Measuring the Welfare Gains from Cardinal-Preference Mechanisms in School Choice

Hulya Eraslan
Rice University & NBER

Jeremy T. Fox
Rice University & NBER

YingHua He
Rice University

Yakym Pirozhenko
Rice University

August 2022

Abstract

We compare cardinal-preference and ordinal-preference mechanisms for assignment problems such as school choice, with a focus on a variant of the pseudomarket mechanism of Hylland and Zeckhauser (1979) as well as an envy-free mechanism related to Nguyen, Peivandi and Vohra (2015). We introduce and theoretically analyze a variant of the pseudomarket mechanism that has an equilibrium selection rule. We also introduce a computer algorithm to compute the implied stochastic assignment. The envy free mechanism is a linear program and hence simpler to compute. We estimate cardinal preferences over schools using data on student submissions of rank ordered lists of schools in Seattle. Using these estimated preferences, we measure the welfare gains from using the pseudomarket and envy-free mechanisms instead of the ordinally-efficient probabilistic serial mechanism. We find that the psuedomarket captures 19% of the possible gains over the ordinal mechanism using a utilitarian benchmark. The envy-free mechanism captures 58% of the possible gains over the ordinal mechanism.

1 Introduction

The allocation of indivisible objects without monetary transfers is common in daily life. Examples include student placement in public schools (where an object corresponds to a school seat and each object has multiple copies) and the allocation of work or living space (where each object has exactly one copy). Due to the lack of monetary transfers, objects in these environments are often allocated by a centralized mechanism that maps agents' reported preferences to an allocation outcome. The outcome, known as an assignment, can be either deterministic or random. A deterministic assignment

dictates who gets what object. A stochastic assignment prescribes the probability shares of objects that each agent obtains and thus is a lottery over a set of deterministic assignments.

The standard allocation mechanisms used in practice and studied in the literature are ordinal: agents are asked to rank schools or offices; the profile of submitted rankings determines the assignment. This applies to many celebrated mechanisms, such as deferred acceptance, random serial dictatorship, and top trading cycles. However, Miralles (2008) and Abdulkadiroglu et al. (2011) point out that one can implement Pareto dominant assignments by eliciting agents' cardinal utilities. Cardinal utilities give relative intensities of preferences over objects and rates of substitution among probability shares in objects (e.g., Debreu, 1959). Furthermore, Liu and Pycia (2012) and Pycia (2014) show that sensible ordinal mechanisms are asymptotically equivalent in large markets, while mechanisms eliciting cardinal utilities maintain their efficiency advantages. The data on Boston and NYC school choice corroborates both the rough equivalence of ordinal mechanisms (see e.g., Pathak and Sonmez 2008 and Abdulkadiroglu et al. 2009) and the inefficiency of ordinal mechanisms (Abdulkadiroglu, Agarwal and Pathak 2017). Relatedly, Abdulkadiroglu et al. estimated that deferred acceptance gives 80% of the gains from an infeasible, utilitarian benchmark, for public school choice in New York City. This leaves 20% of their utilitarian benchmark for other mechanisms to possibly capture. For analysis of ordinal mechanisms see the seminal work of Abdulkadiroglu and Sonmez (2003) and Bogomolnaia and Moulin (2001).

Naturally, with more inputs, we expect a cardinal mechanism to deliver a better outcome; cardinal preferences are more informative than ordinal ones. The search for desirable cardinal mechanisms is both old and new. In their seminal paper, Hylland and Zeckhauser (1979) propose a **pseudomarket mechanism** that elicits cardinal preferences from agents and delivers an assignment. Internally in the mechanism, each agent has an equal budget of an artificial currency. The mechanism computes a competitive equilibrium for this market, where the items being priced and allocated are probability shares of objects. The resulting random assignment can be implemented by appeal to the Birkhoff-von Neumann Theorem. The pseudomarket mechanism is shown to have desirable properties. Agents do not have incentives to misreport in large markets (Azevedo and Budish, 2019; He, Miralles, Pycia and Yan, 2018). As the allocation is based on a competitive equilibrium, each agent is assigned a probability share profile that maximizes its expected utility subject to its budget constraint at the competitive equilibrium prices. This random assignment is Pareto efficient under certain assumptions described below. The random assignment is also envy-free in the sense that each agent weakly prefers its lottery over other agents' according to its expected utility, because all agents have identical budget constraints.

Despite these encouraging properties, Hylland and Zeckhauser (1979) has not yet been compared to existing, ordinal mechanisms in a field setting. Our paper measures the welfare gains from a pseudomarket mechanism in the classic setting of school choice, meaning assigning at most one indivisible item to each agent. We first describe a modified pseudomarket mechanism that largely addresses the

lack of an equilibrium selection rule in the original mechanism in Hylland and Zeckhauser. We then theoretically analyze our pseudomarket mechanism's properties.

At the time of the initial distribution of our paper, we knew of no computational implementation with more than a handful of agents of the Hylland and Zeckhauser mechanism, despite a recent surge of theoretical interest in extensions of pseudomarket mechanisms. One of our contributions is to describe a computational approach to computing the stochastic assignment that is the key output of the mechanism. In other words, we are the first to computationally operationalize the Hylland and Zeckhauser pseudomarket mechanism. Our empirical application shows that the algorithm works for problems on the order of the school choice market that we study.

A second cardinal mechanism that we study is an envy-free mechanism. Our envy-free mechanism is motivated by the cardinal mechanism developed for the more general course bundle problem by Nguyen, Peivandi and Vohra (2015). As we argue in detail below, the requirements on the allocation from the envy-free restriction are weaker than the requirements from the pseudomarket restriction. We also show that both sets of restrictions are sufficient for the "strategy proof in the large" incentive property in Azevedo and Budish (2019). Further, the weaker envy-free restriction could lead to a higher social welfare value, as we show. Finally, computationally the envy-free mechanism is a linear program and, for the same school choice problem, the envy-free mechanism is easier to implement than our algorithm for the pseudomarket mechanism.

While we know from Hylland and Zeckhauser that pseudomarkets are Pareto efficient under certain additional restrictions, another major contribution of our paper is to quantitatively study other measures of the welfare gains from using a cardinal mechanism over an ordinal mechanism.

To work with an empirically-founded set of preferences, we turn to school choice for ninth grade (the first year of high school) in the Seattle Public Schools. To estimate aspects of students' cardinal preferences over schools, we work with a setting where students are using an existing ordinal mechanism where students are facing lotteries.

Conventional approaches to estimating the distribution of student preferences rely on strong assumptions. One common assumption is that student preferences are independently distributed from students' families' locations and siblings' previous school attendance. This is unlikely to be the case in Seattle due to preferences across siblings and endogenous location choice. Instead, we estimate the identified set of cardinal preferences for each student consistent with that student's submitted rank ordered list and our estimates of the lotteries facing each student, based on enrollment data. The empirical approach of estimating an identified set of preferences for each agent was also used in the empirical work on random coefficients in demand models in Bajari and Benkard (2005). We describe our estimation approaches in more detail below.

We use these estimated cardinal preferences over high schools to compare the performance of three school choice mechanisms, two cardinal-preference mechanisms and one ordinal-preference mechanism. One cardinal mechanism is our variant of the pseudomarket of Hylland and Zeckhauser (1979). The

second is our simplification of the envy-free mechanism developed for course bundles by Nguyen, Peivandi and Vohra (2015). The ordinal mechanism is the so-called ordinally efficient mechanism of Bogomolnaia and Moulin (2001), which also gives a stochastic assignment. By comparing the pseudomarket and envy-free mechanisms to an ordinally efficient mechanism, we in some sense compare cardinal mechanisms to the best member of the class of ordinal mechanisms. All three mechanisms in our comparison are strategy proof in the large and our implementations assume truth-telling behavior by the students.[1]

We do not include the actual Seattle mechanism in the comparison because it lacks desirable theoretical properties, such as encouraging truthtelling in large markets. As stated before, not being strategy proof allows us to identify cardinal preferences from the lotteries facing students.

We measure the welfare gains from the pseudomarket mechanism over the probabilistic serial mechanism. Our headline results impose that the capacity of the school choice program for Seattle ninth grade is equal to the number of students who participated. We find that 71% of the students prefer the pseudomarket assignment over probabilistic serial, that 12% are indifferent up to numerical error between the two mechanisms, and that 17% prefer the assignment under probabilistic serial. Likewise, 67% of students prefer the envy-free assignment over probabilistic serial, 10% are indifferent, and 22% prefer probabilistic serial.

Further, we introduce a new gain measure for measuring the welfare gain from one mechanism over another. Our definition of the mechanism gain is given later in the paper. For now, we mention an interpretation. For a particular student, compute the mechanism gain measure using the Seattle ninth grade setting and say it is 0.05. Then an interpretation of the gain of 0.05 for this student is that the difference in expected utilities between the two mechanisms is equivalent to a move of 0.05 probability from the student's lowest-choice school to the student's highest choice school.

For the pseudomarket mechanism, the mean gain over probabilistic serial is under the exact capacity scenario for Seattle described above is 0.016, so equivalent to increasing a student's probability of attending his or her highest-choice school by 0.016 and decreasing the probability of his or her lowest-choice school by the same amount. For the envy-free mechanism, the mean gain over probabilistic serial is higher at 0.045. For comparison, a utilitarian benchmark without any incentives for student truthtelling has a 0.077 gain over probabilistic serial. Therefore, we find that the pseudomarket mechanism captures 21% of the possible gain over an ordinally efficient ordinal mechanism and the envy-free mechanism captures 58% of the possible gain. Note that our gain measure always compares the focal cardinal mechanism to the ordinal mechanism, not to a naive mechanism not involving school choice, unlike the comparisons of ordinal mechanisms to utilitarian benchmarks in Abdulkadiroglu, Agarwal and Pathak (2017).

Both cardinal mechanisms outperform probabilistic serial in terms of mechanism gain, intuitively

[1]The probabilistic serial mechanism is also exactly strategy proof in large enough markets (Kojima and Manea, 2010).

because cardinal mechanisms solicit more information than ordinal mechanisms. The envy-free mechanism imposes weaker restrictions than the pseudomarket mechanism, as we have alluded to and will discuss in more detail. These weaker restrictions explain its higher mechanism gain. We unpack these results and show results for other capacity and estimation scenarios in the second half of the paper.

We have mentioned several related papers already. There is a growing interest in theoretical work studying extensions to pseudomarkets imposing unit or bundle size demand constraints, including work such as Budish (2011), Budish, Che, Kojima and Milgrom (2013), He, Miralles, Pycia and Yan (2018) and Echenique, Miralles and Zhang (2020). Budish (2011) is particularly noteworthy as it is a pseudomarket that to our knowledge is used to assign students to course bundles at Wharton. Budish's mechanism has been implemented in a real setting at scale and evaluated in the lab (Budish and Kessler, 2016; Budish, Cachon, Kessler and Othman, 2015). Our understanding is that the mechanism uses ordinal preferences over course bundles rather than the cardinal preferences studied here. To our knowledge, no prior work on pseudomarkets imposing unit or bundle demand constraints has implemented any cardinal preference mechanism on the computer at scale and particularly not for the mechanism of Hylland and Zeckhauser (1979).

There is also theoretical work on pseudomarkets without imposing unit or bundle demand constraints, such as Pratt and Zeckhauser (1990) and much work in computer science. We know of less attention paid to our envy-free mechanism, which we view as a special case of the mechanism in Nguyen, Peivandi and Vohra (2015).

