The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Preferences, Selection, and Value Added: A Structural Approach Applied to Turkish Exam High Schools

Ş. Pelin Akyol
Verónica Frisancho
Kala Krishna
Cemile Yavaş

CESIFO WORKING PAPER NO. 4302
CATEGORY 5: ECONOMICS OF EDUCATION
JUNE 2013

Presented at CESifo Area Conference on Applied Microeconomics, March 2013

CESifo *Working Paper No. 4302*

Preferences, Selection, and Value Added: A Structural Approach Applied to Turkish Exam High Schools

Abstract

What do applicants care about when choosing a school in Turkey? Are their preferences vertical or horizontal? Which school attributes seem to matter? Do selective schools contribute to their students' success, or is their performance attributable to the higher ability of the students they accept? By taking a structural approach we answer all these questions in one go. We find that students seem to infer quality from past performance in the University entrance exam. There also seems to be a consumption value of attending elite schools and schools add very different value and this is unrelated to their selectivity.

JEL-Code: I200, I210.

Ş. Pelin Akyol
Department of Economics
The Pennsylvania State University
University Park / PA / USA

Verónica Frisancho
The Inter American Development Bank
Washington DC 20009 / USA
vfrisancho@iadb.org

Kala Krishna
Department of Economics
The Pennsylvania State University
University Park / PA / USA
kmk4@psu.edu

Cemile Yavaş
Department of Economics
The Pennsylvania State University
University Park / PA / USA
cemile@psu.edu

June 12, 2013

"The C student from Princeton earns more than the A student from Podunk not mainly because he has the prestige of a Princeton degree, but merely because he is abler. The golden touch is possessed not by the Ivy League College, but by its students."

Shane Hunt, "Income Determinants for College Graduates and the Return to Educational Investment," Ph.D. thesis, Yale University, 1963, p. 56.

1 Introduction

In much of the world, elite schools are established, and often subsidized, by the government in order to bring out the potential of the best and brightest. These schools are fiercely competitive and entry is based on performance in open competitive exams. Such exams often create enormous stress as applicants attempt to improve their performance in the entrance exam in the belief that acceptance will have positive consequences for their lives. Such students, it is argued, will be able to perform at a higher level by going to the elite school where they are challenged by more difficult material and exposed to better peers. What actually happens? Graduates of these elite high schools, without a doubt, do better in college entrance exams, and are more likely to be placed at the best university programs available. But is this due to selection or value added? It is quite possible that the success of the graduates of elite schools creates the (possibly false) belief that they have value added, which in turn results in better students sorting into them, with the result that the graduates of elite schools do better, perpetuating this belief system.

The usual way of ranking schools is in terms of how hard they are to get into (in terms of a performance measure like the SATs in the US[1]) or in terms of how well students that graduate from them do (in terms of some other performance measure like wages, eminence in later life, or admission into further schooling). However, schools may do well in both of

[1]Schools sometimes are less than honest: some inflate their statistics on the performance of their entering class. Some game the system by staying small, and thereby having high SATS and looking very selective. See "Academic integrity should count in rankings" in the Kansas City Star, 2/8/2013. http://www.centredaily.com/2013/02/12/3499088/editorial-academic-integrity-should.html

these dimensions merely because they admit good students and not because they provide value added and thereby *improve* the performance of the students they admit.[2]

In this paper we ask: What do applicants care about when choosing a school in Turkey? To what extent are their preferences vertical or horizontal? Which school attributes matter the most? Do selective schools in Turkey contribute to their students' success, or is their performance due to the higher ability of the students they accept? Do they have different effects on different students that enroll? Though our data is on Turkey, these questions are of universal interest. Parents and students allocate considerable time, energy and money to try to get into elite schools. Hence, it is important to see if these schools actually add value or not. If some do, while others do not, should we not be reporting and rewarding those that add value, rather than those who look good on paper merely because of their excellent students?

We develop a way to answer these questions by taking a more structural approach than much of the literature. Using data on the high school entrance exam and the size of each high school class, we estimate a nested logit model of preferences over high schools and use it to model assignment to schools, taking into account that exam schools only admit the best students who apply. We do this in multiple steps. First, by using information on the cutoff score in each exam high school, we back out the mean valuation of each school in preferences, conditional on the correlation in shocks within a nest. Second, we pin down the correlation in shocks within a nest using information on the maximum scores in each school. This twist, to our knowledge, is new. The idea is quite simple. If shocks are perfectly correlated within a nest, then preferences are purely vertical and the minimum score in the most selective school in the nest cannot be lower than the maximum score in the second most selective school in the nest. Thus, the extent of overlap in the scores between schools within a nest, ranked by their selectivity, identifies the correlation in preference shocks in the nest.

[2]There has recently been considerable effort in determining value added by a school as part of the accountability in the no child left behind legislation. See Darling et al. (2012) for a critique of the approach usually taken.

Third, we regress our estimated mean valuation of a school on the schools characteristics, instrumenting to correct for endogeneity bias, to see what applicants care about in a school. We use the data on the overall distribution of scores in the high school entrance exam, along with the estimated preferences from step 1 and 2 above to allocate students to high schools and obtain the simulated distribution of students scores in the high school entrance exam in each school. Then, by using information on university entrance exam scores and the simulated high school entrance exam scores in a school, we estimate the value added by a school.

Our results suggest that students care about school's past performance in the university entrance exam and value elite science schools highly. However, value added differs a good deal across schools, and across quintiles of students. Our results are consistent with there being a consumption value of attending elite schools and with applicants finding it difficult to infer the value added by a school and using the type of a school and performance of its students in its place.

Turkey is a good place to look for answers to the questions of interest for a number of reasons. To begin with, the data are rich and available from public sources. It is usually hard to find data good enough to allow selection and value added to be effectively de-coupled. Privacy concerns are part of the reason why data is limited or accessible only to a privileged few.[3] Secondly, the Turkish system of admissions is exam driven. Admissions are rationed on the basis of performance in open competitive national central exams at the high school and university level. This provides a way to measure the performance of high schools: namely compare the distributions of students performance in the high school entering exam versus the university entrance exam. Third, as education is highly subsidized in public institutions, educational options outside the country or at private institutions are much more expensive so that these exams are taken seriously by the applicants.[4] When the stakes are high, as in

[3]For example, the College and Beyond data used in Dale and Krueger (2002) is not publicly available.

[4]Many experiments, especially non-natural ones, rely on performance measures or evaluations that do not matter for the student, which makes the effects hard to interpret.

Turkey, it is less likely that outcomes are driven just by noise.

We proceed as follows. In the next sub-section we relate our work to the literature. In Section 2 we provide the needed background regarding the Turkish system and the data. Section 3 deals with estimating preferences. It lays out the model, the estimation strategy, and the results in this regard. Section 4 deals with estimating value added by the school.[5] Section 5 concludes.

1.1 The Literature

There is a large literature that deals with school choice and school effects in the US, as well as in other developed and developing countries. In the US the consensus is that attending a better school does not have much of an impact on students' academic achievement. Abdulkadiroglu, Angrist, and Pathak (2011), and Dobbie and Fryer (2011b) investigate the effect of attending Boston and New York exam schools by using a Regression-Discontinuity approach. They look at students who were just below the cutoff and those that were just above and find no significant effect of being above the cutoff and thereby going to exam schools. However, it may be that selective schools are good for better students, but not marginal ones, and if this were the case, their approach would wrongly conclude that they have no effect.

Cullen, Jacob and Levitt (2006) and Cullen, Jacob and Levitt (2005) use data from randomized lotteries that determine the allocation of students in Chicago public schools system. Students who win the lottery attend the better schools. They find that winning this lottery does not improve students' academic performance. Clark (2010) investigates the effect of attending a selective high school (Grammar School) in UK (where selection is based on a test given at age 11 and primary school merit) and finds no significant effect on performance in courses taken by students, though the probability of attending university is positively affected.

[5] Here we are severely limited by the data, and consequently, we place far less weight on these results.

Dale and Krueger (2002) and (2011) look at the effect of attending elite colleges on labor market outcomes. Their work is among the most careful and well cited on the topic. Much of the work in this area controls for selection on observables by using a two step Heckman approach or matching estimators. Unobservables are typically controlled for by allowing the error terms in the selection and outcome variables to be correlated.[6] What is unique about their work is that they control for selection by controlling for the colleges to which the student *applied* and was *accepted*. The former provides an indication of how the student sees himself while the latter provides a way of controlling for how the colleges rank the student. Intuitively, the effect of selective schools on outcomes is identified by the performance of students who go to a less selective school despite being admitted to a more selective one, relative to those who go to the more selective one. Of course, if this choice is based on unobservables, this estimate would be biased.[7] They find a small but temporary positive effect of selective school attendance on average, though they find that students from disadvantaged backgrounds, less-educated or low-income families, black and Hispanic students, do seem to gain from attending elite colleges.

In contrast to these results, Pop-Eleches and Urquiola (2010) and Jackson (2010) estimate the effect of elite school attendance in Romania and Trinidad and Tobago, respectively. They find a large *positive* effect on students' exam performance for University admission.

From the school choice literature, Hastings, Kane, and Staiger (2006), and Burgess et. al (2009) investigate what parents care about in a school using data from Charlotte-Mecklenburg School District and Millennium Cohort Study (UK), respectively. Hastings, Kane, and Staiger (2006) take a relatively structural approach and estimate a mixed logit model of preferences. A major contribution of their work is to use information on the stated preferences for schools and compare these to what was available to them to back out the weight placed on factors like academics, distance from home and so on. They are then able

[6]See Frisancho and Krishna (2012) for an application using Indian data.

[7]For example, if confident students go to the selective school and less confident ones do not, and confident students do better, the effect of selective schools would be overestimated.

to see whether the impact of a school differs according to "type". For example, whether students who put a high value on academics do better in a good school than students who place a high value on being close to the school. If such differences are large, the reduced form effects estimated of attending good schools could be biased if such selection is not properly accounted for. Burgess et. al (2009) also compare the first choice school to the set what was available and infer that the stated first choice is preferred to this set. Their approach differs from that of Hastings et. al as it is reduced form and so does not impose any structure on preferences.

Although we don't estimate peer effect separately, our estimate of the school's value added includes peer effects. Ding and Lehrer (2007) estimate peer effects using data from a county in China, where students are allocated to high schools based on a criteria which is mainly based on students' entrance exam scores.[8] They find a positive peer effect on students' college entrance exam score. Several other papers [Hanusek et. al (2003), Hoxby (2000), Kang (2007), Zabel (2008), and Zimmerman (2003)] also study peer effects on academic achievements.[9] Duflo, Dupas and Kremer (2011) suggest that the behavior of teachers is crucial. They use data from a randomized experiment in Kenya to investigate how tracking students affects outcomes, and find that tracking helps both high achieving and low achieving students if teachers adjust their instruction level but not otherwise.

In sum, the evidence available suggests that selective schools/tracks can have a positive impact on disadvantaged groups who would otherwise have had a low quality education, or in developing countries with an overall low level of education. However, their average impact seems to be limited in developed countries.

Our work is more tangentially related to work in labor economics on affirmative action and the performance of minorities admitted under preferences to elite universities. Attempts to empirically evaluate the "mismatch hypothesis" in the U.S. provide mixed evidence. Rothstein and Yoon (2009) and Sander (2004) find evidence of mismatch in law school. Loury

[8]This differ slightly from Turkey system where allocation solely depends on exam score.
[9]Epple and Romano (2010) presents a detailed survey about peer effect.

and Garman (1993, 1995) find that blacks in the U.S. get considerable earning gains from attending more selective schools but these gains are offset for black students by lower performance both in terms of grades and probability of graduation. Frisancho Robles and Krishna (2012) look at affirmative action in India where admission cutoffs for backward castes can be much lower than those for the general category. They use data on performance in school and wages after school for a class in an elite engineering institution and show that there is strong evidence of mismatch and an absence of catchup.

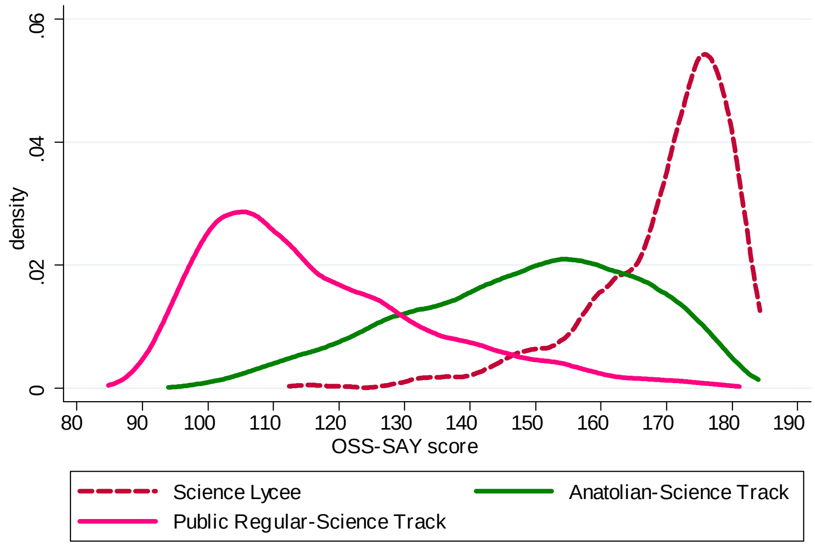
Our contribution to the literature is threefold. First, much of the work described above is reduced form rather than structural. An advantage of the slightly more structural approach taken here is that we can estimate preferences, understand what seems to drive them, look at sorting over schools, as well as try and estimate the value added created by a school. In other words, we examine the whole process and not just one of its components. Second, despite the lack of panel data, i.e., not having the high school entrance exam score and the college entrance exam score for each student, we show how one can use fairly limited data on each high school, along with data on university entrance exam takers, and the model, to get around this deficiency. Third, we allow students in different quintiles to benefit differently from each school's inputs. We find that schools have different effects on students according to their initial score quintile.[10] However, we would like to emphasize that the value added results are the weakest part of the paper as we do not have university entrance exam results for all students, only for a sample of them and we do not have a panel of students.

2 Background

In Turkey, competitive exams are everywhere. Unless the student chooses to go to a regular public high school, he must take a centralized exam at the end of 8th grade to get into "exam schools". These are analogous to magnets schools in the U.S., though the

[10] Such results are consistent with the work of Arcidiacono (2012) who in a study using data from Duke University shows that males and Asians lose ground during their time at Duke relative to their female and non-Asian counterparts.

Figure 1: Distribution of ÖSS-SAY score



Source: ÖSYM data on 2002 ÖSS applicants

competition for placement into them is national and widespread, rather than local as in the U.S. After high school, there is an open competitive university entrance exam every year and there are so many retakers that only a third of the 1.5 million students taking the exam in a given year do so for the first time. Most students go to cram schools (*dershanes*) to prepare for the university entrance exam. Much of the high school education also focuses on preparing students for this exam.

As admission into top schools is seen as essential to future success, private tutoring schools are a big part of the education sector in Turkey. In addition to the resources spent on them, dershanes also weaken the formal schooling system as schools also focus on studying for the exam rather than on fostering ability development and creative learning. If elite schools, in fact, have little value added, then the system itself may have adverse welfare effects. This is especially so if such schools are subsidized relative to the alternatives, as they often are.[11] In this case, students waste effort is trying to capture rents which constitutes wasted resources

[11]The best teachers are allocated to these schools, their facilities are better, and class size is smaller than that of regular schools. In addition, Caner and Ökten (2012) shows that school subsidies are regressive as they go to better off agents who tend to do better in exams and so go to better schools.

and a welfare loss.[12]

Students from elite schools do perform much better in university entrance exams. Figure 1 shows the distribution of average scores (ÖSS-SAY) in the university entrance exam of science track students coming from different kinds of high schools. Science high schools are clearly doing better, followed by the almost as selective Anatolian high schools, while regular Public schools seem to do the worst. However, this says little about the contribution of elite schools in terms of value added.

