Identification and Estimation in Manipulable Assignment Mechanisms*

Nikhil Agarwal† Paulo Somaini‡

March 19, 2014

PRELIMINARY AND INCOMPLETE
(Please do not quote or cite)

Abstract

We develop methods for empirically analyzing preferences and agent welfare using data from single unit assignment mechanisms that are not necessarily truthfully implementable. Our approach views the report made by an agent as a choice of a probability distribution over her assignments. When the finite market mechanism converges uniformly to a limit mechanism, the probability of assignment as a function of the report can be consistently estimated. We verify that this condition is satisfied for a large class of mechanisms that we call report-specific priority + cutoff mechanisms. This class includes the Boston Mechanism and the Deferred Acceptance mechanism. Certain welfare comparisons, such as pareto comparisons of mechanisms, are possible using data generated by two different mechanisms and estimates only of the probabilities of assignment. We then study identification of a latent utility preference model under the assumption that agents play a limit Bayesian Nash Equilibrium (limit equilibria are approximate equilibria in finite markets). Preferences are non-parametrically identified under either sufficient variation in choice environments or sufficient variation in a special regressor. We propose a second-step based on Gibbs' sampling to recover the full distribution of preferences. We apply our techniques to data from elementary school admissions in Cambridge, MA. In extensions, we relax the assumption that all agents are sophisticated.

JEL: C14, C57, C78, D47, D82

Keywords: Manipulable mechanism, school choice, preference estimation, identification

*We are grateful to Parag Pathak for sharing data on admissions in Cambridge Public Schools, and for several helpful discussions. We thank Steve Berry, Victor Chernozhukov, Phil Haile, Liran Einav, Whitney Newey, Ariel Pakes and Joe Shapiro for helpful discussions and suggestions. Vivek Bhattacharya provided excellent research assistance.

†Agarwal: Cowles Foundation for Economic Research, Yale University and Department of Economics, Massachusetts Institute of Technology, email: agarwaln@mit.edu

‡Somaini: Department of Economics, Massachusetts Institute of Technology and National Bureau of Economic Research, email: psomaini@mit.edu

1 Introduction

Several public school districts in the United States and abroad implement formal school choice mechanisms that solicit rank-order lists to assign students to schools (Abdulkadiroglu and Sonmez, 2003; Pathak and Sonmez, 2008). Admissions are typically based on student priorities, a lottery, and importantly, an expressed desire to attend a school. These data therefore provide a promising opportunity to estimate student preferences to evaluate economic questions. Recent empirical studies have used such estimates of student preferences to evaluate student welfare under alternative matching mechanisms (Abdulkadiroglu et al., 2014), implications for student achievement (Hastings et al., 2009), and school competition (Nielson, 2013). However, with rare exceptions, mechanisms used in the real world are susceptible to gaming (Pathak and Sonmez, 2008), making it difficult to directly interpret reported lists in as true preference orderings.

This paper proposes a general method for empirically analyzing preferences for schools using data from manipulable mechanisms. Previous empirical work has typically assumed that observed rank order lists are truthful representation of the students' preferences (Hastings et al., 2009; Abdulkadiroglu et al., 2014; Ayaji, 2013), allowing a direct extension of discrete choice demand methods with such data (c.f. McFadden, 1973; Beggs et al., 1981; Berry et al., 1995, 2004). The assumption is usually motivated by properties of the mechanism or by arguing that strategic behavior may be limited if there was a sudden change in the choice environment. This approach may not be valid for commonly used mechanisms such as the (old) Boston Mechanism or a deferred acceptance mechanism with limit on the length of the rank-order list. Students have a strategic incentives to manipulate their reports since applying to popular schools can be risky, and positions on the rank-lists are scarce. Anecdotal evidence from Boston suggests that parent groups and forums to exchange information about the competitiveness of various schools are fairly active (Pathak and Sonmez, 2008).

O

Our baseline approach is based on assuming that student reports are described by a limit Bayesian Nash Equilibrium. This in a large matching mechanism

A notable empirical exception to assuming truthful revelation of preferences and the closest paper to the proposed work is He (2012). This paper studies strategic behavior in the Boston mechanism in Beijing, and estimates preferences under the assumption that agents' reports are undominated using a full-likelihood approach. The approach relies deriving a limited number of restrictions implied by rationality using the specific number of schools and ranks that can be submitted in Beijing, and that the mechanism treats all agents sym-

metrically. Our approach allows for a general class of mechanisms that includes mechanisms with student priority groups and considers a broad range of restrictions implied by rationality. The proposed method does not require the researcher to analytically derive implications of rationality for use in estimation. Further, our aim is to characterize the identified set or show point identification under the restrictions imposed on the data and directly study the properties of an appropriate estimator, aspects which are not studied in He (2012).

We construct a method for estimating preferences by assuming that the observed rank order lists maximizes each students' expected utility. The key insight is that a reported preference list generates a probability of assignment to the various schools in the list (where the probabilities depend on the reports of other students). The reported rank-order list then reveals comparisons of expected utilities across different reports. We propose a two-step approach where we first estimate these probabilities and then use them to construct revealed preference inequalities. Since this assumes a strong form of rationality by students, we plan to study relaxations in which more limited features of the assignment probabilities and revealed preferences are used to estimate (possibly a set of) preference parameters.

There are two main motivations for estimating student preferences in this environment. First, a large, primarily theoretical, literature has taken a mechanism design approach to the student assignment problem to study algorithms that use submitted preferences list to a centralized clearinghouse to determine student placements (Gale and Shapley, 1962; Shapley and Scarf, 1974; Abdulkadiroglu and Sonmez, 2003). Such mechanisms are currently used in several school districts, including Beijing, Boston, Chicago, Denver, London, New Orleans and New York. Therefore, the preferences of the students are fundamental primitives that influence mechanism comparisons.

While the theoretical properties of various mechanisms have been well-studied and used to guide recent redesigns of matching markets (Roth and Peranson, 1999; Abdulkadiroglu et al., 2006, 2009), their empirical analysis has been relatively limited. Importantly, prospective analysis of the assignment effects of a change in the school choice mechanism is rare (see Pathak and Shi, 2013, for an exception). Such prospective analysis requires empirical estimates of preferences since these are the fundamental primitives determining student assignment.

One significant barrier to such prospective analysis is that the fundamental primitives are difficult to estimate since a large number of school choice mechanisms are susceptible to manipulation (Pathak and Sonmez, 2008, 2013). If the preference lists are not truthful, then the standard approach in the discrete choice literature assuming that higher ranked options are preferred to lower ranked options is not applicable. The proposed research aims to fill this gap. In our application, we will compare the welfare outcomes under the Boston mechanism

and a student proposing deferred acceptance mechanism. The relative benefits of these two mechanisms has been debated in the theoretical literature (Miralles, 2009; Abdulkadiroglu et al., 2011b; Featherstone and Niederle, 2011).

Second, recent work on measuring the effects of school assignment on student achievement simultaneously models the preferences for schools with test-score gains (Hastings et al., 2009; Walters, 2013; Nielson, 2013). This work has been motivated by the fact that without data from a randomized assignment of students to schools (Krueger, 1999; Chetty et al., 2011, among others) or quasi-experimental variation in assignment (Cullen et al., 2006; Deming et al., 2011; Abdulkadiroglu et al., 2011a, among others), a researcher must confront the possibility that students differentially sort into schools that result in idiosyncratic achievement gains. Even with a quasi-experimental design, it is unclear that estimates of Local Average Treatment Effects (LATE) are consistent for larger scale policy changes. Walters (2013) shows that estimates of charter school achievement gains obtained from lottery based research designs are biased for the achievement effect of assigning all students to charter schools if a selected group of students participate in a lottery.

Methods for estimating preferences in a broader class of data-environments expand our ability to study the effects of school assignment on student achievement. As shown in Walters (2013), a structural approach to the assignment problem can be used to extrapolate LATEs measured using quasi-experimental variation. Due to this inability to extrapolate outside the precise margins measured, some recent studies have obtained treatment effects of assignment into a students "first" choice school (Abdulkadiroglu et al., 2012). A selection on preferences is implicit this statement. We hope to extend our methods to combine these local estimates along the preference list to simulate achievement effects of a larger scale intervention.

The project contributes to the growing literature on methods for analyzing preferences in matching markets, both in settings with transferable and non-transferable utility. Many recent advances have been made in using pairwise stability as an equilibrium assumption on the final matches to recover preference parameters (Choo and Siow, 2006; Fox, 2010b,a; Chiappori et al., 2011; Agarwal, 2013; Agarwal and Diamond, 2014). The data environment considered here is significantly different and pairwise stability need not be a good approximation for assignments from manipulable mechanisms.

Our approach to studying large sample properties of our estimator and defining a limit mechanism is motivated by recent theoretical work studying matching markets by Kojima and Pathak (2009); Azevedo and Leshno (2013). Some of our proposed results rely on and extend results in Azevedo and Leshno (2013). While one may conjecture that in a large market, agents act as price-takers and cannot manipulate outcomes, Azevedo and Budish (2012) show that many mechanisms retain strategic incentives in a large market.

