School Vouchers and the Joint Sorting of Students and Teachers

Michela M. Tincani

April, 2014

School Vouchers and the Joint Sorting
of Students and Teachers

Michela M. Tincani[1]

May 12, 2014

Abstract

Countries around the world are adopting market-oriented school choice reforms. Evidence shows that they affect both student and teacher sorting across school sectors. Previous studies have analyzed student and teacher sorting in isolation from each other. This is the first paper to unify parental school choice and teacher sorting in an equilibrium framework. Using data from the large-scale Chilean voucher plan, this paper extends the existing literature in three ways. First, it evaluates how much of the treatment effect of Chilean voucher schools is due to teacher quality. Second, it examines the welfare implications of school specialization in different types of students. Third, it evaluates the impact of school choice expansion with endogenous public and private school teacher quality. There are three main results. First, better teacher quality accounts for 19 percent of the private school effectiveness in Chile. Second, assortative matching of students to teachers by ability can be welfare improving for low-ability students if the schools that have less able teachers are also those that specialize in the weakest students. Third, under the Chilean plan, highly skilled teachers are attracted into private schools from outside of teaching, with only limited cream skimming of teachers from public schools.

1 Introduction

Countries around the world are adopting market-oriented school choice reforms to improve student achievement. Previous studies indicate that they affect the sorting of both students and teachers across school sectors.[1] Moreover, teacher effectiveness has been found to vary by school and by student type.[2] Therefore, to understand the welfare implications of school choice it is important to examine how teacher stratification combines with student stratification to determine student outcomes.

Chile is unusual in having a large-scale school voucher plan and rich availability of data. Hence, it provides an excellent environment to study this research question. Using Chilean data, this paper develops and estimates a structural model of parental school choice with endogenous teacher labor supply and sorting across schools. While a large number of papers has studied student sorting in the presence of school choice, they have typically abstracted from teacher labor supply.[3] At the same time, the few papers that analyze teacher stratification under school choice do not analyze student outcomes or student sorting.[4] By unifying parental school choice and teacher labor supply in an equilibrium framework, I extend the existing literature in three ways.

First, by accounting for self-selection of both students and teachers into public and choice schools, I am able to identify how teacher quality combines with student quality to determine outcomes, and to evaluate how much of the treatment effect of attending a Chilean voucher school is due to teacher quality differences.

Second, I show that the welfare implications of the student-teacher match depend on whether different schools specialize in different types of students. For example, if public school teachers target the weakest students, it could be welfare improving for the weakest students to attend public schools even if they have lower quality teachers.

---

[1] The literature on student sorting is very large. Evidence on student stratification can be found in, for example, Hsieh & Urquiola (2006) and Urquiola (2005). The literature on teacher sorting under school choice is smaller and more recent. Jackson (2012) and Hensvik (2012) study how teacher stratification and other teacher outcomes are affected by a charter school program in North Carolina and a voucher program in Sweden, respectively.

[2] See Jackson (2013), Aaronson, et al. (2007), Clotfelter, et al. (2006), Dee (2005), Hanushek, et al. (2005).

[3] See Evans & Schwab (1995), Neal (1995), Grogger, et al. (2000), Altonji, et al. (2005b), Ladd (2002), Epple & Romano (1998), and Ferreyra (2007).

[4] See Jackson (2012) and Hensvik (2012).

Third, I evaluate how teacher quality, one of the most important determinants of school quality (Rivkin, et al. 2005), reacts endogenously to the expansion of a voucher program. For example, as the private school sector expands, the quality of public schools could worsen if private schools cream skim the best teachers away from public schools.

These three contributions require the joint analysis of student and teacher sorting. For this purpose, I develop a structural model and I combine a number of different datasets. The model assumes that parents consider teacher quality when choosing a school, and that potential teachers, i.e., individuals who are choosing between working in private or public school, working in a non-teaching occupation or staying at home, consider both wages and non-pecuniary job characteristics when making labor supply decisions.[5] The assumed production of student achievement has a flexible specification that allows for the effect of teachers to vary by school sector and by student type.

To take the model to data, I use a number of Chilean datasets from 2006. Data on the pool of potential teachers, including their characteristics, occupational choices and wages, come from the CASEN (*Encuesta de Caracterización Socioeconómica Nacional*), a representative sample of all Chileans, and from the ELD (*Encuesta Longitudinal Docente*) a teacher survey. Data on students come from the SIMCE (*Sistema de Medición de Calidad de la Educación*), which provides administrative test scores and background information on 4th and 10th graders.[6] I link the datasets by geographical location to obtain a multi-market dataset. This provides variation in opportunity wages by geographical location that is important to help identify teacher quality. The model is estimated using the Method of Simulated Moments (McFadden 1989, Pakes & Pollard 1989), and it fits the key aspects of the data very well, including the choice distributions of parents and of potential teachers. The good fit of the model along these important dimensions, and validation of other important implications of the model that were not matched by construction, help build confidence about the lessons that we learn from the estimated model and the counterfactual experiments I perform.

---

[5]A number of studies find that non-pecuniary job characteristics are important determinants of teacher labor supply, see, for example, Boyd, et al. (2005), Jackson (2012), Stinebrickner (2001b), Bonhomme, et al. (2012).

[6]SIMCE administers each year standardized tests in math and Spanish that all students of selected grades are required to take. The schools' average test results are published annually and parents can compare the performance of locally available schools. Hastings & Weinstein (2008) show that when parents have information on schools' test scores, they make school choice decisions that are beneficial to their child's achievement.

In Chile, there is positively assortative matching between students and teachers, with private schools attracting higher-achieving students and better teachers.[7] The institutional environment underlying this outcome is one in which all students receive a voucher that can be used to attend private or public school. While the voucher covers entirely tuition in public schools, private schools are allowed to charge tuition that exceeds the value of the voucher. In addition, as in many education systems around the world, in Chile public school teacher wages are determined by rigid wage formulae that reward seniority.[8] In contrast, private schools are free to set their own wages, and evidence suggests that they reward teacher skills. In spite of its simplicity, the model is able to capture the patterns of stratification on both sides of the market that have been documented using data from different years and data sources.

First, I use my estimated model to simulate the treatment effect of attending private school for those who attend (TT), the parameter of interest of many studies of school choice. I then simulate the counterfactual treatment effect of attending private school if private schools had teachers of the same quality as public schools. The difference between the two treatment effects measures the contribution of teachers. I find that the difference in teacher quality accounts for 19 percent of the effect of attending private school on math and Spanish test scores.

The second set of results derives from parameter estimates and from counterfactual simulations. Parameter estimates indicate that the technology of test score production is such that increasing teacher quality in public schools benefits the weaker students the most. This is not the case in private schools.[9] This is evidence that public school teachers specialize in the weak students. To perform an out-of-sample validation of this finding, I obtain a survey that was administered by the Chilean Ministry of Education to all 8th grade mathematics teachers. The answers provide additional strong evidence that public school teachers devote more attention to the weakest students in the classroom than private school teachers do. Hence, this finding is corroborated by data not used in the model estimation.

---

[7] The positive selection of students into private schools has been documented by many studies. See, for example, Hsieh & Urquiola (2006). The positive sorting of teachers has been documented in Behrman, et al. (2014b), Correa, et al. (2014), and Bravo, et al. (2010).

[8] Examples include Sweden, the Netherlands, Italy, and some states in the United States. See e.g. Sahlgren (2013) for a comprehensive list.

[9] Even though there is assortative matching of teachers to students by ability, students of all types are found in both school sectors. Therefore, there is enough variation in student quality within each sector to identify teacher effectiveness by student type.

I run counterfactual simulations to investigate how this public school specialization affects student welfare and test scores. The common wisdom is that assortative matching by ability is bad for achievement gaps and for disadvantaged students. However, I show that under the configuration of school specialization observed in Chile, policies aimed at reducing assortative matching could actually hurt the students that they intend to help. One of the lessons that we learn from this paper is that it is not possible to draw welfare implications by only using data on matching patterns, as the knowledge of how teacher effectiveness varies by student type and by school sector is key in interpreting the data.

The third and final result evaluates the impact of expanding the private voucher sector. Because the model is estimated from data on a large-scale voucher plan, to perform this counterfactual I simulate eliminating the voucher school sector and I compare outcomes under the two scenarios. I find that, through the teacher-student matching channel, the Chilean voucher system has a positive impact on both advantaged and disadvantaged students. For example, test scores of students at the bottom 25 percent of the income distribution increase by 0.20 standard deviations (sd), and those of students at the top increase by 0.75 sd. This is because private schools attract highly skilled teachers from outside of teaching, while exerting only a limited amount of teacher cream skimming from public schools. Moreover, parents move their children across schools in a way that, on average, benefits child achievement.

The rest of the paper is organized as follows. Section 2 explains how the three contributions of this paper relate to previous research. Section 3 presents the model, and section 4 discusses its features. Section 5 describes the institutional background of the Chilean voucher plan and the data. Section 6 explains the estimation approach and identification, while section 7 presents evidence on the model fit. Section 8 presents the empirical results, and section 9 concludes. The tables and technical appendices follow.

2 Related Literature

First Contribution: Decomposing the Effect of Treatment on the Treated and the Importance of Teacher Quality

a. Student sorting without teacher sorting. A large literature on student sorting estimates the treatment effect of choice schools (e.g., private, charter, voucher schools). The sorting of students across school sectors is analyzed to control for selection bias. Examples include Evans & Schwab (1995), Neal (1995), Grogger et al. (2000), and Altonji et al. (2005b), which estimate the effect of private Catholic schools in the United States.[10] Rouse & Barrow (2009) survey the evidence on small-scale voucher experiments in the United States, while Sapelli & Vial (2002) estimate the effect of attending a private voucher school in Chile, and find positive TT.[11] I adopt a strategy similar to Dynarski, et al. (2009) to account for student selection. Because these papers abstract from teacher sorting, they do not measure how differences in teacher quality contribute to the treatment effect.

b. Teacher sorting without student sorting. Jackson (2012) is one of the first papers to study labor supply decisions of teachers when school choice expands. He studies charter schools in North Carolina. Using data on test scores, he derives measures of teacher effectiveness within a teacher value-added framework.[12] This allows him to compare the effectiveness of public and charter school teachers. Under the assumption that teacher effectiveness does not vary by school or student type, it is possible to use his parameter estimates to measure the importance of teacher quality differences in determining the effect of attending a charter school, and it is not necessary to keep track of student sorting. However, when teacher effectiveness varies by school or student type (as evidenced in Jackson (2013) and Aaronson et al. (2007), among others), student sorting cannot be abstracted from.

Second Contribution: School Specialization and Welfare Analysis of Matching Patterns

Consistent with the finding on the specialization of Chilean public schools, Aaronson et al. (2007) find that teacher quality is particularly important for lower-ability students in Chicago public high schools.[13]

---

[10]These schools are found to have large positive effects on those who attend them, especially in terms of high school graduation and college attendance rates.

[11] See also Hoxby (2003).

[12]Hensvik (2012) and Behrman et al. (2014b) study teacher supply under voucher programs in Sweden and Chile respectively. Unlike Jackson (2012), they do not use data on student test scores. Hanushek et al. (2005) use a strategy similar to Jackson (2012), but they do not use data from a school choice program.

[13]In addition, a survey from the Thomas B. Fordham Institute shows that more than 80 percent of public school teachers in 2008 report that struggling students get more one-on-one attention than gifted students. *Source*:

The parameter estimates indicate evidence on assortative matching. This finding is consistent with the two separate literatures on Chilean teachers (Bravo et al. 2010, Behrman et al. 2014b, Correa et al. 2014) and on Chilean students (Hsieh & Urquiola 2006, McEwan, et al. 2008). This gives me confidence on the external validity of the model within the Chilean context.

Assortative matching has also been documented in other school systems. For example, Clotfelter et al. (2006) find that in North Carolina, more highly qualified teachers tend to be matched with more advantaged students. Similar patterns have been documented in New York State (Lankford, et al. 2002, Lankford 1999), San Diego (Betts, et al. 2003) and Texas (Rivkin et al. 2005). To find the determinants of schools' hiring behavior and teachers' occupational decisions, Boyd, et al. (2006) estimate a model of matching between teachers and schools. Like the studies cited above, their paper does not include data on student achievement. Therefore, student welfare analysis cannot be performed.

Third Contribution: Evaluating School Choice Expansion with Endogenous Public and Private School Qualities

Altonji, et al. (2014) also develop and estimate a structural model to evaluate how school quality changes endogenously as school choice expands.[14] Their measure of school quality is peer quality, rather than teacher quality. Using the National Education Longitudinal Study of 1988, they find that the effect of endogenous changes of peer quality on own achievement (i.e., the student cream skimming effect) is negligible. Even though Altonji et al. (2014) is one of the first papers to estimate this effect, the literature on school choice has long been concerned with it. The issue of teacher cream skimming has received much less attention. This is surprising, given that the evidence on peer effects in education is mixed (e.g. Sacerdote (2011), Epple & Romano (2011)), while there is broad agreement on the importance of teachers (Rivkin et al. 2005).

3   Model

Parents and potential teachers make school and occupational decisions to maximize their utility.

---

http://www.districtadministration.com/article/are-gifted-students-slighted-schools.

[14]Methodologically, their paper and this paper are related to a few papers on the general general equilibrium effect of voucher programs, e.g. Epple & Romano (2008), Epple & Romano (1998), and Ferreyra (2007).

