NOTES AND COMMENTS

ON THE GENERIC (IM)POSSIBILITY OF FULL SURPLUS EXTRACTION IN MECHANISM DESIGN

By Aviad Heifetz and Zvika Neeman[1]

A number of studies, most notably Crémer and McLean (1985, 1988), have shown that in generic type spaces that admit a common prior and are of a fixed finite size, an uninformed seller can design mechanisms that extract all the surplus from privately informed bidders. We show that this result hinges on the nonconvexity of such a family of priors. When the ambient family of priors *is* convex, generic priors do *not* allow for full surplus extraction provided that for at least one prior in this family, players' beliefs about other players' types do not pin down the players' own preferences. In particular, full surplus extraction is generically impossible in finite type spaces with a common prior. Similarly, generic priors on the universal type space do not allow for full surplus extraction.

Keywords: Surplus extraction, information rents, universal type space, genericity, prevalence, shyness, face.

1. INTRODUCTION

Does relevant private information necessarily confer a positive economic rent to those who possess it? Surprisingly, the answer given by the literature to this question is negative. A number of studies, including, most notably, Crémer and McLean (1985, 1988), have shown that under standard assumptions—the existence of a common prior, a fixed finite number of types, risk neutrality, and no limited liability—an uninformed principal facing privately informed players can generically implement any decision rule he could implement were that private information accessible to him. An uninformed seller, for example, is generically able to extract the full surplus of any number of privately informed bidders in an auction. As these "full-surplus-extraction" results imply that the players' private information is (generically) irrelevant, they have been said to "cast doubt on the value of the current mechanism design paradigm as a model of institutional design" (McAfee and Reny (1992, p. 400)).

Since full-surplus-extraction results make heavy use of the assumption that the type spaces are of a fixed finite size, it is natural to ask how crucial this assumption is for obtaining these results. This assumption is problematic because there is no a priori finite bound on the number of types needed for modeling

a situation involving asymmetric information. Indeed, the universal type space (Mertens and Zamir (1985)) that embeds all such models has a continuum of types, and its subspaces that admit a common prior can have an arbitrarily large number of types. If any of the priors in some relevant family of priors $\mathcal{P}$ could just as well serve as a plausible model of a situation involving asymmetric information, would full surplus extraction "typically" be possible in $\mathcal{P}$? This is the question addressed in this paper.

The starting point for our argument is Neeman's (2004) observation that full surplus extraction is possible only if the type space has the "beliefs-determine-preferences" (henceforth, BDP) property, which requires that almost every possible belief of every player about other players' types pins down the player's own preferences. As we show, a nondegenerate convex combination of a BDP prior and a non-BDP prior yields a non-BDP prior. This implies that the collection of priors that permit full surplus extraction (henceforth, FSE priors) is "small" provided that the ambient family of priors $\mathcal{P}$ is convex and contains at least one non-BDP prior.

"Smallness" is established in both a "geometric" and a "measure-theoretic" sense. For the geometric perspective, we show that if $\mathcal{P}$ is convex and contains at least one non-BDP prior, then the subset of FSE priors is contained in a proper face of $\mathcal{P}$. Furthermore, if $\mathcal{P}$ is the set of all priors on finite type spaces or the entire collection of priors on the universal type space, then the proper face containing the FSE priors has an infinite codimension in $\mathcal{P}$. For the measure-theoretic perspective, we show that the set of FSE priors is *shy* in such a $\mathcal{P}$. Shyness is a notion of smallness for convex subsets of infinite-dimensional topological vector spaces (in our case, the set of common priors) that generalizes the notion of Lebesgue measure zero in finite-dimensional spaces. The result applies both in case $\mathcal{P}$ is the collection of all priors on the universal type space and in case it is the collection of all such priors with a finite support.

This paper makes a contribution to the growing literature on robust mechanism design that has stemmed from Robert Wilson's view that further progress in game theory depends "on successive reduction in the base of common knowledge required to conduct useful analyses of practical problems" (Wilson (1987)). As shown by Neeman (2004), full surplus extraction hinges on there being common belief that a player's belief pins down the player's preferences. Once this assumption is relaxed, the full surplus of the players cannot be extracted. The argument presented in this paper describes the conditions under which this is generically the case.

The rest of the paper proceeds as follows. For simplicity, instead of considering surplus extraction in a general mechanism design problem with interdependent types, we confine our attention to the classic problem of the design of a revenue maximizing auction in a private-values setting (the general case is treated in Heifetz and Neeman (2005)). Section 2 provides the required definitions. Section 3 is devoted to the statement and derivation of the results. Section 4 concludes with a discussion. Several more technical definitions and proofs are relegated to the Appendix.

2. SURPLUS EXTRACTION IN SINGLE OBJECT AUCTIONS WITH PRIVATE VALUES

We consider the problem of a seller who wishes to design a revenue maximizing auction of a single object. The seller faces n risk neutral bidders who each have a privately known (private) valuation for the object. The value of the object for the seller is normalized to zero. Each bidder may refuse to participate in the seller's auction, but if she agrees to participate, then she is bound by the auction's outcome.

Let $N = \{1, \dots, n\}$ denote the set of bidders or players. Bidder i's, $i \in N$, valuation, or willingness to pay for the object, is denoted by $v_i \in V_i$. The set of bidder i's valuations V_i, $i \in N$, is assumed to be a complete, separable, metric space (in particular, V_i may be finite). The payoff to a bidder with valuation v_i who wins the object with probability q and who pays an expected amount m is given by $q \cdot v_i - m$. We refer to v_i as bidder i's *preference* or *preference type*. Let $V \equiv V_1 \times \cdots \times V_n$. The set V is the *basic space of uncertainty* for this problem.

