Strategyproofness in the Large as a Desideratum for Market Design*

Eduardo M. Azevedo† and Eric Budish‡

January 2, 2012

Abstract

We distinguish between two ways a mechanism can fail to be strategyproof (SP). A mechanism may have profitable manipulations that *persist with market size*; and, a mechanism may have profitable manipulations that *vanish with market size*. We say that a mechanism is *strategyproof in the large (SP-L)* if all of its profitable manipulations vanish with market size.

Our main result is as follows. Suppose we are given some mechanism that has Bayes-Nash equilibria but is not SP-L. Suppose that the given mechanism is (semi-)anonymous, that agents have private values, and that a mild continuity condition is satisfied. Then, we show by construction that there exists another mechanism that is SP-L, and that implements approximately the same outcomes as the original mechanism, with the approximation error vanishing in the large-market limit. Thus, while SP often severely limits what kinds of mechanisms are possible, SP-L is often approximately costless, and hence may be a useful second-best. We illustrate with examples from assignment, matching and auctions.

*For helpful discussions we are grateful to Nabil Al-Najjar, Susan Athey, Aaron Bodoh-Creed, Gabriel Carroll, Jeff Ely, Alex Frankel, Drew Fudenberg, Jason Hartline, John Hatfield, Richard Holden, Ehud Kalai, Emir Kamenica, Fuhito Kojima, Scott Kominers, Jacob Leshno, Jon Levin, Paul Milgrom, Roger Myerson, David Parkes, Parag Pathak, Nicola Persico, Canice Prendergast, Ilya Segal, Eran Shmaya, Lars Stole, Glen Weyl, and especially Al Roth. We are grateful to seminar audiences at Ohio State, the MFI Conference on Matching and Price Theory, UCLA, Chicago, AMMA 2011, Boston College, the NBER Conference on Market Design, Duke / UNC, Michigan, Carnegie Mellon / Pittsburgh, Montreal, and Berkeley.

†Harvard University, azevedo@fas.harvard.edu.

‡University of Chicago Booth School of Business, eric.budish@chicagobooth.edu.

1 Introduction

Strategyproofness – i.e., that playing the game truthfully is a dominant strategy – is perhaps the predominant notion of incentive compatibility in practical market design. There are at least five important reasons why strategyproofness is so heavily emphasized relative to other forms of incentive compatibility, such as Bayes-Nash. First, strategyproof mechanisms are detail free for the designer, in the sense of Wilson (1987) (cf. Bergemann and Morris (2005)). Second, and relatedly, strategyproof mechanisms are strategically simple for participants; participants need not form beliefs about others' preferences or behavior in order to play the game optimally (Fudenberg and Tirole (1991); Roth (2008)). Third, with this simplicity comes a measure of fairness; agents' outcomes do not depend on their ability to "game the system" (Friedman (1991); Pathak and Sönmez (2008); Abdulkadiroğlu et al. (2006)). Fourth, strategyproof mechanisms may generate information about participants' true preferences that may be useful to policy makers (Roth (2008)). Fifth, Bayesian approaches simply have not yet proved tractable for a number of important market design problems.

However, in a wide variety of economic contexts, impossibility theorems indicate that strategyproofness severely limits what kinds of mechanisms are possible. These range from Gibbard (1973) and Satterthwaite's (1975) dictatorship theorem for general social choice problems, to Hurwicz's (1972) impossibility theorem for general equilibrium settings, to the Green and Laffont (1977) VCG theorem for allocation settings with quasi-linear preferences, to Roth's (1982) impossibility theorem for strategyproof stable matching, to Papai's (2001) dictatorship theorem for multi-unit demand assignment problems, to Abdulkadiroğlu et al.'s (2009) impossibility theorem for strategyproof and efficient school assignment.

This creates a conundrum for market designers. Strategyproofness is the only form of incentive compatibility that the literature finds fully satisfying, yet for many problems there are no good strategyproof mechanisms.

This paper proposes a criterion of approximate strategyproofness, and suggests that it may be a useful second-best alternative in environments where strategyproof mechanisms are unattractive. Our criterion is based on a conceptual distinction between two ways a mechanism might fail to be strategyproof. First, a mechanism might have profitable manipulations that *persist with market size*. Second, a mechanism might have profitable manipulations that *vanish with market size*. While both kinds of manipulability are undesirable, we suggest that manipulations that persist with market size are especially problematic, and are avoidable. If a mechanism only has manipulations that vanish with market size, we will say that it is *strategyproof in the large (SP-L)*.

Whether or not a mechanism is SP-L is simple to check, and generates an intuitively appealing classification of non-strategyproof mechanisms. Many well-known non-strategyproof mechanisms that are thought to work well in practice are SP-L. Examples include the Walrasian mechanism, double auctions, uniform-price auctions, and deferred-acceptance algorithms. Many other non-strategyproof mechanisms that have been shown to have important incentives problems in practice are *not* SP-L, i.e., they are manipulable even in large markets. Examples include the pay-as-bid treasury auction criticized by Friedman (1964, 1991), the Boston mechanism for school choice criticized by Abdulkadiroğlu and Sönmez (2003), the bidding points auction for course allocation criticized by Sönmez and Ünver (2010), and the priority-match algorithm for two-sided matching criticized by Roth (2002). Furthermore, both Friedman's critique of pay-as-bid auctions and Roth's critique of priority-match algorithms explicitly suggested alternative mechanisms that are not strategyproof but that are SP-L: uniform-price auctions and deferred-acceptance algorithms, respectively. This too speaks to the intuitive appeal of the criterion.

Our main result is as follows. Suppose we are given some mechanism that has Bayes-Nash equilibria. Suppose that the mechanism is (semi-)anonymous, which is a common feature of practical market-design settings; that agents have private values, in the sense that they know their own preferences over outcomes without observing other agents' private information; and that the mechanism satisfies a condition called quasi-continuity, which we will describe in more detail below. We show that there necessarily exists another mechanism that is strategyproof in the large, and that implements approximately the same outcomes as the original mechanism, with the approximation error vanishing in the large-market limit. An interpretation of our result is that while restricting attention to SP mechanisms can be very costly in terms of design objectives, restricting attention to SP-L mechanisms is approximately costless. This justifies consideration of SP-L as a second-best alternative to SP.

Our proof is by construction of a specific SP-L mechanism, from a given mechanism that has Bayes-Nash equilibria. The construction works as follows. Agents report their types to our mechanism. Our mechanism then calculates the empirical distribution of these types, and then "activates" the Bayes-Nash equilibrium strategy of the original mechanism associated with this empirical distribution. If agents all report their preferences truthfully, this construction will yield the same outcome as the original mechanism in the large-market limit, because the empirical distribution of reported types converges to the underlying true distribution. The subtle part of our construction is what happens if some agents systematically misreport their preferences, e.g., they make mistakes. Suppose the true prior is μ, but for some reason the agents other than agent i systematically misreport their preferences, according to distribution m. In a finite market, with sampling error, the empirical distribution of the other agents' reports is say $\hat{m}$. As the market grows large, $\hat{m}$ is converging to m, and also i's influence on the empirical distribution is vanishing. Thus in the limit, our construction will activate the Bayes-Nash equilibrium strategy associated with m. This is the "wrong" prior – but agent i does not care. From his perspective, the other agents are reporting according to m, and then playing the Bayes-Nash equilibrium strategy associated with m, so i too wishes to play the Bayes-Nash equilibrium strategy associated with m. This is exactly what our constructed mechanism does on i's behalf in the limit. Hence, no matter how the other agents play, i wishes to report his own type truthfully in the limit, i.e., the constructed mechanism is SP-L.

Our construction resembles a revelation principle construction, in that it takes a mechanism in which agents play the game directly and transforms it into a mechanism in which agents just report their type, and then let the center play optimally on their behalf. However, we emphasize that our mechanism is fundamentally distinct. In a traditional revelation mechanism, the mechanism designer knows the true prior (e.g., μ), and then plays the Bayes-Nash equilibrium strategy associated with this true prior on agents' behalf. It is then a Bayes-Nash equilibrium for agents to report their types truthfully. Our mechanism has two advantages relative to this benchmark. First, our mechanism is prior free: neither the agents nor the mechanism need know the underlying distribution of preferences a priori, because the mechanism infers the prior from the empirical. Second, our mechanism provides dominant-strategy incentives in the limit, whereas a traditional revelation mechanism provides just Bayes-Nash incentives even in the limit.

A difficult technical issue throughout the analysis concerns points of discontinuity in a mechanism. The argument sketched above relied implicitly on an assumption of continuity local to m: as the empirical distribution $\hat{m}$ is converging to m, agent i's utility is converging to what he would receive in the Bayes-Nash equilibrium associated with m. However, many familiar mechanisms have points at which agents' outcomes are not locally continuous. As an example, consider the uniform-price auction. Typically, a small change in the distribution of opponents' bids will have only a small effect on agent i's payoff. However, if i is the marginal bidder, a small change could discontinuously cause i to change from being a winner of the auction to being a loser of the auction. Or, if prices are discrete and demand is exactly equal to supply at some price p, then a small decrease in demand could cause the market clearing price to decrease discontinuously.

Our analysis accommodates discontinuities in two related ways. First, our main result does not require that a mechanism be everywhere continuous, but rather that it satisfy a condition we call *quasi-continuity*. The quasi-continuity condition allows for the kinds of discontinuities that arise in the uniform-price auction. Roughly, the requirement is that discontinuities be "knife edge", in the sense that on either side of a discontinuity is a region of local continuity. Second is the way we define SP-L itself. A mechanism is strategyproof if, for any profile of the other agents' reports, agent i maximizes his utility by reporting his own preferences truthfully. We say that a mechanism is SP-L if, in the large-market limit, for any *probability distribution* of the other agents' reports, agent i maximizes *expected* utility by reporting his preferences truthfully. When a mechanism is continuous, by a law of large numbers argument, there is no distinction in the limit between expected utility from a probability distribution of reports and realized utility from a specific profile of reports. If a mechanism has discontinuities, however, there can be such a distinction. For instance, in the uniform-price auction, an agent who reports her preferences truthfully might wish ex post to revise her report, in the event that the empirical realization of reports is exactly the knife-edge case where she can have a discontinuous influence on price. We classify the uniform-price auction as SP-L because the likelihood of this event vanishes with market size, for *any* probability distribution over the other agents' reports (cf. Example 1 below).

If we assume a stronger form of continuity, we can get stronger results. Specifically, if we assume that mechanisms are uniformly continuous as defined by Kalai (2004), then we can show that any SP-L mechanism has the property that, in a large enough market, no agent ever gains more than ϵ in any realization by misreporting her preferences. This is a stronger form of ex post robustness than that obtained by Kalai (2004) for Bayes-Nash equilibria, both because the likelihood of having an ϵ deviation is exactly zero in a large enough finite market rather than converging to zero in the limit, and because agents need not know the prior, coordinate on a specific equilibrium, etc.

Related Literature Our paper is related to a large literature that has studied how market size can ease incentive constraints. An early paper in this tradition is Roberts and Postlewaite (1976) on the Walrasian mechanism, which can be seen as a response to Hurwicz's (1972) critique that the Walrasian mechanism is not strategyproof. Other papers in this tradition include Jackson and Manelli (1997), Kovalenkov (2002), and Al-Najjar and Smorodinsky (2007) on the Walrasian mechanism, Rustichini et al. (1994) on double auctions with private values, Pesendorfer and Swinkels (2000), Cripps and Swinkels (2006), and Reny and Perry (2006) on double auctions with common-value components, Immorlica and Mahdian (2005), Kojima and Pathak (2009), and Lee (2011) on deferred acceptance algorithms, and Kojima and Manea (2010) on the Bogomolnaia and Moulin (2001) probabilistic serial mechanism. Each of these papers provides a defense of a *specific mechanism* based on its incentive properties in large markets. Our paper aims to justify strategyproofness in the large as a *general desideratum* for practical market design. Note that in the context of any of the specific mechanisms named above, our analysis is much less instructive than are previous analyses tailored to the specific mechanism.

Technically, our paper is most closely related to Kalai (2004). Kalai's Theorem 1 shows that Bayes-Nash equilibria are approximately ex post Nash in a class of large continuous and anonymous games.[1] In words, if a large number of agents with private information about their types play some BNE, then ex post – i.e., after seeing each agent's chosen action – agents will have vanishingly little incentive to revise their play. The difference between our Theorem 1 and Kalai's Theorem 1 is that Kalai shows that a given BNE is approximately ex post Nash, whereas we use the BNE of a given mechanism to create a new mechanism that is approximately strategyproof. In our new mechanism players need not have common knowledge of the prior, or of what equilibrium is being played, nor need they be strategically sophisticated in any way.

There are several other well-known technical ideas that our paper is related to. First is the revelation principle (Myerson (1979)); see our discussion of how our main result is related to but distinct from the revelation principle in Section 4.2. Second is the idea that there can be equivalence, in specialized environments, between what is implementable in Bayes-Nash equilibrium and what is implementable in dominant strategies. The revenue equivalence theorem in auction theory is an early example of such a result, since there exist dominant-strategy auctions that maximize revenue. See Manelli and Vincent (2010) and Goeree and Kushnir (2011) for recent equivalence results in auction and social-choice settings, respectively, and see also Gershkov et al. (2011) for a provocative discussion of these issues. Third is the idea of using the empirical distribution of agents' actions to infer the underlying distribution of preferences; see Goldberg et al (2001) and Segal (2003) for applications of this idea in the context of monopoly pricing.

[1]Recent work by Azrieli and Shmaya (2011) shows that continuity is the crucial assumption in Kalai (2004), and that anonymity can be relaxed. See also Deb and Kalai (2011) and Carmona and Podczeck (2011) for recent extensions of aspects of Kalai (2004). Recent work by Bodoh-Creed (2010) shows that Kalai-like assumptions imply a close relationship between games with a continuum of players and games with a large finite number of players.

Next, our paper is related to the literature on the role of strategyproofness in practical market design. Wilson (1987) famously argued that practical market designs should aim to be detail free, and Bergemann and Morris (2005) formalized the sense in which strategyproof mechanisms are robust in the sense of Wilson. Several recent papers have argued that strategyproofness can be viewed as a design objective and not just as a constraint: papers on this theme include Abdulkadiroğlu et al. (2006), Abdulkadiroğlu et al. (2009), Pathak and Sönmez (2008), and Roth (2008). Our paper contributes to this literature by showing that approximate strategyproofness is approximately costless in large markets, relative to other kinds of incentive compatibility. Also, the distinction we draw between manipulations that persist and manipulations that vanish highlights that many mechanisms in practice are manipulable in a preventable way.

Last, our paper is conceptually related to Parkes et al. (2001), Day and Milgrom (2008), Pathak and Sönmez (2011), and Carroll, each of which seeks to say something more useful about non-strategyproof mechanisms than simply that they are not strategyproof.[2] Parkes et al. (2001) and Day and Milgrom (2008) propose cardinal measures of a combinatorial auction's manipulability, and seek to design an auction that minimizes manipulability subject to other design objectives. Carroll, too, proposes a cardinal measure of manipulability, and explores his measure in the context of voting problems. He derives comparisons amongst voting rules, and asymptotic lower bounds on the manipulability of any rule satisfying other desiderata. Pathak and Sönmez (2011) propose a partial order by which to compare non-strategyproof mechanisms based on their vulnerability to manipulations. Mechanism a is said to be more manipulable than mechanism b if, for any problem instance where b is manipulable by at least one agent, so too is a. This criterion helps to explain several recent policy decisions in which school authorities switched from one manipulable mechanism to another. We view our approach as complementary to these alternative approaches. An advantage of our approach is that it yields an explicit design desideratum, namely that mechanisms be strategyproof in the large.

Organization of the paper The rest of this paper is organized as follows. Section 2 describes the environment and some key assumptions. Section 3 defines strategyproofness in the large and related concepts, and presents several examples. Section 4 presents the main theoretical result. Section 5 discusses various extensions. Section 6 concludes. Proofs are in the appendix.

[2] See also Milgrom (2011) Section IV for a general discussion of these issues.

2 Environment

2.1 Preliminaries

There is a **finite type space**, T, and a **finite outcome space**, X_0. The outcome space describes the outcome possibilities for an individual agent; e.g., in the context of school assignment, X_0 is the set of schools to which a student might be assigned. An agent's type determines her preferences over this outcome space. Specifically, for each $t_i \in T$ there is a von Neumann-Morgenstern expected utility function $u_{t_i} : X \to [0, 1]$, where $X = \Delta X_0$ denotes the set of lotteries over outcomes. Preferences are **private values** in the sense that an agent's utility from her outcome depends only on her own type.

Our interest is in mechanisms that are well defined for various market sizes and various distributions of types, **holding fixed** T and X_0. The set of possible market sizes is simply $\mathbb{N}$, with $n \in \mathbb{N}$ denoting the number of agents in a particular economy. For each $n \in \mathbb{N}$, there is a set $Y_n \subseteq (X_0)^n$ that denotes which allocations are **feasible.** The sequence $(Y_n)_{\mathbb{N}}$ encodes how the constraints relevant to the problem at hand vary with market size. For instance, in Examples 1 and 2 in Section 3.1, the relevant constraints are capacity constraints, and we assume that capacity grows linearly with market size.

We assume for most of the analysis that agents' types are independently and identically distributed (**iid**).[3] Hence the set of possible preference distributions is ΔT, with $\mu \in \Delta T$ denoting the preference distribution (or "prior") in a particular economy. We denote the set of priors with **full support** as $\bar{\Delta} T$.

Distributions $x \in X \equiv \Delta X_0$ and $\mu \in \Delta T$ may be represented as vectors of probabilities, with coordinates representing the probability assigned to each point in X_0 or T. We use this representation to define convex combinations over such distributions denoted $\alpha x + (1 - \alpha)x'$, $\alpha \in [0, 1]$, and the sup norm of the difference between two distributions denoted $|x - x'|$.[4]

2.2 Mechanisms

We define a mechanism as follows:

[3] See Section 5.4 for a discussion on relaxing the iid assumption.

[4] The same conventions apply to distributions of actions, $m \in \Delta A$, as defined in the next subsection.

Definition 1. *A* ***mechanism*** $\{(\Phi^n)_{\mathbb{N}}, A\}$ *consists of a* ***finite action space*** A *and a sequence of allocation functions*

$$\Phi^n : A^n \to \Delta((X_0)^n), \tag{2.1}$$

each of which satisfies feasibility: for any $n \in \mathbb{N}$ *and* $a \in A^n$*, the support of* $\Phi^n(a)$ *lies in the feasible set* Y_n*. The tuple* $\{\Phi^n, A\}$*, for a particular size* n*, is called an* **n-*mechanism***.

We make three important assumptions on the space of possible mechanisms. First, Definition 1 imposes that the action space A is finite, and the same for all market sizes. This assumption parallels our assumption above regarding the outcome space X_0; see Section 3.1 for illustrative examples. Second, Definition 1 imposes that mechanisms are **detail free** in the sense that the allocation functions (2.1) are not allowed to vary with the prior $\mu \in \Delta T$. Of course, how agents choose to play a mechanism may depend on their prior, e.g., in the Bayes-Nash equilibrium of the pay-as-bid auction (cf. Example 1 below).[5]

Third, for the main analysis we assume that mechanisms are anonymous, defined as follows:

Definition 2. *Mechanism* $\{(\Phi^n)_{\mathbb{N}}, A\}$ *is* ***anonymous*** *if, for all* $n \in \mathbb{N}$*, the allocation function* $\Phi^n(\cdot)$ *is invariant to permutations. That is, for any* n*, any* $a \in A^n$*, and any permutation function* $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$*, we have* $\Phi^n(a) = \pi^{-1}(\Phi^n(\pi(a)))$*.*

In words, anonymity means that each agent's outcome is a common function of her own action and the distribution of all actions. Anonymity rules out that an agent's outcome depends on the precise details of who specifically plays what, and it also rules out that two agents who play the same action get different random bundles. Anonymity is a natural feature of many large-market settings, with examples of anonymous mechanisms including the Walrasian mechanism, most well-known single-object and combinatorial-auction formats, and most of the mechanisms that have been proposed for single- and multi-unit assignment problems. In Section 5.1 we show that all of our results obtain if we relax anonymity to semi-anonymity (Kalai (2004)); semi-anonymity accommodates many additional settings in which there are asymmetries amongst classes of participants, e.g., double

[5]By the standard revelation principle (cf. Fudenberg and Tirole (1991); Section 7.2), for any mechanism with a Bayes-Nash equilibrium in which agents misreport their preferences, there exists a direct-revelation mechanism in which reporting one's type truthfully is a Bayes-Nash equilibrium. This direct-revelation mechanism, however, is no longer detail free for the designer; the map between reports and outcomes will have to vary with the prior. For instance, in the direct-revelation mechanism version of the pay-as-bid auction, the amount by which the center shades each type's bid must vary with the prior in order for truthful reporting to be a BNE.

auctions in which there are distinct buyers and sellers, and certain kinds of two-sided matching markets (cf. Azevedo and Leshno (2011), Section 5.4.2).