2.1 The Institutional Structure

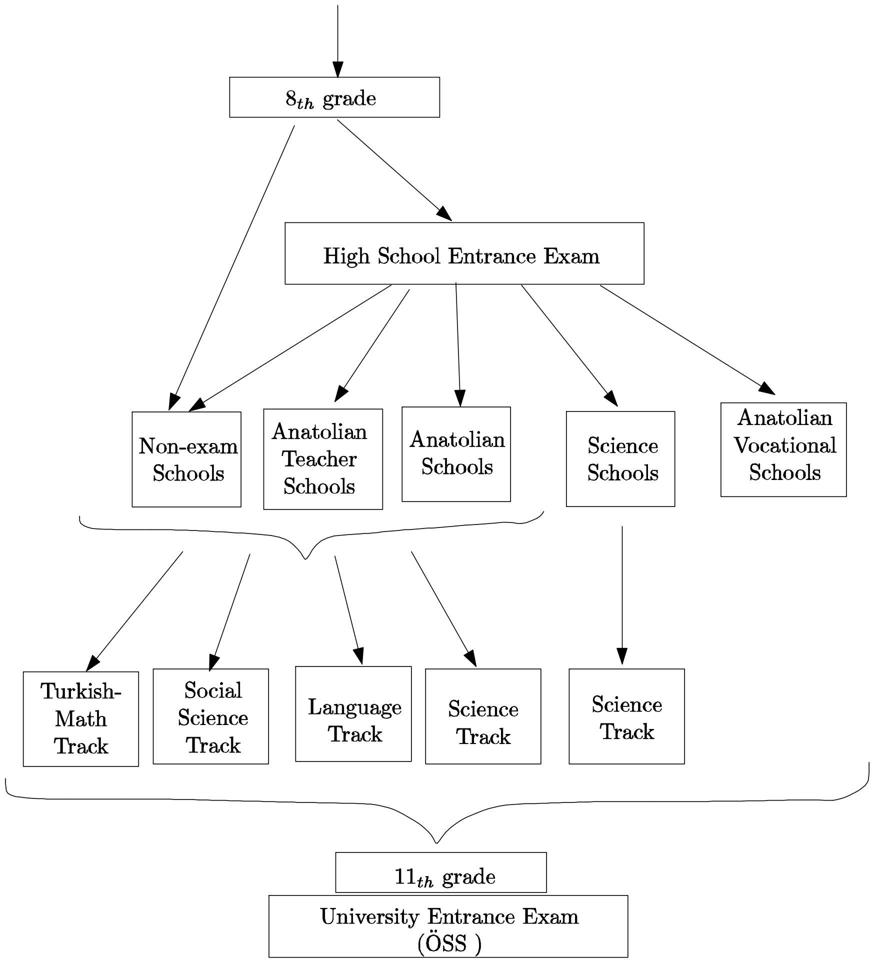
The educational system in Turkey is regulated by the Ministry of Education. All children between the ages of 6 and 14 must go to school. At 14, they take the high school entrance exam (OKS) if they want to be placed in public exam schools which are perceived as being better. Performance in this exam determines the feasible set of schools for a student and students choose their high school from this set. There are four types of public exam schools: Anatolian high schools, Anatolian Teacher Training high schools, Science high schools, and Anatolian Vocational high schools.

Anatolian high schools place a strong emphasis on foreign language education although their specific goals may vary across different types of Anatolian schools. The main goal of Anatolian high schools is to prepare their students for higher education, while teaching them a foreign language at a level that allows them to follow scientific and technological developments in the world. Anatolian Vocational high schools aim to equip their students with skills for certain professions and prepare them both for the labor market and/or higher education. Anatolian Teacher Training high schools train their students to become teachers though they can choose other paths without penalty as well.

The most prestigious of the exam schools are the Science High Schools. These were established to educate the future scientists of the Turkey and initially accepted very few

[12] See Krishna and Tarasov (2012) for more on this.

Figure 2: Education System in Turkey



students. Over time, the success of their students in the university entrance exams, as well as the rigorous education these schools were famous for, created considerable demand for these schools and they spread through the country.

In Public high schools, Anatolian high schools and Anatolian teacher training high schools, students can choose between four tracks: the Science track, the Turkish-Math track, the Social Science track and the Language track. In Science high schools they *must* take the science track. In Anatolian Vocational high schools there are no tracks which puts them a little outside the mainstream. In addition, these students are penalized if they choose to enter university outside the vocational track. All of this is depicted in Figure 2.

After 11^{th} grade, students who wish to pursue higher education take a centralized nationwide University entrance exam (ÖSS), which is conducted by the Student Selection and Placement Center (ÖSYM). This exam is highly competitive and placement of students into colleges is based on their ÖSS score and (adjusted for school) high school grade point average (GPA).

2.2 High School versus University Entrance Exam

In this paper we use high school and university entrance exam scores to estimate the value-added of schools. Thus, it is important to explain what these exams consist of and how similar they are. Both high school and university entrance exams are multiple choice tests that are held once a year. The high school entrance exam is taken by students at the end of eighth grade. There are four tests, Turkish, social science, math, and science, with 25 questions in each test. There is negative marking: for each wrong answer, students get $-\frac{1}{3}$ of a point. Students are given 120 minutes to solve the 100 questions. The University entrance exam is similar. It is a nationwide central exam with four parts, Turkish, social science, math, and science, with 45 questions in each part. Students are given 180 minutes, and there is negative marking: a wrong answer gets $-\frac{1}{4}$ of a point. The questions in both exams are based on the curriculum taught in school and are meant to measure the ability to use the concepts taught in school. Students' scores in the high school entrance exam are calculated as the simple mean score in each test. In the university entrance exam, the weights given to the different tests change according to students' track in high school.

2.3 The Data

The data we use comes from several public sources. To measure students' academic performance at the end of high school, we rely on school level descriptive statistics of scores in the 2002 University entrance exam. This information is published by the Student Selection and Placement Center (ÖSYM) and made available to schools and families so that they are informed about the standing of each school. The data set includes number of students who took University entrance exam from each school, as well as the mean and standard deviation of their scores in each field of the exam. In addition, we were able to get the scores of all students in three science high schools.

The students' initial ability prior to attending high school is assumed to be his performance in the exam high school entrance exam. We obtained data on 2001 cutoff scores[13] for each exam high school from the Ministry of Education's website.[14] We also collected data on the average ÖSS performance of each high school in each part of the exam in 2000 from ÖSYM's Results booklet for that year, which is publicly available from their website. Additional high school characteristics were collected from the Ministry of Education's website (language education, dormitory availability, and location) and the high school's websites (age of the school). We use this to better understand what drives the mean valuation placed on a high school. The allocation process, seats available and preference structure described in the next section are used to back out this mean valuation placed on each high school.

3 The Model

The allocation process used in Turkey is as follows. Seats are allocated according to students' preferences and their performance. The performance measure used is the marks obtained in a centralized exam (conducted once a year). All schools prefer higher scoring students. For each student i, we define the pair (s_i, P_i) where s_i is the score and P_i is the $1 \times m$ vector of schools ordered from the most preferred to the least preferred option.

Each exam school has a fixed quota, q_j, which is exogenously determined. The allocation process basically assigns students to schools according to their stated preferences, with higher scoring students placed before lower scoring ones. Students know past cutoffs for schools

[13] This data was collected using the website http://archive.org/web/web.php, which provides previous versions of the websites.

[14] Ideally, we should look at the high school entrance exam cutoff scores in 1998, which is the year in which most of the students in our ÖSS data got into high school. Unfortunately, this data is not available and the closest year to 1998 for which the data is available is 2001. However, the cutoffs scores are fairly stable as the educational environment in Turkey has been unchanged in the last few decades. In the Appendix we present evidence on the stability of cutoff scores.

when they put down their preferences. They are allowed to put down up to 12 schools though most do not use all twelve choices allowed.

Students do face a location restriction in putting down their Anatolian high school preferences. They are not allowed to put preferences on Anatolian high schools in Ankara, İstanbul, İzmir, and their current city: they have to pick one of these locations and make all their Anatolian high school preferences from their chosen location. We do not incorporate this restriction in our setup. However, we nest the Anatolian Schools so that errors are allowed to be correlated within a city nest.

We model preferences as follows. Suppose that student i's utility from attending school j takes the form

$$U_{ij}(X_j, \xi_j, \varepsilon_{ij}; \beta) = \beta X_j + \xi_j + \varepsilon_{ij}$$

where X_j are the observed school characteristics, ξ_j are the unobserved school characteristics, and ε_{ij} is a random variable which has a Generalized Extreme Value (GEV) distribution. Let δ_j denote school's specific mean valuation where

$$\delta_j = \beta X_j + \xi_j$$

so that

$$U_{ij}(X_j, \xi_j, \varepsilon_{ij}, \beta) = \delta_j + \varepsilon_{ij}$$

This structure of the utility function implies that variation across individual preferences comes from the error term, conditional on the students having the same feasible choice set. If two alternatives j and j' are in the same nest, ε_{ij} and $\varepsilon_{ij'}$ can be correlated. Otherwise, the errors are assumed to be independent.

In general, the cumulative distribution function of $\varepsilon = \langle \varepsilon_{i0}, \varepsilon_{i1}, \ldots, \varepsilon_{iN} \rangle$ is given by

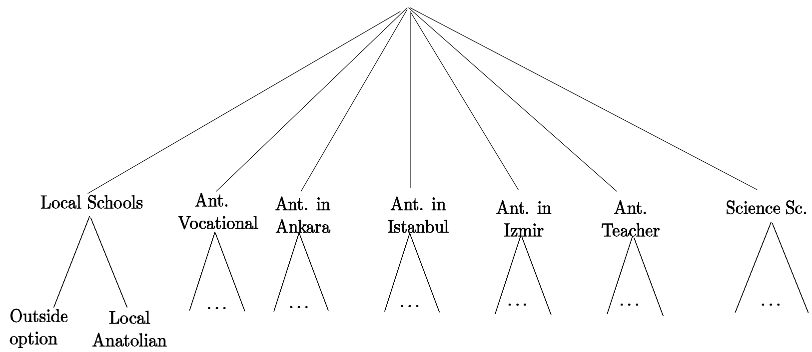
$$H(\varepsilon_{i0}, \varepsilon_{i1}, \ldots, \varepsilon_{iN}) = \exp\left(-\sum_{k=1}^{K}\left(\sum_{j \in B_k} \exp(-\frac{\varepsilon_{ij}}{\lambda_k})\right)^{\lambda_k}\right) \tag{1}$$

where B_k is the set of the alternatives with similar characteristics, K is the number of nests, and λ_k is the measure of degree of independence among the alternatives within nest k (see Train (2009)). As λ_k increases, the correlation between alternatives in nest k decreases. If λ_k is equal to 1, there is no correlation between alternatives within nest k, whereas if λ_k goes to 0, there is perfect correlation among all alternatives in the same nest. In this case, the choice of alternatives for any individual is driven by the δ_j component alone so that there is pure vertical differentiation among schools in a nest.

Given the location restriction for Anatolian high schools discussed above, we partition the set of high schools in Turkey according to their type and location. Figure 2 shows the nesting structure we adopt. At the upper level of the nest, students have seven options: Science high schools, Anatolian Teacher training high schools, Anatolian high schools in Ankara, in İstanbul and in İzmir, Anatolian Vocational high schools, and the local school option. The local school option includes public regular high schools which are modeled as the outside option, and local Anatolian high schools. Since computational intensity will increase with the size of the choice set, we aggregate Anatolian Vocational high schools into five subgroups according to their types. Other nests include all schools in Turkey of that type: 91 Teacher Training high schools, 48 Science high schools, 24 Anatolian high schools in Ankara, 38 Anatolian high schools in İstanbul, 18 Anatolian high schools in İzmir.[15] Thus, overall we have 219+7=226 options. It is worth emphasizing that putting Anatolian High Schools in each city in a separate nest is not the same as restricting students to choosing Anatolian Schools from only one city. The nesting structure we choose merely allows shocks for schools within a city to be correlated. If the correlation is high and a good shock is drawn,

[15] These schools are located in the center of the Ankara, İstanbul, and İzmir. Anatolian high schools located in a town in the provinces are defined as local Anatolian high schools by Ministry of Education.

Figure 3: School Choice in Turkey



then only schools from that city nest will be listed as they will all have high valuations. In this way, we incorporate the restriction that only Anatolian School from one city can be listed.

Each student chooses a school that maximizes his utility given his feasible set of schools, which is determined by her own score, s_i and the cutoff scores of each school

$$\max_{j\in\mathcal{F}_i} U_{ij}(X_j, \xi_j, \varepsilon_{ij}; \beta)$$

where

$$\mathcal{F}_i = \{j: \; c_j \leq s_i\}$$

The feasible set of a student, $\mathcal{F}_i$, includes all the schools whose cutoff score is below the student's score. Given the demand for each school and the number of seats available, the cutoff score, c_j, is endogenously determined in equilibrium.

Let the set of N schools be partitioned into K mutually exclusive sets (nests) where the elements of each of these sets correspond to schools within that nest. For example, B_k, where $k = 1, 2, \ldots, K$, would have as its elements all schools that are in nest k. If there were no rationing, the probability that school j in nest k was chosen by student i would be given by[16]

$$P_{ij}(\boldsymbol{\delta}, \boldsymbol{\lambda}) = \frac{\exp(\frac{\delta_j}{\lambda_k}) \left(\sum\limits_{l \in B_k} \exp(\frac{\delta_l}{\lambda_k}) \right)^{\lambda_k - 1}}{\sum\limits_{n=1}^{K} \left(\sum\limits_{l \in B_n} \exp(\frac{\delta_l}{\lambda_n}) \right)^{\lambda_n}}$$

which would be equivalent to the fraction of students whose best alternative was alternative j.

However, students' choices are constrained by the cutoff scores in each school, c_j, and by their own exam performance, s_i. Suppose that there are $N+1$ choices (including the outside option) and let the cutoff scores for each alternative be ordered in ascending order

$$c_0 = 0 < c_1 < c_2 < \ldots < c_{N-1} < c_N$$

where 0 indexes the outside option. Students whose score is in the interval $[c_m, c_{m+1})$ have the first m schools in their feasible choice set and we call this interval I_m. Similarly, students whose scores are below c_1 have scores in interval I_0 and have their choice set containing only the outside option, while students with $s_i \geq c_N$ get to choose from all the $N+1$ alternatives and have scores in interval I_N. Thus, the probability that student i with a score in interval I_j chooses school t, $t < j$, in nest k from his feasible set will be

$$P_{jt(k)}(\boldsymbol{\delta}, \boldsymbol{\lambda}) = \left\{ \frac{\exp(\frac{\delta_t}{\lambda_k}) \left(\sum\limits_{l \in B_k(I_j)} \exp(\frac{\delta_l}{\lambda_k}) \right)^{\lambda_k - 1}}{\sum\limits_{n=1}^{K_j} \left(\sum\limits_{l \in B_n(I_j)} \exp(\frac{\delta_l}{\lambda_n}) \right)^{\lambda_n}} \quad \text{if } s_i \in I_j \right\}$$

where bold variables denote vectors and where $B_k(I_j)$ denotes the restriction placed on the elements of nest k when the individuals' score is in the interval I_j. λ_k is the extent of independence between alternatives in nest k, and K_j is the total number of nests available to a student whose score is in interval I_j.

