Identification and Estimation in Manipulable Assignment Mechanisms*

Nikhil Agarwal[†] Paulo Somaini[‡]

June 2, 2014

PRELIMINARY AND INCOMPLETE

Abstract

This paper studies estimation and identification of preferences 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. Consistent estimates of these probabilities can be obtained for a large class of mechanisms, which we call report-specific priority + cutoff mechanisms. This class includes the Boston Mechanism and the Deferred Acceptance mechanism. 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). We show that this equilibrium assumption is testable using the available data. Preferences are non-parametrically identified under either sufficient variation in choice environments or sufficient variation in a special regressor. Finally, we illustrate our techniques using data from elementary school admissions in Cambridge, MA.

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 and Phil Haile for several helpful discussions, and Victor Chernozhukov, William Diamond, Liran Einav, Whitney Newey, Ariel Pakes, Larry Samuelson and Joe Shapiro for their comments. 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, a reported ranking of various school options. These data therefore provide a promising opportunity to estimate student preferences to evaluate economic questions. Recent empirical studies have used 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. Table 1 presents a partial list of mechanisms in use at school districts around the world. To our knowlege, only Boston and New Orleans currently employ mechanisms that are not manipulable.[1]

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).[2] The assumption is usually motivated by properties of the mechanism or by arguing that strategic behavior may be limited under a sudden change in the choice environment. This standard approach may not be valid if students have a strategic incentives to manipulate their reports. Anecdotal evidence from Boston suggests that parent groups and forums for exchanging information about the competitiveness of various schools and discussing ranking strategies are fairly active (Pathak and Sonmez, 2008).

Our approach is based on interpreting a student's choice of a report as a choice of a probability distribution over assignments. Intuitively, each rank-order list results can be associated with a probability of getting assigned to each of the schools on that list. The probability of assignment associated with each report depends on the student's priority type, the reports of the other students and a random lottery in the mechanism, if there is one. If agents have

[1] The student proposing deferred acceptance mechanism is a commonly used mechanism that is strategy-proof if students are not restricted to list fewer schools than are available. However, with the exception of Boston since 2005, all implementations of the mechanism known to us, severely restrict the length of the rank-order list. Abdulkadiroglu et al. (2009) and Haeringer and Klijn (2009) show that with this restriction, the mechanism provides incentives for students to drop competitive schools from their rank-order list.

[2] He (2012) is a notable exception that allows for agents to be strategic. We compare our results with this paper in further detail below.

Table 1: School Choice Mechanisms

Mechanism	Manipulable	Examples
Boston Mechanism	Y	Barcelona[1], Beijing[2], Boston (pre 2005), Charlotte-Macklenberg[3], Chicago (pre 2009), Denver, Miami-Dade, Minneapolis, Seattle (pre 1999 and post 2009), Tampa-St. Petersburg.
Deferred Acceptance		
w/ Truncated Lists	Y	New York City[4], Ghanian Schools, various districts in England (since mid '00s)
w/ Unrestricted Lists	N	Boston (post 2005), Seattle (1999-2008)
Serial Dictatorships		
w/ Truncated Lists	Y	Chicago (2009 onwards)
First Priority First	Y	various districts in England (before mid '00s)
Chinese Parallel	Y	Shanghai and several other Chinese provinces[5]
Cambridge	Y	Cambridge[6]
Pan London Admissions	Y	London[7]
Top Trading Cycles	N	New Orleans[8]

Notes: Source Table 1, Pathak and Sonmez (2008) unless otherwise stated. See several references therein for details. Other sources: [1] Calsamiglia and Guell (2014); [2]He (2012); [3]Hastings et al. (2009); [4]Abdulkadiroglu et al. (2009); [5]Chen and Kesten (2013); [6] "Controlled Choice Plan" CPS, December 18, 2001; [7]Pennell et al. (2006);
[8]`http://www.nola.com/education/index.ssf/2012/05/new_orleans_schools_say_new_pu.html` accessed May 20, 2014.

correct beliefs about this probability and are expected utility maximizers, then the chosen report then reveals comparisons of expected utilities with other reports the agent could have chosen. Formally, we require that student behaviour is described by a limit Bayesian Nash Equilibrium. This equilibrium is an approximate equilibrium in a large matching mechanism. We show that this maintained assumption is testable without first estimating the preference distribution. This promises a chance to verify the primary maintained assumption, but also posits potential barriers to using our equilibrium assumption as a basis for estimation in empirical settings. We show, however, that if the mechanism uses rank-order lists from agents and satisfies a stict monotonicity condition in probability of assignments with respect to upgrading a school in the ranks, then any rank-order list with positive probability of assignment to each of the options is consistent with equilibrium play. While negative on the ability to test the equilibrium assumption, the result suggests that our method can be used in many empirical settings for estimating preferences in an internally consistent manner.

In order to learn about preferences from the reports made by an agent, we first estimate the probabilities of assignments associated with each report and student priority type. We present a general convergence condition on the mechanism under which data from a large market can be used to consistently estimate these probabilities without directly estimating preferences or solving for an equilibrium. The ability to do this circumvents difficulties that may arise due to computational difficulties in solving for an equilibrium or multiplicity of equilibria. We describe a new class of mechanisms, called report-specific priority + cutoff mechanisms for which this condition is satisfied. We can show that all mechanisms in Table 1, except the Top Trading Cycles mechanism, are elements of this class.

Since the assignment probabilities as a function of reports and priority types can be consistently estimated, we study identification of preferences treating these probabilities as known to the econometrician. The problem is equivalent to identifying the distribution over preferences over discrete objects with choice data on lotteries over these objects. We follow the discrete choice literature in specifying preferences with a random utility model that allows for student and school unobservables (see Block and Marshak, 1960; McFadden, 1973; Manski, 1977). Using this model, we show conditions under which the distribution of preferences is non-parametrically identified.

We exploit two types of variation to identify the distribution of preferences. First, we use variation in choice environments (as defined by the lotteries available to the agents). Such variation may arise from differences in agent priorities that are excludable from preferences, or if the researcher observed data from two identical populations of agents facing different mechanisms or availability of seats. We characterize the identified set of preference distributions under such variation. Although sufficient variation in choice environments can point identify the preference distribution, we should typically expect set identification. Our second set of identification results relies on the availability of a special regressor that is additively separable in the indirect utility function (Lewbel, 2000). The assumption is commonly made to identify preferences in discrete choice models (Berry and Haile, 2010, 2014, for example). We show that local variation in this regressor can be used to identify the density of distribution of utility in a corresponding region. A special regressor with full support can be used to identify the full distribution of preferences.

Finally, we apply our methods to data from the controlled choice plan for admission into elementary schools in Cambridge, MA. The school district uses a variant of the Boston Mechanism, that is highly manipulable.

Related Literature

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. Indeed, we borrow the concept a limit-equilibrium used in Azevedo and Budish (2012).

The empirical approach of assuming that agent behaviour is described by a Bayesian Nash Equilibrium is similar to He (2012). The paper studies strategic behavior in the Boston mechanism in Beijing, and estimates preferences under the assumption that agents' reports are undominated. The set of undominated reports is derived using a limited number of restrictions implied by rationality, the specific number of schools and ranks that can be submitted in Beijing, and that the mechanism treats all agents symmetrically. The approach fully-specifies the likelihood of reporting each of the undominated stategies.

Compared to He (2012), our approach allows for a more 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).

Previous research has questioned the extent to which agents are sophisticated. For example, Abdulkadiroglu et al. (2006) use particulars of the Boston mechanism to deduce reports that are clearly suboptimal and tabulate the fraction of agents that make one these reports. Recent evidence in Calsamiglia and Guell (2014) suggests that students in Barcelona respondes to a change in strategic incentives when the system of assigning neighborhood priorities was administratively changed. We present a sharp condition for an agent's report to be consistent with equilibrium behaviour that does not depend on details of the mechanism. This allows us to estimate the fraction of agents with reports that are not consistent with equilbrium behaviour. It also shows that the equilibrium restriction we use in our approach is testable in the data. Extensions to relax this assumption are left for future research.

We use techniques and build 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. Since choices over lotteries depend on expected utilities, our data contain direct information on cardinal utilities when the lotteries are not degenerate. 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.

This paper is related to the large, primarily theoretical, literature that has taken a mechanism design approach to the student assignment problem (Gale and Shapley, 1962; Shapley and Scarf, 1974; Abdulkadiroglu and Sonmez, 2003). Theoretical results from this literature has been use to guide redesigns of matching markets (Roth and Peranson, 1999; Abdulkadiroglu et al., 2006, 2009). While preferences are fundamental primitives that influence mechanism comparisons, prospective analysis of a proposed change in the school choice mechanism is rare (see Pathak and Shi, 2013, for an exception). A significant barrier 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). Results in this paper may be useful to allow such analysis in some cases. For instance, our techniques will allow comparing the welfare effects of a change to the Deferred Acceptance mechanism for a school district that uses the Boston mechanism. The relative benefits of these two mechanisms has been debated in the theoretical literature using stylized theoretical models with an assumed distribution of preferences (Miralles, 2009; Abdulkadiroglu et al., 2011; Featherstone and Niederle, 2011).

Our methods may also be useful in extending recent work on measuring the effects of school assignment on student achievement that jointly specifies the preferences for schools and 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, a researcher must confront the possibility that students differentially sort into schools that result in idiosyncratic achievement gains. Additionally, estimates of preferences may be useful in extrapolating lottery based achievement designs if there is selection on the types of students that participate in the lottery (Walters, 2013). Methods for estimating preferences in a broader class of data-environments may expand our ability to study the effects of school assignment on student achievement.

This paper also contributes to the growing literature on methods for analyzing preferences in matching markets. 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.

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.

Overview

Section 2 sets up the notation for describing mechanisms and presents the main convergence condition needed for our analysis. It then presents the class of report-specific priority + cutoff mechanisms and describes the equilibrium concept. Section 3 sets up the preference model, interprets agent choices as a choice over lotteries, and studies identification under varying choice environments and the availability of a special regressor. Section 4 shows a basic consistency theorem for the preference parameters. Section 5 describes the Cambridge data, the Cambridge mechanism and shows that our main convergence condition holds for the Cambridge mechanism. Section 6 concludes.

2 Mechanisms and Limit Behaviour

2.1 Mechanisms and Their Limits

Consider a single-unit assignment mechanism 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 participate in a mechanism by submitting a report $R_i \in \mathcal{R}_i$, which is a rank-order list over the schools $S = \{0, \ldots, J\}$. As is the convention in the school choice literature, write jR_ij' to indicate that j is ranked above 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 often described as 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 $t \in T^n$,

$$\sum_{i=1}^{n} \Phi^n_{ij}(R, t) \leq q^n_j.$$

In this notation, the $i - j$ component of $\Phi^n(R)$, denoted $\Phi^n_{ij}(R, t)$ is the probability that student i is assigned to school j. Hence, the outcome for each student is in the J-simplex ΔS. In the Cambridge school system, there is a lottery that breaks ties between students. Such lotteries are a common 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 $\Phi^n_i((R_i, R_{-i}), t) = \Phi^n_{i'}((R_i, R_{-i}), t)$.*
2. *for all R_i, R_{-i} and permutations π of $-i = \{1, \ldots, i-1, i+1, \ldots, n\}$, we have that $\Phi^n_i((R_i, t_i), R_{-i}, t_{-i}) = \Phi^n_i((R_i, t_i), R_{\pi(-i)}, t_{\pi(-i)})$.*

The restriction to semi-anonymous mechanisms with finitely many priority types rule 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. In a semi-anonymous mechanism, 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 remark highlights that this perspective is without further loss of generality.