In what follows, it will often be useful to view mechanisms from the perspective of a generic agent i. From the perspective of agent i, the mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ can be viewed as a sequence of allocation functions:

$$\Phi_i^n : A \times A^{n-1} \to X, \tag{2.2}$$

where $\Phi_i^n(a_i, a_{-i})$ gives the lottery over outcomes agent i receives from $\Phi^n(\cdot)$ when he plays a_i and the other agents play a_{-i}. Under this perspective anonymity implies, first, that $\Phi_i^n(a_i, a_{-i})$ is invariant to permutations in a_{-i}, and second, that $\Phi_i^n(\cdot) = \Phi_j^n(\cdot)$ for any i and j.

2.3 Limit Mechanisms

We now define a key piece of notation, which is an ex-interim version of the individual allocation function (2.2). Consider a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$, a market size n, and a distribution over actions $m \in \Delta A$. Let:

$$\phi^n(a_i, m) = \sum_{a_{-i}} \Phi_i^n(a_i, a_{-i}) \cdot \Pr(a_{-i} | a_{-i} \sim iid(m)) \tag{2.3}$$

where $\Pr(a_{-i} | a_{-i} \sim iid(m))$ denotes the probability that the action vector a_{-i} is realized given $n-1$ iid draws from the action distribution m. The object $\phi^n(a_i, m)$ describes what a generic agent can expect to receive under mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ when he plays action a_i and the other $n-1$ agents play iid according to m. Since each $\Phi_i^n(a_i, a_{-i})$ is a random outcome in $X \equiv \Delta X_0$, and X is convex, the object $\phi^n(a_i, m)$ is also a random outcome in X. Note that we do not use a subscript i for the function $\phi^n(\cdot, \cdot)$, both to reduce notational clutter and to highlight that the function does not depend on the identity of the agent, due to anonymity.

We use the function $\phi^n(\cdot)$ to define limit mechanisms.

Definition 3. *The function $\phi^\infty : A \times \Delta A \to X$ is the* ***limit of mechanism*** *$\{(\Phi^n)_{\mathbb{N}}, A\}$ if, for all a_i, m:*

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

where ϕ^n is as defined in (2.3).

In words, $\phi^\infty(a_i, m)$ describes what a generic agent who plays a_i receives in the large market limit of mechanism $\{(\Phi^n)_\mathbb{N}, A\}$, when the other agents' play is iid according to m.

An important feature of our method of taking the limit is that each ϕ^n in the sequence converging to ϕ^∞ is random, in the sense that the play of the agent's $n-1$ opponents is stochastic (drawn from distribution m). This is in contrast with Debreu and Scarf's (1963) replicator economy, or with the approach pioneered by Aumann (1964) that looks directly at a continuum economy without explicitly modeling finite economies.

Our approach is more convenient in the present application than the replicator approach for two reasons. First, it allows each ϕ^n in the sequence, as well as the limit ϕ^∞, to be well defined for each possible preference distribution $m \in \Delta T$. By contrast, the replicator approach is only well defined in the limit for rational distributions, and in finite markets of size n the outcome is only well defined for distributions that consist of multiples of $\frac{1}{n}$. Second, if a mechanism has a knife-edge point of discontinuity, in our limit landing on the knife's edge becomes vanishingly likely, whereas under the replicator approach landing on the knife's edge could be a certain event.[6]

While most (if not all) practical market design mechanisms we are aware of have limits according to our definition, we note that it is very easy to construct examples of mechanisms that do not. For instance, if a mechanism acts differently depending on whether n is even or odd it will not have a limit. For the remainder of the analysis we limit attention to mechanisms that have limits.

2.4 Standard Equilibrium Concepts

We briefly state the standard concepts of Bayes-Nash equilibrium and strategyproofness.

A strategy is a map $\sigma : T \to \Delta A$ from types to distributions over actions. With slight abuse of notation, we also write $\sigma(\mu)$ for the distribution over actions induced by drawing types iid according to $\mu \in \Delta T$ and then playing according to $\sigma(\cdot)$.

Definition 4. *A (symmetric)* ***μ-Bayes-Nash equilibrium*** *(μ-BNE) of n-mechanism* $\{\Phi^n, A\}$ *is a strategy* $\sigma^*_\mu(\cdot)$ *such that for all* $t_i \in T$ *and* $a'_i \in A$

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

[6] For instance, if a fair coin is tossed 1,000,000 times, the probability that there are exactly 500,000 heads is very small, but a 500,000 fold replication of {heads, tails} will certainly result in exactly 500,000 heads. See Section 4.1 for a formal definition of what we mean by a "knife edge" point of discontinuity.

In words, the strategy σ^*_μ is a BNE if each agent's expected utility from playing according to σ^*_μ is higher than that from any other action, given that the other agents' types are distributed according to μ and that they also play according to σ^*_μ. Of course, there is no guarantee that $\sigma^*_\mu(t_i)$ is the best action for an agent of type t_i if the other agents play differently, which could occur, e.g., if the other agents make systematic mistakes, or play a different equilibrium, or if their types have a different distribution than μ. If a mechanism is strategyproof these informational requirements are no longer concerns:

Definition 5. *An n-mechanism $\{\Phi^n, A\}$ is **strategyproof** (SP) if $A = T$ and, for all $t_i, t'_i \in T$, and all $t_{-i} \in T^{n-1}$:*

$$u_{t_i}[\Phi^n_i(t_i, t_{-i})] \geq u_{t_i}[\Phi^n_i(t'_i, t_{-i})]$$

In words, a mechanism is strategyproof if the action space is such that agents simply report their types (i.e., $A = T$), and reporting truthfully is a dominant strategy.[7]

3 Strategyproofness in the Large

We are now ready to define strategyproofness in the large:

Definition 6. *Mechanism $\{(\Phi^n)_{\mathbb{N}}, T\}$ is **strategyproof in the large**, or **SP-L**, if, for any full support distribution of types $m \in \bar{\Delta}T$, and any t_i, t'_i:*

$$u_{t_i}[\phi^\infty(t_i, m)] \geq u_{t_i}[\phi^\infty(t'_i, m)]. \tag{3.1}$$

*Else, the mechanism is **manipulable in the large**.*[8]

Conceptually, Definition 6 draws a distinction between two ways a mechanism can fail to be strategyproof.

[7] The definition of strategyproofness can easily be extended to accommodate action spaces that, while not equal to the set of types, nevertheless capture the idea that agents simply report their preferences. For instance, it is often the case in matching applications that the appropriate type space is the set of cardinal preferences, whereas the relevant action space is the set of ordinal preferences (e.g., Abdulkadiroğlu et al. (2011)). Formally, say that mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ is a **preference-reporting mechanism** if the action space A partitions the type space T, and say that a preference-reporting mechanism is strategyproof if it is a dominant strategy to play the action associated with one's type. A direct mechanism in which $A = T$ is just a special case of a preference-reporting mechanism. Any preference-reporting mechanism can be represented as a direct mechanism, by interpreting the report t_i as the action associated with t_i.

[8] For mechanisms that do not have limits, SP-L can be defined by rewriting (3.1) as $\limsup_{n\to\infty} u_{t_i}[\phi^n(t'_i, m)] - u_{t_i}[\phi^n(t_i, m)] \leq 0$.

If a mechanism is SP-L, then it may have profitable manipulations in finite markets, but these manipulations all must vanish in the large market limit. Formally, we call the pair $\{t_i, t_i'\}$ a **profitable manipulation** of mechanism $\{(\Phi^n)_{\mathbb{N}}, T\}$ if there exists n, t_{-i}, such that type t_i profits by misreporting as t_i', i.e., $u_{t_i}[\Phi_i^n(t_i', t_{-i})] > u_{t_i}[\Phi_i^n(t_i, t_{-i})]$. We say that the manipulation $\{t_i, t_i'\}$ **vanishes with market size** if, for all $m \in \bar{\Delta}T$, the manipulation is no longer profitable in the limit, i.e., $u_{t_i}[\phi^\infty(t_i, m)] \geq u_{t_i}[\phi^\infty(t_i', m)]$.

If a mechanism is not SP-L, then it must have profitable manipulations that **persist with market size**. That is, there must be a profitable manipulation $\{t_i, t_i'\}$ where t_i profits from misreporting as t_i' not only in finite markets, but also in the large market limit; i.e., for some m, $u_{t_i}[\phi^\infty(t_i', m)] > u_{t_i}[\phi^\infty(t_i, m)]$.

While both kinds of manipulations are undesirable, we view manipulations that persist with market size as especially problematic. A way to formalize this view is through the concept of the limit budget set. Define mechanism $\{(\Phi^n)_{\mathbb{N}}, T\}$'s **limit budget set at** m as the set:

$$\{\phi^\infty(t_i', m)\}_{t_i' \in T}. \tag{3.2}$$

In words, (3.2) is the set of outcomes possible for a generic agent in the limit of mechanism $\{(\Phi^n)_{\mathbb{N}}, T\}$, given that the aggregate distribution of reports is m. For instance, in a Walrasian mechanism, the aggregate distribution of reports determines prices, and the limit budget set at m is simply the set of consumption bundles an agent can obtain at the prices corresponding to m, by varying her report over the full type set T.[9] If a mechanism is manipulable in the large, then reporting one's type truthfully simply does not select the agent's most preferred outcome from their budget set.

We emphasize how SP-L treats manipulations that arise at points of discontinuity in a mechanism. If the discontinuity is "knife edge",[10] then, since the probability of landing on a knife-edge point of discontinuity vanishes to zero with market size for any full-support distribution of opponent reports, a manipulation that is profitable only at the point of discontinuity will be said to vanish

[9]The concept of limit budget set also has a straightforward interpretation in some matching and assignment models. For instance, in Azevedo and Leshno (2011)'s matching model the limit budget set is governed by a set of statistics called "cutoffs" which describe the level of desirability necessary to achieve each possible match partner, and in Bogomolnaia and Moulin (2001)'s assignment model the limit budget set is governed by a set of statistics called "run-out times" which describe at what time in the algorithm each object exhausts its capacity (cf. Che and Kojima (2010)).

[10]See Section 4.1 for a formal definition of what we mean by a "knife edge" point of discontinuity.

with market size.[11] The uniform-price auction is an example of a mechanism that has such knife-edge manipulations – an agent with multi-unit demand may wish to reduce her quantity demanded if she is the marginal bidder who sets price – and that we classify as SP-L. Intuitively, we are ruling out the case where an agent knows *for sure* that she is pivotal. In the next section we discuss this and several other examples.

3.1 Examples

Our first example considers uniform-price and pay-as-bid auctions, two mechanisms best known for their use in the allocation of government securities. Neither mechanism is strategyproof. We show that the uniform-price auction is SP-L, while the pay-as-bid auction is not.

Example 1 (*Multi-unit auctions*). [12] There are kn units of a homogeneous good, with $k \in \mathbb{Z}_+$. To simplify notation, we assume that agents' preferences take the form of linear utility functions, 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, with $V = \{1, \ldots, \bar{v}\}$ the set of possible values, $Q = \{1, \ldots, \bar{q}\}$ the set of possible capacity limits, and $T = V \times Q$. We let the set of outcomes be $X_0 = V \times Q$ as well, by modeling an outcome as consisting of a per-unit payment and an allotted quantity of the object.

For both uniform-price and pay-as-bid auctions, agents simply report their types (i.e., $A = T$), and a single cutoff price p^* is calculated as a function of the reports $t = ((v_1, q_1), \ldots, (v_n, q_n))$ as follows:[13]

$$p^*(t) = \max_{p \in V} \sum_{i=1}^{n} q_i \cdot \mathbf{1}\{v_i \geq p\} \geq kn$$

i.e., p^* is the highest price at which demand weakly exceeds supply. Allocations of the good are equivalent across the two mechanisms: an agent who reports (v_i, q_i) is allocated q_i units if $v_i > p^*$, is allocated 0 units if $v_i < p^*$, and is rationed if $v_i = p^*$. Payments differ across the two mechanisms. In the uniform-price auction, every agent who is allocated units pays the same per-unit price, p^*. In the pay-as-bid auction, every agent who is allocated units pays a per-unit price equal to her

[11] Without the full support assumption it is possible to construct examples, e.g., with degenerate priors, in which landing on the knife edge is a probability one event even in the limit. See Appendix B for a discussion of this issue in the context of the uniform price and pay-as-bid auctions. In that context, we need an assumption that is weaker than full support but stronger than just ruling out degenerate priors.

[12] Appendix B provides additional details on the uniform-price auction and the pay-as-bid auction. In particular, the appendix shows that the pay-as-bid auction satisfies the quasi-continuity condition defined below in Section 4.1.

[13] The notation $\mathbf{1}\{statement\}$ denotes the indicator function which returns 1 if the statement is true and 0 if the statement is false.

own reported value. It is easy to see that the pay-as-bid auction is not strategyproof. More subtly, neither is the uniform-price auction, because in the finite economy an agent may be able to lower the price $p^*(t)$ by reporting to demand fewer units than she actually does (Ausubel and Cramton (2002)).

Now let us consider the limit economy. In the limit, if the measure of agents' reports is $m \in \bar{\Delta}T$, then for almost all m the cutoff price can be calculated as

$$p^*(m) = \max_{p \in V} \sum_{(v_i,q_i)} m(v_i,q_i) \cdot q_i \cdot \mathbf{1}\{v_i \geq p\} \geq k \tag{3.3}$$

The exception is if the cutoff price p^* that solves (3.3) does so with equality – i.e., there exists a price p^* such that $\sum_{(v_i,q_i)} m(v_i,q_i) \cdot q_i \cdot \mathbf{1}\{v_i \geq p^*\} = k$. In this event, the price in the limit will be p^* with probability one-half and will be $p^* - 1$ with probability one-half. This is due to the stochastic way that we take the limit: as n grows large, the probability that n iid draws from m result in demand strictly greater than supply at p^* is converging to one-half, just as is the probability that n iid draws result in demand strictly less than supply at p^*.

In the limit mechanism, each agent regards price as exogenous to her own report, because they cannot affect m. Note that our method of taking the limit ignores the vanishingly likely possibility that an individual agent can affect the price; i.e., it ignores the "knife edge" case discussed above.[14] It is thus easy to see that the uniform-price auction is SP-L whereas the pay-as-bid auction is not. In particular, in the pay-as-bid auction an agent of type (v_i, q_i) with $v_i > p^* + 1$ can profitably misreport as $(\hat{v}_i = p^* + 1, \hat{q}_i = q_i)$ to get the same quantity at a strictly lower price than if he reports truthfully. □

This example is consistent with Milton Friedman's (1991) observation that "you do not have to be a specialist" to participate in the uniform-price treasury auction, because you can just indicate "the maximum amount you are willing to pay for different quantitites ... if you bid a higher price

[14]To illustrate the knife edge deviation, consider the case where there are n units of the good (i.e., $k = 1$), $Q = \{1, 2\}$, and $V = \{L, H\}$, with $H > L$. If exactly 50% of the population consists of H-value agents with demand for two units, and the remainder of agents have value L, then there is a profitable misreport: by reporting to demand one unit instead of two an H-value agent can lower the price from H to L and hence increase his total profit. This manipulation is knife edge because if instead the proportion of H-value agents is $(50 - \epsilon)$%, then it is strictly more profitable to report truthfully and get two units, and if the proportion is $(50 + \epsilon)$% then the manipulation neither increases nor decreases profits. Notice as well that our definition of SP-L ignores knife edge cases caused by degenerate distributions; e.g., if $k = 2$ and 100% of the population consists of H-value agents with demand for two units, then all such agents have a profitable manipulation even in the limit.

[than the market clearing price], you do not lose as you do under the current [pay-as-bid] method." Friedman seems to be talking about the absence of what we call manipulations that persist, and seems to be less concerned by the vanishing manipulability of the uniform-price auction.[15]

Our next example is the Boston mechanism for school choice, a mechanism that does not explicitly have prices in the description. As mentioned in the introduction, this mechanism was criticized by Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu et al. (2006) for not being strategyproof. We show something stronger, which is that it is not even SP-L.

Example 2 (*The Boston mechanism*)**.** Let X_0 be the set of schools, and let each school $j = 1, \cdots, \#X_0 - 1$ have capacity qn, with $q \in (0,1)$. That is, q is the fraction of the overall student body that each school can accommodate.[16] School $j = \#X_0$ is assumed to have capacity for all students, and may represent either an underdemanded school or being unmatched.

Agents' types are von-Neumann Morgenstern utility functions over the set of schools. That is, functions of the form $u_{t_i} : X_0 \to \{0, 1, \ldots, \bar{u}\}$ for an integer $\bar{u}$. The set of actions A is the set of ordinal preferences over X_0, which is a partition of the type space. Therefore, this is an example of a preference-reporting mechanism as defined in footnote 7.

The Boston mechanism awards as many students as possible their reported first choice school; then, awards as many students as possible their reported second choice school, etc. To keep the description concise we consider a simplification of the Boston mechanism, in which there is only one round. Let $d_j = \sum_{i=1}^{n} \mathbf{1}\{j \text{ is } a_i\text{'s first choice}\}$ denote the number of students who report that school $j \in X_0$ is their first choice. Each such student then receives school j with probability $\min(1, \frac{qn}{d_j^1})$, and is matched to school $j = \#X_0$ with the remaining probability. Let $p_j = \min(1, \frac{qn}{d_j^1})$.

The limit is calculated straightforwardly. If the overall measure of agents' reports is $m \in \bar{\Delta} A$, let m_j denote the measure of students who report that school $j \in X_0$ is their first choice, i.e., $m_j = \sum_{a_i \in A} m(a_i) \cdot (\mathbf{1}\{j \text{ is } a_i\text{'s first choice}\})$. The probability that a student who ranks j first gets

[15] Pathak and Sönmez (2011) provide a complementary perspective on the incentive comparison between the uniform-price and pay-as-bid auctions. Pathak and Sönmez (2011) show that any agent who can profitably manipulate the uniform-price auction in a given finite economy can also profitably manipulate the pay-as-bid auction in that same finite economy. Moreover, the latter manipulation is always larger in utility terms. Thus, Pathak and Sönmez (2011) suggests that the pay-as-bid auction is more manipulable than the uniform-price auction in any given finite economy, whereas our analysis highlights that the pay-as-bid auction's manipulability persists with market size, whereas the uniform-price auction is strategyproof in the large.

[16] There are several different ways one might imagine taking the large-market limit of a school choice problem. The key assumption that our analysis imposes is that X_0 is a finite set. The assumption that the capacity of each element of X_0 grows linearly with n is convenient for the example, but it is not important and easily relaxed. Also, for thinking about the set X_0 in economies of varying sizes, it may be helpful to conceptualize X_0 as the set of "possible types of schools", i.e., as a finite school characteristics space.

it can be calculated as

$$p_j^* = \min(1, \frac{q}{m_j})$$

Notice that in the limit mechanism each agent regards the p_j^*'s as exogenous to their own report. Agent t_i will wish to misreport her first choice school if her first choice is j, but there exists j' where ranking j' first gives her strictly greater expected utility, i.e., $u_{t_i}(j')p_{j'}^* > u_{t_i}(j)p_j^*$. Therefore the mechanism is manipulable in the large. □

There are numerous other examples. For single-unit assignment problems such as in Example 2, Hylland and Zeckhauser's (1979) pseudomarket mechanism is an example of a price-based mechanism that is SP-L, while Bogomolnaia and Moulin's (2001) probabilistic serial mechanism is an example of a mechanism that does not explicitly use prices in the original description but that is SP-L (cf. Kojima and Manea (2010)). For multi-unit assignment problems, the mechanisms found in practice are manipulable in the large, specifically the Bidding Points Auction studied by Sönmez and Ünver (2010), and the Draft Mechanism studied by Budish and Cantillon (Forthcoming). Mechanisms recently proposed in theory are SP-L, specifically the Approximate Competitive Equilibrium from Equal Incomes mechanism proposed by Budish (Forthcoming), the multi-unit generalization of Hylland and Zeckhauser's pseudomarket proposed by Budish et al. (2011), and the Proxy Draft proposed by Budish and Cantillon (Forthcoming). For position auctions, the Generalized Second Price auction studied by Edelman et al. (2007) is manipulable in the large (cf. Example 4 below), whereas any Walrasian procedure that produces a competitive equilibrium price vector as a function of reported values, and then allocates agents their demands (based on their reports) at that price vector, will be SP-L.

The concepts can also be applied to two-sided matching mechanisms, if we generalize the class of mechanisms considered to be the class of semi-anonymous mechanisms (Kalai (2004)), and not just anonymous mechanisms; cf. Section 5.1. Then, techniques in Kojima and Pathak (2009) can be used to show that Gale and Shapley's deferred acceptance algorithm is SP-L in semi-anonymous environments. It is also easy to see that the priority-match algorithm, criticized by Roth (2002) and others, is manipulable in the large.