The proposed two-step estimator leverages insights from the industrial organization literature, specifically the estimation of empirical auctions (Guerre et al., 2000; Cassola et al., 2013) and dynamic games (Bajari et al., 2007). As in the methods used in those contexts, we use a two-step estimation procedure where the distribution of actions from other agents is used to construct probabilities of particular outcome as a function of the agents' own action and a second step is used to recover the primitives of interest. However, the nature of primitives, reports, the mechanism and economic environment are significantly different than in our context.

Our approach to identification results builds on insights from the identification of discrete choice demand (Matzkin, 1992, 1993; Lewbel, 2000; Berry and Haile, 2010). While the primitives are similar, unlike discrete choice demand, each report is a risky prospect that determines a probability of assignment to the schools on the list. In this sense, our paper is similar to Chiappori et al. (2012), although their paper focuses on risk attitudes rather than the value of underlying prizes.

2 Limit Mechanisms

We consider an environment with students indexed by $i \in \{1, \ldots, n\}$ and schools indexed by $j \in \{0, 1, \ldots, J\} = S$. School 0 denotes being unmatched. Each school has q_j seats, with $q_0 = \infty$. Each student has a priority type $t_i \in T$, where T is a finite set. For instance, students in the Cambridge school system and many others are prioritized based on whether live in the school walk-zone or if they have an older sibling in the school. The students in a mechanism by submitting a report $R_i \in \mathcal{R}_i$, which is a rank-order list over the schools $S = \{0, \ldots, J\}$. The set $\mathcal{R}_i$ may depend on the student's priority type t_i, and may be constrained. For example, students in Cambridge can rank up to three schools, and programs are distinguished by paid-lunch status of the student. The rank order lists of all students is denoted $R \in \mathcal{R}_1 \times \ldots \times \mathcal{R}_n = \mathcal{R}^n$, and the vector of priority types is denoted $t \in T^n$. At a minimum, we require the dataset to contain the reported rank order lists and the student's priority types.

A mechanism is an outcome of an algorithm that takes these rank-order lists as inputs. To study properties of a mechanism and our methods, it will be convenient to define this as a function that depends on the number of students n instead of the outcome of an algorithm.

Definition 1. *Fix a set of schools S and a sequence of capacities q_j^n. A* ***mechanism*** $\{\Phi^n\}$ *is a collection of functions $(\Phi^1, \ldots, \Phi^n)$ where*

$$\Phi^n : \mathcal{R}^n \times T^n \to (\Delta S)^n$$

such that for all $R \in \mathcal{R}^n$, *and* $t \in T^n$,

$$\sum_{i=1}^{n} \Phi_{ij}^n(R,t) \leq q_j^n.$$

In this notation, the $i-j$-th component of $\Phi^n(R)$, denoted $\Phi_{ij}^n(R,t)$ is the probability that student i is assigned to school j. Hence, the outcome for each student is in ΔS. In the Cambridge school system, there is a lottery that breaks ties between students. Such lotteries are a natural source of uncertainty faced by students.

We restrict attention to mechanisms that treat students with the same priority types symmetrically.

Definition 2. *A mechanism is* ***semi-anonymous*** *with priorities* T *if*

1. *for all* t, R_i, R_{-i} *and* i, i' *such that* $t_i = t_{i'}$, *we have that* $\Phi_i^n((R_i, R_{-i}), t) = \Phi_{i'}^n((R_i, R_{-i}), t)$.
2. *for all* R_i, R_{-i} *and* $R_{-i}^\pi = \left(R_{\pi(1)}, \ldots, R_{\pi(i-1)}, R_{\pi(i+1)}, \ldots, R_{\pi(n)}\right)$ *where* π *is a permutation of* $\{1, \ldots, i-1, i+1, \ldots, n\}$ *such that* $t_{\pi(i')} = t_{i'}$, *we have that* $\Phi_i^n((R_i, R_{-i}), t) = \Phi_i^n((R_i, R_{-i}^\pi), t)$.

The restriction to semi-anonymous mechanisms rules out admissions systems where a fine metric such as a test score is used to order students.

It is useful to consider the mechanism from an individual student's perspective. The student receives a probabilistic assignment as a function of her report and the joint distribution of priorities and reports of the other agents. The following result, which follows immediately from the definition of an anonymous mechanism, shows that this perspective is without further loss of generality.

Claim 1. *Let* $m(R_{-i}, t_{-i}) = \frac{1}{n-1} \sum_{i' \in -i} \delta_{(R_{i'}, t_{i'})}$ *be the empirical measure of reports of students other than* i. *If* Φ^n *is semi-anonymous, there exists a function* $\phi^n : (\mathcal{R} \times T) \times \Delta(\mathcal{R} \times T) \to \Delta S$, *such that*

$$\phi^n((R_i, t_i), m(R_{-i}, t_{-i})) = \Phi_i^n((R_i, R_{-i}), t).$$

Note that Φ_i^n only restricts $\phi^n((R_i, t_i), m(R_{-i}, t_{-i}))$ at a subset of probability measures m, namely, probability measures of the form $\frac{1}{n} \sum_i \delta_{R_i, t_i}$. We are free to choose ϕ^n at other values. Henceforth, we refer to a specific choice of ϕ^n when discussing a semi-anonymous mechanism.

With the above result in hand, it becomes natural to consider the limit of a mechanism as follows:

Definition 3. *The function* $\phi^\infty : (\mathcal{R} \times T) \times \Delta(\mathcal{R} \times T) \to \Delta S$ *is a* ***limit mechanism*** *of the sequence of semi-anonymous mechanisms* $\{\phi^n\}$ *if for all* R_i *and* $m \in \Delta(\mathcal{R} \times T)$*,*

$$\phi^\infty((R_i, t_i), m) = \lim_{n\to\infty} \phi^n((R_i, t_i), m)$$

Our identification and estimation results procedure will be based on properties of this limit mechanism and agent behavior. The route to econometric analysis relies on understanding whether the properties of a limit mechanism evaluted at m can be approximated by large finite mechanisms. A key property that will allow us to consider consistency properties of our estimator relies on proving the convergence in probability of the outcomes evaluated at the empirical distribution of the reports to its limiting values.

Condition 1 (Convergence)**.** *If* m^{n-1} *is an empirical measure of* $n-1$ *draws of* (R_i, t_i) *from the population measure* $m \in \mathcal{M}$*, then for each* (R, t)*,* $\phi^n((R, t), m^{n-1}) \xrightarrow{p} \phi^\infty((R, t), m)$ *uniformly in* m*.*

This condition can also be derived as a consequence of the continuous mapping theorem if, for example, $\phi^\infty((R, t), m)$ is continuous in m and $\phi^n((R, t), m^{n-1}) \to \phi^\infty((R, t), m)$ uniformly in m. The remainder of this section describes a large class of mechanisms in which the condition is satisfied, and that limit-equilibria of mechanisms that satisfy this condition are approximate equilibria in finite economies. While our baseline model will assume equilibrium play, Section 6 relaxes this strong notion of agent sophistication.

2.1 Report-Specific Priority and Cutoff Mechanisms

Verifying Condition 1 may not be straightforward because matching mechanisms are often described using algorithms instead of functions that take a measure of reports as inputs. In this section, we represent a class of mechanisms using rank-specific cutoffs and show that all mechanisms that admit this representation satisfy Condition 1.

Let $e_i \in [0, 1]^J$ be a vector of student eligibility score that encodes the student priority type t_i and a randomly generated lottery. Here, $e_{ij} > e_{i'j}$ implies that student i has higher eligibility at school j than student i'. Let $p \in [0, 1]^J$ denote a cutoff vector that determines the minimum eligibility score student that a school will accept. Given an eligibilty score e_i and a cutoff vector p, we say that student i's report R_i expresses individual demand

$$D_j^{(R_i, e_i)} = 1\{e_{ij} \geq p_j,\ jR_i\} \prod_{j' \neq j} 1\{jR_ij' \text{ or } e_{j'} < p_{j'}\}.$$

We can now define aggregate demand by integrating over the measure of student priorities

and reports η:

Definition 4 (Aggregate Demand)**.** *For a cutoff vector $p \in [0,1]^J$, the measure of students demanding $j \in S$ is given by*

$$D_j(p) = \eta\left(\{(e_i, R_i) : e_{ij} \geq p_j,\ jR_ii\} \bigcap_{j' \neq j} (\{(e_i, R_i) : jR_ij'\} \cup \{(e_i, R_i) : e_{ij'} < p_{j'}\})\right)$$

and the measure of student demanding no school is given by

$$D_0(p) = \eta\left(\bigcap_j (\{(e_i, R_i) : iR_ij\} \cup \{(e_i, R_i) : e_{ij'} < p_{j'}\})\right).$$

We make the following assumptions on η in the limit continuum economy:

Assumption 1. *The measure over student types η is satisfies*

(i) Non-degenerate lotteries: For each $p, p' \in [0,1]^J$ and R, $\eta(\{(e,R) : p \wedge p' \leq e \leq p \vee p'\}) \leq \kappa\|p - p'\|_\infty$.