Remark 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*

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

Note that Φ^n_i 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 this formulation, 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, t_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 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. The key property that will allow us to proceed with the analysis for a mechanism is that outcomes of the mechanism evaluated at the empirical distribution of the reports converge in probability to the limiting values as the market grows in size.

Condition 1 (Convergence)**.** *Suppose the sequence of empirical measures* m^{n-1} *on* $\mathcal{R} \times T$ *converges in probability to the population measure* $m \in \mathcal{M}$*, then for each* (R, t)*,*

$$\phi^n((R, t), m^{n-1}) \xrightarrow{p} \phi^\infty((R, t), m).$$

Verifying Condition 1 for a mechanism may not be straightforward because matching mechanisms are often described using algorithms instead of functions that take a measure of reports as inputs. A representation of the mechanism as a function may be necessary before proceeding. The next subsection describes a large class of mechanisms in which the condition is satisfied. Researchers interested in applying our results will need to either verify that their mechanism belongs to this class or will need to verify Condition 1 directly.

2.2 Report-Specific Priority and Cutoff Mechanisms

This section introduces a class of mechanisms called report-specific priorities + cutoff mechanisms and shows that all mechanisms that belong to this class satisfy Condition 1. This representation extends the characterization of stable matchings by Azevedo and Leshno (2013) in terms of demand-supply and market clearing to discuss mechanisms. Particularly, we can use the framework to consider mechanisms that produce matchings that are not stable.

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

$$D_j^{(R_i,e_i)} = 1\{e_{ij} \geq p_j,\ jR_i0\} \prod_{j' \neq j} 1\{jR_ij' \text{ or } e_{ij'} < p_{j'}\}.$$

We can now write the aggregate demand by integrating over the measure of student priorities and reports η. For a cutoff vector $p \in [0,1]^J$, the measure of students demanding $j \in S$ is given by

$$D_j(p) = \eta\left(\{(e_i,R_i) : e_{ij} \geq p_j,\ jR_i0\} \bigcap_{j' \neq j} (\{(e_i,R_i) : jR_ij'\} \cup \{(e_i,R_i) : e_{ij'} < p_{j'}\})\right)$$

and the measure of student demanding no school is given by

$$D_0(p) = \eta\left(\bigcap_{j} (\{(e_i,R_i) : 0R_ij\} \cup \{(e_i,R_i) : e_{ij'} < p_{j'}\})\right).$$

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

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

Non-degenerate lotteries is a strengthening of strict preferences in Azevedo and Leshno (2013). The assumption is straightforward to verify with knowledge of the mechanism. For example, it is satisfied if a lottery is used to break ties between multiple students with the same priority type. It also allows for a siutation in which a single tie-breaking lottery that is used by all schools to break ties as well as cases when each school uses an independent lottery. This assumption, however, is not satisfied if the school destrict uses an exam with finitely many possible scores to determine eligibility and does not use a lottery to break ties between students with identical exam scores.

Given an aggregate demand implied by η and school capacities q, we can consider the set of cutoffs that clear the market as follows:

Definition 4 (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 $p_j > 0$.*

We require that the limit economy has a unique market clearing cutoff.

Assumption 2 (Unique Cutoff). *(η, q) admits a unique market clearing cutoff, p^*.*

The assumption restricts the joint distribution of reports and priorities, and the school capacities. Existence of a market clearing cutoff is guaranteed by Corollary A1 of Azevedo and Leshno (2013) for any η. Uniqueness is a restriction on an equilibrium object. Although the assumption is not made on primitives, it is a restriction on features that are observed in the data. Sufficient conditions that imply this assumption are therefore testable in principle. Further, as our next example illustrates, violations of this assumption are knife-edge cases. We refer the reader to Appendix A.2 for a more formal discussion of sufficient conditions for Assumption 2 and conditions under which violations are non-generic. This discussion borrows from results in Azevedo and Leshno (2013) and Berry et al. (2013).

Example 1. *Consider an economy with two schools, a and b, with capacities $q_a = q_b = \frac{1}{2}$. There are two types of students α and β, each with mass $\frac{1}{2}$. Students of type α have priority at school a and types β have priority at school b. Student reports are oppositing so that α-types report school b as preferred to a while students of type β report the reverse. Assume that there is a single a uniformly distributed lottery that is used to break ties.*

Figure 1 illustrates the measure η. Types α have priorities uniformly distributed on the bottom-right diagonal and β have priorities uniformly distributed on the top-left square.

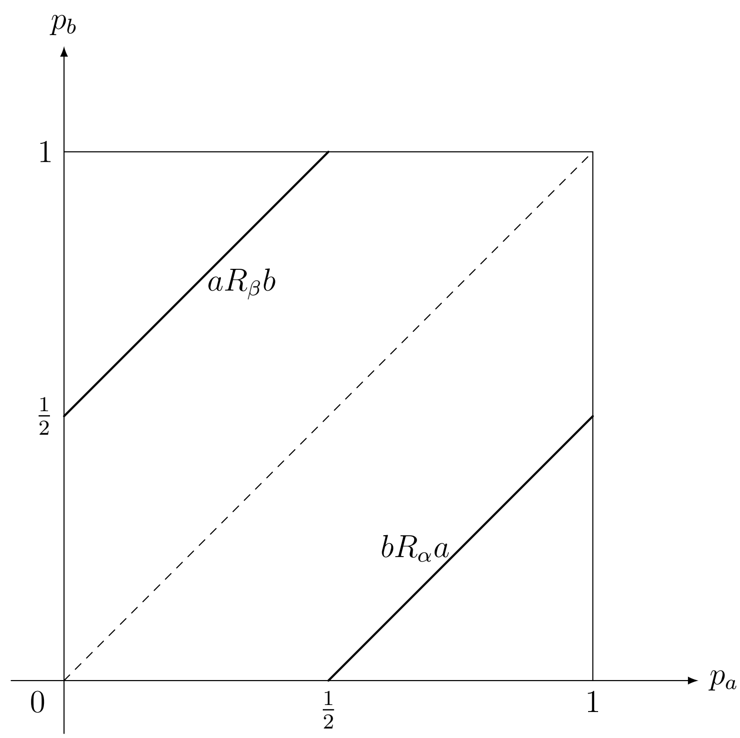


Figure 1: A non-generic example with non-unique cutoffs

The aggregate demand in this economy is given by

$$\begin{aligned} D_j(p_j, p_{-j}) &= \begin{cases} 1 - p_j & \text{if } p_{-j} \geq \frac{1}{2} \\ \frac{1}{2} - p_j + p_{-j} & \text{if } p_{-j} \leq \frac{1}{2}, p_j \leq \frac{1}{2} + p_{-j} \\ 0 & \text{if } p_{-j} \leq \frac{1}{2}, p_j > \frac{1}{2} + p_{-j} \end{cases} \\ D_0(p_a, p_b) &= 1 - \sum_{j \in \{a,b\}} D_j(p_a, p_b) \end{aligned}$$

It is easy to show that $(0, 0)$ *and* $(\frac{1}{2}, \frac{1}{2})$ *are the market clearing cutoffs in this economy.*

First, observe that this multiplicity is non-generic in q*. If* $q_a, q_b \neq \frac{1}{2}$*, then the market clearing cutoffs are unique and given by*

$$(p_j^*, p_{-j}^*) = \begin{cases} (0, 0) & \text{if } q_j, q_{-j} > \frac{1}{2} \\ (1 - q_j, 1 - q_{-j}) & \text{if } q_j, q_{-j} < \frac{1}{2} \\ (\frac{1}{2} - q_j, 0) & \text{if } q_j < \frac{1}{2}, q_j + q_{-j} > 1 \\ (1 - q_j - q_{-j} + \frac{1}{2}, 1 - q_{-j}) & \text{if } q_j < \frac{1}{2}, q_j + q_{-j} < 1 \end{cases}$$

Second, the demand function in this example is also pathological. This is because the demand for the outside good is 0 for all prices $p_j, p_{-j} \leq \frac{1}{2}$*. Hence, the demand for the outside good does not respond to small increases in* p_a *or* p_b *from the cutoff* $(0, 0)$*. Consider a perturbed economy in which* ε *weight is (uniformly) placed on priority types along the dashed line where the* $\frac{1}{2}\varepsilon$ *list only school* a *as acceptable and* $\frac{1}{2}\varepsilon$ *list only school* b *as acceptable. With this modification, the demand function is given by*

$$D_j^\varepsilon(p_j, p_{-j}) = (1 - \varepsilon) D_j(p_j, p_{-j}) + \frac{1}{2}\varepsilon(1 - p_j)$$

It is straightforward to show that if $q_a = q_b = \frac{1}{2}$*, then for any* $\varepsilon \in (0, 1)$*, the only market clearing cutoffs are* $(0, 0)$*.*

Both, the perturbations of capacities and the demand function described above show that the cases in which the market cutoffs are not unique are ones in which the aggregate demand is not-responsive to local changes in the cutoffs. Therefore, ruling out multiplicity of cutoffs is similar to ruling out singularities at the market clearing cutoffs.

We now show that if the the limit economy satisfies Assumptions 1 and 2, then the cutoff is approximated by the market clearing cuttofs of finite economies.

Lemma 1. *Suppose* (η, q) *satisfies Assumption 2. If* p^n *is a sequence of market clearing cutoffs for* (η^n, q^n)*, where* η^n *are a sequence of empirical measures that converges in probability to* η *and* $q^n \to q$*, then* $\|p^n - p^*\| \xrightarrow{p} 0$*.*

Proof. See Appendix A.2 □

The proof is based on first showing 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 and the school j). Combined with a unique cutoff in the limit, a basic consistency theorem can be used to show uniform convergence of p^n to p^*.

We are now ready to present the main definition and results in this section.

Definition 5. *A mechanism ϕ^n is a **report-specific priorities + cutoff** mechanism a there exists a function $f : \mathcal{R} \times \times [0,1]^J \to [0,1]^J$ such that*

(i) f strictly increasing in the last J arguments

(ii) $\phi^n_j((R_i, t_i), m(R_{-i}, t_{-i}))$ is given by

$$\int \cdots \int D^{(R_i, f(R_i, \mathbf{e}_i))}(p^n) \mathrm{d}\eta_{R_1, \mathbf{e}_1|t_1}(R_1, \cdot) \ldots \mathrm{d}\eta_{R_n, \mathbf{e}_n|t_n}(R_n \cdot)$$

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

$$\{(R_1, f(R_1, e_1)), \ldots, ((R_n, t_n, f(R_n, e_n))\}$$

This class of mechanisms use a 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. The representation highlights two ways in which these mechanisms can be manipulable. First, the report of an agent modifies can modify her eligibility. Fixing a cutoff, agents may have the direct incentive to make reports that may not be truthful. Second, even if eligibility does not depend on the report, an agent may (correctly) believe that the cutoff for a school will be high, making it unlikely that she will be eligible. If the rank-order list is constrained in length, she may choose to omit certain competitive schools.