The following table summarizes this informal discussion.

Table 1. Which Non-SP Market Designs are SP-L?

Problem	**Manipulable in the Large**	**Strategyproof in the Large**
Single-unit Assignment	Boston Mechanism	Prob Serial, HZ Pseudomarket
Multi-unit Assignment	Bidding Points Auction	CEEI, Generalized HZ
	HBS Draft Mechanism	Proxy Draft
Multi-unit Auctions	Pay-as-Bid Auctions	Uniform-Price Auctions
	GSP Auction	Walrasian Mechanism
Two-Sided Matching	Priority-Match Algorithm	Deferred-Acceptance Algorithm

We emphasize that, while many of the most familiar examples of SP-L mechanisms use prices – e.g., the Walrasian mechanism, uniform-price auctions, double auctions – there are many examples in Table 1 of SP-L mechanisms that do not use prices, and there are also examples of mechanisms that do use prices but that nevertheless are not SP-L.

4 Main Result

Strategyproofness often severely limits what kinds of mechanisms are possible. Our main result identifies a sense in which SP-L does not. The result requires a quasi-continuity assumption, which we present and discuss in Section 4.1. We then present the main result in Section 4.2, along with a proof sketch. Section 4.3 discusses the relationship between the main result and some well-known related technical ideas. Since the proof is by construction, we provide an example construction in Section 4.4, using the Boston mechanism for school choice discussed in Example 2.

4.1 Quasi-Continuity of Equilibria

We first need some new notation. Given a market size n and a distribution $\bar{m} \in \Delta(A^{n-1})$ over action profiles, we may extend $\Phi_i^n(\cdot,\cdot)$ linearly as:

$$\Phi_i^n(a_i,\bar{m}) = \sum_{a_{-i}} \Phi_i^n(a_i,a_{-i}) \cdot \bar{m}(a_{-i}).$$

Now consider an $n-1$ vector of types t_{-i}, and a strategy $\sigma : T \to \Delta A$. Together, σ and t_{-i} induce a probability distribution over action profiles, i.e., an element of $\Delta(A^{n-1})$. We will denote this induced distribution as $\sigma(t_{-i})$. We highlight that $\sigma(t_{-i})$ denotes a distribution over A^{n-1}. We will then use the notation

$$\Phi_i^n(a_i, \sigma(t_{-i}))$$

to describe the outcome when i plays a_i, and the other players have types t_{-i} and play according to σ.[17] Moreover, given a vector of types t, we denote by $\text{emp}[t] \in \Delta T$ the empirical distribution of types.

Next, we need to define a limit Bayes-Nash equilibrium (Bodoh-Creed 2010).

Definition 7. *Given a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$, with limit $\phi^\infty(\cdot,\cdot)$, the strategy $\sigma_\mu^*(\cdot)$ is a limit μ-BNE if, for all $t_i \in T$ and $a_i' \in A$:*

$$u_{t_i}[\phi^\infty(\sigma_\mu^*(t_i), \sigma_\mu^*(\mu))] \geq u_{t_i}[\phi^\infty(a_i', \sigma_\mu^*(\mu))].$$

Limit equilibria are simply strategy profiles that become arbitrarily close to optimal as the economy grows large. Bodoh-Creed 2010 provides general conditions under which these correspond to the limit of Bayes Nash equilibria as the market grows.Last, we need the concept of a family of limit equilibria. Given a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ with limit $\phi^\infty(\cdot,\cdot)$, we say that $(\sigma_\mu^*)_{\mu\in\Delta T}$ is a family of limit Bayes-Nash equilibria if, for each $\mu \in \Delta T$, σ_μ^* is a limit μ-BNE.

Several results on large games or mechanisms make continuity assumptions, which require that the outcomes of the mechanism do not vary discontinuously with the distribution of agents' play (Kalai 2004; Bodoh-Creed 2010; Liu and Pycia 2011). However, in many important mechanisms, allocations can vary discontinuously with small changes in agents' play,[18] and agents' equilibrium play may likewise vary discontinuously with a small change in beliefs. We therefore introduce the following weakening of continuity, called quasi-continuity. Quasi-continuity allows for a mechanism to have

[17] We highlight that we use the notation

$$\phi^n(a_i, m),$$

where m is a distribution over A to denote the payoff to player i when her opponents' play is independently and identically distributed as m. This is a very different object than

$$\Phi_i^n(a_i, \sigma(t_{-i})),$$

which is i's payoff when her opponents have types given by the vector t_{-i} and play strategy σ.

[18] See for example Azevedo and Leshno (2011) Section 5.5 showing that the set of stable matchings of a large economy may change discontinuously with a small change in the distribution of types in the population, and therefore that the outcomes of any stable matching mechanism must change discontinuously.

discontinuities so long as they are "knife edge", in the sense that on either side of a discontinuity (the sets $\mathcal{B}$ in the definition) is a region of local continuity (the sets $\mathcal{A}_k$ in the definition).

Definition 8. *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** at a prior $\mu_0 \in \bar{\Delta}T$ if for every $\epsilon > 0$, there exists a neighborhood $\mathcal{N}$ of μ_0 such that:*

1. *$\mathcal{N}$ can be decomposed as the union of a finite number of open sets, $\mathcal{A}_1, \dots, \mathcal{A}_K$, and a closed set $\mathcal{B}$. Formally, $\mathcal{N} = \cup_{1\le k\le K}\mathcal{A}_k \cup \mathcal{B}$ with each $\mathcal{A}_k$ open.*
2. *If types are drawn iid according to μ_0, then the probability that the empirical distribution of types lands within distance $1/n$ of $\mathcal{B}$ goes to zero as n grows large. Formally,*

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

3. *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, for each $\mathcal{A}_k$, there exists n_0 such that for any $n > n_0$, and any $\mu, \mu', \text{emp}[t_i, t_{-i}], \text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k$, we have:*

$$|\Phi^n_i(\sigma^*_\mu(t_i), \sigma^*_\mu(t_{-i})) - \Phi^n_i(\sigma^*_{\mu'}(t_i), \sigma^*_{\mu'}(t'_{-i}))| < \epsilon.$$

*The family of equilibria is **continuous at μ_0** if, for the prior μ_0, Conditions 1 and 3 hold with $K = 1$ and $\mathcal{B} = \emptyset$. A family of limit equilibria is (quasi-)continuous if it is (quasi-)continuous for every prior $\mu_0 \in \bar{\Delta}T$.*

Note that continuity and quasi-continuity are defined for families of limit equilibria, and not for mechanisms. Continuity asks that around any prior μ_0, as per condition 3, the allocation that an agent of type t_i receives does not vary too much, as long as opponents have types with an empirical distribution close to μ_0, and play is according to a strategy σ_μ with μ close to μ_0.

Quasi-continuity, while a weaker requirement, has a more involved definition. Intuitively, quasi-continuity allows the family of equilibria to be discontinuous at a prior μ_0. However, it imposes some minimal regularity in how outcomes of the family of equilibria vary close to μ_0. Specifically, quasi-continuity requires that 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 vanisingly small probability. We view quasi-continuity as a mild requirement.

To further clarify this definition, we provide the following example. Consider a uniform price multi-unit auction, as defined in Example 1 from Section 3.1. Take $\mu_0 \in \bar{\Delta}T$, and a family of limit equilibria σ^*_μ such that agents report truthfully for all $\mu \in \bar{\Delta}T$. Assume that, in expectation, the price $1 < p^* < \bar{v}$ clears the market exactly at μ_0. That is,

$$\sum_{(v_i,q_i)} \mu_0(v_i, q_i) \cdot q_i \cdot \mathbf{1}\{v_i \geq p^*\} = k.$$

Take a neighborhood $\mathcal{N}$ of μ_0 small enough such that, for all μ in $\mathcal{N}$, there is excess demand at price $p^* - 1$ and excess supply at price $p^* + 1$. If agents report their types truthfully, and the vector of reported types satisfies $\text{emp}[t] \in \mathcal{N}$, then the market clearing price will either be p^* or $p^* - 1$. Within this set $\mathcal{N}$, price is not a continuous function of the empirical distribution of reports. However, $\mathcal{N}$ can be divided into a set of priors where expected demand at p^* is strictly higher than supply k,

$$\mathcal{A}_1 = \{\mu \in \mathcal{N} : \sum_{(v_i,q_i)} \mu_0(v_i, q_i) \cdot q_i \cdot \mathbf{1}\{v_i \geq p^*\} > k\},$$

a set of priors where demand at p^* is strictly lower than supply k,

$$\mathcal{A}_2 = \{\mu \in \mathcal{N} : \sum_{(v_i,q_i)} \mu_0(v_i, q_i) \cdot q_i \cdot \mathbf{1}\{v_i \geq p^*\} < k\},$$

and a residual set $\mathcal{B} = \mathcal{N} \backslash (\mathcal{A}_1 \cup \mathcal{A}_2)$. $\mathcal{B}$ is then the subset of priors $\mu \in \mathcal{N}$ where the market clears exactly in expectation. Dividing $\mathcal{N}$ in this way, the market clearing price is always p^* when agents have a vector of types with empirical distribution in $\mathcal{A}_1$, and $p^* - 1$ when agents have a vector of types with empirical distribution in $\mathcal{A}_2$. Moreover, as n grows large, the probability that a vector of types drawn randomly according to μ_0 has an empirical distribution closer than $1/n$ to $\mathcal{B}$ is vanishing to zero. This implies that the family of equilibria is quasi-continuous. Appendix B describes a family of equilibria of the pay-as-bid auction in detail and shows that it too satisfies quasi-continuity. Appendix D gives an example where quasi-continuity is not satisfied, and shows that it is important for the proof of our main results.

4.2 Main Result

Our main result is the following:

Theorem 1. *Suppose that there exists a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ with a quasi-continuous family of limit equilibria $(\sigma^*_\mu)_{\mu\in\Delta T}$. Then there exists a direct mechanism $\{(F^n)_{\mathbb{N}}, T\}$ with the following properties:*

1. *$\{(F^n)_{\mathbb{N}}, T\}$ is strategyproof in the large.*
2. *In the limit as $n \to \infty$, truthful play of $\{(F^n)_{\mathbb{N}}, T\}$ and Bayes-Nash equilibrium play of $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ give agents approximately the same outcomes. Specifically:*

   (a) *If $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ is continuous at the true prior μ_0, then in the limit as $n \to \infty$, truthful play of $\{(F^n)_{\mathbb{N}}, T\}$ and Bayes-Nash equilibrium play of $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ give agents the same outcomes. Formally, given $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$, there exists n_0 such that for all $n > n_0$ and all t_i:*

   $$|f^n(t_i, \mu_0) - \phi^n(\sigma^*_{\mu_0}(t_i), \sigma^*_{\mu_0}(\mu_0))| < \epsilon,$$

   where $f^n(\cdot)$ is constructed from $F^n(\cdot)$ according to equation (2.3).

   (b) *If $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ is not continuous at the true prior μ_0, then in the limit as $n \to \infty$, truthful play of $\{(F^n)_{\mathbb{N}}, T\}$ gives agents the same outcomes as a convex combination of equilibrium outcomes under $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$, for priors in an arbitrarily small neighborhood of μ_0. Formally, for every $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$, there exist priors μ_k with $|\mathrm{emp}[\mu_k] - \mathrm{emp}[\mu_0]| < \epsilon$, and n_0, such that for all $n > n_0$ there exist weights π^n_k summing to one such that, for all t_i:*

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

In words, Theorem 1 says that, given any mechanism with Bayes-Nash equilibria satisfying quasi-continuity, there necessarily exists another mechanism that is SP-L and gives agents approximately the same outcomes in the large market limit. If the original mechanism is continuous at the true prior, then for any $\epsilon > 0$ the difference in utilities can be made smaller than ϵ for large enough n. If the original mechanism is not continous at the true prior, then, for large enough n, the utilities under the SP-L mechanism will be within ϵ of a convex combination of the utilities under the original mechanism, for priors in an arbitrarily small neighborhood of the true prior.

Hence, in large markets, under the conditions specified herein, SP-L is approximately costless to satisfy relative to Bayes-Nash incentive compatibility. We view this result as justifying SP-L as a second-best alternative to SP, especially in environments in which SP mechanisms are known to be unattractive.

The proof of Theorem 1 is by construction. We provide a sketch here, with the full details in Appendix A.

Proof Sketch. We begin with the construction of $\{(F^n)_{\mathbb{N}}, T\}$ from $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$. Suppose in a market of size n agents report $t = (t_1, \ldots, t_n)$. We construct $F^n(\cdot)$ from $\Phi^n(\cdot)$ as follows:

$$F^n(t) = \Phi^n(\sigma^*_{\text{emp}[t]}(t)) \tag{4.1}$$

In words, F^n plays action $\sigma^*_{\text{emp}[t]}(t_i)$ for agent i who reports t_i, where emp$[t]$ is not the true distribution of agents' types μ_0 (which is not known to the mechanism) but rather the *empirical distribution of agents' reports*. Intuitively, the empirical distribution emp$[t]$ "activates" a specific Bayes-Nash equilibrium of the original mechanism $\Phi^n(\cdot)$. We will show that the direct mechanism $\{(F^n)_{\mathbb{N}}, T\}$ constructed in this manner is strategyproof in the large and gives agents the same utilities in the limit as the original mechanism.

First, suppose that agents report their preferences truthfully, according to the true prior μ_0. In a finite market of size n there will be sampling error, so the realized empirical will be, say, $\hat{\mu}$. Agent i who reports t_i receives $F^n_i(t_i, t_{-i}) = \Phi^n_i(\sigma^*_{\hat{\mu}}(t_i), \sigma^*_{\hat{\mu}}(t_{-i}))$. As the market grows large, the realized empirical $\hat{\mu}$ converges to the true distribution μ_0, by the law of large numbers. Hence, assuming for now that the original mechanism is continuous at μ_0, agent i's allocation is converging to $\Phi^n_i(\sigma^*_{\mu_0}(t_i), \sigma^*_{\mu_0}(t_{-i}))$, exactly what he receives under Bayes-Nash equilibrium of the original mechanism. Thus, if all agents report truthfully, our mechanism coincides with the original mechanism in the limit, as required.

Now, suppose that the agents other than i misreport their preferences, according to some distribution $m \in \bar{\Delta}T$. As before, in a finite market of size n, there will be sampling error, so the realized empirical will be, say, $\hat{m}$. Agent i will thus receive $F^n_i(t_i, t'_{-i}) = \Phi^n_i(\sigma^*_{\hat{m}}(t_i), \sigma^*_{\hat{m}}(t'_{-i}))$. As the market grows large, the realized empirical $\hat{m}$ will converge towards m, so, assuming continuity at m, agent i's allocation is converging to $\Phi_i^n(\sigma_m^*(t_i), \sigma_m^*(t'_{-i}))$. This is what agent i would receive under the original mechanism *in the Bayes-Nash equilibrium corresponding to prior m*. Even though the other agents are systematically misreporting their preferences, our agent i remains happy to tell the truth, because the other agents are acting *as if* their preferences are distributed according to m, and then playing a strategy that is converging to the Bayes-Nash equilibrium corresponding to m. Thus agent i also wants to play the Bayes-Nash equilibrium strategy corresponding to m – which is exactly what happens when he reports his preferences truthfully to $\{(F^n)_{\mathbb{N}}, T\}$.[19] Hence, in the limit, we get dominant-strategy incentives, i.e., our constructed mechanism is SP-L.

The last step of the proof sketch is to describe what happens in the event that the equilibrium of the original mechanism is not continuous at μ_0 – e.g., the uniform-price auction example described in Section 4.1. This requires a technical lemma (Lemma 1 in the appendix) which says that, for any arbitrary prior $m \in \bar{\Delta}T$, the allocation an agent receives under $\{(F^n)_{\mathbb{N}}, T\}$ can be approximated by a convex combination of the allocations he would receive in the limit Bayes-Nash equilibria of $\{(\Phi^n)_{\mathbb{N}}, A\}$, for priors close to m. The key to the proof of the lemma is that, in a large enough market, a single agent cannot appreciably change the probability that the aggregate profile lands in each region $\mathcal{A}_k$, as defined in Definition 8. This allows us to exploit the continuity *within* each region $\mathcal{A}_k$, and the vanishing likelihood that the aggregate profile lands near the discontinuity region $\mathcal{B}$. □

4.3 Discussion

4.3.1 Relation to the Revelation Principle

It is important to emphasize how our constructed mechanism $\{(F^n)_{\mathbb{N}}, T\}$ differs from a traditional Bayes-Nash direct revelation mechanism (Fudenberg and Tirole 1991 Section 7.2). In a traditional Bayes-Nash DRM, the mechanism needs to know the prior μ_0, i.e., it is not detail free. The mechanism then announces a BNE strategy $\sigma_{\mu_0}^*(\cdot)$, and plays $\sigma_{\mu_0}^*(t_i)$ on behalf of an agent who reports t_i. Our mechanism *infers* a prior from the empirical distribution of agents' play. If agents indeed play truthfully, this inference is exactly correct in the limit, and our detail-free mechanism coincides with the traditional Bayes-Nash DRM. But if the agents other than i misreport, so that the empirical $\hat{m}$ is very different from the prior μ_0, then our mechanism automatically adjusts each

[19] Observe that this step of the argument requires the private values assumption. It is important that i does not care per se about the other players' true types.

agent's play to be the Bayes-Nash equilibrium play in a world where the prior was in fact $\hat{m}$. As a result, an agent who reports his preferences truthfully remains happy to have done so even if the other agents misreport, which is not the case in a traditional Bayes-Nash DRM. To summarize, the two key differences between our mechanism and a traditional Bayes-Nash DRM are: (1) our mechanism is detail free; (2) our mechanism is strategyproof in the large.

4.3.2 Relation to VCG and Random-Sampling Mechanisms

Our proof technique is related to, but distinct from, the ideas behind Vickrey-Clarke-Groves (VCG) and random-sampling mechanisms. In VCG, each agent i faces prices that depend on the empirical distribution of the $n-1$ agents other than himself. In random-sampling mechanisms (Goldberg et al (2001), Segal (2003), Baliga and Vohra (2003)), each agent faces prices that depend on a random sample of the agents other than himself; e.g., a typical approach is to randomly divide the population of n agents into two distinct groups of $n/2$ agents, and have each group face prices that depend on the empirical distribution of the agents in the other group (cf. Hartline and Karlin (2007)).

The advantage of these two approaches is that, since the prices each agent faces depend only on agents other than himself, the mechanisms are able to provide exact dominant-strategy incentives. By contrast, in our approach, since the empirical distribution of all n agents' reports is used to activate a single Bayes-Nash equilibrium, each agent affects which strategy is activated (our analogue of affecting prices), and we are able to provide only approximate incentives.

The disadvantage of the VCG and random-sampling approaches is closely related to the advantage: since the prices each agent faces depend only on agents other than himself, different agents necessarily face different prices. In some settings, such as combinatorial auctions and monopoly pricing (Segal, Goldberg et al, Hartline survey), this is not problematic, but in general this approach can cause violations of feasibility.[20] Suppose we applied the VCG idea to our basic construction above. If in a market of size n the agents report $t=(t_1,\ldots,t_n)$, we might try to construct $\tilde{F}^n(\cdot)$ from $\Phi^n(\cdot)$ as:

$$\tilde{F}_i^n(t_i,t_{-i})=\Phi_i^n(\sigma^*_{\text{emp}[t_{-i}]}(t_i),\sigma^*_{\text{emp}[t_{-i}]}(t_{-i})), \tag{4.2}$$

[20] See, for instance, Parkes et al. (2001) on the feasibility issues that arise in the context of combinatorial exchange (two-sided combinatorial auctions), and Kovalenkov (2002) on the feasibility issues that arise in the context of a Walrasian exchange economy.

that is, for agent i's allocation, we activate the Bayes-Nash equilibrium of $\Phi^n(\cdot)$ that corresponds to the empirical distribution $\text{emp}[t_{-i}]$ of the agents other than i. However, there is no reason to expect that the allocation that results from n applications of (4.2), once for each agent i, will be feasible. We know that $\Phi^n(\cdot)$ itself produces a feasible outcome for any profile of n actions, but in the outcome constructed according to (4.2) there are different action profiles used for different agents, because a different strategy $\sigma^*_{\text{emp}[t_{-i}]}$ gets activated for each i. By contrast, our mechanism constructed according to (4.1) is always feasible, because it inputs just a single action profile into $\Phi^n(\cdot)$.