Parents

Parents care about their child achievement and consumption. Moreover, they have a direct preference for a school sector that is independent of its effect on student achievement. This captures the fact that when less private schools are available (e.g., in rural areas), the average transportation cost associated with the private sector is higher. As a result, the private sector is chosen less often. Formally, the utility of family h in market m selecting school sector $j \in \{M, V\}$ is:

$$
\begin{aligned}
u_{hmj} &= u_h(c_{hmj}, a_{hmj}, \eta_{hj}) + \nu_{hmj}^{pref} \\
&= v_h(c_{hmj}) + a_{hmj} + \eta_{hj} + \nu_{hmj}^{pref}
\end{aligned}
$$

where c_{hmj} is consumption, a_{hmj} is child achievement, and η_{hj} is the direct preference for school j. ν_{hmj}^{pref} is a preference shock distributed as $\sim N(0, \sigma_{perf}^2)$ when $j = M$. It is normalized to be a degenerate random variable equal to zero for $j = V$, because only the difference in utility across choices is identified in a discrete choice model. For the same reason, the direct preference for $j = V$ is set equal to zero, while $\eta_{hM} = \eta(k_h) + \eta_1 primary_h + \eta_2 rural_h$, where k_h is the household's type, discussed below. The utility from consumption is equal to $\tau(k_h) ln(c_h)$, where $\tau(k_h)$ is a parameter that measures the trade-off between consumption and child achievement, and it determines parental willingness to pay for private education.

Consumption is equal to household income if parents select a free public school, and it is equal to income net of tuition payments if they choose a private school. Tuition payments are given by the tuition charged by the school p, minus the voucher subsidy v, minus a fellowship f if the student is eligible for one according to government guidelines. Formally:

$$
c_{hmj} = \begin{cases} y_h & \text{if j=M} \\ y_h - (p - v - f(Z_h)) & \text{if j=V} \end{cases}
$$

where Z_h are household characteristics determining the amount of fellowship the student is eligible for (which can be anywhere between 0 and $p - v$). To account for the fact that in the sample I never observe parents choosing the private sector when their income is smaller than the

tuition fees, I assume that parents are credit constrained. The utility from the voucher school is equal to $-\infty$ when tuition is above income, and it is never chosen.

Student achievement is determined by student characteristics that are observed (X_h) and unobserved ($k_h \in \{1, ..., K\}$) by the econometrician and by the quality of the teachers in the school ($\bar{s}_{jm}$). Formally:

$$a_{hjm} = a_{jm}(X_h, k_h, \bar{s}_{jm}) + \nu_{hj}$$

where ν_{hM} and ν_{hM} are productivity shocks distributed as independent mean-zero random variables with variances $\sigma^2_{\nu M}$ and $\sigma^2_{\nu V}$. They are independent of the preference shock ν^{pref}_{hmj}. Unobserved student characteristics such as ability are modeled as types, in the spirit of Heckman & Singer (1984). A student's type is a discrete random variable with probability mass function $\pi_1, ..., \pi_K$. There may be complementarities between teacher skills and student type. For example, better teachers may be more effective with higher or lower ability students. The model allows for this possibility by letting the effectiveness of teacher skills vary by student type, as can be seen in appendix 11 where the functional forms are reported. Moreover, the test score production technology is allowed to vary by school sector to capture any differences in this complementarity.

Potential Teachers

Potential teachers, i.e., individuals who are making labor supply decisions, care about the wage and non-pecuniary aspects of an occupation. They choose between becoming a public school teacher ($j = M$), becoming a private voucher school teacher ($j = V$), working in the non-teaching sector ($j = NT$), or staying at home ($j = H$). Formally, their utility is:

$$u_{imj} = \begin{cases} u(w_{imj}, \mu_{ij}) & \text{if j=M,V,NT} \\ u(\mu_{ij}) + \epsilon^{pref}_{iH} & \text{if j=H} \end{cases}$$

where w_{imj} is the wage offer obtained by individual i in market m from sector j, μ_{ij} is an occupation specific non-pecuniary preference, and $\epsilon^{pref}_{iH} \sim N(0, \sigma^2_H)$ is a preference shock to the home option. As detailed in appendix 11, containing the functional forms, the non-pecuniary term

for the non-teaching sector has been normalized to zero because it is not separately identified (Heckman & Honore 1990).

Each potential teacher is endowed with a certain level of teaching skills, s_i, which raise the achievement of students when employed in a teaching occupation. Teaching skills are determined by individual characteristics that are observed (X_i) and unobserved ($l_i \in 1, ..., L\}$) by the econometrician. Unobserved characteristics are modeled as types with type proportions $\psi_1, ..., \psi_{\bar{L}}$. Formally, the technology of teaching skill formation is:

$$s_i = exp(a_0(l_i) + a_1' X_i + \epsilon_i^{tech}) \tag{1}$$

where ϵ_i^{tech} is a technological shock.[15]

Wage offers depend on the individual characteristics and type. While wage offers in public schools are determined by rigid governmental formulae that do not depend on an individual's teaching skills, private school wages are assumed to be a linear function of skills. As in a classical Ben-Porath framework (Ben-Porath 1967), the wage in the private sector reflects the product of teaching skills and the price of those skills.[16]

Formally, wage offers in the three sectors are:

$$w_{imj} = \begin{cases} exp(\alpha_{0jm}(l_i) + \alpha_j' X_i + \epsilon_{ij}) & \text{if j=M, NT} \\ r_m s_i = r_m exp(a_0(l_i) + a_1' X_i + \epsilon_i^{tech}) & \text{if j=V} \end{cases}$$

where r_m is the price of teaching skills in market m. The constant in the public school log-wage equation depends on an individual's type l_i to capture variables entering the rigid wage formulae that are not available in the dataset. The wage shocks $\epsilon_i = [\epsilon_{iM}, \epsilon_i^{tech}, \epsilon_{iNT}]'$ are i.i.d., independent of the preference shock, and distributed as $N(0, \Sigma)$, where Σ is a diagonal matrix with elements $\sigma_M^2, \sigma_V^2, \sigma_{NT}^2$. The non-teaching wage offer can be interpreted as the product of the price of non-teaching skills and the amount of non-teaching skills possessed by the individual. Any correlation between teaching and non-teaching skills is captured by the type distribution.

---

[15]See appendix 11 for what variables enter X_i in the empirical implementation.

[16]How the price of skills is determined is outside the scope of this paper. In ongoing work (Tincani 2014b), I extend the model to endogenize the price of teaching skills.

This correlation affects how the wage elasticity of the teacher labor supply is affected by non-teaching opportunities.[17]

Equilibrium

There is an interdependence in the choices of potential teachers and parents. The labor supply decisions of potential teachers determine the amount of mean teaching skills supplied to each school sector ($\bar{s}_{jm}$). In turn, this is the measure of teacher quality used by parents when choosing a school. Refer to appendix 12 for a derivation of the teaching skills being supplied to each sector in equilibrium.

It is easy to establish that an equilibrium exists and is unique. An equilibrium is attained when all parents and all potential teachers choose the option that maximizes their utility. An equilibrium exists because each potential teacher and each household have at least one most-preferred choice. This is true by construction: utilities are well defined. Uniqueness derives from the error structure. Because technology and preference shocks are continuously distributed in the population, the population proportion of potential teachers and of parents who are indifferent between sectors has measure zero.

4   Discussion of Model Features

The specification of the production function is flexible in that it does not impose a specific type of complementarity between student ability and teaching skills. Moreover, the technology is allowed to vary by school type, reflecting the fact that public and private schools may have different teaching philosophies. This flexibility has two advantages: first, results are not driven by ex-ante arbitrary restrictions on the type of complementarity. Second, it is possible to examine ex-post the features of the estimated production function. A limit of this specification is that teacher effectiveness does not react to changes in the composition of students in the school. This assumption is maintained because the data do not provide the exogenous changes to classroom composition that are needed to identify teachers' reactions to classroom composition. In general, how teachers react to classroom composition is a very interesting yet largely unexplored question.

---

[17]If skilled teachers are also skilled in non-teaching occupations, they obtain high wage offers from outside of teaching and schools must offer higher wages to attract them away from those occupations.

An appealing feature of the model is its simplicity. Because parents care about the identity of the teachers in a school, but teachers do not care about the identity of the students, it is not necessary to solve a fixed-point problem to find the equilibrium. This makes the solution of the model computationally straightforward, and it guarantees the existence and uniqueness of the equilibrium.

The results that derive from the parameter estimates are not affected by this assumption. To see why, notice that if teachers have a preference for student characteristics, this preference is captured in the model by the sector-specific non-pecuniary preference. Therefore, the wage parameters, which are used to infer teaching skills, are estimated without bias. In turn, teaching skills are an input into the technology of test score production, which, as a result, is also estimated without bias.

The assumption affects the counterfactual experiments that change the allocation of students across schools. Because the model assumes that teachers do not derive direct utility from student characteristics, it does not predict that teachers change school as a result of the experiments. For example, one of the counterfactual experiments increases the rate at which low-ability students choose public schools. If more able teachers dislike low-ability students, they would move to private schools to avoid the inflow of low-ability students. As a result, the benefit to low-ability students of moving to public schools would be overestimated by the model. The assumption that teachers do not care about students might appear restrictive. However, it is reasonable to expect that teachers face more mobility costs than students. Using a longitudinal dataset of Chilean teachers, Behrman et al. (2014b) document high persistence over time of teachers in each school sector, and they estimate high mobility costs. Therefore, at least in the medium term, I expect the results from these counterfactual experiments to hold.

The model does not include peer effects. Some authors caution that if private schools cream skim the best students from public schools, peer quality in public schools worsens. If peers matter for achievement, this harms the students who are left behind in public schools. In practice, evidence on the empirical relevance of peer effects is mixed (Sacerdote 2011), and recent evidence indicates that cream skimming has a small impact on test scores (Altonji et al. 2014). Estimation of a more general model that allows for peer effects would allow me to compare the importance

of peer effects with that of teacher quality in the context of Chile. However, my data do not provide variation that can identify peer effects (Manski 1993, Moffitt et al. 2001). Moreover, the model would lose its appealing simplicity. The model solution would require solving a fixed-point problem, and estimation would have to address potentially multiple equilibria, in the spirit of Brock & Durlauf (2001). Finally, it has to be noted that the methodological approach in this paper is similar to what has already appeared in the literature: it isolates one channel while abstracting from the other. For example, Altonji et al. (2014) and Epple & Romano (2008) develop structural models that allow for peer effects but not for teacher sorting. Dills (2005) uses a non-structural approach to evaluate the effect on public school students of a change in peer quality due to school choice. However, she abstracts from teacher sorting.

The model is static, and individuals choose between school sectors rather than schools. Both of these assumptions are due to data limitations. Linkable information on potential teachers and students is available only for one year of data, and the sample sizes do not allow me to analyze choices among schools within a sector. Analyzing the joint sorting across school sectors is an important novelty of this paper that fills a gap in the literature. Sorting within school sectors and in a dynamic setting are interesting extensions that should be addressed in future research, but they are not central to the analysis in this paper.

The model assumes that residential sorting is exogenous. While a literature on location choices and public goods exists (Epple & Sieg 1999, Nechyba 2000, Ferreyra 2007), there is not yet a well developed literature on two-sided equilibrium models with two-sided residential sorting. The paper that is closest to this one in terms of modeling matching patterns of schools and teachers is Boyd et al. (2006). As in this paper, they estimate their model on multiple markets and treat the allocation of teachers and schools to markets as exogenous.

Finally, wages in Chilean private schools are not constrained by rigid wage formulae. Therefore, they might change in the second set of counterfactual experiments. For example, if high-ability teachers become less effective on high-ability students, the willingness of their parents to pay for private education might decrease. As a result, the market price of teaching skills in private schools could decrease. This would have feedback effects on the sorting of both teachers and parents. In ongoing work (Tincani 2014b), I extend the model to endogenize wages in pri-

vate schools and I use the extended model to study wage policies in public schools. Simulations from the extended model are very similar to those presented in this paper, indicating that the conclusions of this paper would not change with the inclusion of endogenous wages.[18] While this model extension is important when analyzing wage policies, it is not central to the analysis in this paper.

5 Institutional Background and Data Description

5.1 Institutions

In 1981, Chile introduced a nationwide school voucher plan. Under the plan, each school-aged child receives a voucher that can be spent toward full coverage of tuition fees in a municipal school or coverage (partial or full) in a private subsidized school. The value of the voucher was CLP 27, 391.903 ($\sim$ \$50) in 2006. The voucher cannot be used in private unsubsidized schools, from which this paper abstracts. These schools enroll 6% of students and cater to the wealthiest families. Private voucher schools are allowed to charge a fee that exceeds the value of the voucher, up to a legal cap of CLP 54, 018.768 per month ($\sim$ \$100).

Some children are eligible for a *beca*, a fellowship for private education, that partially or fully covers the tuition fees in excess of the voucher. According to the SIMCE dataset, in 2006 around 60% of all Chilean children enrolled in private subsidized schools received a fellowship. As a result of government guidelines for fellowship assignment, children of lower socioeconomic status and from larger families are eligible for larger fellowships.[19]

Teachers' wages in the municipal sector are determined by rigid formulae that are negotiated between the government and the National Teachers' Association, *Colegio de Profesores*. Wages are subject to seniority increments and other adjustments, such as compensation for working in difficult conditions. Teacher assignment to schools is centralized nationally. Municipal schools, therefore, do not have control over the quality of the incoming pool of teachers. Teachers in

---

[18]Tables with counterfactual simulations under endogenous wages are available upon request.