2.1. Type Spaces

Bidder i's private information is captured by its type $\theta_i \in \Theta_i$. The sets of bidders' types Θ_i, $i \in N$, are assumed to be complete, separable metric spaces. For every measurable space X, let $\Delta(X)$ denote the space of probability measures over X. Each type $\theta_i \in \Theta_i$ is associated with a *preference type* $\hat{v}_i(\theta_i) \in V_i$ that describes θ_i's willingness to pay for the object, and with a *belief type* $\hat{b}_i(\theta_i) \in \Delta(\Theta_{-i})$ that is a probability measure on the space of other bidders' types $\Theta_{-i} \equiv \prod_{j \neq i} \Theta_j$. The space of probability measures $\Delta(\Theta_{-i})$ is endowed with the topology of weak convergence.

We assume that distinct types $\theta_i \neq \theta_i'$ of a given bidder i differ either by their preference type or by their belief type. Each type of each bidder is assumed to know its own willingness to pay for the object and its own beliefs. Since we focus our attention in this paper on a private-values model, each type θ_i's preference type $\hat{v}_i(\theta_i)$ is defined independently of θ_i's belief type $\hat{b}_i(\theta_i)$. This assumption is relaxed in Heifetz and Neeman (2005).[2]

A product space $\Theta \equiv \prod_{i \in N} \Theta_i$ of the players' type spaces is called a private-values type space. Each profile of types $\theta \in \Theta$ is called a state of the world.

2.2. Priors

A probability measure p_i on a private-values type space $\Theta = \prod_{i \in N} \Theta_i$ is called a prior for bidder i if bidder i's belief types $\hat{b}_i(\theta_i)$ are the posteriors

---

[2]The assumption that each type knows its own belief is captured by defining $\hat{b}_i(\theta_i)$ as a probability measure over Θ_{-i} rather than over $\Theta_i \times \Theta_{-i}$. The implied presumption about the bidders' introspective ability is standard and is maintained throughout the paper.

of p_i or if, roughly, $p_i(\cdot|\theta_i.) = \hat{b}_i(\theta_i)$ for p_i-a.e. $\theta_i \in \Theta_i$. Formally, p_i is a prior for bidder i if for every bounded real-valued measurable function $\varphi : \Theta \to \mathbb{R}$,

$$\int_{\Theta_i} \left(\int_{\Theta_{-i}} \varphi(\theta_i, \widetilde{\theta}_{-i}) \, d\hat{b}_i(\theta_i)(\widetilde{\theta}_{-i}) \right) dp_{i|\Theta_i}(\theta_i) = \int_{\Theta} \varphi(\theta) \, dp_i(\theta),$$

where $p_{i|\Theta_i}$ is the marginal of p_i on Θ_i.

A probability measure p on Θ is called a common prior, or prior for short, if it is a prior for every bidder $i \in N$.

For a given collection of type spaces that is closed under finite unions, the set of bidder i's priors on these type spaces is convex: If $p'_i \in \Delta(\Theta')$ and $p''_i \in \Delta(\Theta'')$ are two priors for bidder i, then so is $\alpha p'_i + (1 - \alpha) p''_i \in \Delta(\Theta' \cup \Theta'')$ for every $\alpha \in [0, 1]$. It follows that the set of common priors on any such collection of type spaces is also convex.

Let $\mathcal{P}$ denote a convex family of priors on such a collection of type spaces (equivalently, $\mathcal{P}$ can also denote a convex family of priors on the union of these type spaces). We henceforth refer to $\mathcal{P}$ as the ambient family of models under consideration, with respect to which genericity is to be established. It is natural to assume that $\mathcal{P}$ in indeed convex. If it is conceivable that the seller, who is uninformed, could potentially hold either the belief p' about the bidders' preferences and beliefs or the belief p'', then it is also conceivable that it might hold a belief that is a mixture of p' and p''.

2.3. *BDP Priors*

DEFINITION 1: A prior $p \in \Delta(\Theta)$ satisfies the *beliefs-determine-preferences* property for bidder $i \in N$ if there exists a measurable subset $\Theta_i^p \subseteq \Theta_i$ such that the marginal $p_{|\Theta_i}$ of p on Θ_i assigns probability 1 to Θ_i^p or $p_{|\Theta_i}(\Theta_i^p) = 1$, and no pair of distinct types $\theta_i \neq \theta'_i$ in Θ_i^p hold the same beliefs, i.e., $\hat{b}_i(\theta_i) \neq \hat{b}_i(\theta'_i)$ for every two different types $\theta_i, \theta'_i \in \Theta_i^p$.

This notion of beliefs determine preferences generalizes the one in Neeman (2004). A prior p that satisfies the beliefs-determine-preferences property for bidder i is called a BDP prior for bidder i. A prior p that is a BDP prior for every bidder $i \in N$ is called a BDP prior. Notice that BDP is a property of a prior, *not of a player's type or a state of the world*.

Since any pair of distinct types $\theta_i \neq \theta'_i$ in a private-values type space Θ differ either by belief type or by preference type, there is no pair of distinct types in Θ_i^p that hold identical beliefs but different preferences. If p is a BDP prior, then it follows that for a type in Θ_i^p, knowledge of the type's beliefs "pins down" or implies knowledge of the type's preferences.

The next proposition describes a property of BDP priors that is useful for proving our two main results.