4.4 Example of the Construction

We describe our main result in the context of a specific example, the Boston mechanism for school choice (cf. Example 2 above). Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu et al. (2006) criticized the Boston mechanism on the grounds that it is not strategyproof, and proposed that the strategyproof Gale-Shapley deferred acceptance algorithm be used instead. Indeed, the Gale-Shapley algorithm was eventually adopted for use in practice (cf. Roth (2008)). However, a second generation of papers on the Boston mechanism has argued that it has a Bayes-Nash equilibrium that yields greater student welfare than does the dominant strategy equilibrium of the Gale-Shapley procedure (Abdulkadiroğlu et al. (2011); Miralles (2009); Featherstone and Niederle (2011)). Perhaps, these papers argue, the earlier papers were too quick to dismiss the Boston mechanism in favor of strategyproof deferred acceptance.

Of course, the Bayes-Nash equilibria these second-generation papers construct rely on students having common knowledge of the distribution of other students' preferences; on students being able to coordinate on a specific equilibrium; on students being able to make very precise strategic calculations to determine whether to risk asking for a popular school; etc. Our Theorem 1 says that all of this complexity and non-robustness is unnecessary in a large market. Specifically, there must exist yet another mechanism that implements the same outcomes as these desirable Bayes-Nash equilibria of the Boston mechanism, but that is SP-L. Moreover, our proof of Theorem 1 indicates how to construct such a mechanism.

Interestingly, in the simplified version of the Boston Mechanism that we discussed in Example 2 above, the SP-L mechanism that we construct according to (4.1) closely resembles the Hylland and Zeckhauser (1979) pseudo-market mechanism for single-unit assignment. Specifically, write agent

i's type as a vector of von-Neumann Morgenstern utilities, one for each school: $v_{i1}, \ldots, v_{i|X_0|}$. Any limit Bayes-Nash equilibria of the Boston mechanism are characterized by a set of probabilities p_j^*, one for each school j, such that when each student i asks for the school that maximizes his expected utility – i.e., maximizes his expectation of $p_j^* \cdot v_{ij}$ - these probabilities are in fact correct.[21] Thus our constructed mechanism works as follows. First, students report their types. Next, the mechanism calculates the Bayes-Nash equilibrium associated with the empirical distribution of the reported types, which imply the market-clearing probabilities p_j^*. Last, each student is given the appropriate lottery. This is just like Hylland and Zeckhauser (1979), except that in that paper instead of probabilities p_j^* there are "prices" which are the inverses of these probabilities, i.e., the price of school j is $q_j^* = \frac{1}{p_j^*}$.

We emphasize the two advantages of our constructed mechanism as compared to the Bayes-Nash equilibria of the original Boston mechanism. First, students need not estimate the distribution of other students' preferences, the associated market-clearing probabilities, etc. Second, our mechanism is robust to systematic mistakes by the other students, or miscoordination over which equilibrium to play, etc., because in the limit it provides dominant-strategy incentives.

5 Extensions

This section discusses several extensions of the main results. We show that extensions of Theorem 1 still hold if we relax anonymity to semi-anonymity (5.1), if the given mechanism has a family of complete information Nash equilibria instead of Bayes-Nash equilibria (5.2), and if we consider sequences of families of finite-economy equilibria instead of a family of limit equilibria (5.3). Section 5.4 discusses relaxing the iid assumption to accommodate aggregate uncertainty. Section 5.5 provides an additional result showing that SP-L mechanisms have a strong ex-post robustness property when they are equicontinuous as defined by Kalai (2004). Supporting details for this section are in Appendix C.

5.1 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

[21] See Miralles (2009), which also contains a very nice description of the connection between the Boston mechanism's Bayes-Nash equilibria and Hylland and Zeckhauser (1979).

some notational burden, to the case of semi-anonymous mechanisms, as defined by Kalai (2004).

Assume that agents belong to groups g in a finite set G. Each group has a different set of possible types and actions, so that

$$\begin{aligned} T &= T_{g_1} \cup T_{g_2} \cup \cdots \cup T_{g_G} \\ A &= A_{g_1} \cup A_{g_2} \cup \cdots \cup A_{g_G}. \end{aligned}$$

A semi-anonymous mechanism is defined as $\{(\Phi^n)_{n\in\mathbb{N}}, (A_g)_{g\in G}\}$. As before, 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 characteristics. Agents are then permitted to misreport their preferences over other match partners, but they cannot misrepresent their gender or their characteristics. The following Example provides additional details.

Example 3. (Two-Sided Matching) This example shows that semi-anonymous mechanisms can cover matching mechanisms in two-sided markets (Gale and Shapley 1962). There are men and women, who differ on a set of characteristics. Groups g index both sex and the characteristics, 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 characteristics, and hence 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 either their sex or their characteristics.

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 other 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 the direct mechanism where $A_g = T_g$ for each $g \in G$. Men and women report a vector of types t, and therefore characteristics. This implies a weak preference ordering of each man over each woman and vice versa. Given these preferences, at least one stable matching exists. If there are multiple stable matchings, $\Phi^n(t)$ chooses one uniformly at random.

The only difference between the semi-anonymous case and the original anonymous case is that the set of strategies σ is restricted so that agents of a given group g cannot play actions not in A_g. In Appendix C.1 we extend the definitions of limit BNE and SP-L to this setting, and state and prove an extension of Theorem 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 1. The proof follows from noting that the argument in the anonymous case implies that the approximation formulas in Theorem 1 hold, and then showing that this implies that the constructed semi-anonymous mechanism is SP-L. □

5.2 Complete Information Nash Equilibria

Our construction in Section 4 takes as input a mechanism that has a family of Bayes-Nash equilibria. The same idea can be applied to a mechanism that has a family of complete-information Nash equilibria. (Note, importantly, that assuming that a mechanism has complete-information Nash equilibria is very different from assuming that agents actually have complete information.)

A complete information Nash equilibrium, in an economy where the type profile is known to be t, is defined in the usual way:

Definition 9. *A (symmetric) t-Complete-Information Nash Equilibrium (t-CINE) of n-mechanism* $\{\Phi^n, A\}$ *is a strategy* $\sigma^n_{\text{emp}[t]}(\cdot)$ *such that, for all* $t_i \in t$, $a'_i \in A$*:*

$$u_{t_i}[\Phi^n_i(\sigma^n_{\text{emp}[t]}(t_i), \sigma^n_{\text{emp}[t]}(t_{-i}))] \geq u_{t_i}[\Phi^n_i(a'_i, \sigma^n_{\text{emp}[t]}(t_{-i}))]$$

The important thing to note about Definition 9 is that the strategy in economy t depends only on emp$[t]$, that is, whether a strategy is a CINE is invariant to permutations of t. This follows from our focus on symmetric equilibria of anonymous mechanisms. This observation allows us to define quasi-continuity in a manner analogous to Definition 8, which then yields a theorem statement analogous to Theorem 1. See Appendix C.2 for the precise statements.

The mechanism we construct to prove the CINE version of Theorem 1 is:

$$F^n(t) = \Phi^n(\sigma^n_{\text{emp}[t]}(t)). \tag{5.1}$$

In words, agents report their types to the mechanism, which then computes a symmetric complete information Nash equilibrium strategy in the economy induced by the reports. Note that in general it is *not* a Nash equilibrium for each player to report their preferences truthfully to this mechanism in finite markets. The reason is that, by changing one's report from say t_i to t_i', one changes the profile of reported types from say t to t', and this in turn changes the strategy that is activated from $\sigma^n_{\text{emp}[t]}(\cdot)$ to $\sigma^n_{\text{emp}[t']}(\cdot)$. Thus, i changing his *report* can have the effect of the mechanism changing j's *action*. As the market grows large, i's influence on emp$[t]$ grows small, so under quasi-continuity mechanism (5.1) is SP-L. See Appendix C.2 for details.

An interesting feature of this construction (5.1) is that if agents tell the truth in finite markets, then (5.1) produces outcomes that are *identical* to the outcomes under the complete information Nash equilibria of the original mechanism. By contrast, with Bayes-Nash equilibria our constructed mechanism only approximates the finite market outcomes.

We illustrate the construction (5.1) with the Generalized Second Price (GSP) auction for search-engine advertising slots, as modeled by Edelman et al. (2007) (EOS). EOS showed that the GSP has complete-information Nash equilibria that coincide with the Vickrey-Clarke-Groves mechanism on the equilibrium path. Interestingly, our constructed mechanism coincides with VCG both on and off the equilibrium path and hence actually provides dominant strategy incentives in finite markets. As emphasized above, however, this is not generally the case; typically we will need to consider a large-market limit for (5.1) to provide exact incentives.

Example 4 (Generalized Second Price Auction)**.** We consider a variant of the EOS model of the Generalized Second Price (GSP) auction for search-engine advertising slots. There are $|X_0|$ advertising slots, with click-through rates $\alpha_1 > \alpha_2 > \cdots > \alpha_{|X_0|}$. Each slot has capacity for $q = \lfloor \kappa n \rfloor$

advertisers, with $\kappa \in (0,1)$. That is, κ is the fraction of the population whose advertisements can appear in any particular slot. (In EOS n is exogenously fixed and each slot accommodates exactly one advertiser, i.e., $\kappa = \frac{1}{n}$). There are n bidders, with per-click values in the set $T = \{1, \ldots, \bar{v}\}$.

The GSP works as follows. Each bidder i submits a bid b_i. Next, the bids are ranked in descending order, with the q highest bidders awarded the first advertising slot, the next q highest bidders awarded the second slot, etc. Last, each slot j's price is set equal to the highest bid amongst winners of the $j+1^{th}$ slot (slot $|X_0|$'s price is set equal to the highest bid amongst losing bidders).

Consider a complete information environment in which there are n bidders whose values are $v_1 \geq v_2 \geq \cdots \geq v_n$. Let $p^{VCG,j}$ be the price of the j^{th} slot in the dominant strategy equilibrium of the Vickrey-Clarke-Groves auction, in which all bidders bid truthfully.[22] The EOS complete information Nash equilibrium, translated to our environment, is for bidders $1, \ldots, q$ to bid their values, bidders $q+1, \ldots, 2q$ to bid $\frac{p^{VCG,1}}{\alpha_1}$, bidders $2q+1, \ldots, 3q$ to bid $\frac{p^{VCG,2}}{\alpha_2}$, etc. Notice that in this equilibrium most bidders misreport their values, and that such manipulations persist with market size; this misreporting was pointed out by EOS in their Remark 3. Additionally, note that playing this equilibrium requires each bidder to know enough information about other bidders' values that they can calculate the VCG payments that are an input into their own bids.

Our mechanism constructed according to (5.1) works as follows. First, bidders report their values; call the reported profile $\hat{v} = (\hat{v}_1, \ldots, \hat{v}_n)$, sorted so that $\hat{v}_1 \geq \hat{v}_2 \geq \cdots \geq \hat{v}_n$. Next, execute the EOS complete information Nash equilibrium bids associated with $\hat{v}$; that is, bidders $1, \ldots, q$ bid their values, bidders $q+1, \ldots, 2q$ bid $\frac{p^{V\hat{C}G,1}}{\alpha_1}$, bidders $2q+1, \ldots, 3q$ bid $\frac{p^{V\hat{C}G,2}}{\alpha_2}$, etc., where $p^{V\hat{C}G,j}$ is the VCG price of slot j in an economy where bidders submit the bids $\hat{v}$. Since winners of slot j pay the j^{th} click-through rate α_j multiplied by the highest bid for slot $j+1$, $\frac{p^{V\hat{C}G,j}}{\alpha_j}$, their total payment is simply $p^{V\hat{C}G,j}$. That is, our constructed mechanism exactly coincides with VCG. $\square$

5.3 Finite-Economy Bayes-Nash Equilibria

Theorem 1 starts from a family of limit Bayes-Nash equilibria $(\sigma^*_\mu)_{\mu \in \Delta T}$, and constructs a direct SP-L mechanism that implements approximately the same outcome. This construction could also be obtained based on a sequence of families of finite-economy Bayes-Nash equilibria, $(\sigma^n_\mu)_{\mu \in \Delta T, n \in \mathbb{N} \cup \infty}$.

Appendix C.3 contains the appropriately restated definition of quasi-continuity and of Theorem 1.

[22] It is straightforward to verify that all $\lfloor \kappa n \rfloor$ copies of the j^{th} slot have the same price in the VCG auction, hence the notion of the price of the j^{th} slot is well defined.

The main new assumption we need is that, for all priors μ, the sequence of BNE's $(\sigma_\mu^n)_{\mathbb{N}}$ converges to a limit BNE σ_μ^∞. The specific n-mechanisms we use in the construction to prove the modified Theorem 1 are:

$$F^n(t) = \Phi^n(\sigma^n_{\text{emp}[t]}(t)). \tag{5.2}$$

This is as in Equation (4.1), but using the finite-economy equilibria instead of the limit equilibria. Appendix C.3 contains the formal statement of the modified Theorem 1 and discusses the differences between the proofs.

The reason why we focus on limit equilibria for the statement of the main result is that finite-economy equilibria are often analytically less tractable. This is discussed in detail by Bodoh-Creed (2010), who gives conditions under which limit equilibria correspond to the limits of finite-economy equilibria. In multi-unit auctions, for example, closed form solutions for equilibria are often unavailable, and even showing basic properties of equilibria is a difficult problem (Swinkels (2001); Engelbrecht-Wiggans et al. (2006)). A particular analyst might thus find it more convenient to work with limit equilibria. Moreover, an analyst might even find limit equilibria more compelling in their own right. One argument in favor of limit equilibria is that in a game where exact Nash equilibria are cognitively and computationally very complex, it may be more likely that players reason through an approximate model. On the other hand, depending on the application, an analyst might see exact Nash equilibria of finite economies as a more appropriate solution concept. For these reasons, we do not take a view on which solution concept is in general more appropriate and useful, and provide versions of our main result for either solution concept.

5.4 Aggregate Uncertainty

Throughout the analysis, we have assumed that agents' types are distributed iid. This is a restrictive assumption, as it rules out aggregate uncertainty about the distribution of types in the economy. In this section we make two brief points about incorporating aggregate uncertainty into the analysis.

First, if aggregate uncertainty about the distribution of types creates new equilibria of some given mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$,[23] it is not always the case that our constructed mechanism can approximately implement the same outcomes as those equilibria. To illustrate, suppose that there is a state

[23] This is the case in equilibria that depend on agents having low information (Roth and Rothblum 1999; Featherstone and Niederle 2011; Featherstone 2011), or on exploiting correlation between agents' information (Cremer and McLean 1988).

of the world ω that is L (low) with probability p and H (high) with probability $1-p$, and that agents' types are distributed iid conditional on the aggregate state. Call these conditional distributions μ_L and μ_H, and write the unconditional distribution as $[p : \mu_L; 1-p : \mu_H]$. Suppose agents report their preferences truthfully under our mechanism. In the L state our mechanism will observe an empirical close to μ_L, while in the H state it will observe an empirical close to μ_H. However, our mechanism will not be able to identify the true prior $[p : \mu_L; 1-p : \mu_H]$. If the equilibrium associated with the true prior $[p : \mu_L; 1-p : \mu_H]$ corresponds to a p weighted blend of the equilibria associated with the iid priors $[1 : \mu_L; 0 : \mu_H]$ and $[0 : \mu_L; 1 : \mu_H]$ then our mechanism will correspond with the original, but this is not guaranteed. For instance, in Cremer-McLean mechanisms the equilibria that are possible under aggregate uncertainty are not equivalent to convex combinations of the equilibria that are possible under iid distributions.

Second, we highlight that reporting truthfully in an SP-L mechanism is still close to optimal even if agents have non iid priors over opponents' types. Formally we show that, if agents are given strictly less information than an iid belief over opponent types, in the Blackwell sense, then it is approximately optimal to report truthfully in an SP-L mechanism.

Consider the gain from misreporting for an agent who knows that other agents' actions are distributed according to $m \in \bar{\Delta}T$. If $\{(\Phi^n)_{\mathbb{N}}, T\}$ is SP-L, then this gain must be vanishingly small. That is, given ϵ, t_i, t_i', there exists n_0 such that, for all $n \geq n_0$,

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

Now consider an agent who knows strictly less than agents with any iid beliefs. Following Blackwell, we define a garbling of iid beliefs as a measure $\nu \in \Delta(\Delta A)$. The agent assigns probability $\nu(m)$ that opponents' types are iid according to $m \in \bar{\Delta}T$.

For an agent with such beliefs, the gain from deviating is

$$\int u_{t_i}[\phi^n(t_i', m)] - u_{t_i}[\phi^n(t_i, m)] d\nu(m).$$

We now show it is approximately optimal for such an agent to report truthfully. Given $\epsilon > 0$, from the definition of SP-L, we know that for each $m \in \Delta T$ there must exist $n_0(m)$ such that

$$u_{t_i}[\phi^n(t_i', m)] - u_{t_i}[\phi^n(t_i, m)] < \epsilon/2$$

for all $n \geq n_0(m)$. Now take n_1 such that $n_1 \geq n_0(m)$ for all m in a set $M \subseteq \Delta T$ with measure $\nu(M)$ at least $1-\epsilon/2$. We then have

$$\begin{aligned}\int u_{t_i}[\phi^n(t_i', m)] - u_{t_i}[\phi^n(t_i, m)]d\nu(m) &= \\ \int_M u_{t_i}[\phi^n(t_i', m)] - u_{t_i}[\phi^n(t_i, m)]d\nu(m) &+ \\ \int_{M^c} u_{t_i}[\phi^n(t_i', m)] - u_{t_i}[\phi^n(t_i, m)]d\nu(m) &< \\ \epsilon/2 + \epsilon/2 &= \epsilon.\end{aligned}$$

This result may be formally stated as follows.

Proposition 1. *Consider an SP-L mechanism $\{(\Phi^n)_{n\in\mathbb{N}}, T\}$. For any garbling ν of iid beliefs over opponents' types, and any $\epsilon > 0$, there exists n_0 such that an agent with beliefs ν in a market of size $n \geq n_0$ cannot gain more than $\epsilon > 0$ by misreporting her type.*

5.5 Ex post Robustness of SP-L Mechanisms in Large Finite Markets

Kalai (2004) studies the ex post robustness of Bayes-Nash equilibria in large games. Under an equicontinuity assumption that we provide below, he shows that Bayes-Nash equilibria are ex post robust in the following sense: For any $\epsilon > 0$, the probability that any player will have an ex post deviation that yields a gain of more than ϵ converges to 0 exponentially in market size n.[24]

Here we show that, under Kalai's equicontinuity assumption, SP-L mechanisms satisfy a stronger robustness property: For any $\epsilon > 0$, in a large enough market, no player has an ex post deviation which increases her payoff by more than ϵ. This is a stronger ex post robustness property for two reasons. First, in a large enough market, the probability of an ϵ deviation is exactly zero rather than converging to zero. Second, the fact that the ex post gain from any deviation is small does not depend on a player knowing the distribution of opponent types, nor on agents playing in equilibrium.

Formally, we follow Kalai and define an equicontinuous mechanism as follows.

[24] More specifically, Kalai (2004) considers a sequence of (semi-)anonymous games, with an increasing number of players, that satisfies an equicontinuity condition. He defines an ϵ-ex post Nash equilibrium profile as a profile of types and actions such that no player may gain more than ϵ by changing her action. He defines a profile of (possibly mixed) strategies to be an (ϵ, ρ) ex post strategy profile if with probability at least $1-\rho$ the realized profile of types and strategies is an ϵ ex post Nash equilibrium. Kalai's Theorem 1 shows that, for any sequence of Bayes Nash equilibria $(\sigma^n)_{n\in\mathbb{N}}$ of the games, and any $\epsilon > 0$, there exist constants $\alpha > 0, \beta < 1$ such that σ^n is an $(\epsilon, \alpha\beta^n)$ ex post strategy profile. We refer the interested reader to Kalai (2004) for more details.

Definition 10. *A mechanism* $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ *is equicontinuous if, for all* $\epsilon > 0$ *there exists* $\delta > 0$ *such that for all* $n, n', a_{-i} \in A^{n-1}, a'_{-i} \in A^{n'-1}$ *with*

$$|\operatorname{emp}[a_{-i}] - \operatorname{emp}[a'_{-i}]| < \delta$$

we have that for all a_i

$$|\Phi_i^n(a_i, a_{-i}) - \Phi_i^{n'}(a_i, a'_{-i})| < \epsilon.$$

We then define ϵ-Strategyproofness as follows.

Definition 11. *A direct mechanism* $\{(\Phi^n)_{n\in\mathbb{N}}, T\}$ *is* (ϵ, n)*-strategyproof if for all* $t \in T^n, t'_i \in T$

$$u_{t_i}[\Phi_i^n(t)] \geq u_{t_i}[\Phi_i^n(t'_i, t_{-i})] - \epsilon.$$

It is then straightforward to prove the following result (see Appendix D).