[19]The value of the voucher and the cap on private school tuition can be found in the *Decreto con Fuerza de Ley N° 2, De Education, De 20.08.98* and in the law on shared financing, *Financiamiento Compartido, Ley N° 19.532*. The guidelines for fellowship assignment can be found in articles 24 and 27 of the *Ley de Subvenciones, Decreto con Fuerza de Ley N° 2, 20.08.98*.

private schools, on the other hand, are subject to the Private Labor Code, and their wages can be set freely by private schools. They are allowed to tie wages to teacher quality to attract a high-quality pool of teachers, and below I present evidence suggesting that this is occurring.

5.2 Data Description

5.2.1 Data Sources

I combine three data sources from 2006, the only year for which information on students and teachers in primary and secondary schools is available and linkable. I use the *Encuesta de Caracterización Socioeconómica Nacional* (CASEN) dataset to identify the pool of potential entrants into the teaching profession through a representative sample of individuals holding a college degree, a requirement for teaching.[20] The CASEN survey is a nationally representative survey of the general population from which I extract a sample of $3,520$ individuals holding a college degree, tracking their occupational choices, accepted wages, and characteristics. I drop from the sample individuals who reside in the remote rural areas of Aisén or Tarapacá, for sample size reasons.

To augment the sample of teachers, I use a sample of $3,195$ teachers from the *Encuesta Longitudinal Docente* (ELD) dataset. I extract from ELD the same set of individual characteristics obtained from CASEN, as well as the choice of school sector and accepted wages. As with CASEN, I drop teachers who live in the remote Aisén or Tarapacá areas.

On the students' side, I randomly select a sample of $100,000$ students from the restricted-access version of the *Sistema de Medición de Calidad de la Educación* (SIMCE) dataset, which contains information on all 4th and 10th graders in the country.[21] The dataset contains adminis-

---

[20]Individuals who want to become teachers must obtain a teaching certification. Although the teachers' statute, *Estatuto Docente, Ley N° 10.070*, allows for four ways to become certified to teach, according to the 2006 teacher census (*Idoneidad Docentes*), 95% of all teachers (100% of all teachers in this paper's sample) get certified through one of two channels: i) a college degree in education, ii) a college degree in another area and a special degree in education (2-4 semesters). Importantly, anyone with a college degree can become a teacher, as long as they receive training in education if their college major was not education. Because in CASEN I do not observe the college major, I let the non-pecuniary preference for teaching depend on an individual's unobserved characteristics. This captures in a reduced form way the fact that to accept an offer from the teaching sector, a college graduate without a major in education must pay the (money and time) cost of obtaining training in education. Therefore, everything else being equal, an individual without a major in education has a lower preference for teaching.

[21]The sample size is approximately one third of the population size. Selecting a sample was necessary for computational tractability.

trative information on students' test scores in math and Spanish, used to measure achievement, as well as information on the students' household and choice of school.

The model is estimated on 18 local labor and education markets. In appendix 13, I discuss how market boundaries were determined. Markets are closed, as required for estimation purposes, with 98.8% of teachers working in the market in which they reside, and 99.0% of parents choosing a school in the market in which they reside.[22] Nationally, the voucher sector accounts for 52.99% of student enrollment and 45.16% of teacher employment. However, as shown in appendix 13, there is across-market variation due to different market conditions affecting demand for private education and teacher supply. Variation across markets, such as variation in opportunity wages of teachers, is treated as exogenous and used for identification.

5.2.2 Descriptive Statistics

In private schools, there are students with higher socioeconomic status (SES) and less experienced teachers with higher measures of cognitive skills.

Children in the top 25 percent of the income distribution score, on average, 0.60 standard deviations (sd) higher than children in the bottom 25 percent. There is also a sizable test score gap between municipal and voucher school students. The difference in test score means is equal to 0.33 sd, which is more than one third of the black-white test score gap in the U.S., and larger than the gap between charter and traditional public schools in the U.S. A third of this gap remains after controlling for student characteristics.

As documented also in previous studies (e.g., McEwan et al. (2008)), in the Chilean education system there is considerable school stratification by students' SES. Table 2 shows average household characteristics by type of school in the 2006 SIMCE dataset. Parents of students in private subsidized schools earn almost twice as much as parents of students in municipal schools. Similar patterns are present among virtually all the household characteristics available in the sample.

Teachers in the private voucher sector are on average 8.2 years younger and have 9.0 fewer years of teaching experience than teachers in the municipal sector. They score 0.27 standard

---

[22]I drop from the sample the students and teachers who are observed moving across markets.

deviations higher on the PAA test, the Chilean equivalent of the SAT, a measure of cognitive skills. Interestingly, private voucher school teachers also score 0.19 standard deviations higher on the AEP test (*Asignación Excelencia Pedagógica*), which evaluates teaching ability. However, taking the AEP test is voluntary and only 1.5 percent of teachers take it. Thus, this figure is not computed on a representative sample and must be taken with caution. Still, to the extent that the sample selection bias is the same among municipal and private school teachers, the difference in AEP test scores is free of bias and it indicates that private school teachers are better.

An examination of teaching wages reveals why private school teachers have higher cognitive skills: private school wages reward cognitive skills, while public school wages do not. A panel data fixed-effects regression of log wages in municipal schools on teaching experience, teaching experience squared, nonteaching experience, and standardized PAA scores yields an insignificant coefficient on the PAA score (p-value=0.169).[23] The same regression estimated for voucher schools yields a significant coefficient (p-value=0.009), indicating that a one standard deviation increase in the PAA score increases wages by 4.0 percent in private schools. Similar regression results have been reported in Bravo et al. (2010), who also show that teacher PAA scores are positively correlated with student test scores. To the extent that teacher PAA scores are positively correlated with teaching skills, this suggests that private school wages reward teaching skills, while public school wages do not.[24]

Wages of teachers with up to ten years of teaching experience are 14% higher in private voucher schools than in municipal schools. This wage difference disappears for more experienced teachers as can be seen in table 3.

Finally, non-teaching wages are on average 62.3% higher in the non-teaching sector for equally educated individuals. A college graduate employed in a non-teaching occupation earns monthly, on average, CLP 777,396 ($\sim$ \$1,550), while a college graduate employed in teaching earns on average CLP 479,041 ($\sim$ \$960). A wage difference persists at all ages, reaching peaks of over 80% for individuals younger than 45. In terms of hourly wages, the gap reduces to 18.7%, reflecting the fact that individuals in the non-teaching sector work more hours. Perhaps because of the larger flexibility of the teaching time schedule, around 70 percent of teachers are women.

---

[23]The longitudinal ELD dataset was used to estimate these panel regressions.

[24]Some studies use teacher test scores as measures of teaching ability (e.g., Manski (1987)).

6 Estimation and Identification

6.1 Estimation Approach

The parameter vector θ is estimated by the method of simulated moments (MSM) (McFadden 1989, Pakes & Pollard 1989). The method minimizes the distance between observed outcomes and outcomes simulated from the model. The outcomes are occupational choices of potential teachers, school choices of parents, wages of potential teachers, test scores of children, and fellowship amounts. A list of the moments used can be found in appendix 14.

Let y_i denote an observed outcome for individual i. Let $\Omega_i \times \{1, ..., L\}$ denote the state space of individual i with elements (ω_i, l_i) (where $l_i \in \{i, ..., L\}$ is the person's type). Vector ω_i contains, for example, degrees, age, gender, etc. Let $\hat{y}_i(\omega_i, \theta)$ denote the outcome predicted by the model. This outcome is replaced by the simulator:

$$\tilde{\hat{y}}_i(\omega_i, \theta) = \frac{1}{S} \sum_{s=1}^{S} \sum_{l=1}^{L} Pr(l_i|\theta)\tilde{\hat{y}}_i(\omega_i, l_i, s, \theta)$$

obtained by drawing S simulated shocks from the model's shock distribution under parameter θ and using the model to simulate behavior, and hence an outcome for each individual, simulation, and type: $\tilde{\hat{y}}_i(\omega_i, l_i, s, \theta)$.[25] The simulated outcomes are then averaged across simulations and types. Moment conditions are constructed by taking the difference between the actual and the simulated outcome: $m_i(\theta) = y_i - \tilde{\hat{y}}_i(\omega_i, \theta)$.

The MSM finds the vector θ that minimizes the weighted distance of the empirical moment conditions from zero:

$$\hat{\theta}_{MSM} = \arg \min_{\theta} m(\theta)'_n W_n m_n(\theta) \tag{2}$$

where W_n is a symmetric positive definite weighting matrix such that as $n \to \infty$, $W_n \to W$ in probability with W symmetric and positive definite. Vector $m_n(\theta)$ is the sample average of the individual moment conditions $m_i(\theta)$. Because I use multiple data sources, I adjust the criterion function in (2) and the parameter standard errors to account for the relative sizes of the datasets and of the relative populations of reference (potential teachers and students). I

---

[25]In estimation, S is set equal to 100.

follow the method developed in Bhattacharya (2005), details of which can be found in appendix 15. Appendix 15 also contains the asymptotic properties of the estimator, as well as details of the estimation of the asymptotic variance covariance matrix.

6.2 Identification

6.2.1 Identification of Teacher Skills

The strategy to identify teaching skills exploits the teacher labor supply part of the model and data. Teacher sorting is modeled as a Roy model (Roy 1951) of self-selection into occupations, a workhorse model in labor economics.[26] In this class of models, individual wages in each occupation are determined by the product of the price of occupation-specific skills and the amount of skills possessed by the individual. Bias in wage parameters due to self-selection is accounted for by explicitly modeling occupational choices with exclusion restrictions.[27] Identification of the wage and non-pecuniary preference parameters in this class of models has been proven formally by, for example, Heckman & Honore (1990).

The model of potential teacher labor supply is an extended Roy model of self-selection into occupations with log-normal skills, non-pecuniary preferences, and a non-work option. Exclusion restrictions are given by the fertility variables that affect occupational choice but not wages.[28] Private school wages are assumed to be the product of teaching skills and the price of those skills; therefore, they contain information on the underlying skills of the teachers. This assumption is supported by the evidence presented in section 5. In addition, it is supported by the findings in Bravo (1999), it is consistent with the Chilean institutions, and it is maintained in other studies of the Chilean teacher labor market (Behrman et al. 2014b, Correa et al. 2014).[29] Wages in the non-teaching sector reflect non-teaching skills, while wages in public schools are set by rigid wage

---

[26]It has been used to study self-selection in a number of contexts, for example, immigration and residential choice (Borjas 1987), occupational and industrial choice (Heckman & Sedlacek 1985, Heckman & Sedlacek 1990), optimal taxation with self-selection (Rothschild & Scheuer 2013) and employment in the private and public sectors (Borjas 2002).

[27]Exclusions restrictions are not needed in a fully parametric model.

[28]A Roy model of Chilean teacher choices with only two occupational choices (public or private school) has been estimated in Correa et al. (2014), while the model in Behrman et al. (2014b) is very similar to the labor supply part of this model, but the setting is dynamic. The findings of both papers are consistent with the finding of this paper that private school teachers are more skilled.

[29]Consistent with this assumption, Hoxby (2002) shows that when schools face competition, as Chilean private schools do, teacher characteristics that are valued more by parents are also rewarded more in the labor market.

formulae. The parameters of interest for the identification of teaching skills are the parameters of private school wages. Once these parameters are estimated without bias, they can be used to infer the teaching skills of all potential teachers in the sample through equation 1.

The fact that wage data come from multiple markets presents additional advantages for identification. First, exogenous variation in non-teaching opportunity wages across markets provides exogenous shifts to the teacher labor supply that help identify differences in teacher skills across markets.[30] Second, in the classical Roy model with one market, skills are not separately identified from the price of skills. In contrast, when wage data from multiple markets are available, one can exploit variation in skill prices across markets to separately identified skills from skill prices.[31] For this purpose, I use the identification strategy developed in Heckman & Sedlacek (1985).

I make the identifying assumptions that the distribution of potential teachers' unobserved types is the same across markets and that the function that maps individual teacher characteristics into teaching skills in equation 1 does not depend on the market in which the teacher resides. For example, five years of teaching experience produce the same amount of teaching skills in Santiago as they do in Valparaiso.[32] This assumption is standard in the education literature, where the coefficients on teacher characteristics in the wage equation are not normally tied to the teacher's location (see, for example, Stinebrickner (2001a), Stinebrickner (2001b), and Dolton & Van der Klaauw (1999)). Hence, after controlling for selection bias in private school wages, any residual variation in the constant of log-wages across markets is due only to changes in the price of skills and not to changes in skills. Heckman & Sedlacek (1985) refer to this assumption as the "proportionality hypothesis." Notice that the price of teaching skills must be normalized in one market.[33]

---

[30]Table 15 in appendix 13 shows how non-teaching wages vary across markets.

[31]The local labor markets are defined in a way that results in infinitely high moving costs. Hence, price differences across markets may persist in equilibrium.

[32]Notice the difference between teaching skills, and their effect on student test scores. Here, I am only assuming that the technology of production of teaching skills is the same across markets; I'm not making any assumptions about how those skills affect student test scores in the different markets.