Our main results show that this class of mechanism satisfy the key convergence condition needed to proceed with the rest of our analysis, and that this class contains the most commonly used mechanisms.

Theorem 1. *Assume that $\eta^f(R, \{e : e \leq p\}) = \eta(R, \{e : f(R, e) \leq p\})$ satisfies Assumption 1 and 2. If ϕ^n is a report-specific priority + cutoff mechanism, then ϕ^n satisfies Condition 1*

Proof. See Appendix A.3. □

The proof is a straightforward consequence of Lemma 1. Given that η^f satisfies Assumption 2, the market clearing cutoffs of the finite economy converges to the limit. Assumption 1 is used to ensure that probability that a student with priority t_i and report R_i is assigned to a given choice is a continuous function of the cutoff in the limit economy. Together, these claims imply the result.

We conclude this section by showing that most commonly used mechanisms can be expressed as report-specific lottery + cutoff mechanisms. The main text focusses on the two most 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 (see Abdulkadiroglu et al., 2009; Haeringer and Klijn, 2009, for details).

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 $f_j(R,e) = \dfrac{e_j - \#\{k : kRj\}}{J} + \dfrac{J-1}{J}$ for the Boston Mechanism. This choice of f for Boston upgrades the priority of the student at her first choice relative to all students that list that school lower. $\square$

Remark 2. *The Serial Dictatorship, First Priority First, Chinese Parallel Mechanism and the Pan London Admissions scheme are also report-specific priority + cutoff mechanisms. For completeness, we discuss these mechanisms in Appendix A.4.*

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.

2.3 Limit Equilibrium

Our approach assumes that the students submit rank-ordered lists according to a Bayesian-Nash Equilibrium (of the limit mechanism). We will present a test of this assumption in the next section. Also note that the previous results including the limit properties of assignment probabilities do not rely on having a distribution of reports that is generated by the limit equilibrium.

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 $v_i = (v_{i1}, \ldots, v_{iJ})$ and t as $\sigma_i(v_i, t_i; R)$. We will only consider symmetric equilibria, i.e. $\sigma_i^*(v_i, t_i; R) = \sigma^*(v_i, t_i; R)$ 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 ex-ante payoff from report R_i is given by

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

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

We will consider a limit Bayesian-Nash Equilibrium (BNE) and the set of approximate equilibria in the finite mechanisms.

Definition 6. *The strategy σ^* is a **limit equilibrium** if for all i, $\sigma^*(v_i, t_i; R_i) > 0$ implies* *that*

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

*The strategy σ^* is an ε-**equilibrium** of the n-th game if $\sigma^*(v_i, t_i; R_i) > 0$ implies that for all $R_i' \in \mathcal{R}_i$*

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

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 agents do not have a large incentive to deviate from the prescribed limit equilibrium strategy. Our next result shows that if the mechanism satisfied Condition 1, then the limit equilibria are approximate equilibria of a large mechanism.

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, unlike the equilibria of finite market mechanisms, the limit-equilibrium does not depend on n. Without this simplification, a technical challenge is to show that reports sampled from equilibria that depend on n converge in an appropriate sense (see Menzel, 2012, for example).

The 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). Anecdotal evidence suggests that parent groups and forums discussing ranking strategies are active (Pathak and Sonmez, 2008). While direct evidence showing that agents play equilibrium stategies is limited, Calsamiglia and Guell (2014) observe a strategic response in the distribution of reports to a change in the allocation of neighborhood priorities. 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).

3 Identification and Testable Restrictions

3.1 Preference Model and Choice Over Lotteries

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 $S_b = \{0, 1, \dots, J_b\}$ 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 $z_{ijb} = (z_{i1b}, \dots, z_{iJ_b b})$.

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. Let

$$v_i = (v_{i1}, \dots, v_{iJ_b})$$

be the random vector of indirect utilities for student i and denote its joint distribution with $F_V(v_{i1}, \dots, v_{iJ_b}|\xi_b, z_{ib})$. We normalize the utility of not being assigned, v_{i0}, to zero and make the following assumption of F_V:

Assumption 3. $F_V(v_{i1}, \dots, v_{iJ_b}|\xi_b, z_{ib})$ *admits a density* $f_V(v_{i1}, \dots, v_{iJ_b}|\xi_b, z_{ib})$.

The assumption implies 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 $F_V(v_{i1}, \dots, v_{iJ_b}|\xi_b, z_{ib})$ using data from either a single or multiple large markets. Formally, a market is defined by the tuple

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

where m_b is the joint measure on the space of priority and reports of the other students. In our notation, a market conditions on students on with the same observables z. This treatment may be somewhat counterintuitive as it treats two students with different observables in the same year of the data as parts of separate markets. However, the notation allows for an explicity discussion of what must be held fixed, particularly when pooling data from different markets. Conditioning on the observable quantities z, t, m and ϕ^∞ is without loss because a researcher may always do so, but holding ξ fixed may require us to assume that school unobservables are fixed across various contexts.

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. We assume that the probability that lottery L is chosen in market Γ, denoted $P(L \in \mathcal{L}_\Gamma | \Gamma)$, is observed. 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.

3.2 Equilibrium Behavior and Testable Restrictions

Our empirical methods will be based on the assumption that agent behavior is described by equilibrium play. This section discusses whether this assumption is testable in principle and types of mechanisms for which it may be rejected.

Assumption 4. *The map $\sigma_i(v_i, t_i) \to \Delta\mathcal{R}_i$ that generates the data is a symmetric limit Bayesian Nash Equilibrium.*

This assumption implies that students have consistent beliefs of the probability that they are assigned to each school in S_b as a function of their report $R \in \mathcal{R}$. Further, 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. Therefore, a student's choice set can be treated as known to the econometrician. This reformulation therefore transforms the problem of an student playing against a distribution of other students to a single agent problem choosing from a known set of options.

A 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:

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

This observation immediately yields the result that agents maximize their utility by picking lotteries that are extremal in the set of lotteries.

Proposition 3. *Let the distribution of indirect utilities satisfy Assumption 3. If L is not an extreme point of the convex hull of $\mathcal{L}$, the set of utilities v such that $v \cdot L \geq v \cdot L'$ for all $L' \in \mathcal{L}$ has measure zero.*

Proof. If L is not an extreme point of the convex hull of $\mathcal{L}$, then $N_{\mathcal{L}}(L)$ has Lebesgue-measure zero. Assumption 3 implies therefore implies that $\int 1\{v \in N_{\mathcal{L}}(L)\}\mathrm{d}F_V = 0$. $\square$

Since ties are non-generic, agents who's behavior is consistent with limit-BNE play (typically) pick extremal lotteries. Figure 2 presents an example with 2 schools and 8 lotteries. The lotteries can be represented by elements of the 2-simplex. Lotteries in the interior of the convex hull are suboptimal, i.e., almost all students will find another lottery that yields higher expected utility. The normal cone to a lottery in the boundary is set of utility vectors for which the lottery is optimal.

Proposition 3 also indicates that the fraction of students with behaviour that is not consistent with equilibrium play can be identified. This suggests that Assumption 4 is testable. This ability promises a chance to validate this strong restriction on agent behavior as well as answer a question of independent interest. However, we have not yet exploited the structure of assignment probabilities that result from typical assignment mechanisms. We now present a general sufficient condition under which observed behavior can be rationalized as equilibrium play.

Consider a **ranking mechanism** in which reports correspond to rank-orders over the available options. Therefore, a report is a function $R : \{1, \ldots, K\} \to S \cup \{0\}$ such that (i) for all $k, k' \in \{1, \ldots, K\}$, $R(k) = R(k') \neq 0 \Rightarrow k = k'$ and (ii) $R(k) = 0 \implies R(k') = 0$ if $k' > k$. Let $\mathcal{R}$ be the space of such functions.

Definition 7. *The ranking mechanism ϕ^∞ is **rank-monotonic** for t at m if*

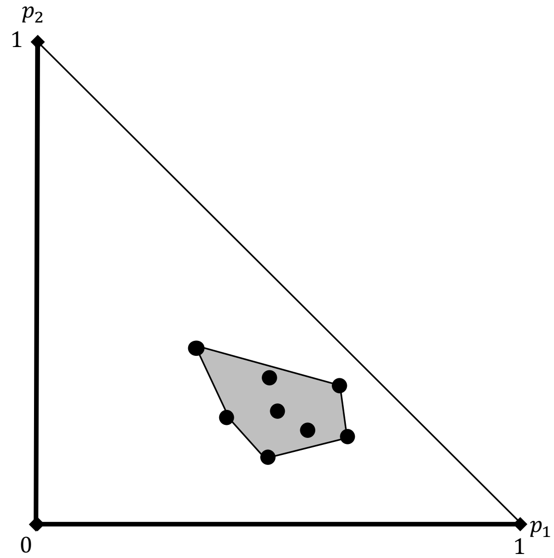
(i) For all $R \in \mathcal{R}$ and $j \in S$, $\phi_j^\infty((R, t), m) = 0$ if $j \notin Im(R)$

(ii) For all $k \in \{1, \ldots, K\}$, if $R(k') = R'(k')$ for all $k' < k$ then

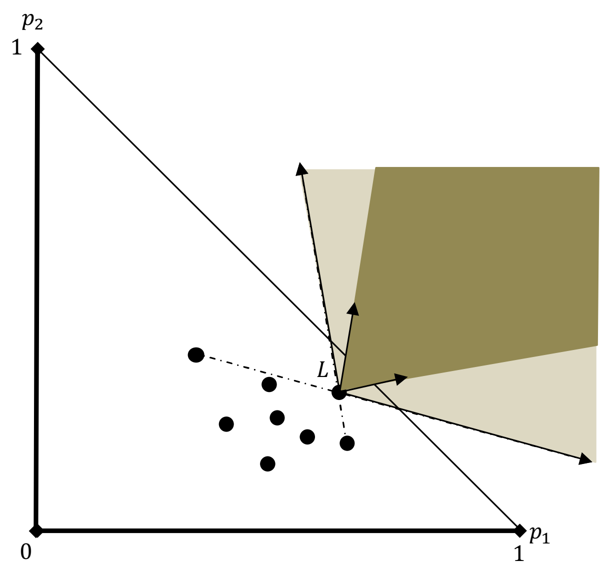
$$\phi^\infty_{R(k)}((R, t), m) \geq \phi^\infty_{R(k)}((R', t), m).$$

*Further, ϕ^∞ is **strictly rank-monotonic** for t at m if the last inequality is strict iff $R(k) \neq R'(k)$ and $\phi^\infty_{R(k)}((R, t), m) > 0$.*

Rank-monotonicity is a natural condition that is often satisfied by single-unit assignment mechanism. Part (i) requires that a student is not assigned to a school which she did not rank, a condition that is satisfied by all mechanisms that we know of. Part (ii) requires the mechanism to obey certain intuitive properties in how rankings correspond to assignment probabilities. Specifically, it requires that assignment at the k-th ranked school does not depend on schools ranked below it, and that ranking a school higher increases a student's chances of getting assigned to it.