Aggregate demand for each school will thus depend on the distribution of scores, $F(s)$,

[16] The derivation of the nested logit probability, P_{ij}, can be found in the Appendix A.1

the minimum entry cutoff scores of all other schools whose cutoff score is higher, and the observed and unobserved characteristics of all schools. Using the equilibrium cutoff scores and students' score distribution we can get the density of students that are eligible for admission in each school.

For simplicity, we will not write the arguments of the aggregate demand function for each school j, $d_j(\boldsymbol{\delta}, \boldsymbol{\lambda})$. The demand for school N, the best school, which is in nest k comes only from those in I_N

$$d_N = P_{NN(k)}(\boldsymbol{\delta}, \boldsymbol{\lambda})[1 - F(c_N)]$$

Only students with scores above c_N have the option to be in school N which gives the term $[1 - F(c_N)]$. In addition, N in nest k has to be their most preferred school; hence the term $P_{NN(k)}(\boldsymbol{\delta}, \boldsymbol{\lambda})$. Similarly, the demand for school j which is in nest s comes from those in I_j,... I_N

$$\begin{aligned} d_j &= P_{Nj(s)}(\boldsymbol{\delta}, \boldsymbol{\lambda})[1 - F(c_N)] \\ &\quad + P_{(N-1)j(s)}(\boldsymbol{\delta}, \boldsymbol{\lambda})[F(c_N) - F(c_{N-1})] \\ &\quad +.. + P_{(j)j(s)}(\boldsymbol{\delta}, \boldsymbol{\lambda})[F(c_{j+1}) - F(c_j)] \\ &= \sum_{w=j}^{N-1} P_{wj(s)}[F(c_{w+1}) - F(c_w)] + P_{Nj(s)}(\boldsymbol{\delta}, \boldsymbol{\lambda})[1 - F(c_N)] \end{aligned}$$

For students above c_j, school j must be the highest ranked by them to be demanded.

3.1 Estimation Strategy and Results

Given the preference parameters and the number of seats in each school, the real world cutoffs are determined by setting the demand for seats, as explained above, equal to their supply and obtaining the market clearing score cutoffs. This is not what we will do. For us, the cutoffs and the number of seats are data. We want to use this data and the nesting structure imposed to obtain the preference parameters. In particular, we want to estimate the coefficients of school characteristics (β) and the parameter vector $\boldsymbol{\lambda}$,where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_K]$, that best fit the data and respect the solution of the model that equates demand (d) with supply (q).

We do this in three steps. In Step 1, we back out the values of δ_j for each school j for a given $\boldsymbol{\lambda}$. In essence, the minimum cutoff in each school denoted by the vector $\mathbf{c} = (c_1, ..c_N)$, the number of seats in each school denoted by the vector $\mathbf{q} = (q_1, ..q_N)$, together with the market clearing conditions of the model, pin down the mean valuation of each school for a given vector, $\boldsymbol{\lambda} = (\lambda_1, ...\lambda_K)$.[17] In step 2, we find $\boldsymbol{\lambda}$ so as to best match the extent of overlap in the scores in schools in the same nest. A higher correlation in the errors within a nest means that there is less of a role for preference shocks to play in choice, so that preferences are driven by the non random terms. This corresponds to having more of a vertical preference structure. As a result there is less overlap in the range of student scores across schools in a nest. If there is perfect correlation, the maximum score in worse school will be less than or equal to the minimum score in the better one. In step 3, we relate our estimates of δ_j to the characteristics of each school to see what drives the preferences for schools.

3.2 Step 1

Our model includes unobserved school characteristics, and these unobserved characteristics enter the demand function nonlinearly, which complicates the estimation process. Berry (1994) proposed a method to transform the demand functions so that unobserved school characteristics appear as school fixed effects. By normalizing the value of the outside option to zero, $\delta_0 = 0$, we have N demand equations with N unknowns. This permits us to get the vector $\boldsymbol{\delta}(\mathbf{q}, \mathbf{c}, \boldsymbol{\lambda})$ for given vector $\mathbf{q}$, conditional on a vector $\boldsymbol{\lambda}$ such that

[17] In spirit, this is like the Hotz Miller inversion commonly used in Industrial Organization Models.

$$\mathbf{q} = \mathbf{d}(\boldsymbol{\delta}(\mathbf{q}, \mathbf{c}, \boldsymbol{\lambda}), \boldsymbol{\lambda})$$

Our setup is more complex than the models presented in Berry (1994) so we cannot solve for $\boldsymbol{\delta}(\mathbf{q}, \mathbf{c}, \boldsymbol{\lambda})$ analytically. Our setup is closer to that in Bresnahan, Stern and Trajtenberg (1997) who solve the system numerically as we do. Inverting the demand function numerically gives us the vector of mean valuation of the alternatives, $\boldsymbol{\delta}(\mathbf{q}, \mathbf{c}, \boldsymbol{\lambda})$.

3.3 Step 2

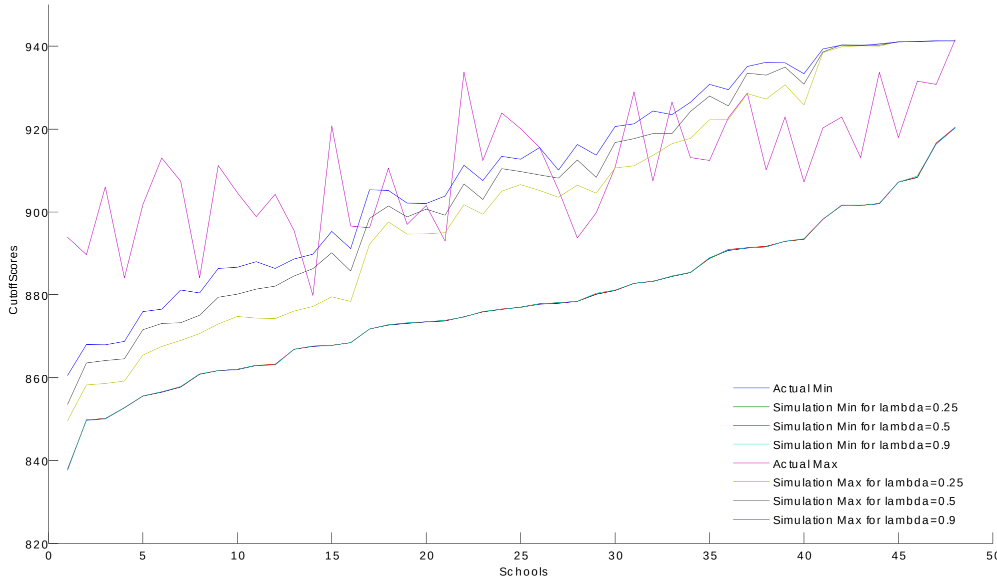
Once we get $\boldsymbol{\delta}(\mathbf{q}, \mathbf{c}, \boldsymbol{\lambda})$, we can specify individual i's utility from alternative j as

$$U_{ij}(\boldsymbol{\lambda}, \mathbf{q}, \varepsilon_{ij}) = \delta_j(\mathbf{q}, \mathbf{c}, \boldsymbol{\lambda}) + \varepsilon_{ij}$$

At this stage, the only unknown in the utility function is the λ. As λ decreases in a nest, the correlation of the value of alternatives within the nest will increase. In the extreme case, when the correlation is perfect, if one agent values a particular school highly so do all other agents which can be interpreted as pure vertical differentiation. In this case, there will be no overlap in the score distributions of different schools *within* the nest. If correlation is low, then some students will choose one school and others will choose another and there will be overlap in the score distributions. The extent of overlap in the minimum and maximum scores within a nest helps to pin down the λ in the nest.

Figure 4 shows how different values of the λ affect the fit of the model to the data for the Science High school nest that we will focus on. For each λ, the simulated minimum scores must lie exactly on top of the actual minimum scores (which is in turquoise blue) as a result of our estimation strategy. The pink line gives the actual maximum scores. The yellow line gives the simulated maximum for $\lambda = .25$, the brown and blue lines depict the simulated maximums for $\lambda = .5$ and $.9$ respectively. Note how the lines move up as λ rises (or correlation falls) so that the extent of overlap increases.

Figure 4: Real and Simulated Cutoff Scores for $\lambda = 0.25, 0.5, 0.9$



We pin down $\boldsymbol{\lambda}$ using a simulation based approach. In 2001, more than 500,000 students took the high school entrance exam. Roughly 100,000 of them were allocated to the 1,745 public exam schools. We ignore private exam high schools as they are less than 5% of the total and there is no data on them. We aggregate the roughly 1,000 vocational high schools into 5 groups according to their types and define the maximum and minimum score of each group as the maximum or minimum of the cutoff scores of the schools in that group, and available number of seats as the total number of seats available in that group. Similarly we aggregate Local Anatolian high schools into one group and we have the outside option which gives 7 choices. We allow each of the 219 Science, Anatolian Teacher and National Anatolian high schools into students' choice set. For computational reasons, we choose to focus on the 50,000 highest ranked students. The score of the student whose rank is 50,000 is lower than the minimum cutoff score of 216 of the 219 schools and lower than the maximum cutoff score of all the 219 schools.

The simulation algorithm works as follows: For a given λ, we obtain $\delta(q, \lambda)$ and simulate the minimum and maximum cutoff scores, $\underline{c}_j$, and $\bar{c}_j$ for each school. Then we find the $\boldsymbol{\lambda}$ that best matches the actual maximum cutoff score. Recall that because of our approach in the first step of the estimation, we can match minimum cutoff scores perfectly.

Simulating the error terms in the nested logit model creates some difficulties: taking a draw from the GEV distribution with the standard Markov Chain Monte Carlo Method is computationally intensive. We use a method proposed by Cameron and Kim (2001) which takes a draw from GEV using a far less computationally intensive procedure.[18]

We draw M sets of error terms ε_{ij} from the distribution function given in equation 1 by using the scale parameters, $\boldsymbol{\lambda}$. For each of the M set of errors drawn, $\varepsilon_k = \langle \varepsilon_{ij}^k \rangle,\ k = 1, ..M$, we allocate students to schools by using the placement rule. After each set of errors drawn we get a distribution of scores for students in each school. Let g_j^k be the set of scores in school j in simulation k, ordered to be increasing.[19]

$$g_j^k(\boldsymbol{\lambda}) = \langle s_{j1}^k(\boldsymbol{\lambda}), s_{j2}^k(\boldsymbol{\lambda}), \ldots, s_{jq_j}^k(\boldsymbol{\lambda}) \rangle$$

After ordering scores in ascending order for each school j and simulation M, we find the expected value of the score for each rank within each school across the M simulations. The expected score of student with rank r in school j is thus:

$$s_{jr}^*(\boldsymbol{\lambda}) = \frac{1}{M} \sum_{k=1}^{M} s_{jr}^k(\boldsymbol{\lambda})$$

Let $g_j^*(\boldsymbol{\lambda})$ be

$$g_j^*(\boldsymbol{\lambda}) = \langle s_{j1}^*(\boldsymbol{\lambda}), s_{j2}^*(\boldsymbol{\lambda}), \ldots, s_{jq_j}^*(\boldsymbol{\lambda}) \rangle$$

and we match this to the data to pin down $\boldsymbol{\lambda}$.

Given this score distribution in each school, we find the $\boldsymbol{\lambda}$ that gives the least square distance between simulated minimum and maximum cutoff scores and observed minimum

[18] This method is explained in Appendix A.2.

[19] In the method proposed by Cameron and Kim (2001), change in λ only affects coefficients, this allows us to keep random seeds drawn from Extreme value distribution over simulations, and only change coefficients.

Table 1: Nesting Parameters: λ

Variable	Coefficient
λ_{loc}	0.958
λ_{voc}	0.986
λ_{ank}	0.795
λ_{ist}	0.837
λ_{izm}	0.777
λ_{teach}	0.999
λ_{sci}	0.897

and maximum cutoff scores. In effect, we are matching the maximum scores as the minimum scores are matched on average given our estimation procedure for obtaining δ.

$$\hat{\boldsymbol{\lambda}} = \arg\min_{\lambda} \frac{1}{N}\sum_j (s^*_{jq_j}(\boldsymbol{\lambda}) - \bar{c}_j)^2 + \frac{1}{N}\sum_j (s^*_{j1}(\boldsymbol{\lambda}) - \underline{c}_j)^2$$

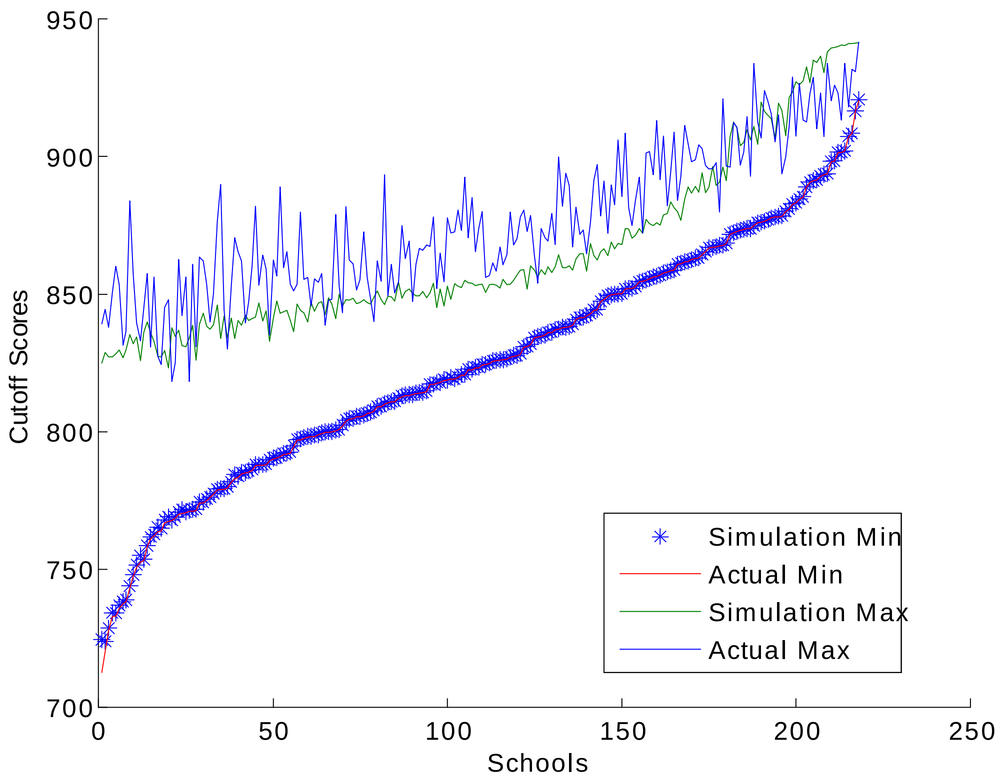
Table 1 shows the $\boldsymbol{\lambda}$ values for each nest that minimize the distance between simulated and real maximum and minimum cutoff scores. As we mentioned before, $\boldsymbol{\lambda}$ is a measure of dissimilarity within a nest. If $\boldsymbol{\lambda}$ is small, student's rank schools in the same nest according to their perceived quality ($\boldsymbol{\delta}$) so that students tend to agree on the ranking of schools. However as $\boldsymbol{\lambda}$ gets bigger, students differ in their preferences and no such ranking exists as their tastes for schools differ.

As expected, the restriction that students choose Anatolian Schools in only one city is reflected in the estimated correlation in the students' taste for schools. This correlation is highest in the Ankara, Izmir, and Istanbul Anatolian high schools nests (as λ is lowest). As expected, science high schools are also more vertically differentiated than local schools, vocational schools and teachers schools which are more horizontally differentiated.

The real and simulated cutoff scores for $\boldsymbol{\lambda}$ presented in Table 1 are given in Figure 5. As it is seen simulated maximum scores match the real maximum cutoffs quite well[20].

[20] Additional figures that show the fit of the Model in each nest can be found in the Appendix A.5.