2 Our Pseudomarket Mechanism

Consider the classic problem of assigning at most (or sometimes, like in public school choice, exactly) one slot in a school to each student. We first outline the assignment problem and our pseudomarket mechanism. We then analyze the mechanism's theoretical properties. We finally use one example to explain why our mechanism differs in some details from the original pseudomarket mechanism in Hylland and Zeckhauser (1979).

2.1 Mechanism Definition

Let there be J schools, indexed by j. Let there be I students, indexed by i. The symbol J can also be the set of schools and the symbol I can also be the set of students. Each school j has a capacity q_j. We study the cases of exact capacity, where $\sum_{j \in J} q_j = I$, and excess capacity, where $\sum_{j \in J} q_j > I$.

The mechanism produces a stochastic assignment matrix $\pi = (\pi_{ij})_{i \in I, j \in J}$, where $\pi_{ij} \in [0, 1]$ is the probability that student i is assigned to school j. We restrict attention to feasible assignments, meaning the sum of assignment probabilities is less than or equal to capacities for each school, in notation $\sum_i \pi_{ij} \leq q_j \, \forall \, j$. Every feasible random assignment in this school choice setting can be decomposed into

a convex combination of deterministic assignments and can therefore be implemented as a lottery over deterministic assignments (Birkhoff 1946; von Neumann 1953). Code for the decomposition is freely available and we do not discuss implementing the assignment lottery further. We let $\pi_i = (\pi_{ij})_{j \in J}$ be the vector of assignment probabilities for student i, which we sometimes call the lottery facing student i. We impose $\sum_{j \in J} \pi_{ij} = 1$ for all students i.

In a cardinal-preference mechanism like a pseudomarket, students submit real number values v_{ij} for the utility of school j to student i. Let $v_i = (v_{ij})_{j \in J}$ be the vector of utilities for all J schools submitted by student i. Inside the mechanism, the utilities are used in expected value calculations. The expected value of a submission v_i of utilities for all J schools by student i is

$$\pi_i \cdot v_i = \sum_{j=1}^{J} \pi_{ij} v_{ij},$$

which simply multiplies each school's utility v_{ij} by the corresponding assignment probability π_{ij}. These utilities v_i are called cardinal preferences because the rank ordering over different lotteries $(\pi_{ij})_{j \in J}$ facing student i are preserved by a positive affine transformations of the submission $(v_{ij})_{j \in J}$. The pseudomarket mechanism we propose can be made to have the same stochastic assignment $\pi = (\pi_{ij})_{i \in I, j \in J}$ under any positive affine transformation of preferences, as we discuss below.

The exact pseudomarket mechanism that we implement produces a stochastic assignment matrix $\pi = (\pi_{ij})_{i \in I, j \in J}$ by solving the following mathematical program, taking as inputs the matrix of student utilities $v = (v_{ij})_{i \in I, j \in J}$ and the vector of school capacities $q = (q_j)_{j \in J}$:

$$\begin{aligned} &\max_{\pi, p} W(\pi, v) \text{ subject to} \\ &(\pi, p) \in \text{Equilibria}(v, q). \end{aligned} \tag{1}$$

Briefly, the mechanism maximizes a social welfare function $W(\pi, v)$ subject to the constraint that the stochastic assignment matrix π and a vector of school-specific prices $p = (p)_{j \in J}$ are one of the equilibria to a pseudomarket. In practice, if there are multiple global solutions (π, p) to (1), we pick the first solution that is found by our numerical algorithm.

Let us define these terms further. First, $W(\pi, v)$ is a social welfare function chosen by the school district that implements the mechanism. One social welfare function is the Nash product, or

$$W(\pi, v) = \prod_{i=1}^{I} (\pi_i \cdot v_i - \tilde{q} \cdot v_i), \tag{2}$$

where $\tilde{q} = \left(q_j / \sum_{k \in J} q_k\right)_{j \in J}$ is the vector of percentage capacities for all schools. Other social welfare

functions can be chosen, such as the sum of (non-logged) expected utilities

$$W(\pi, v) = \sum_{i=1}^{I} \frac{\pi_i \cdot v_i}{\max_j v_{ij} - \min_j v_{ij}}. \tag{3}$$

By inspection, one can verify that an assignment matrix π maximizing both these welfare functions will be invariant to student-specific positive affine transformations.

The term Equilibria (v, q) is the set of competitive equilibria (π, p) to a pseudomarket with a utility matrix of $v = (v_{ij})_{i \in I, j \in J}$ and a vector of capacities $q = (q_j)_{j \in J}$. A competitive equilibrium is a pair (π, p) such that:

1. The assignment matrix π is feasible: $\sum_i \pi_{ij} \leq q_j \, \forall \, j \in J$.

2. At the price vector p, the assignment vector π_i for each student i maximizes that student's expected utility subject to a unit budget constraint:

$$\pi_i \in \arg\max_{\tilde{\pi}_i} (\tilde{\pi}_i \cdot v_i) \tag{4}$$
$$\text{subject to } \tilde{\pi}_i \cdot p \leq 1.$$

Each student's shopping problem is a linear program. Note that we do not impose a "market clearing" condition; such a condition for example might rule out setting some prices to be positive infinity. The social welfare function in practice prevents such an equilibrium from being chosen by the mechanism.

The set of competitive equilibria Equilibria (v, q) is invariant to positive affine transformations. To see this, because the objective function in the individual consumer maximization problem (4) is linear in both the choice π_i and the utilities v_i, it is straightforward to see that the set of solutions to (4) is invariant to a positive affine transformation for student i.

In general, the computation of the mechanism occurs on a school district computer, away from students. Students collectively submit the matrix of utilities v and receive the stochastic assignment π or the deterministic assignment using the final lottery mentioned above. The school district could choose to disclose the computed price vector p in the interests of transparency. Immorlica, Lucier, Mollner and Weyl (2017) propose a raffles mechanism that is more interactive for students than the pseudomarket mechanism defined here.

2.2 Properties of the Pseudomarket Mechanism

2.2.1 Number of Solutions to the Mechanism

The pseudomarket mechanism is a solution to the mathematical program (1). Given the mechanism, it is desirable to know whether there exists at least one solution to the mechanism, so that in part

the set of competitive equilibria Equilibria (v, q) is non-empty.

Proposition 1. *Let the following two conditions hold:*

1. *The social welfare function $W(\pi, v)$ can be written as $\bar{W}(\pi_1 \cdot v_1, \ldots, \pi_I \cdot v_I)$ such that the function $\bar{W}$ is continuous in all I student expected utilities $\pi_i \cdot v_i$.*
2. *There exists a constant $\epsilon > 0$ such that the mechanism restricts attention to competitive equilibria such that the lowest price for a school is bounded above by $1 - \epsilon$ and all school prices are bounded above by a positive constant.*

Then the mathematical program (1) augmented with the conditions above has at least one solution.

In our computational implementation, we do not impose the restriction on prices in the second condition of our proposition because we have never found it to change the stochastic assignment π.

It would be also desirable to know that there is only one solution to the mechanism. With the Nash product social welfare function (2), we have found somewhat stylized examples with multiple stochastic assignments π that are solutions to the same mechanism. We conjecture based on computational examples that all equilibria maximizing the Nash product have the same profile of expected utilities and that there is a unique equilibrium with probability 1 if student utilities v_i are independently sampled from a continuous distribution.

2.2.2 Pareto Efficiency

Pareto efficiency is one of the key properties of an assignment π from the pseudomarket mechanism in Hylland and Zeckhauser (1979). In simulations, our mechanism achieves Pareto efficiency if a school district picks a social welfare function that is strictly increasing in the expected utilities of students. We do not have a formal theoretical result.

Hylland and Zeckhauser impose additional assumptions in order to formally prove that equilibria are Pareto efficient. For the case of exact capacities, Hylland and Zeckhauser prove that equilibria are Pareto efficient when (HZ1) at least one school has zero price and (HZ2) a student with multiple utility maximizing lotteries π_i picks one with the lowest cost. In the case of excess capacity, a supplemental appendix further shows that an equilibrium is Pareto efficient provided HZ2 holds and (HZ3) all under-demanded schools have a price of zero. We could impose these zero price and least cost conditions to prove that our mechanism's stochastic assignment is Pareto efficient, but we have found Pareto efficiency holds for our mechanism without imposing any of HZ1–HZ3 in all examples we have considered. We discuss the conditions HZ1–HZ3 further in Section 2.3.

2.2.3 Incentive Properties

The pseudomarket mechanism is not strategy proof because some students may find it profitable to misreport their preferences (Hylland and Zeckhauser, 1979). However, this incentive to misreport is

not a concern when the market grows larger in sense that the number of students and the capacity of each school increase proportionally while the number of schools is constant. We rely on two results in the literature.

First, in a sequence of regular replica economies, He, Miralles, Pycia and Yan (2018, Theorem 2) show that any student's incentive to misreport vanishes.[2] Second, Azevedo and Budish (2019, online appendix) prove that the pseudomarket mechanism is strategy proof in the large.[3]

In the following, we assume that students report their cardinal preferences truthfully in the pseudomarket mechanism. Note that in our application, we do have a large market that includes all students participating in school choice for Seattle's ninth grade.

2.3 Example of Equilibrium Selection

The main difference between our approach and that in Hylland and Zeckhauser (1979) is our emphasis on equilibrium selection. Hylland and Zeckhauser restrict attention to equilibria such that the requirements HZ1 and HZ2 mentioned before are satisfied for the case of exact capacity and HZ1 and HZ3 are satisfied for the case of excess capacity.

We do not impose the above restrictions and instead rely on social welfare function maximization to both select an equilibrium (or one of perhaps a handful if multiple equilibria maximize the same welfare function) and ensure Pareto efficiency. The following example shows that requiring prices for under-demanded schools to be equal to zero can result in a lower achievable social welfare value.

Consider an example with 3 students and 4 schools, each with one slot:

$$v = \begin{bmatrix} & j_1 & j_2 & j_3 & j_4 \\ i_1 & 3 & 1 & 2 & 4 \\ i_2 & 1 & 3 & 2 & 5 \\ i_1 & 2 & 4 & 5 & 1 \end{bmatrix}, \quad q = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

We compute four distinct equilibria by i) maximizing the Nash product (2), ii) maximizing the sum of expected utilities, iii) maximizing the sum of expected utilities subject to an additional constraint

[2] Regularity in their setting resembles a concept common in the general equilibrium literature (e.g., Dierker, 1974; Hildenbrand, 1974; Jackson, 1992). Also, He et al. (2018, Theorem 2) states that "[t]here always exists a selection of PM prices" that satisfies their definition of asymptotic strategy profess. However, their proof on page 295 only requires that the equilibrium selection does not depend on market size. In a sequence of replica economies, our selection rule, which is a social welfare function, is indeed independent of market size.

[3] Informally, strategy proof in the large for our mechanism means that, as the market grows larger, each student is more likely to take the prices as given and therefore have a dominant strategy to report truthfully. Their result requires the school valuation vectors v_i to take on finite support and also that there is an equilibrium selection rule such that students who report the same utilities v_i receive the same stochastic lottery over schools π_i. We can impose that students who report the same utilities v_i receive the same stochastic lottery over schools π_i as an extra constraint in the mathematical program (1). In the case where there could be multiple stochastic assignments π solving the mathematical program, we need a further equilibrium selection rule to fully define the mechanism in the sense of the definition of a mechanism in Azevedo and Budish.

that prices of under-demanded schools are zero and iv) minimizing the sum of expected utilities. The following table lists equilibrium prices and utility levels. The equilibrium assignments π are given in Appendix B.