(ii) Decreasing Demand: For all $p,\ p'$ such that $p_j < p'_j$ and $p_k = p'_k$ for all $k \neq j$, then $D_j(p|\eta) > D_j(p'|\eta)$.

(iii) Gross Substitutes: For all $p,\ p'$ such that $p_j < p'_j$ and $p_k = p'_k$ for all $k \neq j$, then $D_k(p|\eta) < D_k(p'|\eta)$ for all $k \neq j$.

(iv) Continuity: $D(p|\eta)$ is continuous in p.

Non-degenerate lotteries is a strengthening of strict preferences in Azevedo (2011). It allows for perfectly correlated lotteries, but maintains is satisfied, for example, if η admits a density. Decreasing demand and gross substitutes are satisfied in η has full support. Also satisfied if school preferences are perfectly correlated and preferences and there is full support on the preferences submitted by the students. Implied by assumptions in Azevedo and Leshno (2013).

Definition 5 (Market Clearing Cutoff)**.** *The vector of cutoffs p is a market clearing cutoff for economy (η, q) if for all $j \in S$, $z_j(p|\eta, q) = D_j(p|\eta) - q_j \leq 0$, with equality if $p_j > 0$.*

A unique market clearing cutoff exists in the limit economy under Assumption 1, and is approximated by the market clearing cuttofs of finite economies.

Lemma 1. *Consider an economy $[\eta, q]$. If η satisfies Assumptions 1(i)-(iii), then $[\eta, q]$ admits a unique market clearing cutoff p^*.*

Further, let p^n be a sequence of market clearing cutoffs for (η^n, q^n), where η^n is an empirical measure corresponding to n draws from η and $q^k \to q$. If η satisfies Assumption 1 and $q^n \to q$, then $\sup_{\{p^n : z(p^n|\eta^n, q^n) \leq q^n\}} \|p^n - p^*\| \xrightarrow{p} 0$.

Proof. See Appendix A.2 □

The proof of a unique cutoff is based on results in Azevedo and Leshno (2013). To show that the cutoffs converge, we first show uniform convergence in probability of the aggregate demand as a function of p using empirical process methods. This result is a straightforward consequence of the observation that the set of student-types that express demand for a school j form a VC-class (indexed by the price p). Combined with a unique cutoff in the limit, we can then show uniform convergence of p^n to p^*.

Definition 6. *A mechanism ϕ^n is a **report-specific priorities + cutoff** mechanism a there exists a function $f : \mathcal{R} \times [0,1]^J \to [0,1]^J$ such that*

(i) f strictly increasing in the last J arguments

(ii) $$\phi_j^n((R_i, t_i), m(R_{-i}, t_{-i})) = \int \dots \int D^{(R_i, f(R_i, \mathbf{e}_i))}(p^n) \mathrm{d}\eta_{R_1, \mathbf{e}_1 | t_1}(R_1, \cdot) \dots \mathrm{d}\eta_{R_n, \mathbf{e}_n | t_n}(R_n, \cdot)$$

(iii) p^n are market clearing cutoffs for each profile of reports and lotteries

$$((R_1, f(R_1, e_1)), \dots, (R_n, f(R_n, e_n)))$$

This mechanism finds market clearing cutoffs for an associated economy in which every agent's priority score is modified through f as a function of their report. The agent is then assigned to her most preferred program for which her modified priority score exceeds the cutoff.

Theorem 1. *Assume that $\eta^f(R, \{e : e \leq p\}) = \eta(R, \{e : f(R, e) \leq p\})$ satisfies Assumption 1. If ϕ^n is a report-specific priority + cutoff mechanism, then ϕ^n satisfies Condition 1*

The proof is based on.. The requirement that η^f satisfies Assumption 1 is jointly on the equilibrium play as well as the lotteries.

It turns out that several commonly used mechanisms can be expressed as report-specific lottery + cutoff mechanisms.

Our first set of results show that this formulation is useful for studying real-world mechanisms. We illustrate this using two commonly used mechanisms.

The **Student Proposing Deferred Acceptance Mechanism**: For reports $R_1, \ldots, R_N$ and priorities $t_1, \ldots, t_N$,

Step 1: Students apply to their first listed choice and their applications are *tentatively* held in order of priority and a lottery number until the capacity has been reached. Schools reject the remaining students.

Step k: Students that are rejected in the previous round apply to their highest choice that has not rejected them, and applications are help in order of priority and a lottery number until the capacity has been reached. The remaining students are rejected. Continue if any rejected student has not been considered at all their listed schools.

This mechanism is strategy-proof for the students if the students can rank all J schools (Dubins and Freedman, 1981; Roth, 1982), but provides strategic incentives for students if students are constrained to list $K < J$ schools.

The **Boston Mechanism**: For reports $R_1, \ldots, R_N$ and priorities $t_1, \ldots, t_N$, each school

Step 1: Assign students to their first choice in order of priority and a lottery number until the capacity has been reached. Reject the remaining students.

Step k: Assign students that are rejected in the previous round to their k-th choice in order of priority and a lottery number until the capacity has been reached. Schools reject the remaining students. Continue if any rejected student has not been considered at all their listed schools.

This mechanism is a canonical example for one that provides strategic incentives to students (Abdulkadiroglu et al., 2006).

Proposition 1. *The Deferred Acceptance Mechanism and the Boston Mechanism with lotteries are report-specific priority + cutoffs mechanisms.*

Proof. See Appendix A.4. We use $f(R, e) = e$ for deferred acceptance and $f_j(R, e) = \frac{e_j - \#\{k : kR_ij\}}{J} + 1$ for the Boston Mechanism. □

The function f used for the Boston Mechanism is intuitive. A student gains priority at a school relative to all students that rank that school lower on their list.

We allow for single tie-breaking as well as multiple tie-breaking in these mechanisms.

In addition, it is easy to verify that serial dictatorship mechanisms are also report-specific priority + cutoff mechanisms.

2.2 Limit Equilibrium

Our baseline approach assumes that the students submit rank-ordered lists according to a Bayesian-Nash Equilibrium (of the limit mechanism). However, many of our results including the limit properties of assignment probabilities does not rely on having a distribution of reports that is generated by the limit equilibrium. In an extension, we will consider the case when not all agents are sophisticated. Here, we will assume that sophisticated agents correctly forecast that a certain known fraction of agents are na ive and report their preferences truthfully. They then best respond to this distribution. We defer this to Section 6 and focus on the baseline case for now.

A (mixed) **strategy** is a function $\sigma_i : (\mathbb{R}^J \times T) \to \Delta\mathcal{R}_i$, where the domain is the vector (indirect) utilities for each school and the student's priority type. Denote the weight on the profile R as a function of $u_i = (u_{i1}, \ldots, u_{iJ})$ and t as $\sigma_i(u_i, t_i; R)$. We will only consider symmetric equilibria, i.e. $\sigma_i^*(u_i, t_i; R) = \sigma^*(u_i, t_i; R)$ for all i and R.

Student i's payoff from a particular report is based on the distribution of reports of the other students since that distribution affects the probability of assignment to the various schools. For the game implied by ϕ^n, the ante payoff from report R_i is given by

$$V_i^n(R_i, m_\sigma) = \mathbb{E}\left[\sum_j \phi_{ij}^n((R_i, t_i), m_\sigma^{n-1})u_{ij}\right]. \tag{1}$$

where $\hat{m}_\sigma$ is an empirical measure of $n-1$ iid draws from m_σ, with $m_\sigma(R, t) = f_T(t) \times \int \sigma(u, t; R)\mathrm{d}F_{U|T}$, $F_{U,T}$ is the joint distribution of player utilities and priority type. The expectation is taken with respect to uncertainty about $\hat{m}_\sigma$.

We will consider a limit Bayesian-Nash Equilibrium (BNE) and our first result is that if the limit mechanism approximates the finite mechanisms, then the limit BNE is an approximate BNE for the finite mechanisms.

Definition 7. *The strategy σ^* is a **limit equilibrium** if for all i, $\sigma^*(u_i, t_i; R_i) > 0$ implies that*

$$R_i \in \arg\max_{R_i' \in \mathcal{R}_i} \sum_j \phi_{ij}^\infty((R_i, t_i), m_\sigma)u_{ij}.$$

We will show that under certain conditions on the mechanism, the limit equilibria are approximate equilibria of a large mechanism in the following sense:

Definition 8. *The strategy σ^* is an ε-equilibrium of the n-th game if $\sigma^*(u_i, t_i; R_i) > 0$ implies*

$$V_i^n(R_i', m_{\sigma^*}) - V_i^n(R_i, m_{\sigma^*}) \leq \varepsilon\|u_i\|_\infty.$$

In an ε-equilibrium, no agent can expect a gain of more than a small fraction of her maximum payoff. The notion formalizes the idea that (most) agents do not have a large incentive to deviate from the prescribed strategy.