Figure 5: Real and Simulated Cutoff Scores



3.4 Step 3

Once we pin down the $\boldsymbol{\lambda}$ that gives the best match of the actual and the simulated cutoffs, we get $\hat{\boldsymbol{\delta}}(\mathbf{q}, \hat{\boldsymbol{\lambda}})$. Returning to the definition of $\boldsymbol{\delta}$, we write $\boldsymbol{\delta}$, the vector of mean preferences for schools, as

$$\hat{\boldsymbol{\delta}} = \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\xi}$$

where X is the observed characteristics of the school, and ξ is the school specific component of mean preferences, which is unobserved, or the preference shock. X includes the school's success in the ÖSS in previous year, age/experience, type, education language, dormitory availability, whether it is located a big city (Ankara, İstanbul, or İzmir), the number of seats, and the cutoff score of the school. Including the cutoff score as an explanatory variable for mean valuation allows for the possibility of circular causation discussed above. If there is value placed on being selective per se, i.e., on a lower cutoff score, then it could be that students like selective schools and so there is demand for them, which results in a high cutoff score. The dummy for being a Science or Anatolian High School incorporates the possibility that such schools have a good reputation and this makes people value them. This need not be for what they add in value: it could be for consumption purposes, perhaps for the bragging rights associated with going there.

There is an econometric issue associated with including the cutoff score as an explanatory variable. If ξ, the school specific shock is large and positive, then the cutoff score will be high as well, so that the cutoff will be correlated with ξ. For this reason, adding the cutoff score of a school to the regression above will not give unbiased estimates of the effect of the school characteristics on preferences directly. This is the familiar endogeneity problem.

We can partition X as $[\tilde{X}, \underline{c}]$

$$\hat{\boldsymbol{\delta}} = \tilde{\boldsymbol{\beta}}\tilde{X} + \gamma\underline{c} + \boldsymbol{\xi}$$

To deal with this we need a good instrument for the cutoff score. We need an exogenous variable that shifts cutoff score, but does not effect schools' average valuation δ directly. The first variable that comes to mind that shifts minimum cutoff score is the available number of seats in a school. However, the available number of seats may affect valuation of school directly. In addition, it may be a response to a high $\boldsymbol{\xi}$ which makes it less than optimal.

Fortunately, the model suggests a really lovely valid instrument for the minimum cutoff score. The model predicts that the number of available seats in worse schools than itself has no effect on the demand for a school. However, the number of seats in better schools does affect the demand of a school: more seats in better schools reduces the cutoff score of a school. This result comes from the observation that the demand for a school comes from those who like it the most among the alternatives that are open to them. Changing the cutoff in worse schools has no effect on the alternatives open to a student going to a better school and hence on their demand. In other words, if Podunk University offers more seats, there is no effect on the demand for Harvard since everyone choosing to go to Harvard had, and continues to have Podunk in their choice set. But if Harvard offers more seats, it may well reduce the demand for Podunk University. It could be that someone chose Podunk because they could not get into Harvard. Once Harvard increases its seats and so reduces its own cutoff, Harvard may become feasible for such a student.

To construct the instrumental variable, we need a ranking of schools free of ξ. We will use the schools' success in the verbal and quantitative part of the ÖSS to rank schools.

We construct our instrumental variable as follows:

1. For each school, we find the schools that have better average test scores in both dimensions, verbal and quantitative.
2. We find the total number of seats in all of the schools found in step 1. The available number of seats in the school itself is *not* included.

As we use seats in other schools to instrument for own cutoff we need not worry about any correlation with ξ. Table 2 shows our first stage estimation:

$$\underline{c} = \eta\tilde{X} + \kappa * \text{Seats in better schools} + \varepsilon$$

Note that the cutoff score is higher in big cities, science and teacher high schools which is consistent with their being value placed on these attributed. Also, the number of seats in better schools, the instrumental variable we constructed, has a negative coefficient: more seats in better schools reduces the own cutoff as expected. We also validate the instrument used. According to the model, the number of seats in worse schools should have no effect on own cutoff. Table 9 presented in Appendix A.4 shows the first stage estimation with two instrumental variables, one constructed using higher scoring schools, and the other using lower scoring schools. Only the higher scoring schools instrument is significant.

The first column of Table 3 shows our baseline estimates, where we regress average valuation on the exogenous variables, and do not include the cutoff score of a school. This column suggests that past performance in the University entrance exam (ÖSS scores) and school type drive preferences. The second column of Table 3 shows the results of the regression of average valuation on the exogenous variables and the schools cutoff score. The coefficient on the minimum score is positive and highly significant suggesting that a more selective school is more highly valued. The significance of past scores in the University entrance exam are less significant as would be expected given that the cutoff is positively correlated with the past performance of a school so that including it picks up some of this variation. However, as explained above, cutoffs are not exogenous. As cutoffs are high when the school preference shock, cutoffs are positively correlated with the error term which imparts an upward bias to the coefficient. The third column shows the results when we instrument for cutoffs using seats in better schools as the instrument. As expected, the coefficient on the cutoff score falls and is now no longer significant. This suggests that students *do not* simply look at the selectivity of a school and blindly put greater value on more selective ones. Past performance

Table 2: First Stage Estimation

Variable	Coefficients
Number of Available Seats	-0.0281
	(0.0664)
Average Quantitative Score in 2000 ÖSS	1.158
	(0.685)
Average Verbal Score in 2000 ÖSS	1.087
	(0.806)
Age	-0.0926
	(0.444)
Science High School	67.53***
	(17.55)
Teacher High School	48.46*
	(20.4)
Anatolian High School in Istanbul	33.37
	(18.7)
Anatolian High School in Izmir	17.77
	(22.1)
Education Language- English	31.57
	(17.8)
Education Language- German	13.18
	(17.05)
Dormitory Availability	10.58
	(7.085)
Ankara	48.37**
	(15.41)
Istanbul	35.64***
	(8.921)
Izmir	40.52**
	(13.97)
Seats in better schools	-0.00438*
	(0.00218)
Constant	701.9***
	(41.85)

Note: Standard errors are reported in paranthesis. *, **, *** indicate significance at the .90, .95 and .99, respectively.

Table 3: School Choice: Estimation Results

Variable	(I)	(II)	(III)
Number of Available Seats	0.005	0.00842*	0.006
	(0.007)	(0.004)	(0.005)
Average Quantitative Score in 2000 ÖSS	0.218***	0.0680*	0.172*
	(0.053)	(0.033)	(0.067)
Average Verbal Score in 2000 ÖSS	0.306***	0.0865*	0.239**
	(0.053)	(0.038)	(0.093)
Age	0.026	0.032	0.028
	(0.062)	(0.031)	(0.049)
Science High School	8.422***	3.237***	6.844**
	(1.764)	(0.765)	(2.351)
Teacher High School	4.039*	0.867	3.074
	(1.931)	(0.763)	(1.867)
Anatolian High School in Istanbul	1.928	-0.544	1.176
	(2.152)	(0.728)	(1.835)
Anatolian High School in Izmir	1.715	0.392	1.312
	(2.280)	(0.679)	(1.708)
Education Language- English	2.671	-0.028	1.850
	(2.192)	(1.118)	(2.131)
Education Language- German	1.198	0.330	0.934
	(2.254)	(1.175)	(1.862)
Dormitory Availability	1.617	0.500	1.277
	(0.893)	(0.455)	(0.821)
Ankara	4.235**	0.733	3.169
	(1.494)	(0.485)	(1.736)
Istanbul	4.485***	1.687**	3.634*
	(1.297)	(0.522)	(1.480)
Izmir	3.746*	0.608	2.791
	(1.449)	(0.381)	(1.552)
Minimum Cutoff Score		0.0846***	0.026
		(0.006)	(0.032)
Constant	-22.95***	-76.18***	-39.15*
	(3.121)	(4.363)	(19.700)

Note: Standard errors are reported in paranthesis. *, **, *** indicate significance at the .90, .95 and .99, respectively.

in the University entrance exam becomes more significant suggesting that, conditional on the cutoff, a schools' performance in the university entrance exam is an important determinant of a school's valuation. Thus, students *do* look at the output of a school in forming their valuation of a school. Science High Schools and Anatolian Schools in Istanbul are also valued beyond what they would be based on their selectivity alone.

These results are consistent with the findings of Burgess et. al (2009), and Hastings, Kane, and Staiger (2006) that reach same result by using data from Millennium Cohort Study in UK, and school choice data from the Charlotte-Mecklenburg School District, respectively.

Our results suggest that being in a science high school is greatly valued by students as is being in a big city. As mentioned before, Science High Schools are the most prestigious high schools with a curriculum concentrated on science and mathematics. It could be that attending such schools gives one contacts in the future as well as consumption value in the present coming from the prestige associated with attending such schools.

Macleod and Urquiola (2013) show that a school's reputation can affect wages as the identity of the school attended gives information about a student's ability. This could also rationalize the high valuation placed on Science High Schools. It could also be that the high valuation of Science High Schools comes from the students use of school type as a proxy for school quality. In the next section we ask if this is valid. We look at the value added of each school by estimating the *effect* of the high school on the students' performance in the university entrance exam.

4 High School's Value Added

Using the estimates we have so far, we can recover the distribution of high school entrance exam scores for each school. This gives us the initial "ability" distribution in each high school. Our goal in this section is to estimate the school effect on student's academic performance. Here we are limited by the data. We do not have a panel, so we cannot match the score the student obtained in the High School Entrance Exam to what he obtained in the University Entrance Exam. Rather, we infer the effects of schools on student performance by comparing High School Entrance Exam score distributions to the University Entrance Exam score distribution for a given school. We have the simulated distributions for the scores in the High School Entrance Exam based on our estimated preference parameters and the model of allocation of students to schools. We only have data on the population mean and standard deviation of University Entrance Exam scores in each school. We use this information to fit a parametric (beta) distribution of scores for each school. For three schools, we actually have the actual distribution of scores in the University entrance exam[21]. We use the data on these three schools for a more detailed analysis of the school effect and as a validation of our simulation intensive approach.

4.1 The Approach

We have the distribution of High school entrance exam scores based on our estimates and the mean, and standard deviation of the University entrance exam scores, as well as the number of students in each high school. By making a parametric assumption on the distribution of University Entrance Exam scores in each school, we estimate parameters of the distribution by matching mean, and standard deviation.

We assume that University entrance exam scores are distributed according to a beta distribution with supports $[79.042, 184.993]$, where 79.042 is the minimum score and 184.993 is the maximum score in 2002 University entrance exam. Scores in school j have a beta distribution with parameters α_j and β_j.

$$s_j^c \sim Beta(\alpha_j, \beta_j, 79.042, 184.993)$$

Once we estimate the parameters of the distribution, we take S random samples from

[21] The data was obtained from the web sites of these school.

this distribution of the same size as the number of students who took the ÖSS in that school. We find the average simulated score according to rank of the scores over all simulated samples. Namely, let $s_j^{c^r} = [s_{1j}^r, s_{1j}^r, \ldots, s_{n_j j}^r]$ be the rth simulated sample from the school j's distribution, where s_{1j}^r is the lowest score, and $s_{n_j j}^r$ highest score and n_j is the number of students in this school. Then the scores in school j are given to be

$$s_j^c = [\frac{1}{S}\sum_{r=1}^{S} s_{1j}^r, \frac{1}{S}\sum_{r=1}^{S} s_{2j}^r, \ldots, \frac{1}{S}\sum_{r=1}^{S} s_{n_j j}^r]$$

We can compare score distributions in the High school and University entrance exam. As we do not have a panel we cannot really say how a school affects a student. However, we can parameterize and measure the change in the distribution of scores in the high school entrance exam and the distribution of scores in the university entrance exam and interpret this distance as the school effect.

The first thing to note is that difference in the mean score in the High School and University entrance exams varies across schools suggesting that not all schools add the same value. However, there is a lot of variance in the scores so that it is hard to say that they differ significantly.

We incorporate the school effect in a number of ways. In what follows, we present three models that differ in the assumptions made about the influence of the school on academic performance. The first model allows the school to shift the scores of each student by a constant and allows for a rescaling. The second allows both these parameters to vary above and below the mean. The third allows both these parameters to differ by quintile.

Let G_j denote the empirical cumulative distribution function of ÖSS scores for high school j. Let s_i^{hs} denote student i's score in the High school entrance exam. Define parameter vector θ_j to capture school j's effect on performance so that University entrance exam scores, s_i^c, are determined in the following way:

$$s_i^c = h(s_i^{\mathrm{hs}}, \theta_j) + \varepsilon_{ij}$$

where $\varepsilon_{ij} \sim N(0, \sigma_j^2)$

For different functions h, we estimate θ_j to minimize the distance between the empirical distribution G_j and the empirical distribution that function h generates for s_i^c in school j, $\hat{G}_j$. To implement our minimum distance estimator, we are going to rely on the Cramér-von Mises criterion as a measure of distance $d[.,.]$:

$$d[G_j, \hat{G}_j] = \int_{-\infty}^{+\infty} [G_j(s) - \hat{G}_j(s)]^2 dG_j(s)$$

1. Model 1

$$s_i^c = \alpha_j + \beta_j s_i^{\text{hs}} + \varepsilon_{ij}$$

2. Model 2

$$s_i^c = \alpha_{jk} + \beta_{jk} s_i^{\text{hs}} + \varepsilon_{ij}$$

   where the subindex k differs according to whether the high school entrance exam score is above or below the mean score in school j.

3. Model 3

$$s_i^c = \alpha_{jq} + \beta_{jq} s_i^{\text{hs}} + \varepsilon_{ij}$$

For each school j, the subindex q denotes the quintile to which the high school entrance exam score belongs.

Using each method we estimate parameter values that minimize our measure of distance. As a check that the model structure we impose is consistent with the data, we test for whether the distribution we estimated and the distribution of college scores that we observed from the data comes from the same underlying continuous distribution. We follow Anderson and Darling (1952) and Anderson (1962) to test the hypothesis:

$$H_0 : \hat{s}_c = s_c$$

$$H_1 : \hat{s}_c \neq s_c$$

Next, we present the results of our estimation.

4.2 Results

In this section, we will first present score distributions of all Science High Schools in the University and High School entrance exam. Then we focus on the three Science Schools for which we have more data. We focus on Science High School students as they follow a single track and because we have the ÖSS scores for all the students in three of these schools. Science track students focus on science and math tests so we use their quantitative score (ÖSS-SAY) in the ÖSS as their performance measure in the University Entrance Exam.

We transform simulated score distributions for the *subset* L of schools studied, i.e., Science High Schools, $(F^{\{L\}}(s))$ into percentiles:

$$F^{\{L\}}(s_i) = i, i = 1, ..100$$

so that $s_i = F^{\{L\}^{-1}}(i)$ is uniformly distributed over $[0, 100]$.

This then defines a distribution for school j as

$$F^j(s_i) = F^j(F^{\{L\}-1}(i)) = H(i) \text{ for } i \in [0, 100].$$

Simply put, we ask the following question: if all science high schools are uniformly distributed over the interval $[0, 100]$, what does the score distribution of a particular school look like? We argue that a school adds value by moving its score distribution into higher percentiles. Thus, if a high school admits students with scores in the 20th percentile (in the High School Entrance Exam) and graduates students with scores in the 40th percentile (in the University Entrance Exam) of all Science High Schools, it must be adding value!

