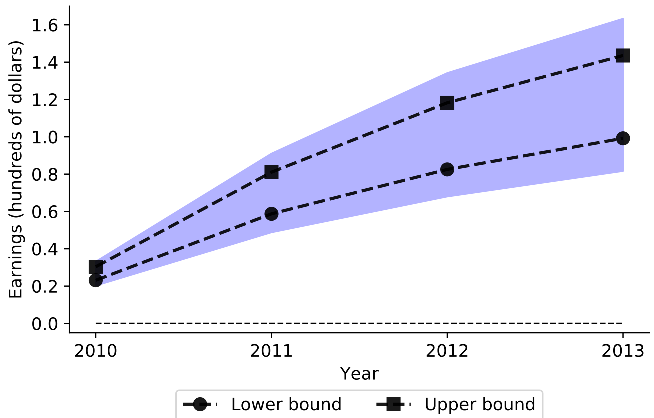
Notes: We show the fitted values of a local polynomial regressions of monthly earnings (2013 average in hundreds of dollars) and math test scores (in standard deviation units).

Figure 3: Labor market returns to math and language test scores during high school

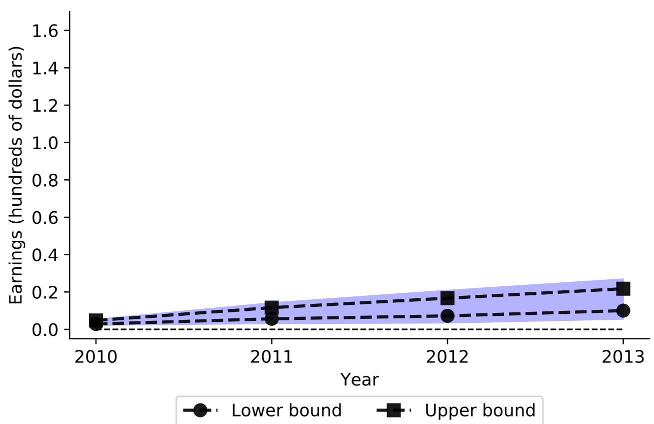


Notes: We show OLS estimates of a regression of 2013 monthly average earnings (in hundred of dollars) on a set of dummy variables indicating test scores quintiles. The baseline category is the first quintile of each test. Each point in the graph represents the effect of scoring in each quintile relative 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 attending voucher and private schools relative to public schools on earnings (in hundreds of US$)



(a) Voucher



(b) Private

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. The shaded area shows Imbens and Manski (2004) confidence intervals.

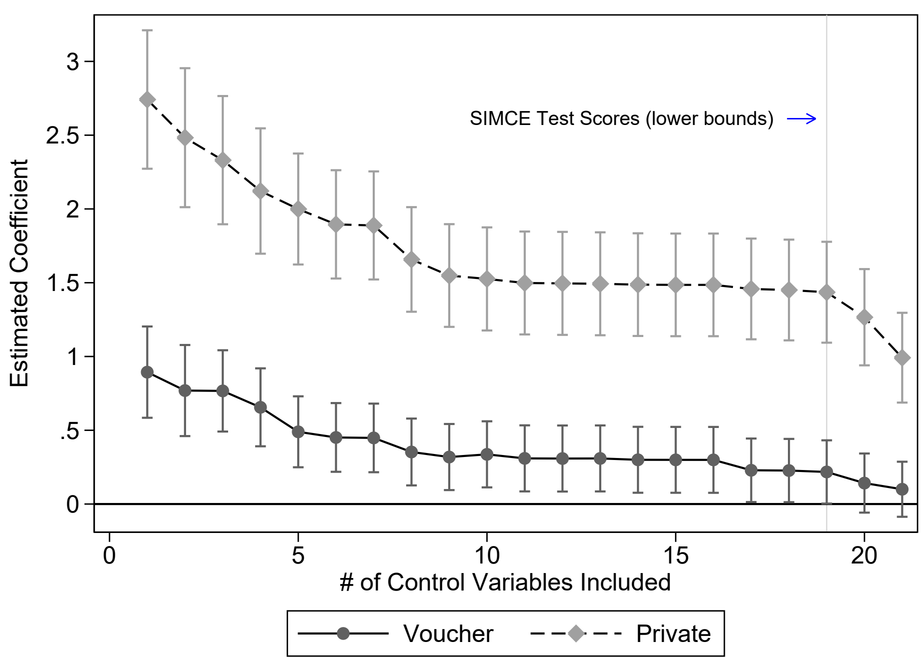
Appendices

Appendix A Evidence supporting biases as bounds

A.1 Coefficient stability

Figure A1 analyzes how sensitive our estimates are to some control variables. The figure presents estimated coefficients associated to voucher and private school in regressions with increasingly sets of control variables. For both voucher and private school earnings effects, we see stable estimates for most of the regressions, when a small set of control variables are included, the marginal control variable decreases the estimated coefficients by a large amount. However, even though we keep including powerful sets of control variables, estimated coefficients remain stable at some points, suggesting that selection on unobservables have little influence on our estimates. When we add tests scores to our regressions, estimated coefficients once again fall. This drop can be explained because when we include tests scores part of the causal effects of schools are absorbed in the estimated coefficients for language and math scores.