PROPOSITION 1: *Let Θ', Θ'' be two type spaces and $\Theta = \Theta' \cup \Theta''$. A nondegenerate convex combination $p = \alpha p' + (1-\alpha) p'' \in \Delta(\Theta)$ of two priors $p' \in \Delta(\Theta')$ and $p'' \in \Delta(\Theta'')$ is BDP if and only if both p' and p'' are BDP. In particular, a nondegenerate convex combination of a BDP prior and a non-BDP prior (or of two non-BDP priors) is a non-BDP prior.*

For example, two different common priors p' and p'', which are represented by the matrices below (where the entries in each matrix are positive and sum to one, and either $v_1 \neq \tilde{v}_1$ or $v_2 \neq \tilde{v}_2$), are BDP if and only if $\frac{a'}{c'} \neq \frac{b'}{d'}$ and $\frac{a''}{c''} \neq \frac{b''}{d''}$, respectively:

(1)

p'	$\theta_2' = (v_2, b_2')$	$\widetilde{\theta}_2' = (\tilde{v}_2, \tilde{b}_2')$
$\theta_1' = (v_1, b_1')$	a'	b'
$\widetilde{\theta}_1' = (\tilde{v}_1, \tilde{b}_1')$	c'	d'

p''	$\theta_2'' = (v_2, b_2'')$	$\widetilde{\theta}_2'' = (\tilde{v}_2, \tilde{b}_2'')$
$\theta_1'' = (v_1, b_1'')$	a''	b''
$\widetilde{\theta}_1'' = (\tilde{v}_1, \tilde{b}_1'')$	c''	d''

A nondegenerate convex combination of these two priors, $p = \alpha p' + (1-\alpha) p''$, which is represented by the matrix below, is BDP if and only if both p' and p'' are BDP:

(2)

p	$\theta_2 = (v_2, b_2')$	$\widetilde{\theta}_2 = (\tilde{v}_2, \tilde{b}_2')$	$\theta_2' = (v_2, b_2'')$	$\widetilde{\theta}_2' = (\tilde{v}_2, \tilde{b}_2'')$
$\theta_1 = (v_1, b_1')$	$\alpha a'$	$\alpha b'$	0	0
$\widetilde{\theta}_1 = (\tilde{v}_1, \tilde{b}_1')$	$\alpha c'$	$\alpha d'$	0	0
$\theta_1' = (v_1, b_1'')$	0	0	$(1-\alpha) a''$	$(1-\alpha) b''$
$\widetilde{\theta}_1' = (\tilde{v}_1, \tilde{b}_1'')$	0	0	$(1-\alpha) c''$	$(1-\alpha) d''$

The proof of Proposition 1, which is relegated to the Appendix, has two parts. The more straightforward part consists of showing that a convex combination of a BDP and a non-BDP prior is non-BDP. The more delicate part consists of showing that a convex combination of two BDP priors is BDP.

2.4. *Full Surplus Extraction*

By the revelation principle, no loss of generality is implied by assuming that the seller employs an incentive compatible and individually rational "direct-revelation" auction mechanism $\langle q_i : \Theta \to [0, 1], m_i : \Theta \to [0, 1] \rangle_{i \in N}$ in which each bidder i is asked to report its type $\theta_i \in \Theta_i$ and then to participate in a lottery in which he or she pays an amount $m_i(\theta)$, and wins the object with probability $q_i(\theta)$.

A mechanism $\langle q_i, m_i \rangle_{i \in N}$ is incentive-compatible if every type $\theta_i \in \Theta_i$ of every bidder $i \in N$ maximizes its expected payoff by truthfully reporting its type or if, for every $\theta_i \in \Theta_i$,

$$\int_{\Theta_{-i}} \left(q_i(\theta_i, \widetilde{\theta}_{-i}) \hat{v}_i(\theta_i) - m_i(\theta_i, \widetilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i)(\widetilde{\theta}_{-i})$$

$$\geq \int_{\Theta_{-i}} \left(q_i(\theta_i', \widetilde{\theta}_{-i}) \hat{v}_i(\theta_i) - m_i(\theta_i', \widetilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i)(\widetilde{\theta}_{-i})$$

for every $\theta_i' \in \Theta_i$.

A mechanism $\langle q_i, m_i \rangle_{i \in N}$ is individually rational if every type $\theta_i \in \Theta_i$ of every bidder $i \in N$ prefers to participate in the mechanism rather than to opt out or if, for every $\theta_i \in \Theta_i$,

$$\int_{\Theta_{-i}} \left(q_i(\theta_i, \widetilde{\theta}_{-i}) \hat{v}_i(\theta_i) - m_i(\theta_i, \widetilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i)(\widetilde{\theta}_{-i}) \geq 0.$$

DEFINITION 2: A prior p permits the *full surplus extraction* from a set $K \subseteq N$ of bidders if there exists an incentive-compatible and individually rational mechanism $\langle q_i, m_i \rangle_{i \in N}$ that generates an expected payment to the seller that is equal to the full surplus generated by the bidders in K, i.e.,

$$\sum_{i \in K} \int_{\Theta} m_i(\theta) \, dp(\theta) = \int_{\Theta} \max_{i \in K} \{ \hat{v}_i(\theta_i) \} \, dp(\theta).$$

A prior that permits the full surplus extraction from the K bidders is called a full-surplus-extraction prior for K.

We show that BDP is necessary for full surplus extraction. Specifically, we show that if a prior p permits the extraction of bidder i's full surplus, then p is a BDP prior for player i.

PROPOSITION 2: *A prior p that is a FSE prior for bidder i is a BDP prior for bidder i.*

PROOF: Suppose that p is a FSE prior for bidder i. Let $\langle q_i, m_i \rangle_{i \in N}$ be an incentive-compatible and individually rational mechanism that extracts the full surplus of bidder i. Observe that this implies that bidder i must win the object with p-probability 1 under the mechanism $\langle q_i, m_i \rangle_{i \in N}$ and that bidder i's individual rationality constraint must be binding with p-probability 1 under the mechanism $\langle q_i, m_i \rangle_{i \in N}$.