Figure 6: ÖSS Score Distribution vs. High School Entrance Exam Score Distribution

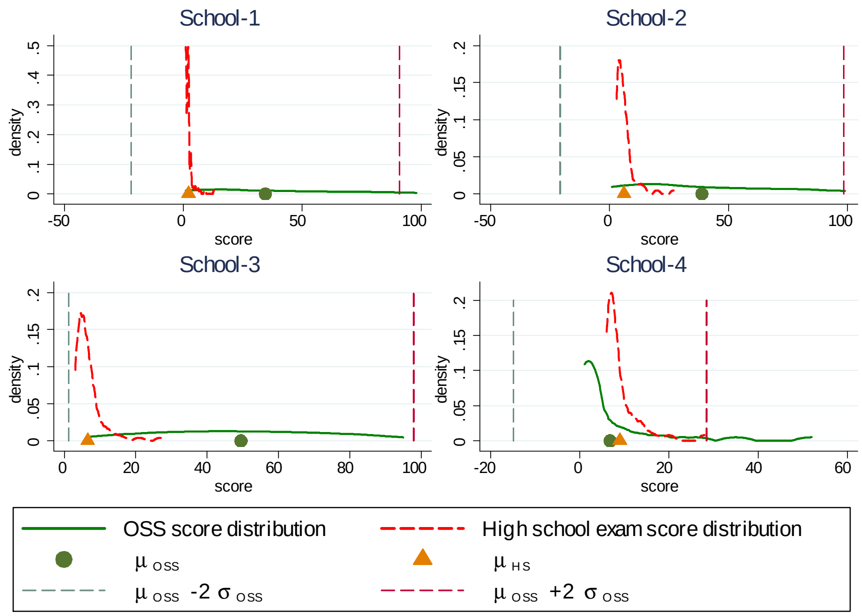
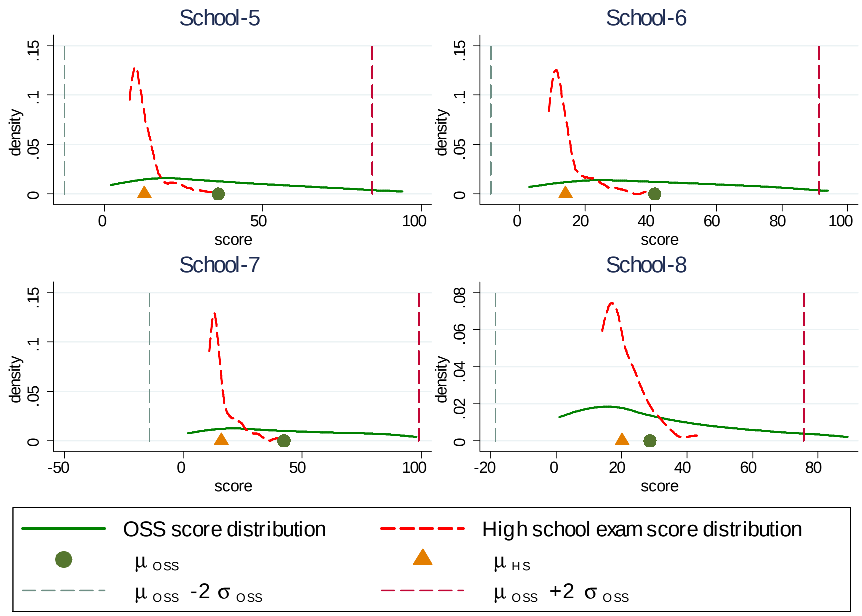
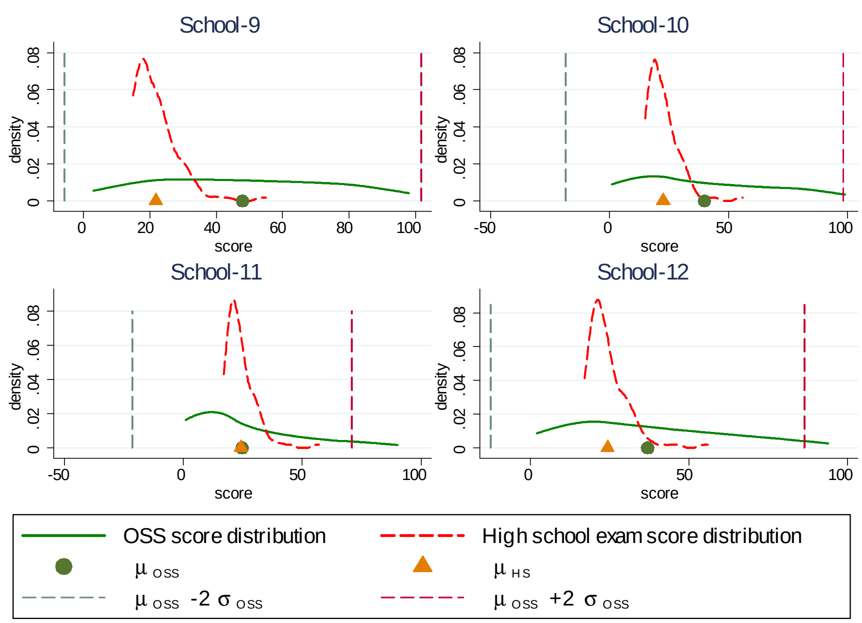
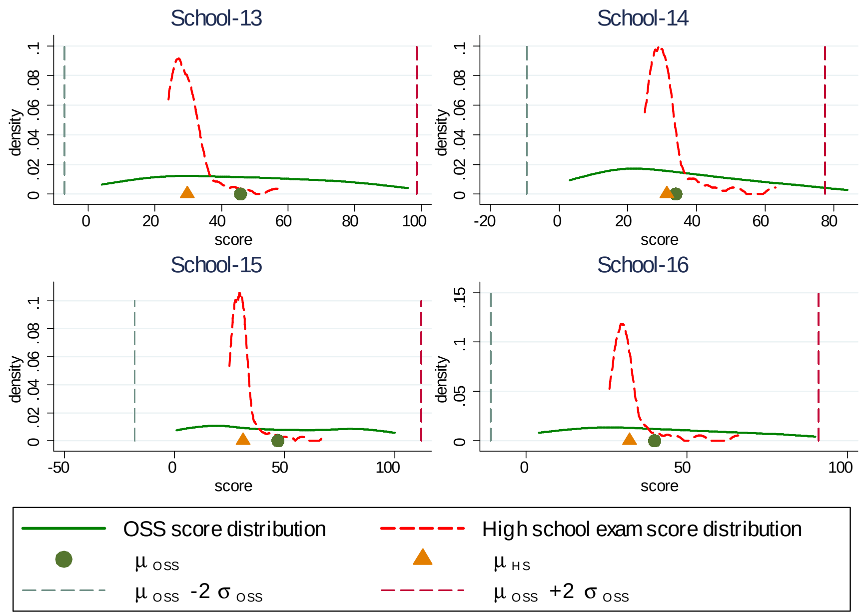


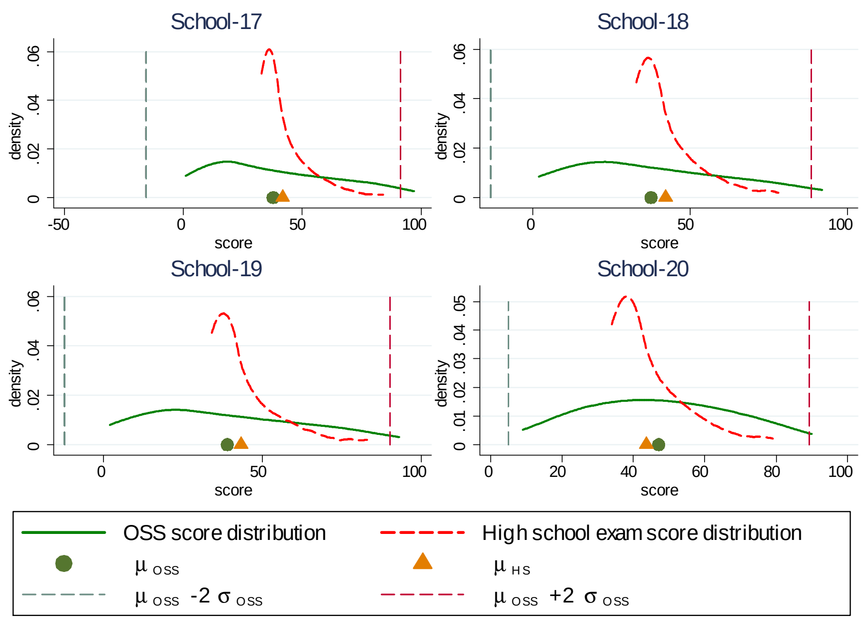
Figure 6 presents the distributions of scores in High school and University Entrance exam. For privacy reasons we do not name the schools, but in all figures, they are ordered according to their minimum cutoff scores in the 2001 High school entrance exam (OKS) so that School-1 has the lowest cutoff score, while School-46 has the highest cutoff score. As these figures show there is no evidence that more selective schools have greater value added. Students from more selective schools perform better in absolute terms. However if we compare their performance relative to where they started, we see no evidence of greater value added in more selective schools.

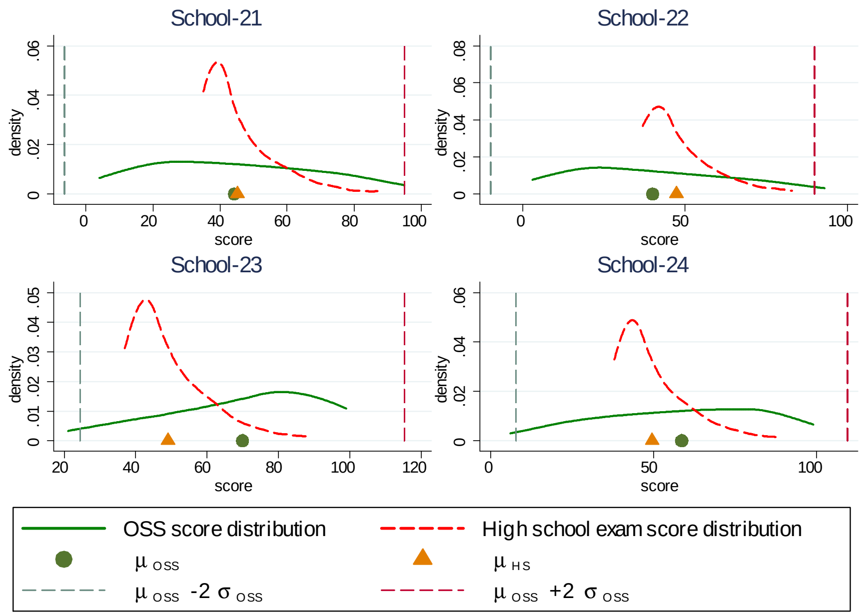
Figure 6 shows the score distributions for all schools: it shows how schools affect their students' ÖSS score relative to their high school entrance exam score. For example, school number 9 has a mean score in the OKS in the 21th percentile, but a mean score in the ÖSS in the 44th percentile. However, we need to do a bit more to see if the value-added of a school varies according to the rank of the students Here we will focus on the three schools that we have data on all students in University entrance exam, and use our model to estimate school value-added that varies by quintile. Figure 7, Figure 8 and Figure 9 present the results for

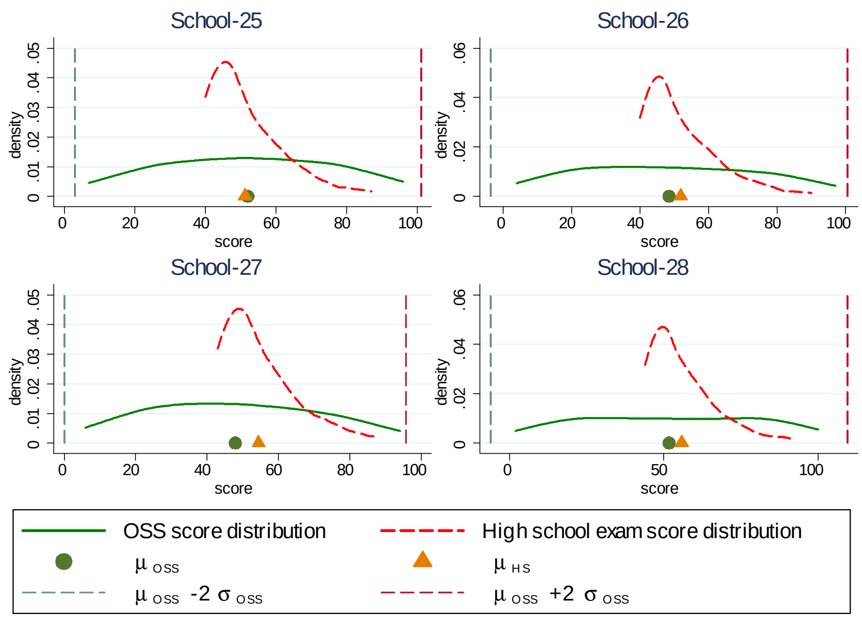

School-5
School-6
School-7
School-8
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS

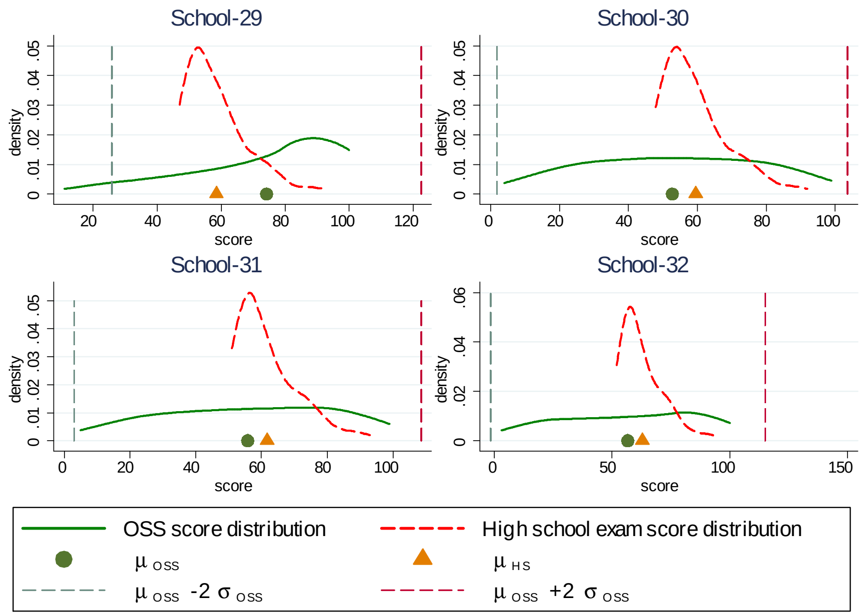

School-9
School-10
School-11
School-12
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS

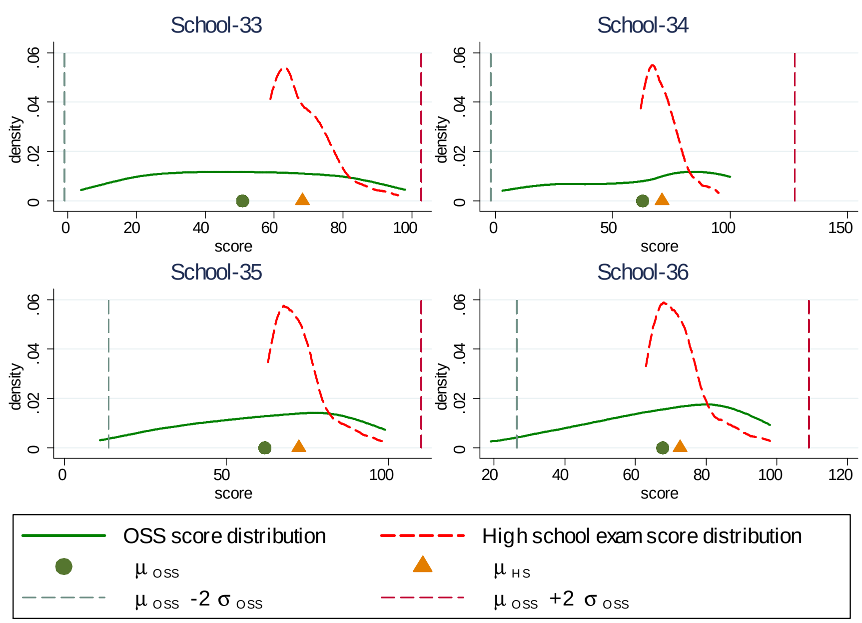

School-13
School-14
School-15
School-16
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS

