Supplementary Appendix to "Strategy-proofness in the Large" (for Online Publication)

Eduardo M. Azevedo*and Eric Budish†

August 25, 2015

B Relaxing Continuity to Quasi-Continuity

In this section we relax the continuity requirement in Theorem 2 to a condition we call quasi-continuity. Roughly, quasi-continuity relaxes continuity to allow for discontinuities so long as they are knife-edge. This relaxation is useful because some mechanisms are discontinuous at some knife-edge configurations. For example, the families of equilibria of pay-as-bid and uniform-price auctions described in Appendix D are not continuous, but are quasi-continuous.

Definition B.1. *Consider a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ with limit $\phi^\infty(\cdot,\cdot)$, and a family of limit Bayes-Nash equilibria $(\sigma^*_\mu)_{\mu\in\Delta T}$. The family of equilibria is **quasi-continuous** if, for every $\mu_0 \in \bar{\Delta}T$ and every $\epsilon > 0$, there exists a neighborhood $\mathcal{N}$ of μ_0 that can be decomposed as $\mathcal{N} = \cup_{1\le k\le K}\mathcal{A}_k \cup \mathcal{B}$ with each $\mathcal{A}_k$ open, such that:*

1. *If types are drawn i.i.d. according to μ_0, then the probability that the realized empirical distribution of types is within distance $1/n$ of $\mathcal{B}$ goes to zero as n grows large. Formally,*

$$\lim_{n\to\infty} \Pr\{\text{distance}(\text{emp}[t], \mathcal{B}) \le 1/n | t \in T^n, t \sim iid(\mu_0)\} = 0.$$

2. *Within each set $\mathcal{A}_k$, in a large enough market, agents' outcomes are continuous with respect to changes in the empirical distribution of opponents' types and the strategy that agents use. Formally, there exists n_0 such that for each $\mathcal{A}_k$, for any $n \ge n_0$, and*

*Wharton, eazevedo@wharton.upenn.edu.
†University of Chicago Booth School of Business, eric.budish@chicagobooth.edu.

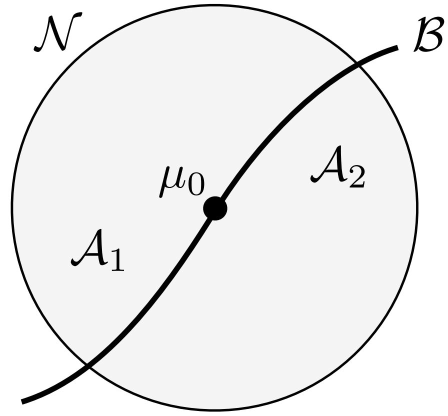

Figure B.1: The Figure illustrates the quasi-continuity definition, with $K = 2$. Around the prior $\mu_0 \in \bar{\Delta}T$, there exists a neighborhood $\mathcal{N}$ that can be decomposed into the sets $\mathcal{A}_1$ and $\mathcal{A}_2$, where equilibrium outcomes vary continuously, and a small "knife edge" set $\mathcal{B}$ where equilibrium outcomes may be discontinuous.

any $\mu, \mu', \text{emp}[t_i, t_{-i}], \text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k$, *we have:*

$$\|\Phi_i^n(\sigma_\mu^*(t_i), \sigma_\mu^*(t_{-i})) - \Phi_i^n(\sigma_{\mu'}^*(t_i), \sigma_{\mu'}^*(t'_{-i}))\| < \epsilon.$$

In words, quasi-continuity allows the family of equilibria to be discontinuous at prior μ_0, but it requires that the discontinuity is knife-edge in the following sense: a small enough neighborhood $\mathcal{N}$ of μ_0 can be decomposed as a finite number of subsets $\mathcal{A}_k$ where the outcomes vary continuously, and a set $\mathcal{B}$ where the empirical distribution of a randomly drawn type profile lands with vanishingly small probability. This decomposition is illustrated in Figure B.1. Heuristically, $\mathcal{B}$ is a small discontinuity set, and is surrounded by sets $\mathcal{A}_k$ where outcomes vary continuously. Note that quasi-continuity requires that, within each region $\mathcal{A}_k$, outcomes vary continuously with both types and strategies. In contrast, continuity as defined in Definition 9 only requires that outcomes vary continuously with strategies, which is less restrictive.

Theorem B.1. *Given any mechanism* $\{(\Phi^n)_\mathbb{N}, A\}$ *with a quasi-continuous family of limit Bayes-Nash equilibria* $(\sigma_\mu^*)_{\mu \in \Delta T}$*, there exists a direct, SP-L mechanism* $\{(F^n)_\mathbb{N}, T\}$ *with the following properties.*

1. *If the original mechanism is continuous at a prior* $\mu_0 \in \bar{\Delta}T$ *then, in the limit, truthful play of the direct mechanism produces the same outcomes as equilibrium play of the*

original mechanism. Formally, for any $t_i \in T$, we have

$$f^{\infty}(t_i, \mu_0) = \phi^{\infty}(\sigma^*_{\mu_0}(t_i), \sigma^*_{\mu_0}(\mu_0)),$$

where f^{∞} is the limit of the direct mechanism.

2. *For any prior $\mu_0 \in \bar{\Delta}T$, in the large market limit, truthful play of the direct mechanism produces the same outcomes as a convex combination of equilibrium play of the original mechanism under priors that are close to μ_0. Formally, for any $\epsilon > 0$, there exists n_0, an integer K, numbers π^n_k with $\sum_{k=1,\cdots,K} \pi^n_k = 1$, and priors μ_k with $\|\mu_k - \mu_0\| < \epsilon$ such that, for all $n \geq n_0$ and $t_i \in T$, we have*

$$\|f^n(t_i, \mu_0) - \sum_{k=1,\cdots,K} \pi^n_k \cdot \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| < \epsilon,$$

where f^n is the function representing the direct mechanism from an interim perspective, as defined in equation (3.1).

Theorem B.1 says that, if the original mechanism is quasi-continuous rather than continuous, there exists an SP-L mechanism that approximates the Bayes-Nash mechanism, but in a weaker sense than in the continuous case: the SP-L mechanism approximates a convex combination of outcomes of the original Bayes-Nash mechanism, for a set of priors arbitrarily close to the prior μ_0.

B.1 Proof of Theorem B.1

Let F be defined as in equation (5.2). The proof of Theorem B.1 is based on the following approximation lemma.

Lemma B.1. *Fix a prior $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$. Then there exists a neighborhood $\mathcal{N} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_K \cup \mathcal{B}$, with each $\mathcal{A}_k$ open, and n_0, with the following property. For each $k = 1, \cdots, K$ we can take a prior μ_k in $\mathcal{A}_k$ with $\|\mu_k - \mu_0\| < \epsilon$ such that, for all $n \geq n_0$, there exist positive weights π^n_k with $\sum_{1 \leq k \leq K} \pi^n_k = 1$, such that, for all t_i,*

$$\|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi^n_k \cdot z_k(t_i)\| < \epsilon, \tag{B.1}$$

where

$$z_k(t_i) = \phi^{\infty}(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)). \tag{B.2}$$

The lemma states that the bundle received by an agent playing t_i in the direct mechanism can be approximated by a convex combination of the bundles type t_i receives in the limit Bayes-Nash equilibria of the original mechanism, with the elements in the convex combination corresponding to priors in a small neighborhood of μ_0, with one prior for each region $\mathcal{A}_k$.

The lemma deals with one of the key difficulties in the proof. Namely, the lemma implies that each agent can only have a small effect on the aggregate outcome of the constructed mechanism, in the sense that the weights π_k^n do not depend on agent i's report t_i. Note also that the approximation formula is not generally true without the quasi-continuity condition, as shown with an example in Appendix B.2. Before proving the lemma, we use it to prove Theorem B.1.

Proof of Theorem B.1. If the family of Bayes-Nash equilibria is continuous at a prior $\mu_0 \in \bar{\Delta}T$, then the proof of Theorem 2 implies that truthful play of f under μ_0 produces the same outcomes as equilibrium play of the original mechanism. That is, Part 1 of of the theorem statement follows from the proof of Theorem 2.

Consider now a prior $\mu_0 \in \bar{\Delta}T$, and $\epsilon > 0$. Recall that, by assumption, the family of limit Bayes-Nash equilibria is quasi-continuous, but may not be continuous. We will show that there exists n_0 such that, for all $n \geq n_0$, the gain from misreporting is lower than ϵ for any type (which proves that the direct mechanism is SP-L) and that Part 2 of the theorem statement holds.

The proof is based on the following approximation. By Lemma B.1 (using $\frac{\epsilon}{2|X_0|}$ as the constant), there exists a neighborhood $\mathcal{N} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_K \cup \mathcal{B}$ of μ_0, priors $\mu_k \in \mathcal{N}_k$, weights π_k^n, and n_0 with the following properties. For all t_i' in T, and $n \geq n_0$,

$$\begin{aligned} \sum_{k=1}^{K} \pi_k^n &= 1, \qquad \text{(B.3)} \\ \|\mu_k - \mu_0\| &< \epsilon, \text{ and} \\ \|f^n(t_i', \mu_0) - \sum_{k=1}^{K} \pi_k^n \cdot z_k(t_i')\| &< \frac{\epsilon}{2|X_0|} \leq \frac{\epsilon}{2}, \end{aligned}$$

where z_k is given by equation (B.2).

Step 1: the gain from misreporting is no greater than ϵ.