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

We now show that in all strictly rank-monotonic ranking mechanisms, all agents that pick a report that gives them a positive probability of assignment at each of their options are behaving in a manner consistent with a limit equilibrium.

Theorem 2. *Assume that the ranking mechanism ϕ^∞ is strictly rank-monotonic at m for priority type t. The report $R \in \mathcal{R}$ corresponds to an extremal lottery $L_R \in \{\phi^\infty((R,t),m) : R \in \mathcal{R}\}$ if $\phi_j^\infty((R,t),m) > 0$ for all $j \in Im(R)$.*

Proof. See Appendix B.1. □

There are two ways to interpret this result. On the one hand, it indicates that our ability to test Assumption 4 is restricted to special cases where we have degenerate mechanisms or when agents rank schools where they have zero chances of getting accepted. On the other hand, this result also indicates that it is quite likely that we can rationalize the behavior of most agents as optimal. While negative on our ability to test equilibrium behavior, the result is positive from an applied perspective that is interested in a method for estimating preferences while viewing the equilibrium assumption as an approximation. An inability to rationalize agent behavior as consistent with Assumption 4 would result in barriers to proceeding with using the assumption as a basis for estimation.

In what follows, we assume that agent behavior is consistent with Assumption 4. For simplicity of exposition, we assume that all lotteries in $\mathcal{L}$ are extremal. With the notation developed in this section, the probability of choosing lottery L from set $\mathcal{L}_\Gamma$ is given by:

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

3.3 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

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

In this section, we will consider results that fix (ξ, z) and therefore drop this from the notation. As a reminder, conditioning on z is without loss since it is observed, but this implies that the researcher assumes that the variation considered holds school unobservables ξ fixed.

The next result characterizes what can be learned about $F_V(v)$ from observing data from several large markets in $\mathcal{T}$. Let $\mathcal{N} = \{\text{int}(N_{\mathcal{L}_\Gamma}(L))\}_{\Gamma\in\mathcal{T},L\in\mathcal{L}_\Gamma}$ be the collection of (the interiors of) normal cones to lotteries faced by agents in the markets $\mathcal{T}$. 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. $\mathbb{R}^J \in \mathcal{D}(\mathcal{N})$ and $\mathcal{N} \subset \mathcal{D}(\mathcal{N})$
2. For all $N \in \mathcal{D}(\mathcal{N})$, $N^c \in \mathcal{D}(\mathcal{N})$
3. For all countable sequences of sets $N_k \in \mathcal{D}(\mathcal{N})$ 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 $\Gamma \in \mathcal{T}$ and $L \in \mathcal{L}_\Gamma$, the quantity*

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

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

Proof. See Appendix B.2. □

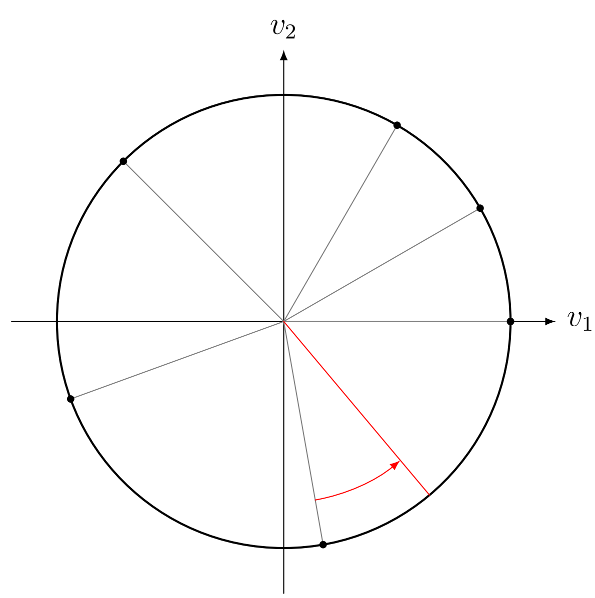


Figure 3: Variation in Choice Environments

The result follows from basic measure theory and characterizes the features of $F_V(v)$ that are identified under such variation in choice environments without any further restrictions. In particular, with the free normalization $\|v_i\| = 1$, the result implies that the mass accumulated on the projection of the sets in $\mathcal{D}(\mathcal{N})$ on the $J-1$ dimensional sphere, $\mathbb{S}^J$, is identified. Typically, this implies only partial identification of $F_V(v)$, but extensive variation in the lotteries could result in point identification.[3]

One way to interpret this result is that enough variations in the set of lotteries faced by an individual can be used to identify preferences. Figure 3 illustrates this visually by that variation in lottery sets sweep out an arc along utilities (normalized to be) on the circle. Of course, we do not expect that typical variation in the data will be rich enough to use non-parametric estimation methods based on such variation.

3.4 Identification With Preference Shifters

In this section we assume that the indirect utilities are given by

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

where $\epsilon_i \perp z^1_{ijb}$. In the school choice context, this independence assumption may be made on a characteristic that varies by student and school.[4] The term z^1_{ijb} is sometimes referred to as a special regressor (Lewbel, 2000; Berry and Haile, 2010, 2014). Let $\zeta(\xi, z^2)$ be the support of z^1 conditional on (ξ, z^2). For simplicity of notation, we will drop ξ, z^2 with the reminder that these are variables that the researcher needs to condition on. 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 8. *A convex cone C is **simplicial** if it is spanned by J linearly independent vectors $\{v_1, v_2, ..., v_J\} = V$ so that $C = \{v \in \mathbb{R}^J : v = Va$ for some $a \geq 0\}$.*

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

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.[5] This identification result exploits local variation of d.

Theorem 3. *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 $\zeta = \mathbb{R}^J$.*

[3] Specifically, the $\pi - \lambda$ theorem implies that $F_V(v)$ is identified if and only if the Dynkin-system $\mathcal{D}(\mathcal{N})$ contains a π-system that generates the Borel σ-algebra.

[4] For instance, Abdulkadiroglu et al. (2014) assume that distance to school is independent of student preferences. The assumption is violated if unobserved determinants of student preferences simultaneously detemine residential choices.

[5] 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.

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. See Appendix B.3. □

The result follows by using local variation around z^1 to identify the density of ν evaluated at z^1. Intuitively, we can use local changes in z^1 to shift the distribution of cardinal utilities to favor certain lotteries over others. Since simplicial cones are spanned by linearly independent vectors, we can decompose the change in how often a lottery is chosen into the principle directions to identify the density.

Also note that the local nature of this identification result articulates precisely, the fact that identification of the density at a point does not rely on observing extreme values of z^1. Of course, identification of the tails of the distribution of ν will rely on support on extreme values of z^1. 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 $J = 2$ and L is extremal, then the normal cone $N_{\mathcal{L}}(L)$ is simplicial. For $J \geq 3$, 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 z^1 has full support on $\mathbb{R}^J$ as long as the tails of g are exponentially decreasing. Formally, assume that the density of ν belongs to the set

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

Theorem 4. *Assume that $g \in \mathcal{G}$ and there is a lottery L such that $N_{\mathcal{L}}\left(L\right)$ 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. See Appendix B.4. □

The proof is based on Fourier-deconvolution techniques since the distribution of v if given by a location family parametrized by z^1. The key insight is that fourier tranform of an exponential density restricted to any salient cone is non-zero on any open set. This allows us to learn about g from observing how choices over lotteries change with z^1. However, because the result is based on deconvolution techniques, it requires stronger support restrictions than in Theorem 3. Nonetheless, the conditions on $\mathcal{G}$ are quite weak, and are satisfied for commonly used distributions with additive errors such as normal distributions, or generalized extreme value distributions. It is also satisfied if ν has bounded support.[6]

[6]When ν has bounded support, the support conditions on ζ can also be relaxed. In this case, we can allow for ζ to be a corresponding bounded set.

4 Estimation

Non-parametric estimation of random utility models can be computationally prohibitive and imprecise in finite samples, particularly if there are several choices. Following the discrete choice literature, we parametrize the distribution of indirect utilities $F_V(v|z,\xi)$ with $F_{V;\theta}(v|z,\xi)$ where θ belongs to a compact set $\Theta \in \mathbb{R}^K$. The identification results in the previous section can therefore be interpreted as articulating the fact that the parametric assumptions are made for tractability rather than essential maintained assumptions.

We consider a two-step estimator where in the first step we replace $\phi^\infty((R,t),m)$ with a consistent estimate $\hat{\phi}(R,t)$. For example, $\hat{\phi}(R,t) = \phi^n((R,t),m^{n-1})$ where m^{n-1} is the empirical measure on the reports and priority types of $n-1$ agents in the sample. Condition 1 implies that $\hat{\phi}(R,t) \xrightarrow{p} \phi^\infty((R,t),m)$. Our second step is defined as an extremum estimator:

$$\hat{\theta} = \inf_{\theta \in \Theta} Q_n(\theta, \hat{\phi})$$

Theorem 5 (Consistency)**.** *Suppose there exists a function Q_0 such that*

1. *θ are elements of a compact set*

2. *$\|\hat{\phi}(R,t) - \phi^\infty((R,t),m)\|_\infty \xrightarrow{p} 0$*

3. *$\sup_{\theta,\phi} |Q_n(\theta,\phi) - Q_0(\theta,\phi)| \xrightarrow{p} 0$*

4. *$Q_0(\theta,\phi)$ is jointly continuous in θ and ϕ*

5. *$Q_0(\theta,\phi_0)$ is uniquely minimized at θ_0*

then $\hat{\theta} \xrightarrow{p} \theta_0$.

Proof. Hypotheses 1-4 and the continuous mapping theorem imply that $\sup_{\theta\in\Theta} |Q_n(\theta,\hat{\phi}) - Q_0(\theta,\phi_0)| \xrightarrow{p} 0$. The conclusion follows by 1, 5, and Newey and McFadden (1994), Theorem 2.1. □

The objective function Q_n could be based on a likelihood or a method of moments.

5 Elementary School Admissions in Cambridge

This section describes the application. For now, we only describe the Cambridge Elementary School admissions system and the data avialable for our study.

5.1 Data

We have obtained data from the Cambridge Public School's (CPS) Controlled Choice Plan for the academic years 2004-2005 to 2008-2009. The CPS system has 12 schools and about 400 students participating in it each year. Two schools are divided into bilingual and regular programs in which bilingual eligible students are are considered only for the bilingual program. The other 10 schools are divided into paid lunch and free/reduced lunch programs. Student eligible for federal free or reduced lunch are only considered for the corresponding program. A goal of the Controlled Choice Plan is to maintain a ratio of paid lunch students in each school to be close to the district wide average. It implements this by setting an overall number of seats in a school and a maximum for each of the programs categorized by paidlunch. The sum of the program seats may exceed the total number of seats. Our dataset contains both the total number of seats available in the lottery as well as the seats available in each of the programs.