Figure A1: Estimated effects of voucher and private schools for different set of control variables



Notes: The figure shows estimated effects of voucher and private schools on earnings for different set of control variables. From left to right, we increase the number of control variables added to our regressions. The final two points corresponds to estimates with test scores as control variables (lower bounds on the causal effects).

A.2 Regressions on an omitted variable

Table A1 performs the first selection-on-unobservable test. We use a dummy for grade retention.[33] In the sample, 22% of students have been retained at least once.[34] 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 A1 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 be retained. When we add observed characteristics, the estimated coefficients associated to school dummies are cut down by half. When 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.

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

[34] 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.

Table A1: Testing selection based on unobservables: regression of omitted baseline variable on high school dummies

Variables	Unconditional	Upper bound	Lower bound
Voucher	-0.069***	-0.031***	-0.017***
	(0.010)	(0.007)	(0.006)
Private	-0.128***	-0.061***	-0.018**
	(0.010)	(0.009)	(0.007)
Exogenous characteristics		✓	✓
Family Background		✓	✓
Test scores			✓
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 grade retention. 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 (*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$).

A.3 Selection models

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 after controlling for test scores.[35] 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,

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

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

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 A2 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 A2 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 equations. In one of the models, the estimated ρ ends up being not significant after including test scores.[36]

[36] 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 A2: 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.

A.4 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:

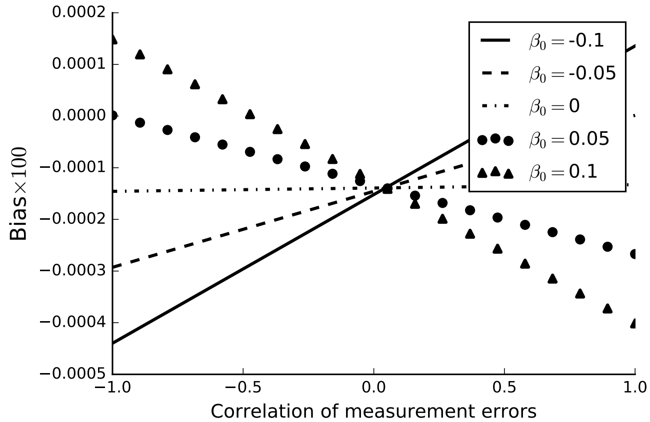
$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_{OLS} = \beta - \Sigma_x^{-1}\Omega\beta,$$

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).

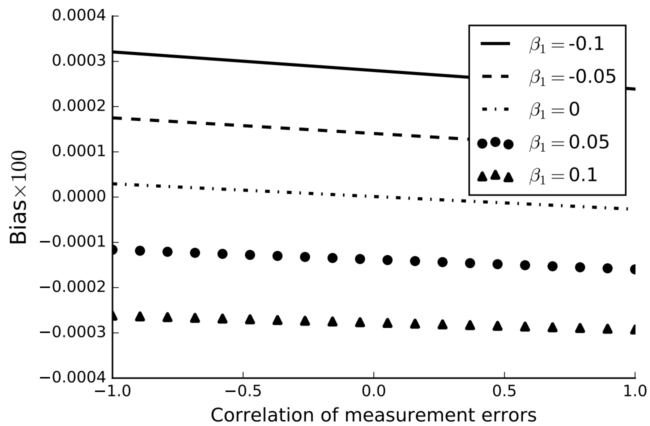
To analyze the consequences of measurement error in our regression models, we use our baseline specification and simulate the bias on the coefficient associated with private-fee-paying school dummy under different parameterizations. Figure A2 presents our results. The top panel depicts the bias as a 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 measurement errors is negative, an unlikely case given the nature of our variables.

Figure A2: 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).