Consider, for any pair of types t_i and t_i', the gain of a type t_i player from deviating to t_i', when opponents play i.i.d. according to μ_0. Using the approximation formula, we can bound the gain from deviating, for $n \geq n_0$, by

$$\begin{aligned} u_{t_i}[f^n(t_i', \mu_0)] - u_{t_i}[f^n(t_i, \mu_0)] &\leq \\ \sum_{k=1}^{K} \pi_k^n \cdot \{u_{t_i}[z_k(t_i')] - u_{t_i}[z_k(t_i)]\} & \\ +|u_{t_i}[f^n(t'_i, \mu_0)] - \sum_{k=1}^{K} \pi_k^n \cdot u_{t_i}[z_k(t'_i)]| & \\ +|u_{t_i}[f^n(t_i, \mu_0)] - \sum_{k=1}^{K} \pi_k^n \cdot u_{t_i}[z_k(t_i)]| &< \\ 0 + \epsilon/2 + \epsilon/2 &= \epsilon. \end{aligned}$$

The first inequality follows from rearranging the LHS, and then taking absolute values of the two last terms in the RHS. As for the second inequality, the first term is weakly negative by equation (B.2) and the fact that $\sigma^*_{\mu_k}$ is a limit equilibrium. The second and third terms are smaller than $\epsilon/2$ by the bounds (B.3), the fact that utility is always between 0 and 1, and that the set of random bundles has $|X_0|$ dimensions. Since $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$ are arbitrary, we have that the constructed mechanism is SP-L.

Step 2: outcomes of the constructed mechanism $\{(F^n)_{\mathbb{N}}, T\}$ approximate a convex combination of equilibrium outcomes of $\{(\Phi^n)_{\mathbb{N}}, A\}$ under $(\sigma^*_\mu)_{\mu \in \Delta T}$ at μ_0.

By the triangle inequality we have

$$\begin{aligned} & \|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi_k^n \cdot \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| \\ \leq \quad & \|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi_k^n \cdot \phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| \\ & + \sum_{k=1}^{K} \pi_k^n \cdot \|\phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) - \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\|. \end{aligned}$$

The first term on the RHS of the inequality is bounded by $\epsilon/2$, by the bound (B.3). By the definition of the limit, and the fact that the π_k^n sum to 1, we may take n_0 to be large enough such that the second term is also bounded by $\epsilon/2$. Moreover, this bound can be taken uniform for all $t_i \in T$. Therefore, we have that

$$\|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi_k^n \cdot \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| < \epsilon/2 + \epsilon/2 = \epsilon, \tag{B.4}$$

as desired.

□

B.1.1 Proof of Lemma B.1

In the proof we will use the following notation. If t' is a vector of types, and $\mathcal{S} \subseteq \Delta T$, we will say that $t' \in \mathcal{S}$ iff $\mathrm{emp}[t'] \in \mathcal{S}$. Throughout the proof, we use the shorthand $\hat{\mu} = \mathrm{emp}[t]$. The expression

$$\Pr(t'_{-i} | t'_{-i} \sim \mu)$$

denotes the probability that the vector of types t'_{-i} is realized if each player's type is drawn i.i.d. according to the distribution μ.

Proof of Lemma B.1. We begin by constructing the neighborhood in the statement of the lemma. By the quasi-continuity condition, we take a neighborhood $\mathcal{N} = \cup_{k=1}^{K} \mathcal{A}_k \cup \mathcal{B}$ and n_0 such that Conditions 1 and 2 of Definition B.1 are satisfied with $\epsilon/5$ in place of ϵ. We take $\mathcal{N}$ to be convex, which is without loss of generality.[1] We let the μ_k be any priors in $\mathcal{A}_k$ such that $\|\mu_k - \mu_0\| < \epsilon$. With that, to prove the lemma we have to show that there exist weights π_k^n such that the approximation formula (B.1) holds. The proof involves three steps.

Step 1: approximation of $F^n(t)$ for vectors of types with empirical distribution of types in each region $\mathcal{A}_k$.

Claim B.1. The integer n_0 can be taken such that, for all $n \geq n_0$, $k = 1, \cdots, K$, and $t \in \mathcal{A}_k$ we have

$$\|F_i^n(t) - z_k(t_i)\| < 3\epsilon/5.$$

Proof. We begin with the term $F_i^n(t)$, and derive two inequalities that together yield the desired bound. The first inequality bounds the distance between $F_i^n(t)$ and $\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))$. By definition, we have that

$$\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) = \sum_{t'_{-i}} \Pr(t'_{-i} | t'_{-i} \sim \mu_k) \cdot \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i})). \tag{B.5}$$

[1] To see this, note that, if $\mathcal{N}$ is a non-convex neighborhood satisfying the requirements, we can take a ball $\mathcal{N}' \ni \mu_0$ contained in $\mathcal{N}$, and define the sets $\mathcal{A}'_k = \mathcal{A}_k \cap \mathcal{N}'$ and $\mathcal{B}' = \mathcal{B} \cap \mathcal{N}'$. It follows that Conditions 1 and 2 hold for the new sets, as each $\mathcal{A}'_k \subseteq \mathcal{A}$ and $\mathcal{B}' \subseteq \mathcal{B}$.

Therefore, we have that

$$\begin{aligned}
& \|F_i^n(t) - \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| \qquad \text{(B.6)} \\
= \ & \|F_i^n(t) - \sum_{t_{-i}} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i}))\| \\
\leq \ & \sum_{t'_{-i} \in T^{n-1}} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \|F_i^n(t) - \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i}))\| \\
= \ & \sum_{t'_{-i}:\text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \|F_i^n(t) - \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i}))\| + \\
& \sum_{t'_{-i}:\text{emp}[t_i, t'_{-i}] \notin \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \|F_i^n(t) - \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i}))\|.
\end{aligned}$$

The first equality follows by substituting the definition of ϕ^n from equation (B.5). The inequality follows from the triangle inequality and the fact that the probabilities must sum to 1. The last equality simply breaks the sum into two parts, the t'_{-i} for which $\text{emp}[t_i, t'_{-i}]$ is in $\mathcal{A}_k$, and the ones for which it is not.

Consider now the expression in the RHS of inequality (B.6). From the way we construct $F^n(\cdot)$ and using the convention that $\hat{\mu} = \text{emp}[t]$, the first term equals

$$\begin{aligned}
& \sum_{t'_{-i}:\text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \|F_i^n(t) - \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i}))\| \\
= \ & \sum_{t'_{-i}:\text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \|\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i), \sigma^*_{\hat{\mu}}(t_{-i})) - \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i}))\|.
\end{aligned}$$

Condition 2 in Definition B.1 implies that, for all $n \geq n_0$, this expression is bounded above by $\epsilon/5$. As for the second term in the RHS of inequality (B.6), by the weak law of large numbers, we may take n_0 large enough such that the total probability that $\text{emp}[t_i, t'_{-i}] \notin \mathcal{A}_k$ is lower than $\epsilon/5$ for all $n \geq n_0$. This bounds the second term by $\epsilon/5$. Substituting these bounds in inequality (B.6) then yields

$$\|F_i^n(t) - \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| < \epsilon/5 + \epsilon/5 = 2\epsilon/5. \qquad \text{(B.7)}$$

Finally, by the definition of the limit, we may take n_0 such that, for all $n \geq n_0$,

$$\begin{aligned}
& \|\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) - z_k(t_i)\| \qquad \text{(B.8)} \\
= \ & \|\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) - \phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| < \epsilon/5.
\end{aligned}$$

Note that these bounds are uniform for all $t \in \mathcal{A}_k$. Moreover, since K is finite, n_0 can be taken such that the bounds hold for all $k = 1, \cdots, K$. The claim then follows from inequalities (B.7) and (B.8), as, for $n \geq n_0$,

$$\begin{aligned} & \|F_i^n(t) - z_k(t)\| \\ \leq \ & \|F_i^n(t) - \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| + \|\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) - z_k(t_i)\| \\ < \ & 2\epsilon/5 + \epsilon/5 = 3\epsilon/5, \end{aligned}$$

completing the proof. □

The next step shows that the probability that a vector (t_i, t_{-i}) falls within region $\mathcal{A}_k$, when t_{-i} is distributed randomly, does not vary too much with t_i in large markets. This is a key step in our argument, as it implies that an individual agent cannot appreciably change the probability that t falls within each $\mathcal{A}_k$, and therefore cannot have a large effect on the aggregate allocation.

Step 2: approximation of the probability that the empirical distribution of types is in region $\mathcal{A}_k$.

Claim B.2. The integer n_0 can be taken such that, for all $n \geq n_0$ there exist weights $\pi_1^n, \ldots, \pi_K^n$ such that $\sum_{k=1}^K \pi_k^n = 1$ and

$$|\Pr\{(t_i, t_{-i}) \in \mathcal{A}_k | t_{-i} \in T^{n-1}, t_{-i} \sim \mu_0\} - \pi_k^n| < \epsilon/5K$$

for all k and all t_i.

Proof. Let $\epsilon' = \epsilon/5K$. We begin by constructing numbers $\bar{\pi}_k^n$ that are approximately equal to the weights π_k^n in the statement of the claim. Let

$$\bar{\pi}_k^n = \Pr\{t' \in \mathcal{A}_k | t' \in T^n, t' \sim \mu_0\}$$

be the probability that a vector of n types drawn independently according to μ_0 is in $\mathcal{A}_k$. We will show that, for large n, for any type t_i, the $\bar{\pi}_k^n$ are very close to the probability

$$\Pr\{(t_i, t_{-i}) \in \mathcal{A}_k | t_{-i} \in T^{n-1}, t_{-i} \sim \mu_0\}$$

that a vector with agent i's type fixed at t_i and the other types drawn i.i.d. is in $\mathcal{A}_k$. To see this, consider the difference between the probability of a vector of types falling within region

$\mathcal{A}_k$ when i's type is fixed as t_i, minus the probability of falling within $\mathcal{A}_k$ when i's type is drawn randomly. This difference equals

$$\begin{aligned} & \Pr\{(t_i, t'_{-i}) \in \mathcal{A}_k | t'_{-i} \in T^{n-1}, t'_{-i} \sim \mu_0\} - \bar{\pi}_k^n \\ = & \Pr\{(t_i, t'_{-i}) \in \mathcal{A}_k \text{ and } t' \notin \mathcal{A}_k | t' \in T^n, t' \sim \mu_0\} \\ - & \Pr\{(t_i, t'_{-i}) \notin \mathcal{A}_k \text{ and } t' \in \mathcal{A}_k | t' \in T^n, t' \sim \mu_0\}. \end{aligned} \tag{B.9}$$

This expression equals the probability of drawing a vector $t' \in T^n$ where changing agent i's type from t'_i to t_i moves the vector of types from outside $\mathcal{A}_k$ to inside $\mathcal{A}_k$, minus the probability of choosing a vector where changing i's type from t'_i to t_i moves the vector from inside $\mathcal{A}_k$ to outside $\mathcal{A}_k$. We now show that the probability of such vectors being drawn is very small in a sufficiently large market.