[33]In Heckman & Sedlacek (1985), years play the role of markets in this paper: the constant in the log-skills is assumed to be constant across years, but skill prices are allowed to vary by year. See note 17 in their paper. A consequence of the normalization is that skills are identified only up to scale. This does not affect any of the counterfactual experiments considered in this paper, because the choice of normalizing constant does not affect the estimated impact of teacher skills on achievement ($\beta_{1j}(k)\bar{s}_{jm}$ in equation 3 in appendix 11). The results that derive from comparing the magnitudes of production function parameters across student types and across school

6.2.2 Identification of Unobserved Student Ability

Self-selection into schools is governed by the price elasticity of parents, which is determined by parental willingness to trade off child achievement for consumption ($\tau(k_h)$ in the model). This is correlated with students' unobserved ability through the unobserved type k_h. Therefore, private school students differ from public school students in terms of unobserved ability, because self-selection induces different distributions of unobserved types across the two school sectors. A selection-correction method must be adopted.

The scholarship assignment rule can be exploited to account for self-selection, because it provides tuition fee variation that is uncorrelated with parental willingness to pay and with unobserved student characteristics. Anand, et al. (2009) document this feature of Chilean scholarships and exploit it to control for self-selection bias.

As in a standard two-step correction procedure, the model in this paper contains both the selection equation (parental choice of school) and the output equation (achievement production function). The selection equation depends on the tuition payment in private school. This in turn depends on the fellowship assignment formula, which contains exclusion restrictions. These are variables that affect the amount of fellowship received, but do not directly affect achievement, such as, for example, family size.

Altonji, et al. (2005a), in their study of instrumental variables to control for self-selection bias in education, conclude that a promising approach is to use "tuition levels and tuition discounts based on number of children with the idea of using the dependence of tuition on family size as a source of identifying variation." This is the source of identifying variation used in this paper. Dynarski et al. (2009) use the same instrument to identify the price elasticity of private school attendance in the United States.

6.2.3 Identification of the Achievement Production Function Parameters

An important question in this paper is how teaching skills and student characteristics combine to produce achievement in the two school sectors. To identify the parameters of the pro-

---

sectors are not affected either. To see why, notice that if c is the normalizing constant, then $c\beta_{1j}(k)$ is identified for all $j \in \{M, V\}$ and student types $k \in \{1, ..., K\}$. Therefore, the ratios $\frac{\beta_{1j}(k)}{\beta_{1i}(m)}$ are identified for all combinations of schools j, i and student types k, m.

duction of achievement, one must observe variation in teacher skills and variation in student characteristics and relate them to variation in student outcomes. However, teacher skills and student ability are not directly observed. I overcome this obstacle by using the fact that they are identified in the model.

The key insight is the following: if only the demand for education part of the model and data were available, the overall impact of teaching skills on achievement could be identified as a test score residual that is not explained by student characteristics.[34] However, this residual could also include other school characteristics not related to teachers. Moreover, it would not be possible to examine how the impact of teachers varies with teacher and student characteristics.

On the other hand, if only the labor supply part of the model and data were available, it would be possible to estimate teaching skills through the Roy model, but this would only tell a part of the story. The effect of teaching skills on student achievement would be unknown. This is a typical feature of teacher labor supply studies. This literature has developed separately from the literature on Cognitive Achievement Production Functions. As noted by Hanushek & Rivkin (2006), this is problematic, because knowledge of the wage elasticity of the supply of teacher characteristics is not sufficient to infer the impact of teachers on student outcomes.

This paper joins demand for education and supply of teachers into an equilibrium framework that enables me to: first, identify the teaching skills supplied to each school through the teacher labor supply part of the model and data; and second, estimate the impact of the identified teaching skills on student achievement through the student part of the model and data. Technically, I accomplish this by developing an algorithm that embeds within the estimation the equilibrium restrictions of the model. An outer loop searches over the parameter space for the parameter that minimizes the objective function of the MSM. At each parameter iteration, an inner loop solves for the equilibrium supply of teaching skills to each school and plugs the simulated teaching skills into the achievement production function. Therefore, at each parameter iteration, it is *as if* teaching skills were observed, and their impact can be separated from the impact of other school characteristics. Details of the algorithm can be found in appendix 16.

---

[34]Unobserved student characteristics are accounted for by the self-selection correction described in the previous section.

7 Model Fit

Table 1 presents evidence on the model fit. Simulations of the choice distributions of parents and potential teachers are very close to the data, within, respectively, 0.9 and 1.1 percentage points. Figures 1 and 3 show visually how accurate the model predictions are for the choice distributions of parents and potential teachers. Figure 2 shows the accuracy of the model in predicting tuition payments (net of the voucher and fellowship) in private schools, which depend on the endogenous selection of households into private schools because different households are eligible for different fellowship amounts. Wages simulated from the model are within about 5 percent of actual wages. Mean test scores are slightly under-predicted. However, the simulated test score gap by school type is close to the actual one, and the gap by income is within 7.5 percent of the actual one. Figure 4 shows that the distributions of actual and simulated test scores by school type are close, especially for public schools.

8 Empirical Results

8.1 Parameter Estimates

Parameter estimates of potential teacher wage offers and utility are reported in tables 4 and 5, of parental utility and test score production in tables 6 and 7, and of the fellowship formula in table 8. The estimation allows for three unobserved types of parents/students and of potential teachers to capture unobserved heterogeneity in preferences and ability.

First, in markets with more appealing non-teaching options, teachers are of lower ability. The correlation between simulated teaching skills and simulated non-teaching wage offers is negative in both sectors, and more negative for public schools (-0.37) than private schools (-0.10). This indicates that private schools are better able to retain highly skilled teachers when appealing teaching options exist. This is consistent with evidence from the United Kingdom and the United States. Dolton & Van der Klaauw (1999) show that "higher opportunity wages increase the tendency among teachers to switch careers and leave the profession voluntarily," while Corcoran, et al. (2004) find that the rise in employment opportunities for talented women in the United States is responsible for the decline in the quality of the teacher labor force.

Table 1: Model Fit

		Actual	Model
Parents			
Proportion Enrolled in Voucher Schools			
	Overall	51.26%	51.92%
	Primary	49.67%	50.39%
	Secondary	53.18%	53.76%
	Urban	53.28%	54.10%
	Rural	30.01%	29.12%
Mean Tuition (1,000 CLP)		15.25	15.56
Mean Test Scores			
	Municipal Schools	-0.1850	-0.2354
	Voucher Schools	0.1661	0.0829
	Gap Municipal-Voucher	0.3511	0.3182
	Income Gap (top-bottom quartile)	0.6065	0.6522
Potential Teachers			
Proportion Enrolled in			
	Municipal Schools	9.48%	10.65%
	Voucher Schools	7.81%	7.28%
	Non-Teaching Occupations	70.28%	68.97%
	Home	12.44%	13.10%
Mean Accepted Wages (1,000 CLP) in			
	Teaching	4790	4564
	Municipal Schools	5095	4828
	Voucher Schools	4415	4178
	Non-Teaching Occupations	7774	8075

Second, teachers in public schools have lower skills than teachers in private schools. Among potential teachers, type 2 have the lowest teaching skills, as can be seen from parameters $\hat{\alpha}_0^1, \hat{\alpha}_0^2$, and $\hat{\alpha}_0^3$ in the second column of table 4. They represent 41.46 percent of the population and they are more likely to select into public schools, where they represent 45.03 percent of the teacher body, than into private schools, where only 6.50 percent of teachers are of this type. I estimate that if public school students were taught by teachers of the same quality as private school teachers, keeping everything else equal, their test scores would increase, on average, by 0.71 sd.

Third, there are differences in achievement production by student type. Students of type 2 are low-ability students: they obtain the lowest test scores in both types of schools, as shown in table 9, and they account for 47.94 percent of the student population. They are more likely to select into public schools, where they represent 57.53 percent of the student body.

Together, these findings are evidence of positive assortative matching between teachers and students in Chile: in private schools, where more advantaged and higher ability students are found, teachers are more skilled.

Technology of Test Score Production

Parameter estimates indicate that the technology of test score production is such that increasing teacher quality in public schools benefits weaker students the most. I refer to this as sub-modularity. The same is not true in private schools. In public schools, low-ability students (type 2) benefit more than higher ability students from having more skilled teachers. This can be seen in the first column of table 6, where $\hat{\beta}_1^2 > \hat{\beta}_1^1$ and $\hat{\beta}_1^2 > \hat{\beta}_1^3$. In private schools, on the other hand, type 2 students are not those who benefit the most from having better teachers. They benefit slightly more than type 3 students, and considerably less than type 1 students, as can be seen in the third column of table 6. On average, type 2 students (low ability) benefit less than the more able types 1 and 3 grouped together.[35] In the remainder, I refer to this feature of private schools as super-modularity, with the understanding that it holds on average.

The finding of sub-modularity in public schools is suggestive of public school teachers fo-

---

[35]To see this, I create an ability dummy that groups types 1 and 3 together, and estimate the production function in voucher schools using this ability dummy. The interaction between teaching skills and the ability dummy is positive, indicating that, on average, high-ability students benefit more than low-ability students from having more able teachers.

cusing more on the students who are struggling than private school teachers do. I investigate whether this is compatible with evidence on classroom behavior. I obtained a survey administered by SIMCE in 2011. The data were not used in estimation; therefore, they can serve as an out-of-sample corroboration of the model's finding. Math teachers were asked to describe their classroom behavior and their expectations for their students' achievement. Their answers, reported in tables 10 and 11, indicate that public school teachers devote more attention to the lower achieving and lower SES students in the classroom than private subsidized school teachers do. For example, public school teachers are 6 percentage points more likely to always make sure that all students understand, and they are 14 percentage points more likely to explain to everybody the solutions to the tests. As a result, they are more likely to expect that difficult and low SES students will do well in school.

8.2 Simulations

Decomposing the Treatment Effect of Voucher Schools: the Importance of Teacher Quality

The focus of many papers in the school choice literature is to estimate the effect of attending a choice school for those who attend, or the treatment on the treated. Using my estimated model, I simulate the counterfactual test score that a student observed in the voucher school would have obtained in the municipal school. I also simulate the test score that a voucher school student would obtain in the voucher school if voucher school teachers were of the same quality of public school teachers. Specifically, I substitute $\bar{s}_{Vm}$ with $\bar{s}_{Mm}$ in the test score production of the voucher schools in every market m:

$$\tilde{a}_{hVm} = a_{Vm}(X_h, k_h, \bar{s}_{Mm}) + \nu_{hV} \quad \forall m.$$

I then use these counterfactual test scores to estimate two parameters: the treatment on the treated (TT) and the hypothetical treatment on the treated if the voucher school did not have better teachers ($\tilde{TT}$). I estimate a TT of 1.12 sd and a $\tilde{TT}$ of 0.91. The portion of TT that exceeds $\tilde{TT}$ is due to the teacher quality difference between voucher and public schools:

$$1 - \frac{\tilde{TT}}{TT} = 0.19.$$

Welfare Implications of Assortative Matching with School Specialization

Assortative matching of students to teachers is generally considered bad for the achievement of low SES students. However, studies that document assortative matching do not typically have information on test scores and on the production technology. The next counterfactual experiment shows that when public schools specialize in the weakest students, matching assortatively by ability is beneficial to low-ability students.

I simulate the effect of moving from a counterfactual scenario of no school specialization (i.e., both public and voucher schools adopt super-modular technologies) to the baseline case where public schools specialize in the weakest students. To simulate a super-modular production function in the public school, I reshuffle the β_1^{type} parameters of the public school production technology across types. The counterfactual production technology is such that low-ability students benefit less from able teachers than high-ability students: $\beta_1^2 < \beta_1^1 < \beta_1^3$.

Table 12 reports the results of the experiment. Student welfare is obtained by simulating the utility of every student in the sample. Utility depends both on test scores and on consumption, which is lower in private schools because of tuition fee payments. Column one reports test scores, welfare, and student sorting in the baseline scenario. Column two reports outcomes under the counterfactual scenario in which public schools do not specialize in the weakest students. Not surprisingly, student sorting is more assortative by ability in the baseline scenario (column one), because when public schools specialize in the low-ability students, the low-ability students are more likely to select them.

Column three reports outcomes when public schools specialize in the weak students, but student sorting is fixed at the no-specialization case. When matching becomes more assortative under school specialization (i.e., moving from column three to column one), the test scores and welfare of low-ability students improve. In fact, the largest welfare improvements from the more assortative matching accrue to the disadvantaged students. Therefore, assortative matching is not bad for disadvantaged students if the schools with lower ability teachers are also those

26

that specialize in low-ability students. Hence, it is not possible to draw welfare implications on student-teacher matching by using only data on student and teacher characteristics without knowledge of the technology of test score production. This has important policy implications. When the technologies in public and private schools are like the ones in Chile, a policy that reduces assortative matching could decrease the welfare of the students that it intends to help.

Finally, one possible concern of the current configuration of school specialization in Chile is that the focus of public school teachers on struggling students may harm more able students. In the presence of widespread school choice with fellowships for private education, whether this happens is theoretically ambiguous, because able students have the option to leave the public sector. In run a set of counterfactual experiments that compare the outcomes of all four possible combinations of school specialization in the two type of schools (both sub-modular, both super-modular, private school sub-modular and public school super-modular, and vice-versa). Results are reported in table 13. Sub-modularity in public schools harms high-ability students. In general, under each configuration there are winners and losers. For example, high-ability students are better off when both schools adopt super-modular technologies, while low-ability students are better off when both adopt sub-modular technologies.