Appendix B Full Regression Table

Table B1: Full Regression Table

Variables	(1)	(2)	(3)	(4)
Voucher	0.100	-0.027	0.066	0.218**
	(0.095)	(0.103)	(0.096)	(0.109)
Private	0.992***	0.543***	1.029***	1.436***
	(0.155)	(0.174)	(0.158)	(0.174)
PreK age 4	-0.155**	-0.103*	-0.132**	-0.192***
	(0.061)	(0.061)	(0.062)	(0.062)
PreK age 5	-0.126	-0.053	-0.123	-0.180**
	(0.086)	(0.087)	(0.088)	(0.088)
Studies sometimes	-0.504***	-0.485***	-0.472***	-0.683***
	(0.095)	(0.094)	(0.095)	(0.097)
Studies only for exams	-0.512***	-0.497***	-0.478***	-0.750***
	(0.095)	(0.094)	(0.095)	(0.096)
Almost never studies	-0.958***	-0.855***	-0.912***	-1.164***
	(0.125)	(0.124)	(0.128)	(0.127)
Doesn't know study habits	-0.587***	-0.530***	-0.522***	-0.775***
	(0.176)	(0.175)	(0.180)	(0.179)
Grade retention (one year)	-0.521***	-0.487***	-0.547***	-1.030***
	(0.063)	(0.063)	(0.064)	(0.065)
Grade retention (2 years +)	-0.744***	-0.676***	-0.768***	-1.345***
	(0.086)	(0.087)	(0.089)	(0.086)
Male	2.595***	2.571***	2.599***	2.835***
	(0.085)	(0.082)	(0.086)	(0.088)
Father in same house	0.112	0.124	0.099	0.232
	(0.195)	(0.195)	(0.199)	(0.196)
Mother in same house	0.032	0.040	0.039	0.176
	(0.122)	(0.122)	(0.126)	(0.123)
Both in the same house	0.199	0.154	0.199	0.078
	(0.211)	(0.211)	(0.217)	(0.213)
Attends parental meetings	0.234***	0.217***	0.233***	0.238***
	(0.057)	(0.057)	(0.058)	(0.057)
Father knows school scores	0.055	-0.129*	0.048	0.226***
	(0.077)	(0.076)	(0.078)	(0.084)
Father expectations: high school	0.043	0.074	0.072	0.111
	(0.119)	(0.118)	(0.124)	(0.120)
Father expectations: two-year college	0.477***	0.516***	0.492***	0.621***
	(0.128)	(0.127)	(0.132)	(0.130)
Father expectations: four-year college	0.798***	0.698***	0.811***	1.460***
	(0.130)	(0.127)	(0.134)	(0.130)
Mother's age	-0.008	-0.010*	-0.008	-0.009*
	(0.005)	(0.005)	(0.005)	(0.005)
Father's age	0.006	0.004	0.005	0.009*
	(0.004)	(0.004)	(0.005)	(0.005)
Mother: high school	-0.033	-0.090	-0.046	0.021
	(0.065)	(0.065)	(0.067)	(0.066)
Mother: vocational school	0.230***	0.165*	0.213**	0.341***
	(0.088)	(0.088)	(0.089)	(0.088)
Mother: two-year college	-0.234	-0.395**	-0.276	-0.016
	(0.183)	(0.183)	(0.183)	(0.185)

Table B2: Full Regression Table (continued)

Variables	(1)	(2)	(3)	(4)
Mother: four-year college	0.199	0.026	0.118	0.418***
	(0.158)	(0.159)	(0.157)	(0.159)
Mother: undergraduate studies	-0.433***	-0.630***	-0.496***	-0.051
	(0.143)	(0.144)	(0.143)	(0.147)
Mother: graduate studies	-0.315*	-0.408**	-0.377*	-0.185
	(0.188)	(0.188)	(0.193)	(0.188)
Father: high school	0.014	-0.020	0.013	0.050
	(0.060)	(0.059)	(0.061)	(0.060)
Father: vocational school	0.116	0.065	0.122	0.221***
	(0.083)	(0.083)	(0.086)	(0.085)
Father: two-year college	0.198	0.035	0.143	0.433**
	(0.177)	(0.176)	(0.181)	(0.180)
Father: four-year college	0.153	0.012	0.126	0.364**
	(0.166)	(0.167)	(0.167)	(0.166)
Father: undergraduate studies	0.056	-0.148	-0.005	0.402***
	(0.127)	(0.125)	(0.127)	(0.129)
Father: graduate studies	-0.069	-0.295	-0.085	0.281
	(0.270)	(0.262)	(0.272)	(0.277)
Student works	-0.029	0.032	-0.013	-0.053
	(0.096)	(0.096)	(0.098)	(0.097)
Log of family income	0.062	0.067*	0.045	0.045
	(0.038)	(0.038)	(0.039)	(0.039)
Books 10 - 50	-0.059	-0.078	-0.072	0.040
	(0.062)	(0.062)	(0.064)	(0.062)
Books 50 - 100	-0.166**	-0.221***	-0.184**	0.062
	(0.084)	(0.085)	(0.086)	(0.085)
Books 100+	-0.368***	-0.459***	-0.400***	-0.002
	(0.103)	(0.104)	(0.105)	(0.104)
Student language score	0.078**	0.033	0.069*	
	(0.039)	(0.039)	(0.039)	
Student math score	1.114***	0.918***	1.108***	
	(0.047)	(0.044)	(0.048)	
School math score		1.182***		
		(0.273)		
School language score		-0.237		
		(0.277)		
Log of educational resources			0.205**	
			(0.080)	
Observations	111,395	111,395	107,282	111,395
Dependent mean (US$/100)	6.511	6.511	6.511	6.511