Consider the case where $(t_i, t'_{-i}) \notin \mathcal{A}_k$, but $(t'_i, t'_{-i}) \in \mathcal{A}_k$. One possibility is that $(t_i, t'_{-i}) \notin \mathcal{N}$. By the law of large numbers, we may take n_0 large enough such that for $n \geq n_0$ the probability of this happening is less than $\epsilon'/8$. The other possibility is that $(t_i, t'_{-i}) \in \mathcal{N}$, but $(t_i, t'_{-i}) \notin \mathcal{A}_k$. In this case, the line segment connecting $\text{emp}[t_i, t'_{-i}]$ and $\text{emp}[t']$ contains a point in $\mathcal{B}$, because $\mathcal{N}$ is convex and each $\mathcal{A}_k$ is open. This means that the distance between $\text{emp}[t']$ and $\mathcal{B}$ is at most $1/n$. By Condition 1 of Definition B.1, we may take n_0 such that this probability is less than $\epsilon'/8$. This argument implies that we may take n_0 such that, for all $n \geq n_0$,

$$\Pr\{(t_i, t'_{-i}) \notin \mathcal{A}_k, t' \in \mathcal{A}_k | t' \in T^n, t' \sim \mu_0\} < \epsilon'/8 + \epsilon'/8 = \epsilon'/4.$$

An analogous argument proves that n_0 can be chosen such that, for $n \geq n_0$,

$$\Pr\{(t_i, t'_{-i}) \in \mathcal{A}_k, t' \notin \mathcal{A}_k | t' \in T^n, t' \sim \mu_0\} < \epsilon'/4.$$

Substituting these two inequalities into the RHS of equation (B.9) yields that

$$|\Pr\{(t_i, t'_{-i}) \in \mathcal{A}_k | t'_{-i} \in T^{n-1}, t'_{-i} \sim \mu_0\} - \bar{\pi}_k^n| < \epsilon'/4 + \epsilon'/4 = \epsilon'/2. \tag{B.10}$$

Note, however, that the $\bar{\pi}_k^n$ do not necessarily sum to 1, as it may be the case that $t' \notin \cup_k \mathcal{A}_k$. To complete the proof, define

$$\pi_k^n = \bar{\pi}_k^n / \sum_{k'=1}^{K} \bar{\pi}_{k'}^n. \tag{B.11}$$

We have that the probability that $t' \notin \cup_k \mathcal{A}_k$ converges to 0. Therefore, we may take n_0 such that for $n \geq n_0$

$$|1 - 1/\sum_{k'=1}^{K} \bar{\pi}_{k'}^n| < \epsilon'/2. \tag{B.12}$$

We now apply these bounds to prove the claim. We have

$$\begin{aligned}
& |\Pr\{(t_i, t'_{-i}) \in \mathcal{A}_k | t'_{-i} \in T^{n-1}, t'_{-i} \sim \mu_0\} - \pi_k^n| \\
\leq \; & |\Pr\{(t_i, t'_{-i}) \in \mathcal{A}_k | t'_{-i} \in T^{n-1}, t'_{-i} \sim \mu_0\} - \bar{\pi}_k^n| + |\pi_k^n - \bar{\pi}_k^n| \\
< \; & \epsilon'/2 + |\pi_k^n - \bar{\pi}_k^n| \\
= \; & \epsilon'/2 + |\bar{\pi}_k^n / (\sum_{k'=1}^{K} \bar{\pi}_{k'}^n) - \bar{\pi}_k^n| \\
= \; & \epsilon'/2 + |1 - 1/(\sum_{k'=1}^{K} \bar{\pi}_{k'}^n)| \cdot \bar{\pi}_k^n \\
< \; & \epsilon'/2 + \epsilon'/2 = \epsilon/5K.
\end{aligned}$$

The series of steps in the above derivation were as follows. The second line follows from the triangle inequality. The third line uses the bound from inequality (B.10). The fourth line uses the definition of π_k^n from equation (B.11). Finally, the fifth line follows from multiplying $\bar{\pi}_k^n$ out of the right term, and the sixth line comes from inequality (B.12) and $\bar{\pi}_k^n \leq 1$. □

Step 3: completing the proof.

Finally, we apply the results from Steps 1 and 2 to prove the lemma. We have

$$f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi_k^n \cdot z_k(t_i) = \sum_{t_{-i} \in T^{n-1}} \Pr(t_{-i} | t_{-i} \sim \mu_0) \cdot F_i^n(t) - \sum_{k=1}^{K} \pi_k^n \cdot z_k(t_i).$$

This sum can be decomposed depending on whether $\hat{\mu} = \text{emp}[t]$ is in each of the $\mathcal{A}_k$ sets or outside the union of the $\mathcal{A}_k$ sets. We have

$$\begin{aligned}
& f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi_k^n \cdot z_k(t_i) \qquad \text{(B.13)} \\
= \; & \sum_{k=1}^{K} (\{ \sum_{t_{-i} : \hat{\mu} \in \mathcal{A}_k} \Pr(t_{-i} | t_{-i} \sim \mu_0) \cdot F_i^n(t) \} - \pi_k^n \cdot z_k^n(t_i)) + \sum_{t_{-i} : \hat{\mu} \notin \cup_k \mathcal{A}_k} \Pr(t_{-i} | t_{-i} \sim \mu_0) \cdot F_i^n(t).
\end{aligned}$$

We begin by looking at the terms where $\hat{\mu}$ is in one of the $\mathcal{A}_k$. We will show that for each k these terms are small. We have that, for each k,

$$
\begin{aligned}
&\|(\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot F_i^n(t)) - \pi_k^n\cdot z_k(t_i)\| \qquad \text{(B.14)}\\
=\ &\| \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot (F_i^n(t) - z_k(t_i))\\
&+(\{\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\} - \pi_k^n)\cdot z_k(t_i)\|\\
\leq\ & \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot \|F_i^n(t) - z_k(t_i)\|\\
&+|\{\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\} - \pi_k^n|.
\end{aligned}
$$

The equality in the second line follows from rearranging the expression. The inequality in the third line follows from the triangle inequality, and the fact that the norm of the vector $z_k(t_i)\in X$ is weakly less than 1.

Consider now the right hand side of inequality (B.14). By Claim B.1, we may take n_0 such that for all $n\geq n_0$ and t_{-i} such that $\hat{\mu}\in\mathcal{A}_k$,

$$\|F_i^n(t) - z_k(t_i)\| < 3\epsilon/5.$$

By Claim B.2, n_0 may be taken such that, for $n\geq n_0$, the second term is bounded by

$$|\{\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\} - \pi_k^n| < \frac{1}{5K}\epsilon.$$

Substituting these two bounds in inequality (B.14) we have that, for all $n\geq n_0$,

$$
\begin{aligned}
&\|\{\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot F_i^n(t)\} - \pi_k^n\cdot z_k^n(t_i)\|\\
<\ & \frac{3}{5}\epsilon\cdot \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0) + \frac{1}{5K}\epsilon.
\end{aligned}
$$

Summing over all k we get

$$\sum_{k=1}^{K}\|\{\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot F_i^n(t)\} - \pi_k^n\cdot z_k^n(t_i)\| < \frac{3}{5}\epsilon + K\frac{1}{5K}\epsilon = 4\epsilon/5.$$

Using the triangle inequality, the sum operator can be brought into the norm, yielding the inequality

$$\|\sum_{k=1}^{K}(\{\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k}\Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot F_i^n(t)\}-\pi_k^n\cdot z_k^n(t_i))\|<4\epsilon/5. \tag{B.15}$$

The argument above bounds the terms in the RHS of equation (B.13) that correspond to t within the sets $\mathcal{A}_k$. To bound the last term, note that we may take n_0 to be large enough so that, for all $n \geq n_0$, the probability that $t \notin \cup_k \mathcal{A}_k$ is strictly less than $\epsilon/5$. That is,

$$\sum_{t_{-i}:\hat{\mu}\notin\cup_k\mathcal{A}_k}\Pr(t_{-i}|t_{-i}\sim\mu_0)<\epsilon/5. \tag{B.16}$$

Plugging equations (B.15) and (B.16) into equation (B.13) we obtain

$$\|f^n(t_i,\mu_0)-\sum_{k=1}^{K}\pi_k^n\cdot z_k(t_i)\|<\epsilon,$$

completing the proof of Step 3, and hence the lemma.

□

B.2 An Example without Quasi-Continuity

This section shows, by example, that it is necessary to impose regularity conditions on a family of limit equilibria to obtain the results in Theorem B.1. We consider an example of a non quasi-continuous family of equilibria, and show that the mechanism constructed in the proof of Theorem B.1 does not satisfy any of the implications of the theorem. Namely, the constructed mechanism is not SP-L, and outcomes of the constructed mechanism do not approximate the outcomes of the original mechanism, nor a convex combination of the outcomes. We consider the case of BNE of the finite mechanism to highlight that, even for finite economy BNE, the construction used to prove our Theorems 2 and B.1 depends on the quasi-continuity condition. It is a simple matter to extend the logic of the example to limit BNE.