Impact of Voucher Sector with Endogenous Teacher Quality

I evaluate the impact of school choice expansion with endogenous teacher quality. To do so, I compare student outcomes in the current system with counterfactual student outcomes in a system where there are no voucher schools. Because the private school employment option is not available without a voucher sector, the pool of teachers is different in the counterfactual scenario if those who currently teach in private school would not have become teachers in the absence of a private school sector. This counterfactual should be interpreted as a thought experiment explaining one channel of operation of school vouchers. It is not an evaluation of the voucher plan overall because, as explained in section 4, the model does not include the alternative channel of peer effects.

Table 14 presents the results of the experiment. When the voucher sector is introduced, test scores improve for students at all income levels on average. For example, average test scores of students at the top 25 percent of the income distribution increase by 0.75 sd, while average test

27

scores of students at the bottom 25 percent increase by 0.20 sd. Students at the top are more likely to select into private schools. For example, students below the median income have a 38.9 percent probability of switching to private schools, whereas students above the median income switch at a rate of 65.1 percent.

Disadvantaged students are not harmed on average by the voucher plan for two reasons. First, the overall pool of teachers changes. More skilled individuals enter teaching when there is a private school employment option. Private schools attract highly skilled individuals from outside of teaching and conduct only a limited amount of teacher cream skimming from public schools. Therefore, the quality of public school teachers suffers only a small decline. Second, the students who benefit the most from having skilled teachers are likely to switch to private schools.

The teacher-student sorting channel of vouchers has a positive impact on test scores of students at all income levels, on average. This is because the voucher sector attracts highly skilled individuals from outside of teaching, and it expands the school options of parents, who choose schools in a way that benefits the achievement of their children.

9   Conclusions

Chile's long experience with a large-scale school voucher plan provides a unique opportunity to study school choice. Previous research has documented that school choice affects the sorting of both students and teachers across school sectors, but the literatures on parental school choice and on teacher sorting have remained mostly separate so far. In this paper, I show that studying the sorting decisions of parents and teachers together provides new insights into the impact of school choice on student outcomes.

My results indicate that, for the students who attend private schools in Chile, 19 percent of the improvement in test scores with respect to the test scores they would have obtained in public schools is attributable to the better teachers that they have in private schools. In fact, Chilean private schools successfully recruit more highly skilled teachers by offering higher wages to reward higher skills. An important implication of this fact is that private schools exert only

limited cream skimming of the best teachers from public schools, because the teaching body of the private schools is composed mostly of individuals who would not have become teachers had employment in private schools not been an option. A key consequence of this sorting of higher ability teachers into private schools is that the expansion of the private school sector in Chile led to an improvement in the quality of the pool of teachers. This, combined with the fact that parents choose schools in a way that benefits the achievement of their children, results in improved test scores for students at all income levels. Test scores of students belonging to the bottom 25 percent of the income distribution increase on average by 0.20 sd, while test scores of students at the top increase on average by 0.75 sd. It would be interesting to extend the analysis in this paper to investigate how the quality of public school teachers would change if public school wages were not regulated by rigid government formulae, a feature common to many education systems around the world.[36]

My estimated model implies that private and public Chilean schools specialize in teaching to, respectively, high- and low-ability students. This finding is also validated by a teacher survey not used in my estimation. Given this specialization, introducing policies that reduce assortative matching of students to teachers could harm the students that they are intended to help. In fact, if the schools that have less able teachers are also those that specialize in the weakest students, moving low-ability students to schools with better teachers, who do not target their ability type, would not necessarily benefit the low-ability students.

More generally, it is important to study how teacher characteristics, student characteristics, and endogenous specialization patterns combine to produce student test scores. It would be interesting to relax the model assumption that school specialization is independent of the composition of students. For example, it is possible that teachers target the ability level that is most represented in class. If this is the case, specialization patterns would change as student composition across schools changes. At the same time, it is possible that teachers need re-training to change their focus of instruction. This would introduce frictions that hinder endogenous changes in school specialization. In general, how instruction relates to classroom composition and type

---

[36]Examples include the United States, Sweden, the Netherlands, and Italy. In ongoing work (Tincani 2014b), I extend the model to endogenize wage determination in private schools. This extension permits the analysis of wage policies in public schools, which induce an endogenous reaction of wages in the private sector.

of school is an interesting but largely unexplored question.[37] One of the lessons that we learn from this paper is that it is not advisable to abstract from this relationship when analyzing the welfare implications of changing the allocation of students to teachers.

---

[37]Using data from a randomized tracking experiment in Kenya, Duflo, et al. (2011) conclude that teachers focus on the highest achieving students in the class. However, the authors warn that their finding could reflect incentives that are idiosyncratic to the Kenyan system. Using data from a randomized experiment in Mexico, Behrman, et al. (2014a) estimate a game in effort between teachers and students that yields endogenous changes in teacher behavior following changes in student composition. In ongoing work on peer effects in education (Tincani 2014a), I find preliminary evidence that public school teachers in Chile do not increase their focus on high-ability students when the proportion of high-ability students in the classroom increases exogenously.

Table 2: Household Characteristics by Type of School

Household's characteristics	M	V	Difference
Avg parents' educ (yrs)	9.66	11.92	2.26***
Mother's educ (yrs)	9.60	11.84	2.24***
Hh monthly income (CLP)	169,771	312,320	142,549***
Hh head not working (frac)	9.08%	4.68%	4.40%***
Hh head low-skilled job (frac)	44.50%	22.21%	22.30%***

Source: SIMCE 2006. Three stars indicate a p-value< 0.001 in the t-test of the null hypothesis that the difference in means is zero.

Table 3: Average Monthly Teaching Wages by Teaching Experience and Type of School

texp	wage M	wage V	ratio wage V/wage M
(years)	(2006 CLP)	(2006 CLP)	
≤ 10	368,816.2	423,417.7	1.148
11-20	472,502	472,967.1	1.001
21-30	540,992	544,536.9	1.007
31+	585,682.5	583,352.7	0.996

Source: ELD 2006. The first column indicates teaching experience in years, the second and third columns contain average wages in the municipal and voucher schools, and the last column contains the ratio of the third to second column. 1 USD=545.50 CLP
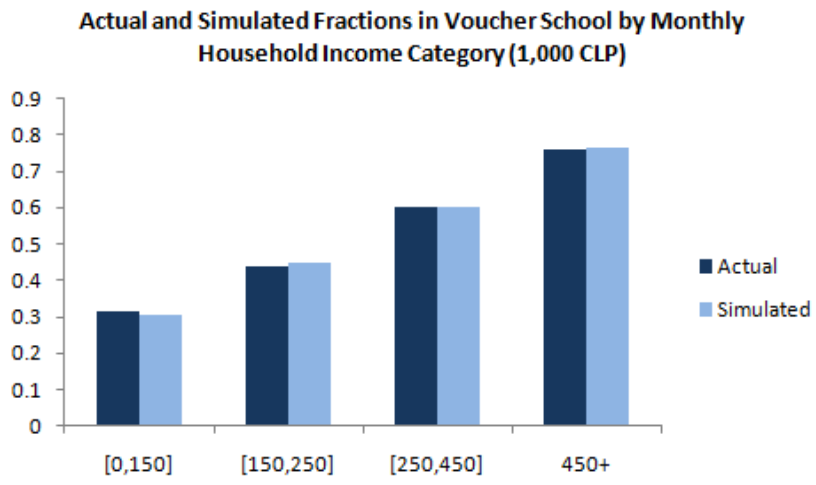
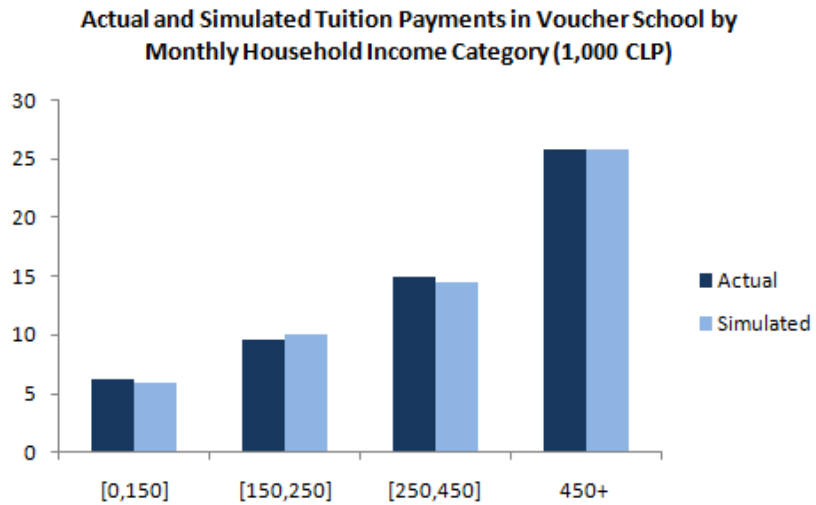## 10 Tables and Figures

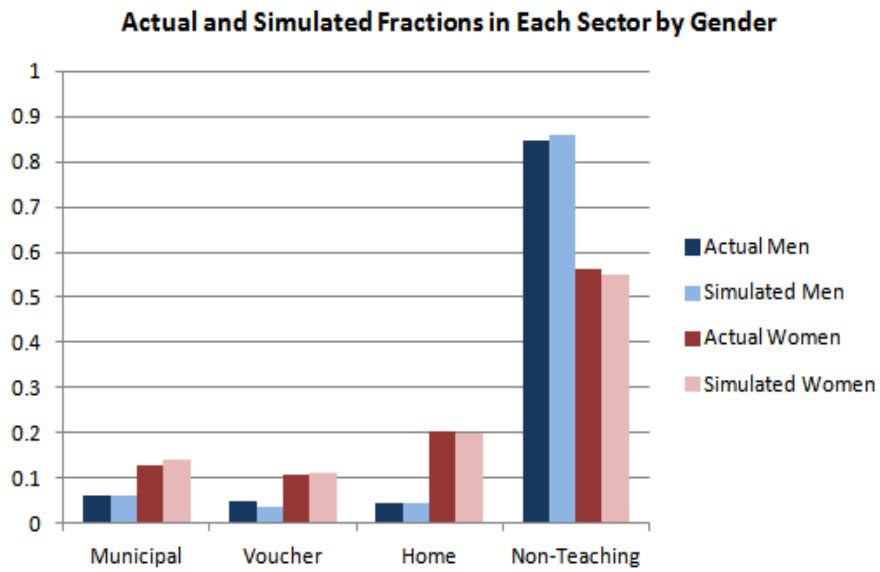Figure 1: Parental Choices by Income

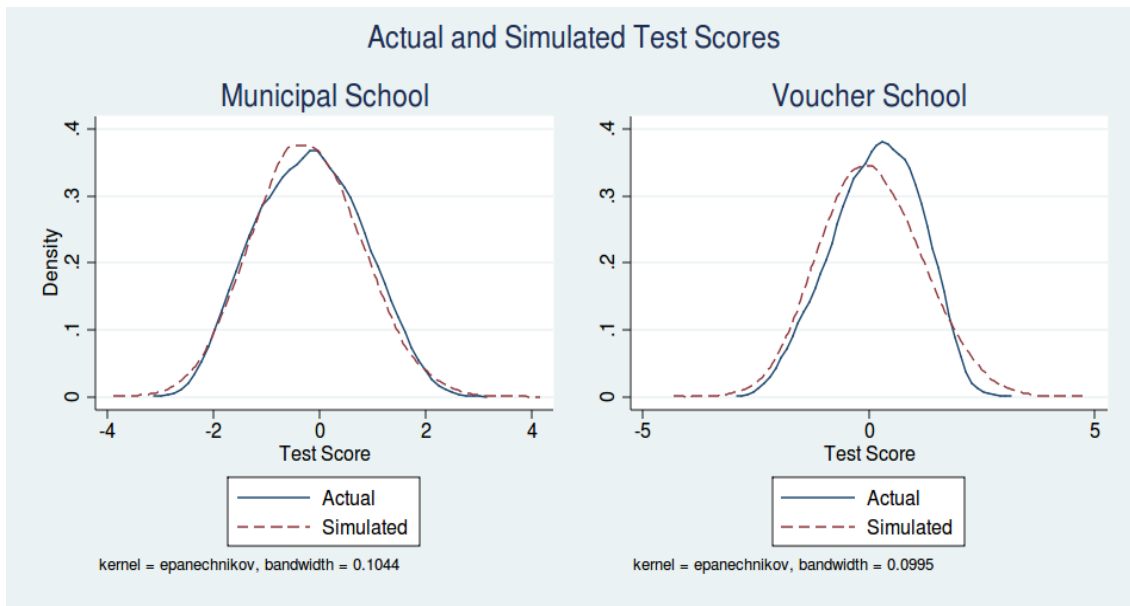Figure 2: Tuition Payments by Income

Figure 3: Occupational Choices by Gender

Figure 4: Test Scores

Table 4: **Parameters of Log-Wage Offer Functions**

Parameter	Description	Municipal	Voucher	Non-Teaching
α_0^1	Intercept, type 1	2.45e-02	6.42e-02	1.38
			(6.33e-04***)	
$\alpha_0^2 - \alpha_0^1$	Intercept type 2 minus type 1	4.72e-02	-1.04	-1.18
		(7.72e-04***)	(4.15e-05***)	
$\alpha_0^3 - \alpha_0^1$	Intercept type 3 minus type 1	-1.27e-02	-1.93e-02	-2.04
		(3.09e-03***)	(2.47e-03***)	
α_1	Age	3.99e-02	8.63e-02	1.01e-02
		(1.08e-03***)	(4.20e-04***)	(3.66e-03***)
α_2	Age Squared	-1.51e-04	-1.65e-03	-4.00e-04
		(2.69e-01)	(2.42e-02)	(9.85e-02)
α_3	Female Dummy	-1.43e-01	-1.71e-01	-1.38e-01
		(2.51e-04***)	(2.39e-04***)	(3.14e-04***)
α_4	Has Professional Certificates	4.25e-01	3.61e-01	-3.13e-02
		(9.41e-05***)	(1.24e-04***)	(1.19e-03***)
α_5	Graduate Degree	4.03e-01	2.71e-01	1.19e-01
		(9.05e-05***)	(1.58e-04***)	(3.29e-04***)
$\log(\sigma)$	Wage shock	-1.22	-8.09e-01	-4.00e-01
		(3.28e-05***)	(4.94e-05***)	(1.33e-04***)

Stars indicate the significance level of a two-sided Wald test. Three stars: 1 percent; two stars: 5 percent; one star: 10 percent; no stars: p-value above 10 percent. The parameters with no standard errors vary by market, only their mean across markets is reported for ease of exposition. Estimates by market and their standard errors are available. The intercept in the voucher sector is parameter a_0 of equation 4 in appendix 12, which does not vary by market.