Proposition 2. *If* $\{(\Phi^n)_{n\in\mathbb{N}}, T\}$ *is SP-L and equicontinuous, then given* $\epsilon > 0$ *there exists* n_0 *such that for all* $n > n_0$*, the mechanism* $\{(\Phi^n)_{n\in\mathbb{N}}, T\}$ *is* (ϵ, n) *strategyproof.*

Bayes-Nash equilibria always have some strictly positive probability of an ϵ ex-post deviation, because the empirical realization of agents' types could land very far from the underlying distribution of agents' types. By contrast, in an SP-L mechanism, agents do not really care whether or not the empirical is close to the true underlying distribution. This is the basic reason why SP-L mechanisms satisfy a stronger ex-post robustness condition than Bayes-Nash mechanisms.

6 Conclusion

This paper proposes strategyproofness in the large (SP-L) as a second-best alternative to strategyproofness (SP). Our main results show that, while it is well known that SP often severely limits what kinds of mechanisms are possible, there is a sense in which SP-L does not. Specifically, in our class of environments, SP-L is approximately costless to satisfy relative to other forms of incentive compatibility such as Bayes-Nash equilibrium or complete information Nash equilibrium, with the approximation error vanishing to zero in the large-market limit.

We view our results as providing formal justification for focusing on SP-L mechanisms when confronting a new market design problem for which there are no good SP solutions. Additionally, in some environments our proxy-like method of constructing an SP-L mechanism from a given non SP-L mechanism may be of direct use.

We conclude the paper with a few informal arguments in support of SP-L as a desideratum for market design, as well as some caveats.

Empirical Evidence on SP-L There are several empirical studies of mechanisms which are manipulable in the large, and which have been shown to have important incentives problems in practice. These include Jegadeesh (1993) and others on the 1991 pay-as-bid treasury auction scandals, Abdulkadiroğlu et al. (2006, 2009) on the Boston mechanism for school choice, Budish and Cantillon (Forthcoming) on Harvard Business School's course-allocation draft mechanism, Sönmez and Ünver (2010), Krishna and Ünver (2008) and Budish (Forthcoming) on the Bidding Points Auction, Edelman and Ostrovsky (2007) on pay-as-bid keyword auctions, Cramton and Katzman (2010) and Merlob et al. (2010) on a proposed Medicare auction for durable equipment, Roth (2002) and others on non-stable matching algorithms such as the priority match, and potentially others. By contrast, to the best of our knowledge, there are no empirical examples of market designs that are SP-L but which have been shown to be harmfully manipulated in large finite markets.

To the extent that this pattern is indeed true, it suggests that perhaps the relevant distinction for practice, in contexts with a large number of participants, is not "SP vs. not SP", but rather "SP-L vs. not SP-L." Or, more conservatively, "SP vs. SP-L vs. not SP-L."

Several Arguments for SP Design are also Arguments for SP-L Design In traditional mechanism design, incentives are viewed as a constraint, not an objective. A number of recent papers in the market design literature have suggested, either formally or informally, that strategyproofness be viewed as an explicit design objective. Many of these arguments can be interpreted as supporting SP-L design as well.

One such argument is that strategyproof mechanisms eliminate any unmodeled costs of calculating an optimal response; e.g., Roth (2008) argues that good markets are "sufficiently simple to participate in" and make it "safe to participate straightforwardly". Any SP-L mechanism has the following property: for any conjecture m about the distribution of opponents' play, and any cost $c > 0$ associated with calculating an optimal response, there exists n_0 such that in markets with $n > n_0$ participants, each agent maximizes her expected utility by simply reporting her preferences truthfully, and avoiding the cost c of strategizing.

A second such argument is that strategyproof mechanisms are fair, in the sense that they do not penalize participants who are strategically unsophisticated (Abdulkadiroğlu et al. (2006); Pathak and Sönmez (2008)). By an analogous argument to that in the previous paragraph, in a large enough market SP-L mechanisms are approximately fair, in the sense that the cost of being strategically unsophisticated can be bounded above by c.

Last, strategyproof mechanisms are prior free for the designer, and hence satisfy what has come to be known as the Wilson doctrine (Bergemann and Morris (2005)). SP-L mechanisms share this feature with SP mechanisms.

Caveats

We conclude with two important caveats on SP-L. First, there is no simple answer to the question of how large a market is large enough to ignore vanishing deviations.[25] We view the limit representation of a mechanism as a useful if imperfect abstraction for many interesting market design problems, just as the assumption of price-taking behavior is a useful abstraction in other parts of economics. In any specific context, the analyst's case for using an SP-L mechanism can be strengthened with empirical, experimental, or computational evidence suggesting that the gains from misreporting are small and/or rare; see, for instance, Roth and Peranson (1999).

Our second caveat relates to the difficulty of determining and reporting one's type. Many of the mechanisms that we have criticized as being manipulable in the large, and hence strategically complicated for participants, have the virtue that their message spaces are quite simple. For instance, in the Boston mechanism, it may be unrealistic to expect that a student will be able to accurately estimate the equilibrium p_j^*'s (cf. Example 2), but it seems realistic to expect that a student could determine which school to ask for as her first choice *given* the p_j^*'s. SP-L mechanisms are strategically simple, but require agents to report a potentially unrealistic amount of information about their preferences: for instance, in the SP-L mechanism we construct based on the Boston mechanism's equilibria (cf. Section 4.3), students' reports of their types consist of their von Neumann-Morgenstern utilities for each possible school, including schools to which they are highly unlikely to be assigned. An

[25]Indeed, even in analyses of the convergence properties of specific mechanisms, rarely is the analysis sufficient to answer the question of, e.g., "is 1000 participants large?" Convergence is often slow, or includes a large constant term. A notable exception is Rustichini et al. (1994), who are able to show, in the context of a double auction with unit demand and uniformly distributed values, that 6 buyers and sellers is large enough to approximate efficiency to within one percent.

interesting question that we leave for future research is how to define SP-L, or a criterion that is similar in spirit, in environments where reporting one's type is unrealistic.

A Proof of Theorem 1

As described in (4.1), let

$$F^n(t) = \Phi^n(\sigma^*_{\mathrm{emp}[t]}(t))$$

and define $f^n(t_i, m)$ from $F^n_i(t_i, t_{-i})$ as in (2.3). The core of the proof is the following approximation result.

Lemma 1. *Fix a prior μ_0 and $\epsilon > 0$. Let $\mathcal{N}$ be a neighborhood as in Definition 8. Let μ_k be priors $\mu_k \in \mathcal{A}_k$ for each $k = 1, \ldots, K$, with $|\mu_k - \mu_0| < \epsilon$. Then there exists n_0, such that for all $n > n_0$: there exist positive weights π^n_k with $\sum_{1 \le k \le K} \pi^n_k = 1$, such that for all t_i*

$$|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi^n_k \cdot z_k(t_i)| < 6\epsilon,$$

where

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

The Lemma states that the bundle received by an agent playing t_i in the constructed mechanism can be approximated by a convex combination of the bundles received when playing the original equilibrium within each region $\mathcal{A}_k$. Each $z_k(t_i)$ is defined as the bundle an agent receives when playing t_i when opponents' types and the prior on which equilibrium is selected are each within $\mathcal{A}_k$. The key assertion that the approximation Lemma makes is that the π^n_k do not depend on t_i. That is, irrespective of the type an agent reports, the approximation weights can be taken to be the same. This reflects the fact that a single agent has a very small effect on the probability of the distribution of types falling within each region $\mathcal{A}_k$.

We now prove the Lemma, and then use it to prove Theorem 1.

Proof of Lemma 1.

The proof of the Lemma involves three steps. Throughout the proof we use the shorthand $\hat{\mu} = \mathrm{emp}[t_i, t_{-i}]$. The first step shows that the appoximation formula holds within each region $\mathcal{A}_k$.

Step 1.

There exists n_0 such that, for all $n > n_0$ and $t \in \mathcal{A}_k$ we have

$$|\Phi^n_i(\sigma^*_{\hat{\mu}}(t_i), \sigma^*_{\hat{\mu}}(t_{-i})) - \phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))| < 4\epsilon.$$

Using the $z_k(t_i)$ notation, this is

$$|\Phi_i^n(\sigma_{\hat{\mu}}^*(t_i), \sigma_{\hat{\mu}}^*(t_{-i}) - z_k(t_i)| < 4\epsilon.$$

Proof. First note that, by Condition 3 of Definition 8 we may take n_1 such that for $n \geq n_1$

$$|\Phi_i^n(\sigma_{\hat{\mu}}^*(t_i), \sigma_{\hat{\mu}}^*(t_{-i})) - \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i}))| < \epsilon. \tag{A.1}$$

Note that the left term $\Phi_i^n(\sigma_{\hat{\mu}}^*(t_i), \sigma_{\hat{\mu}}^*(t_{-i}))$ is the term whose distance to $z_k(t_i)$ we wish to bound. We will do so by showing that $\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i}))$ is close to $\phi^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(\mu_k))$, and then showing that $\phi^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(\mu_k))$ is close to $z_k(t_i)$.

By definition we have that

$$\phi^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(\mu_k)) = \sum_{t'_{-i}} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t'_{-i})). \tag{A.2}$$

We will now bound the distance between $\phi^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(\mu_k))$ and $\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i}))$. For all $t \in \mathcal{A}_k$ we have

$$\begin{aligned}
& |\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i})) - \phi^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(\mu_k))| \\
= \quad & |\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i})) - \textstyle\sum_{t_{-i}} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t'_{-i}))| \\
\leq \quad & \textstyle\sum_{t'_{-i}} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot |\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i})) - \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t'_{-i}))| \\
= \quad & \textstyle\sum_{t'_{-i}:\mathrm{emp}[t_i,t'_{-i}] \in \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot |\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i})) - \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t'_{-i}))| \\
+ \quad & \textstyle\sum_{t'_{-i}:\mathrm{emp}[t_i,t'_{-i}] \notin \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot |\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i})) - \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t'_{-i}))|.
\end{aligned} \tag{A.3}$$

The first equality follows by substituting the definition of ϕ^n from Equation (A.2). 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} such that for which $\mathrm{emp}[t_i, t'_{-i}]$ is in $\mathcal{A}_k$, and the ones for which it is not. Consider now the expression on the right side of the last equality. Note that we may take n_1 such that the first term is bounded by

$$\sum_{t'_{-i}:\mathrm{emp}[t_i,t'_{-i}] \in \mathcal{A}_k} \Pr(t'_{-i}|t'_{-i} \sim \mu_k) \cdot |\Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t_{-i})) - \Phi_i^n(\sigma_{\mu_k}^*(t_i), \sigma_{\mu_k}^*(t'_{-i}))| < \epsilon,$$

which follows from Condition 3 in Definition 8. As for the second term, by the law of large numbers, we may take n_2 large enough such that the total probability mass that $\text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k$ is greater than $1 - \epsilon$. This bounds the second term by

$$\sum_{t'_{-i}:\text{emp}[t_i,t'_{-i}] \notin \mathcal{A}_k} \Pr(t'_{-i} | t'_{-i} \sim \mu_k) \cdot |\Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t_{-i})) - \Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t'_{-i})) < \epsilon,$$

Substituting these bounds in inequality (A.3) then yields

$$|\Phi_i^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(t_{-i})) - \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))| < \epsilon + \epsilon = 2\epsilon. \tag{A.4}$$

Finally, by the definition of the limit we may take n_3 such that for all $n > n_3$

$$|\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) - \phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))| < \epsilon. \tag{A.5}$$

If we take $n_0 = \max\{n_1, n_2, n_3\}$ the Lemma then follows from Inequalities (A.1), (A.4) and (A.5). □

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 says 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.

There exists n_0 such that, for all $n > n_0$ there exist weights $\pi_1^n, \ldots, \pi_K^n$ such that $\sum_k \pi_k^n = 1$ and

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

for all k and all t_i.

Proof. We begin by constructing numbers which will be approximately equal to the π_k^n in the statement of this step. Let

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

be the probability that a vector of n types drawn independently according to μ_k is in $\mathcal{A}_k$. We will show that for large n these $\bar{\pi}_k^n$ are very close to the probabilities $\Pr\{(t_i, t_{-i}) \in \mathcal{A}_k | t_{-i} \sim \mu_0\}$. For any type t_i, 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, versus when i's type is drawn randomly, is

$$\begin{aligned}\Pr((t_i, t'_{-i}) \in \mathcal{A}_k | t'_{-i} \in T^{n-1}, t'_{-i} \sim \mu_0) - \bar{\pi}^n_k &= \\ \Pr((t_i, t'_{-i}) \in \mathcal{A}_k, t' \notin \mathcal{A}_k | t' \in T^n, t' \sim \mu_k) & \\ - \Pr((t_i, t'_{-i}) \notin \mathcal{A}_k, t' \in \mathcal{A}_k | t' \in T^n, t' \sim \mu_k). & \end{aligned} \tag{A.6}$$

This is just the difference between the probability of choosing a vector t' where changing a single type (i's) from t'_i to t_i moves the vector of types from inside $\mathcal{A}_k$ to outside $\mathcal{A}_k$, and the probability of choosing a vector where changing i's type from t'_i to t_i moves the vector from outside $\mathcal{A}_k$ to inside $\mathcal{A}_k$. We now show that the probability of such vectors being drawn can be taken to be very small.

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 $n > n_0$ 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 that case, the segment $[(t_i, t'_{-i}), t']$ must have a point that lies in $\mathcal{B}$, as we assumed $\mathcal{N}$ to be convex. This means that the distance between t' and $\mathcal{B}$ is at most $1/n$. By Condition 2 of Definition 8, we may take n_0 such that this probability is less than $\epsilon/8$. This argument then yields that

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

An analogous argument proves that we may assume that for $n > n_0$

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

Substituting these two inequalities in Equation (A.6) yields that

$$|\Pr((t_i, t'_{-i}) \in \mathcal{A}_k | t'_{-i} \in T^{n-1}, t'_{-i} \sim \mu_0) - \bar{\pi}^n_k| < \epsilon/4 + \epsilon/4 = \epsilon/2. \tag{A.7}$$

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

$$\pi^n_k = \bar{\pi}^n_k / \sum_{k'} \bar{\pi}^n_{k'}. \tag{A.8}$$

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 $n > n_0$

$$|1 - 1/\sum_{k'} \bar{\pi}^n_{k'}| < \epsilon/2. \tag{A.9}$$

Putting this together, we may finish the proof of Step 2.

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

The series of steps in the above derivation were as follows. The second line derives from the triangle inequality. The third line uses the bound from Inequality (A.7). The fourth line uses the definition of π^n_k from equation (A.8). Finally, the fifth line is algebra, and the sixth line comes from the bound in inequality (A.9). □

Step 3.

Finally, we apply the results from steps 1 and 2 to prove the Lemma, obtaining the desired approximation formula. That is, there exists n_0 such that for all $n \geq n_0$

$$|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi^n_k \cdot z_k(t_i)| < 6\epsilon.$$

Proof. We may write

$$f^n(t_i, \mu_0) - \sum_k \pi^n_k \cdot z_k(t_i) = \sum_{t_{-i}} \Pr(t_{-i} | t_{-i} \sim \mu_0) \cdot \Phi^n_i(\sigma^*_{\hat{\mu}}(t_i), \sigma^*_{\hat{\mu}}(t_{-i})) - \sum_k \pi^n_k \cdot z_k(t_i).$$

This sum can be decomposed depending on whether $\hat{\mu}$ is in each of the $\mathcal{A}_k$ sets or not. We have

$$f^n(t_i,\mu_0) \quad - \textstyle\sum_k \pi_k^n \cdot z_k(t_i) = \displaystyle\sum_k \Big(\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - \pi_k^n\cdot z_k^n(t_i)\Big) + \sum_{t_{-i}:\hat{\mu}\notin\cup_k\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})). \tag{A.10}$$

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,

$$\begin{aligned}
&\Big|\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - \pi_k^n\cdot z_k(t_i)\Big| \\
&\quad\le \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot|\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - z_k(t_i)| \\
&\qquad + \Big|\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0) - \pi_k^n\Big|\cdot|z_k(t_i)| \\
&\le \max_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} |\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - z_k(t_i)|\cdot \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0) \\
&\qquad + \Big|\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0) - \pi_k^n\Big|
\end{aligned} \tag{A.11}$$

The first inequality follows from the triangle inequality. The second inequality bounds each term $|\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - z_k^n(t_i)|$ by the maximum value of these terms, and it bounds $|z_k(t_i)|$ by 1.

Consider now the right side of the last inequality. By step 1, we may take n_0 such that for all $n \ge n_0$,

$$\max_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} |\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - z_k(t_i)| < 4\epsilon.$$

By step 2, for $n \ge n_0$ for a suitable n_0 the second term is bounded by

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

Substituting these two bounds in inequality (A.11) we have that for all $n \ge n_0$

$$\begin{aligned}
&\Big|\textstyle\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - \pi_k^n\cdot z_k^n(t_i)\Big| \\
\le\ & 4\epsilon\cdot\textstyle\sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0) + \frac{\epsilon}{K}.
\end{aligned}$$

Summing over all k we get

$$\sum_k | \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - \pi_k^n\cdot z_k^n(t_i)| \leq 5\epsilon$$

and then using the triangle inequality we can bring the summation inside to get

$$|\sum_k \sum_{t_{-i}:\hat{\mu}\in\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0)\cdot\Phi_i^n(\sigma^*_{\hat{\mu}}(t_i),\sigma^*_{\hat{\mu}}(t_{-i})) - \pi_k^n\cdot z_k^n(t_i)| \leq 5\epsilon. \tag{A.12}$$

The argument above bounds the terms in equation (A.10) that correspond to t within the sets $\mathcal{A}_k$. To bound the other term, note that we may take n_0 to be large enough so that for all $n \geq n_0$ the total probability that $t \notin \cup_k\mathcal{A}_k$ is strictly less than ϵ. That is,

$$\sum_{t_{-i}:\hat{\mu}\notin\cup_k\mathcal{A}_k} \Pr(t_{-i}|t_{-i}\sim\mu_0) < \epsilon. \tag{A.13}$$

Plugging in equations (A.12) and (A.13) in equation (A.10) we obtain

$$|f^n(t_i,\mu_0) - \sum_{k=1}^{K}\pi_k^n\cdot z_k(t_i)| < 6\epsilon,$$

completing the proof of Step 3, and hence the Lemma. □

With the Lemma in hand, it is a simple matter to establish Theorem 1.

Proof. (Theorem 1)

Part 1.

To see that the constructed mechanism is SP-L, consider the gain for type t_i from deviating to $\hat{t}_i$ when opponents play μ_0. That is,

$$u_{t_i}[f^n(\hat{t}_i,\mu_0)] - u_{t_i}[f^n(t_i,\mu_0)].$$

By the approximation Lemma, and the boundedness of u, given $\epsilon > 0$, there exists n_0, π_k^n, μ_k, and

z_k as in the statement of the Lemma such that, for all $n > n_0$:

$$|u_{t_i}[f^n(\hat{t}_i, \mu_0)] - \sum_k \pi_k^n \cdot u_{t_i}[z_k(\hat{t}_i)]| < \epsilon/2 \tag{A.14}$$

$$|u_{t_i}[f^n(t_i, \mu_0)] - \sum_k \pi_k^n \cdot u_{t_i}[z_k(t_i)]| < \epsilon/2. \tag{A.15}$$

Also, by the definition of $z_k(\cdot)$, we have that

$$u_{t_i}[z_k(t_i)] \geq u_{t_i}[z_k(\hat{t}_i)]. \tag{A.16}$$

Therefore, we may bound the gain from deviating for $n > n_0$ by

$$\begin{aligned} u_{t_i}[f^n(\hat{t}_i, \mu_0)] - u_{t_i}[f(t_i, \mu_0)] &\leq \\ \sum_k \pi_k^n \cdot \{u_{t_i}[z_k(\hat{t}_i)] - u_{t_i}[z_k(t_i)]\} & \\ +|u_{t_i}[f^n(\hat{t}_i, \mu_0)] - \sum_k \pi_k^n \cdot u_{t_i}[z_k(\hat{t}_i)]| & \\ +|u_{t_i}[f^n(t_i, \mu_0)] - \sum_k \pi_k^n \cdot u_{t_i}[z_k(t_i)]| &< \\ 0 + \epsilon/2 + \epsilon/2 &= \epsilon. \end{aligned}$$

The first inequality follows from the triangle inequality, and the second inequality from the bounds in inequalities (A.14), (A.15), (A.16).

Part 2(b).

Part 2(b) follows from the approximation Lemma. Given $\mu_0 \in \Delta T, \epsilon > 0$, by the Lemma we may take n_0, μ_k such that all $|\mu_k - \mu| < \epsilon$, and for all $n \geq n_0$

$$|f^n(t_i, \mu_0) - \sum_k \pi_k^n \cdot \phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))| < \epsilon/2. \tag{A.17}$$

By the definition of the limit, we may take n_0 such that for all k, t_i, and $n > n_0$

$$|\phi^\infty(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k)) - \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))| < \epsilon/2. \tag{A.18}$$

By the triangle inequality and inequalities (A.17) and (A.18) we have that

$$|f^n(t_i, \mu_0) - \sum_k \pi_k^n \cdot \phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))| < \epsilon/2 + \epsilon/2 = \epsilon. \quad \text{(A.19)}$$

Part 2(a).