Consider a set of two objects $O = \{o_1, o_2\}$. The set of bundles is $X_0 = O \times \{0, -10\}$, so that a bundle x_0 specifies an object $x_0(1) = o_1$ or o_2, and a transfer $x_0(2) = 0$ or -10 of a numeraire. Therefore, agents either receive no transfer, or pay a fine of 10. The set of types is $T = O = \{o_1, o_2\}$, with an agent's type denoting her favorite object. Utility is given by

$$u_{t_i}(x_0) = \mathbf{1}\{x_0(1) = t_i\} + x_0(2).$$

That is, an agent has utility 1 for receiving an object matching her type, and quasilinear utility on the transfer. Consider the set of actions

$$A = O \times \{f, nf\}.$$

An action a_i specifies an object preference $a_i(1) = o_1$ or o_2, and a message $a_i(2) = f$ (standing for fine) or nf (standing for no fine). We define the mechanism $\{\Phi^n, A\}$ as follows.

- If all $a_j(2) = nf$, $j = 1, \cdots, n$, then $\Phi_i^n(a) = (a_i(1), 0)$. That is, if all agents choose the no fine option, then each agent receives her favorite object and no one is fined.

- If some $a_j(2) = f$, $j = 1, \cdots, n$, then some agents will be fined, depending on whether the number of agents asking for object o_1 is odd or even.

  - If $\#\{j : a_j(1) = o_1\}$ is odd, then agents asking for object o_1 are fined:

  $$\Phi_i^n(a) = (a_i(1), -10 \cdot \mathbf{1}\{a_i(1) = o_1\}).$$

  - If $\#\{j : a_j(1) = o_1\}$ is even, then agents asking for object o_2 are fined:

  $$\Phi_i^n(a) = (a_i(1), -10 \cdot \mathbf{1}\{a_i(1) = o_2\}).$$

We now define a sequence of families of BNE. Let n_0 be a sufficiently large number, and $\delta > 0$ a small positive constant. Let μ_0 be the distribution putting equal weight on o_1 and o_2. Define now the following subset of $\mathbb{N} \times \bar{\Delta}T$,

$$S = \{(n, \mu) \in \mathbb{N} \times \bar{\Delta}T : n \cdot \mu(o_1) \text{ is an odd integer}, n \geq n_0, \|\mu - \mu_0\| < \delta\}.$$

That is, S is the set of all pairs of a number of players and a distribution over types such that, in a type profile with n types and empirical distribution of types μ, the number of players with $t_i = o_1$ is odd. Moreover, n has to be larger than n_0 and μ sufficiently close to μ_0.

Consider now the following sequence of families $(\sigma_\mu^n)_{\mu \in \Delta T, n \in \mathbb{N}}$ of BNE of this mechanism.

- If $(n, \mu) \in S$, then σ_μ^n specifies that agents play actions that match their types $a_i(1) = t_i$, and send the fine message $a_i(2) = f$.
- Otherwise agents play actions that match their types $a_i(1) = t_i$, but send the no fine message $a_i(2) = nf$.

Note that, for suitably chosen n_0 and δ, this is a family of equilibria. If $(n, \mu) \notin S$, then it is optimal for agents to request their favorite object, as no agents are fined. If $(n, \mu) \in S$, then sending the f or nf message is immaterial, as fines are always activated since all other players send the f message in equilibrium. Moreover, it is optimal to request one's favorite object ($a_i(1) = t_i$), as the probability that agents requesting objects o_1 or o_2 are fined are both approximately equal to $1/2$.

Note also that this family of BNE is not quasi-continuous at μ_0. To see this formally, take a neighborhood $\mathcal{N}$ of μ_0 small enough such that the set

$$\{\mu \in \mathcal{N} : \exists n \text{ with } (n, \mu) \in S\} \tag{B.17}$$

is relatively dense with respect to $\mathcal{N}$. Any open subset $\mathcal{A}_k$ of this neighborhood therefore contains infinite points in this set (B.17), and infinite points outside this set (B.17). In particular, for any n_0, there exist $n \geq n_0$ and μ and emp$[t]$ in $\mathcal{A}_k$ where type o_1 agents are not fined, i.e., $\Phi_i^n(o_1, \sigma_\mu^n(t_{-i})) = (o_1, 0)$, and similarly there exist $n \geq n_0$ and μ and emp$[t]$ in $\mathcal{A}_k$ where type o_1 agents are fined, i.e., $\Phi_i^n(o_1, \sigma_\mu^n(t_{-i})) = (o_1, -10)$. This violates Condition 2 of the definition of quasi-continuity.

Define now the direct mechanism $((F^n)_{n \in \mathbf{N}}), T)$ such that

$$F^n(t) = \Phi^n(\sigma^n_{\text{emp}[t]}(t)).$$

This is the construction used in the proof of Theorem 2. We now show that this mechanism is neither SP-L, nor does it approximate outcomes of the σ_μ^n equilibria. Consider a type profile t such that $(n, \text{emp}[t]) \notin S$. Then the no fine equilibrium is played and

$$F_i^n(t) = (t_i, 0).$$

That is, the mechanism simply assigns the requested object to each agent, and there are no fines. However, if $(n, \text{emp}[t]) \in S$, we have

$$F_i^n(t) = (t_i, -10 \cdot \{t_i = o_1\}).$$

That is, the mechanism assigns the requested object to each agent, but only fines the $t_i = o_1$ types. This happens because the equilibria σ^n_μ where agents send the fine message are played exactly at the profiles where the number of o_1 reports is odd, and therefore where agents reporting o_1 are fined. In contrast, agents reporting o_2 are never fined. Note that, if types are distributed according to μ_0, the probability that $(n, t) \in S$ converges to $1/2$ as the number of players grows. We have that the constructed mechanism has a limit

$$\begin{aligned} f_i^\infty(t_i, \mu_0) &= \frac{1}{2}(o_1, 0) + \frac{1}{2}(o_1, -10) \text{ if } t_i = o_1 \\ &\quad (o_2, 0) \text{ if } t_i = o_2. \end{aligned}$$

In particular, the constructed mechanism is not SP-L, as a type $t_i = o_1$ agent would prefer to report being a type o_2. Moreover, the above allocation does not approximate a convex combination of allocations received in the sequence of families of equilibria.

C Semi-Anonymity

Our main analysis considers anonymous mechanisms, where agents' outcomes depend on their own report and the distribution of all reports. The analysis generalizes straightforwardly, though at some notational burden, to the case of semi-anonymous mechanisms, as defined by Kalai (2004). In this setting, agents are divided into a number of groups, and agents within each group can be treated differently by the mechanism.

In this section, agents belong to **groups** g in a finite set G. The set of types is partitioned into subsets

$$T = T_{g_1} \cup T_{g_2} \cup \cdots \cup T_{g_G}.$$

A **semi-anonymous mechanism** is defined as $\{(\Phi^n)_{n\in\mathbb{N}}, (A_g)_{g\in G}\}$, where the A_g are the sets of actions available to each group g, and

$$A = A_{g_1} \cup \cdots \cup A_{g_G}$$

is the set of actions. As in the anonymous case, the Φ^n are functions

$$\Phi^n : A^n \to \Delta({X_0}^n).$$

The difference with respect to anonymous mechanisms is that agents in group g are restricted to play strategies in A_g. That is, if $t_i \in T_g$ then the support of any strategy $\sigma(t_i)$

is contained in A_g. In a matching setting, for example, the groups may specify whether an agent is a man or a woman, and the agent's traits. Agents are then permitted to misreport their preferences over other match partners, but they cannot misrepresent their gender or their traits. **Limit mechanisms** are defined as in Section 3.1. In particular, we define limit mechanisms with respect to a single distribution $\mu \in \Delta T$, and not distributions of types within groups. Alternatively, one could assume that the number of agents in each group grows in a specific way, and that types are drawn i.i.d. within each group. We now formally define a two-sided matching mechanism, to clarify the definition.

Example C.1. (Two-Sided Matching) This example shows that semi-anonymous mechanisms include matching mechanisms in two-sided markets (Gale and Shapley, 1962). Agents are men and women, who differ on a set of traits. Groups g index both sex and the traits, so that the set of groups is

$$G = \{m_1, m_2, \cdots, m_M\} \cup \{w_1, w_2, \cdots, w_W\}.$$

That is, there are M groups of men and W groups of women. Men and women within each group have the same traits, and are equally good marriage partners. However, within each group, agents may differ in their preferences over the other groups. The way in which the semi-anonymous framework differs from the anonymous setting is that men and women may misreport their preferences, but cannot misreport their sex nor traits.

Formally, agent i's type is

$$t_i = (g_{t_i}, u_{t_i}),$$

where $g_{t_i} \in G$ is the agent's group, and u_{t_i} is a strictly positive utility function over the groups of the opposite sex. The set of outcomes $X_0 = G \cup \emptyset$. That is, each agent only cares about which type of man (woman) she (he) is matched to, or whether she (he) is unmatched. Utilities of each type t_i are given by $u_{t_i}(g)$ if she is matched to someone of the opposite sex. We extend u_{t_i} so that it is 0 if the agent is unmatched or matched to a group of the same sex.

Consider now a stable matching mechanism, using a tie-breaking lottery, as in school choice mechanisms (Abdulkadiroğlu et al., 2009). The mechanism is direct, so that $A_g = T_g$ for each $g \in G$. Men and women report a vector of types t, and therefore traits. This implies a weak preference ordering of each man over each woman and vice versa. The mechanism assigns a lottery number l_i to each agent, uniformly and independently distributed between 0 and 1. Lottery numbers are used to break ties between preferences. That is, preferences are refined to strict preferences, by using the lottery numbers to break ties. Conditional on a vector of lotteries l and a vector of reported types t, the mechanism implements a stable matching $x^n(t,l)$. The function $x^n(t,l)$ is taken to be symmetric, to conform to the semi-anonymity assumption. The mechanism is then defined as

$$\Phi^n(t) = \int_{l\in[0,1]^n} x^n(t,l)\, dl.$$

$\square$

Define a semi-anonymous mechanism as SP-L if no agent wants to misreport as a different type within the same group.