Equilibrium	p_1	p_2	p_3	p_4	$\pi_1 \cdot v_1$	$\pi_2 \cdot v_2$	$\pi_3 \cdot v_3$	$\sum_{i \in I} v_i \cdot \pi_i$	$\sum_{i \in I} \log(v_i \cdot \pi_i - v_i \cdot \tilde{q})$
i)	$\frac{82877}{85822}$	$\frac{32531}{39010}$	1	$\frac{8391}{7802}$	$\frac{53}{16}$	$\frac{35}{8}$	5	$\frac{203}{16}$	$\log\left(\frac{169}{64}\right)$
ii)	$\frac{28436}{29001}$	$\frac{7972}{9667}$	1	$\frac{10232}{9667}$	$\frac{13}{4}$	$\frac{9}{2}$	5	$\frac{51}{4}$	$\log\left(\frac{21}{8}\right)$
iii)	0	0	1	2	$\frac{7}{2}$	4	5	$\frac{25}{2}$	$\log(\frac{5}{2})$
iv)	0	$\frac{11}{10}$	$\frac{33}{20}$	$\frac{11}{5}$	$\frac{38}{11}$	$\frac{31}{11}$	$\frac{42}{11}$	$\frac{111}{11}$	$\log\left(\frac{567}{10648}\right)$

By definition, equilibrium i) has the highest Nash product and equilibrium ii) has the highest sum of expected utilities. The equilibria i), ii) and iii) are Pareto efficient, while equilibrium iv) is not, as it is Pareto dominated by iii). Prices in equilibrium iv) are weakly greater than prices in equilibrium iii), causing students to shift demand to the less desirable school 1, as shown in Appendix B. Equilibria i) and ii) assign a positive price to each school and hence fail the requirement for under-demanded schools to have zero prices. Moreover, the utility levels in equilibrium ii) cannot be attained in any equilibrium imposing zero prices for under-demanded schools.

Equilibrium restrictions HZ1, HZ2 and HZ3 proposed in Hylland and Zeckhauser are attractive because they provide simple sufficient conditions for Pareto efficiency. However, imposing these restrictions rules out some Pareto efficient utility profiles, sometimes results in lower social welfare to the school district, and finally makes the numerical problem of finding a mechanism solution more challenging.

All four equilibria listed above fall into the set of equilibria considered in Proposition 1. In contrast, only equilibrium iii) meets the requirements in Hylland and Zeckhauser. The inefficient equilibrium iv) satisfies both restrictions HZ1 and HZ2; however, under-demanded schools 3 and 4 have positive prices and so HZ3 is not satisfied.

2.4 Numerical Algorithm for Mechanism

While there has been much recent theoretical research into pseudomarkets, to our knowledge (at the time of the initial distribution of our paper) no article in economics or computer science has shown how to compute an equilibrium to a pseudomarket, like Hylland and Zeckhauser (1979), with more than say ten students where each student is assigned to at most one school.[4] Our reading is that the computer science literature has focused more on the problem of finding a Competitive Equilibrium with Equal Incomes when there is no unit demand constraint, meaning a student could attend multiple

[4]Vazirani and Yannakakis (2020) prove that the problem is in the complexity class FIXP. Hosseini and Vazirani (2021) state that "Although HZ is based on the attractive idea of using a pricing mechanism, from the viewpoint of use in applications, it has a serious drawback, namely lack of computational efficiency, due to which solving instances of size even 4 or 5 is difficult."

schools at once (e.g., Pratt and Zeckhauser, 1990; Aziz, 2015).

Our pseudomarket mechanism is the mathematical program (1), where the definition of a competitive equilibrium includes consumer utility maximization in (4). Combining the main mathematical program with the definition of a competitive equilibrium gives:

$$\max_{\pi,p} W(\pi, v) \text{ subject to}$$

$$\sum_i \pi_{ij} \leq q_j \, \forall \, j \in J \text{ (capacity)}$$

$$\pi_i \in \arg\max_{\tilde{\pi}_i} (\tilde{\pi}_i \cdot v_i) \text{ (utility maximization)}$$

$$\text{subject to } \tilde{\pi}_i \cdot p \leq 1 \, \forall \, i \in I. \tag{5}$$

We make the constraints that each $\pi_{ij} \in [0, 1]$ and that $\sum_{j \in J} \pi_{ij} = 1$ implicit, although of course they are imposed.

This mathematical program is an example of a bilevel optimization problem, where there is an outer maximization problem subject to one or more inner maximization problems. The Stackelberg leader/follower model is one example of a bilevel optimization problem. In our mechanism's case, there are as many inner optimization problems as students, I. We consider applications like school choice with many students, so solving this mathematical program is more difficult than examples often found in methodological papers on solving bilevel optimization problems, which often have one or two inner optimization problems.

Sinha, Malo and Deb (2018) and Schmidt and Beck (2021) are recent reviews of methods for solving bilevel optimization methods. While we tried several alternatives, in the end the method of single level reduction had subjectively the best combination of speed and reliability of finding a solution. The method of single level reduction transforms the mathematical program (5) into the problem of maximizing the welfare function subject to the capacity constraints and the Karush-Kuhn-Tucker (KKT) constraints for the I student-specific utility maximization problems. Because the student problems are linear programs, Slater's condition states that satisfying the KKT conditions are necessary and sufficient for a proposed solution π_i to maximize student utility. Note also that the KKT conditions are robust to the existence of multiple solutions π_i for a particular student, given a price vector. An alternative where we attempted to compute the demand correspondence for student i would suffer because computing a correspondence can be harder and because the correspondence may be discontinuous in price.

The mathematical program with the KKT conditions is described in full detail in Appendix C. The KKT conditions include $J+1$ complementarity constraints for each student. This makes the optimization problem in the appendix a mathematical program with complementarity constraints (MPCCs)

Table 1: Percentage of Pseudomarket Starting Values Where KNITRO Converges for Synthetic Examples

# students	5 schools	10 schools	20 schools
100	0.11	0.30	0.21
500	0.10	0.03	0.09
1000	0.07	0.025	0.06
1500	0.06	0.08	0.01
2000	0.09	0.05	0.08

where there are $I \cdot (J+1)$ complementarity constraints. There is a separate literature on solving MPCCs (e.g., Hoheisel, Kanzow and Schwartz, 2013). While there are specialized packages such as PATH for solving MPCCs, our experiments found that the interior point/barrier solver KNITRO performed the best, as confirmed for MPCCs in the online benchmarks of Mittelman (2021). The algorithm KNITRO uses to solve MPCCs is described in Leyffer, López-Calva and Nocedal (2006).

Our computational strategy is to try KNITRO with many starting values for each assignment problem. In a synthetic example, we used 160 starting values for KNITRO for each of several combinations of the number of students, N, and the number of schools, M. Table 1 shows the percentage of these 160 runs for each market structure that converge. While most starting values do not converge, enough do to make this approach feasible. Figure 1 shows the mean time KNITRO takes on a Linux server for to finish for these same runs, broken out separately for runs where KNITRO reports convergence and runs where it reports non-convergence. A starting value that leads to convergence will converge quickly, while KNITRO may take a long time to report non-convergence. One recommendation is to use many starting values and terminate runs after a certain number of iterations, as runs taking a long time are less likely to converge.

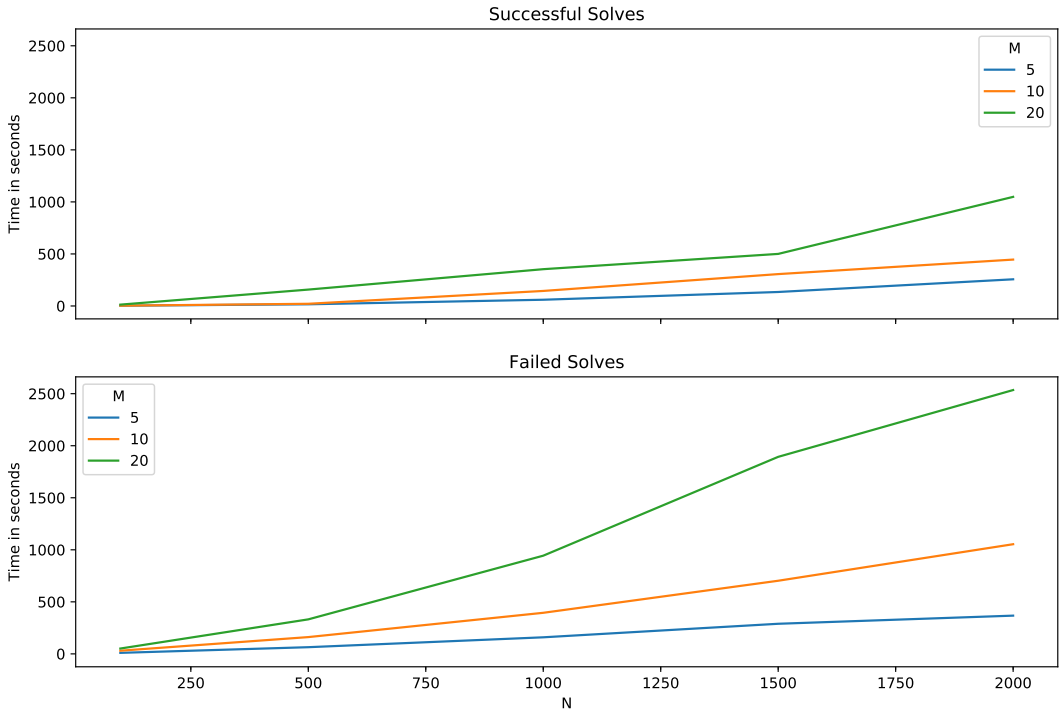
3 Envy-Free Mechanism

The second cardinal mechanism that we consider is an envy-free mechanism that is a special case of the mechanism for college students picking bundles of courses in Nguyen, Peivandi and Vohra (2015). Of course, picking a single school is an equivalent assignment problem to picking a bundle of only a single course. Many contributions in Nguyen et al. are specific to bundles of two or more courses and are not needed for the assignment one school to each student.

The mechanism is to pick an assignment matrix π to maximize social welfare, as discussed earlier, subject to envy-freeness constraints, or

$$\begin{aligned} &\max_{\pi} W(\pi, v) \text{ subject to} \\ &\pi \in \text{Envy Freeness}(v). \end{aligned} \tag{6}$$

Figure 1: Time Needed for Converged and Non-Converged KNITRO Runs for Pseudomarket



Note that there are no prices and capacities are used only to restrict assignment matrices π. The envy-freeness constraint is satisfied when, for all pairs of distinct students i, j, i does not envy j's assignment lottery, meaning the expected utility of i's own bundle is higher to i than i's expected utility for j's bundle. In notation, i not envying j means that

$$\pi_i \cdot v_i \geq \pi_j \cdot v_i. \tag{7}$$

Nguyen et al. presumably impose envy freeness in part because Azevedo and Budish (2019) show it is one of several sufficient conditions for a mechanism to be strategy proof in the large, a property also shared by the pseudomarket mechanism. We impose envy freeness for the same incentive properties.

Indeed, one property of pseudomarket equilibria is that they are envy free. With equal budgets for all students, each student can afford the bundle assigned to any other student by the mechanism. By individual utility maximization, each student prefers his or own her own lottery to those assigned to other students. The pseudomarket is also more restrictive than envy freeness as a student does not prefer any other affordable lottery, rather than just the lotteries assigned to other students in the same school district. The constraints in the envy-freeness mathematical program are therefore weaker than the constraints in the pseudomarket. Using weaker constraints means that the envy-free mechanism will produce weakly higher social welfare than the pseudomarket.