Finally for Part 2(a), note that we may take $\mathcal{N} = \mathcal{A}_1$ and $\mu_1 = \mu_0$ in the continuous case. Therefore $\pi_k^n = 1$, and equation (A.19) becomes

$$|f^n(t_i, \mu_0) - \phi^n(\sigma^*_{\mu_0}(t_i), \sigma^*_{\mu_0}(\mu_0))| < \epsilon.$$

□

B Multi-Unit Auctions

This section provides further detail on the uniform-price and pay-as-bid multi-unit auctions described above in Example 1. We will derive the limit of these mechanisms, show that uniform price auctions are SP-L, derive a particular family of limit equilibria of the pay-as-bid auction, and show that it satisfies our quasi-continuity condition.

B.1 The Mechanisms

As discussed in the text, there are kn units of a homogeneous good, with $k \in \mathbb{Z}_+$. Agents' preferences take the form of linear utility functions, 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, with $V = \{0, 1, \ldots, \bar{v}\}$ the set of possible values, $Q = \{0, 1, \ldots, \bar{q}\}$ the set of possible capacity limits, and $T = V \times Q$. We can denote the set of outcomes, X_0, by $X_0 = V \times Q$ as well, by modeling an outcome as consisting of a per-unit payment, bounded above by $\bar{v}$, and a quantity allocated, bounded above by $\bar{q}$. Utility is then given by

$$u_{t_i}[(v, q)] = v_i \cdot \min\{q, q_i\} - v \cdot q.$$

In both auctions, agents simply report their types, $A = T$. To define the market clearing price given a vector of reports $t = ((v_1, q_1), \ldots, (v_n, q_n))$, let

$$D(p|t) = \sum_{i:v_i \geq p} q_i.$$

The market clearing price is defined then as

$$p^*(t) = \max\{p \in V : D(p|t) \geq k\}.$$

i.e., p^* is the highest price at which demand weakly exceeds supply. The market clearing price is defined as 0 if there is always excess supply. Allocations of the good are equivalent across the two mechanisms: an agent who reports (v_i, q_i) is allocated q_i units if $v_i > p^*$, is allocated 0 units if $v_i < p^*$, and is rationed if $v_i = p^*$. Agents bidding p^* will be rationed if

$$D(p^*(t)|t) > kn.$$

In that case, we define the rationing probability as

$$\pi^*(t) = \frac{D(p^*(t)|t) - D(p^*(t)+1|t)}{kn - D(p^*(t)+1|t)}. \tag{B.1}$$

The exact form of the rationing will be immaterial, as long as no agent receives more than q_i units, and the expected number of units each agent receives is $\pi^*(t) \cdot q_i$. For concreteness, we assume that agents are rationed by random serial dictatorship, where they are randomly put in a line, and sequentially take q_i units, or as many as there are left, until there are no units left.

Payments differ across the two mechanisms. In the uniform-price auction, every agent who is allocated units pays the same per-unit price, p^*. In the pay-as-bid auction, every winner pays her bid. So, for the discriminatory price auction, for example,

$$\begin{aligned}
\Phi^n_{d,i}(t_i, t_{-i}) &= (v_i, q_i) \text{ if } v_i > p^*(t) \\
\Phi^n_{d,i}(t_i, t_{-i}) &= (0, 0) \text{ if } v_i > p^*(t).
\end{aligned}$$

If $v_i = p^*(t)$, the agent will be rationed. In this case the expected bundle she receives is

$$E\Phi^n_{d,i}(t_i, t_{-i}) = (v_i, \pi^*(t) \cdot q_i) \text{ if } v_i = p^*(t).$$

The exact lottery over deterministic bundles is of course a more complicated object, given by the serial dictatorship procedure.

For the uniform price auction, the allocations are similar, but with agents paying the bid of the marginal winner:

$$\begin{aligned} \Phi^n_{u,i}(t_i, t_{-i}) &= (p^*(t), q_i) \text{ if } v_i > p^*(t) \\ \Phi^n_{u,i}(t_i, t_{-i}) &= (0, 0) \text{ if } v_i > p^*(t). \end{aligned}$$

$$E\Phi^n_{u,i}(t_i, t_{-i}) = (p^*(t), \pi^*(t) \cdot q_i) \text{ if } v_i = p^*(t).$$

B.2 Large Economies

For a given measure m, the lottery $\phi^n(t_i|m)$ can be quite complicated. It must take into account a probability distribution over market clearing prices, and the possibility that a bid of v_i is rationed when $p^*(t) = v_i$. Fortunately, the limit allocation $\phi^\infty(t_i|m)$ is quite simple. To describe it, given $m \in \Delta T$, define the demand function

$$D(p|m) = \sum_{t_i : v_i(t_i) \geq p} q_i(t_i) \cdot m(t_i).$$

That is, the average mass of agents with values of at least p. We define the market clearing price as the highest price at which demand weakly exceeds supply.

$$p^*(m) = \max\{p \in V : D(p|m) \geq k\}.$$

Throughout this section we will concentrate on distributions of actions within the set

$$\mathcal{M} = \{m \in \Delta A : p^*(m) = 0 \text{ or } \exists q_i : (p^*(m) - 1, q_i) \in \text{support}[m]\}.$$

We focus on this set because equilibrium analysis will only depend on such distributions.

There are two cases to consider. If

$$D(p^*(m)|m) > k,$$

then by the law of large numbers the probability that the marginal winning bidder will have a bid $v_i = p^*(m)$ is converging to 1. In this case, some of these agents will be rationed. Define the rationing probability as

$$\pi^*(m) = \frac{D(p^*(m)|m) - D(p^*(m)+1|m)}{k - D(p^*(m)+1|m)}.$$

With the assumption that rationing is by random serial dictatorship, it is very unlikely, in a large economy, that any individual agent will be the last one to receive the good. All other rationed agents either receive q_i units or 0. So, in the limit, rationed agents receive q_i units of the good with probability $\pi^*(m)$, and 0 units with probability $1 - \pi^*(m)$.

We highlight that we have defined $D(p|t), p^*(t), \pi^*(t)$ for a given profile of types or actions, and $D(p|m), p^*(m), \pi^*(m)$ for distributions over actions. The definitions of these objects differ. We used the same symbols for the functions as they are analogous, to save on notation. In what follows, the argument of the functions makes clear whether we are considering for example $D(p|t)$ for a profile of types t or $D(p|m)$ for a distribution over actions m.

Consider now the case where average demand exactly equals supply

$$D(p^*(m)|m) = k.$$

This is the case where, in a continuum, there is enough of the object to exactly give the goods to the agents with value $p^*(m)$. In a large economy, this means that about half the time the marginal winning bidder will have a bid of $p^*(m)$, and about half the time a lower value. Define the next lowest bid after $p^*(m)$ as

$$p^*_{-1}(m) = \max\{p^*(m) - 1, 0\}.$$

That is, $p^*_{-1}(m)$ is the highest price that (i) is lower than $p^*(m)$, and (ii) there are bids of p in the support of m. In case $p^*(m) = 0$, we have $p^*_{-1}(m)$ defined as 0 also.

In large economies, the marginal winning bid will be $p^*(m)$ with probability close to 50%, and the next lowest bid $p^*_{-1}(m)$ with remaining probability of about 50%. Note also that, in either case, almost all agents with bids $v_i = p^*(m)$ will not be rationed. On the other hand, almost all agents with bids $v_i = p^*_{-1}(m)$ will be rationed (unless $0 = p^*(m) = p^*_{-1}(m)$).[26]

Note that, in either case, the limit allocation can be described in the same way. Bids higher than $p^*(m)$ are never rationed, and bids lower than $p^*(m)$ almost never win any objects. Bids of exactly $p^*(m)$ are rationed with probability $\pi^*(m)$, which can be 1 in the second case where demand exactly clears supply.

The limit mechanisms are as follows. For the pay-as-bid auction,

$$\begin{aligned} E\phi_d^\infty(t_i|m) &= (v_i, q_i) \text{ if } v_i > p^*(m) \\ &\quad (v_i, \pi^*(m)\cdot q_i) \text{ if } v_i = p^*(m) \\ &\quad 0 \text{ if } v_i < p^*(m). \end{aligned} \tag{B.2}$$

This is just the allocation we described above, where bids of exactly $p^*(m)$ are possibly rationed, and higher (lower) bids always win (lose) .

For the uniform price auction, the allocation of the objects is the same, but payments differ. If

$$D(p^*(m)|m) > k,$$

then the marginal price is almost certainly $p^*(m)$ and we have

$$\begin{aligned} E\phi_u^\infty(t_i|m) &= (p^*(m), q_i) \text{ if } v_i > p^*(m) \\ &\quad (p^*(m), \pi^*(m)\cdot q_i) \text{ if } v_i = p^*(m) \\ &\quad 0 \text{ if } v_i < p^*(m). \end{aligned} \tag{B.3}$$

In the case where demand exactly equals supply, the marginal price will be either p^*_{-1} or p^* with probability 50%. That is, if

$$D(p^*(m)|m) = k$$

[26]Note that these conclusions rely on our restriction that $m \in \mathcal{M}$. If for example we had m placing all of its mass on a single action such as $(v_i, q_i) = (5, 10)$, and demand being satisfied exactly, then it would no longer be the case that the marginal winning bid would be random. This is immaterial for our equilibrium analysis, where only action distributions $m \in \mathcal{M}$ will arise.

we have

$$\phi_u^\infty(t_i|m) = \frac{1}{2}(p^*(m), q_i) + \frac{1}{2}(p^*(m) - 1, q_i) \text{ if } v_i \geq p^*(m) \qquad \text{(B.4)}$$
$$0 \text{ if } v_i < p^*(m).$$

B.3 The Uniform Price Auction is SP-L

Recall that for SP-L we restrict attention to priors μ with full support. In particular, under truth-telling, it will always be the case that the distribution of actions satisfies $m = \mu \in \mathcal{M}$, so that the formulae for the limit mechanisms in the previous section may be used.

It is immediate from Equations (B.3) and (B.4) for ϕ_u^∞ that the uniform price auction is strategyproof. Whatever $\mu \in \bar{\Delta}T$ is, it is always weakly optimal for an agent of type t_i to report truthfully. Note that this is true even though we know that this mechanism isn't exactly strategyproof. It has vanishing ex ante deviations, as an agent might have ex ante incentives to reduce capacity in a small economy. And it has ex post deviations in knife edge cases in large economies, where an agent happens to know it is pivotal. But it is still SP-L, as no matter what μ is, reporting truthfully is always optimal under ϕ_u^∞.

B.4 Equilibria of the Pay-as-Bid Auction

We now derive one family of equilibria $(\sigma_\mu)_{\mu\in\bar{\Delta}T}$ of the discriminatory price auction. There are other families of equilibria, but we focus on this particular family for concreteness, and because it is similar to equilibria of a model where types and bids are distributed according to a continuous distribution over an interval. Since we are only interested in establishing quasi-continuity of this family, we restrict attention to priors $\mu \in \bar{\Delta}T$.

Given a prior μ over types, define

$$\bar{p} = p^*(\mu)$$
$$\bar{\pi} = \pi^*(\mu)$$
$$\bar{p}_{-1} = \max\{\bar{p} - 1, 0\}.$$

That is, $\bar{p}$ would be the market clearing price under truthfull reporting, and $\bar{\pi}$ the associated rationing probability. Consider the strategy σ_μ where for $\bar{p} > 0$:

- Agents with $v_i > \bar{p}$ play $(\bar{p}, q_i)$.
- Agents with $v_i = \bar{p}$ play $(\bar{p}, q_i)$ with probability $\bar{\pi}$, and $(\bar{p}_{-1}, q_i)$ with probability $1 - \bar{\pi}$.
- Agents with $v_i < \bar{p}$ play $(\bar{p}_{-1}, q_i)$.

That is, all agents report their capacities truthfully. With respect to the price, the agents with values above $\bar{p}$ bid $\bar{p}$. Those with values exactly equal to $\bar{p}$ mix between bidding $\bar{p}$ and the lower value $\bar{p}_{-1}$. And the ones with values lower than $\bar{p}$ simply play $\bar{p}_{-1}$. We highlight that, under the assumption that $\mu \in \bar{\Delta}T$, whenever $\bar{p} > 0$ there exists a positive mass of agents bidding $\bar{p}_{-1}$.

For $\bar{p} = 0$ strategies are slightly different:

- Agents with $v_i > 0$ play $(0, q_i)$.
- Agents with $v_i = 0$ play $(0, q_i)$ with probability $\bar{\pi}$, and $(0, 0)$ with probability $1 - \bar{\pi}$.

We now argue that these strategies constitute a limit equilibrium. Consider the case $\bar{p} > 0$. First note that, if agents follow these reports, the resulting measure of bids $m = \sigma_\mu(\mu)$ clears the market exactly at prices $\bar{p}$ in the limit. That is

$$D(\bar{p}|m) = k.$$

Therefore, in a large finite economy, the realized market clearing price $p^*(t)$ will be $\bar{p}$ or $\bar{p}_{-1}$ with probability roughly equal to 50%. Moreover, an agent bidding $\bar{p}_{-1}$ will be rationed almost certainly, while an agent bidding $\bar{p}$ will almost certainly receive the quantity he asked for. From Equation (B.2) we have that the limit allocation received by an agent bidding t_i is simply

$$\begin{aligned} \phi_d^\infty(t_i|m) &= (v_i, q_i) \text{ if } v_i \geq \bar{p} \\ & \quad 0 \text{ if } v_i < \bar{p}. \end{aligned} \tag{B.5}$$

The case $\bar{p} = 0$ is similar. It is also the case that agents with valuations $v_i = 0$ mix so that the probability of being rationed with a bid of $\bar{p} = 0$ is negligible. Therefore, Equation (B.5) also describes the equilibrium allocation when $\bar{p} = 0$.

From Equation (B.5), it follows that σ_μ is a limit equilibrium. No agent will ever want to bid more than $\bar{p}$, as bidding $\bar{p}$ is enough to win q_i objects with near certainty. The agents with $v_i > \bar{p}$ are best responding, as they are willing to pay $\bar{p}$ to win the object. Likewise, the agents with $v_i = \bar{p}$ are indifferent between winning or not, so they are best responding too. Finally, the agents with $v_i < \bar{p}$ would not be willing to pay $\bar{p}$ to win, so they are best responding.

B.5 Quasi-Continuity of Equilibria of the Discriminatory Price Auction

To prove that the family of equilibria $(\sigma_\mu)_{\mu\in\bar{\Delta}T}$ is quasi-continuous, we will establish two useful Lemmas.

The first Lemma considers a distribution of actions m where all agents bid one of two prices, $p^*(m)$ and $p^*(m)-1$. We consider this case because all distributions of actions $m = \sigma_\mu(\mu)$ in equilibrium have this form. The Lemma shows that whatever the realized profile of actions a, as long as its empirical distribution is close to m, the allocation does not vary too much.

Lemma 2. *Consider $m \in \Delta T$ such that:*

$$\begin{aligned} \textit{support}(m) &\subseteq \{(v,q) : v = p^*(m) \text{ or } p^*(m)-1\} \\ D(p^*(m)|m) &= k \\ D(p^*(m)-1|m) &> k, \end{aligned}$$

Given $\epsilon > 0$, there exists $\delta > 0$ such that for all $a \in A^n$ with

$$|\operatorname{emp}[a] - m| < \delta$$

then

$$|\Phi^n_{d,i}(a_i, a_{-i}) - \phi^\infty(a_i, m)| < \epsilon.$$

Proof. First note that under such m, by the formulae for ϕ^∞, we have

$$\begin{aligned} \phi^\infty(a_i, m) &= (v_i, q_i) \text{ for } v_i \geq p^*(m) \\ & \quad 0 \text{ for } v_i < p^*(m). \end{aligned} \tag{B.6}$$

We must show that this allocation is close to $\Phi^n_{d,i}(a_i, a_{-i})$. Denote $a_i = (v_i, q_i)$.

Note that, by our definition of demand,[27]

$$D(p|a) = D(p|\operatorname{emp}[a]) \cdot n.$$

Consider the case where $0 < p^*(m) < \bar{v}$. Since

$$D(p^*(m)+1, m) < D(p^*(m)|m) = k < D(p^*(m)-1, m),$$

we may take δ to be small enough such that

$$D(p^*(m)+1, a) < k < D(p^*(m)-1, a).$$

This guarantees that $p^*(a)$ is either $p^*(m)$ or $p^*(m)-1$. In particular, if an agent bids $v_i > p^*(m)$ she receives the good for sure, and if she bids $v_i < p^*(m)-1$ she never receives the good. This proves Equation (B.6) in all cases, except $v_i = p^*(m)-1$ and $p^*(m)$.

Consider the case where $v_i = p^*(m)$. We have to show that

$$|\Phi^n_{d,i}(a_i, a_{-i}) - (v_i, q_i)| < \epsilon, \tag{B.7}$$

that is, that the probability that such agent i is rationed is sufficiently small. In case the market clearing price $p^*(a) = p^*(m)-1$, this is evidently true, as i is rationed with 0 probability. In the case where $p^*(a) = p^*(m)$, the rationing probability $\pi^*(a)$ is given by Equation (B.1). Since this varies continuously with the empirical distribution of a, and equals 0 for $\operatorname{emp}[a] = m$, we may take δ small enough such that Inequality (B.7) is satisfied. The case $v_i = p^*(m)-1$ is analogous. This completes the proof in the case where $0 < p^*(m) < \bar{v}$.

The case $p^*(m) = \bar{v}$ follows basically the same argument. It is also the case that for small enough δ we have $p^*(a) = p^*(m)$ or $p^*(m)-1$, and the rest of the argument is analogous. The case $p^*(m) = 0$ is even simpler, as for δ small enough we always have that $p^*(a) = 0$. The argument above carries over easily to this case. □

The second Lemma is a key step to establishing quasi-continuity. It shows that, given a prior μ, if

[27] Recall that we defined $D(p|a)$ for a vector of actions differently than $D(p|m)$ for a distribution over actions. We refer the reader to the first two Subsections of this Section for these definitions.

the empirical distribution of a vector of types t is close to μ, and agents play strategies σ' which are close to $\sigma_\mu(\mu)$, then in large markets the outcome is approximately the same as in the limit equilibrium with prior μ.