Suppose that p is not a BDP prior for bidder i. It follows that there exist two disjoint measurable subsets of bidder i's types, $A_i, A_i' \subseteq \Theta_i$, that each have a positive p-probability

$$p_{|\Theta_i}(A_i) > 0, \quad p_{|\Theta_i}(A_i') > 0,$$

and the same range of beliefs

$$\hat{b}_i(A_i) = \hat{b}_i(A_i') \subseteq \Delta(\Theta_{-i}),$$

but different valuations. That is, if $\theta_i \in A_i$ and $\theta_i' \in A_i'$ are such that

$$\hat{b}_i(\theta_i') = \hat{b}_i(\theta_i),$$

then

$$\hat{v}_i(\theta_i') < \hat{v}_i(\theta_i).$$

In particular, for every type $\theta_i \in A_i$ there exists a type $\theta_i' \in A_i'$ such that $\hat{b}_i(\theta_i') = \hat{b}_i(\theta_i)$ but $\hat{v}_i(\theta_i') < \hat{v}_i(\theta_i)$. It follows that

$$\int_{\Theta_{-i}} \left(q_i(\theta_i, \tilde{\theta}_{-i}) \hat{v}_i(\theta_i) - m_i(\theta_i, \tilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i)(\tilde{\theta}_{-i})$$

$$\geq \int_{\Theta_{-i}} \left(q_i(\theta_i', \tilde{\theta}_{-i}) \hat{v}_i(\theta_i) - m_i(\theta_i', \tilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i)(\tilde{\theta}_{-i})$$

$$= \int_{\Theta_{-i}} \left(q_i(\theta_i', \tilde{\theta}_{-i}) \hat{v}_i(\theta_i) - m_i(\theta_i', \tilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i')(\tilde{\theta}_{-i})$$

$$> \int_{\Theta_{-i}} \left(q_i(\theta_i', \tilde{\theta}_{-i}) \hat{v}_i(\theta_i') - m_i(\theta_i', \tilde{\theta}_{-i}) \right) d\hat{b}_i(\theta_i')(\tilde{\theta}_{-i})$$

$$\geq 0.$$

The first inequality follows from the incentive compatibility constraint for type θ_i; the following equality follows from the fact that $\hat{b}_i(\theta_i') = \hat{b}_i(\theta_i)$; the next strict inequality follows from the fact that $\hat{v}_i(\theta_i') < \hat{v}_i(\theta_i)$ and that $q_i(\theta_i', \tilde{\theta}_{-i}) = 1$ for p-almost every type $\theta_i' \in A_i'$; and the last inequality follows from the individual rationality constraint for type θ_i'. It therefore follows that bidder i's individual rationality constraint is not binding for p-almost every type $\theta_i \in A_i$. A contradiction.   *Q.E.D.*

3. THE SET OF FSE PRIORS IS SMALL

In this section we show that within a convex family $\mathcal{P}$ of priors that contains at least one non-BDP prior (henceforth, NBDP prior) for bidder i, the subset $\mathcal{F}$ of FSE priors for bidder i is small in two different senses. The first sense is geometric: The set of FSE priors is contained in a proper face of the convex body of priors $\mathcal{P}$. The second sense is measure-theoretic: The set $\mathcal{F}$ of FSE priors is finitely shy in $\mathcal{P}$.

3.1. *The Set of FSE Priors Is Small in a Geometric Sense*

DEFINITION 3—Rockafellar (1970): A convex subset F of a convex set C is called a *face* if whenever $f \in F$ is a convex combination of $x, y \in C$, then $x, y \in F$.

DEFINITION 4: A face F of C is called a proper face if F is a proper subset of C. If C is a convex subset of a vector space X, then the codimension of F in C is the dimension of the minimal subspace Y of X such that C is contained in the subspace spanned by F and Y.

THEOREM 1: *Let P be a convex family of priors that includes at least one NBDP prior for bidder i. Then the subset B of BDP priors for bidder i is a proper face of P.*

PROOF: To show that $\mathcal{B}$ is a proper face of $\mathcal{P}$ we have to show (i) that the set $\mathcal{B}$ is convex, (ii) that if a nondegenerate convex combination $p = \alpha p' + (1 - \alpha) p''$ belongs to $\mathcal{B}$, then so do p' and p'', and (iii) that $\mathcal{B}$ is a proper subset of the set of priors $\mathcal{P}$.

Statement (i) follows directly from Proposition 1. The contrapositive of (ii) is "if either p' or p'' is NBDP, then so is $p = \alpha p' + (1 - \alpha) p''$ provided $\alpha \in (0, 1)$." This also follows directly from Proposition 1. Finally, (iii) follows from the fact that $\mathcal{P}$ contains a NBDP prior for bidder i.   Q.E.D.

COROLLARY 1: *Let $\mathcal{P}$ be a convex family of priors that includes at least one NBDP prior for bidder i. Then the subset $\mathcal{F}$ of full-surplus-extraction priors for bidder i is contained in a proper face of $\mathcal{P}$.*

PROOF: The proof follows immediately from Theorem 1 and Proposition 2.   Q.E.D.

REMARK 1: In particular, Corollary 1 applies to two important families of priors:

— The family of priors $\mathcal{P}_u$ on the universal type space. The universal type space is the type space into which any other type space can be mapped in a beliefs-preserving way. The fact that each bidder is assumed to know its own valuation of the object implies that the universal type space here is a special case of the standard universal type space analyzed in Mertens and Zamir (1985), Brandenburger and Dekel (1993), and Heifetz (1993). For the sake of completeness, we describe in the Appendix the properties of this private-values universal space, which we denote by T.[3]

---

[3]Ely and Peski (2006) have recently suggested that the basic space of uncertainty in the construction of the universal space should be larger, and consist not only of payoffs (as reflected by

— The family of priors $\mathcal{P}_f$ on finite type spaces. Corollary 1 thus implies that consideration of the *convex* family of all priors on finite type spaces leads to a reversal of Crémer and McLean's (1985, 1988) result, which was obtained for the (nonconvex) family of priors on type spaces with a some *pre-given* finite number of types $n_i \geq 2$ for each bidder i. See Section 4.1 for additional discussion of the relationship of our results to those of Crémer and McLean.