We implement the envy-free mechanism using the sum of expected utilities, (3), as the social welfare function. With the envy-freeness constraints (7) also being linear in π, (6) is a linear program. The main computational concern is that there are $N \cdot (N-1)$ envy-free constraints (7), where recall N is the number of students. Still, linear programs are easier to solve than MPCCs and we believe that the envy-free mechanism can be used for even the largest school districts in the United States, like New York City.[5] Only one run of a linear programming solver like Gurobi is typically needed for each assignment problem; global convergence happens each run.

4 Estimation of Cardinal Preferences

The goal of the remainder of the paper is to measure the welfare gains from using the two cardinal-preferences mechanisms instead of the ordinal-preference mechanism called probabilistic serial. Rather than exploring only arbitrarily chosen preferences in an artificial situation, we estimate the cardinal utilities of real students applying to ninth grade, the first year of high school, in the Seattle Public Schools.

Cardinal preferences are preferences over lotteries and so we need students facing lotteries over schools in order to have a hope at estimating cardinal utilities over schools. We sought out data from

[5] One could drop many of the $N \cdot (N-1)$ envy-free constraints and presumably get around the same solution, as long as many constraints are still kept.

Seattle because it uses a variant of the Boston mechanism and so the mechanism is not strategy proof. We do not seek to analyze the Seattle mechanism; rather we use data from it to identify cardinal preferences in order to evaluate the pseudomarket and probabilistic serial mechanisms using cardinal preferences estimated from field data.

There are 11 high schools in Seattle participating in school choice and 964 students participated in the year we have data for, 2018. Participation is voluntary; if a student prefers his or her assigned default school than there is no need to participate in school choice.

4.1 Estimating Assignment Probabilities in Seattle

In Seattle, and as in other ordinal school choice mechanisms, current eighth grade students submit ranked ordered lists of schools m_i, in this case before the winter holidays. Students can have priorities for various schools based on geographic residence and sibling status, which our data measure and we call x_i. Our data then show enrollment in ninth grade for these students. Let $y_{ij} = 1$ if student i is enrolled in school j in ninth grade and let $y_i = (y_{ij})_{j \in J}$.

We use the data on (m_i, x_i, y_i) to estimate the statistical model

$$\Pr\left(y_{ij} = 1 \mid x_i, m_i\right) \approx h_j\left(x_i, m_i\right).$$

We estimate this classification problem (school enrollment is a discrete outcome) using a random forest method, with several tuning parameters chosen by cross validation to mitigate overfitting. We then assume our estimated probabilities $h\left(x_i, m_i\right)$ are the probabilities that a student with priorities x_i would use when picking a ranked ordered list m_i to submit to the Seattle mechanism. In the language of the pseudomarket mechanism, $h\left(x_i, m_i\right) = \left(h_j\left(x_i, m_i\right)\right)_{j \in J}$ is the estimate of the lottery π_i facing student i from the Seattle mechanism.[6]

This approach of using ninth grade enrollment data to estimate the probabilities that a student faces when submitting a list of schools allows for uncertainty from events after the formal school choice mechanism is run, like a family moving within the district and a student appealing his or her initial assignment. It does not allow the student to have more information than the researcher or school district at the time the list is submitted. If at the time of list submission a student knew he or she would move and get different geographic priorities, then $h\left(x_i, m_i\right)$ would not account for the student's private information.

Our baseline approach to estimating preferences assumes that students know the probabilities $h\left(x_i, m\right)$ for each submission m. In a robustness exploration, we make weaker assumptions in estimation. Relaxing the knowledge of assignment probabilities has also been done in other work on mechanisms that are not strategy proof (e.g., Agarwal and Somaini, 2018; Kapor et al., 2020).

[6] Other papers in the school choice literature estimate the lotteries facing students directly from the data (e.g., Agarwal and Somaini, 2018; Kapor, Neilson and Zimmerman, 2020).

Our coming preference estimates depend on our estimates of $h(x_i, m_i)$ and different techniques used in the literature would result in new estimates of $h(x_i, m_i)$ and hence cardinal preferences. The by now common alternative approach of computing $h(x_i, m_i)$ using formal simulations of Seattle's mechanism is not feasible as the mechanism is run on a computer by the district's information technology department and the IT department was not willing to share its code or the formal definition of the mechanism, which is similar but not identical to the historic Boston mechanism.

4.2 Partial Identification of Student Utilities for Schools

We need cardinal preferences over schools in order to evaluate expected utility calculations in the pseudomarket and to compare the welfare gains between the pseudomarket and probabilistic serial mechanisms. In Seattle, our baseline approach assumes students submit a ranked ordered list m_i that maximizes expected utility

$$v_i \cdot h(x_i, m) = \sum_j v_i h_j(x_i, m).$$

If m_i is the submitted list and m another, arbitrary list of schools, list m_i must have given a weakly higher expected utility:

$$v_i \cdot h(x_i, m_i) \geq v_i \cdot h(x_i, m).$$

Let a choice situation be given by M possible lists m and a corresponding matrix of probabilities $h^i = (h_j(x_i, m))_{j \in J, m \in M}$, one column for each submission m. Say the researcher has data on a cross section of matrices h^i and chosen lists m_i.

A more traditional approach of point identifying a random utility model relies on two strong assumptions: the lotteries facing students vary continuously in a small set and students' preferences for schools are distributed independently from the lottery h_i.[7] These two assumptions are not well satisfied in the Seattle data. On the lack of continuous variation in lotteries, our priority measures x_i of neighborhood and sibling enrollment are discrete. Further, there are two key reasons why independence may not hold. First, sibling preferences are part of x_i and likely correlated with v_i, as a parent with strong preferences for a particular school for an older sibling likely has strong preferences for the same school for a younger sibling. For example, the school could be near the parent's commuting route. Second, school preferences likely play a role in choice of residence, so that those students who live nearer a school have a higher preference for it.

For these reasons, we adopt a partial identification approach that removes the need for continuous variation in lotteries and independence between the menu of lotteries and preferences. For each student i, the identified set of preferences V_i is the set of school utilities where the submitted list m_i gives

[7] Agarwal and Somaini (2018, Theorem D.1) point identify the probability of certain sets of school utilities using a finite number of lottery matrices h_i.

higher expected utility than all other submitted lists m in the set of all such lists, M. In other words,

$$V_i = \{v_i \mid v_i \cdot h(x_i, m_i) \geq v_i \cdot h(x_i, m) \; \forall \, m \in M\}.$$

The identified set of preferences for student i is hard to further characterize on a computer. One challenge is the high number of possible lists in M. For each student, we pick a subset of lists $\tilde{M}_i \subseteq M$ based on our intuition of what counterfactuals lists will shrink the non-sharp identified set for student i,

$$\tilde{V}_i = \left\{v_i \mid v_i \cdot h(x_i, m_i) \geq v_i \cdot h(x_i, m) \; \forall \, m \in \tilde{M}_i\right\}. \tag{8}$$

We describe how we pick the counterfactual lists $\tilde{M}_i$ in Appendix D.

During our mechanism simulations, we follow the approach in the empirical work of Bajari and Benkard (2005) on the demand for computers and impose a uniform distribution over each $\tilde{V}_i$. In Appendix D, we propose a rejection-based MCMC sampler that allows us to draw from the uniform distribution over each $\tilde{V}_i$ (e.g., Bélisle, Boneh and Caron, 1998). The main challenge in our implementation of the sampler is imposing our location normalization that $\min_j v_{ij} = 0$ and our scale normalization that $\sum_j v_{ij} = 1 \, \forall \, i$. For each replication of all of the pseudomarket, envy-free and probabilistic serial mechanisms, we draw one vector of school utilities v_i for each student i from $\tilde{V}_i$, using the sampler.

In a robustness exploration, we use rules other than a uniform distribution to take values from the revealed-preference set $\tilde{V}_i$.

4.3 Weaker Preference Restriction

As mentioned previously, the assumption that the students know the lotteries $h(x_i, m_i)$ is controversial in the literature. If students, do not know $h(x_i, m_i)$, then they cannot compute their own expected utilities and even continuous variation in lotteries facing students would likely not lead to the point identification of a distribution of cardinal utilities for schools.

Nevertheless, the partial identification approach can be extended to weaker assumptions, resulting in a larger identified set for each student. Recall that each student is assigned to a default school based on his or her residence. While there are many models of behavior other than expected utility maximization, we use the assumption that all schools listed on the submitted list are preferable to the default school and all schools not listed are not preferred to the default school. We only use students for which the maximum length of the submitted list, five, is not used. This weaker preference restriction only informs us directly about ordinal preferences. Combining this restriction with the previous sampling procedure means that we sample cardinal preferences, with appropriate location and scale normalizations, uniformly from the set of cardinal preferences consistent with the restriction that all schools listed on the submitted list are preferable to the default school and all schools not listed are

Table 2: Seattle High Schools, Submitted Preferences, Capacities and Estimated Top School Preferences for One Replication

High School	# **Submitted First Choice**	# **Estimated Top Utility**	**Capacity** (**Under Exact**)
Ballard	51	66	79
Cleveland	243	194	233
Franklin	55	82	72
Garfield	85	73	58
Roosevelt	120	161	84
Chief Sealth	17	24	31
West Seattle	50	49	44
Ingraham	198	173	165
Ranier Beach	8	10	47
Nathan Hale	54	58	84
Center	83	74	66

not preferred to the default school.

4.4 Estimates of Preferences for Seattle

Table 2 lists the public high schools in Seattle. For each school, the first column is the number of students who listed that school as their first choice in their submissions to the school choice program. Cleveland is the school most often listed first, while Ranier Beach was rarely listed first.

In a non-strategy proof mechanism, sophisticated students might not submit their true first choices to the mechanism. In the second column, we use the estimation method described above to take one uniform draw from the revealed preference set of school utilities for each student, under the baseline assumption that students know the lotteries they are facing in the Seattle mechanism. The second column of Table 2 reports the number of students who have each school as their true top preference under these estimates. Schools like Cleveland and Ingraham are less popular than the submitted lists indicate, which is explained in terms of the model by students using the lottery information to list schools with a higher chance of admittance.

The third column of Table 2 lists the capacities of the school for the school choice program that we use in our mechanism comparisons. The capacities are set proportional to the fall enrollment for students participating in school choice. Under the exact capacity scenario in the table, the total number of seats in the school choice program is equal to the number of students participating in school choice. Cleveland has a higher capacity than the number of students who are estimated to have it first in their preferences, consistent with the larger number of students listing Cleveland first being driven by strategic behavior.

5 Mechanism Gain Measure

We will compare the pseudomarket and envy-free cardinal mechanisms to the ordinal mechanism probabilistic serial. Before presenting our comparisons, we introduce a new mechanism gain measure for the purposes of comparing the outcomes of mechanisms.

5.1 Gain Definition

For concreteness, we focus on the comparison of the pseudomarket cardinal mechanism, called PM, to the probabilistic serial ordinal mechanism, called PS. Then let π_i^{PM} be the pseudomarket assignment (probabilities for each school) for student i and similarly let π_i^{PS} be the probabilistic serial assignment for student i. Then the mechanism gain to student i of the pseudomarket over the probabilistic serial mechanism is defined to be

$$\frac{v_i \cdot (\pi_i^{\text{PM}} - \pi_i^{\text{PS}})}{\max_j v_{ij} - \min_j v_{ij}}.$$

The numerator of the gain measure is the difference in expected utilities between the two mechanisms, recalling that v_i is the vector of school utilities for student i. The denominator is the difference in the utilities of the most and least preferred schools for student i.