Elementary schools in the CPS system assign about 41% of the seats through a partnerships with pre-schools (junior kindergarten or montessori) or an appeals process for special needs students. The remaining seats are assigned through a "lottery process." We now describe this process.

5.2 The Cambridge Controlled Choice Mechanism

The process prioritizes students at a given school based on two criteria:

1. Students with siblings that are attending that school get the highest priority

2. Students receive priority at the two schools closest to their residents

Students can submit a ranking of up to three school programs that they are eligible at. A variant of the Boston Mechanism assigns students as follows:

Step 0: Generate a single lottery for each student

Step 1: Each school considers all students that listed it first and arranges the students in order of priority, breaking ties using the lottery.

   1. The top student that has not been considered in this round is assigned to the paid lunch program if she is not eligible for a federal lunch subsidy and there is an open seat in the paid lunch program. If she is eligible for a federal lunch subsidy, then she is assigned to the free/subsidized lunch program as long as seats are remaining.

2. This step is iterated until all students are considered.

Step k: Students not assigned to their $k-1$-st choice are assigned as in Step 1.

We observe these submitted rank-order lists, seats available and the student priorities. Strictly speaking, even though this mechanism is very similar to the Boston Mechanism, it is not a report-specific priority + cutoff mechanism as defined in this paper because there are two cutoffs, one for each type of program, in a school. The school cutoffs may bind even if both individual program cutoffs do not since the capacity at a school is typically lower than the sum of the capacity at each of the programs. The following result shows that Condition 1 is still satisfied by the Cambridge mechanism.

Proposition 4. *For any m, the Cambridge mechansim satisfies Condition 1 generically in q.*

Proof. See Appendix C. □

Our proof explicitly considers the three rounds of the Cambridge mechanism from the perspective of a single student and keeps track of the set of students that are rejected in each round. It constructs the set of lottery draws for which a student is rejected or assigned in any given round. Holding fixed the draws of the other students, the assignment indicator discretely jumps at certain lottery draws. We integrate over the lottery draws to smooth these jumps.

Fixing the limit measure m, this technique fails for knife-edge cases of q because students with certain report-priority combination may be pivotal due to finite-sample noise, but have measure 0 in the limit. For example, it may be that at the limit m, a school capacity is exactly exhausted by students with sibling priorities reporting the school first. Assume that no students with only proximity priority report this school. However, due to finite sample noise, some realizations of m^{n-1} may result in positive probabilities that additional students are assigned to this school. If the measure of students with proximity priority in this school is small enough, this uncertainty does not vanish in the limit. Constructing consistent estimated of the counterfactual assignment probabilities of an agent of this priority-type deviating to this report would not be possible. However, this problem is solved by perturbing q so that in the limit, a very small fraction of students with sibling priority are rejected or if a small fraction of seats remain after assigning these students.

6 Conclusion

We develop a general method for analyzing preferences from reports made to a single unit assignment mechanism that may not be truthfully implementable. We view the choice of report as a choice from available assignment probabilities. The available probabilities can be consistently estimated under a weak condition on the convergence of a sequence of mechanism to a limit. The condition is verified for a broad class of school choice mechanisms including the Boston mechanism or the Deferred Acceptance mechanism. We then characterize the identified set of preference distributions under the assumption that agents play a (limit) Bayesian Nash Equilibirum. The set of preference distributions are typically not point identified, but may be with sufficient variation in the lottery set. We then obtain point identification if a special regressor is available.

The methods in this paper rely on sophisticated agents participating in the mechanism. We discuss some extensions but leave a formal treatment of estimation and identification issues for future research.

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., Che, Y.-K., and Yasuda, Y. (2011). 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., Gandhi, A., and Haile, P. A. (2013). Connected Substitutes and Invertibility of Demand. *Econometrica*, 81(5):2087–2111.

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

Berry, S. T. and Haile, P. A. (2014). Identification in Differentiated Products Markets Using Market Level Data. *Econometrica, forthcoming.*

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.

Block, H. and Marshak, J. (1960). Random Orderings and Stochastic Theories of Responses. In Olkin, I., Ghurye, S., Hoeffding, W., Mado, W., and Mann, H., editors, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press.

Calsamiglia, C. and Guell, M. (2014). The Illusion of School Choice: Empirical Evidence from Barcelona.

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.

Chen, Y. and Kesten, O. (2013). From Boston to Chinese parallel to deferred acceptance: Theory and experiments on a family of school choice mechanisms. *Discussion Papers, Research Unit: Market Behavior*.

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.

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.

Haeringer, G. and Klijn, F. (2009). Constrained school choice. *Journal of Economic Theory*, 144(5):1921–1947.

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.

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

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, 8(3):229–254.

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.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, volume 4, pages 2111–2245. Elsevier.

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.

Pennell, H., West, A., and Hind, A. (2006). Secondary School Admissions in London.

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 Lemma 1

Existence and Uniqueness of Cutoffs

We introduce two definitions before discussing existence and uniqueness. The first definition is a notion of substitutes in a neighborhood around the market clearing price. This borrows from the notion of connected substitutes introduced in Berry et al. (2013); Berry and Haile (2010) to show conditions when demand is invertible.

Definition 9. *The aggregate demand function satisfies **local connected substitutes** at $p^* \in [0,1]^J$ if there exists an $\varepsilon > 0$, such that for all $p \in [0,1]^J$ with $\|p - p^*\| < \varepsilon$, we have that*

1. *for all $j \in \{1, \ldots, J\}$ and $k \neq j$, $D_j(p)$ is nondecreasing in p_k*
2. *for all non-empty subsets $K \subset \{1, \ldots, J\}$, there exist $k \in K$ and $l \notin K$ such that $D_l(p)$ is strictly increasing in p_k*

Definition 10 (Azevedo and Leshno (2013))**.** *The demand function $D(p|\eta)$ is **regular** if the image $D(\bar{P}|\eta)$, where*

$$P = \{p \in [0,1]^J : D(\cdot|\eta) \text{ is not continuously differentiable at } p\}$$

has Lebesgue measure 0.

We now observe that Assumption 2 is satisfied (generically satisfied) if the demand function satisfies connected substitutes (regular).

Proposition 5. *Every economy (η, q) admits at least one market clearing cutoff.*

Further, for a fixed η, let Q be the set of capacities such that for (η, q) has multiple market clearing cutoffs.

1. *$Q \cap \{q : \sum_j q_j < \eta(\mathcal{R} \times T \times [0,1]^J)\}$ has Lebesgue measure zero if η is regular*
2. *Q is empty if $D(p|\eta)$ satisfies local connected substitutes for any market clearing cutoff p^*. In particular, Q is empty if $D(p|\eta)$ satisfies local connected substitutes at every cutoff p.*

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

Uniqueness of the cutoff result is a generalization of Theorem 1 of Azevedo and Leshno (2013). Statement 1 is a consequence of Azevedo and Leshno (2013), Theorem 1(2).

Statement 2 is a strengthening of Azevedo and Leshno (2013), Theorem 1(1). 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 p^* be the market clearing cutoff such that $D(p|\eta)$ satisfies local connected substitutes at p^*. Let $C^+ = \{j \in S : p_j^* < p_j^+\}$ and $C^- = \{j \in S : p_j^* > p_j^-\}$. We will show that $C^+ = \emptyset$ i.e. $p^+ = p^*$. The proof to show that $C^- = \emptyset$ is symmetric and together, these claims imply that $p^+ = p^- = p^*$.

Towards a contradiction, assume that $C^+ \neq \emptyset$. Since $D(p|\eta)$ satisfies local connected substitutes at p^* (Definition 9), there exists $\varepsilon \in (0,1)$, $k \in C$ and $l \notin C$ such that,

$$D_l(p^*|\eta) > D_l(p^\varepsilon|\eta),$$

where $p_k^\varepsilon = (1-\varepsilon)p_k^+ + \varepsilon p_k^*$ and $p_j^\varepsilon = p_j^*$ if $j \neq k$. Hence, we have that

$$\sum_{j \in S \setminus C^+} D_j(p^*|\eta) > \sum_{j \in S \setminus C^+} D_j(p^\varepsilon|\eta) \geq \sum_{j \in S \setminus C^+} D_j(p^+|\eta),$$

where the implication on the summation and the second inequality are implied by the definition of aggregate demand. Since this inequality contradicts the Rural Hosptals Theorem, it must be that $C^+ = \emptyset$. □

Remark 3. *The condition that $D(p|\eta)$ satisfies local connected substitutes for all $p \in [0,1]$ is testable. Note that connected substitutes is implied by strict gross substitutes.*

Finally, we prove Lemma 1. The result is similar in spirit to Azevedo and Leshno (2013), Theorem 2. It differs from their results in that we are considering a random sequence of economies.

Proof of Lemma 1

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^*. Define

$$Q_n(p) = \left\| \begin{bmatrix} \max\{z(p|\eta^n, q^n), 0\} \\ p \odot z(p|\eta^n, q^n) \end{bmatrix} \right\|,$$

where $\odot$ represents element-wise multiplication. Note that p^n is a market clearing cutoff iff $Q_n(p) = 0$. Let Q_0 be the limit. By the continuous mapping theorem, $\sup_p |Q_n(p) - Q_0(p)| \xrightarrow{p} 0$. Also, $Q_0(p)$ is continuous since Assumption 1 implies that $D(p|\eta)$ is continuous. Assumption 2 implies that $Q_0(p)$ is uniquely minimized at p^*. By Theorem 2.1 of Newey and McFadden (1994), $\|p^n - p^*\| \xrightarrow{p} 0$.

A.3 Proof of Theorem 1

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

Definition 11. *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 satisfies Assumption 1.

Lemma 3. *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} converges in probability to m and consequently, η^{n-1} on reports and lotteries converges in probability to 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 3 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

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

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 $p_j^n = \frac{\underline{e}_j - k_j}{J} + \frac{J-1}{J}$ 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.

Serial Dictatorship:

The Serial Dictatorship Mechanism orders the students according to a single priority and then assigns the top student to her top ranked choice. The k-th student is then assigned to her top ranked choice that has remaining seats. It is straightforward to show that this mechanism is equivalent to a Deferred Acceptance mechanism in which all students have identical priorities at all schools. Hence, it is a report-specific priority + cutoff mechanism.