Definition C.1. *The semi-anonymous mechanism* $\{(\Phi^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ *is* ***strategy-proof in the large (SP-L)*** *if, for any* $m \in \bar{\Delta}T$, $g \in G$, *and* $t_i, t_i' \in T_g$,

$$u_{t_i}[\phi^\infty(t_i, m)] \geq u_{t_i}[\phi^\infty(t_i', m)]. \tag{C.1}$$

Equivalently, the mechanism is SP-L if, for any $m \in \bar{\Delta}T$ *and* $\epsilon > 0$, *there exists* n_0 *such that, for all* $n \geq n_0$, $g \in G$, *and* $t_i, t_i' \in T_g$,

$$u_{t_i}[\phi^n(t_i, m)] \geq u_{t_i}[\phi^n(t_i', m)] - \epsilon.$$

Otherwise, the mechanism is ***manipulable in the large.***

The sufficient conditions for a mechanism to be SP-L also have straightforward extensions. The extension of the EF-TB condition is that no agent envies another agent in the same group, and with lower lottery number.

Definition C.2. *A direct semi-anonymous mechanism* $\{(\Phi^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ *is* ***envy-free but for tie-breaking*** *(EF-TB) if for each* n *there exists a function* $x^n : (T\times[0,1])^N \to \Delta(X_0^n)$, *symmetric over its coordinates, such that*

$$\Phi^n(t) = \int_{l\in[0,1]^n} x^n(t,l)dl$$

and, for all i, j, n, t, *and* l, *if* $l_i \geq l_j$, *and if* t_i *and* t_j *belong to the same group, then*

$$u_{t_i}[x_i^n(t,l)] \geq u_{t_i}[x_j^n(t,l)].$$

With this definition, an extension of Theorem 1 to semi-anonymous mechanisms follows from essentially the same proof.[2] This implies that the stable matching procedure in example C.1 is SP-L, because an agent envying another agent with a lower lottery number would violate the stability condition.

We now extend the definition of limit BNE to this setting, and state and prove an extension of Theorem B.1. The conclusions of the theorem are unchanged, and the only difference is that it considers a family of limit equilibria of a semi-anonymous mechanism, and not an anonymous mechanism. The proof uses a construction identical to that in Theorem B.1. The proof follows from noting that the argument in the anonymous case implies that the approximation formulas in Theorem B.1 hold, and then showing that this implies that the constructed semi-anonymous mechanism is SP-L.

We must first extend the concept of a limit BNE. The difference with respect to the anonymous case is that in the semi-anonymous case it is only necessary to rule out deviations where agents of group g play other actions in A_g, or, in a direct mechanism, report being a different type in T_g. A **strategy** is defined as a map $\sigma : T \to \Delta A$ such that if $t_i \in T_g$ then the support of $\sigma(t_i)$ is contained in A_g.

Definition C.3. *Given a semi-anonymous mechanism $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$ with limit $\phi^\infty(\cdot,\cdot)$, and a probability distribution over types $\mu \in \Delta T$, the strategy $\sigma^*_\mu : T \to \Delta A$ is a **limit Bayes-Nash Equilibrium at prior** μ if, for all $g \in G$, $t_i \in T_g$ and $a'_i \in A_g$:*

$$u_{t_i}[\phi^\infty(\sigma^*_\mu(t_i), \sigma^*_\mu(\mu))] \geq u_{t_i}[\phi^\infty(a'_i, \sigma^*_\mu(\mu))].$$

The definition of **(quasi-) continuous** families of limit equilibria is identical to the anonymous case. With these definitions, the statement of the semi-anonymous version of Theorem B.1 is similar to the anonymous case, with the key difference being the larger class of mechanisms considered.

Theorem C.1 (Extension of Theorem B.1 to semi-anonymous mechanisms)**.** *Given any semi-anonymous mechanism $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$ with a quasi-continuous family of limit Bayes-Nash equilibria $(\sigma^*_\mu)_{\mu\in\Delta T}$, there exists a direct, SP-L, semi-anonymous mechanism $\{(F^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ with the following properties.*

[2]Lemma A.1 holds as is, since it is a statement about the empirical distribution of randomly drawn vectors of types, and therefore does not rely on the definition of a mechanism. Lemma A.2 holds for any two types t_i and t'_i in the same group, using the same proof, as for any such pairs of types the EF-TB condition in the semi-anonymous case implies the same properties as in the anonymous case. Given the two lemmas, the argument in the proof of Theorem 1 in Appendix A.1 holds as is, as long as we take t'_i to be in the same group as t_i, which is all that is needed for the definition of SP-L for semi-anonymous mechanisms.

1. *If the original mechanism is continuous at a prior $\mu_0 \in \bar{\Delta}T$ then, in the limit, truthful play of the direct mechanism produces the same outcomes as equilibrium play of the original mechanism. Formally, for any $t_i \in T$, we have*

$$f^\infty(t_i, \mu_0) = \phi^\infty(\sigma^*_{\mu_0}(t_i), \sigma^*_{\mu_0}(\mu_0)),$$

where f^∞ is the limit of the direct mechanism.

2. *For any prior $\mu_0 \in \bar{\Delta}T$, in the large market limit, truthful play of the direct mechanism produces the same outcomes as a convex combination of equilibrium play of the original mechanism under priors that are close to μ_0. Formally, for any $\epsilon > 0$, there exists n_0, an integer K, numbers π^n_k with $\sum_{k=1,\cdots,K} \pi^n_k = 1$, and priors μ_k with $\|\mu_k - \mu_0\| < \epsilon$ such that, for all $n \geq n_0$ and $t_i \in T$, we have*

$$\|f^n(t_i, \mu_0) - \sum_{k=1,\cdots,K} \pi^n_k \cdot \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))\| < \epsilon,$$

where f^n is the function representing the direct mechanism from an interim perspective, as defined in equation (3.1).

Proof. Let F be defined as in equation (5.2). Parts 1 and 2 of the theorem statement follow from the same argument as in the proof of Theorem B.1. This is the case because $\{(\Phi^n)_\mathbb{N}, \cup_{g\in G} A_g\}$ is an anonymous mechanism. Moreover, parts 1 and 2 do not rely on agents playing optimally under σ^*_μ, and the definitions of (quasi-) continuity of a family of limit equilibria are the same in the anonymous and semi-anonymous case. Likewise, Lemma B.1 holds in the semi-anonymous setting, with essentially the same proof.

It only remains to be proven that the direct mechanism is SP-L. To establish this, we employ a small modification of the original argument, as now the σ^*_μ are limit equilibria of a semi-anonymous mechanism. As in the proof of Theorem B.1, take $\mu_0 \in \bar{\Delta}T$, and $\epsilon > 0$. By Lemma B.1 (with $\frac{\epsilon}{2|X_0|}$ as the constant), there exist a neighborhood $\mathcal{N}$ of μ_0, priors μ_k and weights π^n_k for $k = 1, \cdots, K$, and n_0 such that, for all $n \geq n_0$ and t'_i in T we have

$$\begin{aligned} \sum_{k=1}^K \pi^n_k &= 1, \\ \|\mu_k - \mu_0\| &< \epsilon, \text{ and} \\ \|f^n(t'_i, \mu_0) - \sum_{k=1}^K \pi^n_k \cdot z_k(t'_i)\| &< \frac{\epsilon}{2|X_0|} \leq \frac{\epsilon}{2}, \end{aligned} \tag{C.2}$$

where

$$z_k(t_i') = \phi^{\infty}(\sigma^*_{\mu_k}(t_i'), \sigma^*_{\mu_k}(\mu_k)).$$

Take now any two types t_i and t_i' in the same group. We have that

$$\begin{aligned} & u_{t_i}[f^n(t_i', m)] - u_{t_i}[f^n(t_i, m)] \\ \leq \; & \sum_{k=1,\ldots,K} \pi_k^n \cdot \{u_{t_i}[z_k(t_i')] - u_{t_i}[z_k(t_i)]\} \\ & + |u_{t_i}[f^n(t_i', \mu_0)] - u_{t_i}[\sum_{k=1}^{K} \pi_k^n \cdot z_k(t_i')]| + |u_{t_i}[f^n(t_i, \mu_0)] - u_{t_i}[\sum_{k=1}^{K} \pi_k^n \cdot z_k(t_i)]|. \end{aligned}$$

Consider now the RHS of this inequality. From the definition of z_k and of a limit equilibrium, we have that the first sum is nonpositive. Moreover, by the bound (C.2), the fact that utility is in $[0,1]$, and that the set of random bundles has X_0 dimensions, the second and third terms are each bounded by $\epsilon/2$. Therefore, we have that

$$u_{t_i}[f^n(t_i', m)] - u_{t_i}[f^n(t_i, m)] < 0 + \epsilon/2 + \epsilon/2 = \epsilon.$$

Since this holds for all t_i and t_i' in the same group, we have that the constructed semi-anonymous mechanism is SP-L. □

D Details for Table 1

This Section provides supporting details for the classification of non-SP mechanisms presented as Table 1. For each mechanism we provide a formal definition of the mechanism in our setting, a formal proof of the classification, and relevant references. The analysis of multi-unit auctions in Section D.1.1 is especially detailed and illustrates why the interim approach to taking the large-market limit is crucial for obtaining the classification.

D.1 Anonymous Mechanisms.

D.1.1 Multi-Unit Auctions

We consider multi-unit auctions for identical goods, such as government bond auctions. The two most common formats are uniform-price auctions and pay-as-bid auctions. While neither mechanism is SP (Ausubel and Cramton, 2002), Milton Friedman famously argued in favor of the uniform-price auction on incentives grounds (Friedman, 1960, 1991). We will show that uniform-price auctions are SP-L, whereas pay-as-bid auctions are manipulable in the large.