Including the denominator makes the gain measure invariant to a positive affine transformation of student i's school utilities. Including the denominator allows different papers with different normalizations on student utilities to report the gain measure in comparable units. If a particular paper is happy with its own choice of normalizations and wants to work with expected utilities directly, the denominator can be dropped.

We often highlight the mean gain calculated across students and across replications of the two mechanisms being compared, each replication using new school utility draws for each student. We also look into heterogeneity of gains across students for the same replication and how those gains are related to students' school utilities.

5.2 Equivalent Lottery Interpretation of Gain Measure

The gain measure has a convenient interpretation in terms of "equivalent lotteries" that involve only two schools, the top school for student i and the worst school for student i. First, we find the equivalent lottery over the the top school for student i and the worst school for student i that gives the same expected utility as the pseudomarket assignment for all the schools in the school district, all the high schools in Seattle in our application.

Let $\bar{\pi}_i^{(1,\text{PM})}$ be the probability that i is assigned to i's top school (top across the entire district) in this i-specific equivalent lottery giving i the same utility as the pseudomarket assignment in the district as a whole, and let $1 - \bar{\pi}_i^{(1,\text{PM})}$ be the probability that i is assigned to i's worst school in this equivalent lottery. Likewise, let $\bar{\pi}_i^{(1,\text{PS})}$ be the top-school probability for the equivalent lottery over

two schools giving the same expected utility as the probabilistic serial assignment for the entire school district.

Some simple algebra shows that

$$\frac{v_i \cdot \pi_i^{\mathrm{PM}}}{\max_j v_{ij} - \min_j v_{ij}} = \bar{\pi}_i^{(1,\mathrm{PM})}.$$

The same algebra applies to the probabilistic serial equivalent lottery. To relate the equivalent lotteries to the gain measure, consider that

$$\frac{v_i \cdot \pi_i^{\mathrm{PM}}}{\max_j v_{ij} - \min_j v_{ij}} - \frac{v_i \cdot \pi_i^{\mathrm{PS}}}{\max_j v_{ij} - \min_j v_{ij}} = \frac{v_i \cdot (\pi_i^{\mathrm{PM}} - \pi_i^{\mathrm{PS}})}{\max_j v_{ij} - \min_j v_{ij}} = \bar{\pi}_i^{(1,\mathrm{PM})} - \bar{\pi}_i^{(1,\mathrm{PS})}.$$

Therefore, the gain measure satisfies

$$\frac{v_i \cdot (\pi_i^{\mathrm{PM}} - \pi_i^{\mathrm{PS}})}{\max_j v_{ij} - \min_j v_{ij}} = \bar{\pi}_i^{(1,\mathrm{PM})} - \bar{\pi}_i^{(1,\mathrm{PS})}.$$

To understand the implied interpretation, say the gain measure is 0.05. In this case, a gain or normalized differences in expected utilities between the cardinal and ordinal mechanisms for student i of 0.05 is equal to an increase in top school probabilities of 0.05, when the needed 0.05 reduction occurs for the worst choice school.

5.3 Computational Bounds on the Gain Measure

Cardinal mechanisms solicit more information from students than ordinal mechanisms, as ordinal preferences can be constructed from cardinal preferences. Under truthtelling behavior, one might naively conjecture that cardinal mechanisms will always outperform ordinal mechanisms. We show computationally that this is not the case. The mean gain across students of the pseudomarket mechanism over the probabilistic serial mechanism can be both positive and negative and so it is an empirical issue requiring preference estimates from a real setting to resolve.

Consider a synthetic market with four students and four schools. Each school has one slot. The remaining primitives of the market are the student utilities for the schools. With four students and four schools, there are sixteen scalar school utilities. We numerically search over these sixteen school utilities to maximize the mean gain of the four students from the pseudomarket mechanism over the probabilistic serial mechanism. For each guess of the sixteen school utilities by an outer optimization routine, we compute a solution to both mechanisms and then report the gain measure to the outer optimization routine. We use a similar procedure to minimize the gain measure.

The upper bound on the gain is 0.29. The lower bound on the gain is -0.19. There is an asymmetry

in favor of the cardinal mechanism: the absolute value of the upper bound is higher than the absolute value of the lower bound. There is also a large difference of 0.48 between the upper and lower bounds. Using the interpretation in terms of equivalent lotteries, the difference in gains of 0.48 is equivalent to a move in probability from a student's worst school to the student's top school of 0.48. These results indicate that both the magnitude and sign of the gain of a cardinal mechanism over an ordinal mechanism are empirical matters requiring preference estimates.

5.4 No-Incentives Utilitarian Benchmark

Borrowing a concept from Abdulkadiroglu, Agarwal and Pathak (2017), we define the no-incentives utilitarian benchmark to be the assignment π that maximizes social welfare $W(\pi, v)$ subject to only feasibility constraints, like school capacity constraints. There are no pseudomarket or envy-free truthtelling incentives so this type of mechanism could not be implemented in practice. For a full benchmark, Abdulkadiroglu et al. compare this infeasible-in-practice mechanism to a naive approach of assigning each student to the closest school subject to capacity constraints. Those authors then calculate how much of the improvement in social welfare of the infeasible, no-incentive-constraint utilitarian approach over the naive assignment is achieved by ordinal mechanisms such as deferred acceptance.

We are interested in comparing cardinal mechanisms to ordinal mechanisms, not to the naive approach of assigning each student to the closest school. So we calculate the mechanism gain comparing the infeasible, no-incentive-constraint utilitarian "cardinal mechanism" to probabilistic serial. We view this as the maximum gain possible over ordinal mechanisms achievable by any cardinal mechanism. We then divide the gains for each of the pseudomarket and envy-free mechanisms, both compared to probabilistic serial, by this maximum gain from the no-incentive utilitarian benchmark. This tells us the percentage of the total possible gain that is achieved by the particular cardinal mechanism in question over ordinal mechanisms.

6 Mechanism Comparisons for Seattle Preferences

We now compare the two cardinal mechanisms to the ordinally-efficient ordinal mechanism probabilistic serial using estimated preferences from students in Seattle.

6.1 Baseline Results

Our baseline cases imposes that the total capacity in the school choice program is equal to the number of students participating in school choice. We also impose that students maximize expected utility with knowledge of the lotteries facing them in the Seattle mechanism. Finally, we take uniform draws

Table 3: Baseline Mechanism Comparisons

	Envy Free			Sum Expect Util PM			Nash Product PM		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Indifferent	0.10	0.08	0.13	0.12	0.09	0.14	0.12	0.10	0.15
Prefer Cardinal	0.67	0.63	0.71	0.71	0.67	0.75	0.71	0.66	0.75
Prefer Ordinal	0.22	0.17	0.27	0.17	0.13	0.21	0.18	0.13	0.22
Mean Gain	0.045	0.036	0.052	0.016	0.012	0.020	0.016	0.011	0.021
Gain No Indiff.	0.049	0.041	0.057	0.018	0.014	0.023	0.018	0.013	0.024
No-Incentives Gain	0.077	0.066	0.085		Same			Same	

of school utility vectors from the revealed-preference set of each student. We alter each of these assumptions soon.

Table 3 reports these numbers, which we view as the most important results in the paper. We run 400 replications, each corresponding to a new draw of the vector of school utilities for each student. For each replication, we run four mechanisms: probabilistic serial, our envy free mechanism, a pseudomarket with the sum of the expected utilities in (3) as the social welfare function, and a pseudomarket with the Nash product in (2) as the social welfare function. The first three columns compare the envy-free mechanism to the probabilistic serial ordinal mechanism, the second group of three columns compares the pseudomarket with the sum of the expected utilities welfare function to the ordinal mechanism, and the last three columns compare the pseudomarket with the Nash product to the ordinal mechanism. For each measure, the table reports the mean across students within a replication and then across replications, the minimum across replications of the mean-across-students within a replication, and the maximum across replications. Looking as the mins and maxes across replications, the variability of all the measures in the table across the 400 replications is not high.

Consider the comparison of the envy free mechanism to the probabilistic serial ordinal mechanism in Table 3. 10% of students are indifferent between the two mechanisms. These could include students who prefer unpopular schools, who might not be affected as much by mechanism choice. 67% of students prefer their assignments in the cardinal envy free mechanism over their assignments from the ordinal mechanism. The fraction preferring the cardinal assignment varies from a low of 63% to a high of 71% over the 400 replications.

The mechanism gain measure is 0.045, ranging from a min of 0.036 and a max of 0.052 across replications. The number 0.045 means that the mean student's expected utility improvement from the cardinal mechanism over the ordinal mechanism is equal to moving a probability of 0.045 from the student's least preferred school (out of eleven) to the student's most preferred school. This is one of the central results of the paper. Dividing the envy-free mechanism gain of 0.045 by the no-incentive utilitarian benchmark of 0.077 implies that the envy-free mechanism captures 58% of the gain that any cardinal mechanism could hope to have over an ordinally efficient ordinal mechanism, probabilistic serial.

Table 4: Baseline Mechanism Comparisons

	Envy Free			Sum Expect Util PM			Nash Product PM		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Indifferent	0.12	0.09	0.27	0.14	0.10	0.31	0.13	0.10	0.31
Prefer Cardinal	0.72	0.56	0.77	0.61	0.33	0.76	0.62	0.18	0.75
Prefer Ordinal	0.16	0.12	0.21	0.25	0.12	0.54	0.25	0.13	0.68
Mean Gain	0.032	0.023	0.038	0.006	-0.012	0.018	0.006	-0.032	0.017
Gain No Indiff.	0.036	0.029	0.042	0.007	-0.013	0.020	0.007	-0.037	0.017
No-Incentives Gain	0.047	0.037	0.056		Same			Same	

The table reports the mean gain excluding those students who are indifferent between the mechanisms, likely because they prefer schools that are unpopular. For the comparison of the envy-free and ordinal mechanisms, the mean gain excluding the indifferent is 0.049.

Still on Table 3, turn next to the comparison of the pseudomarket with the sum of expected utilities social welfare function to the same ordinal mechanism as before. Slightly more students are indifferent compared to the envy-free case, a few more students prefer the cardinal mechanism than in the envy free case (71% vs 67% for envy free), and fewer students therefore prefer the ordinal mechanism's assignments. However, the mean game of the pseudomarket over the ordinal mechanism, at 0.016, while still positive, is not as high as we saw with the envy-free mechanism. As before, the interpretation is of 0.016 is that the mean student's expected utility improvement from the cardinal mechanism over the ordinal mechanism is equal to moving a probability of 0.016 from the student's least preferred school (out of eleven) to the student's most preferred school. The magnitude of the decrease of the mean gain compared to envy-free is perhaps surprising but the direction of the change is not. As discussed before, the pseudomarket imposes stricter constraints on social welfare maximization than the envy-free mechanism. Dividing the mean gain of 0.016 by the benchmark gain of 0.077 means that the pseudomarket captures 21% of the possible gains from any cardinal mechanism.

Table 3 also shows that the pseudomarket with the Nash product social welfare function performs almost identically on the considered measures to the pseudomarket with the sum of expected utilities social welfare function.

6.2 Other Capacity and Preference Estimation Scenarios

6.2.1 Excess Capacity

Table 4 considers the same student preference estimation as the baseline case but increases all school capacities by 10% over the exact capacity case studied in Table 3. Otherwise the two tables have the same structure.