Lemma 3. *Consider $\mu \in \bar{\Delta}T, \epsilon > 0$. There exists δ and n_0 such that for any t_i, t_{-i} and σ' with*

$$\begin{aligned} |\operatorname{emp}[t] - \mu| &< \delta \\ |\sigma_\mu - \sigma'| &< \delta \end{aligned}$$

we have that for all $n \geq n_0$

$$|\Phi^n_{d,i}(\sigma'(t_i), \sigma'(t_{-i})) - \phi^\infty(\sigma_\mu(t_i), \sigma_\mu(\mu))| < \epsilon.$$

Proof. Let $m = \sigma_\mu(\mu)$. Therefore, m satisfies the assumptions of the previous Lemma. Therefore there exists δ_1 such that for all $a \in A^n$ with

$$|\operatorname{emp}[a] - m| < \delta_1$$

we have

$$|\Phi^n_{d,i}(a) - \phi^\infty(a_i, m)| < \epsilon/3. \tag{B.8}$$

Fix now a_i, and consider the probability that a vector $a = [a_i, a_{-i}]$, with $a_{-i} \in A^{n-1}$ drawn according to $\sigma'(t)$ satisfies $|\operatorname{emp}[a] - m| > \delta_1$. That is

$$\sum_{a_{-i} \in A^{n-1}} \Pr\{|\operatorname{emp}[a] - m| > \delta_1 | a_{-i} \sim \sigma'(t_{-i})\}.$$

If we take δ small enough so that σ' is sufficiently close to σ_μ, and $\operatorname{emp}[t]$ sufficiently close to μ, we can apply the law of large numbers, and take n_0 such that for all $n \geq n_0$ this probability is bounded by

$$\sum_{a_{-i} \in A^{n-1}} \Pr\{|\operatorname{emp}[a] - m| > \delta_1 | a_{-i} \sim \sigma'(t_{-i})\} < \epsilon/3. \tag{B.9}$$

Consider now the expression

$$|\Phi^n_{d,i}(a_i, \sigma'(t_{-i})) - \phi^\infty(a_i, \sigma_\mu(\mu))|$$

which we wish to bound. We may decompose it as

$$\begin{aligned}
& |\Phi^n_{d,i}(a_i, \sigma'(t_{-i})) - \phi^\infty(a_i, \sigma_\mu(\mu))| \\
= \quad & |\textstyle\sum_{a_{-i} \in A^{n-1}} \Pr\{a_{-i} | a_{-i} \sim \sigma'(t_{-i})\} \Phi^n_{d,i}(a) - \phi^\infty(a_i, \sigma_\mu(\mu))| \\
\leq \quad & \textstyle\sum_{a_{-i} \in A^{n-1}} \Pr\{a_{-i} | a_{-i} \sim \sigma'(t_{-i})\} \cdot |\Phi^n_{d,i}(a) - \phi^\infty(a_i, \sigma_\mu(\mu))| \\
= \quad & \textstyle\sum_{|\operatorname{emp}[a] - m| < \delta_1} \Pr\{a_{-i} | a_{-i} \sim \sigma'(t_{-i})\} \cdot |\Phi^n_{d,i}(a) - \phi^\infty(a_i, \sigma_\mu(\mu))| \\
+ \quad & \textstyle\sum_{|\operatorname{emp}[a] - m| \geq \delta_1} \Pr\{a_{-i} | a_{-i} \sim \sigma'(t_{-i})\} \cdot |\Phi^n_{d,i}(a) - \phi^\infty(a_i, \sigma_\mu(\mu))|
\end{aligned}$$

The second line follows from simply writing down the left term $\Phi^n_{d,i}(\sigma'(t_i), \sigma'(t_{-i}))$ as a sum over realized profiles of actions a. The third line follows from the triangle inequality. The fourth and fith lines break this sum into profiles of actions a that have an empirical close or far from m. Finally, substituting Inequalities (B.8) and (B.9) we get that

$$|\Phi^n_{d,i}(a_i, \sigma'(t_{-i})) - \phi^\infty(a_i, \sigma_\mu(\mu))| < \epsilon/3 + \epsilon/3 = 2\epsilon/3. \tag{B.10}$$

Moreover, we may take these bound to be uniform over all a_i. To complete the proof, note that we can take δ small enough such that

$$|\phi^\infty(\sigma'(t_i), \sigma_\mu(\mu)) - \phi^\infty(\sigma_\mu(t_i), \sigma_\mu(\mu))| < \epsilon/3. \tag{B.11}$$

Using these last two bounds we have that

$$\begin{aligned}
& |\Phi^n_{d,i}(\sigma'(t_i), \sigma'(t_{-i})) - \phi^\infty(\sigma_\mu(t_i), \sigma_\mu(\mu))| \\
\leq \quad & |\Phi^n_{d,i}(\sigma'(t_i), \sigma'(t_{-i})) - \phi^\infty(\sigma'(t_i), \sigma_\mu(\mu))| \\
+ \quad & |\phi^\infty(\sigma'(t_i), \sigma_\mu(\mu)) - \phi^\infty(\sigma_\mu(t_i), \sigma_\mu(\mu))| \\
< \quad & 2\epsilon/3 + \epsilon/3 = \epsilon.
\end{aligned}$$

The first inequality follows from the triangle inequality, and the second inequality follows from Inequalities (B.10) and (B.11). □

We are now ready to show that the family σ_μ is quasi-continuous. Fix $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$. Let

$$\mathcal{N} = \{\mu \in \Delta T : |\mu - \mu_0| < \delta\},$$

where δ will be determined to satisfy the requirements of Definition (8). Throughout, we take δ to be small enough such that $\mathcal{N} \subseteq \bar{\Delta}T$.

We begin with the simplest case where

$$D(p^*(\mu_0)|\mu_0) > k.$$

In this case, $p^*(\mu) = p^*(\mu_0)$ for all $\mu \in \mathcal{N}$, as long as we take δ to be small enough. That is, small changes in μ do not change the market clearing price. We may simply take $\mathcal{A}_1 = \mathcal{N}$. Given ϵ and μ_0, take δ_1 and n_0 as per Lemma 3. We then have that for all t_i, t_{-i} and σ' with

$$\begin{aligned} |\operatorname{emp}[t] - \mu| &< \delta_1 \\ |\sigma_{\mu_0} - \sigma'| &< \delta_1 \end{aligned}$$

we have that for all $n \geq n_0$

$$|\Phi^n_{d,i}(\sigma'(t_i), \sigma'(t_{-i})) - \phi^\infty(\sigma_{\mu_0}(t_i), \sigma_{\mu_0}(\mu_0))| < \epsilon/2. \tag{B.12}$$

Note that since $p^*(\mu)$ is constant in $\mathcal{N}$, and therefore the rationing probability $\pi^*(\mu)$ varies continuously with μ in $\mathcal{N}$ by Equation (B.1). Consequently, we may take δ small enough such that

$$|\sigma_{\mu_0} - \sigma_\mu| < \delta_1$$

for all $\mu \in \mathcal{N}$. If we take $\delta \leq \delta_1$, then Inequality (B.12) is satisfied for any $t, \mu \in \mathcal{N}$. This implies that for any t_i and $[t_i, t_{-i}], [t_i, t'_{-i}] \in \mathcal{N}, \mu, \mu' \in \mathcal{N}$ we have that

$$|\Phi^n_{d,i}(\sigma_\mu(t_i), \sigma_\mu(t_{-i})) - \Phi^n_{d,i}(\sigma_{\mu'}(t_i), \sigma_{\mu'}(t'_{-i}))| < \epsilon.$$

This completes the proof in the case $D(p^*(\mu_0)|\mu_0) > k$.

Consider now the case where

$$D(p^*(\mu_0)|\mu_0) = k.$$

In this case, σ_μ may change discontinuosuly with μ. Assume for now that $p^*(\mu) > 0$. Note that, due to the assumption that $\mu \in \mathcal{P}$, we have that

$$D(p^*(\mu_0) + 1|\mu_0) < D(p^*(\mu_0)|\mu_0) = k < D(p^*(\mu_0) - 1|\mu_0).$$

In particular, we may take δ small enough such that for all $\mu \in \mathcal{N}$ we have

$$D(p^*(\mu_0) + 1|\mu) < k < D(p^*(\mu_0) - 1|\mu).$$

Therefore, for any such μ either $p^*(\mu) = p^*(\mu_0)$ or $p^*(\mu) = p^*(\mu_0) - 1$. We then define the two sets

$$\begin{aligned}\mathcal{A}_1 &= \{\mu \in \mathcal{N} : p^*(\mu) = p^*(\mu_0) - 1, D(p^*(\mu_0) - 1 \neq k\} \\ \mathcal{A}_2 &= \{\mu \in \mathcal{N} : p^*(\mu) = p^*(\mu_0), D(p^*(\mu_0) \neq k\}.\end{aligned}$$

By the argument used for the case where $D(p^*(\mu_0)|\mu_0) > k$, we have that strategies σ_μ vary continuously within each $\mathcal{A}_k$. As before, we may take δ small enough such that the quasi-continuity condition 3 of Definition (8) holds within each $\mathcal{A}_k$. The last step is to show that the set

$$\mathcal{B} = \mathcal{N} \setminus \cup_p \mathcal{A}_p$$

satisfies condition 2 of Definition (8). That is

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

Note that, if distance(emp$[t], \mathcal{B}) < 1/n$, then there must be p such that

$$|D(p|t) - kn| < \bar{q}.$$

Otherwise, no agent would be pivotal in moving the aggregate distribution of types enough to change the equilibrium. However, given our assumption that μ_0 has full support, if t is drawn iid according to μ_0 then for any $p \in V \setminus \{0\}$ $D(p|t)$ follows a multinomial distribution. By standard arguments, the variance of this distribution is of the order of $\sqrt{n}$. Therefore, the probability that $|D(p|t) - kn| < \bar{q}$ for a fixed k is converging to 0 as n grows. For $p = 0$, this probability is also small as we assumed $p^*(m) > 0$.

Finally, we have yet to consider the case where

$$\begin{aligned} p^*(\mu_0) &= 0 \\ D(p^*(\mu_0)|\mu_0) &= k. \end{aligned}$$

This case is quite simple, as for δ small enough the strategies σ_μ with $\mu \in \mathcal{N}$ are very similar to σ_{μ_0}. For any such μ we have $p^*(\mu) = 0$. The strategies σ_μ then only change in the probability of agents bidding $(0, q_i)$ versus $(0, 0)$, which varies continuously with μ, and we omit the details.

C Details for Extensions Section

C.1 Semi-Anonymous Mechanisms

This Section states and proves an extension of Theorem 1 to the case of semi-anonymous mechanisms, referred to in Section 5.1. We must first define the concepts of a limit BNE and SP-L. Both definitions are straightforward generalizations of the anonymous case. 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.

Definition 12. *Given a semi-anonymous mechanism $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$, the strategy $\sigma^*_\mu(\cdot)$ is a limit μ-BNE if, for all $g \in G$, $t_i \in T_g$ and $a'_i \in A_g$:*

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

Definition 13. *Mechanism $\{(\Phi^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ is **strategyproof in the large**, or **SP-L**, if, for all $g \in G$, $t_i, t'_i \in T_g$, and $m \in \bar{\Delta}T$:*

$$u_{t_i}[\phi^\infty(t_i, m)] \geq u_{t_i}[\phi^\infty(t'_i, m)]. \tag{C.1}$$

The definition of continuous and quasi-continuous families of limit equilibria are then identical to the anonymous case. With these definitions, the statement of Theorem 1 for the semi-anonymous case is identical, save for the broader class of mechanisms, to that for the anonymous case.

Theorem 2. *Consider a semi-anonymous mechanism $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$ with a quasi-continuous family of limit equilibria $(\sigma^*_\mu)_{\mu\in\Delta T}$. Then there exists a direct semi-anonymous mechanism $\{(F^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ with the following properties*

1. *$\{(F^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ is strategyproof in the large.*
2. *If $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$ is continuous at the prior μ_0, then in the limit as $n \to \infty$, truthful play of $\{(F^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ and Bayes-Nash equilibrium play of $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$ give agents the same utilities. Formally, given $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$, there exists n_0 such that for all $n > n_0$ and all t_i:*
$$|u_{t_i}[f^n(t_i,\mu_0)] - u_{t_i}[\phi^n(\sigma^*_{\mu_0}(t_i), \sigma^*_{\mu_0}(\mu_0))]| < \epsilon,$$
where $f^n(\cdot)$ is constructed from $F^n_i(\cdot)$ according to Equation (2.3).
3. *If $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$ is not continuous at the true prior μ_0, then in the limit as $n \to \infty$, truthful play of $\{(F^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$ gives agents the same utilities as a convex combination of equilibrium outcomes under $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$, for priors in a neighborhood of μ_0. Formally, for every $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$, there exist priors μ_k with $|\mathrm{emp}[\mu_k] - \mathrm{emp}[\mu_0]| < \epsilon$, and n_0, such that for all $n > n_0$ there exist weights π^n_k summing to one such that, for all t_i:*
$$|u_{t_i}[f^n(t_i,\mu_0)] - \sum_{k=1,\ldots,K} \pi^n_k \cdot u_{t_i}[\phi^n(\sigma^*_{\mu_k}(t_i), \sigma^*_{\mu_k}(\mu_k))]| < \epsilon.$$

Proof. Our proof considers the direct semi-anonymous mechanism $\{(F^n)_{\mathbb{N}}, (T_g)_{g\in G}\}$, constructed from $\{(\Phi^n)_{\mathbb{N}}, (A_g)_{g\in G}\}$, exactly as in the proof of Theorem 1 (cf. 4.1). The sets of types are the same as in the original mechanism.

Part 2 of the Theorem follows from the same argument as in the proof of Theorem 1. This is so as the argument deriving the approximation formulas in the proof of Theorem 1 does not use the fact that agents play strategies restricted to be in the action set of their groups.

To establish the first part of the Theorem, we employ a small modification of the original argument, as now the σ^*_μ are assumed to be limit equilibria of a semi-anonymous mechanism. Given $\epsilon > 0$,

$g \in G$, $t_i, t_i' \in T_g$, and $m \in \bar{\Delta}T$, by the third part of the proof of Theorem 1 there exists n_0 such that for all $n \geq n_0$

$$u_{t_i}[f^n(t_i', m)] - u_{t_i}[f^n(t_i, m)] \quad < \\ \sum_{k=1,\ldots,K} \pi_k^n \cdot \{u_{t_i}[x_k^n(t_i')] - u_{t_i}[x_k^n(t_i)]\} \quad + \epsilon/2$$

where the weights π_k^n sum to 1 and for any $t_i'' \in T_g$

$$x_k^n(t_i'') = \phi^n(\sigma_{\mu_k}^*(t_i''), \sigma_{\mu_k}^*(\mu_k)).$$

Moreover, since $\sigma_{\mu_k}^*$ are limit equilibria, n_0 may be taken such that

$$u_{t_i}[x_k^n(t_i')] - u_{t_i}[x_k^n(t_i)] < \epsilon/2.$$

Therefore,

$$u_{t_i}[f^n(t_i', m)] - u_{t_i}[f^n(t_i, m)] < \epsilon,$$

and the first part of the Theorem follows. □

C.2 Complete Information Nash Equilibria

Quasi-continuity is defined analogously to Definition 8 (the difference is Condition 3):

Definition 14. *Consider a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ with limit $\phi^\infty(\cdot,\cdot)$, and a sequence of families of Complete Information Nash equilibria $(\sigma^n_{\text{emp}[t]})_{t \in T^n, n \in \mathbb{N} \cup \infty}$. The sequence of families of equilibria is **quasi-continuous** if, for every $\mu_0 \in \bar{\Delta}T$ and $\epsilon > 0$, there exists a neighborhood $\mathcal{N}$ of μ_0 such that:*

1. *$\mathcal{N}$ can be decomposed as the union of a finite number of open sets, $\mathcal{A}_1, \ldots \mathcal{A}_K$, and a closed set $\mathcal{B}$. Formally, $\mathcal{N} = \cup_{1 \leq k \leq K} \mathcal{A}_k \cup \mathcal{B}$ with each $\mathcal{A}_k$ open.*

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

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

3. *Within each set $\mathcal{A}_k$, in a large enough market, agents' outcomes are continuous with respect to changes in the strategy that all agents use. Formally, for each $\mathcal{A}_k$, there exists n_0 such that for any $n > n_0$, and any $\text{emp}[t_i, t_{-i}], \text{emp}[t'] \in \mathcal{A}_k$, we have:*

$$|\Phi_i^n(\sigma^n_{\text{emp}[t]}(t_i), \sigma^n_{\text{emp}[t]}(t_{-i})) - \Phi_i^n(\sigma^n_{\text{emp}[t']}(t_i), \sigma^n_{\text{emp}[t']}(t_{-i}))| < \epsilon.$$

The family of equilibria is ***continuous at μ_0*** *if, for the prior μ_0, Conditions 1 and 3 hold with $K = 1$ and $\mathcal{B} = \emptyset$.*

The main theorem statement is analogous to Theorem 1 (note that Part 2 of the statement is much simpler):

Theorem 3. *Suppose that there exists a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ with a quasi-continuous sequence of families of Complete Information Nash equilibria $(\sigma^n_{\text{emp}[t]})_{t \in T^n, n \in \mathbb{N} \cup \infty}$. Then there exists a direct mechanism $\{(F^n)_{\mathbb{N}}, T\}$ with the following properties*

1. *$\{(F^n)_{\mathbb{N}}, T\}$ is strategyproof in the large.*
2. *For any size market n, truthful play of $\{(F^n)_{\mathbb{N}}, T\}$ and Complete Information Nash equilibrium play of $\{(\Phi^n)_{n \in \mathbb{N}}, A\}$ give agents the same utilities.*

As described in the main text, the mechanism constructed to prove Theorem 3 is:

$$F^n(t) = \Phi^n(\sigma^n_{\text{emp}[t]}(t)). \tag{C.2}$$

Part (2) of the Theorem statement is immediate from inspection of (C.2).

For Part (1) of the Theorem statement, fix arbitrary $m \in \bar{\Delta}T$. Initially, suppose that $\{(\Phi^n)_{\mathbb{N}}, A\}$ is continuous at m. Choose arbitrary t_i, t'_i, and $n > n_0$, and then choose arbitrary t_{-i} so that $\text{emp}[t_i, t_{-i}]$ and $\text{emp}[t'_i, t_{-i}]$ are in the neighborhood of m as defined in Definition 14 quasi-continuity. Then, Nash equilibrium implies that

$$u_{t_i}[\Phi_i^n(\sigma_{\text{emp}[t]}(t_i), \sigma_{\text{emp}[t]}(t_{-i}))] \geq u_{t_i}[\Phi_i^n(\sigma_{\text{emp}[t]}(t'_i), \sigma_{\text{emp}[t]}(t_{-i}))] \tag{C.3}$$

and continuity implies that

$$u_{t_i}[\Phi_i^n(\sigma_{\text{emp}[t]}(t'_i), \sigma_{\text{emp}[t]}(t_{-i}))] \geq u_{t_i}[\Phi_i^n(\sigma_{\text{emp}[t'_i, t_{-i}]}(t'_i), \sigma_{[t'_i, t_{-i}]}(t_{-i}))] - \epsilon \tag{C.4}$$

Combining (C.3) and (C.4) yields:

$$u_{t_i}[\Phi_i^n(\sigma_{\text{emp}[t]}(t_i), \sigma_{\text{emp}[t]}(t_{-i}))] \geq u_{t_i}[\Phi_i^n(\sigma_{\text{emp}[t_i', t_{-i}]}(t_i'), \sigma_{[t_i', t_{-i}]}(t_{-i}))] - \epsilon$$

Hence, for realizations of t_{-i} close enough to m, type t_i's gain from misreporting as t_i' is bounded above by ϵ. Additionally, by the law of large numbers, there exists n_0, such that, for $n \geq n_0$, the probability that t_{-i} is close enough to m (i.e., that both $\text{emp}[t_i, t_{-i}]$ and $\text{emp}[t_i', t_{-i}]$ are in the relevant neighborhood) is greater than $1 - \epsilon$. Hence, t_i's total gain from misreporting, if his opponents' reports are distributed according to m, is at most 2ϵ in a large enough market.

To complete the argument that $\{(F^n)_{\mathbb{N}}, T\}$ is SP-L, we need to address the case where $\{(\Phi^n)_{\mathbb{N}}, A\}$ is not continuous at m. In this case, the same argument as given above for the continuous case works within each of the open sets $\mathcal{A}_k$ in the neighborhood of m, as defined in Definition 14. That is, within each set $\mathcal{A}_k$, we can bound t_i's gain from misreporting to 2ϵ in a large enough market. Since, by Step 2 of the proof of Theorem 1, agent i regards the probability that $\text{emp}[t_i, t_{-i}] \in \mathcal{A}_k$ as approximately exogenous in a large enough market, and by Condition (2) of Definition 14 the probability that by misreporting as t_i' he can change which of the sets $\mathcal{A}_k$ or $\mathcal{B}$ the empirical distribution lands in goes to zero as the market grows large, agent i's gain from misreporting is vanishing to zero as the market grows large for the case where $\{(\Phi^n)_{\mathbb{N}}, A\}$ is not continuous at m as well. This completes the argument that $\{(F^n)_{\mathbb{N}}, T\}$ is SP-L.

C.3 Finite Economy Bayes-Nash Equilibria

As described in Section 5.4, the main Theorem could have been stated using exact Nash equilibrium. In this Section we provide a formal definition of quasi-continuity in this setting, state the Theorem, and outline how the proof differs from the proof of Theorem 1. Given a sequence of families of Nash equilibria $(\sigma_\mu^n)_{\mu \in \Delta T, n \in \mathbb{N} \cup \infty}$, we extend the definition of (quasi-)continuity as follows (the difference is Condition 4).

Definition 15. *Consider a mechanism $\{(\Phi^n)_{\mathbb{N}}, A\}$ with limit $\phi^\infty(\cdot,\cdot)$, and a sequence of families of Bayes-Nash equilibria $(\sigma_\mu^n)_{\mu \in \Delta T, n \in \mathbb{N} \cup \infty}$. The sequence of families of equilibria is **quasi-continuous** if, for every $\mu_0 \in \Delta T$ and $\epsilon > 0$, there exists a neighborhood $\mathcal{N}$ of μ_0 such that:*

1. *$\mathcal{N}$ can be decomposed as*

$$\mathcal{N} = \cup_{1 \leq k \leq K} \mathcal{A}_k \cup \mathcal{B}$$

where the $\mathcal{A}_k$ are open sets.