Table 5: **Parameters of Non-pecuniary Utility of Potential Teachers**

Parameter	Description	Home	Municipal	Voucher
μ_0^1	Intercept type 1	-3.30e+03	-8.00e-01	-9.59e-01
		(1.45e-08***)	(5.54e-05***)	(4.52e-05***)
$\mu_0^2 - \mu_0^1$	Intercept type 2 minus type 1	-2.15e+03	-1.05e-01	5.12e-01
		(1.64e-08***)	(3.26e-04***)	(8.56e-05***)
$\mu_0^3 - \mu_0^1$	Intercept type 3 minus type 1	6.51e+02	-1.55e-01	3.40e-01
		(5.70e-08***)	(2.25e-04***)	(1.35e-04***)
μ_1	Gender	1.64e+03		
		(2.33e-08***)		
μ_2	Gender*N children	3.61e+02		
		(1.10e-07***)		
μ_3	Age	-1.72e+01		
		(2.17e-06***)		
μ_4	N children	1.43e+01		
		(2.44e-06***)		
μ_5	Has children aged 0-2	-1.19e+01		

Parameter	Description	Home	Municipal	Voucher
		(3.21e-06***)		
μ_6	Has children aged 3-6	1.73e+02		
		(2.70e-07***)		
μ_7	Age squared	3.31e-01		
		(1.30e-04***)		
μ_{0Teach}	Non-pecuniary utility from teaching if female	1.00		
		(3.92e-05***)		

Stars indicate the significance level of a two-sided Wald test. Three stars: 1 percent; two stars: 5 percent; one star: 10 percent; no stars: p-value above 10 percent.

Table 6: **Technology of Achievement Production**

Parameter	Description	Municipal	Voucher
β_0^1	Intercept, type 1	-1.85	-1.88
$\beta_0^2 - \beta_0^1$	Intercept type 2 minus type 1	-6.10e-02	-3.00
		(3.58e-02*)	(6.62e-04***)
$\beta_0^3 - \beta_0^1$	Intercept type 3 minus type 1	5.71e-01	-1.13
		(3.52e-03***)	(1.75e-03***)
β_1^1	Teachers' skills, type 1	3.40e-01	2.11e-01

Parameter	Description	Municipal	Voucher
		(5.89e-03***)	(9.06e-03***)
$\beta_1^2 - \beta_1^1$	Teachers' skills, type 2 minus type 1	3.74e-02	-1.95e-01
		(5.32e-02)	(1.08e-02***)
$\beta_1^3 - \beta_1^1$	Teachers' skills, type 3 minus type 1	-2.10e-01	-2.11e-01
		(8.34e-03***)	(9.66e-03***)
β_2^1	Parental education, type 1	5.72e-02	1.03e-01
		(3.15e-02*)	(2.05e-02***)
$\beta_2^2 - \beta_2^1$	Parental education, type 2 minus type 1	-4.71e-02	1.38e-01
		(4.44e-02)	(1.31e-02***)
$\beta_2^3 - \beta_2^1$	Parental education, type 3 minus type 1	4.27e-02	1.19e-02
		(5.36e-02)	(1.18e-01)
β_3^1	Hh income pro capite, type 1	1.55e-01	9.78e-01
		(1.56e-02***)	(2.22e-03***)
$\beta_3^2 - \beta_3^1$	Hh income pro capite, type 2 minus type 1	6.28e-03	2.33e-01
		(2.93e-01)	(8.39e-03***)
$\beta_3^3 - \beta_3^1$	Hh income pro capite, type 3 minus type 1	-1.48e-01	2.81e-01

Parameter	Description	Municipal	Voucher
		(1.24e-02***)	(7.60e-03***)
β_4	Hh income pro capite squared	-4.67e-02	-2.52e-01
		(4.07e-02)	(7.13e-03***)
$\log(\sigma_\nu)$	Shock	-3.22e-02	-1.95e-01
		(6.69e-02)	(8.78e-03***)

Stars indicate the significance level of a two-sided Wald test. Three stars: 1 percent; two stars: 5 percent; one star: 10 percent; no stars: p-value above 10 percent. The parameters β_0^1 have some geographical variation, only the mean is reported. (For identification reasons, they are restricted to be identical in some markets).

Table 7: **Direct Parental Preference for Municipal School and Weight on Consumption**

Parameter	Description	Estimate
η^1	Intercept of preference for Municipal, type 1	-1.12
		(1.53e-03***)
$\eta^2 - \eta^1$	Type 2 minus type 1	7.53e-01
		(2.33e-03***)
$\eta^3 - \eta^1$	Type 3 minus type 1	-7.58e-02
		(2.56e-02***)
η_1	*primary* in preference for Municipal	5.02e-01

Parameter	Description	Estimate
		(4.21e-03***)
η_2	*rural* in preference for Municipal	3.73e-01
		(5.65e-03***)
τ^1	Weight on consumption, type 1	1.18e-01
		(1.70e-02***)
$\tau^2 - \tau^1$	Type 2 minus 1	1.87e-01
		(9.04e-03***)
$\tau^3 - \tau^1$	Type 3 minus type 1	5.57
		(3.26e-04***)
$\log(\sigma_\eta)$	Preference shock	-4.52
		(4.85e-04***)

Stars indicate the significance level of a two-sided Wald test. Three stars: 1 percent; two stars: 5 percent; one star: 10 percent; no stars: p-value above 10 percent.

Table 8: **Fellowship Formula**

Parameter	Description	Estimate
b_0	Intercept	4.48e-01
		(4.99e-03$^{(***)}$)
b_1	*primaria*	6.67e-02

Parameter	Description	Estimate
		(3.07e-03***)
b_2	Family size	1.05e-01
		(1.64e-02***)
b_3	*rural*	-3.25e-01
		(5.96e-03***)
b_4	monthly income	-5.42e-02
		(3.56e-02)
$\log(\sigma_{me})$	Measurement error	-5.91
		(3.21e-04***)

Stars indicate the significance level of a two-sided Wald test. Three stars: 1 percent; two stars: 5 percent; one star: 10 percent; no stars: p-value above 10 percent.

Table 9: Mean Hypothetical Test Scores by Student Type and School Type

School	Type 1	Type 2	Type 3
Municipal	-.6064383	-1.088188	-.0897732
Voucher	.7560885	-1.460198	-.9736694

Table 10: Teacher behavior by school sector

	Municipal (% teachers)	Voucher (% teachers)
Always explain material until all students understand	65	59
Always explain to all students exam solutions on blackboard	62	48
Always explain to all students homework solutions on blackboard	68	59

Source: SIMCE 2011 Math teachers survey.

Table 11: Teacher expectations by school sector

	Municipal (% teachers)	Voucher (% teachers)
Absolutely certain that low SES students will learn	66	64
Absolutely certain that students with low motivation will learn	30	24
Absolutely certain that misbehaving students will learn	26	23
Absolutely certain that students with emotional problems will learn	25	21

Source: SIMCE 2011 Math teachers survey.

Table 12: Simulating Public School Specialization in Weakest Students

Outcome	(1) Baseline sub-M&super-V	(2) Counterfactual super-M&super-V	(3) Counterfactual sub-M&super-V sorting fixed at super-M&super-V
Test scores high ability	0.40	0.63	0.42
Test scores low ability	-0.58	-0.86	-0.62
Test scores top 25	0.33	0.36	0.33
Test scores bottom 25	-0.32	-0.35	-0.33
Test score gap	0.65	0.71	0.66
Welfare high ability	2.43	2.68	2.42
Welfare low ability	-0.40	-0.66	-0.43
Welfare top 25	3.47	3.51	3.46
Welfare bottom 25	-0.44	-0.45	-0.47
Fraction in V high ability	0.60	0.57	0.57
Fraction in V low ability	0.42	0.54	0.54
Fraction in V top 25	0.71	0.73	0.73
Fraction in V bottom 25	0.39	0.44	0.44

Table 13: Simulating All Configurations of Specialization

| | No Specialization | | Specialization | |
| | (1) | (2) | (3) | (4) |
Outcome	super-M&super-V	sub-M&sub-V	sub-M&super-V	super-M&sub-V
Test scores high ability	0.63*	0.11	0.40	0.35
Test scores low ability	-0.86	-0.12*	-0.58	-0.29
Test scores top 25	0.36	0.49	0.33	0.56*
Test scores bottom 25	-0.35	-0.30	-0.32	-0.27*
Test score gap	0.71	0.79	0.65*	0.83
Welfare high ability	2.68*	2.19	2.43	2.37
Welfare low ability	-0.66	0.08*	-0.40	-0.08
Welfare top 25	3.51	3.67	3.47	3.71*
Welfare bottom 25	-0.45	-0.39*	-0.44	-0.39*
Fraction in V high ability	0.57	0.60	0.60	0.52
Fraction in V low ability	0.54	0.63	0.42	0.73
Fraction in V top 25	0.73	0.80	0.71	0.78
Fraction in V bottom 25	0.44	0.49	0.39	0.51

start indicates the best outcome in the row.

Table 14: Simulating Elimination of Voucher Sector

| | Counterfactual | Baseline |
Outcome	w/o vouchers	w vouchers
Test scores M	-0.48	-0.24
Test scores V	-	0.08
Test scores top 25	-0.42	0.33
Test scores bottom 25	-0.52	-0.32
Test score gap	0.10	0.65

References

D. Aaronson, et al. (2007). 'Teachers and student achievement in the Chicago public high schools'. *Journal of Labor Economics* **25**(1):95–135.

J. G. Altonji, et al. (2005a). 'An evaluation of instrumental variable strategies for estimating the effects of catholic schooling'. *Journal of Human Resources* **40**(4):791–821.

J. G. Altonji, et al. (2005b). 'Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools'. *Journal of Political Economy* pp. 151–184.

J. G. Altonji, et al. (2014). 'Estimating the cream skimming effect of school choice'. *forthcoming Journal of Political Economy* .

P. Anand, et al. (2009). 'Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile'. *Economics of Education Review* **28**(3):370–381.

J. Behrman, et al. (2014a). 'Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools'. *forthcoming Journal of Political Economy* .

J. Behrman, et al. (2014b). 'The Impact of School Voucher Systems on Teacher Quality in Public and Private Schools: The Case of Chile'. *mimeo, U. of Pennsylvania* .

Y. Ben-Porath (1967). 'The production of human capital and the life cycle of earnings'. *The Journal of Political Economy* pp. 352–365.

J. R. Betts, et al. (2003). *Determinants of student achievement: New evidence from San Diego.* Public Policy Institute of California San Francisco.

D. Bhattacharya (2005). 'Asymptotic inference from multi-stage samples'. *Journal of econometrics* **126**(1):145–171.

S. Bonhomme, et al. (2012). *Job Characteristics and Labor Turnover: Assessing the Role of Preferences and Opportunities in Teacher Mobility.* Centre for Economic Policy Research.

G. J. Borjas (1987). 'Self-Selection and the Earnings of Immigrants'. *The American Economic Review* **77**(4):531–553.

G. J. Borjas (2002). 'The wage structure and the sorting of workers into the public sector'. Tech. rep., national Bureau of Economic Research.

D. Boyd, et al. (2005). 'Explaining the short careers of high-achieving teachers in schools with low-performing students'. *The American economic review* **95**(2):166–171.

D. Boyd, et al. (2006). 'Analyzing the determinants of the matching public school teachers to jobs: Estimating compensating differentials in imperfect labor markets'. Tech. rep., National Bureau of Economic Research.

D. Bravo (1999). 'Competencia y calidad de la educación en Chile: una revisión de la literatura'. *Educación Particular Subvencionada. Un aporte a la modernización de la Educación Chilena. Santiago, CONACEP* .

D. Bravo, et al. (2010). '¿ Se premia la habilidad en el mercado laboral docente?¿ Cuánto impacta en el desempeño de los estudiantes?'. *Serie Documentos de Trabajo* .

W. A. Brock & S. N. Durlauf (2001). 'Discrete choice with social interactions'. *The Review of Economic Studies* **68**(2):235–260.

C. T. Clotfelter, et al. (2006). 'Teacher-student matching and the assessment of teacher effectiveness'. *Journal of Human Resources* **41**(4):778–820.

S. Corcoran, et al. (2004). 'Women, the labor market, and the declining relative quality of teachers'. *Journal of Policy Analysis and Management* **23**(3):449–470.