In fact, for the two families of priors $\mathcal{P}_u$ and $\mathcal{P}_f$ mentioned in Remark 1, Corollary 1 can be strengthened as follows:

PROPOSITION 3: *For $\mathcal{P}_u$ and $\mathcal{P}_f$, the subset of full-surplus-extraction priors for bidder i is contained in a proper face of infinite codimension in $\mathcal{P}_u$ and $\mathcal{P}_f$, respectively.*

PROOF: Given Proposition 1 and Theorem 1, it remains to show that the codimension of the set of BDP priors $\mathcal{B}$ in $\mathcal{P}_u$ and $\mathcal{P}_f$ is infinite. This follows from the fact that there are infinitely many (in fact, a continuum of) finite-support NBDP priors that are not convex combinations of other priors.

To see this, consider two different bidders i, j, and two distinct valuations for each, $v_i \neq v_i' \in V_i$ and $v_j \neq v_j' \in V_j$. There is a continuum of priors $p_{r,s}$ with $r, s \in (0, 1)$ that are each described by the matrix

$p_{r,s}$	v_j	v_j'
v_i	rs	$r(1-s)$
v_i'	$(1-r)s$	$(1-r)(1-s)$

Each prior $p_{r,s}$ is such that with probability 1 each bidder has the same belief about the other bidder's types irrespective of its own valuation, so $p_{r,s}$ is not a BDP prior. (If there are more bidders, then the definition of $p_{r,s}$ can be extended by choosing some particular valuation for each of those extra bidders, and having $p_{r,s}$ assign probability 1 to that combination of valuations for each of the four combinations of valuations of i and j.) Moreover, each prior $p_{r,s}$ is not a convex combination of other common priors in $\mathcal{P}$, because if $(r, s) \neq (r', s')$, then the priors $p_{r,s}$ and $p_{r',s'}$ on the universal space T have disjoint supports. *Q.E.D.*

REMARK 2: In a finite-dimensional space, a proper face of a convex set is nowhere dense (namely, its closure has an empty interior). This is not necessarily the case in infinite-dimensional spaces. In particular, if the space of

---

valuations in our auction setting), but also of conditional beliefs about payoffs. This extension captures bidders' beliefs about correlations across their types, which may affect the range of implementable outcomes. Corollary 1 applies just as well to the family of priors on the universal type space in Ely and Peski's construction.

finite-support priors $\mathcal{P}_{fu}$ on the universal space is equipped with the topology of weak convergence, then both the subset of finite-support BDP priors *and* its complement, the subset of finite-support NBDP priors, are *dense* in $\mathcal{P}_{fu}$. Because $\mathcal{P}_{fu}$ is dense in the space of all priors $\mathcal{P}_u$ on the universal space (Mertens Sorin, and Zamir (1994, p. 156)), it follows that both the sets of BDP and of NBDP priors are also dense in $\mathcal{P}_u$. In particular, neither set is open and dense in $\mathcal{P}_u$.

3.2. *The Set of FSE Priors Is Shy*

A natural definition of genericity is that of full Lebesgue measure. Unfortunately, there is no direct analogue for Lebesgue measure in infinite-dimensional spaces. Unlike the Lebesgue measure in a finite-dimensional Euclidean space $\mathbb{R}^k$, which is spread uniformly across the space, in infinite-dimensional spaces there does not exist any (countably additive) translation invariant measure. For example, in an infinite-dimensional separable Banach space, any open ball of radius $r > 0$ contains an infinite sequence of disjoint open balls of radius $\frac{r}{4}$, so if a translation-invariant measure were to assign a positive measure to these balls, then the r ball would be assigned an infinite measure for any $r > 0$.[4] Therefore, in infinite-dimensional spaces, probabilities or measures are not satisfactory devices for determining whether events are "typical."

Recently, a general notion of largeness, which coincides with full Lebesgue measure in finite-dimensional spaces, has been proposed. An event E in a finite-dimensional Euclidean space $\mathbb{R}^k$ has Lebesgue measure zero if and only if there exists a positive measure μ on $\mathbb{R}^k$ such that E and all its translations $\{x + y : x \in E\}$, $y \in \mathbb{R}^k$, have μ-measure zero. Christensen (1974) and Hunt, Sauer, and Yorke (1992) have relied on this observation and defined a Borel subset of a complete metric topological vector space to be *shy* if there exists a positive measure μ on the space such that the set and all its translations have μ-measure zero.[5] They called the complement of a shy set *prevalent*. They showed that shy sets satisfy the properties one would expect "small" or "negligible" events to satisfy. In particular, a subset of a shy set is shy, every translation of a shy set is shy, a countable union of shy sets is shy, and no open set is shy.

Anderson and Zame (2001) have adapted the Christensen (1974) and Hunt, Saver, and Yorke (1992) definition to the case in which the relevant parameter

---

[4]Furthermore, confining attention to full-support *quasi-invariant* measures, which preserve null sets under translations (such as the Gaussian measures on the Euclidean spaces), is unhelpful either. Under fairly general conditions, it can be shown that if there does not exist a nontrivial full-support invariant measure on an infinite-dimensional space, then neither does there exist such a quasi-invariant measure (see, e.g., Yamasaki (1985)).

[5]For more on prevalence, see the recent survey by Ott and Yorke (2005).

set is a convex subset C of a topological vector space X. Since we are interested in determining the genericity of the set of FSE priors within a convex family of priors, this is the definition we employ.