There are kn units of a homogeneous good. To simplify notation, we assume that agents assign a constant per-unit value to the good, up to a capacity limit. Specifically, each agent i's type t_i consists of a per-unit value v_i and a maximum capacity q_i. The set of possible values is $V = \{1, \ldots, \bar{v}\}$, the set of possible capacity limits is $Q = \{0, 1, \ldots, \bar{q}\}$ with $1 < k < \bar{q}$, and $T = V \times Q$. The set of outcomes is $X_0 = (\{1, 2, \cdots, \bar{v}\} \times \{1, 2, \cdots, \bar{q}\}) \cup \{0\}$, with an outcome consisting either of a per-unit payment and an allotted quantity, or 0 to denote that the agent receives no units and makes no payment.

We first describe the uniform-price auction. Bids consist of a per-unit value and a maximum capacity, so the action set $A = T$. Given a vector of n bidders' reports t, let $D(p;t)$ denote the demand for the object at price p.[3] The market-clearing price is

$$p^*(t) = \max\{p \in V : \frac{D(p;t)}{n} \geq k\}. \tag{D.1}$$

That is, $p^*(t)$ is the highest price at which demand weakly exceeds supply. The uniform-price auction allocates each bidder i her demanded quantity at $p^*(t)$, with the exception that bids with $v_i = p^*(t)$ are rationed with equal probability. Formally, $\Phi_i^n(t)$ allots each bidder the following number of units of the good,

Reported Value	Expected Number of Units
$v_i < p^*(t)$	0
$v_i = p^*(t)$	$\bar{r} \cdot q_i$
$v_i > p^*(t)$	q_i

at a price per unit of $p^*(t)$. The probability of a bid being rationed $\bar{r}$ is set to clear the market.[4]

We now analyze the large-market limit of the uniform-price auction. Let $\rho^*(m)$ denote the price that clears supply and *average demand* given bid distribution m:

$$\rho^*(m) = \max\{p \in V : E[D(p;t_i)|t_i \sim m] \geq k\}. \tag{D.2}$$

[3]Formally, $D(p;t) = \sum_{i=1}^{n} q_i \cdot 1\{v_i \geq p\}$, where the notation $1\{\cdot\}$ denotes the indicator function.

[4]Because preferences are linear up to the capacity limit, the exact form of the rationing is immaterial. The rationing probability is

$$\bar{r} = \frac{kn - D(p^*(t)+1;t)}{D(p^*(t);t) - D(p^*(t)+1;t)}.$$

Generically, expected demand at price $\rho^*(m)$ strictly exceeds supply, that is,

$$E[D(\rho^*(m); t_i)|t_i \sim m] > k.$$

In this generic case, as the market grows large, the realized price as defined in (D.1) equal $\rho^*(m)$ with probability converging to one. Therefore, the limit mechanism allocates each bidder their demand at $\rho^*(m)$, with the exception that bidders with value exactly equal to $\rho^*(m)$ are rationed, and with all winning bidders paying $\rho^*(m)$ per unit. Formally, $\phi^\infty(t_i, m)$ gives player i

Reported Value	Expected Number of Units
$v_i < \rho^*(m)$	0
$v_i = \rho^*(m)$	$\bar{r} \cdot q_i$
$v_i > \rho^*(m)$	q_i

at a per unit price of $\rho^*(m)$, and the rationing probability $\bar{r}$ is set so that the market clears on average.[5] Note that, in this generic case, the price in the limit is deterministic and is exogenous from the perspective of each individual bidder.

In addition to the generic case, there is a knife-edge case, in which expected demand at $\rho^*(m)$ is exactly equal to supply, that is, $E[D(\rho^*(m); t_i)|t_i \sim m] = k$. In this case, focusing for now on m with full support, the price is stochastic even in the large-market limit. Given large n, the realized per-capita demand at price $\rho^*(m)$ will be weakly greater than per-capita supply k with probability of about $\frac{1}{2}$, and will be strictly smaller than per-capita supply k with probability of about $\frac{1}{2}$.[6] Therefore, the price in the limit will be $\rho^*(m)$ with probability of $\frac{1}{2}$, and $\rho^*(m) - 1$ with probability of $\frac{1}{2}$. $\phi^\infty(t_i, m)$ assigns to player i the following expected number of units,

Reported Value	Expected Number of Units
$v_i < \rho^*(m)$	0
$v_i \geq \rho^*(m)$	q_i

and prices are $\rho^*(m)$ or $\rho^*(m) - 1$ with equal probability. Note that bids of $\rho^*(m)$ are not

[5] That is, $\bar{r}$ satisfies

$$\bar{r} = \frac{k - E[D(\rho^*(m) + 1; t_i')|t_i' \sim m]}{E[D(\rho^*(m); t_i')|t_i' \sim m] - E[D(\rho^*(m) + 1; t_i')|t_i' \sim m]}.$$

[6] The intuition is that if a fair coin is tossed $n \to \infty$ times, the probability that at least $n/2$ of the tosses are heads converges to $1/2$, just as the probability that less than $n/2$ of the tosses are heads converges to $1/2$, with both probabilities independent of the outcome of the i^{th} toss.

rationed in the limit. This is so because, in this knife-edge case, average demand is exactly equal to average supply. Moreover, in both cases the price in the limit is exogenous from the perspective of each individual bidder. Even though the price is sometimes $\rho^*(m)$ and sometimes $\rho^*(m) - 1$, the probability that bidder i is pivotal in determining which of the two prices occurs converges to zero.

The argument that the uniform-price auction is SP-L is now straightforward. Choose any type t_i and any full support distribution $m \in \bar{\Delta}T$. The description of ϕ^∞ above implies that truthful reporting is a dominant strategy in the limit, hence Definition 4 is satisfied.

Note that this argument would not go through had we used a stronger notion of approximate strategy-proofness based on realizations of opponents' reports rather than probability distributions. In any size market, it is always possible to construct a profile of opponent bids t_{-i} where, ex-post, bidder t_i can profitably lower the market-clearing price by shading his quantity demanded. Similarly, our argument would not go through if SP-L required equation (3.2) to obtain for all probability distributions $m \in \Delta T$, rather than for all full support probability distributions $m \in \bar{\Delta}T$. Full support ensures that the probability that any particular bidder is pivotal goes to zero as the market grows large. See Swinkels (2001; Section 5) for an elegant example, with limited support, in which bidders remain pivotal with probability one even in very large markets.

Last, we turn to the pay-as-bid auction. The pay-as-bid auction allocates units of the good in exactly the same way as the uniform-price auction. The difference is that winning bidders pay their bid instead of the market-clearing price $p^*(t)$. Clearly, bidders will gain from misreporting their value, even in the large-market limit. If the distribution of opponent bids is m and the limit price is $\rho^*(m)$, then a bidder of type $t_i = (v_i, q_i)$ with $v_i > \rho^*(m) + 1$ strictly prefers to misreport as $t_i' = (\rho^*(m) + 1, q_i)$: he receives the same allocation in the limit but pays a strictly lower price per unit. Hence, the pay-as-bid auction is not SP-L.

D.1.2 Single-Unit Assignment

In single-unit assignment problems, each agent is to be assigned at most one indivisible object, and there are no transfers. We refer the reader to Kojima and Manea (2010) and references therein for a detailed description of the environment and applications.

Formally, we define single-unit assignment as follows. Denote the set of object types by X_0. In a market of size n there are $\{q_{x_0} \cdot n\}$ units of object type x_0 available.[7] An agent of type $t_i \in T$ has a strict utility function u_{t_i} over X_0. It is assumed that X_0 includes a

[7] A bracketed expression denotes the nearest integer to the real number within brackets.

null object $\emptyset$, in supply $n - \sum_{x_0' \neq \emptyset}\{q_{x_0'} \cdot n\} \geq 0$, so that the total quantity of objects equals n. The utility of the null object is normalized to 0. Therefore, we assume that all agents strictly prefer any other object (termed a proper object) to the null object.

Boston Mechanism

The Boston mechanism is a mechanism used in many cities to allocate seats in public schools. Abdulkadiroğlu and Sönmez (2003) show that the Boston mechanism is not SP, and Abdulkadiroğlu et al. (2006) document that it was extensively manipulated in practice. We now formally define the Boston mechanism and show that it is not SP-L. This complements an example given by Kojima and Pathak (2009), in a formally different environment, where the Boston mechanism can be manipulated in a large market.

We now define the Boston mechanism. Fix a vector of reports t. To be consistent with the literature we will use the terminology of schools (the objects) and students (the agents). The mechanism first assigns to each student a lottery number l_i, uniformly and independently distributed in $[0, 1]$. The mechanism then proceeds in rounds, following the algorithm below.

1. The mechanism begins in $\texttt{round} = 1$. All students are initially unassigned.
2. Students that are still present in the mechanism take turns, in the order of their lottery number, with higher lottery numbers going first. In her turn student i is permanently assigned to her $\texttt{round}^{\text{th}}$ choice, as given by u_{t_i}, if there are still seats in that school, or remains unassigned otherwise.
3. If all students have been assigned, finish, otherwise increase `round` by 1 and go to Step 2.