First Priority First:

The First Priority First mechanism assigns students to their top ranked choice if seats are available, with tie-breaking according to priorities and lotteries. Rejected students are then processed for the remaining seats according to the Deferred Acceptance mechanism. Arguments identical to the ones above show that the First Priority First mechanism is a report-specific priority + cutoff mechanism for

$$f_j(R, e) = \frac{e_j + 1\{jRj' \quad \forall j' \neq j\}}{2}.$$

Chinese Parallel (Chen and Kesten, 2013):

The chinese parallel mechanism operates in t rounds, each with t_c-subchoices. In each round, rejected students apply with the next t_c highest choices that have not yet rejected her. Within each round, the algorithm implements a deferred acceptance procedure in which applications are held tentatively until no new proposals are made. Assignments are finalized after all t_c choices have been considered. It is straightforward to show that the Chinese Parallel mechanism is a report-specific priority + cutoff mechanism for

$$f_j(R,e) = \frac{e_j - \left\lfloor \dfrac{\#\{k : kR_ij\}}{t_c} \right\rfloor}{\left\lfloor \dfrac{J}{t_c} \right\rfloor} + \frac{\left\lfloor \dfrac{J-1}{t_c} \right\rfloor}{\left\lfloor \dfrac{J}{t_c} \right\rfloor}.$$

Pan London Admissions (Pennell et al., 2006):
The Pan London Admissions system uses the Student Proposing Deferred Acceptance Mechanism except that a subset of schools upgrade the priority of students that rank the school highly. Suppose school j upgrades students that rank it first. For such schools, we set

$$f_j(R,e) = \frac{e_j + 1\{jRj' \quad \forall j' \neq j\}}{2},$$

and $f_j(R,e) = e$ otherwise. With this modification, the Pan London Admissions scheme is a report-specific priority + cutoff mechanism.

B Proofs: Identification

B.1 Proof of Theorem 2

We show that if $\phi^\infty_j((R^*,t),m) > 0$ for all $j \in \mathrm{Im}(R^*)$ then $\phi^\infty_j((R^*,t),m)$ is extremal. Towards a contradiction, assume that there exist λ_R for $R \in \mathcal{R}$ such that

$$\begin{aligned} \lambda_R &\geq 0 \\ \sum \lambda_R &= 1 \\ \phi^\infty\left(\left(R^*,t\right),m\right) &= \sum \lambda_R \phi^\infty\left(\left(R,t\right),m\right). \end{aligned}$$

We show that $\lambda_R = 0$ if $R\left(k\right) \neq R^*\left(k\right)$ by induction on $k \in \{1,\ldots,K\}$. For the base case, set $k=1$. Part (ii) of Definition 7 implies that for all $R \in \mathcal{R}$,

$$\phi^\infty_{R^*(1)}\left(\left(R^*,t\right),m\right) \geq \phi^\infty_{R^*(1)}\left(\left(R,t\right),m\right).$$

Assume that this inequality is strict for some $\tilde{R}$. Since

$$\phi^{\infty}_{R^*(1)}((R,t),m) = \sum \lambda_R \phi^{\infty}_{R^*(1)}((R,t),m) > 0$$

and $\lambda_R \geq 0$ with $\sum \lambda_R = 1$, it must be that $\lambda_{\tilde{R}} = 0$. Therefore, if $\lambda_R > 0$, then $R(1) = R^*(1)$.

Assume that, $\lambda_R = 0$ if $R(k') \neq R^*(k')$ for all $k' < k$. Note that part (ii) of Definition 7 implies that for all R with $\lambda_R > 0$,

$$\phi^{\infty}_{R^*(k)}((R^*,t),m) \geq \phi^{\infty}_{R^*(k)}((R,t),m).$$

If this inequality is strict, then

$$\phi^{\infty}_{R^*(k)}((R^*,t),m) = \sum \lambda_R \phi^{\infty}_{R^*(k)}((R,t),m),$$

$\lambda_R \geq 0$ and $\sum \lambda_R = 1$ imply that that $\lambda_{\tilde{R}} = 0$ unless $\tilde{R}(k) = R^*(k)$ or $\phi^{\infty}_{R^*(k)}((R^*,t),m) = 0$ and $\phi^{\infty}_{R^*(k)}((R,t),m) = 0$. The second possibility is ruled out by the hypothesis of the theorem.

By strong induction, we have proved the result.

B.2 Proof of Lemma 2

The identified set of conditional distributions $F_V(v)$ is given by

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

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

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

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

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

is identified.

B.3 Proof of Theorem 3

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 $\varepsilon = Va$:

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

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

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

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

B.4 Proof of Theorem 4

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 $\varepsilon \in \mathrm{int}(T_{\mathcal{L}}(L))$, 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 4 below.

Lemma 4. *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 $\varepsilon \in \mathrm{int}(\Gamma^o)$, and let $\hat{f}_{\varepsilon,\Gamma}(\alpha)$ be its Fourier Transform. $\hat{f}_{\varepsilon,\Gamma}$ is an entire function. 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).

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

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 $\xi = \left(V_K^{-1}\right)' \alpha$,

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

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. □

C Verifying Condition 1 for the Cambridge Mechanism

We first find a representation of the Cambridge Mechanism as a function

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

We first find a representation of the Cambridge Mechanism as a function

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

C.1 Representation

C.1.1 Priorities and Lotteries

Each student receives an independent priority draw ν_i from a uniform $[0,1]$ distribution. We modify this random priority by the sibling and proximity priority t_i. Let $f : [0,1] \times \mathcal{T} \to [0,1]^J$, such that for each $j = 1, ..., J$:

$$e_{ij} = f_j(\nu_i, t_i) = \frac{\nu_i + t_{ij}}{T} \in [0,1]$$

where T is the maximum priority points a student can have. In Cambridge, $t_{ij} = 1$ if student i has only proximity priority at program j, $t_{ij} = 2$ if student i has only sibling priority at program j, and $t_{ij} = 3$ if student i has both proximity and sibling priority at program j.

C.1.2 Economy

Let Π be a partition of the programs in Cambridge into a set of schools in Cambridge and let $q \in \mathbb{R}_+^{J+|\Pi|}$ be a vector of program and school capacities. Typically, for any $\pi \in \Pi$, $\sum_{j\in\pi} q_j < q_\pi$.

Consider a n-student economy where the vacancies are represented by $q^n \in \mathbb{R}_+^{J+|\Pi|}$, the measure of report-priority shares of all but the focal student is given by

$$m^{n-1} = \frac{1}{n-1}\sum_{i=1}^{n-1} \delta_{R_i,t_i}$$

and η^{n-1} includes the realization of random priority draws of the $n-1$ students

$$\eta^{n-1} = \frac{1}{n-1}\sum_{i=1}^{n-1} \delta_{R_i,t_i,e_i}$$

where η^{n-1} agrees with m^{n-1} on the marginals on R and t.

C.1.3 Sub-Economies in Rounds $k \in \{1,2,3\}$

With a slight abuse of notation, let $R_{[k]}$ be program in position k in report R. We will use a map $s(\eta, q|k) \mapsto (\eta', q')$ that takes a measure over reports, priority types, random priorites, and a capacity in each round and maps it to a measure over remaining reports, priority-types and random priorities in the next round.

To define $s(\eta, q|k)$, we introduce some additional notation. Let $D_{j,k}\left(p|\eta\right) = \eta\left(\left\{(R,e) : R_{[k]} = j, e_j \geq p\right\}\right)$ be the measure of types that ranked school j in the k-th round and have eligbility at least p in that round. Note that $D_{j,k}\left(p|\eta\right)$ is nonincreasing. Define the excess capacity at elegibility p as:

$$\begin{aligned}
\tilde{z}_j\left(p;\eta,q|k\right) &= q_j - D_{j,k}\left(p|\,\eta\right) \\
\tilde{z}_{\pi_j}\left(p;\eta,q|k\right) &= q_\pi - q_j - \sum_{l\in\pi_j/\{j\}} \min\left\{q_l, D_{l,k}\left(p|\,\eta\right)\right\}. \\
z_j\left(p;\eta,q|k\right) &= \tilde{z}_j\left(p;\eta,q|k\right) + \min\left(0, \tilde{z}_{\pi_j}\left(p;\eta,q|k\right)\right)
\end{aligned}$$

z_j is nondecreasing in p.

A student is not assigned to a school if the measure of students that have (weakly) higher eligbility exceeds the school or the program's capacity. Therefore, the set of students that are not assigned in step k can be written as

$$r(\eta, q|k) = \left\{(R,e):\ z_{R_{[k]}}(e;\eta,q|k) < 0\right\}.$$

Define η' as the restriction of η to $r(\eta, q|k)$.

The capacities that remain after step k, are given by:

$$q_j' = \max\{q_j - D_{j,k}(0|\eta), 0\}$$

since all students, i.e. measure $D_{j,k}(0|\eta)$, are assigned if there are seats available.

C.1.4 Cambridge Mechanism

Let $(\eta_1, q_1) = (\eta, q)$ and $(\eta_k, q_k) = s(\eta_{k-1}, q_{k-1}|k)$. Define the following function:

$$\varphi_{(R,t)}(\nu;\eta,q,k) = 1\left[\left(R, \frac{\nu+t}{T}\right) \in r(\eta_k, q_k|k)^c \cap_{k'<k} r(\eta_{k'}, q_{k'}|k')\right].$$

This function returns 1 if a student that reports R and has priority (ν, t) is assigned to program $R_{[k]}$ when the measure over reports and priorities is given by η and the vector of capacities is q.

For a fixed student priority-type, report and lottery-draw, (R, t, e) define

$$\eta^n = \frac{1}{n}\left[(n-1)\eta^{n-1} + \delta_{R,t,e}\right].$$

Note that the finite economy and limit economy mechanisms are given by

$$\begin{aligned}
\phi^n_{R_{[k]}}\left((R,t), m^{n-1}, q^n\right) &= \int E\left[\varphi_{(R,t)}(\nu;\eta^n,q^n,k)\middle|\, m^{n-1},\nu\right] d\nu \\
\phi^\infty_{R_{[k]}}((R,t), m, q) &= \int \varphi_{(R,t)}(\nu;\eta,q,k)\, d\nu
\end{aligned}$$

where the limit measure η if given by

$$\eta(R, e < p) = \sum_{t=0}^{T} m(R,t) \min_j (p_j T - t_j). \tag{3}$$

C.2 Main Results

We make the following assumption about the genericity of vacancies:

Assumption 5 (Generic Vacancies)**.** *For $k = 1, 2, 3$, let*

$$m_k(R,t) = \eta_k\left(\{(R',e') : R' = R, t \le Te' \le 1+t\}\right)$$

where $(\eta_k, q_k) = s(\eta_{k-1}, q_{k-1}|k-1)$ and $(\eta_1, q_1) = (\eta, q)$. If $m(R,t) = 0$ then for each k,

$$\min\begin{bmatrix} q_{k,R_{[k]}} - \sum m_k(R',t')\,1\left(R'_{[k]} = R_{[k]}, t'_{R_{[k]}} > t_{R_{[k]}}\right), \\ q_{k,\pi_{R_{[k]}}} - \sum_{l\in\pi_{R_{[k]}}} \min\left\{q_{k,l}, \sum m_k(R',t')\,1\left(R'_{[k]} = l, t'_{R_{[k]}} > t_{R_{[k]}}\right)\right\} \end{bmatrix} \neq 0$$

For each (R,t), there is no open set in $\mathbb{R}_+^{J+|\Pi|}$ such that every q in that set violates Assumption 5. Fix a q such that this assumption is satisfied. We restate Proposition 4.