Proposition 2. *Let σ^* be a symmetric equilibrium of $\phi^\infty((R_i, t_i), m)$. If Condition 1 holds, then there exists an n_0 such that for all $n > n_0$, σ^* is an ε-equilibrium of ϕ^n.*

Proof. See Appendix A.1. □

This result justifies our focus on the limit-BNE as an approximation to large-sample play. The main advantage is that the limiting equilibria are more tractable for econometric analysis because the equilibrium itself does not depend on n. Without this simplification, a technical challenge is to show that reports sampled from equilibria that depend on n converge (see Menzel, 2012, for example).

A focus on equilibrium play implies that students submit the report that maximizes their expected utility with correct notions of the distribution of play by other students. This approach is a natural starting point for analyzing mechanisms that are not dominant-strategy and is commonly taken in the empirical analysis of auction mechanisms (Guerre et al., 2000; Cassola et al., 2013, among others). While there is little systematic evidence on the extent of optimal strategic behavior, anecdotal evidence suggests that parent groups and forums discussing the chances of getting into various schools and ranking strategies are active (Pathak and Sonmez, 2008). However, the assumption implies a strong degree of rationality and knowledge, particularly if parents vary in their level of sophistication as postulated by Pathak and Sonmez (2008, 2013). Section 6 presents proposed approaches to weaken this equilibrium assumption.

3 Identification

We follow the treatment in Berry and Haile (2010) for describing student preferences. For simplicity, we do not consider endogeneous characteristics.

Student i in market b picks a report $R_i \in \mathcal{R}_{b,i}$ when faced with the limit-mechanism ϕ_b^∞. This mechanism assigns her to a school in $S_b = \{0, 1, \ldots, J_b\}$ where 0 denotes remaining unassigned. Each school-student pair has observable characteristics z_{ijb}, some of which may vary only at the school level, and may include the student priority types t_i. Additionally, we allow for school unobservables $\xi_b(z_{ib})$ that can depend on the vector of observables $z_{ijb} = (z_{i1b}, \ldots, z_{iJ_bb})$.

We use a random utility model to represent student preferences. Let χ denote the support of (ξ_{jb}, z_{ijb}). Each student has an indirect utility function $v_i : \chi \to \mathbb{R}$. This formulation allows

for heterogeneous preferences conditional on observables. To formalize the realized utility function of each student as a random draw, let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space and assume that each

$$v_i(\xi_{jb}, z_{ijb}) = v(\xi_{jb}, z_{ijb}, \omega_i)$$

for some $\omega_i \in \Omega$ and measurable function v. ??? Say something about the restriction on ω Let the random variable $v_{ijb} = v(\xi_{jb}, z_{ijb}, \omega_i)$ and denote the joint distribution of

$$v = (v_{i1}, \ldots, v_{iJ_b})$$

with $F_V(v_{i1}, \ldots, v_{iJ_b}|\xi_b, z_{ib})$. We normalize the utility of not being assigned, v_{i0}, to zero and make the following assumption:

Assumption 2. *Change to admits a density... For all $p \in \Delta^{|S_b|}$ and $\alpha \in \mathbb{R}$, $\mathbb{P}(p \cdot v = \alpha|\xi_b, z_{ib}) = 0$.*

The assumption requires that the probability that any two lotteries over assignments to S_b yield the same expected utility is zero.

Our objective will be to identify $F_V(v_{i1}, \ldots, v_{iJ_b}|\xi_b, z_{ib})$ or functionals thereof using data from either a single or multiple large markets. Formally, a market is defined by the tuple

$$\Gamma_b = (\xi_b, z_{ib}, t_{ib}, m_b, \phi_b^\infty)$$

where m_b is the joint measure on the space of priority and reports of the other students. ??? Say something about interpretations of the various objects.

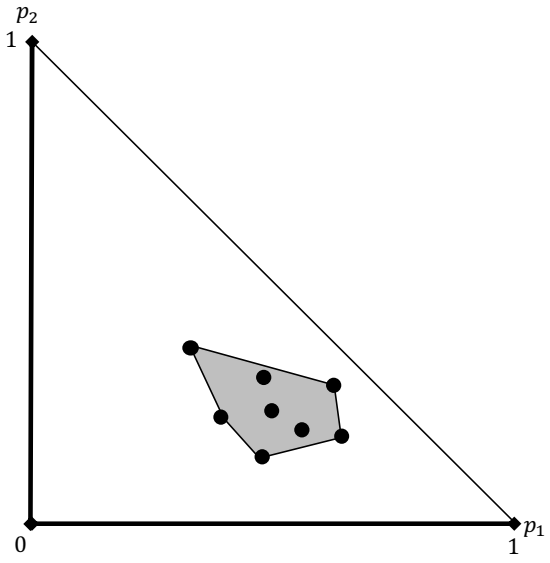
3.1 Choice Over Lotteries

Condition 1 implies that $\phi_b^\infty((R,t), m_b)$ is identified and can be consistently estimated with knowledge of the mechanism and the measure m_b. We will cast the student's choice of report $R \in \mathcal{R}_b$ as a choice over these lotteries. For simplicity of notation, let

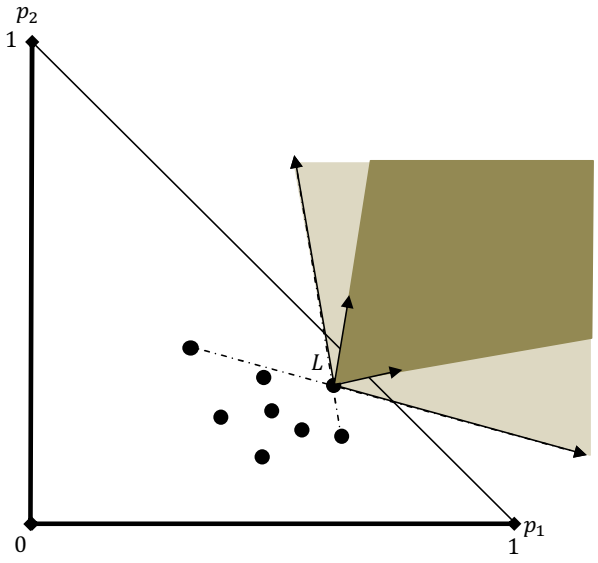
$$\mathcal{L}_\Gamma = \{L_R = \phi_b^\infty((R,t), m_b) : R \in \mathcal{R}_{t,b}\}$$

be the set of lotteries that an agent with priority type t chooses from in market Γ. In what follows, we will often be considering choices of agents in a single market Γ, and when clear from the context, we omit the subscript Γ.

Our baseline assumption is that students have consistent beliefs of the probability that they are assigned to each school in S_b as a function of their report $R \in \mathcal{R}$. Hence, a



(a) The shaded area is the convex hull of $\mathcal{L}$. There are 3 lotteries that are suboptimal for all students.



(b) The tangent cone to $\mathcal{L}$ at L is spanned by the two long directed arrows. The normal cone is spanned by two short directed arrows. These vectors are orthogonal to the tangent vectors.

Figure 1: In this example, $J = 2$ and $\mathcal{L}$ has 8 elements.

student with utility vector v maximizes expected utility by picking lottery L_R if and only if $L_R \cdot v \geq L \cdot v$ for all $L \in \mathcal{L}$. The set of students that choose lottery L_R therefore have utilities that belong to the normal cone to $\mathcal{L}$ at L_R:

$$N_{\mathcal{L}}(L_R) = \left\{ v \in \mathbb{R}^J : \forall L \in \mathcal{L}, \langle v, L_R - L \rangle \geq 0 \right\}.$$

Claim 2. *The exists utility vector $v \in \mathbb{R}^J$ such that $L_R \cdot v \geq L_{R'} \cdot v$ for all $L_{R'} \in \mathcal{L}$ if and only if L'_R is on the boundary of the convex hull of $\mathcal{L}$.*

Further, Assumption 2 implies that if L is not an extreme point of the convex hull of $\mathcal{L}$, then L is optimal only for a measure zero set of students.

For the baseline model, since the probability of picking an non-extremal lottery is zero, the assumption that agents are playing a limit Bayesian Nash equilibrium is testable. For now, assume that all lotteries in $\mathcal{L}$ are extremal, and relax this assumption in extensions.

The following sections studying identification treat the probability that a student picks $L \in \mathcal{L}$ in market Γ as data and study what can be learned from these objects. This probability is given by:

$$\begin{aligned} P(L \in \mathcal{L}|\Gamma) &= h_{N_{\mathcal{L}}(L)}(z, \xi) \\ &= \mathbb{P}(v \in N_{\mathcal{L}}(L)|\xi, z) \\ &= \int 1\{v \in N_{\mathcal{L}}(L)\} \mathrm{d}F_V(v|\xi, z). \end{aligned}$$

3.2 Identification Under Varying Choice Environments

In some cases, a researcher is willing to exclude certain elements of the priority structure t from preferences, or may observe data from multiple years in which the set of schools are the same, but either the mechanism is different or schools have different number of seats offered. Such variation can result in variation in the lotteries that a student picks from that is independent of preferences. More formally, consider the collection of markets

$$\mathcal{T}(\xi, z) = \{\Gamma_b = (\xi_b, z_{ib}, t_{ib}, m_b, \phi_b^\infty) : (\xi_b, z_{ib}) = (\xi, z)\}.$$

Say something like "Conditioning on z is without loss, but conditioning on ξ implies tha ability to keep school unobservables fixed."