It turns out that for our purposes it is not necessary to use Anderson and Zame's general definition of shyness, but rather a simpler and stronger notion called finite shyness. Let λ_H denote the Lebesgue measure on a finite-dimensional subspace $H \subseteq X$.

DEFINITION 5—Anderson and Zame (2001): Let C be a completely metrizable convex subset of the topological vector space X. A universally measurable[6] subset $E \subseteq C$ is *finitely shy* in $C \subseteq X$ if there exists a finite-dimensional subspace $H \subseteq X$ such that $\lambda_H(C + p) > 0$ for some $p \in X$ and $\lambda_H(E + x) = 0$ for every $x \in X$. An arbitrary subset $F \subseteq X$ is finitely shy in C if it is contained in a finitely shy universally measurable set.

Anderson and Zame (2001) showed that if a set E is finitely shy in C, then it is also shy in C. A subset $Y \subseteq C$ is said to be prevalent in C if its complement $C \backslash Y$ is shy in C.

Consider a convex family of priors $\mathcal{P}$. Positive multiples of priors in $\mathcal{P}$ constitute a convex cone of (positive) measures. Taking the differences of pairs of such measures yields the vector space of signed measures that are generated by $\mathcal{P}$, denoted $\mathcal{M}$. We assume that the vector space $\mathcal{M}$ is endowed with a topology that satisfies the following two properties:

(i) The mappings

$$(p, p') \rightarrow p + p',$$
$$(\alpha, p) \rightarrow \alpha p$$

are continuous in $p, p' \in \mathcal{P}$ and $\alpha \in \mathbb{R}$ (these continuity requirements make $\mathcal{M}$ a topological vector space).

(ii) A subset $A \subseteq \mathbb{R}$ is Borel if and only if for every pair of priors $p, p' \in \mathcal{P}$, the one-dimensional set of weighted averages

$$\{\alpha p + (1 - \alpha) p' : \alpha \in A\}$$

is a Borel subset of $\mathcal{M}$.

These two properties are satisfied by a large variety of topologies on $\mathcal{M}$, including the topology of weak convergence and the topology of the total variation norm, but not by extremely strong topologies such as the totally disconnected topology in which every subset of $\mathcal{M}$ is open.

---

[6]A subset $E \subseteq X$ is universally measurable if it is measurable with respect to the completion of every regular Borel probability measure on X.

THEOREM 2: *Let $\mathcal{P}$ be a completely metrizable convex set of priors with a topology that satisfies requirements* (i) *and* (ii) *above. Suppose that $\mathcal{P}$ contains at least one NBDP prior for bidder i. If the subset $\mathcal{B} \subset \mathcal{P}$ of BDP priors for bidder i is universally measurable, then both $\mathcal{B}$ and the set of FSE priors for bidder i, $\mathcal{F}$, are finitely shy in $\mathcal{P}$.*

PROOF: Theorem 1 implies that $\mathcal{B}$ is a proper face of $\mathcal{P}$. Lemma 1 implies that $\mathcal{B}$ is finitely shy in $\mathcal{P}$. Proposition 2 implies that also $\mathcal{F} \subset \mathcal{B}$ is finitely shy in $\mathcal{P}$.   *Q.E.D.*

Lemma 1 states that a proper face of a convex set is finitely shy in the set.

LEMMA 1: *Let C be a convex subset of the topological vector space X. Suppose that X is endowed with a topology that satisfies the following property*: *for every $c \neq c' \in C$ and $A \subseteq R$, the one-dimensional set*

$$\{\alpha(c - c') : \alpha \in A\}$$

is a Borel subset of X if and only if A is a Borel subset of $\mathbb{R}$. Let F be a Borel set that is a proper face of C. Then F is finitely shy in C.

PROOF: Fix some $g \in C \setminus F$ and $f \in F$. Consider the one-dimensional subspace of X,

$$H = \{\alpha(g - f) : \alpha \in \mathbb{R}\}.$$

Observe that $\alpha(g - f) + f = \alpha g + (1 - \alpha)f \in C$ if $\alpha \in [0, 1]$ and hence $\lambda_H(C - f) \geq 1 > 0$. However, $\lambda_H(F + x) = 0$ for every $x \in X$. Indeed, $H \cap (F + x)$ is either empty or a singleton. Assume by contradiction that

$$f_1 + x = h_1 = \alpha_1(g - f),$$
$$f_2 + x = h_2 = \alpha_2(g - f),$$

where $h_1, h_2 \in H$, $f_1, f_2 \in F$, and $\alpha_1 > \alpha_2$. Then

$$f_1 - f_2 = (\alpha_1 - \alpha_2)g - (\alpha_1 - \alpha_2)f$$

or

$$\frac{1}{1 + (\alpha_1 - \alpha_2)} \cdot f_1 + \frac{(\alpha_1 - \alpha_2)}{1 + (\alpha_1 - \alpha_2)} \cdot f$$

$$= \frac{1}{1 + (\alpha_1 - \alpha_2)} \cdot f_2 + \frac{(\alpha_1 - \alpha_2)}{1 + (\alpha_1 - \alpha_2)} \cdot g,$$

where the left-hand side is a convex combination of $f_1, f \in F$, and hence in F, while the right-hand side is a convex combination of f_2 and g. Since F is a

face, this implies that $f_2, g \in F$. A contradiction to the assumption that $g \in C \setminus F$.   *Q.E.D.*

REMARK 3: Theorem 2 applies in particular when $\mathcal{P}$ is the convex family of all priors $\mathcal{P}_u$ on the universal space or the convex subfamily of priors with a finite support $\mathcal{P}_{fu}$ on the universal space, provided that $\mathcal{P}$ is endowed with a topology at least as strong as the topology of weak convergence, that satisfies properties (i) and (ii), and with which $\mathcal{P}$ is completely metrizable.[7] Theorem 2 is applicable because by Lemma 2 in the Appendix, the set $\mathcal{B}$ of BDP priors for bidder i is a Borel (and therefore universally measurable) subset of the set of priors $\mathcal{P}_u$ and hence also of its subset of finite-support priors $\mathcal{P}_{fu}$.[8]

REMARK 4—$\mathcal{B}$ Is a Null Set: In a recent paper, Perry and Reny (2003) define a set to be null if it is a countable union of finitely shy sets $\{A_n\}_{n=1}^{\infty}$, where for each n there exists a *one-dimensional* subspace $H_n = \{\alpha x_n : \alpha \in \mathbb{R}\}$ such that $\lambda_{H_n}(A_n + x) = 0$ for every $x \in X$. Inspection of the proof of Lemma 1 reveals that the set $\mathcal{B}$ of BDP priors for bidder i is a null set according to this definition.