Note that the algorithm must finish, as eventually all students are assigned either to a proper school or to the null school $x_0 = \emptyset$. Therefore, conditional on a vector of types t and lottery numbers l the mechanism produces a well-defined outcome $x^n(t, l)$. Before lottery draws, the mechanism is defined as

$$\Phi^n(t) = \int_{l \in [0,1]^n} x^n(t, l) dl.$$

We now show that the Boston mechanism is not SP-L. Consider an economy with two proper schools, $x_0 = A, B$, and the null school $x_0 = \emptyset$, corresponding to being unmatched. That is, $X_0 = \{A, B, \emptyset\}$. Let $q_A = q_B = 1/6$. Consider a distribution $m \in \bar{\Delta}T$ such that 2/3 of the agents prefer school A, while only 1/3 prefer school B. Then, in a large market, the proper schools are filled in the first round with probability close to 1. Therefore, an agent has a negligible chance of getting her second choice. The chance of getting her first choice is $(1/6)/(2/3) = 1/4$ for school A and $(1/6)/(1/3) = 1/2$ for school B. That is, the limit mechanism is

$$\begin{aligned} \phi^{\infty}(t_i, m) &= \frac{1}{4} \cdot A + \frac{3}{4} \cdot \emptyset \text{ if } u_{t_i}[A] > u_{t_i}[B] \\ & \quad \frac{1}{2} \cdot B + \frac{1}{2} \cdot \emptyset \text{ otherwise.} \end{aligned} \tag{D.3}$$

Note in particular that an agent who prefers school A faces a tradeoff when reporting her preferences. If she announces that she prefers school A, she will be assigned to it with 1/2 the chance she has of receiving school B. Therefore, it is not optimal for an agent with $u_{t_i}[A] > u_{t_i}[B] > u_{t_i}[A]/2$ to report truthfully.

Probabilistic Serial Mechanism

The probabilistic serial mechanism has been proposed as a solution to the assignment problem by Bogomolnaia and Moulin (2001). The mechanism works as follows. With time running continuously, agents "eat" probability shares of their favorite object, out of all objects still available. After probability shares of all objects are assigned, the objects are randomly assigned to agents according to these probabilities. We refer the reader to Kojima and Manea (2010) page 110 for a formal definition of the mechanism, as their analysis includes ours as a particular case.

Bogomolnaia and Moulin (2001) show that the mechanism is EF. Consequently, Theorem 1 guarantees that it is SP-L. Note that the fact that this mechanism is SP-L is a particular case of Kojima and Manea's Theorem 1.

Hylland and Zeckhauser Pseudo-Market Mechanism

Hylland and Zeckhauser (1979) proposed a pseudo-market mechanism for single-unit assignment, in which agents are endowed with equal budgets of an imaginary currency which they use to purchase probability shares of the objects. The mechanism works as follows. First, agents report their types, t. Second, the mechanism allocates each agent an equal budget $B > 0$ of an artificial currency. Third, the mechanism computes a competitive equilibrium price vector $p^* \in \mathbb{R}_+^{|X_0|}$. That is, a vector of prices, one for each object type, such that when each agent is allocated her most-preferred affordable bundle of probability shares, based on her reported preferences, the market clears. Last, each agent is allocated her most-preferred affordable bundle at p^*, given her reported preferences. We refer the reader to the original paper for full details.

Hylland and Zeckhauser (1979) prove existence of competitive equilibrium prices in a setting that is strictly more general than ours (in particular, they allow for indifferences).

For each market size n and each possible reported vector of types $t \in T^n$, choose one such price vector in an anonymous manner, and use this price vector to define the resulting allocation $\Phi^n(t)$. As Hylland and Zeckhauser (1979) observe on page 307, since each agent has the same budget and faces the same prices, such a mechanism is EF. Consequently, Theorem 1 guarantees that it is SP-L.

D.1.3 Multi-Unit Assignment

In multi-unit assignment problems, each agent is to be assigned a finite number of indivisible objects. Transfers of a numeraire are not allowed. A prototypical application is the allocation of courses to students at business schools. For further details we refer the reader to Budish (2011).

Denote the finite set of object types by J. Each object j is available in supply $\{q_j \cdot n\}$. A bundle $x_0 \in X_0 = \mathcal{P}(J)$ specifies a subset of the object types.[8] A type t_i specifies a utility function u_{t_i} over bundles. We will adopt the terminology of course allocation, denoting object types by courses, and agents by students.

HBS Draft Mechanism

The mechanism used by Harvard Business School to allocate MBA courses was studied empirically by Budish and Cantillon (2012). Using survey data, they showed that students often misreport their preferences. Here we formally define the mechanism and show that it is not SP-L.

The HBS draft mechanism does not allow students to express preferences over bundles of courses. Instead, students submit a preference ordering over single courses. To examine the possibility of truthful reporting, we restrict our attention to preferences over bundles that are responsive to preferences over individual courses, with preferences over individual courses strict. We will say that a student of type t_i prefers course j_A to course j_B if she prefers a bundle consisting only of course j_A to a bundle consisting only of course j_B, that is, $u_{t_i}(\{j_A\}) > u_{t_i}(\{j_B\})$.

The HBS draft mechanism works as follows. First, each student is assigned a lottery number l_i, uniformly distributed in $[0, 1]$. In the first round, students take turns ordered by their lottery number, with higher lottery numbers going first. At her turn, student i chooses her favorite course out of the ones that are still available. In round two, the same procedure is repeated, but with students with lower lottery numbers going first. The procedure is repeated in the following rounds, with higher lottery numbers going first in the odd rounds

[8] $\mathcal{P}(J)$ denotes the power set of J.

and last in the even rounds. The mechanism ends after k rounds, where k is the number of courses required per student.

To see that this mechanism is not SP-L, consider the following example based closely on Example 1 of Budish and Cantillon (2012). There are 4 proper courses, $J = \{j_A, j_B, j_C, j_D\}$, of which students require $k = 2$ courses each. Each course has capacity for $\frac{2}{3}$ of the population, that is $q_j = \frac{2}{3}$ for each $j \in J$. Consider a probability distribution over students' reports where $\frac{1}{3}$ of the population lists courses in the order j_A, j_B, j_C, j_D, $\frac{1}{3}$ lists courses in the order j_B, j_A, j_C, j_D, and $\frac{1}{3}$ lists courses in the order j_A, j_C, j_D, j_B. Given this distribution of reports, the probability that course j_A reaches capacity either in the end of the first round, or early in the second round converges to 1, as the market grows large. Therefore, a student that ranks course j_A as her first choice has probability close to 1 of receiving it, while a student who ranks j_A second has probability close to 0 of receiving it. In contrast, course j_B is very likely to reach capacity either late in the second round, or early in the third round, in a large market. Consequently, a student who ranks course j_B either first or second is very likely to receive it. For this reason, a student whose true preference order is j_B, j_A, j_C, j_D profits by misreporting as j_A, j_B, j_C, j_D. By doing so, the student receives both j_A and j_B, her two favorite courses, rather than courses j_B and j_C if she reports truthfully.[9]

The Bidding Points Auction Mechanism

The bidding points auction mechanism is used by several business schools to allocate MBA courses. It has been described by Sönmez and Ünver (2010) and Krishna and Ünver (2008), who demonstrated that the mechanism is flawed in several important ways, despite its widespread use. We now define the bidding points auction mechanism and show that it is not SP-L.

The mechanism works as follows. Students report vectors of bids, with one bid per course. Students can only spend up to a budget of B points, so that the set of actions is the set of all vectors of bids that sum to at most B. We restrict the bids to be integers, so that

$$A = \{a_i \in \{0, 1, \cdots, B\}_+^J : \sum_j a_{i,j} \leq B\}.$$

Given a vector of bids, the mechanism starts with the highest bid and allocates the course to the student, as long as the course still has capacity. Ties are broken randomly.

To examine the possibility of truthful reporting, we assume that students' preferences are

[9]This particular profitable misreport is valid for any cardinal preferences consistent with the ordinal preferences j_B, j_A, j_C, j_D. In other examples the profitability of a particular misreport might depend on cardinal preference information.

additive, meaning that their utility for a bundle of courses is the sum of their utilities from the component courses in that bundle. This allows us to interpret a student's bid vector as an expression of their individual course preferences, and allows us to interpret the bidding points auction as a direct mechanism with $T = A$.

Consider the case where there are three courses, $J = \{j_A, j_B, j_C\}$. Consider an agent who likes the three courses j_A, j_B, j_C equally, and derives no utility of being unmatched. That is,

$$\begin{aligned} u_{t_i}(j_A) &= u_{t_i}(j_B) = u_{t_i}(j_C) = B/3, \\ u_{t_i}(\emptyset) &= 0. \end{aligned} \tag{D.4}$$

Consider a distribution of play m, such that, in the large-market limit, the last accepted bid for the courses j_A, j_B, j_C is $2B/3$ with very high probability. In that case, the agent should not report her true preferences, with bids equal to her utility. If bids are given by equation (D.4), then with very high probability the agent does not receive any course. If instead she bids B for one of the courses she likes, and 0 for the others, she receives at least one of the courses. Therefore, the mechanism is not SP-L.

Approximate Competitive Equilibrium from Equal Incomes (A-CEEI)

Budish (2011) proposed a pseudo-market mechanism for multi-unit assignment problems. Budish's setting is a strict generalization of ours. For that reason, we do not repeat all formal definitions, and refer the reader to the original paper for further details. In our setting, the A-CEEI mechanism can be defined as follows. First, assign each student a lottery number l_i uniformly and identically distributed in $[0, 1]$. Then give each student a budget in an imaginary currency of $1 + l_i \cdot \beta_{(n)}$, where $\beta_{(n)}$ is a strictly positive constant that is weakly decreasing in n, as defined in Budish (2011) page 1081. Budish's Theorem 1 guarantees that given these budgets there exists an approximate competitive equilibrium of the economy where agents purchase courses using the imaginary currency. The A-CEEI mechanism selects one such equilibrium, anonymously, and gives each agent his equilibrium allocation. This defines a function $x^n(\cdot, \cdot)$ giving an assignment of bundles $x^n(t, l) \in X_0^n$, for each vector of types t and lottery draws l. The A-CEEI mechanism is defined as

$$\Phi^n(t) = \int_{l \in [0,1]^n} x^n(t, l) dl.$$

To show that this mechanism is SP-L, we use Theorem 1. By the definition of approximate competitive equilibrium (Budish's Definition 1), after lotteries are drawn, no agent envies another agent with a lower lottery number. Therefore, the CEEI mechanism is EF-TB, and therefore SP-L.