Proposition 6. *Assume that (m,q) satisfies Assumption 5 above. If m^{n-1}, q^n are empirical sequences such that $m^{n-1} \xrightarrow{p} m$, and $q^n \xrightarrow{p} q$, then for each $k \in \{1,2,3\}$ and (R,t)*

$$\phi^n_{R_{[k]}}\left((R,t), m^{n-1}, q^n\right) \xrightarrow{p} \phi^\infty_{R_{[k]}}\left((R_1,t_1), m, q\right).$$

C.3 Proof of Proposition 6

We first state some preliminary results:

Let $\triangle$ be the symmetric difference operator. Consider the VC class of sets

$$\mathcal{V} = \{V : \exists (R,p,k) \in \mathcal{R} \times [0,1] \times \{1,2,3\}, V = v(R,p,k)\},$$

where $v(R,p,k) = \left\{(R,e) : e_{R[k]} < p\right\}$.

Lemma 5. *If $\sup_{V\in\mathcal{V}} |\eta^n(V) - \eta(V)| \xrightarrow{p} 0$, $\sup_j \left|q^n_j - q_j\right| \xrightarrow{p} 0$ and $D_{j,k}(p|\eta)$ is continuous in p for all j and k, then (i) $\sup_{p,j,k} |D_{j,k}(p|\eta^n) - D_{j,k}(p|\eta)| \xrightarrow{p} 0$, (ii) $\sup_{\nu,j,k} |z_j(\nu; t, \eta^n, q^n|k) - z_j(\nu; t, \eta, q|k)|$ 0 where each $z_j(\nu; t, \eta, q|k)$ is continuous and nondecreasing in ν, (iii) $r(\eta, q|k) = \bigcup_{R\in\mathcal{R}} V_R$ where each $V_R \in \mathcal{V}$, (iv) $\eta^n(r(\eta^n, q^n|k) \triangle r(\eta, q|k)) \xrightarrow{p} 0$, and (v) if η' as the restriction of η to $r(\eta, q|k)$ then $D_{j,k}(p|\eta')$ is continuous in p for all j and k.*

Proof. **Parts (i - iii):** For every $p \in [0,1]$,

$$\begin{aligned} D_{j,k}(p|\eta) &= \eta\left(\left\{(R,e) : R_{[k]} = j, e_j \ge p\right\}\right) \\ &= \sum_{R:R_{[k]}=j} \eta(v(R,1,k)) - \eta(v(R,p,k)). \end{aligned}$$

Hence, part (i) follows from uniform convergence of η^n over sets in $\mathcal{V}$. Part (ii) follows from the continuous mapping theorem: $z_j(\cdot; t, \eta, q|k)$ is continuous with respect to functions $D_{l,k}(\cdot|\,\eta)$, where both types of functions are vectors in vector spaces endowed with the sup-norm. Continuity of $z_j(\nu; t, \eta, q|k)$ follows directly continuity of the min function and of $D_{l,k}(\cdot|\,\eta)$ for every l. Part (iii) is easily verified noting that $r(\eta, q|k) = \bigcup_j \bigcup_{R:R_{[k]}=j} \upsilon(R, p_j, k)$ where $p_j = 0$ if $z_j(0; \eta, q|k) \geq 0$ and otherwise,

$$p_j = \sup\{e \in [0,1] : z_j(e; \eta, q|k) < 0\}.$$

Part (iv): The definitions of $r(\eta, q|k)$ and $r(\eta^n, q^n|k)$ imply:

$$\begin{aligned} & \eta^n\left(r(\eta^n, q^n|k) \triangle r(\eta, q|k)\right) \\ = & \sum_j \eta^n\left(\left\{(R,e) : R_{[k]} = j, (e_j < p_j \vee z_j(e; \eta^n, q^n|k) \geq 0) \wedge (e_j \geq p_j \vee z_j(e; \eta^n, q^n|k) < 0)\right\}\right) \end{aligned} \tag{4}$$

where $\vee$ and $\wedge$ are logical AND and OR respectively. It is enough to show convergence in probability for each term in the summation.

Pick an N such that for all $n > N$ with probability greater than $1 - \varepsilon$,

$$\sup_{k,e} |z_j(e; \eta, q|k) - z_j(e; \eta^n, q^n|k)| \leq \frac{\varepsilon}{2} \tag{5}$$

and

$$\sup_{p_1 \leq p_2, R'} \eta^n\left(\{(R,e) : R = R', p_1 \leq e_j \leq p_2\}\right) \leq |p_1 - p_2| + \frac{\varepsilon}{8}. \tag{6}$$

Existence of such an N is guaranteed by part (ii) of the Lemma above and since

$$\sup_{p_1 \leq p_2, R'} \eta\left(\{(R,e) : R = R', p_1 \leq e_j \leq p_2\}\right) \leq |p_1 - p|.$$

We first show that equation (6) implies that

$$\eta^n\left(\left\{(R,e) : R_{[k]} = j, z_j(e; \eta^n, q|k) \in [a,b]\right\}\right) \leq \frac{\varepsilon}{4} + b - a. \tag{7}$$

Let $\underline{e}_n = \inf\{e : z_j(e;\eta^n, q|k) > a\}$, $\overline{e}_n = \sup\{e : z_j(e;\eta^n, q|k) < b\}$. We have that

$$\begin{aligned}
& \eta^n\left((R,e) : R_{[k]} = j, z_j(e;\eta^n, q^n|k) \in [a,b]\right) \\
\leq\ & \eta^n\left((R,e) : R_{[k]} = j, e \in [\underline{e}_n, \overline{e}_n]\right) \\
=\ & \eta^n\left((R,e) : R_{[k]} = j, e \in (\underline{e}_n, \overline{e}_n)\right) + \eta^n\left((R,e) : R_{[k]} = j, e \in \{\overline{e}_n, \overline{e}_n\}\right) \\
\leq\ & \lim_{e\downarrow\underline{e}_n} D_{j,k}(e|\eta^n) - \lim_{e\uparrow\overline{e}_n} D_{j,k}(e|\eta^n) + \eta^n\left((R,e) : R_{[k]} = j, e \in \{\overline{e}_n, \overline{e}_n\}\right) \\
\leq\ & \lim_{e\downarrow\underline{e}_n} D_{j,k}(e|\eta^n) - \lim_{e\uparrow\overline{e}_n} D_{j,k}(e|\eta^n) + \frac{\varepsilon}{4} \\
=\ & \lim_{e\uparrow\overline{e}_n} \tilde{z}_j(e;\eta^n, q^n|k) - \lim_{e\downarrow\underline{e}_n} \tilde{z}_j(e|;\eta^n, q^n|k) + \frac{\varepsilon}{4} \\
\leq\ & \lim_{e\uparrow\overline{e}_n} z_j(e;\eta^n, q^n|k) - \lim_{e\downarrow\underline{e}_n} z_j(e|;\eta^n, q^n|k) + \frac{\varepsilon}{4} \\
\leq\ & b - a + \frac{\varepsilon}{4}
\end{aligned}$$

where the first inequality follows by the definition of $\underline{e}_n$ and $\overline{e}_n$; the second inequality follows from the definition of $D_{j,k}(e|\eta^n)$ and because it is decreasing; the third inequality follows from equation (6); the last inequality follows from the definition of $\tilde{z}_j$ and the final inquality follows from the fact that for all $e \in (\underline{e}_n, \overline{e}_n)$, $z_j(e;\eta^n, q^n|k) \in (a,b)$ and that $z_j(e;\eta^n, q^n|k)$ is monotonically increasing.

Now consider a term in the summation in equation (4). If $z_j(p_j;\eta^n, q^n|k) < 0$, this term is bounded by

$$\eta^n\left(\{(R,e) : e_j \geq p_j, z_j(e_j;\eta^n, q^n|k) \in [z_j(p_j;\eta^n, q^n|k), 0]\}\right).$$

If $z_j(p_j;\eta^n, q^n|k) \geq 0$, the term is bounded by

$$\eta^n\left(\{(R,e) : e_j < p_j, z_j(e;\eta^n, q^n|k) \in [0, z_j(p_j;\eta^n, q^n|k)]\}\right).$$

Hence, equations (7) and (5) imply that

$$\begin{aligned}
& \eta^n\left(\left\{(R,e) : R_{[k]} = j, (e_j < p_j \vee z_j(e;\eta^n, q^n|k) \geq 0) \wedge (e_j \geq p_j \vee z_j(e;\eta^n, q^n|k) < 0)\right\}\right) \\
\leq\ & |z_j(p_j;\eta, q|k) - z_j(p_j;\eta^n, q^n|k)| + 2 \times \frac{\varepsilon}{4} \\
\leq\ & \varepsilon.
\end{aligned}$$

Since equations (5) and (6) (consequently, equation (7)), hold for all $n > N$ with probability at least $1 - \varepsilon$, we have the desired result.

Part (v): Follows because

$$\begin{aligned} D_{j,k}(p|\eta') &= \eta'\left(\left\{(R,e) : R_{[k]} = j, e_j \geq p\right\}\right) \\ &= \eta\left(\left\{(R,e) : R_{[k]} = j, e_j \geq p\right\} \cap r(\eta, q|k)\right) \\ &= \eta\left(\left\{(R,e) : R_{[k]} = j, p_j > e_j \geq p\right\}\right) \\ &= \begin{cases} D_{j,k}(p|\eta) - D_{j,k}(p_j|\eta) & \text{if } p_j < p \\ 0 & \text{if } p_j \geq p \end{cases} \end{aligned}$$

and continuity of $D_{j,k}(p|\eta)$. $\square$

Define the function

$$\zeta_{(R,t)}\left(\nu;\eta,q,k\right) = \min\left\{ z_{R_{[k]}}\left(\frac{\nu + t_{R_{[k]}}}{T};\eta_k, q_k\middle|\, k\right), - \max_{k'<k} z_{R_{[k']}}\left(\frac{\nu + t_{R_{[k']}}}{T};\eta_{k'}, q_{k'}\middle|\, k'\right)\right\}.$$

If $\zeta_{(R,t)}\left(\nu;\eta,q,k\right) > 0$ both terms are positive. Program $R_{[k]}$ could enroll every unasigned student that ranked it in position k and that has a priority score higher than $\frac{\nu+t_{R_{[k]}}}{T}$ without exhausting program or school capacity. At the same time, if for some $k' < k$, program $R_{[k']}$ had enrolled every unasigned student that ranked it in position k' and had a priority score higher than $\frac{\nu+t_{R_{[k']}}}{T}$ it would have exceeded the total program or school capacity. Therefore a student with report and priority (R,t,ν) such that $\zeta_{(R,t)}\left(\nu;\eta,q,k\right) > 0$ is assigned to school $R_{[k]}$ in round k. Notice that $\zeta_{(R,t)}\left(\nu;\eta,q,k\right) > 0$ implies $\varphi_{(R,t)}\left(\nu;\eta,q,k\right) = 1$ and $\zeta_{(R,t)}\left(\nu;\eta,q,k\right) < 0$ implies $\varphi_{(R,t)}\left(\nu;\eta,q,k\right) = 0$.

Lemma 6. *If* $\sup_{V\in\mathcal{V}}\left|\eta^n\left(V\right) - \eta\left(V\right)\right| \xrightarrow{p} 0$ *and* $\sup_j\left|q_j^n - q_j\right| \xrightarrow{p} 0$ *where* η *is defined as in (3), then* $\sup_{\nu,R,t,k}\zeta_{(R,t)}\left(\nu;\eta^n,q^n,k\right) \xrightarrow{p} \zeta_{(R,t)}\left(\nu;\eta,q,k\right)$.