The next result characterizes what can be learned about $F_V(v|z, \xi)$ from observing data from several large markets in $\mathcal{T}(\xi, z)$. Let $\mathcal{N}(\xi, z) = \{\text{int}(N_{\mathcal{L}_\Gamma}(L))\}_{\Gamma \in \mathcal{T}(\xi,z), L \in \mathcal{L}_\Gamma}$ be the collection of (the interiors of) normal cones to lotteries faced by agents in the markets

$\mathcal{T}(\xi, z)$. For a collection of sets $\mathcal{N}$, let $\mathcal{D}(\mathcal{N})$ be the smallest collection of subsets of $\mathbb{R}^J$ such that

1. $\mathbb{R}^J, \mathcal{N} \in \mathcal{D}(\mathcal{N})$
2. For all $N \in \mathcal{D}(\mathcal{N})$, $N^c \in \mathcal{D}(\mathcal{N})$
3. For all sequences $N_k \in \mathcal{D}(\mathcal{N})$ such that $N_{k_1} \cap N_{k_2} = \emptyset$, $\bigcup_k N_k \in \mathcal{D}(\mathcal{N})$

The collection $\mathcal{D}(\mathcal{N})$ is sometimes called the minimal Dynkin system containing $\mathcal{N}$.

Lemma 2. *Given $P(L \in \mathcal{L}_\Gamma|\Gamma)$ for each $\Gamma \in \mathcal{T}(\xi, z)$ and $L \in \mathcal{L}_\Gamma$, the quantity*

$$h_D(\xi, z) = \int 1\{v \in D\} \mathrm{d}F_V(v|\xi, z)$$

is identifed for each $D \in \mathcal{D}(\mathcal{N}(\xi, z))$.

Proof. The identified set of conditional distributions $F_V(v|\xi, z)$ is given by

$$\mathscr{F}_I = \left\{ F_V \in \mathscr{F} : \forall L \in \mathcal{L}_\Gamma \text{ and } \Gamma \in \mathcal{T}(\xi, z), P(L \in \mathcal{L}_\Gamma|\Gamma) = \int 1\{v \in N_{\mathcal{L}_\Gamma}(L)\} \mathrm{d}F_V(v|\xi, z) \right\}.$$

Note that for any two distributions F_V and $\tilde{F}_V$ in $\mathscr{F}$, the collection of sets

$$\mathscr{L}(F_V, \tilde{F}_V) = \left\{ A \in \mathcal{F} : \int 1\{v \in A\} \mathrm{d}F_V(v|\xi, z) = \int 1\{v \in A\} \mathrm{d}\tilde{F}_V(v|\xi, z) \right\}$$

is a Dynkin system. Since $\mathcal{D}(\mathcal{N}(\xi, z))$ is the minimal Dynkin system where all elements of $\mathscr{F}_I$ agree, $\mathcal{D}(\mathcal{N}(\xi, z)) \subseteq \mathscr{L}(F_V, \tilde{F}_V)$ for any two elements F_V and $\tilde{F}_V$. Hence, for all $D \in \mathcal{D}(\mathcal{N}(\xi, z))$, we have that

$$h_D(\xi, z) = \int 1\{v \in D\} \mathrm{d}F_V(v|\xi, z) = \int 1\{v \in D\} \mathrm{d}\tilde{F}_V(v|\xi, z)$$

is identified. □

The result follows from basic measure theory and characterizes the features of $F_V(v|\xi, z)$ that are identified under such variation in choice environments without any further restrictions. Typically, this implies only partial identification of $F_V(v|\xi, z)$, but extensive variation in the lotteries could result in point identification. In particular, with the free normalization $\|v_i\| = 1$, for any set C in the $J-1$ dimensional sphere, $\mathbb{S}^J$, a basic uniqueness result from measure thoery implies that $F_V(v|\xi, z)$ is identified if and only if the Dynkin-system $\mathcal{D}(\mathcal{N})$ is the Borel σ-algebra.

3.3 Identification With Preference Shifters

In this section we assume that the indirect utility function has the form

$$v(\xi_{jb}, z_{ijb}, \omega_i) = \nu(\xi_{jb}, z^2_{ijb}, \omega_i) - z^1_{ijb} \tag{2}$$

where $\omega_i \perp z^1_{ijb}$. The term z^1_{ijb} is sometimes referred to as a special regressor (Lewbel, 2000; Berry and Haile, 2010). Let $\zeta(\xi, z^2)$ be the support of z^1 conditional on (ξ, z^2). Say something about dropping notation on ξ, z^2. Since $f_V(v|z^1)$ is a location family, this implies that $f_V(v|z^1) = g(v + z^1)$ where g is the density of ν.

Before proceeding, we introduce two definitions:

Definition 9. *A convex cone C is **simplicial** if it is spanned by J linearly independent vectors $\{v_1, v_2, ..., v_J\} = V$ so that $C = \{v \in \mathbb{R}^J : v = Va \text{ for some } a \geq 0\}$.*

*A cone C is **salient** if $v \in C \implies -v \notin C$ for all $x \neq 0$.*

Note that every J-dimensional polygonal convex cone C has a simplicial decomposition $\{C_q\}_{q=1}^Q$ satisfying: i) each C_q is a simplicial convex cone, ii) $C = \cup_{q=1}^Q C_q$, iii) and $C_q \cap C_{q'}$ is empty or a lower dimension convex cone.

The first identification result holds for lotteries whose normal cone to $\mathcal{L}$ is simplicial, or equivalently, their tangent cone to the set $\mathcal{L}$ is simplicial.[1] This identification result exploits local variation of d.

Theorem 2. *Let C be a convex simplicial cone. If $h_C(z^1)$ is known on an open set containing z^1, then $g(z^1)$ is identified. Hence, $f_V(v|z^1)$ is identified everywhere if $\zeta = \mathbb{R}^J$.*

In particular, the result holds if $C = N_{\mathcal{L}_\Gamma}(L)$ for some Γ and $L \in \mathcal{L}_\Gamma$ or if $C \in \mathcal{D}(\mathcal{N})$.

Proof. Let V be a matrix of linearly independent vectors such that (the closure of) $C = \{v : v = Va \text{ for some } a \leq 0\}$ and $|\det V| = 1$. Evaluating h_C at Vx:

$$h_C(Vx) = \int_{\mathbb{R}^J} 1\{\varepsilon - Vx \in C\} g(\varepsilon) \, \mathrm{d}\varepsilon.$$

After the change of variables $\varepsilon = Va$:

$$\begin{aligned} h_{L,\mathcal{L}}(Vx) &= \int_{\mathbb{R}^J} 1\{V(a-x) \in N_{\mathcal{L}}(L)\} g(Va) d\varepsilon \\ &= \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_J} g(Va) \, da \end{aligned}$$

[1] The normal cone to the set $\mathcal{L}$ at point L is simplicial if and only if the tangent cone to the set $\mathcal{L}$ at point L is simplicial.

where the second inequality follows because $1\{V(a-x) \in N_{\mathcal{L}}(L)\} = 1\{a-x\} \leq 0$. Then:

$$\frac{\partial^J h_{L,\mathcal{L}}(Vx)}{\partial x_1 ... \partial x_J} = g(Vx)$$

and $g(\varepsilon)$ is identified by $\frac{\partial^J h_{L,\mathcal{L}}(Vx)}{\partial x_1 ... \partial x_J}$ evaluated at $x = V^{-1}\varepsilon$. □

The result follows by using local variation around z^1 to identify the density of ν evaluated at z^1. Simplicial cones allow us to do this since they are spanned by linearly independent vectors. Hence, to identify tails of the distribution of ν, we need to use extreme values of z^1, and full support in z^1 allows us to get the full distribution of ν. Doing this only requires one convex cone generated by a lottery, and therefore, observing additional lotteries with simplicial cones generates testable restrictions.

It turns out that if $J = 2$, then all normal cones $N_{\mathcal{L}}(L)$ are simplicial if L is an extremal lottery in $\mathcal{L}$. For $J > 3$, this need not be the case. In particular, it may be that $\mathcal{D}(\mathcal{N})$ does not contain a simplicial convex cone.

Fortunately, we can still identify g if $\zeta = \mathbb{R}^J$ as long as the tails of g are exponentially decreasing. Formally, assume that the density of ν belongs to the set

$$\mathcal{G} \equiv \left\{ g \in L^1\left(\mathbb{R}^J\right) : e^{c|x|} g(x) \in L^1\left(\mathbb{R}^J\right) \text{ for some } c > 0 \right\}.$$

Theorem 3. *Assume that $g \in \mathcal{G}$ and there is a lottery L such that $N_{\mathcal{L}}(L)$ is a salient convex cone with a non-empty interior. If $\zeta = \mathbb{R}^J$, then g is identified from*

$$h_{N_{\mathcal{L}}(L)}\left(z^1\right) = P(L \in \mathcal{L} | z^1).$$

Proof. Define the linear operator A:

$$A \circ g(z) = \int_{N_{\mathcal{L}}(L)} g(v+z)\, \mathrm{d}v.$$

We need to show that A in injective on $\mathcal{G}$. The proof is by contradiction. Suppose that there are $g', g'' \in \mathcal{G}$ such that $Ag' = Ag''$ but $g' - g'' \neq 0$.