The Generalized Hylland and Zeckhauser Pseudo-Market

Budish et al. (2013) have proposed an extension of the Hylland and Zeckhauser pseudo-market mechanism that can be used for multi-unit assignment problems. In the simplest setting they consider, students have additive preferences over courses. We therefore assume that T only includes additive preferences. With this assumption, their setting is a strict generalization of ours. Budish et al. (2013) then formally define the mechanism. It works similarly to the Hylland and Zeckhauser mechanism, with students purchasing probability shares of courses using a fake currency. The mechanism then calculates a competitive equilibrium allocation of probability shares. Finally, the mechanism implements a lottery over allocations that gives each agent her equilibrium probability share. Budish et al.'s Theorem 6 and Corollary 3 guarantee that the mechanism is well-defined, as both an equilibrium exists and can be implemented by a lottery over feasible assignments. Budish et al.'s Theorem 8 shows that the mechanism is envy-free. Along with our Theorem 1, this implies that the mechanism is SP-L.

D.1.4 Exchange Economies

Walrasian Mechanism

A Walrasian mechanism implements competitive equilibrium allocations in an exchange economy. Several contributions in the literature have considered approximate incentive compatibility of Walrasian mechanisms in large markets, including the classic paper by Roberts and Postlewaite (1976). We refer the reader to Jackson and Manelli (1997) for an overview and references. We note that this example has an infinite set of bundles X_0, which does not fit the framework in the body of the paper. However, the mechanism fits the more general framework considered in Appendix A.1.2, which allows us to use Theorem 1 to classify it as SP-L.

We consider an exchange economy with J goods. A type $t_i = (e_{t_i}, v_{t_i})$ specifies

- An endowment vector $e_{t_i} \in \mathbb{R}^J_+$.
- A continuous utility function v_{t_i} over bundles of goods in $\mathbb{R}^J_+$, taking values in $[0, 1]$.

Assume that the finite set of types T is such that, for any finite n and type vector $t \in T^n$, there always exists at least one competitive equilibrium where all agents of the same type receive the same bundle. This is guaranteed under standard assumptions on the set of utility functions and endowment vectors.

Given a type t_i, we define the utility function u_{t_i} over net trades $x_0 \in \mathbb{R}^J$ as

$$\begin{aligned} u_{t_i} &= v_{t_i}(e_{t_i} + x_0) \text{ if } e_{t_i} + x_0 \in \mathbb{R}^J_+ \\ & \quad -\infty \text{ if } e_{t_i} + x_0 \notin \mathbb{R}^J_+. \end{aligned}$$

We let X_0 be $\mathbb{R}^J$, the set of all possible vectors of net trades.

Having defined X_0 and T, we now define the mechanism. For all n, t, $\Phi^n(t)$ anonymously selects a competitive equilibrium allocation of an economy with the n agents of types in the vector t, such that agents of the same type receive the same bundle, and assigns each agent i her vector of net trades in that equilibrium.

Note that the Walrasian mechanism is EF, as each agent receives her preferred vector of net trades given prices. Furthermore, while X_0 is not finite, it does satisfy the more general assumptions in Remark 1. Namely, X_0 is a measurable subset of Euclidean space, utility is measurable and bounded above by 1, and the utility of telling reporting truthfully is at least 0. Therefore, by Theorem 1, the Walrasian mechanism is SP-L.

D.2 Semi-Anonymous Mechanisms

Semi-anonymity generalizes anonymity to allow a mechanism to treat agents differently if they belong to identifiably distinct groups. Examples include treating men and women differently in a matching mechanism, and treating buyers and sellers differently in a double auction. While the body of the paper deals with the notationally simpler case of anonymous mechanisms, semi-anonymous mechanisms are analyzed in Appendix C. This subsection classifies some of these mechanisms.

D.2.1 Double Auctions

Double auctions have been extensively studied as a simplified model of price formation. We consider auctions where buyers and sellers submit bids, and prices are given as the average of marginal winning and losing bids. See for example Rustichini et al. (1994) for further details and references.

Types t_i specify whether an agent is a potential buyer or seller, and a value. That is, types specify the agent's group, which is $g_{t_i} = b$(uyer) or s(eller), and her value for the object, which is v_{t_i}. Sellers are endowed with a unit of the object, while buyers are not. The set of types is $T = G \times V$, with $G = \{b, s\}$ and $V = \{1, \cdots, \bar{v}\}$. A bundle x_0 specifies whether the agent trades or not, with a dummy $d_{x_0} = 0$ or 1, and the price of the transaction

$$p_{x_0} \in P = \{(p' + p'')/2 : p', p'' \in V\}.$$

We have $X_0 = \{0,1\} \times P$. Buyers and sellers have quasilinear utility. The utility of a bundle is 0 if the agent does not trade. If the bundle prescribes a trade, utility is $v_{t_i} - p_{x_0}$ for a buyer, and $p_{x_0} - v_{t_i}$for a seller.

The mechanism works as follows. Given t, let $n_s(t)$ be the number of sellers, and therefore the number of objects. The market clearing price is the average of the $n_s(t)^{\text{st}}$ and $n_s(t)+1^{\text{st}}$ highest valuations. The mechanism assigns bundles x_0 with this price to all agents. The objects are assigned to the agents with the $n_s(t)$ highest valuations, with uniform tie-breaking for agents tied with the lowest winning valuation. Formally, the mechanism $\Phi^n(t)$ assigns bundles x_0 specifying trade to all buyers with valuations higher than the price, all sellers with valuations lower than the price, and randomly rations agents with valuations equal to the price.

Note that the mechanism is envy-free. This is so because all agents pay the same price, and therefore do not envy the price paid by other agents. Moreover, at this price, agents who trade with probability 1 would rather trade than not trade, and likewise agents that trade with probability 0 would rather not trade. Agents that are rationed are indifferent between trading or not trading, and therefore the mechanism is envy-free.[10] Therefore, double auctions are SP-L.

D.2.2 Matching

This setting is defined formally in Section C, Example C.1. That section also defines stable matching mechanisms, which are shown to be SP-L using a semi-anonymous version of the EF-TB condition.

Priority Match

Priority match mechanisms are described by Roth (1991), who proved that these mechanisms can produce unstable outcomes. Roth also documented that labor market clearinghouses using priority matching mechanisms were very likely to fail, and hypothesized that the reason why they failed is that they produce unstable outcomes.

The priority match works as follows. Given a man i (woman) and a woman (man) j define the rank of i on j's preferences as 1 plus the number of men (women) who are strictly

[10] Note that agents are only rationed in the case of a tie between the marginal winning and losing bids, and therefore both of these bids equal the price.

preferred to i. Assign to the pair i, j the priority $p_{i,j}$ equal to the rank of the man in the woman's preferences, times the rank of the woman in the man's preferences. The mechanism then proceeds by matching pairs with the lowest priorities first, breaking ties randomly.

To see that the priority match mechanism is not SP-L, consider the case where there is a single trait for men. Then women are indifferent over all men. In this case, the priority match mechanism coincides with the Boston mechanism, which is not SP-L.

It is interesting to note that Roth (1991) conjectured that the reason why stable matching mechanisms seem to succeed in practice, while priority matching mechanisms lead to unravelling and market failures, is stability. Our analysis, however, shows that stable matching mechanisms are SP-L, while priority matching mechanisms are not. Therefore, Roth's empirical finding can be phrased equivalently as saying that SP-L mechanisms succeed while non SP-L mechanisms fail.

References

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

_ , Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez, "Changing the Boston Mechanism: Strategyproofness as Equal Access," *Mimeo, Harvard University*, 2006.

_ , _ , and _ , "Strategy-Proofness Versus Efficiency in Matching with Indifferences: Redesigning the Nyc High School Match," *The American Economic Review*, 2009, *99* (5), 1954–1978.

Ausubel, Lawrence M. and Peter Cramton, "Demand Reduction and Inefficiency in Multi-Unit Auctions," 2002. Mimeo, University of Maryland.

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.

Budish, Eric B. and Estelle Cantillon, "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," *American Economic Review*, 2012, *102(5)*, 2237–71.

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

Friedman, Milton, *A Program for Monetary Stability*, Vol. 541 of *The Miller Lectures*, Fordham University Press, 1960.

_ , "How to Sell Government Securities," *Wall Street Journal*, 1991, p. A8.

Gale, David and Lloyd Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69* (1), 9–15.

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

Jackson, Matthew O. and Alejandro M. Manelli, "Approximately Competitive Equilibria in Large Finite Economies," *Journal of Economic Theory*, 1997, *77* (2), 354–376.

Kalai, Ehud, "Large Robust Games," *Econometrica*, 2004, *72* (6), 1631–1665.

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

_ **and Parag A. Pathak**, "Incentives and Stability in Large Two-Sided Matching Markets," *The American Economic Review*, 2009, *99* (3), 608–627.

Krishna, Aradhna and Utku Ünver, "Improving the Efficiency of Course Bidding at Business Schools: Field and Laboratory Studies," *Marketing Science*, 2008, *27* (2), 262–282.

Roberts, Donald J. and Andrew Postlewaite, "The Incentives for Price-Taking Behavior in Large Exchange Economies," *Econometrica*, 1976, pp. 115–127.

Roth, Alvin E., "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.," *American Economic Review*, 1991, *81*, 415–40.

Rustichini, Aldo, Mark A. Satterthwaite, and Steven R. Williams, "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica*, 1994, pp. 1041–1063.

Sönmez, Tayfun and Utku Ünver, "Course Bidding at Business Schools," *International Economic Review*, 2010.

Swinkels, Jeroen M., "Efficiency of Large Private Value Auctions," *Econometrica*, 2001, *69* (1), 37–68.