4. DISCUSSION

4.1. *Comparison with Crémer and McLean's Results*

Crémer and McLean (1988) showed that within the set of models with a *fixed* finite number of types $n_i \geq 2$ for each player i (equivalently, within the set of priors that are supported on a *fixed* finite number of types $n_i \geq 2$ for each player i), the set of priors that permit full surplus extraction from any bidder is generic. Our argument cannot be phrased in this more limited setting, because the set of priors that are supported on a fixed finite number of types *is not convex*. For example, the mixture $p = \alpha p' + (1 - \alpha) p''$ of the common priors that are represented by the two matrices in (1) is *not* the prior that is represented by the matrix

	$\theta_2''' = (v_2, b_2'')$	$\widetilde{\theta}_2''' = (\tilde{v}_2, \tilde{b}_2''')$
$\theta_1''' = (v_1, b_1''')$	$\alpha a' + (1-\alpha)a''$	$\alpha b' + (1-\alpha)b''$
$\widetilde{\theta}_1''' = (\tilde{v}_1, \tilde{b}_1''')$	$\alpha c' + (1-\alpha)c''$	$\alpha d' + (1-\alpha)d''$

---

[7]Prior $\mathcal{P}_u$ is a complete metric space with, for instance, the topology of weak convergence, as well as with the total-variation norm. In contrast, its subspace of finite-support priors $\mathcal{P}_{fu}$ is a complete metric space with the total-variation norm, but we do not know if it is completely metrizable also with the topology of weak convergence. Since complete metrizability of the ambient convex set C is a prerequisite for *defining* shyness of its subsets, the definition itself might therefore apply to the finite-support priors $\mathcal{P}_{fu}$ only with a smaller range of topologies than the range with which it applies to $\mathcal{P}_u$. (Such a subtlety did not arise in the purely geometric argument of the previous subsection, which relied entirely on the linear structure of the spaces and did not involve any choice of topology.)

[8]We do not know if Theorem 2 can be proved also for the set of priors $\mathcal{P}_f$ on finite type spaces (not necessarily within the universal space).

but rather the prior that is represented by the matrix (2), which is supported on eight states rather than on four.

4.2. *Approximate and Robust Full Surplus Extraction*

A number of results suggest that although it might be impossible to implement a given social choice rule, it might nevertheless be possible to implement a rule that is ε close to it, for any $\varepsilon > 0$. When this is the case, the social choice rule is said to be "virtually implementable" (see, e.g., McAfee and Reny (1992), Abreu and Matsushima (1992)). Our results imply that full surplus extraction also fails to be generically virtually implementable. This is because if a prior p is not BDP for bidder i—as is the case for generic priors—then at most $1 - \varepsilon_p^i$ of the surplus can be extracted from this bidder for some $\varepsilon_p^i > 0$. Hence, for $0 < \varepsilon < \varepsilon_p^i$, it is not the case that $1 - \varepsilon$ of the surplus can be extracted from this bidder.

We conjecture, but have been so far unable to prove, that, for every small $\varepsilon > 0$, both the set of priors in which it is possible to extract at least $1 - \varepsilon$ of the available surplus and the set of priors in which it is impossible to extract at least $1 - \varepsilon$ of the available surplus are not small in the sense that neither of them is shy.[9]

A conceptually distinct question is how much surplus can be extracted in a robust way. Suppose that for a given prior p, the principal has designed an optimal mechanism μ_p that extracts as much surplus as possible. If it turns out that the principal has misspecified the prior slightly, would the mechanism μ_p extract nearly as much of the surplus as could be extracted with the correct prior? We conjecture that the answer to this question is negative, and furthermore, that the extent of surplus extraction by a fixed optimal mechanism is generally discontinuous in the prior. That is, we conjecture that arbitrarily close to any prior p, there exists another prior p' such that the mechanism μ_p extracts a much smaller portion of the surplus than the portion of surplus that μ_p extracts under p.[10]

Finally, it is also interesting to know how the portion of the extractable surplus varies with the prior p when the mechanism is allowed to vary optimally

---

[9]For the case of public good provision, Neeman (2004) describes an example where if beliefs do not determine preferences, then the probability that a public good can be provided decreases to zero with the number of players, while efficiency requires that the public good be provided with probability 1. It therefore follows that in such a setting the total surplus that can be extracted from the players converges to zero at the same time that the total surplus that could be generated by the players remains uniformly bounded away from zero.

[10]Consider the sequence of common priors

$$
p_n = \left(
\begin{array}{c|c|c}
 & v = 1 & v = 2 - \frac{1}{n} \\
\hline
v = 1 & \frac{1}{3} & \frac{1}{6} \\
\hline
v = 2 - \frac{1}{n} & \frac{1}{6} & \frac{1}{3} \\
\end{array}
\right).
$$

with the prior. Since the set of FSE priors is dense (even though, as we showed, it is nongeneric), the portion of the extractable surplus is discontinuous at any NFSE prior. The percentage of the surplus that can possibly be extracted as a function of the prior is thus discontinuous "almost everywhere" (i.e., on a prevalent set of priors). We conjecture that the extractable surplus may nevertheless be continuous at the nongeneric subset of environments described by the FSE priors themselves.