The last row of Table 4 is perhaps most instructive. The infeasible, no-incentives utilitarian benchmarks are lower than in Table 3. The mean gain for the benchmark is 0.047 in Table 4 and 0.077

in Table 3. With excess capacity, the room for any mechanism to outperform the ordinal mechanism has gone down. Given this context, the mean gain of envy free over the ordinal mechanism of 0.032 captures 0.032/0.047 or 68% of the possible gain from any mechanism based on cardinal utilities. By this measure, the performance of the pseudomarket under excess capacity is much less impressive, with a mean gain of 0.006 over the ordinal mechanism capturing 13% of the possible gains of any mechanism. Across the 400 replications, many of the mean-across-students gains are negative for the pseudomarket, while the lowest gain is 0.023 for the envy-free market. In this particular excess capacity scenario, the envy-free mechanism shows a big improvement over the ordinal mechanism while the improvement of the pseudomarket mechanism varies across replications and is smaller in magnitude, with the best gain of the 400 replications for the pseudomarket mechanism being smaller than the worst gain for the envy-free mechanism. Despite these findings using the gain measure, 72% of students prefer the envy-free mechanism to the ordinal mechanism and 61% or 62% of the students prefer the pseudomarket mechanism to the ordinal mechanism.

6.2.2 Weaker Preference Restriction in Estimation

Return to the case of exact capacities. As described in Section 4.3, we use a weaker preference restriction in estimation that does not require students to have knowledge of the lotteries generated by the existing Seattle mechanism. This results in a larger identified set of cardinal preferences. As before, we use a sampler to take new, uniformly distributed draws of school utility vectors for each of 400 replications of the four mechanisms being studied.

Table 5 presents the results of the mechanism comparisons using these larger identified sets of student preferences. Looking at the no-incentives utilitarian benchmark in the last row first, we see that at a mean gain of 0.033 over the ordinal mechanism, there is less scope for the best performing cardinal mechanism to improve on ordinal mechanism with these draws from larger preference sets. The benchmark gain of 0.033 is less than half of the value of the value of 0.077 for the baseline case in Table 3. Given this context, it is not surprising that the mean gains of the real cardinal mechanisms over the ordinal mechanism are lower by about the same fraction. The mean gain of the envy free mechanism over the ordinal mechanism is 0.022 and the mean gain of the pseudomarket mechanism is 0.009, over the ordinal mechanism. The mins and maxes in Table 5 indicates that all 400 replications for all cardinal mechanisms show a positive mean gain across students over the ordinal mechanisms. 79% of the students prefer the envy-free mechanism to the ordinal mechanism while 78% prefer the pseudomarket mechanisms to the ordinal mechanism.

6.2.3 Non-Uniform Distribution Approaches to Drawing School Utilities

Following the work on demand estimation by Bajari and Benkard (2005), the previous tables use a uniform distribution to draw school utility vectors from the student-specific revealed-preference sets

Table 5: Mechanism Comparisons for Weaker Preference Restrictions

	Envy Free			Sum Expect Util PM			Nash Product PM		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Indifferent	0.05	0.02	0.10	0.06	0.03	0.11	0.06	0.04	0.11
Prefer Cardinal	0.79	0.72	0.84	0.78	0.67	0.84	0.78	0.67	0.83
Prefer Ordinal	0.17	0.11	0.23	0.16	0.11	0.22	0.16	0.12	0.23
Mean Gain	0.022	0.019	0.027	0.009	0.006	0.012	0.009	0.006	0.012
Gain No Indiff.	0.024	0.020	0.029	0.010	0.006	0.012	0.034	0.006	0.013
No-Incentives Gain	0.033	0.028	0.038		Same			Same	

in order to perform the mechanism comparisons. Here we explore other approaches to taking school school utilities from the identified sets and perform model comparisons.

Across many draws of vectors of school utilities from a sampler, for each student we take the draw corresponding to the lowest "utility of the top school minus the lowest utility of the second choice school". Our preliminary findings indicate a much higher mean gains of all cardinal mechanisms over the ordinal mechanism using this approach. What is likely happening is that students are better able to substitute to their second choice school in this scenario, as they are nearly indifferent between the first and second choice schools.

Likewise, we constructed a market and by taking the draw corresponding to the highest, instead of lowest, "utility of the top school minus the utility of the second choice school". The performance of the mechanisms is quite similar to the baseline results in Table 3.

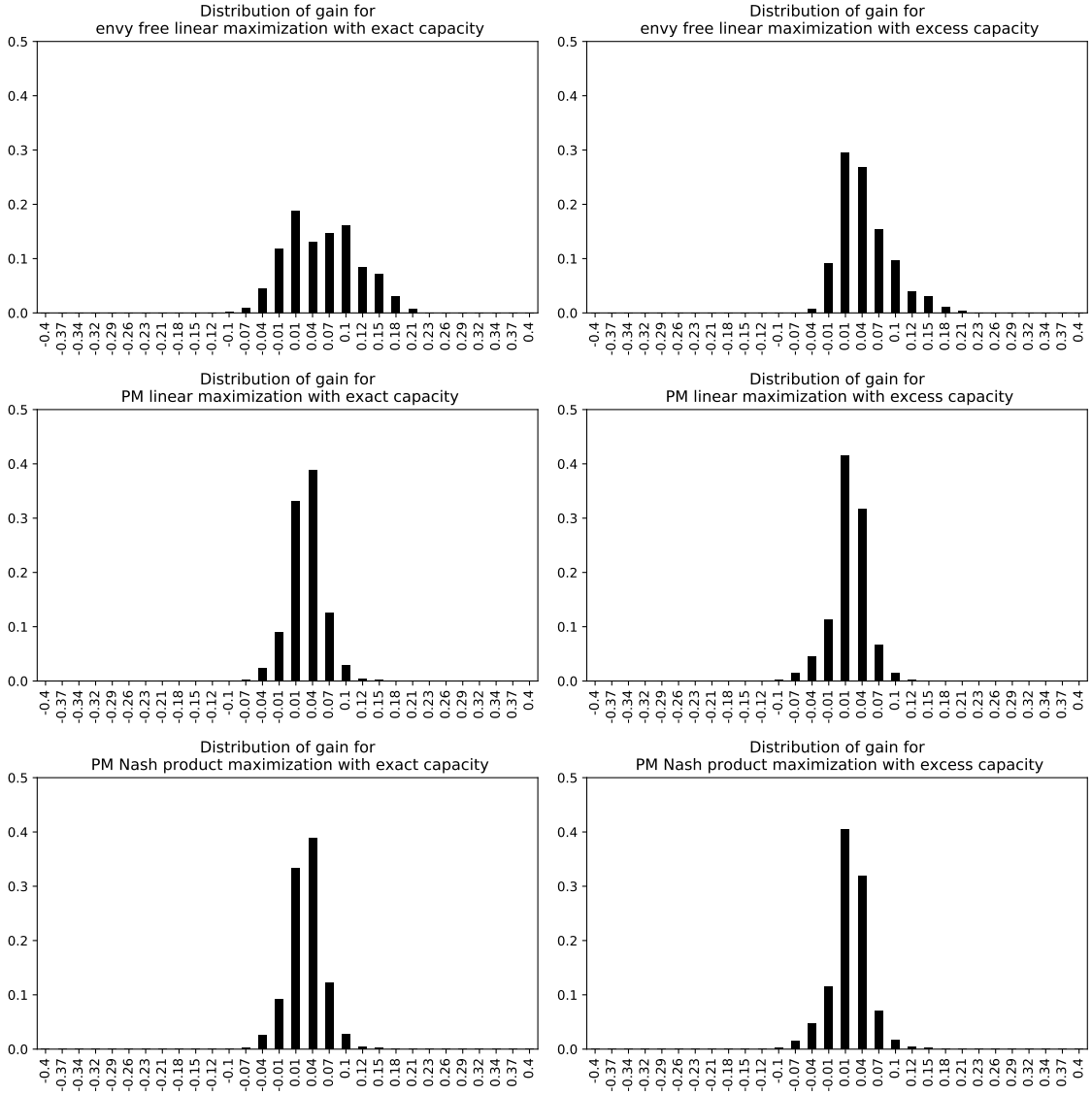
6.3 Student Heterogeneity in Gains

It is important to explore heterogeneity in the gains from students in the same replication. For this purpose, return to the baseline preferences results studied in Section 6.1 and the excess capacity extension using the same preference estimates in Section 6.2.1.

Figure 2 shows histograms of the gains of students from the various cardinal mechanisms considered over the probabilistic serial ordinal mechanism. There is a good deal of heterogeneity in the gain across students for all the six scenarios shown. More students gain than lose as shown earlier. However, there clearly are some losers who would prefer the ordinal assignment. Further, one can see that the envy-free mechanism has a thicker right tail than the psuedomarket mechanism. A noticeable fraction students quite strongly benefit from the envy free mechanism over the ordinal mechanism, more so than the comparison of the pseudomarket and ordinal mechanisms.

There is some intuition that students with a higher utility difference between their first and second most preferred schools will gain more from a cardinal mechanism compared to an ordinal mechanism. Let $v_{i(1)} = \max_j v_{ij}$ be the utility of the most preferred school for student i and let $v_{i(2)}$ the utility of i's second most preferred school. The utility difference or gap is $v_{i(1)} - v_{i(2)}$. For the purposes of

Figure 2: Histograms of Student Gains of Cardinal Over Ordinal Mechanisms



visual analysis, we sort students by the percentiles of the gaps, or the percentiles of $v_{i(1)} - v_{i(2)}$.

Figure 3 examines heterogeneity in the mechanism gain for students by the percentiles of $v_{i(1)} - v_{i(2)}$. For each percentile gap on the horizontal, the corresponding student is identified in each of the 400 replications and the gain of those student from a cardinal mechanism over the ordinal mechanism is averaged over the replications. The vertical axes in Figure 3 translate these mean gains into percentiles using the entire set of students and their gains for reference.

For all cardinal mechanism comparisons to the ordinal mechanism and all capacity scenarios, the panels in Figure 3 tell the same story. There is a flat trend in the level of the mechanism gain up until the 60th percentile or so of the students' utility differences between their top two schools. The mechanism gain starts increasing beyond the 60th percentile of the utility differences between schools and slope is increasing with the percentile as well. So the students with the highest utility differences between their top two choices on average have quite high percentile gains from participating in a cardinal mechanism over the probabilistic serial ordinal mechanism.

7 Conclusion

Cardinal mechanisms solicit more information on preferences than ordinal mechanisms. Therefore, cardinal mechanisms may outperform ordinal mechanisms on welfare metrics. This paper measures the extent of the outperformance using estimated cardinal preferences from a real setting, Seattle ninth grade.