Since the cone $N_{\mathcal{L}}(L)$ is salient, its dual $T_{\mathcal{L}}(L)$ has a nonempty interior. Let $\varepsilon \in interior(T_{\mathcal{L}}(L))$, with $|\epsilon|$ sufficiently small so that $g_\varepsilon(z) = g(z) e^{2\pi\langle \varepsilon, z\rangle} \in L^1$. Note that $1\{z \in N_{\mathcal{L}}(L)\} e^{-2\pi\langle \varepsilon, z\rangle} \in L^1$ for every $\varepsilon \in \mathrm{int}(T_{\mathcal{L}})$ because $\langle \varepsilon, z\rangle > 0$.

Let $g = g' - g''$. Since $A \circ (g' - g'') = 0$ and $\zeta = \mathbb{R}^J$, we have that for all $z \in \mathbb{R}^J$,

$$A \circ g(z) \;=\; e^{-2\pi\langle \varepsilon, z\rangle} \int 1(v \in N_{\mathcal{L}}(L))\, e^{-2\pi\langle \varepsilon, v\rangle} e^{2\pi\langle \varepsilon, v+z\rangle} g(v+z) dv = 0.$$

Since $e^{-2\pi\langle\varepsilon,z\rangle} > 0$ for all z, $A \circ g = 0 \iff \hat{f}_{\varepsilon,N_{\mathcal{L}}(L)}(\xi) \cdot \overline{\hat{g}}_{\varepsilon}(\xi) = 0$, where $\hat{f}_{\varepsilon,N_{\mathcal{L}}(L)}$ is the Fourier Transform of $f_{\varepsilon,N_{\mathcal{L}}(L)}(x) = 1\{x \in N_{\mathcal{L}}(L)\}e^{-2\pi\langle\varepsilon,x\rangle}$ and $\overline{\hat{g}}_{\varepsilon}$ is the conjugate of the Fourier Transform of $g_{\varepsilon}(x)$, both continuous functions in L^1. Since $\hat{g}_{\varepsilon}$ is continuous, the set where $\hat{g}_{\varepsilon} \neq 0$ is open. Further, since $g \neq 0$, the support of $\hat{g}_{\varepsilon}$ is non-empty. It follows that there is an open Z_{ϵ} where $\hat{g}_{\varepsilon}$ is different from zero, and therefore, $\hat{f}_{\varepsilon,N_{\mathcal{L}}(L)}(\xi) = 0$ for all $\xi \in Z_{\epsilon}$. This contradicts the fact that $\hat{f}_{\varepsilon,N_{\mathcal{L}}(L)}$ is an entire function, as shown in Lemma 3. □

Lemma 3. *Let $f_{\varepsilon,\Gamma}(x) = \chi_{\Gamma}(x)\, e^{-2\pi\langle\varepsilon,x\rangle}$ for some polygonal, full-dimensional, salient, convex cone Γ and $\varepsilon \in \text{int}(\Gamma^o)$, and let $\hat{f}_{\varepsilon,\Gamma}(\alpha)$ be its Fourier Transform. There is no open subset of $\mathbb{R}^J$ where $\hat{f}_{\varepsilon,\Gamma}$ is zero.*

Proof. Let K be a full-dimensional simplicial convex cone such that $\Gamma \subset K$; and $\{\Gamma_1 ... \Gamma_Q\}$ a simplicial triangulation of Γ. K exists because Γ is salient. Let V_q be a matrix $[v_{q1}, v_{q_2}, ..., v_{qn}]$ with the linear independent vectors that span cone Γ_q arranged as column vectors. $x \in \Gamma_q \iff x = V_q\alpha$ for some $0 \leq \alpha \in \mathbb{R}^J \iff V_q^{-1}x \geq 0$. Normalize V_q so that $\det|V_q| = 1$. Let $f_{\varepsilon,\Gamma}(x) = \chi_{\Gamma}(x)\, e^{-2\pi\langle\varepsilon,x\rangle}$. This is an integrable function (if ε is in the dual of the cone Γ).

$$
\begin{aligned}
\hat{f}_{\varepsilon,\Gamma}(\xi) &= \int_{\Gamma} \exp\left(-2\pi i \langle \xi - i\varepsilon, x\rangle\right) dx \\
&= \sum_{Q} \int_{\Gamma_q} \exp\left(-2\pi i \langle \xi - i\varepsilon, x\rangle\right) dx \\
&= \sum_{Q} \int_{\mathbb{R}^J} \chi_{\left[x:V_q^{-1}x\geq 0\right]} \exp\left(-2\pi i \langle \xi - i\varepsilon, x\rangle\right) dx \\
&= \sum_{Q} \int_{\mathbb{R}^J_+} \exp\left(-2\pi i \langle \xi - i\varepsilon, V_q a\rangle\right) da \\
&= \sum_{Q} \int_{\mathbb{R}^J_+} \exp\left(-2\pi i \left\langle V_q'\xi - iV_q'\varepsilon, a\right\rangle\right) da \\
&= \sum_{q=1..Q} \prod_{j=1..J} \int_{\mathbb{R}_+} \exp\left(-2\pi i \left(v_{qj}'\xi - iv_{qj}'\varepsilon\right) a\right) da \\
&= \sum_{q=1..Q} \prod_{j=1..J} \int_{\mathbb{R}_+} \exp\left(-a\left[2\pi\left(v_{qj}'\varepsilon\right) + 2\pi i\left(v_{qj}'\xi\right)\right]\right) da \\
&= \sum_{q=1..Q} \prod_{j=1..J} \frac{1}{2\pi} \frac{1}{\left[\left(v_{qj}'\varepsilon\right) + i\left(v_{qj}'\xi\right)\right]}
\end{aligned}
$$

Let V_K be the corresponding matrix for K. $\kappa_{qj} = V_K^{-1}v_{qj} \geq 0$ for all $q \in \{1..Q\}$ and

$j \in \{1..J\}$. Consider $\xi = \left(V_K^{-1}\right)' \alpha$,

$$\begin{aligned}
\hat{f}_{\varepsilon,\Gamma}\left(\left(V_K^{-1}\right)'\alpha\right) &= \left(\frac{1}{2\pi i}\right)^J \sum_{q=1..Q} \prod_{j=1..J} \frac{1}{\left[\left(\kappa_{qj}'\alpha\right) - i\left(v_{qj}'\varepsilon\right)\right]} \\
&= \left(\frac{1}{2\pi i}\right)^J \sum_{q=1..Q} \prod_{j=1..J} \frac{\left(\kappa_{qj}'\alpha\right) + \left(v_{qj}'\varepsilon\right) i}{\left[\left(\kappa_{qj}'\alpha\right)^2 + \left(v_{qj}'\varepsilon\right)^2\right]}
\end{aligned}$$

This is an entire function for every $\varepsilon \in \Gamma^o / \{0\}$. Therefore, if it is zero in an open subset of $\mathbb{R}^J$ is zero everywhere. Each term in the summation has a positive denominator and a numerator that is a polynomial function of α with positive coefficients. It follows that there is no open subset of $\mathbb{R}^J$ where $\hat{f}_{\varepsilon,\Gamma}$ is zero. □

4 Welfare Analysis

5 Estimation

6 Extensions

7 Application: Elementary School Admissions in Cambridge

7.1 Data and Description of the Cambridge School System

7.2 The Controlled Choice Plan in Cambridge

References

Abdulkadiroglu, A., Agarwal, N., and Pathak, P. A. (2014). The Welfare Effects of Coordinated School Assignment: Evidence from the NYC High School Match.

Abdulkadiroglu, A., Angrist, J., Dynarski, S. M., Kane, T. J., and Pathak, P. A. (2011a). Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots. *The Quarterly Journal of Economics*, 126(2):699–748.

Abdulkadiroglu, A., Angrist, J., and Pathak, P. A. (2012). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica, forthcoming.*

Abdulkadiroglu, A., Che, Y.-K., and Yasuda, Y. (2011b). Resolving Conflicting Preferences in School Choice: The Boston Mechanism Reconsidered. *American Economic Review*, 101(1):399–410.

Abdulkadiroglu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-Proofness Versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match.

Abdulkadiroglu, A., Pathak, P. A., Roth, A. E., and Sonmez, T. (2006). Changing the Boston School Choice Mechanism. *NBER Working Papers 11963*.

Abdulkadiroglu, A. and Sonmez, T. (2003). School Choice: A Mechanism Design Approach. *The American Economic Review*, 93(3):729 – 747.

Agarwal, N. (2013). An Empirical Model of the Medical Match.

Agarwal, N. and Diamond, W. F. (2014). Identification and Estimation in Two-sided Matching Markets.

Ayaji, K. (2013). School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana.

Azevedo, E. and Budish, E. (2012). Strategyproofness in the large.

Azevedo, E. and Leshno, J. (2013). A Supply and Demand Framework for Two-Sided Matching Markets.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating Dynamic Models of Imperfect Competition. *Econometrica*, 75(5):1331–1370.