Proof. We first show that if $D_{j,k}(p|\eta)$ is continuous in p for all j and k, $\|s(\eta^n,\ q^n|k) - s\left(\eta,\ q|k\right)\|_\infty \xrightarrow{p} 0$ where

$$\|s(\eta^n,\ q^n|k) - s\left(\eta,\ q|k\right)\|_\infty = \max\left\{\sup_j\left|q_j'^n - q_j'\right|, \sup_{V\in\mathcal{V}}\left|\eta'^n\left(V\right) - \eta'\left(V\right)\right|\right\}.$$

Since q_j' is jointly continuous in q_j and $D_{j,k}\left(0|\eta\right)$, $q_j'^{,n} \xrightarrow{p} q_j'$ by the continuous mapping theorem.

$$\begin{aligned}
& \sup_{V\in\mathcal{V}} \left|\eta'^{n}(V)-\eta'(V)\right| \\
= \; & \sup_{V\in\mathcal{V}} \left|\eta^{n}\left(r(\eta^{n},q^{n}|k)\cap V\right)-\eta\left(r(\eta,q|k)\cap V\right)\right| \\
\leq \; & \sup_{V\in\mathcal{V}} \left|\eta^{n}\left(r(\eta,q|k)\cap V\right)-\eta\left(r(\eta,q|k)\cap V\right)\right| \\
& +\sup_{V\in\mathcal{V}} \left|\eta^{n}\left(r(\eta^{n},q^{n}|k)\cap V\right)-\eta^{n}\left(r(\eta,q|k)\cap V\right)\right|
\end{aligned}$$

The first term converges in probability to zero because $r(\eta^n,q^n|k)\in\mathcal{V}$ (Lemma 5, part iii) and $\mathcal{V}$ is closed under finite intersections. The second term is bounded by: $\eta^n\left(r(\eta^n,q^n|k)\bigtriangleup r(\eta,q|k)\right)$, which is shown to converge in probability to zero (Lemma 5, part iv). Moreover, for all j and k, $D_{j,k}(p|\eta')$ is continuous in p (Lemma 5, part v).

Notice that $D_{j,k}(p|\eta_1)$ is continuous in p for all j and k. By induction, $\sup_{V\in\mathcal{V}}\left|\eta^n_{k-1}(V)-\eta_{k-1}(V)\right|\xrightarrow{p}0$ and $\sup_j\left|q^n_{k-1,j}-q_{k-1,j}\right|\xrightarrow{p}0$ implies that for all $k=2,3$: $\sup_{V\in\mathcal{V}}\left|\eta^n_k(V)-\eta_k(V)\right|\xrightarrow{p}0$, $\sup_j\left|q^n_{k,j}-q_{k,j}\right|\xrightarrow{p}0$ and $D_{j,k}(p|\eta_k)$ is continuous in p. The result now follows from the the continuous mapping theorem and Lemma 5, part ii, since $\zeta_{(R,t)}(\cdot;\eta,q,k)$ is continuous in $z_j(\cdot;\eta,q|k)$ for all t,j,k. □

We are now ready for the main result

Proof. For each (R,t), there is no open set in $\mathbb{R}_+^{J+|\Pi|}$ such that every q in that set violates Assumption 5. Fix a q such that this assumption is satisfied. For this q, it is enough to show the result for fixed (R,t,k) since it belongs to a finite set. $\zeta_{(R,t)}(\nu;\eta,q,k)$ for notational convenience.

Let

$$\mathcal{E}_k=\left\{\nu:\zeta_{(R,t)}(\nu;\eta,q,k)=0\right\},$$

where $j=R_{[k]}$. We first show that $|\mathcal{E}_k|\leq 2$. Since

$$\zeta_{(R,t)}(\nu;\eta,q,k)=\min\left\{z_{R_{[k]}}\left(\frac{\nu+t_{R_{[k]}}}{T};\eta_k,q_k\middle|k\right),-\max_{k'<k}z_{R_{[k']}}\left(\frac{\nu+t_{R_{[k']}}}{T};\eta_{k'},q_{k'}\middle|k'\right)\right\},$$

where both components inside the min are monotonic, continuous functions of ν, it is easy to show that $\mathcal{E}_k$ is the union of at most two convex sets. Further, $\mathcal{E}_k$ is closed since $\zeta_{(R,t)}(\nu;\eta,q,k)$ is continuous in ν. Suppose that there is there is a k and an open interval $(\underline{\nu},\overline{\nu})\subseteq\mathcal{E}_k$. Then, for all $\nu\in(\underline{\nu},\overline{\nu})$, $D_j\left(\frac{\nu+t_j}{T}\middle|\eta\right)$ is constant. This only occurs if $m(R,t)=0$, which implies a violation of the generic vacancies condition. Since $\mathcal{E}_k\subseteq\mathbb{R}$, we have that $|\mathcal{E}_k|\leq 2$ and $\left|\cup_{k'\in\{1,..,k\}}\mathcal{E}_{k'}\right|<\infty$.

Fix $\varepsilon > 0$. Construct an open set U that covers $\cup_{k' \in \{1,..,k\}} \mathcal{E}_{k'}$ and has Lebesgue measure less than $\frac{\varepsilon}{2}$. Consider the difference,

$$\begin{aligned}
& \left| \phi^n_{R_{[k]}} \left((R,t), m^{n-1}, q^n \right) - \phi^\infty_{R_{[k]}} \left((R,t), m, q \right) \right| \\
= & \left| \int E \left[\varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(\nu; \eta, q, k \right) \right| m^{n-1}, q^n, \nu \right] d\nu \right| \\
\leq & \int E \left[\left| \varphi_{(R,t)} \left(e; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(\nu; \eta, q, k \right) \right| \right| m^{n-1}, q^n, \nu \right] d\nu \\
\leq & \sup_{\nu \notin U} E \left[\left| \varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(\nu; \eta, q, k \right) \right| \right| m^{n-1}, q^n, \nu \right] P \left(\nu \notin U \right) \\
& + \sup_{\nu \in U} E \left[\left| \varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(e; \eta, q, k \right) \right| \right| m^{n-1}, q^n, \nu \right] P \left(\nu \in U \right) \\
< & \sup_{\nu \notin U} E \left[\left| \varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(e; \eta, q, k \right) \right| \right| m^{n-1}, q^n, \nu \right] + \frac{\varepsilon}{2}
\end{aligned}$$

where the last inequality follows from the fact that $P(\nu \in U) < \frac{\varepsilon}{2}$ and

$$\sup_{\nu \in U} E \left[\left| \varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(\nu; \eta, q, k \right) \right| \right| m^{n-1} q^n, , \nu \right] \leq 1.$$

We now show that there exists N such that for all $n > N$:

$$\mathbb{P} \left(\sup_{\nu \notin U} E \left[\left| \varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(\nu; \eta, q, k \right) \right| \right| m^{n-1}, q^n, \nu \right] \geq \frac{\varepsilon}{2} \right) < \varepsilon. \tag{8}$$

This would complete the proof as it implies that

$$\mathbb{P} \left(\left| \phi^n_{R_{[k]}} \left((R,t), m^{n-1}, q^n \right) - \phi^\infty_{R_{[k]}} \left((R,t), m, q \right) \right| > \varepsilon \right) < \varepsilon.$$

Let $\zeta_\varepsilon = \inf_{\nu \notin U} \left| \zeta_{(R,t)} \left(\nu; \eta, q, k \right) \right| > 0$, since $\left| \zeta_{(R,t)} \left(\nu; \eta, q, k \right) \right| > 0$ and $\zeta_{(R,t)} \left(\nu; \eta, q, k \right)$ is continuous with respect to ν. By Lemma 6, there exists N such that for all $n > N$,

$$\mathbb{P} \left(\sup_{\nu \notin U} \left| \zeta_{(R,t)} \left(\nu; \eta, q, k \right) - \zeta_{(R,t)} \left(\nu; \eta^n, q^n, k \right) \right| > \zeta_\varepsilon \right) < \frac{\varepsilon^2}{2}.$$

Note that for all $\nu \notin U$, $\left| \zeta \left(\nu; \eta, q, k \right) \right| > \zeta_\varepsilon$. Therefore for all $\nu \notin U$,

$$\varphi_{(R,t)} \left(\nu; \eta^n, q^n, k \right) - \varphi_{(R,t)} \left(\nu; \eta, q, k \right) \neq 0 \Rightarrow \left| \zeta \left(\nu; \eta^n, q^n, k \right) - \zeta \left(\nu; \eta, q, k \right) \right| > \zeta_\varepsilon$$

since the antecedent requires $\zeta_{(R,t)} \left(\nu; \eta^n, q^n, k \right) \geq 0$ and $\zeta_{(R,t)} \left(\nu; \eta, q, k \right) < -\zeta_\varepsilon$ or $\zeta_{(R,t)} \left(\nu; \eta^n, q^n, k \right) \leq$

0 and $\zeta_{(R,t)}(\nu;\eta,q,k) > \zeta_\varepsilon$. By set inclusion, for all $n > N$,

$$\mathbb{P}\left(\sup_{\nu\notin U}\left|\varphi_{(R,t)}(\nu;\eta^n,q^n,k)-\varphi_{(R,t)}(\nu;\eta,q,k)\right|\neq 0\right)<\frac{\varepsilon^2}{2}.$$

Since $\sup_{\nu\notin U}\left|\varphi_{(R,t)}(\nu;\eta^n,q^n,k)-\varphi_{(R,t)}(\nu;\eta,q,k)\right| \in \{0,1\}$, the above inequality implies that for all $n > N$,

$$\begin{aligned}
\frac{\varepsilon^2}{2} &> E\left[\sup_{\nu\notin U}\left|\varphi_{(R,t)}(\nu;\eta^n,q^n,k)-\varphi_{(R,t)}(\nu;\eta,q,k)\right|\right]\\
&= E\left(E\left[\sup_{\nu\notin U}\left|\varphi_{(R,t)}(\nu;\eta^n,q^n,k)-\varphi_{(R,t)}(\nu;\eta,q,k)\right|\middle| m^{n-1},q^n\right]\right)\\
&\geq E\left(\sup_{\nu\notin U}E\left[\left|\varphi_{(R,t)}(\nu;\eta^n,q^n,k)-\varphi_{(R,t)}(\nu;\eta,q,k)\right|\middle| m^{n-1},q^n\right]\right),
\end{aligned}$$

where the equality follows from the law of iterated expectations and the weak inequality from the additional restriction that the optimal ν cannot depend on the realization of η^n. Markov inequality implies:

$$\mathbb{P}\left(\sup_{\nu\notin U}E\left[\left|\varphi_{(R,t)}(\nu;\eta^n,q^n,k)-\varphi_{(R,t)}(\nu;\eta,q,k)\right|\middle| m^{n-1},q^n\right]\geq\frac{\varepsilon}{2}\right)<\varepsilon$$

which is exactly equation (8). $\square$