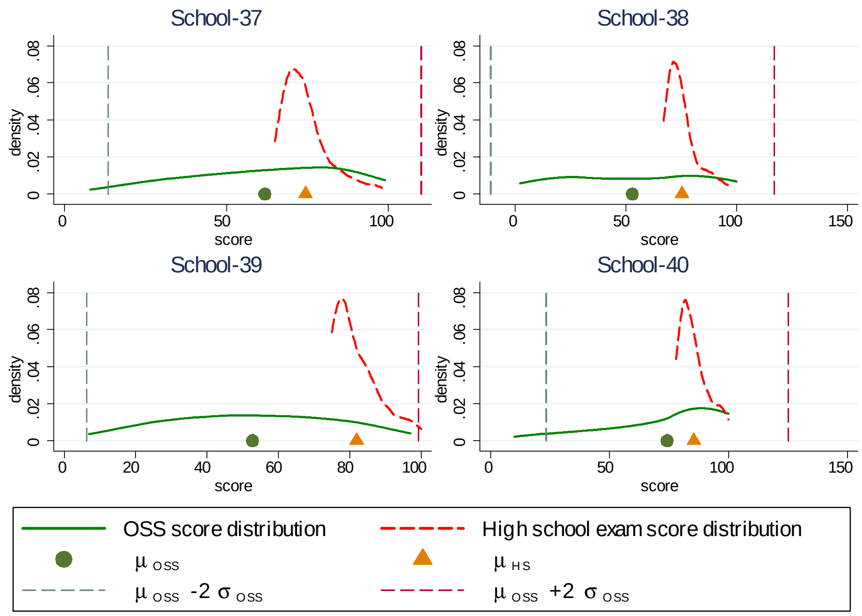

School-17
School-18
School-19
School-20
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS

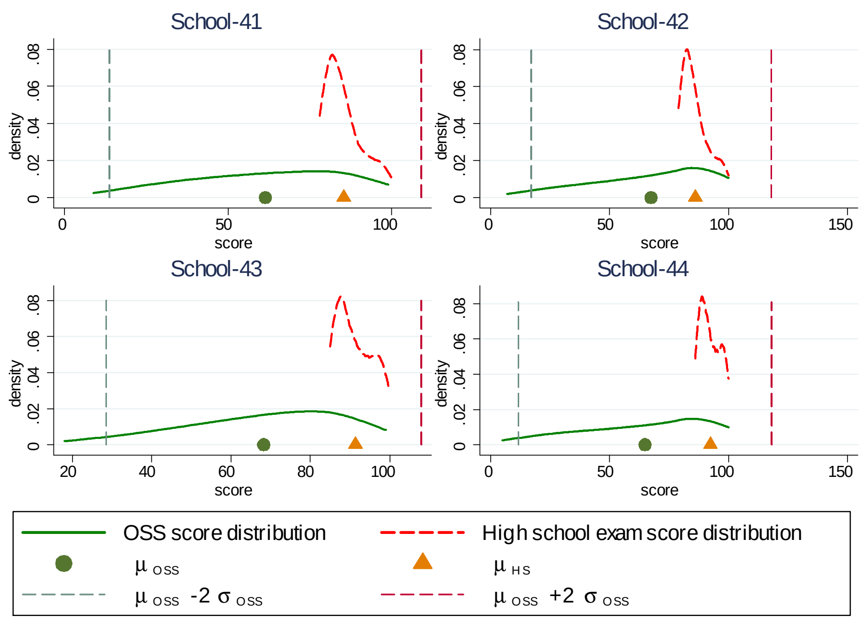

School-21
School-22
School-23
School-24
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS

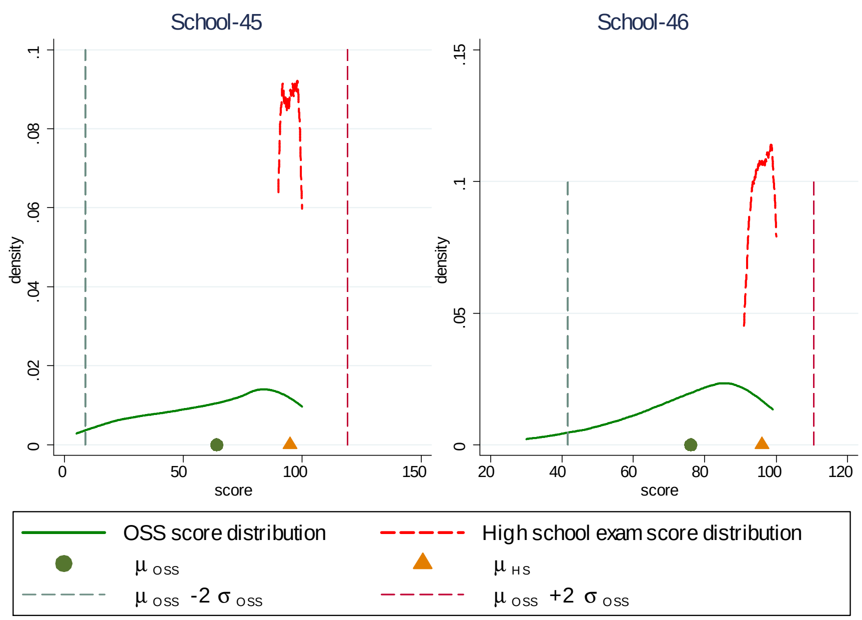

School-25
School-26
School-27
School-28
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS


School-29
School-30
School-31
School-32
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS


School-33
School-34
School-35
School-36
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS


School-37
School-38
School-39
School-40
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS


School-41
School-42
School-43
School-44
density
score
OSS score distribution
High school exam score distribution
μ OSS
μ HS
μ OSS -2 σ OSS
μ OSS +2 σ OSS



model 1, model 2 and model 3, respectively. These schools correspond to School-3, School-40, and School-43 in Figure 6, so that School-3 has the lowest cutoff score, while School-43 has the highest cutoff score. Note that scores are normalized within the subset of these three schools.

Model 1 assumes that schools have same effect on their students regardless of their initial ranking within their school. Figure 7 shows high school score distribution, estimated and real ÖSS exam score distribution and the average change in scores from OKS exam to ÖSS exam for each school. These results suggest that while School-3 and School-40 have positive effect on their students' performance when they take the ÖSS exam, and school-43 has a negative effect.

However,we reject the null hypothesis that the score distribution we observed from the ÖSYM data and the score distribution we estimate from our model comes from the same underlying distribution for all three schools. This suggests that schools have different effects on students from different portions of the initial distribution. Therefore, we extend our model to Model 2, and allow for different school effects for students below and above the average score within each high school.

Figure 7: Model 1: ÖSS Score Distributions vs. High School Entrance Exam Score Distribution

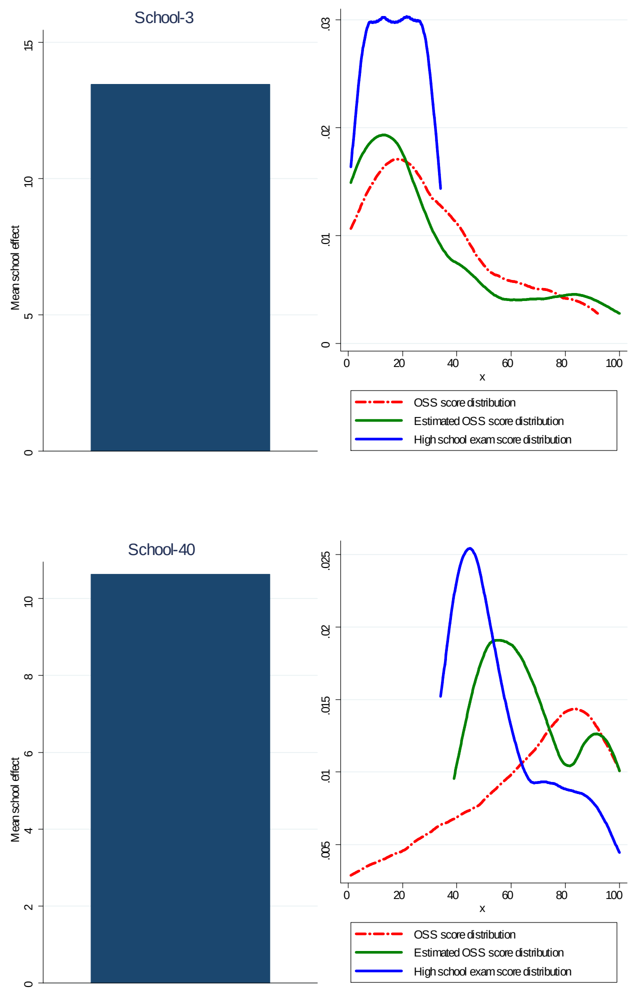


Figure 8 shows the school effect for both groups of students in each school. As it is clearly seen from the Figure 8, Model 2 also cannot explain the ÖSS exam score distribution in School-40 and School-43. Therefore we extend our model further and allow different effects on students in each quintile of the initial score distribution within school.

Figure 9 shows the predictions of Model 3 where we allow schools to have different effect on students coming from different quintiles. We cannot reject the null hypothesis (that the score distribution we observed from the ÖSYM data and the score distribution we estimate from our model comes from the same underlying distribution for all three schools) making

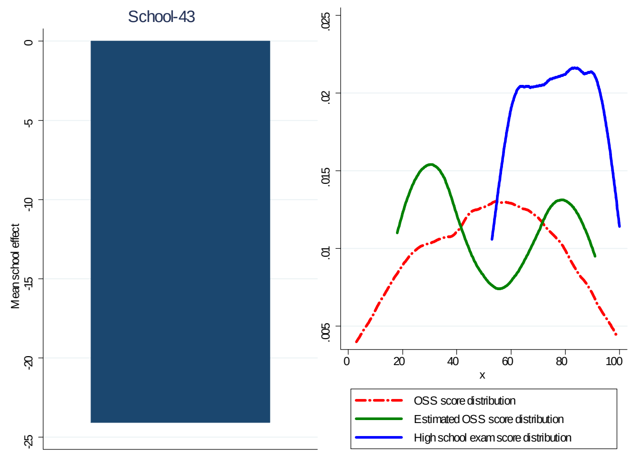
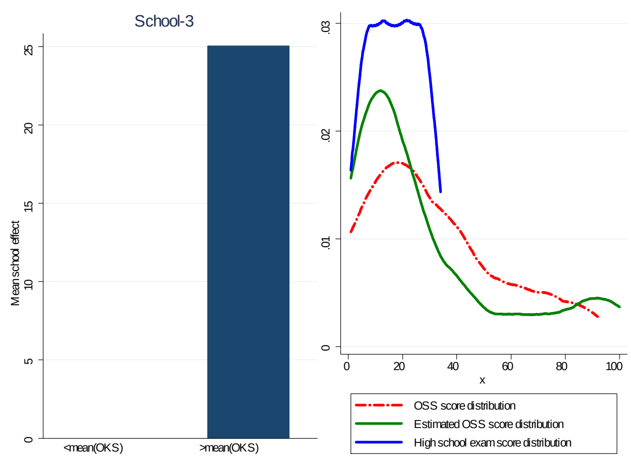
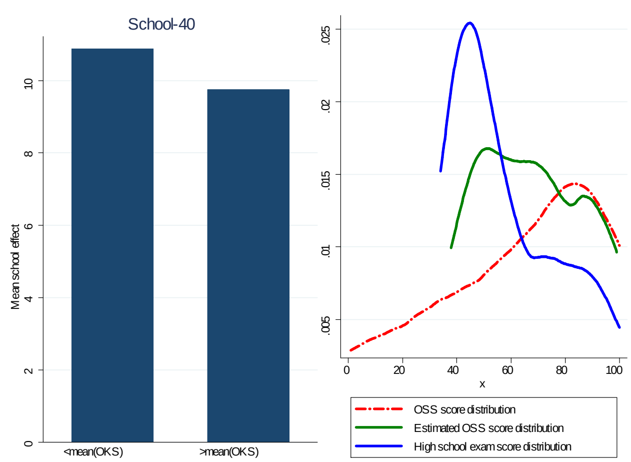


Figure 8: Model 2: ÖSS Score Distributions vs. High School Entrance Exam Score Distribution





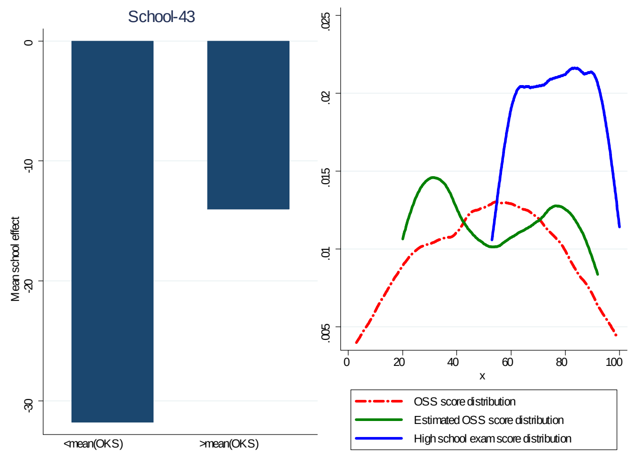
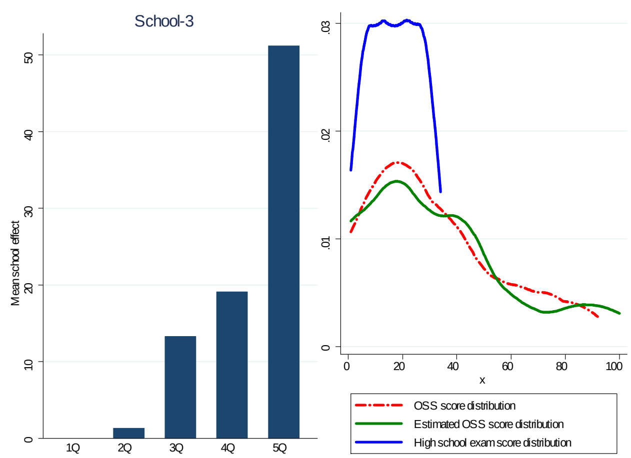
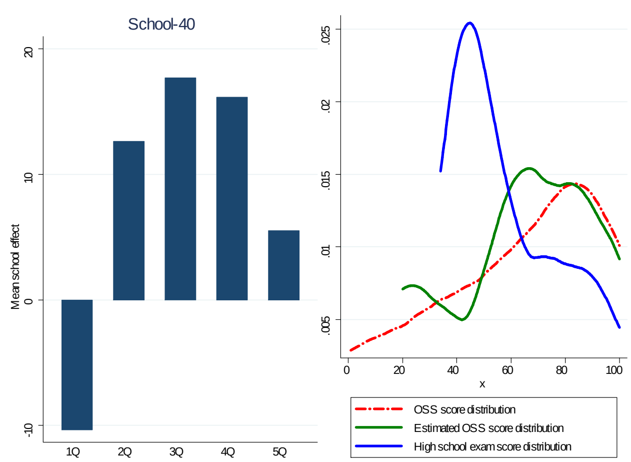
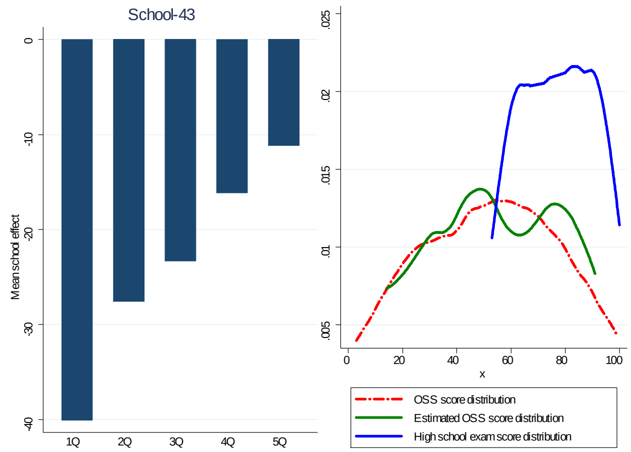


Figure 9: Model 3: ÖSS Score Distributions vs. High School Entrance Exam Score Distribution







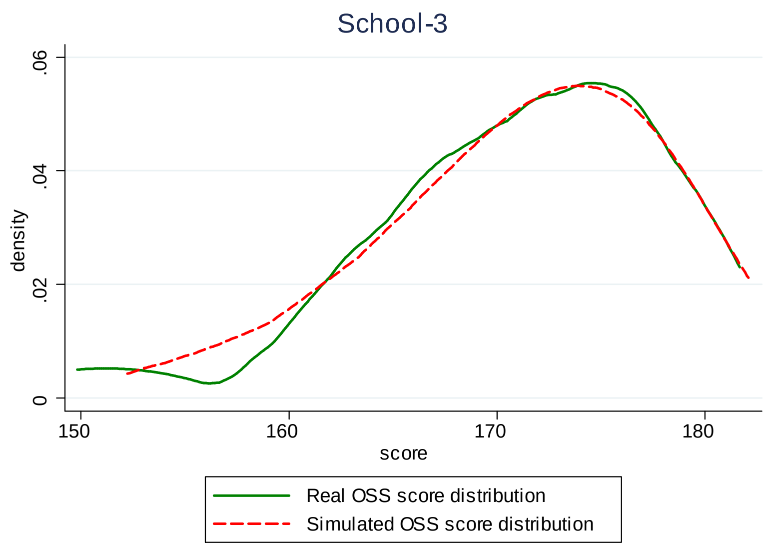
it our preferred model.