Beggs, S., Cardell, S., and Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Econometrics*, 17(1):1–19.

Berry, S. T. and Haile, P. A. (2010). Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers.

Berry, S. T., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841 – 890.

Berry, S. T., Levinsohn, J., and Pakes, A. (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy*, 112(1):68–105.

Cassola, N., Hortaçsu, A., and Kastl, J. (2013). The 2007 Subprime Market Crisis Through the Lens of European Central Bank Auctions for Short Term Funds. *Econometrica*, 81(4):1309–1345.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics*, 126(4):1593–1660.

Chiappori, P.-A., Salanie, B., Salanie, F., and Gandhi, A. (2012). From Aggregate Betting Data to Individual Risk Preferences.

Chiappori, P.-A., Salanié, B., and Weiss, Y. (2011). Partner Choice and the Marital College Premium.

Choo, E. and Siow, A. (2006). Who Marries Whom and Why. *Journal of Political Economy*, 114(1):175–201.

Cullen, J. B., Jacob, B. A., and Levitt, S. (2006). The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, 74(5):1191–1230.

Deming, D. J., Hastings, J. S., Kane, T. J., and Staiger, D. O. (2011). School Choice, School Quality and Postsecondary Attainment.

Dubins, L. and Freedman, D. (1981). Machiavelli and the Gale-Shapley Algorithm. *The American Mathematical Monthly1*, 88(7):485–494.

Featherstone, C. and Niederle, M. (2011). School choice mechanisms under Incomplete Information: an Experimental Investigation.

Fox, J. T. (2010a). Estimating Matching Games with Transfers.

Fox, J. T. (2010b). Identification in matching games. *Quantitative Economics*, 1(2):203–254.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal Nonparametric Estimation of First-Price Auctions. *Econometrica*, 68(3):525–574.

Hastings, J. S., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous Preferences and the Efficacy of Public School Choice.

He, Y. (2012). Gaming the Boston School Choice Mechanism in Beijing. mimeo, Columbia University.

Kojima, F. and Pathak, P. A. (2009). Incentives and Stability in Large Two-Sided Matching Markets. *American Economic Review*, 99(3):608 – 27.

Krueger, A. B. (1999). Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics*, 114(2):497–532.

Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97(1):145–177.

Matzkin, R. L. (1992). Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models. *Econometrica*, 60(2):239–70.

Matzkin, R. L. (1993). Nonparametric identification and estimation of polychotomous choice models. *Journal of Econometrics*, 58(1-2):137–168.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. pages 105 – 142.

Menzel, K. (2012). Inference for large games with exchangeable players.

Miralles, A. (2009). School Choice: The Case for the Boston Mechanism.

Nielson, C. (2013). Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students.

Pathak, P. A. and Shi, P. (2013). Simulating Alternative School Choice Options in Boston.

Pathak, P. A. and Sonmez, T. (2008). Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism. *American Economic Review*, 98:1636–1652.

Pathak, P. A. and Sonmez, T. (2013). School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation. *American Economic Review*, 103(1):80–106.

Roth, A. E. (1982). The Economics of Matching: Stability and Incentives. *Mathematics of Operations Research*, 7(4):617–628.

Roth, A. E. and Peranson, E. (1999). The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review*, 89(4):748–780.

Shapley, L. and Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37.

Walters, C. (2013). A Structural Model of Charter School Choice and Academic Achievement.

A Proofs: Mechanisms

A.1 Proof of Proposition 2

Proof. For a strategy σ^*, a particular realization of the reports of the other agents given by the empirical measure m^{n-1} from $n-1$ iid draws from m_{σ^*} where $m_{\sigma^*}(R,t) = f_T(t) \times \int \sigma^*(u,t;R) dF_{U|T}$. Condition 1 implies that $\phi^n((R_i,t_i),m^{n-1}) \xrightarrow{p} \phi^\infty((R_i,t_i),m_{\sigma^*})$. Fix $\varepsilon > 0$ and pick n_0 such that for all $n > n_0$ and $(R,t) \in \mathcal{R} \times T$,

$$P\left(\|\phi^n((R,t),\hat{m}) - \phi^\infty((R,t),m)\|_\infty > \frac{\varepsilon}{8|S|} \right) < \frac{\varepsilon}{16|S|}.$$

Since $\|\phi^n((R,t),\hat{m}) - \phi^\infty((R,t),m)\|_\infty$ is bounded by 2, we have that

$$\mathbb{E}\left[\|\phi^n((R,t),\hat{m}) - \phi^\infty((R,t),m)\|_\infty \right] < \frac{\varepsilon}{4|S|}.$$

Note that the choice of n_0 did not depend on u_i or t_i.

Now, we show that no agent of type t_i and utility u_i can expect a gain of more than $\varepsilon\|u_i\|_\infty$ by deviating from σ^*. For $n > n_0$ and each (R_i,t_i),

$$\begin{aligned} |V_i^n(R_i,m_\sigma) - V_i^\infty(R_i,m_\sigma)| &\leq \mathbb{E}\left| \sum_j \phi_j^n((R_i,t_i),\hat{m}) u_{ij} - \sum_j \phi_j^\infty((R_i,t_i),m) u_{ij} \right| \\ &\leq 2|S|\|u_i\|_\infty \mathbb{E}\left[\|\phi^n((R_i,t_i),\hat{m}) - \phi^\infty((R_i,t_i),m)\|_\infty \right] \\ &\leq \frac{\varepsilon}{2}\|u_i\|_\infty \end{aligned}$$

Since σ^* is a limit equilibrium, $\sigma^*(u_i,t_i;R_i) > 0$ implies that for all R_i',

$$\begin{aligned} & V_i^\infty(R_i,m_\sigma^*) \geq V_i^\infty(R_i',m_\sigma^*) \\ \Rightarrow \quad & V_i^n(R_i,m_\sigma^*) \geq V_i^n(R_i',m_\sigma^*) - \varepsilon\|u_i\|_\infty \end{aligned}$$

for all $n > n_0$. □

A.2 Proof of Lemma 1

Part 1: Existence and Uniqueness of Cutoffs

Existence follows from Corollary A1 of Azevedo and Leshno (2013).

The uniqueness result is a generalization of Theorem 1 of Azevedo and Leshno (2013) and Proposition 2.2 ??? Azevedo (2011). By the Lattice Theorem (Azevedo and Leshno, 2013), there exists minimum and maximum market clearing cutoffs $p^- \leq p^+$. Note that

the measure of students matched with program j at cutoff p is given by $D_j(p|\eta)$, and the measure of students unmatched is given by $D_0(p|\eta)$. Hence, by the Rural Hospitals theorem (Azevedo and Leshno (2013)), for all $C \subseteq S \cup \{0\}$,

$$\sum_{j \in C} D_j(p^+|\eta) = \sum_{j \in C} D_j(p^-|\eta).$$

Let $C = \{j \in S : p_j^- < p_j^+\}$. By Assumptions 1(ii) and 1(iii) we have that

$$\sum_{j \in S \setminus C} D_j(p^+|\eta) < \sum_{j \in S \setminus C} D_j(p^-|\eta).$$

Therefore, it must be that $C = \emptyset$ i.e. $p^+ = p^-$ and the cutoff vector is unique.

Part 2: Convergence of Cutoffs

Define the class $\mathcal{B} = \{\{(e_i, R_i) : e_{ij} \geq p_j, R_i = R\} : p, j, R\}$. Note that $\mathcal{B}$ is a VC class since it is collection of half-spaces, which are VC classes. Hence, the class of sets

$$\mathcal{V} = \left\{ v_{pj} = \{(e_i, R_i) : e_{ij} \geq p_j,\ jR_i i\} \bigcap_{j' \neq j} (\{(e_i, R_i) : jR_i j'\} \cup \{(e_i, R_i) : e_{ij'} < p_{j'}\}) : p, j \right\}$$

is a VC-class since it is a subset of finite unions and intersections of $\mathcal{B}$. Hence,

$$\sup_P \|D(p|\eta) - D(p|\eta^n)\|_\infty = \sup_{V \in \mathcal{V}} |\eta^n(V) - \eta(V)| \xrightarrow{p} 0.$$

Since $D_j(p|\eta) = \eta(v_{pj})$, $D(p|\eta^n) - q^n \xrightarrow{p} D(p|\eta) - q$ uniformly in p.

Let the unique market clearing cutoff for (η, q) be p^*. For $\delta > 0$, let

$$\delta_z = \inf\{\gamma : \|z(p|\eta, q)\| > \gamma \Rightarrow \|p - p^*\| > \delta\}.$$

Note that $\delta_z > 0$ as p is an element of a compact space, $z(p|\eta, q)$ is continuous in p and p^* is the unique market clearing cutoff. Pick n_0 such that for all $n > n_0$,

$$\mathbb{P}\left(\sup_p \|z(p|\eta^n, q^n) - z(p|\eta, q)\|_\infty > \delta_z\right) < \varepsilon.$$

Since $z(p^n|\eta^n, q^n) \leq 0$, for all $n > n_0$,

$$\mathbb{P}\left(z(p^n|\eta, q) > \delta_z\right) < \varepsilon \Rightarrow \mathbb{P}\left(\|p^n - p^*\| > \delta\right) < \varepsilon$$

where the impliciation follows from set inclusion.