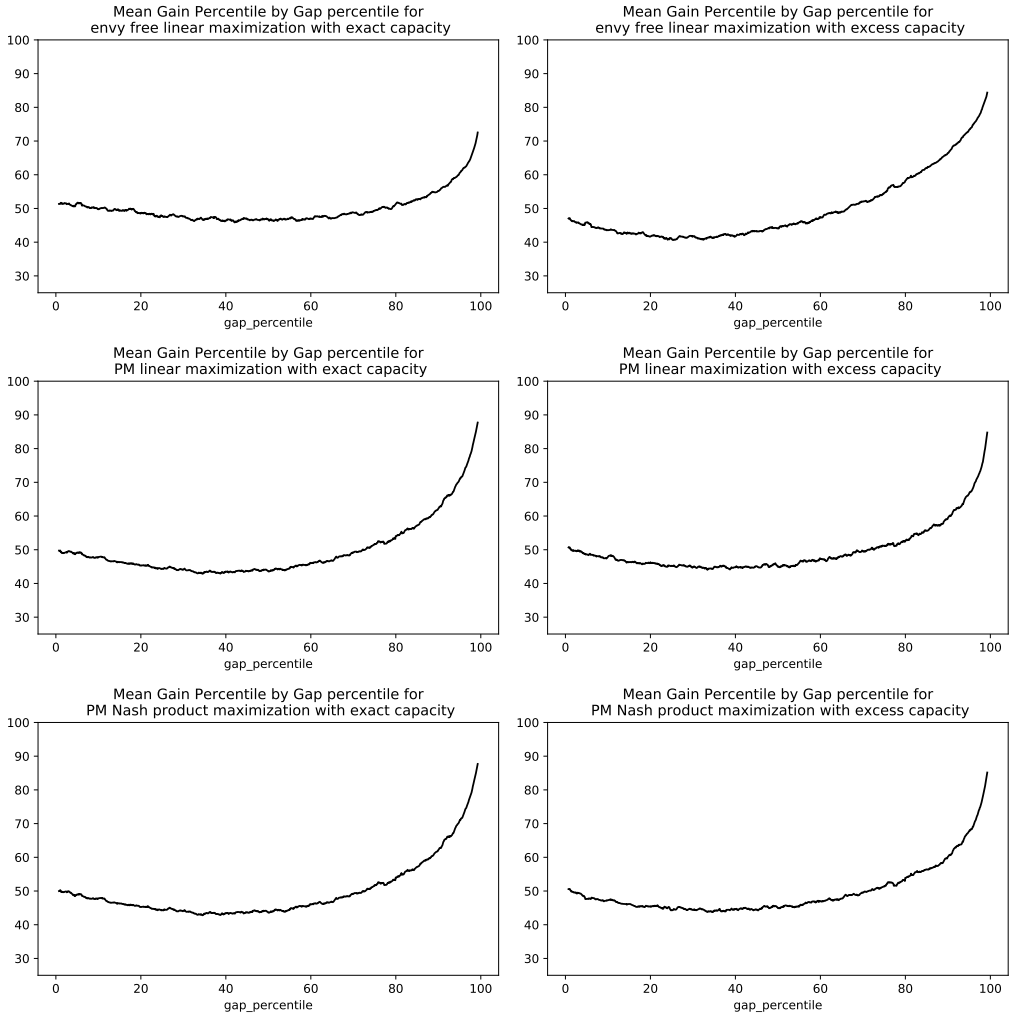
We evaluate two cardinal mechanisms, a variant of the pseudomarket mechanism of Hylland and Zeckhauser (1979) and the envy-free mechanism, which we see as a special case of the course bundles mechanism in Nguyen, Peivandi and Vohra (2015). Both of these mechanisms are strategy proof in the large. Our version of the pseudomarket differs from Hylland and Zeckhauser because we restrict attention to equilibria that maximize a social welfare function chosen by a school district and drop other some conditions in Hylland and Zeckhauser that seem to not be needed in our approach for Pareto efficiency. We are the first paper to develop numerical algorithms to implement a pseudomarket for single item assignment at scale. Our algorithm works well for problems on the scale of Seattle's ninth grade choice program and somewhat larger problems as well.

If both cardinal mechanisms are seen as mathematical programs maximizing social welfare, the envy-free mechanism imposes fewer restrictions than the pseudomarket problem. Therefore, the social welfare from the envy-free mechanism will be higher than that from the pseudomarket mechanism.

Our main goal is to measure the welfare gain of both mechanisms over ordinal mechanisms. We introduce a new gain measure for this purpose. We then use the gain measure to compare both cardinal mechanisms to the ordinally efficient mechanism known as probabilistic serial, which is also strategy proof in the large.

Our baseline case is where the schools have a total capacity equal to the number of students,

Figure 3: Gain Percentile for Students by the Percentile of the Utility Difference (Gap) Between Their First and Second Most Preferred Schools



students are fully informed about the lotteries they face in the mechanism and we take a uniform draw from the revealed-preference set of school utilities. We find that 71% of students prefer their pseudomarket assignments to that from the ordinal mechanism. Likewise, 67% of students prefer their envy-free assignments to the ordinal mechanism. The mean gain of the pseudomarket mechanism over the ordinal mechanism is 0.015. The mean gain of the envy-free mechanism over the ordinal mechanism is 0.045. Using as a reference the gain over the ordinal mechanism from a utilitarian benchmark that lacks incentives for students to tell the truth, the pseudomarket mechanism captures 21% of the possible gain over the ordinal mechanism and the envy-free mechanism captures 58% of the possible gain over the ordinal mechanism.

We explore various perturbations to our baseline case. Increasing the capacity of schools while keeping the number of students fixed gives more students their top choices and lowers the gain of using one mechanism over another. Dropping the assumption that students know their preferences and using only the weak assumption that schools listed in the data are preferred to the default school, we find that the mean gain drops in magnitude by roughly a half but that more students prefer their cardinal mechanism assignments to their ordinal mechanism assignments. We also changed the assumption of a uniform draw from the revealed preference set of each student.

References

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

_ , Nikhil Agarwal, and Parag A. Pathak, "The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match," *American Economic Review*, 2017, *107* (12), 3635–89.

_ , Parag A. Pathak, and Alvin E. Roth, "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match," *American Economic Review*, 2009, *99* (5), 1954–78.

_ , Yeon-Koo Che, and Yosuke Yasuda, "Resolving Conflicting Preferences in School Choice: the Boston Mechanism Reconsidered," *American Economic Review*, 2011, *101* (1), 399–410.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis using Strategic Reports: An application to a school choice mechanism," *Econometrica*, 2018, *86* (2), 391–444.

Azevedo, Eduardo M and Eric Budish, "Strategy-proofness in the Large," *Review of Economic Studies,*, 2019, *86* (1), 81–116.

Aziz, Haris, "Competitive equilibrium with equal incomes for allocation of indivisible objects," *Operations Research Letters*, 2015, *43* (6), 622–624.

Bajari, Patrick and C. Lanier Benkard, "Demand Estimation With Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *The Journal of Political Economy*, 2005, *113* (6), 1239–1276.

Bélisle, Claude, Arnon Boneh, and Richard J Caron, "Convergence properties of hit–and–run samplers," *Stochastic Models*, 1998, *14* (4), 767–800.

Birkhoff, Garrett, "Three Observations on Linear Algebra," *Universidada Nacional de Tucuman, Revista*, 1946, *Serie A* (vol. 5), 147–151.

Bogomolnaia, Anna and Herve Moulin, "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 2001, *100* (2), 295–328.

Budish, Eric, "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes," *Journal of Political Economy*, 2011, *119* (6), 1061–1103.

_ **and Judd B. Kessler**, "Bringing Real Market Participants' Real Preferences into the Lab: An Experiment that Changed the Course Allocation Mechanism at Wharton," Working Paper 22448, National Bureau of Economic Research July 2016.

Budish, Eric B, Gerard Cachon, Judd B Kessler, and Abraham Othman, "Course Match: A Large-Scale Implementation of Approximate Competitive Equilibrium from Equal Incomes for combinatorial allocation," *Chicago Booth Research Paper*, 2015.

Budish, Eric, Yeon-Koo Che, Fuhito Kojima, and Paul Milgrom, "Designing Random Allocation Mechanisms: Theory and Applications," *American Economic Review*, 2013, *103* (2), 585–623.

Debreu, Gerard, "Topological methods in cardinal utility theory," Technical Report, Cowles Foundation for Research in Economics, Yale University 1959.

Dierker, Egbert, *Topological methods in Walrasian economics* Lecture Notes in Economics and Mathematical Systems, Springer Berlin Heidelberg, 1974.

Echenique, Federico, Antonio Miralles, and Jun Zhang, "Constrained Pseudomarket Equilibrium," November 2020. Working paper.

He, Yinghua, Antonio Miralles, Marek Pycia, and Jianye Yan, "A Pseudo-Market Approach to Allocation with Priorities," *American Economic Journal: Microeconomics*, 2018, *10* (3), 272–314.

Hildenbrand, Werner, *Core and equilibria of a large economy*, Princeton, NJ: Princeton University Press, 1974.

Hoheisel, Tim, Christian Kanzow, and Alexandra Schwartz, "Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints," *Mathematical Programming*, 2013, *137* (1-2), 257–288.

Hosseini, Mojtaba and Vijay V Vazirani, "Nash-Bargaining-Based Models for Matching Markets, with Implementations and Experimental Results," *arXiv preprint arXiv:2105.10704*, 2021.

Hylland, Aanund and Richard Zeckhauser, "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 1979, *87* (2), 293–314.

Immorlica, Nicole, Brendan Lucier, Joshua Mollner, and E. Glen Weyl, "Raffles," 2017. Microsoft Research working paper.

Jackson, Matthew O., "Incentive Compatibility and Competitive Allocations," *Economics Letters*, 1992, *40* (3), 299–302.

Kapor, Adam, Christopher Neilson, and Seth Zimmerman, "Heterogeneous Beliefs and School Choice," *American Economic Review*, 2020, *110* (5).

Kojima, Fuhito and Mihai Manea, "Incentives in the probabilistic serial mechanism," *Journal of Economic Theory*, 2010, *145* (1), 106–123.

Leyffer, Sven, Gabriel López-Calva, and Jorge Nocedal, "Interior methods for mathematical programs with complementarity constraints," *SIAM Journal on Optimization*, 2006, *17* (1), 52–77.

Liu, Qingming and Marek Pycia, "Ordinal Efficiency, Fairness, and Incentives in Large Markets," *Working paper*, 2012.

Miralles, Antonio, "School Choice: The Case for the Boston Mechanism," *Working paper*, 2008.

Mittelman, Han, "MPEC Benchmark (Math. Progr. w. Equilibrium Constraints)," July 2021.

Nguyen, Thanh, Ahmad Peivandi, and Rakesh Vohra, "Assignment Problems with Complementarities," *Working paper*, 2015.

Pathak, Parag A. and Tayfun Sonmez, "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 2008, *98* (4), 1636–52.

Pratt, John Winsor and Richard Jay Zeckhauser, "The fair and efficient division of the Winsor family silver," *Management Science*, 1990, *36* (11), 1293–1301.

Pycia, Marek, "The Cost of Ordinality," *Working paper*, 2014.

Schmidt, Martin and Yasmine Beck, "A Gentle and Incomplete Introduction to Bilevel Optimization," November 2021. Working paper.

Sinha, Ankur, Pekka Malo, and Kalyanmoy Deb, "A review on bilevel optimization: from classical to evolutionary approaches and applications," *IEEE Transactions on Evolutionary Computation*, 2018, *22* (2), 276–295.

Vazirani, Vijay V and Mihalis Yannakakis, "Computational complexity of the Hylland-Zeckhauser scheme for one-sided matching markets," *arXiv preprint arXiv:2004.01348*, 2020.

von Neumann, John, "A Certain Zero-Sum Two-Person Game Equivalent to the Optimal Assignment Problem," in Harold William Kuhn and Albert William Tucker, eds., *Contributions to the Theory of Games*, Vol. II, Princeton, N. J.: Princeton University Press, 1953, pp. 5–12.

A Proofs

A.1 Proposition 1

Proposition 1 gives sufficient conditions for existence of an equilibrium maximizing a social welfare function that is strictly increasing in the expected utilities of all students.

All equilibria in Hylland and Zeckhauser (1979) satisfy our definition of an equilibrium.[8] Therefore the set of equilibria is always nonempty.

Because equilibrium assignments π are contained in a compact set and the Nash product is a continuous function, there exists a sequence of equilibria (π^n, p^n) such that

$$\lim_{n\to\infty} W(\pi^n, v) = \sup_{(\pi,p)\in \text{Equilibria}(v,q)} W(\pi, v).$$

The technical challenge is to show that the supremum is attained. The result would follow immediately if a limit of a sequence of equilibria was itself an equilibrium. The following example shows that this is not always the case.