J. A. Correa, et al. (2014). 'Self-Selection in the Market of Teachers'. *mimeo* .

T. S. Dee (2005). 'A teacher like me: Does race, ethnicity, or gender matter?'. *American Economic Review* pp. 158–165.

A. K. Dills (2005). 'Does cream-skimming curdle the milk? A study of peer effects'. *Economics of Education Review* **24**(1):19–28.

P. Dolton & W. Van der Klaauw (1999). 'The turnover of teachers: A competing risks explanation'. *Review of Economics and Statistics* **81**(3):543–550.

E. Duflo, et al. (2011). 'Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya.'. *American Economic Review* **101(5)**:1739–74.

S. Dynarski, et al. (2009). 'Cheaper by the dozen: Using sibling discounts at Catholic schools to estimate the price elasticity of private school attendance'. Tech. rep., National Bureau of Economic Research.

D. Epple & R. Romano (1998). 'Competition between private and public schools, vouchers, and peer-group effects'. *American Economic Review* pp. 33–62.

D. Epple & R. Romano (2008). 'Educational Vouchers and Cream Skimming'. *International Economic Review* **49**(4):1395–1435.

D. Epple & R. Romano (2011). 'Peer effects in education: A survey of the theory and evidence'. *Handbook of social economics* **1**(11):1053–1163.

D. Epple & H. Sieg (1999). 'Estimating Equilibrium Models of Local Jurisdictions'. *Journal of Political Economy* **107**(4):645–681.

W. N. Evans & R. M. Schwab (1995). 'Finishing high school and starting college: Do Catholic schools make a difference?'. *The Quarterly Journal of Economics* **110**(4):941–974.

M. M. Ferreyra (2007). 'Estimating the Effects of Private School Vouchers in Multidistrict Economies'. *The American Economic Review* **97**(3):pp. 789–817.

J. Grogger, et al. (2000). 'Further Evidence on the Effects of Catholic Secondary Schooling [with Comments]'. *Brookings-Wharton papers on urban affairs* pp. 151–201.

E. Hanushek, et al. (2005). 'The market for teacher quality'. *NBER Working Paper* .

E. Hanushek & S. Rivkin (2006). 'Teacher quality'. *Handbook of the Economics of Education* **2**:1051–1078.

J. S. Hastings & J. M. Weinstein (2008). 'Information, school choice, and academic achievement: Evidence from two experiments'. *The Quarterly Journal of Economics* **123**(4):1373–1414.

J. Heckman & B. Honore (1990). 'The empirical content of the Roy model'. *Econometrica: Journal of the Econometric Society* pp. 1121–1149.

J. Heckman & G. Sedlacek (1985). 'Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market'. *The journal of political economy* **93**(6):1077–1125.

J. Heckman & G. Sedlacek (1990). 'Self-selection and the distribution of hourly wages'. *Journal of Labor Economics* pp. 329–363.

J. Heckman & B. Singer (1984). 'A method for minimizing the impact of distributional assumptions in econometric models for duration data'. *Econometrica: Journal of the Econometric Society* pp. 271–320.

L. Hensvik (2012). 'Competition, Wages and Teacher Sorting: Lessons Learned from a Voucher Reform*'. *The Economic Journal* **122**(561):799–824.

C. Hoxby (2003). 'School Choice and School Productivity. Could school choice be a tide that lifts all boats?'. *in The Economics of School Choice, University of Chicago Press* .

C. M. Hoxby (2002). 'Would School Choice Change the Teaching Profession?'. *The Journal of Human Resources* **37**(4):pp. 846–891.

C. Hsieh & M. Urquiola (2006). 'The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program'. *Journal of public Economics* **90**(8-9):1477–1503.

C. K. Jackson (2012). 'School competition and teacher labor markets: Evidence from charter school entry in North Carolina'. *Journal of Public Economics* **96**(5):431–448.

C. K. Jackson (2013). 'Match quality, worker productivity, and worker mobility: Direct evidence from teachers'. *Review of Economics and Statistics* .

H. Ladd (2002). 'School vouchers: a critical view'. *The Journal of Economic Perspectives* **16**(4):3–24.

H. Lankford (1999). 'A descriptive analysis of the New York State and New York City teaching force'. *Report prepared for the New York Supreme Court case Campaign for Fiscal Equity v. New York State* .

H. Lankford, et al. (2002). 'Teacher sorting and the plight of urban schools: A descriptive analysis'. *Educational evaluation and policy analysis* **24**(1):37–62.

C. Manski (1987). 'Academic ability, earnings, and the decision to become a teacher: Evidence from the National Longitudinal Study of the High School Class of 1972'.

C. F. Manski (1993). 'Identification of endogenous social effects: The reflection problem'. *The review of economic studies* **60**(3):531–542.

P. McEwan, et al. (2008). 'School Choice, Stratification, and Information on School Performance: Lessons from Chile'. *Economia* pp. 1–42.

D. McFadden (1989). 'A method of simulated moments for estimation of discrete response models without numerical integration'. *Econometrica: Journal of the Econometric Society* pp. 995–1026.

R. A. Moffitt et al. (2001). 'Policy interventions, low-level equilibria, and social interactions'. *Social dynamics* **4**:45–82.

D. Neal (1995). 'The effect of Catholic secondary schooling on educational attainment'. *Journal of labor Economics* pp. 98–123.

T. J. Nechyba (2000). 'Mobility, targeting, and private-school vouchers'. *American Economic Review* pp. 130–146.

A. Pakes & D. Pollard (1989). 'Simulation and the asymptotics of optimization estimators'. *Econometrica* **57**(5):1027–1057.

S. Rivkin, et al. (2005). 'Teachers, schools, and academic achievement'. *Econometrica* pp. 417–458.

C. Rothschild & F. Scheuer (2013). 'Redistributive taxation in the roy model'. *The Quarterly Journal of Economics* **128**(2):623–668.

C. E. Rouse & L. Barrow (2009). 'School vouchers and student achievement: Recent evidence and remaining questions'. *Annu. Rev. Econ.* **1**(1):17–42.

A. Roy (1951). 'Some thoughts on the distribution of earnings'. *Oxford Economic Papers* pp. 135–146.

B. Sacerdote (2011). 'Peer effects in education: How might they work, how big are they and how much do we know thus far?'. *Handbook of the Economics of Education* **3**:249–277.

G. H. Sahlgren (2013). 'Incentivising excellence: school choice and education quality'. *London: Centre for Market Reform of Education and Institute of Economic Affairs* .

C. Sapelli & B. Vial (2002). 'The performance of private and public schools in the Chilean voucher system'. *Cuadernos de economía* **39**(118):423–454.

T. Stinebrickner (2001a). 'Compensation policies and teacher decisions'. *International Economic Review* **42**(3):751–780.

T. R. Stinebrickner (2001b). 'A dynamic model of teacher labor supply'. *Journal of Labor Economics* **19**(1):196–230.

M. Tincani (2014a). 'Social Incentives in the Classroom: Evidence from a Natural Experiment'. *mimeo, University College London* .

M. Tincani (2014b). 'Teacher Labor Markets, School Vouchers and Student Cogntive Achievement: Evidence from Chile'. *mimeo, University College London* .

M. Urquiola (2005). 'Does school choice lead to sorting? Evidence from Tiebout variation'. *American Economic Review* pp. 1310–1326.

11 Functional Forms

The fellowship formula is:

$$f(Z_h) = b_o + b_1 primary_h + b_2 fam\_size_h + b3 rural_h + b_4 y_h$$

where $primary_h = 1$ if the child is in primary school, and $= 0$ if she is in secondary school; fam_size_h is the family size; $rural_h = 1$ if the family lives in a rural area; and y_h is household monthly income. In estimation, I assume that the fellowship is measured with error: $\tilde{f} = f + me$ with $me \sim N(0, \sigma_{me}^2)$.

Child achievement in school sector j is equal to:

$$a_{hjm} = \beta_{0j}(k_h) + \beta_{1j}(k_h)\bar{s}_{jm} + \beta_{2j}(k_h)peduc_h + \beta_{3j}(k_h)py_h + \beta_{4j}py_h^2 + \nu_{hj} \tag{3}$$

where $peduc_h$ is parental education in years (average between mother's and father's education) and py_h is household monthly income pro capite.

The choice-specific utilities of potential teachers are:

$$u_{imj} = \begin{cases} ln(w_{imj}) + \mu_{0j}(l_i) + \mu_{0Teach}female_i & \text{for j=M,V} \\ ln(w_{imj}) & \text{for j=NT} \\ \mu_{0j}(l_i) + \mu_1 female_i + \mu_2 female_i * nk_i + \\ + \mu_3 age_i + \mu_4 nk_i + \mu_5 nk02_i + \mu_6 nk36_i + \mu_7 age_i^2 + \epsilon_{iH}^{pref} & \text{if j=H} \end{cases}$$

where nk is the number of children, $nk02$ and $nk36$ is the number of children between the ages of zero and two and of three and six, respectively.

In the empirical implementation, the variables entering the wage offer function are $X_i = [age_i, age_i^2, female_i, cert_i, grad\_deg_i]$ where $cert_i$ is a dummy for whether the individual holds professional certificates and $grad\_deg_i$ is a dummy for whether the individual has graduate degrees (master's or Ph.D).

12 Derivation of Equilibrium Teacher Quality by School Sector

To compute the mean teaching skills supplied to the voucher sector in each m, I derive the density of teaching skills conditional on the voucher school being chosen, which in general is different from the population density of teaching skills. Recall that the teaching skills of individual i are:

$$s_i = exp\left(a_0(l_i) + a_1'X_i + \epsilon_i^{tech}\right) \tag{4}$$

with $\epsilon_i^{tech} \sim N(0, \sigma_V^2)$. That is, conditional on type, skills are log-normally distributed. Conditional on $X_i = x$, the density of teaching skills depends both on the density of the shock ϵ_i^{tech} and on the type probability ψ_{l_i}:[38]

$$f^s(s_i|x) = \frac{\psi_{l_i}}{s_i \sigma_V \sqrt{2\pi}} exp\left\{-\frac{(lns_i - a_0(l) - a'x)^2}{2\sigma_V^2}\right\}.$$

The population density is obtained by integrating over the distribution of x in market m, $f_m^x(x)$:

$$f_m^s(s_i) = \int \frac{\psi_{l_i}}{s_i \sigma_V \sqrt{2\pi}} exp\left\{-\frac{(lns_i - a_0(l) - a'x)^2}{2\sigma_V^2}\right\} f_m^x(x)dx.$$

To derive the density of teaching skills in the voucher school, define $A(q, \epsilon_i^{tech}, l_i)$ to be the subset of $\mathbb{R}^3$ that is such that if $\epsilon_i^{-tech} = [\epsilon_i^M \quad \epsilon_i^{NT} \quad \epsilon_i^H]' \in A(q, \epsilon_i^{tech}, l_i)$, an individual with characteristics q, shock realization ϵ_i^{tech}, and type realization l_i chooses the voucher school. Letting $Pr_m(V)$ denote the proportion of individuals choosing sector V in market m, the density of teaching skills in sector V may be written as:

$$g_m^V(s_i|\text{sector V chosen}) = \frac{1}{Pr_m(V)}\psi_{l_i}\int_{\epsilon_i^{-tech} \in A} f_m^s(s_i)f^{-tech}\left(\epsilon_i^{-tech}\right)d\epsilon_i^{-tech}$$

where I let $\int_{\epsilon_i^{-tech} \in A}$ denote multiple integration with respect to $\epsilon_i^M, \epsilon_i^{NT}, \epsilon_i^H$ over the area $\epsilon_i^{-tech} \in A(q, \epsilon_i^{tech}, l_i)$ and where the joint density of the shocks in sectors M, NT and H is:

---

[38]If $ln(x) \sim N(0, \sigma^2)$, x has density $\frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}$ with $x \geq 0$.

$$f^{-tech}\left(\epsilon_i^{-tech}\right) = \frac{1}{\sigma_M \sigma_{NT} \sigma_H} \phi\left(\frac{\epsilon_i^M}{\sigma_M}\right) \phi\left(\frac{\epsilon_i^{NT}}{\sigma_{NT}}\right) \phi\left(\frac{\epsilon_i^H}{\sigma_H}\right).$$

The density of teaching skills in the municipal school, $g_m^M(s_i|\text{sector M chosen})$, can be derived in a similar way.[39]

The mean skills supplied to each sector in market m are obtained using the conditional densities g_m^M, g_m^V:

$$\bar{s}_{Mm} = \sum_{l_i} \psi_{l_i} \int s_i g_m^M(s_i|\text{sector M chosen}) d\epsilon_i^{tech}$$

$$\bar{s}_{Vm} = \sum_{l_i} \psi_{l_i} \int s_i g_m^V(s_i|\text{sector V chosen}) d\epsilon_i^{tech}. \tag{5}$$

13 Market Boundaries

To define market boundaries, I analyzed mobility of parents and teachers.[40] Market boundaries must be such that the mobility across them is close to zero. On the one hand, larger markets guarantee small across-market mobility. On the other hand, the larger the markets, the fewer of them there are. Choosing market size, therefore, presents a trade-off in terms of sample sizes: a large within-market sample size is obtained by having a small number of large markets, whereas a large across-market sample size is obtained by having a large number of small markets. The design of market boundaries attempts to strike a balance between across- and within-market sample sizes, while guaranteeing that the markets are closed.