A.3 Proof of Theorem 1

We first introduce a simpler class of mechanisms and prove a lemma.

Definition 10. *A mechanism ϕ^n is a **lottery + cutoff** mechanism if for each profile of reports and scores $(R, e) = ((R_1, e_1), \ldots, (R_n, e_n)) \in \mathcal{R}^n \times ([0,1]^J)^n$, there are market clearing cutoffs $p^n : \mathcal{R}^n \times ([0,1]^J)^N \to [0,1]^J$, such that*

$$\phi^n((R_i, t_i), m(R_{-i}, t_{-i})) = \int \ldots \int D^{(R_i, \mathbf{e}_i)}(p^n) \, \mathrm{d}\eta_{R_1, \mathbf{e}_1 | t_1}(R_1, \cdot) \ldots \mathrm{d}\eta_{R_n, \mathbf{e}_n | t_n}(R_n, \cdot)$$

where η_L is a measure on $[0,1]^J$ satisfying strict preferences.

Lemma 4. *Suppose η satisfies Assumption 1. If ϕ^n is a lottery + cutoff mechanism, then ϕ^n satisfies Condition 1.*

Proof. Since ϕ^n is a lottery + cutoff mechanism, $\phi^n((R_i, t_i), m^{n-1}) = E[D^{(R_i, \mathbf{e}_i)}(p^n) | R_i, t_i, m^{n-1}]$ where expectation is taken with respect to the lottery draws conditional on t. Lemma 1 implies that the market clearing cutoffs $p^n \xrightarrow{p} p$ since m^{n-1} is an empirical measure from m and consequently, η^{n-1} on reports and lotteries is an empirical measure from $n-1$ iid draws from η.

Note that $E[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, m] = E[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, p]$ since the distribution of reports m^{n-1} only affects an agent's demand through p^n. Hence, if $g(p) = E[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, p]$ is continuous in p,

$$\|\phi^n(R, m^{n-1}) - \phi^\infty(R, m)\|_\infty = \| \mathbb{E}[D^{(R_i, \mathbf{e}_i)}(p^n) | R_i, t_i, m^{n-1}] - \mathbb{E}[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, m]\|_\infty \xrightarrow{p} 0$$

by the continuous mapping theorem since $p^n \xrightarrow{p} p$.

To show that $E[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, p]$ is continuous in p, for $\varepsilon > 0$, pick δ such that for all p' with $\|p' - p\| < \delta$ implies $\eta(R_i, \{e_i \in E_{t_i} : p \wedge p' \leq e_i \leq p \vee p'\}) < \varepsilon$ where $E_{t_i} \subseteq [0,1]^J$ is the set of priority scores consistent with t_i. Assumption 1(i) implies that such a δ exists. We now s that for all p' such that $\|p' - p\| < \delta$, $\| \mathbb{E}[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, p] - \mathbb{E}[D^{(R_i, \mathbf{e}_i)}(p') | R_i, t_i, p']\|_\infty < \varepsilon$.

$$\begin{aligned} & \| \mathbb{E}[D^{(R_i, \mathbf{e}_i)}(p) | R_i, t_i, p] - \mathbb{E}[D^{(R_i, \mathbf{e}_i)}(p') | R_i, t_i, p']\|_\infty \\ = & \sup_j \left| \int D_j^{(R_i, \mathbf{e}_i)}(p) \mathrm{d}\eta_{R_i, \mathbf{e}_i | t_i}(R_i, \cdot) - \int D_j^{(R_i, \mathbf{e}_i)}(p) \mathrm{d}\eta_{R_i, \mathbf{e}_i | t_i}(R_i, \cdot) \right| \end{aligned}$$

For each j, we have

$$\begin{aligned}
&\left|\int D_j^{(R_i,\mathbf{e}_i)}(p)\mathrm{d}\eta_{R_i,\mathbf{e}_i|t_i}(R_i,\cdot) - \int D_j^{(R_i,\mathbf{e}_i)}(p)\mathrm{d}\eta_{R_i,\mathbf{e}_i|t_i}(R_i,\cdot)\right| \\
= \;& 1\{jR_ii\}\left|\int\left(1\{\mathbf{e}_{ij} > p_j\}\prod_{k\in\{k:kR_ij\}} 1\{\mathbf{e}_{ik} < p_k\} - 1\{\mathbf{e}_{ij} > p'_j\}\prod_{k\in\{k:kR_ij\}} 1\{\mathbf{e}_{ik} < p'_k\}\right)\mathrm{d}\eta_{R_i,\mathbf{e}_i|t_i}(R_i,\cdot)\right| \\
\leq \;& \eta_{R_i,\mathbf{e}_i|t_i}(R_i,\{e_i \in E_{t_i} : p\wedge p' \leq e \leq p\vee p'\}) < \varepsilon
\end{aligned}$$

□

We now show that Theorem 1 is a Corollary to Lemma 4 by observing that ϕ^n is a lottery + cutoff mechanism. To see this, note that p^n is a market clearing cutoff for the economy $((R_1, f(R_1, e_1)), \ldots, (R_n, f(R_n, e_n)))$ and that

$$\begin{aligned}
\phi_j^n((R_i,t_i),m(R_{-i},t_{-i})) &= \int\ldots\int D^{(R_i,f(R_i,\mathbf{e}_i))}(p^n)\mathrm{d}\eta_{R_1,\mathbf{e}_1|t_1}(R_1,\cdot)\ldots\mathrm{d}\eta_{R_n,\mathbf{e}_n|t_n}(R_n,\cdot) \\
&= \int\ldots\int D^{(R_i,e_i)}(p^n)\mathrm{d}\eta^f_{R_1,\mathbf{e}_1|t_1}(R_1,\cdot)\ldots\mathrm{d}\eta^f_{R_n,\mathbf{e}_n|t_n}(R_n,\cdot)
\end{aligned}$$

A.4 Proof of Proposition 1

Deferred Acceptance:

Let $\underline{e}_j$ be supremum of the priority scores of the rejected students. We claim that $p^n = \underline{e}$ are the cutoffs with the desired properties (if a school does not reject any students, set $p_j = 0$).

Let $\underline{e}_j^r$ be the supremum the priority scores of students that were rejected in round r. Observe that for each school, $\underline{e}_j^r \leq \underline{e}_j^{r+1}$. If the algorithm terminates in round k, then $\underline{e}_j^k = \underline{e}_j$.

Assume that student i is assigned to school j' and consider any school j with jR_jj'. Let r be round in which student i was rejected by j. By definition, it must be that $e_{ij} < \underline{e}_j^r$. Therefore, $e_{ij} < \underline{e}_j$ and we have that each student is assigned to $D^{(R_i,,e_i)}(p^n)$.

Finally, the aggregate demand must not exceed q_j by construction of p^n.

Boston Mechanism:

We show that the Boston Mechanism is report-specific priority + cutoff mechanims for

$$f_j(R,e) = \frac{e_j - \#\{k : kR_ij\}}{J} + 1$$

by constructing market cutoffs p^n for each profile $((R_1,e_1),\ldots,(R_N,e_N))$ such that (i) the assignment of each agent is given by $D^{(R_i,f(R_i,\mathbf{e}_i))}(p^n)$ and (ii) p^n clears the market for the economy $((R_1, f(R_1, e_1)), \ldots, (R_N, f(R_N, e_N)))$.

Note that if a school rejects a student in round k, then it rejects students in all further rounds since it is full at the end of that round. Let k_j denote that round, and let $\underline{e}_j$ be supremum of the lotteries of the rejected students in round k_j. We claim that $p_j^n = \frac{\underline{e}_j - k_j}{J} + 1$ are the cutoffs with the desired properties (if a school does not reject any students, set $k_j = J$ and $p_j = 0$).

We first show that the assignment of each student in the Boston mechanism is given by $D^{(R_i, f(R_i, e_i))}(p^n)$. Assume that student i is assigned to school j' and consider any school j with jR_ij'. Since jR_ij', it must be that the student was rejected at j, and could not have applied to j before round k_j. If student applied to k_j after round j, then $e_{ij} - \#\{k : kR_ij\} < \underline{e}_j - k_j$ since $|e_{ij} - \underline{e}_j| \leq 1$. If $\#\{k : kR_ij\} = k_j$, then $e_{ij} < \underline{e}_j$. In either case, $f_j(R_i, e_i) < p_j$. Therefore, the student is assigned to $D^{(R_i, f(R_i, e_i))}(p^n)$.

Next, we show that p^n clears the market for economy $((R_1, f(R_1, e_1)), \ldots, (R_N, f(R_N, e_N)))$. As noted earlier, each agent is assigned to $D^{(R_i, f(R_i, e_i))}(p^n)$. By construction of p^n, the aggregate demand must be less than q_j, and $p_j^n = 0$ if aggregate demand is strictly less than q_j.

B Proofs: Identification