In this section we were mainly interested in the relation between perceived valuation (δ) and value-added of the schools. The perceived valuation of the schools follows the same order as their minimum cutoff score, that is School 43 is the highest valued school, while School 3 is the lowest valued school. However, Figure 9 suggests that School 43 has limited value added. Applicants mistakenly infer quality from selectivity! Cullen, Jacob,Levitt (2005) reach the same conclusion when they investigate the impact of school choice on students' outcomes in Chicago Public School system where students can apply for open enrollment in schools outside their district. Alstadsæter (2009), and Jacob, McCall, and Stange (2011) show the importance of the role of consumption value in students' school choice in different contexts. It is also important to note that we are only investigating the effect of exam schools on academic achievement. However, students attending exam schools may have other benefits that are valuable for them, but unobservable to us.

4.3 Validity Check

Score distributions presented in Figure 6 are based on our simulated ÖSS scores, since we don't observe all students in all schools. However we could get the ÖSS scores of all students for three schools, School-3, School-40, School-43. By using the information on these schools,

Figure 10: Real vs. Simulated ÖSS Score Distribution

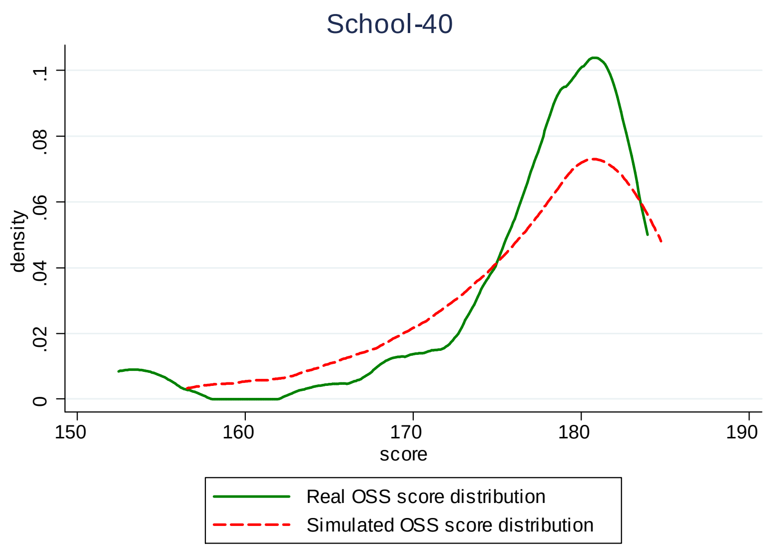


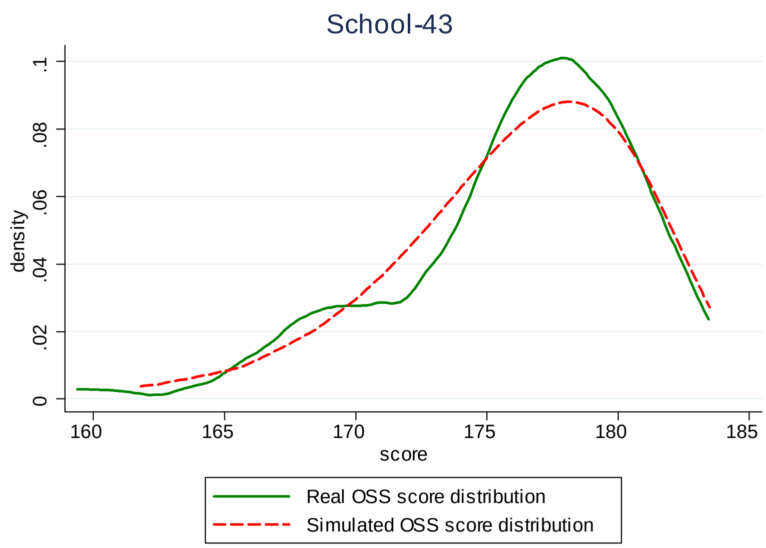
we check whether the simulated distributions (using means and standard deviations and imposing a Beta distribution) are similar to the actual distributions. Figure 10 shows real and simulated distributions for these schools, and Table 9 presents the descriptive statistics of real and simulated distributions. We also test the hypothesis that simulated and real distributions are equal by using Kolmogorov-Smirnov two-sample test. We cannot reject the null hypothesis for each of the three schools. This shows that our parametric assumptions seem reasonable.

5 Conclusion

In this paper, we employ the data collected from public sources on High school and University entrance exams to investigate the school characteristics that students seem to care about while choosing a school, and the effect of exam schools on students' performance in the ÖSS.

By introducing a multi step structural model, we tackle the problem of having limited data on individual students, and are able to recover preferences, score distributions for each


School-40
density
0
.02
.04
.06
.08
.1
150
160
170
180
190
score
Real OSS score distribution
Simulated OSS score distribution


School-43
density
0
.02
.04
.06
.08
.1
160
165
170
175
180
185
score
Real OSS score distribution
Simulated OSS score distribution

school and get some idea of their value added. We find some groups of schools are more vertically differentiated (those at the top of the pecking order) than others. We find evidence that more selective schools don't necessarily help their students to improve their scores. Elite schools seem to get better students because everyone wants to go to them, even when they need not add value to the student in terms of their performance in the University entrance exam. This may be because of a consumption value of going to such schools: bragging rights, or networks formed in such schools that are of value later. In this case, especially because higher income students are more likely to be able to get into such schools, it is hard to defend the subsidies received by elite schools.

References

Abdulkadiroğlu, A. & Angrist, J. D. & Pathak, P. A. (2011). "The Elite Illusion: Achievement Effect at Boston and New York Exam Schools", NBER Working Paper No. 17264

Alstadsæter (2009). "Measuring the Consumption Value of Higher Education", CESIFO Working Paper no. 2799

Anderson, T. W. & Darling, D. A. (1952). "Asymptotic Theory of Certain "Goodness of Fit" Criteria based on Stochastic Processes", The Annals of Mathematical Statistics, Vol. 23, No. 2.

Anderson, T. W. (1962). "On the Distribution of the Two-Sample Cramér-von Mises Criterion", The Annals of Mathematical Statistics, Vol. 33, No. 3.

Arcidiacono, P. & Aucejo, E. M. & Spenner, K. (2012). "What Happens After Enrollment? An Analysis of the Time Path of Racial Differences in GPA and Major Choice", Working paper

Azevedo, E. and Leshno, J. (2011). "A Supply and Demand Framework for Two-Sided Matching Markets". Mimeo. Harvard University.

Berry, S. (1994). "Estimating Discrete-Choice Models of Product Differentiation", RAND Journal of Economics, Vol. 25 pp. 242-262.

Bresnahan, T.F. & Stern, S. & Trajtenberg, M. (1997). "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," RAND Journal of Economics, Vol. 28 No. 0, S17-S44.

Burgess, S. & Greaves, E. & Vignoles, A. & Wilson, D. (2009). "What Parents Want: School Preferences and School Choice", Centre for Market and Public Organisation, Working Paper No. 09/222

Cameron , A. C.& Kim, N. (2001). "Simulation Methods for Nested Logit Models" presented in 2001 Econometric Society Australasian Meeting.

Caner, A. & Okten, C. (2012). "Higher Education in Turkey: Subsidizing the Rich or the Poor?," IZA Working Paper No. 7011

Clark, D. (2010). "Selective Schools and Academic Achievement", The B.E. Journal of Economic Analysis & Policy, Vol. 10, Issue 1, pp. 1935-1682.

Cullen, J. B. & Jacob, B. A. & Levitt, S. (2005). "The Impact of School Choice on Students Outcomes: An Analysis of the Chicago Public Schools" Journal of Public Economics, 89, pp. 729 - 760.

Cullen, J. B. & Jacob, B. A. & Levitt, S. (2006). "The Effect of School Choice on Participants: Evidence from Randomized Lotteries" Econometrica, Vol. 74, No. 5, pp. 1191-1230.

Dale, S. B. & Krueger, A. B. (2002). "Estimating the Payoff to Attending a more Selective College: An Application of Selection on Observables and Unobservables" Quarterly Journal of Economics, 2002, Vol. 107, pp. 1491-1527.

____________________________(2011). "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data", NBER Working Paper No.17159.

Darling-Hammond, L. & Amrein-Beardsley, A. & Haertel, E. & Rothstein, J. (2012). “Evaluating Teacher Evaluation”. Phi Delta Kappan, March 2012, vol. 93 no. 6, pp. 8-15.

Ding, W. & Lehrer, S. (2007). "Do Peers Affect Student Achievement in China's Secondary Schools?" The Review of Economics and Statistics 89 (2), 300–312.

Dobbie, W. & Fryer, R. G. (20011b). "Exam high schools and Academic Achievement: Evidence from New York City" NBER Working paper No. 17286

Duflo, E. & Dupas, P. & Kremer, M. (2011). "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya", American Economic Review, Vol. 101, pp. 1739–1774

Epple, D. & Romano, R. (2010). “Peer Effects in Education: A Survey of the Theory and Evidence,” Handbook of Social Economics, Vol. 1B, Edited by: Jess Benhabib, Alberto Basin, and Matthew O. Jackson, 1053-1163.

European Commission, (2009/2010). " Organization of the Education System in Turkey".

Gourieroux, C., Monfort, A. (1996) , "Statistics and Econometric Models: Volume 2, Testing, Confidence Regions, Model Selection and Asymptotic Theory" Cambridge University Press

Frisancho, V. & Krishna, K. & Yavas, C. (2011). “Learning Gains among Repeat Takers of the Turkish College Entrance Exam: Disentangling Selection and Learning Effects”, Working paper. Retrieved from `http://www.personal.psu.edu/vcf104/blogs/resume/1_{C}atchUp_{J}MP.pdf`.

Frisancho, V. & Krishna, K. (2012). “"Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch”, Working paper. Retrieved from `http://www.econ.psu.edu/~kkrishna/Papers/AA_IndiaJDE.pdf`.

Hanushek, E. A. & Kain, J.F. & Markman, J.M. & Rivkin, S.G. (2003). "Does peer ability affect student achievement?" Journal of Applied Econometrics 18 527–544.

Hastings, J. S. & Kane, T. J. & Staiger, D. O. (2006). "Parental Preferences and School Competition: Evidence from a Public School Choice Program", NBER Working Paper No. 11805

Hoxby, C.M. (2000). "Peer effects in the classroom: Learning from gender and race variation", NBER working paper No. 7867

Hsieh, C. & Urquiola, M. (2006). "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program", Journal of Public Economics, Vol. 90, pp.1477– 1503

Jacob, B. A. & McCall, B. & Stange, K. (2011). "The Consumption Value of Postsecondary Education", Working paper.

Jackson, C. K. (2010). "Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago", The Economic Journal, 120(549): pp. 1399-1429.

Kang, C. (2007). "Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea", Journal of Urban Economics, Vol. 61, pp. 458-495

Loury, L. D. & Garman, D. (1995). "College Selectivity and Earnings", Journal of Labor Economics, University of Chicago Press, 13(2), pp. 289-308

Macleod, W. B. & Urquiola, M. (2013). "Anti-Lemons: Reputation and Educational Quality", Working paper. Retrieved from http://www.columbia.edu/~msu2101/MacLeod-Urquiola(2013).pdf

Pop-Eleches, C. & Urquiola, M. (2010). "Going to a better school: Effects and Behavioral Responses", American Economic Review, forthcoming.

ÖSYM (2002). "Ortaöğretim Kurumlarına Göre 2002 Öğrenci Seçme Sınavı Sonuçları Kitabı "

Rothstein, J. M. and Yoon, A. (2009). “Mismatch in Law School”, Working paper.

______________. (2008). “Affirmative Action in Law School Admissions: What Do Racial Preferences Do?”, University of Chicago Law Review 75(2), Spring, pp. 649-714.

Sander, R. H. (2004). “A Systemic Analysis of Affirmative Action in American Law Schools”, Stanford Law Review, 57(2): 367-483.

Train, K. (2009). "Discrete Choice Methods with Simulation", Cambridge University Press.

Zabel, J. (2008), "The Impact of Peer Effects on Student Outcomes in New York City Public Schools", American Education Finance Association.

Zimmerman, D.J. (2003). "Peer effects in academic outcomes: Evidence from a natural experiment", Review of Economics and Statistics 85 pp. 9–23.

A Appendix

A.1 The Nested Logit Model

Suppose that individual i's choice set,C, contains $N+1$ alternatives. These alternatives are partitioned into K nests according to certain characteristics. Therefore we can write the choice set as:

$$C = \{B_1, B_2, \ldots, B_k\}$$

Let utility of the individual i from alternative j in nest k be

$$U_{ij} = \delta_{kj} + \varepsilon_{ij}$$

where δ_{kj} is the mean valuation of the alternative j. We can decompose δ_{kj} as:

$$\delta_{kj} = W_k + V_j$$

where W_k is the valuation related only to the nest characteristics and V_j is the valuation related with alternative j's attributes.

Let λ_k be the scale parameter of the nest k, which is inversely related with correlation of error terms within nest k.