The unique geographical configuration of Chile aided in the design of boundaries: the country occupies a narrow but long coastal strip, where mobility between northern and southern regions is hindered. With a total area of $291,933$ square miles ($756,102$ km^2), Chile is larger than all U.S. states except Alaska and larger than all countries in the European Union, its size being comparable to that of Turkey. Yet, it extends $2,653$ miles ($4,270$ km) from north to south,

---

[39]First define the proportion of potential teachers choosing the municipal school in market m, $Pr_m(M)$. Then define the area $B(q, \epsilon_i^{tech}, l_i)$ that is such that if $[\epsilon_i^M \quad \epsilon_i^{NT} \quad \epsilon_i^H]' \in B(q, \epsilon_i^{tech}, l_i)$, an individual with characteristics q, shock realization ϵ_i^{tech}, and type realization l_i chooses the municipal school.

[40]Home and work location are available only in the ELD dataset of teachers. I assume that the mobility of non-teachers is similar.

and it averages only 110 miles (177 km) from east to west. I exploit this unique geographical configuration to identify closed labor and educational markets. Table 15 reports the region in which each market lies, as well as average wages in the non-teaching sector (opportunity wages), the fraction of students attending voucher schools, and the fraction of teachers employed in voucher schools. On average, in each market there are 135 municipal and 115 public schools per school level (primary, secondary). Movement of individuals across markets is negligible.

Table 15: Markets, Opportunity Wages and Market Shares of Private Schools

Market	Region	Avg Non-Teaching wages (1,000 CLP, 1CLP≈ $500USD$)	% Students in V	% Teachers in V
1	Arica and Parinacota	441	62	53
2	Coquimbo	662	45	41
3	Libertador G. B. O'Higgins	1005	38	36
4	Atacama	771	38	31
5	Maule	710	39	30
6	Biobío	712	33	22
7	Biobío	633	43	38
8	Los Ríos	614	40	28
9	Los Lagos	581	35	32
10	Los Lagos	416	47	48
11	Antofagasta	652	34	29
12	Libertador G. B. O'Higgins	437	37	24
13	La Araucanía	665	54	46
14	La Araucanía	474	64	55
15	Región Metropolitana (Santiago)	864	67	60
16	Valparaíso	679	56	48
17	Biobío	481	48	33
18	Magallanes and Antártica	828	41	30

14   List of Moment Conditions

I compute 607 moments, 321 pertaining to parents and 286 to potential teachers.

14.1   Parents' Moments: Matching Choices, Test Scores and Fellowship Amounts

I use the following categories:

- family size $nfam_h$: $[2,3], [4,6], \geq 7$

- monthly income in terms of CLP100, 000 y_h: $[0, 0.5]$, $(0.5, 1.5]$, $(1.5, 2.5]$, $(2.5, 3.5]$, $(3.5, 4.5]$, $(4.5, 5.5]$, $(5.5, 7]$, $(7, 9]$, $(9, 11]$, > 11

- average parental education in years $peduc_h$: $[0-6.5]$, $(6.5, 8]$, $(8, 9.5]$, $(9.5, 10.5]$, $(10.5, 11.5]$, $(11.5, 12]$, $(12, 12.5]$, $(12.5, 13]$, $(13, 14]$, > 14

- monthly income in terms of CLP100, 000 divided by family size, $\frac{y_h}{nfam_h}$: $[0, 0.15]$, $(0.15, 0.25]$, $(0.25, 0.36]$, $(0.36, 0.45]$, $(0.45, 0.50]$, $(0.50, 0.70]$, $(0.70, 0.84]$, $(0.84, 1.13]$, $(1.13, 1.75]$, > 1.75

I partition the state of observable exogenous variables and build an indicator for whether an observation belongs to a certain element of the partition. The moment conditions are obtained by multiplying the difference between actual and predicted outcomes by this indicator. The moment conditions are built on the following outcomes (number of moment conditions in parentheses):

- Test scores by sector and by:

  - market (18x2=36)

  - monthly income per capita and parental education (10x10x2=200)

- Fraction choosing voucher school by:

  - market (18)

  - parental education (10)

  - monthly income (10)

  - number of individuals in the family (3)

  - elementary school (2)

  - rurality of the household's residence (2)

- Private school tuition payments made by parents by:

  - elementary school, number of individuals in the household, rurality of the residence (2x3x2=12)

- monthly income (10)

- market (18)

Total number of parents' moments: 321.

14.2 Potential Teachers' Moments: Matching Choices and Accepted Wages

I use the following categories:

- coarse age, age_i: $[20-30]$, $[31-40]$, $[41, 50]$, ≥ 51

- fine age, age_i: $[20, 31]$, $(31, 36]$, $(36, 39]$, $(39, 45]$, $(45, 48]$, $(48, 52]$, $(52, 56]$, > 56

- number of children in the household, $nkids_i$: 0,1,2,$\geq$ 3

- number of children aged 0-2, $nkids2_i$: 0, ≥ 1

- number of children aged 3-6, $nkids3-6_i$: 0, ≥ 1

I partition the state of observable exogenous variables and build an indicator for whether an observation belongs to a certain element of the partition. The moment conditions build on the following outcomes (number of moment conditions in parentheses):

- Accepted wages by sector (3 working options) and by:

  - age, gender, professional certifications (3x4x2x2=48)

  - graduate degree (3x2=6)

  - market (3x18=54)

- Fractions in sector M, V and NT (exclude one sector to avoid multicollinearity and hence singularity of the variance-covariance matrix of the moment conditions) by:

  - professional certifications (3x2=6)

  - age, gender, graduate degree (3x4x2x2=48)

  - market (3x18=54)

- gender, number of kids (3x2x4=24)

- number of kids up to 2 years of age, age (3x2x4=24)

- number of kids of age 3 to 6 (3x2=6)

- Accepted wages in the teaching occupations (2) by finer age category (2x8=16)

Total number of potential teachers' moments: 286.

15 Criterion Function and Asymptotic Properties of the Estimator when Combining Multiple Data Sources

Consider the population moment condition based on outcome y_i:

$$E\left[(y_i - \hat{y}_i(\omega_i, \theta))I_i(\omega_i, y_i \text{ non-missing})\right]$$

and suppose that there are M moment conditions $\{m_i^1, ..., m_i^M\}$ with

$$m_i^m = (y_i^m - \hat{y}_i^m(\omega_i^m, \theta))I_i(\omega_i^m, y_i^m \text{ non-missing}).$$

Let m_i be a vector that stacks all moment conditions for individual i. Assume that the population is divided in two strata: the stratum of students, with mass H_A, and the stratum of college graduates, with mass H_B. The M population moment conditions are:

$$H_A E_A[m_i] + H_B E_B[m_i]$$

where $E_A[\cdot]$ and $E_B[\cdot]$ represent within-stratum expectations.

Let n_A be the sample size of students and n_B be the sample size of potential teachers, and let $m_i(\theta)$ be the $M \times 1$ vector of empirical moment conditions computed at a parameter value θ. The sample analog of the population moment conditions is:

$$H_A \frac{1}{n_A} \sum_{i \in A} w_i m_i(\theta) + H_B \frac{1}{n_B} \sum_{i \in B} w_i m_i(\theta)$$

where w_i are weights provided with the datasets that are used to reweight the sample back to random sampling proportions, and that are normalized to sum to n_A and n_B.[41] Let $n = n_A + n_B$ and pre-multiply the sample moments by $\frac{n}{n}$. Denote the vector of empirical moments based on a sample of size n by $m_n(\theta)$:

$$m_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} (H_A a_A w_i m_i(\theta) I(i \in A) + H_B a_B w_i m_i(\theta) I(i \in B))$$

where $a_A = \frac{n}{n_A}$, $a_B = \frac{n}{n_B}$ and $I(\cdot)$ is an indicator function equal to 1 if the expression in parentheses is true.

The method of simulated moments finds the vector θ that minimizes the weighted distance of the empirical moment conditions from zero:

$$\hat{\theta}_{MSM} = arg \min_{\theta} m_n(\theta)' W_n m_n(\theta) \tag{6}$$

where W_N is an $M \times M$ symmetric positive definite weighting matrix such that as $n \to \infty$, $W_n \to W$ in probability with W symmetric and positive definite.

To derive the asymptotic properties of the estimator, I let $n_A, n_B \to \infty$ with $\frac{n_A}{n} \to a_A < \infty$ and $\frac{n_B}{n} \to a_B < \infty$ as in Bhattacharya (2005), who studies the asymptotic properties of the generalized method of moments with a stratified sample. The MSM estimator defined in (6) is consistent and asymptotically normal:

$$\sqrt{n}(\hat{\theta} - \theta) \Rightarrow N(0, Q)$$

with $Q = (\Gamma' W_n \Gamma)^{-1} \Gamma' W_n V W_n \Gamma (\Gamma' W_n \Gamma)^{-1}$ and $\Gamma = E[\frac{\partial m(\theta)}{\partial \theta'}]$. V is the variance covariance matrix of the moment vector.[42]

---

[41]For SIMCE observations the weights are all equal to one because the SIMCE sample is a simple random sample.

[42]The optimal weighting matrix is the inverse of the variance covariance matrix of the moment conditions, $W_n^* = V^{-1}$. The asymptotic variance reduces to $(\Gamma' V^{-1} \Gamma)^{-1}$ when the optimal weighting matrix is used. I cannot adopt the optimal weighting matrix because the variance covariance matrix is a high-order sparse matrix that cannot be numerically inverted. The inverse of the variance covariance matrix must be obtained to compute the standard errors of the efficient MSM estimator. This negative result is standard in numerical methods. I adopt a weighting matrix that contains the variances of the moments on the main diagonal and zeros elsewhere. This matrix is easily invertible.

To estimate consistently the asymptotic variance of the estimator, I substitute V with a consistent estimate $\hat{V}$ computed at $\hat{\theta}_{MSM}$. The estimator includes a stratum correction that accounts for the sampling design.[43] The estimator of the variance covariance matrix is:

$$
\begin{aligned}
\hat{V} \quad = \quad & \sum_{i \in A} \left(\frac{H_A}{n_A} w_i \right)^2 m_i(\hat{\theta}_{MSM}) m_i(\hat{\theta}_{MSM})' + \sum_{i \in B} \left(\frac{H_B}{n_B} w_i \right)^2 m_i(\hat{\theta}_{MSM}) m_i(\hat{\theta}_{MSM})' \qquad (7) \\
& - \frac{1}{n_A} \left(\sum_{i \in A} \frac{H_A}{n_A} w_i m_i(\hat{\theta}_{MSM}) \right) \left(\sum_{i \in A} \frac{H_A}{n_A} w_i m_i(\hat{\theta}_{MSM}) \right)' \\
& - \frac{1}{n_B} \left(\sum_{i \in B} \frac{H_B}{n_B} w_i m_i(\hat{\theta}_{MSM}) \right) \left(\sum_{i \in B} \frac{H_B}{n_B} w_i m_i(\hat{\theta}_{MSM}) \right)'
\end{aligned}
$$

where $m_i(\hat{\theta}_{MSM})$ is the $M \times 1$ vector of individual-level moment conditions computed at $\hat{\theta}_{MSM}$. To estimate consistently the matrix of moments' partial derivatives, I use:

$$
\hat{\Gamma} = H_A \frac{1}{n_A} \sum_{i \in A}^{N} w_i \frac{\partial m_i}{\partial \theta'} |_{\hat{\theta}_{MSM}} + H_B \frac{1}{n_B} \sum_{i \in B}^{N} w_i \frac{\partial m_i}{\partial \theta'} |_{\hat{\theta}_{MSM}}
$$

where the differentiation is numerical. Letting Δ_t denote a vector of the same size as the parameter vector with zeros everywhere and $\delta > 0$ as its tth element, the derivative of the mth element of $m_i(\theta)$ with respect to the tth element of θ is computed as:

$$
\frac{\partial \hat{m}_i^m(\theta)}{\partial \theta_t} |_{\theta = \hat{\theta}_{MSM}} = \frac{\hat{m}_i^m(\theta + \Delta_t) - \hat{m}_i^m(\theta)}{\delta} |_{\theta = \hat{\theta}_{MSM}}.
$$

16 Estimation Algorithm Embedding Equilibrium Restriction

- Choose an initial guess for the parameter: $\theta^{(0)}$.

- Draw unobserved types for each potential teacher and student.

- Use $a^{(0)} = [a_0^{(0)}(l_i) \quad a_1^{(0)}]'$ to compute teaching skills for each potential teacher i: $s_i\left(a^{(0)}\right)$.

- Calculate the optimal occupational choice of each teacher in each market and use these individual choices to calculate the mean skills supplied to each school sector $j \in \{M, V\}$ in

---

[43]The correction term is derived and discussed in Bhattacharya (2005). Intuitively, ignoring the fact that observations come from two separate strata would over-estimate the between-strata variances.

each market m: $\bar{s}_{jm}$. This is the non-linear function of $a^{(0)}$ in 4 in appendix 12. Simulate also accepted wages.

- Plug the values for mean teacher skills into the production functions for achievement in M and V.

- Simulate achievement of each student in each school, and simulate optimal parental choice of school.

- Compute value of objective function of the Method of Simulated Moments using simulated and real data.

- Update guess $\theta^{(0)}$ to $\theta^{(1)}$ (using the Generating Set Search optimization algorithm in HOPSPACK) and repeat until objective function is minimized.

Notice that even though the simulated skills $\bar{s}_{jm}$ vary at each parameter iteration, they are a non-linear function of $a^{(0)}$. Therefore, the coefficient on teaching skills in the cognitive achievement production function is separately identified from the vector $a^{(0)}$.