Example. Consider a single-student economy with two unit-capacity schools and preferences given by $v_1 = (0, 1)$, so the student prefers school 2. For each n, $\left(p^n = \left(1, 1 + \frac{1}{n}\right), \pi_1^n = (1, 0)\right)$ describes an equilibrium, as the price of school 2 is above the budget and the student must purchase only school 1. However, $(p = (1, 1) = \lim_{n\to\infty} p^n, \pi_1 = (1, 0) = \lim_{n\to\infty} \pi_1^n)$ is not an equilibrium because the prices of the two schools are the same and the student prefers school 2.

The conditions in Proposition 1 guarantee that a limit of equilibria is itself an equilibrium. By Berge's theorem of the maximum, $u_i(p)$ is continuous at the price vector p if there exists at least one

[8] Hylland and Zeckhauser imposed a zero price for all under-demanded schools, when there is excess capacity. This condition is compatible with the restriction on prices in the statement of Proposition 1.

school with price less than 1, where we define

$$u_i(p) = \max \left\{ v_i \cdot x : \forall j \in J, x_j \geq 0, \sum_{j \in J} x_j = 1, x \cdot p \leq 1 \right\}.$$

We first show the following lemma:

Lemma 1. *Consider a student i with a preference vector v_i. The indirect utility function*

$$u_i(p) := \max \left\{ v_i \cdot x : \forall j \in J, x_j \geq 0, \sum_{j \in J} x_j = 1, x \cdot p \leq 1 \right\}$$

is continuous at price vector p if $\exists j$ s.t. $p_j < 1$.

Proof. The claim is clear if the individual budget constraint does not bind at p, as in this case there exists an open neighborhood U of p s.t. $u_i(p^{'}) = \max_j v_{ij}$ on U, i.e. u_i is locally constant around p. Let p be a price vector s.t. $p_j < 1$ and suppose that the budget constraint binds at p. By choice of p, there exists $p^{'}$ s.t. $p_j^{'} < 1$.

We define two price vectors $\underline{p}, \overline{p}$ by $\underline{p}_j = \min\{p_j, p_j^{'}\}$ and $\overline{p}_j = \max\{p_j, p_j^{'}\}$. We note that $u_i(\underline{p}) \leq u_i(p), u_i(p^{'}) \leq u_i(\overline{p})$, and so $|u_i(p) - u_i(p^{'})| \leq |u_i(\overline{p}) - u_i(\underline{p})|$. Let $\underline{x} \in \arg\max\{v_i \cdot x : x \geqq 0, x \cdot 1 = 1, x \cdot \underline{p} \leq 1\}$ and $t = \frac{1-\overline{p}_i}{\overline{p} \cdot \underline{x} - \overline{p}_j}$; we note that $\tilde{x} = t\underline{x} + (1-t)e_j$ is affordable under prices $\overline{p}$. Hence,

$$|u_i(\overline{p}) - u_i(\underline{p})| \leq (1-t)|v_i \cdot (\underline{x} - e_j)| \leq (1-t) \times \sqrt{2\|v_i\|},$$

where the last inequality follows by Cauchy-Schwartz. Choosing $p^{'}$ sufficiently close to p, we can make $\overline{p} \cdot \underline{x}$ arbitrarily close to 1, hence $(1-t)$ arbitrarily close to 0 and the right-hand side arbitrarily small. It follows that u_i is continuous at p. □

A.1.1 Proof of Proposition 1

Let $\epsilon \in (0,1)$ and consider the subset of equilibria

$$\text{Equilibria}^{\epsilon}(v,q) = \left\{ (\pi, p) \in \text{Equilibria}(v,q) : \min_{j \in J} p_j \leq 1 - \epsilon \text{ and } \max_{j \in J} p_j \leq |I| \right\}.$$

Let $W_{\epsilon}^{*} = \sup_{(\pi,p) \in \text{Equilibria}^{\epsilon}(v,q)} W(\pi, v)$. By the definition of a supremum, there exists a sequence of equilibria $(\pi^n, p^n) \in \text{Equilibria}^{\epsilon}(v,q)$ such that $\lim_{n \to \infty} W(\pi^n, v) = W_{\epsilon}^{*}$. Since equilibrium assignments are contained in a compact set and prices are bounded above by $|I|$, there exists a convergent subsequence (π^{n_k}, p^{n_k}) of (π^n, p^n) with $\lim_{n_k \to \infty}(\pi^{n_k}, p^{n_k}) = (\pi^{\star}, p^{\star})$. Because $\min_{j \in J} p_j^{n_k} \leq 1 - \epsilon$ for all n_k, we have $\min_{j \in J} p_j^{\star} < 1$. As indirect utilities are continuous in the price vector p if at least one

school's price is less than 1, the indirect utilities converge: $u_i(p^{n_k}) \to u_i(p^\star)$. In addition, for each student i the expected utilities converge $\pi_i^{n_k} \cdot v_i \to \pi_i^\star \cdot v_i$. It follows that $\pi_i^\star$ maximizes student i's utility at prices $p^\star$, and we conclude that $(\pi^\star, p^\star)$ is an equilibrium. (Feasibility of the assignment $\pi^\star$ is easy to check).

Recall that $\lim_{n\to\infty} W(\pi^n, v) = W_\epsilon^*$. Because the social welfare function is continuous in expected utilities by the statement of the proposition, $W_\epsilon^* = W(\pi^\star, v)$. Hence, $(\pi^\star, p^\star)$ is an equilibrium maximizing the social welfare function.

Hylland and Zeckhauser (1979) restrict attention to equilibria such that at least one price is zero and prices for all under-demanded schools are zero, the latter under excess capacity. Hence, under unit budgets, the set of equilibria considered by Hylland and Zeckhauser is a subset of Equilibria$^\epsilon(v, q)$ for every $\epsilon \in (0, 1)$.

B Details of Equilibrium Selection Example from Section 2.3

The following are the stochastic assignments π from the example of equilibrium selection in Section 2.3:

i)

	j_1	j_2	j_3	j_4
i_1	$\frac{11}{16}$	0	0	$\frac{5}{16}$
i_2	0	$\frac{5}{16}$	0	$\frac{11}{16}$
i_3	0	0	1	0
prices	$\frac{82877}{85822}$	$\frac{32531}{39010}$	1	$\frac{8391}{7802}$

ii)

	j_1	j_2	j_3	j_4
i_1	$\frac{3}{4}$	0	0	$\frac{1}{4}$
i_2	0	$\frac{1}{4}$	0	$\frac{3}{4}$
i_3	0	0	1	0
prices	$\frac{28436}{29001}$	$\frac{7972}{9667}$	1	$\frac{10232}{9667}$

iii)

	j_1	j_2	j_3	j_4
i_1	$\frac{1}{2}$	0	0	$\frac{1}{2}$
i_2	0	$\frac{1}{2}$	0	$\frac{1}{2}$
i_3	0	0	1	0
prices	0	0	1	2

iv)

	j_1	j_2	j_3	j_4
i_1	$\frac{6}{11}$	0	0	$\frac{5}{11}$
i_2	$\frac{1}{11}$	$\frac{10}{11}$	0	0
i_3	$\frac{4}{11}$	$\frac{1}{11}$	$\frac{6}{11}$	0
prices	0	$\frac{11}{10}$	$\frac{33}{20}$	$\frac{11}{5}$

C Full Mathematical Program for Mechanism

The statement of the mathematical problem for the mechanism involves $|J| + 1$ complementarity constraints per student. We adopt the standard notation from the optimization literature and write $f(x) \leq 0 \perp g(x) \leq 0$ to indicates that constraints $f(x) \leq 0$ and $g(x) \leq 0$ are complementary. This is mathematically equivalent to $f(x) \leq 0$ and $g(x) \leq 0$ and $f(x)g(x) = 0$, and allows us to take advantage of optimization software that specializes in problems with complementarity constraints.

Given prices, the optimization problem for student $i \in I$ can be stated as

$$\begin{aligned}\max_{\pi_i} \pi_i \cdot v_i \text{ s.t. } \pi_i \cdot p \leq 1 &\text{(budget constraint)}\\ \sum_{j \in J} \pi_{ij} = 1 &\text{(unit demand constraint)}\\ \forall j, \pi_{ij} \geq 0 &\text{(non-negative demand constraint).}\end{aligned}$$

We introduce Lagrange multipliers $\lambda_i \geq 0$ associated to the budget constraints, $\eta_i \in \mathbb{R}$ associated to the unit demand constraint and $\forall j$, $\mu_{ij} \geq 0$ associated to the non-negative demand constraint. The Karush-Kuhn-Tucker conditions imply that π_i solves the individual problem if and only if the following conditions are met:

$$\begin{aligned}1 - \pi_i \cdot p \geq 0 \perp \lambda_i \geq 0\\ \forall j, \pi_{ij} \geq 0 \perp \mu_{ij} \geq 0\\ \forall j, v_{ij} + \mu_{ij} = \lambda_i p_j + \eta_i\\ \sum_{j \in J} \pi_{ij} = 1.\end{aligned}$$

Slater's condition shows that the KKT conditions are necessary and sufficient for π_i to be a solution of the student's original linear program.

The problem of finding an equilibrium can be stated as a constraint maximization problem

$$\begin{aligned}\max_{\pi, p, \lambda, \eta, \mu} \sum_{i \in I} \log(v_i \cdot \pi_i - v_i \cdot \tilde{q}) \text{ s.t.}&\\ \forall j, \sum_{i \in I} \pi_{ij} \leq q_j &\text{(capacity constraints)}\\ \forall i, \text{ KKT conditinos for } i \text{ hold} &\text{(individual optimality constraints)}\\ \forall j, p_j \geq 0. &\text{(prices are non-negative)}\end{aligned}$$

In practice, it is useful to impose an additional constraint $\forall i$, $\pi_i \cdot v_i \geq \tilde{q} \cdot v_i$ to ensure that the Nash product objective function is well-defined along the solution path.

The statement of the problem assumes that there does not exist a fully indifferent student, i.e. $\forall i$, $\exists\, j, k \in J$ such that $v_{ij} \neq v_{ik}$. If such as student was present, we could remove her from the assignment problem, compute an equilibrium and then assign the indifferent student arbitrarily.

D Sampler for School Utilities

We propose an algorithm to draw vectors of school utilities v_i uniformly from the revealed-preference set $\tilde{V}_i$, which is equation (8).

A hit-and-run sampler is one MCMC algorithm to draw points uniformly from a convex polytope (e.g., Bélisle, Boneh and Caron, 1998). Given a valid vector of school utilities v_i^k for iteration k of the algorithm, a direction $\lambda^{k+1} \in \mathbb{R}^J$ is chosen uniformly. The algorithm then finds the sharp bounds $\underline{\alpha}^{k+1}$ and $\bar{\alpha}^{k+1}$ such that

$$\left(v_i^k + \lambda^{k+1}\alpha\right) \in \tilde{V}_i, \, \forall \, \alpha \in \left[\underline{\alpha}^{k+1}, \bar{\alpha}^{k+1}\right].$$

Then the algorithm picks α^{k+1} uniformly in $\left[\underline{\alpha}^{k+1}, \bar{\alpha}^{k+1}\right]$. The next vector of school utilities is

$$v_i^{k+1} = \lambda^{k+1}\alpha^{k+1}.$$

The algorithm can be run for any number of iterations. It can be graphically checked that the resulting set of sampled vectors of school utilities appears uniform over $\tilde{V}_i$.