The probability alternative i is chosen conditional on nest k being chosen is given by:

$$P(j|B_k) = \frac{\exp(\frac{V_j}{\lambda_k})}{\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k})}$$

The probability of nest k being chosen depends on the nest characteristics W_k, and inclusive value I_k, which depends on all the alternatives in the nest k.

$$P(B_k) = \frac{\exp(W_k + \lambda_k I_k)}{\sum\limits_{n=1}^{K} \exp(W_n + \lambda_n I_n)} \quad \text{where} \quad I_k = \log(\sum_{l\in B_k} \exp(\frac{V_l}{\lambda_k}))$$

We can write $P(j)$ as:

$$\begin{aligned} P(j) &= P(j|B_k)P(B_k) \\ &= \frac{\exp(\frac{V_j}{\lambda_k})}{\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k})} \frac{\exp(W_k + \lambda_k I_k)}{\sum\limits_{n=1}^{K} \exp(W_n + \lambda_n I_n)} \end{aligned}$$

Replace I_k by $\log(\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k}))$

$$\begin{aligned} P(j) &= \frac{\exp(\frac{V_j}{\lambda_k})}{\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k})} \frac{\exp(W_k + \lambda_k \log(\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k})))}{\sum\limits_{n=1}^{K} \exp(W_n + \lambda_n \log(\sum\limits_{l\in B_n} \exp(\frac{V_l}{\lambda_n})))} \\ &= \frac{\exp(\frac{V_j}{\lambda_k})}{\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k})} \frac{(\exp(W_k))(\sum\limits_{l\in B_k} \exp(\frac{V_l}{\lambda_k}))^{\lambda_k}}{\sum\limits_{n=1}^{K} (\exp(W_n))(\sum\limits_{l\in B_n} \exp(\frac{V_l}{\lambda_n}))^{\lambda_n}} \end{aligned}$$

Multiply both sides by $\frac{\exp(\frac{W_k}{\lambda_k})}{\exp(\frac{W_k}{\lambda_k})}$:

$$\begin{aligned}
P(j) &= \frac{\exp(\frac{W_k}{\lambda_k})}{\exp(\frac{W_k}{\lambda_k})} \frac{(\exp(W_k))(\exp(\frac{V_j}{\lambda_k}))(\sum_{l \in B_k} \exp(\frac{V_l}{\lambda_k}))^{\lambda_k - 1}}{\sum_{n=1}^{K} (\exp(W_n))(\sum_{l \in B_n} \exp(\frac{V_l}{\lambda_n}))^{\lambda n}} \\
&= \frac{\exp(\frac{W_k}{\lambda_k})}{\exp(\frac{W_k}{\lambda_k})} \frac{(\exp(\frac{W_k}{\lambda_k})^{\lambda_k})(\exp(\frac{V_j}{\lambda_k}))(\sum_{l \in B_k} \exp(\frac{V_l}{\lambda_k}))^{\lambda_k - 1}}{\sum_{n=1}^{K} (\exp(\frac{W_n}{\lambda_n})^{\lambda_n})(\sum_{l \in B_n} \exp(\frac{V_l}{\lambda_n}))^{\lambda_n}} \\
&= \frac{\exp(\frac{W_k}{\lambda_k})^{\lambda_k - 1}(\exp(\frac{V_j}{\lambda_k} + \frac{W_k}{\lambda_k}))(\sum_{l \in B_k} \exp(\frac{V_l}{\lambda_k}))^{\lambda_k - 1}}{\sum_{n=1}^{K} (\exp(\frac{W_n}{\lambda_n})^{\lambda_n})(\sum_{l \in B_n} \exp(\frac{V_l}{\lambda_n}))^{\lambda_n}}
\end{aligned}$$

Therefore

$$P(j) = \frac{(\exp(\frac{\delta_{kj}}{\lambda_k}))(\sum_{l \in B_k} \exp(\frac{\delta_{kl}}{\lambda_k}))^{\lambda_k - 1}}{\sum_{n=1}^{K} (\sum_{l \in B_n} \exp(\frac{\delta_{nl}}{\lambda_n}))^{\lambda_n}}$$

A.2 Cameron and Kim (2001)

Suppose that ε_1 and ε_2 are jointly distributed with bivariate extreme value distribution

$$H(\varepsilon_1, \varepsilon_2) = \exp\left(- \left(\exp(-\frac{\varepsilon_1}{\lambda}) + \exp(-\frac{\varepsilon_2}{\lambda}) \right)^{\lambda} \right)$$

Camoren and Kim (2001) propose that

$$\varepsilon_1 = a\xi + bv_1 + c$$

$$\varepsilon_2 = a\xi + bv_2 + c$$

where ξ, v_1, v_2 are independently distributed with univariate extreme value distribution, and a, b and c are the weights that match the moments of bivariate extreme value distribution.

$$E(\varepsilon_i) = E(a\xi + bv_1 + c) = a\gamma + b\gamma + c = \gamma$$

$$Var(\varepsilon_i) = a^2\frac{\pi^2}{6} + b^2\frac{\pi^2}{6} = \frac{\pi^2}{6}$$

$$Corr(\varepsilon_1, \varepsilon_2) = [1 - \lambda^2] = \frac{a^2}{a^2 + b^2}$$

This result in

$$a = \sqrt{1 - \lambda^2}$$

$$b = \sqrt{1 - a^2}$$

$$c = (1 - a - b)\gamma$$

where γ is the euler constant.

Table 4: Correlation in Minimum Cutoff Scores

Min Score	2000	2001	2002	2003	2004
2000	1.00	0.97	0.97	0.96	0.96
2001	0.97	1.00	0.97	0.96	0.95
2002	0.97	0.97	1.00	0.98	0.97
2003	0.96	0.96	0.98	1.00	0.98
2004	0.96	0.95	0.97	0.98	1.00

Source: Science and Anatolian High School's cutoff scores in 2000 - 2004 from Ministry of Education Website

This method is generalized to the multivariate extreme value distribution,

$$H(\varepsilon_{i0}, \varepsilon_{i1}, \ldots, \varepsilon_{iN}) = \exp\left(-\sum_{k=1}^{K}\left(\sum_{j\in B_k}\exp(-\frac{\varepsilon_{ij}}{\lambda_k})\right)^{\lambda_k}\right)$$

such that

$$\varepsilon_j \ = a_k\xi + b_k v_j + c_k$$

where

$$a_k = \sqrt{1-\lambda_k^2},\ b_k = \sqrt{1-a_k^2},\ c_k = (1-a_k-b_k)\gamma$$

A.3 Stability of Exam Schools' Cutoff Scores

Following tables show the correlation of cutoff scores over five year periods from 2000 to 2004. As the Tables 4 and 5 show correlation between minimum cutoff scores over years are no less than 0.95. The correlation between maximum cutoff score are lower than it is in minimum cutoff scores, but it is still around 0.8. Similarly we also look at how the rank of schools with respect to their minimum and maximum scores are correlated over years. Table 6 shows how schools' rank with respect to their minimum cutoff scores correlated over five years period. Similarly, Table 7 shows the corresponding table for the maximum cutoff scores. These tables show that exam schools' cutoff scores are stable in Turkey.

Table 5: Correlation in Maximum Cutoff Scores

Max Score	2000	2001	2002	2003	2004
2000	1.00	0.82	0.83	0.83	0.82
2001	0.82	1.00	0.80	0.82	0.78
2002	0.83	0.80	1.00	0.87	0.85
2003	0.83	0.82	0.87	1.00	0.86
2004	0.82	0.78	0.85	0.86	1.00

Source: Science and Anatolian High School's cutoff scores in 2000 - 2004 from Ministry of Education Website

Table 6: Correlation in Rank of Minimum Cutoff Scores

Rank of Min Score	2000	2001	2002	2003	2004
2000	1.000	0.953	0.946	0.943	0.946
2001	0.953	1.000	0.973	0.969	0.968
2002	0.946	0.973	1.000	0.985	0.979
2003	0.943	0.969	0.985	1.000	0.979
2004	0.946	0.968	0.979	0.979	1.000

Source: Science and Anatolian High School's cutoff scores in 2000 -2004 from Ministry of Education Website

Table 7: Correlation in Rank of Maximum Cutoff Scores

Rank of Max Score	2000	2001	2002	2003	2004
2000	1.000	0.785	0.800	0.793	0.771
2001	0.785	1.000	0.829	0.837	0.798
2002	0.800	0.829	1.000	0.858	0.838
2003	0.793	0.837	0.858	1.000	0.847
2004	0.771	0.798	0.838	0.847	1.000

Source: Science and Anatolian High School's cutoff scores in 2000 -2004 from Ministry of Education Website

Table 8: Validity Check: Instrumental Variable

Variable	Coefficients
Number of Available Seats	-0.0308
	(0.0661)
Average Quantitative Score in 2000 ÖSS	0.657
	(0.962)
Average Verbal Score in 2000 ÖSS	0.189
	(1.534)
Age	-0.119
	(0.443)
Science High School	68.34***
	(17.18)
Teacher High School	48.51*
	(19.98)
Anatolian High School in Istanbul	33.22
	(18.48)
Anatolian High School in Izmir	16.27
	(21.62)
Education Language- English	30.86
	(17.51)
Education Language- German	12.79
	(16.78)
Dormitory Availability	10.2
	(7.03)
Ankara	48.49**
	(15)
Istanbul	35.54***
	(9.186)
Izmir	40.94**
	(13.52)
Seats in better schools	-0.00500*
	(0.00242)
Seats in worse schools)	0.00166
	(0.0019)
Constant	737.1***
	(66.68)

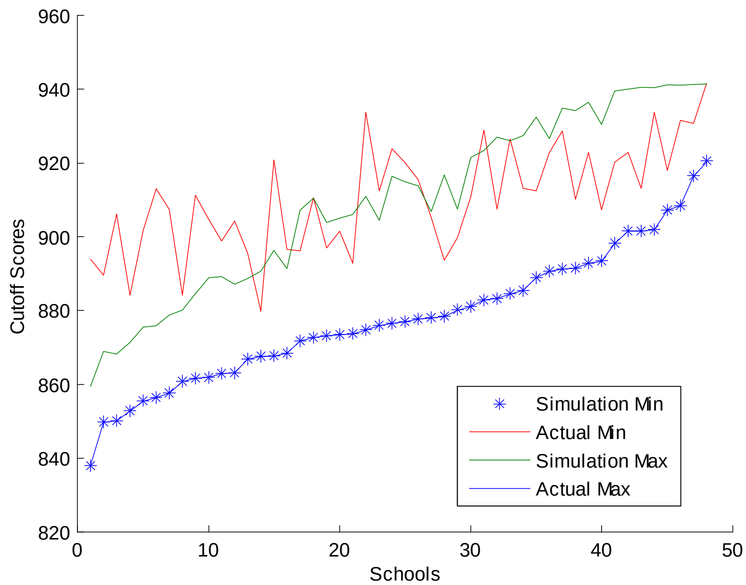
Note: Standard errors are reported in paranthesis. *, **, *** indicate significance at the .90, .95 and .99, respectively.

A.4 Additional Tables and Figures

Table 9: Summary Statistics

					Quantiles									
	Mean		Std		Min		0.25		0.5		0.75		Max	
School	Real	Simulated	Real	Simulated	Real	Simulated	Real	Simulated	Real	Simulated	Real	Simulated	Real	Simulated
School-3	171.18	171.18	7.13	6.97	149.82	152.23	166.36	166.99	171.98	172.22	175.92	176.43	181.67	182.12
School-40	177.41	177.41	6.74	6.53	152.45	156.53	176.04	174.26	179.68	179.22	181.34	182.36	183.94	184.74
School-43	176.17	176.16	4.62	4.56	159.39	161.81	174.01	173.58	177.04	176.90	179.49	179.53	183.45	183.49

Figure 11: Model Fit: Science High Schools Nest



A.5 Model Fit

Figure 12: Model Fit: Teacher High Schools Nest

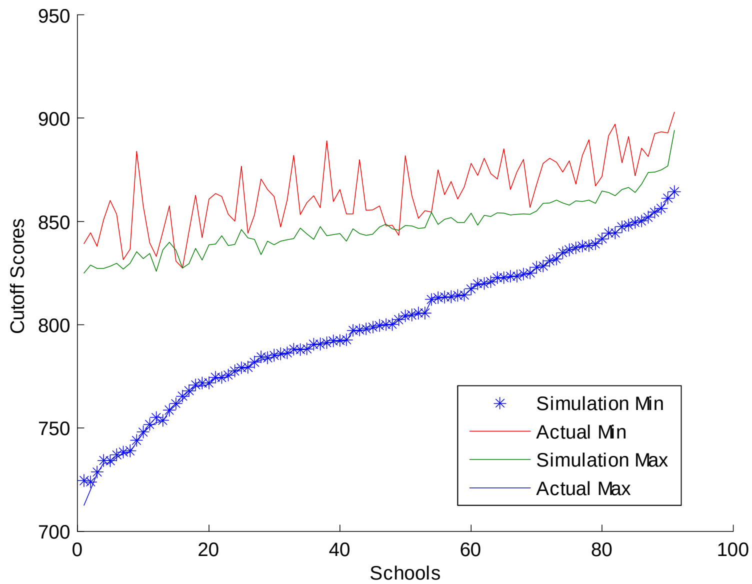


Figure 13: Model Fit: Ankara Anatolian High Schools Nest

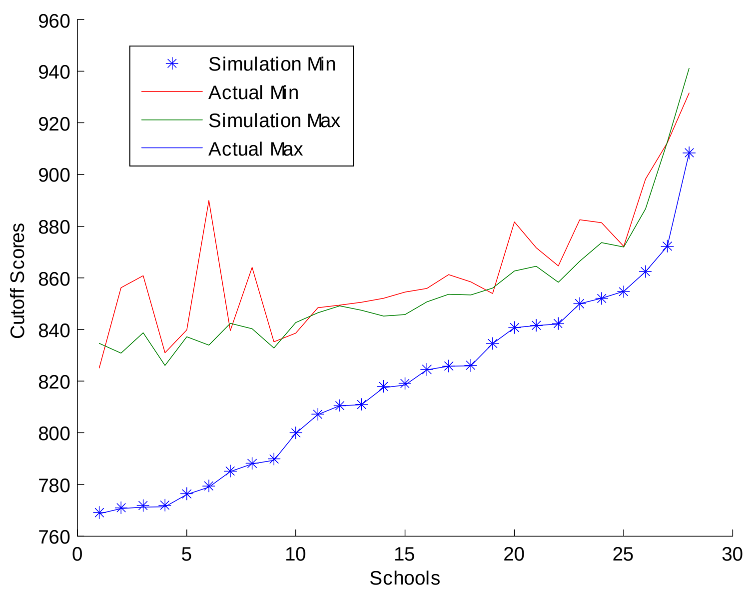


Figure 14: Model Fit: Istanbul Anatolian High Schools Nest

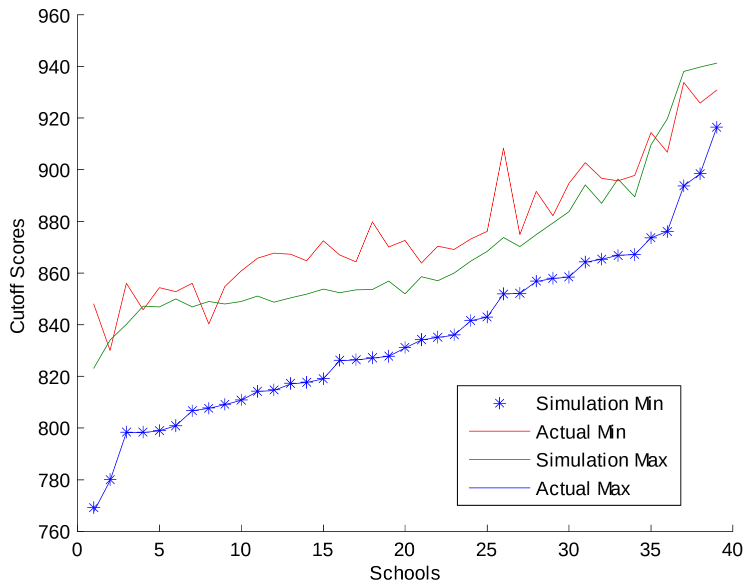


Figure 15: Model Fit: Izmir Anatolian High Schools Nest