2. $\lim_{n\to\infty} \Pr\{\text{distance}(\text{emp}[t], \mathcal{B}) \leq 1/n | t \in T^n, t \sim iid(\mu_0)\} = 0$, *where $t \sim iid(\mu_0)$ denotes a vector of n types t with each component drawn iid according to μ_0.*

3. *For each set $\mathcal{A}_k$, there exists n_0 such that for any $n > n_0$, any $\mu, \mu', \text{emp}[t_i, t_{-i}], \text{emp}[t_i, t'_{-i}] \in \mathcal{A}_k$ we have:*

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

4. *In addition, for all $\mu \in \Delta T, t_i \in T$*

$$\lim_{n\to\infty} \phi^n(\sigma_\mu^n(t_i), \sigma_\mu^n(\mu)) = \phi^\infty(\sigma_\mu^\infty(t_i), \sigma_\mu^\infty(\mu)).$$

The Theorem statement is as follows.

Theorem (Alternative Statement of Theorem 1)**.** *Consider a mechanism $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ with a quasi-continuous sequence of families of limit equilibria $(\sigma_\mu^n)_{\mu\in\Delta T, n\in\mathbb{N}\cup\{\infty\}}$. Then there exists a direct mechanism $\{(F^n)_\mathbb{N}, T\}$ with the following properties*

1. *$\{(F^n)_\mathbb{N}, T\}$ is strategyproof in the large.*

2. *If $\{(F^n)_\mathbb{N}, T\}$ is continuous at the true prior μ_0, then in the limit as $n \to \infty$, truthful play of $\{(F^n)_\mathbb{N}, T\}$ and Bayes-Nash equilibrium play of $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ give agents the same utilities. Formally, given $\mu_0 \in \Delta T$ and $\epsilon > 0$, there exists n_0 such that for all $n > n_0$ and all t_i:*

$$|u_{t_i}[f^n(t_i, \mu_0)] - u_{t_i}[\phi^n(\sigma_{\mu_0}^n(t_i), \sigma_{\mu_0}^n(\mu_0))]| < \epsilon.$$

3. *If $\{(F^n)_\mathbb{N}, T\}$ is not continuous at μ_0, then in the limit as $n \to \infty$, truthful play of $\{(F^n)_\mathbb{N}, T\}$ gives agents the same utilities as a convex combination of equilibrium outcomes under $\{(\Phi^n)_{n\in\mathbb{N}}, A\}$ and priors in a neighborhood of μ_0. Formally, for every $\mu_0 \in \Delta T$, $\epsilon > 0$ there exist priors μ_k with $|\text{emp}[\mu_k] - \text{emp}[\mu_0]| < \epsilon$, and n_0, such that for all $n > n_0$ there are weights π_k^n summing to one such that, for all t_i:*

$$|u_{t_i}[f^n(t_i, \mu_0)] - \sum_{k=1,\ldots,K} \pi_k^n \cdot u_{t_i}[\phi^n(\sigma_{\mu_k}^n(t_i), \sigma_{\mu_k}^n(\mu_k))]| < \epsilon.$$

Note that, unlike Theorem 1 in the text, the SP-L mechanism we construct in this case is

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

The proof of this Theorem is largely analogous to that of the Theorem for limit Nash equilibria. For conciseness, we will discuss the points where the proofs diverge, and how to adjust the proof of Theorem 1, instead of giving a complete proof. The proof is also based on an approximation Lemma.

The statement of the Lemma differs slightly.

Lemma 4. *Fix a prior μ_0 and $\epsilon > 0$. Let $\mathcal{N}$ be a neighborhood as in Definition 15. Let μ_k be priors $\mu_k \in \mathcal{A}_k$ for each $k = 1, \ldots, K$, with $|\mu_k - \mu_0| < \epsilon$. Then there exists n_0, such that for all $n > n_0$: there exist positive weights π^n_k with $\sum_{1 \le k \le K} \pi^n_k = 1$, such that for all t_i*

$$|f^n(t_i, \mu_0) - \sum_{k=1}^{K} \pi^n_k \cdot z_k(t_i)| < 6\epsilon,$$

where

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

The proof of the alternate Lemma is largely similar to the proof of Lemma 1. In fact, the steps are basically the same, but replacing $\sigma^*$by σ^n or σ^∞ as appropriate. The only step of the proof that differs significantly is deriving the analogue of Inequality (A.5). Mutatis mutandis, this inequality would be showing that we may take n_3 large enough such that

$$|\phi^n(\sigma^n_{\mu_k}(t_i), \sigma^n_{\mu_k}(\mu_k)) - \phi^\infty(\sigma^\infty_{\mu_k}(t_i), \sigma^\infty_{\mu_k}(\mu_k))| < \epsilon. \tag{C.5}$$

This is still true. However, it does not follow from the definition of the limit, as in the proof of Lemma (1). Instead, it is a consequence of Condition 4 in the definition of a quasi-continuous sequence of families of equilibria. The rest of the proof follows straightforwardly, with the modifications we described.

C.4 Proof of Proposition 2 on Ex-Post Robustness

Step 1.

Given $\epsilon > 0$, there exists n_0 such that for all $n, n' \geq n_0, t_i, t_{-i}$ we have

$$|\Phi_i^n(t_i, t_{-i}) - \phi^{n'}(t_i, \operatorname{emp} t_{-i})| < \epsilon.$$

Proof. Let $\hat{\mu} = \operatorname{emp} t_{-i}$. We may write

$$\phi^{n'}(t_i, \hat{\mu}) = \sum_{t'_{-i}} \Pr(t'_{-i} | t'_{-i} \in T^{n'-1}, t'_{-i} \sim \hat{\mu}) \cdot \Phi_i^{n'}(t_i, t'_{-i}). \tag{C.6}$$

By the definition of equicontinuity, we may take $\delta > 0$ such that for all t'_{-i} with

$$|\operatorname{emp} t'_{-i} - \hat{\mu}| < \delta$$

we have

$$|\Phi_i^n(t_i, t_{-i}) - \Phi_i^{n'}(t_i, t'_{-i})| < \epsilon/2. \tag{C.7}$$

Moreover, we may take n_0 such that for all $n \geq n_0$, by the law of large numbers,

$$\sum_{|\operatorname{emp} t'_{-i} - \hat{\mu}| \geq \delta, t'_{-i} \in T^{n'-1}} \Pr(t'_{-i} | t'_{-i} \in T^{n'-1}, t'_{-i} \sim \hat{\mu}) < \epsilon/2. \tag{C.8}$$

Consider now the difference

$$|\Phi_i^n(t_i, t_{-i}) - \phi^{n'}(t_i, \hat{\mu})|.$$

From Equation (C.6), we have that

$$|\Phi_i^n(t_i, t_{-i}) - \phi^{n'}(t_i, \hat{\mu})| = |\Phi_i^n(t_i, t_{-i}) - \sum_{t'_{-i}} \Pr(t'_{-i} | t'_{-i} \in T^{n'-1}, t'_{-i} \sim \hat{\mu}) \cdot \Phi_i^{n'}(t_i, t'_{-i})|.$$

By the triangle inequality we have that

$$\begin{aligned}
&|\Phi_i^n(t_i,t_{-i}) - \phi^{n'}(t_i,\hat{\mu})| \\
\leq & \sum_{|\operatorname{emp} t'_{-i} - \hat{\mu}| < \delta, t'_{-i} \in T^{n'-1}} \Pr(t'_{-i} | t'_{-i} \in T^{n'-1}, t'_{-i} \sim \hat{\mu}) \cdot |\Phi_i^n(t_i,t_{-i}) - \Phi_i^{n'}(t_i,t'_{-i})| \\
+ & \sum_{|\operatorname{emp} t'_{-i} - \hat{\mu}| \geq \delta, t'_{-i} \in T^{n'-1}} \Pr(t'_{-i} | t'_{-i} \in T^{n'-1}, t'_{-i} \sim \hat{\mu}) \cdot |\Phi_i^n(t_i,t_{-i}) - \Phi_i^{n'}(t_i,t'_{-i})|.
\end{aligned}$$

Plugging in Inequalities (C.7) and (C.8) we have

$$\begin{aligned}
|\Phi_i^n(t_i,t_{-i}) - \phi^{n'}(t_i,\hat{\mu})| &< \\
\epsilon/2 + \epsilon/2 &= \epsilon.
\end{aligned}$$

Moreover, note that the above bounds in Inequalities (C.7) and (C.8) may be taken uniform in t_i, t_{-i}. Therefore the overall bound is uniform. This completes this step. □

Step 2.

Given $\epsilon > 0$, there exists n_0 such that for all $n \geq n_0, t_i, t_{-i}$ we have

$$|\Phi_i^n(t_i,t_{-i}) - \phi^{\infty}(t_i, \operatorname{emp} t_{-i})| < \epsilon. \tag{C.9}$$

Proof. By Step 1, we may take n_0 such that for all $n, n' \geq n_0, t_i, t_{-i}$ we have

$$|\Phi_i^n(t_i,t_{-i}) - \phi^{n'}(t_i, \operatorname{emp} t_{-i})| < \epsilon/2.$$

Taking the limit as $n' \to \infty$ we have

$$|\Phi_i^n(t_i,t_{-i}) - \phi^{\infty}(t_i, \operatorname{emp} t_{-i})| \leq \epsilon/2 < \epsilon.$$

□

Step 3.

Use Step 2 to complete the proof of the Proposition.

Proof. By Step 2, we may take n_0 large enough such that for all $n \geq n_0, t$:

$$|\Phi_i^n(t_i, t_{-i}) - \phi^\infty(t_i, \operatorname{emp} t_{-i})| < \epsilon/2. \tag{C.10}$$

Consider now the gain for agent i to, given t_{-i}, perfom an ex post deviation to $\hat{t}_i$. We have

$$\begin{aligned} u_{t_i}[\Phi_i^n(\hat{t}_i, t_{-i})] - u_{t_i}[\Phi_i^n(t)] &\leq |u_{t_i}[\Phi_i^n(\hat{t}_i, t_{-i})] - u_{t_i}[\phi^\infty(\hat{t}_i, \operatorname{emp} t_{-i})]| \\ &+ |u_{t_i}[\phi^\infty(t_i, \operatorname{emp} t_{-i})] - u_{t_i}[\Phi_i^n(t_i, t_{-i})]| \\ &+ (u_{t_i}[\phi^\infty(\hat{t}_i, \operatorname{emp} t_{-i})] - u_{t_i}[\phi^\infty(t_i, \operatorname{emp} t_{-i})]). \end{aligned}$$

By the boundedness of u we have

$$\begin{aligned} u_{t_i}[\Phi_i^n(\hat{t}_i, t_{-i})] - u_{t_i}[\Phi_i^n(t)] &\leq |\Phi_i^n(\hat{t}_i, t_{-i}) - \phi^\infty(\hat{t}_i, \operatorname{emp} t_{-i})| \\ &+ |\phi^\infty(t_i, \operatorname{emp} t_{-i}) - \Phi_i^n(t_i, t_{-i})| \\ &+ (u_{t_i}[\phi^\infty(\hat{t}_i, \operatorname{emp} t_{-i})] - u_{t_i}[\phi^\infty(t_i, \operatorname{emp} t_{-i})]). \end{aligned}$$

Plugging in Inequality (C.10), we have

$$\begin{aligned} u_{t_i}[\Phi_i^n(\hat{t}_i, t_{-i})] - u_{t_i}[\Phi_i^n(t)] &< \epsilon/2 + \epsilon/2 \\ &+ (u_{t_i}[\phi^\infty(\hat{t}_i, \operatorname{emp} t_{-i})] - u_{t_i}[\phi^\infty(t_i, \operatorname{emp} t_{-i})]). \end{aligned}$$

Since the original mechanism is SP-L, we have that the last term is nonpositive. Therefore,

$$u_{t_i}[\Phi_i^n(\hat{t}_i, t_{-i})] - u_{t_i}[\Phi_i^n(t)] < \epsilon.$$

This completes the proof. □

D An Example that is not Quasi-Continuous

This Section gives an example of a sequence of families of BNE that is not quasi-continuous. Moreover, the construction used in the proof of Theorem C.3 does not produce an SP-L mechanism. We consider the case of BNE of the finite mechanism to highilght that, even for finite economy BNE, the construction used to prove our Theorems 1 and C.3 does not work.

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 are fined 10 units. The set of is types $T = O = \{o_1, o_2\}$, 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 $a_i(1)$, 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 $\Phi^n_i(a) = (a_i(1), 0)$. 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:
  $$\Phi^n_i(a) = (a_i(1), -10 \cdot \mathbf{1}\{a_i(1) = o_1\}).$$
  - If $\#\{j : a_j(1) = o_1\}$ is even, then agents asking for object o_2 are fined:
  $$\Phi^n_i(a) = (a_i(1), -10 \cdot \mathbf{1}\{a_i(1) = o_2\}).$$

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 limit 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 sequence of families of BNE is not quasi-continuous at μ_0. For μ and $\operatorname{emp} t$ in any small neighborhood of μ_0, the allocation $\Phi_i^n(t_i, \sigma_\mu^n(t_{-i}))$ varies discontinuously with t_{-i} and μ. To see this formally, take a neighborhood $\mathcal{N}$ of μ_0 small enough such that the set

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

is relatively dense with respect to $\mathcal{N}$. Any open subset $\mathcal{A}_k$ of this neighborhood therefore contains infinite points in this set, and infinite points outside this set. In particular, there exist n, μ and $\operatorname{emp}[t]$ in $\mathcal{A}_k$, where type o_1 agents are not fined, $\Phi_i^n(o_1, \sigma_\mu^n(t_{-i})) = (o_1, 0)$ and n, μ and $\operatorname{emp}[t]$ where they are fined, $\Phi_i^n(o_1, \sigma_\mu^n(t_{-i})) = (o_1, -10)$. Therefore, the conditions for quasi-continuity are not satisfied.

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

$$F^n(t) = \Phi^n(\sigma_{\operatorname{emp}[t]}(t)).$$

This is the construction used in the proof of our main theorems. 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_μ^n 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. 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

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

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, as would be the case if the sequence of families of equilibria were quasi-continuous, by Theorem C.3.

References

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

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

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

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

Al-Najjar, N.I. and R. Smorodinsky, "The Efficiency of Competitive Mechanisms under Private Information," *Journal of Economic Theory*, 2007, *137* (1), 383–403.

Aumann, R.J., "Existence of Competitive Equilibria in Markets with a Continuum of Traders," *Econometrica*, 1964, pp. 1–17.

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

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

Azrieli, Y. and E. Shmaya, "Lipschitz Games," 2011. Mimeo, Ohio State University.

Bergemann, D. and S. Morris, "Robust Mechanism Design," *Econometrica*, 2005, *73* (6), 1771–1813.

Bodoh-Creed, A., "Approximation of Large Games," 2010. Mimeo, Cornell University.

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

Budish, E., "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes," *Journal of Political Economy*, Forthcoming.

_ **and E. Cantillon**, "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," *American Economic Review*, Forthcoming.

_ **, Y.K. Che, F. Kojima, and P. Milgrom**, "Implementing Random Assignments: A Generalization of the Birkhoff-Von Neumann Theorem," 2011. Mimeo, Stanford University.

Carmona, G. and K. Podczeck, "Ex-Post Stability of Bayes-Nash Equilibria of Large Games," *Games and Economic Behavior*, 2011.

Carroll, G., "A Quantitative Approach to Incentives: Application to Voting Rules," *Mimeo, MIT*, 2011.

Che, Y.K. and F. Kojima, "Asymptotic equivalence of probabilistic serial and random priority mechanisms," *Econometrica*, 2010, *78* (5), 1625–1672.

Cramton, P. and B.E. Katzman, "Designed to Fail: The Medicare Auction for Durable Medical Equipment," 2010. Mimeo, University of Maryland.

Cremer, J. and R.P. McLean, "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 1988, pp. 1247–1257.

Cripps, M.W. and J.M. Swinkels, "Efficiency of Large Double Auctions," *Econometrica*, 2006, *74* (1), 47–92.

Day, R. and P. Milgrom, "Core-Selecting Package Auctions," *International Journal of Game Theory*, 2008, *36* (3), 393–407.

Deb, J. and E. Kalai, "Stability in Large Bayesian Games with Heterogeneous Players," 2011. Mimeo, Northwestern University.

Debreu, G. and H. Scarf, "A Limit Theorem on the Core of an Economy," *International Economic Review*, 1963, *4* (3), 235–246.

Edelman, B. and M. Ostrovsky, "Strategic Bidder Behavior in Sponsored Search Auctions," *Decision support systems*, 2007, *43* (1), 192–198.

_ , _ , and M. Schwarz, "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *The American Economic Review*, 2007, *97* (1), 242–259.

Engelbrecht-Wiggans, R., J.A. List, and D.H. Reiley, "Demand Reduction in Multi-Unit Auctions with Varying Numbers of Bidders: Theory and Field Experiments," *International Economic Review*, 2006, *47* (1), 203–231.

Featherstone, C.R., "Rank Eciency: Investigating a Widespread Ordinal Welfare Criterion," *Mimeo, Harvard Business School*, 2011.

_ and M. Niederle, "School choice mechanisms under incomplete information: an experimental investigation," *Mimeo, Harvard Business School*, 2011.

Friedman, M., "Comment on" Collusion in the Auction Market for Treasury Bills"," *Journal of Political Economy*, 1964, *72.*

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

Fudenberg, D. and J. Tirole, *Game Theory*, MIT Press, 1991.

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

Gershkov, A., B. Moldovanu, and X. Shi, "Bayesian and Dominant Strategy Implementation Revisited," 2011. Mimeo, University of Toronto.

Gibbard, A., "Manipulation of Voting Schemes: A General Result," *Econometrica:*, 1973, pp. 587–601.

Green, J. and J.J. Laffont, "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica:*, 1977, pp. 427–438.

Hurwicz, L., "On Informationally Decentralized Systems," *Decision and organization: A volume in honor of Jacob Marschak*, 1972, *12*, 297.

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

Immorlica, N. and M. Mahdian, "Marriage, Honesty, and Stability," in "Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms" Society for Industrial and Applied Mathematics 2005, pp. 53–62.

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

Jegadeesh, N., "Treasury Auction Bids and the Salomon Squeeze," *Journal of Finance*, 1993, pp. 1403–1419.

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

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

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

Kovalenkov, A., "Simple Strategy-Proof Approximately Walrasian Mechanisms," *Journal of Economic Theory*, 2002, *103* (2), 475–487.

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

Lee, S.M., "Incentive Compatibility of Large Centralized Matching Markets," 2011. Mimeo, Caltech.

Liu, Q. and M. Pycia, "Ordinal Efficiency, Fairness, and Incentives in Large Markets," *Mimeo, UCLA*, 2011.

Manelli, A.M. and D.R. Vincent, "Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model," *Econometrica*, 2010, *78* (6), 1905–1938.

Merlob, B., K. Peters, C.R. Plott, A. Pradhana, and Y. Zhang, "An Evaluation of the Proposed Procurement Auction for the Purchase of Medicare Equipment: Experimental Tests of the Auction Architecture," 2010. Mimeo, Caltech.

Milgrom, P., "Critical Issues in the Practice of Market Design," *Economic Inquiry*, 2011.

Miralles, A., "School Choice: The Case for the Boston Mechanism," *Auctions, Market Mechanisms and Their Applications*, 2009, pp. 58–60.

Myerson, R.B., "Incentive Compatibility and the Bargaining Problem," *Econometrica: journal of the Econometric Society*, 1979, pp. 61–73.

Papai, S., "Strategyproof and Nonbossy Multiple Assignments," *Journal of Public Economic Theory*, 2001, *3* (3), 257–271.

Parkes, D.C., J. Kalagnanam, and M. Eso, "Achieving Budget-Balance with Vickrey-Based Payment Schemes in Exchanges," in "International Joint Conference on Artificial Intelligence," Vol. 17 Citeseer 2001, pp. 1161–1168.

Pathak, P.A. and T. Sönmez, "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *The American Economic Review*, 2008, *98* (4), 1636–1652.

_ **and** _ , "School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation," 2011. NBER Working Paper No. 16783.

Pesendorfer, W. and J.M. Swinkels, "Efficiency and Information Aggregation in Auctions," *American Economic Review*, 2000, pp. 499–525.

Reny, P.J. and M. Perry, "Toward a Strategic Foundation for Rational Expectations Equilibrium," *Econometrica*, 2006, *74* (5), 1231–1269.

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

Roth, A.E., "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 1982, pp. 617–628.

_ , "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica*, 2002, *70* (4), 1341–1378.

_ , "What Have We Learned from Market Design?*," *The Economic Journal*, 2008, *118* (527), 285–310.

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

_ **and U.G. Rothblum**, "Truncation Strategies in Matching Markets in Search of Advice for Participants," *Econometrica*, 1999, *67* (1), 21–43.

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

Satterthwaite, M.A., "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 1975, *10* (2), 187–217.

Segal, I., "Optimal Pricing Mechanisms with Unknown Demand," *The American economic review*, 2003, *93* (3), 509–529.

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

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

Wilson, R., "Game-Theoretic Analyses of Trading Processes,," in "Advances in Economic Theory: Fifth World Congress," Cambridge, UK: Cambridge University Press, 1987, chapter 2, pp. 33–70.