4.3. *Type Spaces Without a Common Prior*

In this paper we restrict our attention to type spaces that admit a common prior. There are two reasons for this restriction. First, this is a standard practice in economic modeling, synonymous with the so-called Harsanyi doctrine. More importantly, the universal type set T_i of each bidder (see the Appendix) has the product structure $V_i \times \Delta(T_{-i})$. It therefore violates the BDP property in the most extreme possible way—every belief type of the bidder can be associated with each and every one of its possible valuations. Thus FSE is impossible in the universal type space.

4.4. *Fubini's Theorem*

The fact that no invariant measure exists on infinite-dimensional vector spaces prevents the notion of prevalence from satisfying *all* the properties that full Lebesgue measure satisfies in finite-dimensional spaces. For example, it fails to satisfy the analogue of Fubini's theorem. There could be a subset E of an infinite-dimensional space $Y \times Z$ such that the sections $\{z \in Z : (y, z) \in E\}$ are shy in Z for every $y \in Y$, while the sections $\{y \in Y : (y, z) \in E\}$ are prevalent in Y for every $z \in Z$. This, of course, is impossible if $Y \times Z$ is finite dimensional.

For instance, if we take $Y = \Delta(\mathbb{R})$, $Z = \mathbb{R}$, and $E = \{(\mu, z) \in \Delta(\mathbb{R}) \times \mathbb{R} : \mu(z) > 0\}$, then for every $\mu \in \Delta(\mathbb{R})$ the section $\{z \in \mathbb{R} : \mu(z) > 0\}$ (i.e., the atoms of μ) is at most countable and hence shy, while for every $z \in \mathbb{R}$ the section $\{\mu \in \Delta(\mathbb{R}) : \mu(z) > 0\}$ is prevalent because its complement is a proper face of $\Delta(\mathbb{R})$.

The total surplus that is generated by this sequence converges to $\frac{5}{3}$. However, for every n, the mechanism that extracts full surplus from the limit prior

$$p = \left(\begin{array}{c|c|c} & v=1 & v=2 \\ \hline v=1 & \frac{1}{3} & \frac{1}{6} \\ \hline v=2 & \frac{1}{6} & \frac{1}{3} \end{array} \right)$$

cannot extract more than $\frac{1}{3}$ from any element of the sequence because it excludes bidders with valuation $2 - \frac{1}{n}$.

The fact that this section is a countable intersection of weakly open and dense sets

$$\{\mu \in \Delta(\mathbb{R}): \mu(z) > 0\} = \bigcap_{n=1}^{\infty}\left\{\mu \in \Delta(\mathbb{R}): \mu\left(z - \frac{1}{n}, z + \frac{1}{n}\right) > 0\right\}$$

implies that it is a second-category set with the topology of weak convergence. Thus, this example illustrates that the analogue of Fubini's theorem also fails in infinite-dimensional spaces for this alternative, topological, notion of "largeness" and not only for the measure-theoretic notion of prevalence.

Dept. of Economics and Management, The Open University of Israel; aviadhe@openu.ac.il; http://www.openu.ac.il/Personal_sites/Aviad-Heifetz.html
and
Dept. of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, U.S.A.; and Center for Rationality and Economics Department, Hebrew University of Jerusalem, Israel 91904; zvika@BU.edu; http://people.bu.edu/zvika.

APPENDIX

The Private-Values Universal Type Space

Given a basic space of uncertainty $V \equiv V_1 \times \cdots \times V_n$ and a set of bidders N as defined in Section 2.1, there exists a *private-values universal type space*

$$T = \prod_{i \in N} T_i$$

into which every other private-values type space can be mapped in a beliefs-preserving way. The proof of existence follows from a slight adaptation of the arguments contained in Mertens and Zamir (1985), Brandenburger and Dekel (1993), and Heifetz (1993). That is, for every type space $\Theta \equiv \prod_{i \in N} \Theta_i$, there exists a unique set of measurable mappings

$$\eta_i: \Theta_i \to T_i, \quad i \in N,$$

that satisfy

$$\hat{v}_i(\eta_i(\theta_i)) = \hat{v}_i(\theta_i)$$

and

$$\hat{b}_i(\eta_i(\theta_i))(A) = \hat{b}_i(\theta_i)(\eta_{-i}^{-1}(A))$$

for every measurable set $A \subseteq T_{-i}$, where $\eta_{-i} : \Theta_{-i} \to T_{-i}$ is defined by

$$\eta_{-i}((\theta_j)_{j \neq i}) = (\eta_j(\theta_j))_{j \neq i}.$$

The universal type set T_i of bidder $i \in N$ is isomorphic to the product space

$$V_i \times \Delta(T_{-i})$$

by the mapping

$$\tau_i \to (\hat{v}_i(\tau_i), \hat{b}_i(\tau_i)).$$

Thus, in what follows we use the terms T_i and $V_i \times \Delta(T_{-i})$ interchangeably.

PROOF OF PROPOSITION 1: Suppose that $p = \alpha p_1 + (1 - \alpha)p_2$, where $\alpha \in (0, 1)$. We first show that p is non-BDP if either p_1 or p_2 is non-BDP. Suppose, without loss of generality that p_1 is non-BDP. It follows that for some bidder $i \in N$ there exist two disjoint measurable subsets of i's types, $A_i, A_i' \subseteq \Theta_i$, that each have a positive p_1-probability,

$$p_{1|\Theta_i}(A_i) > 0, \quad p_{1|\Theta_i}(A_i') > 0,$$

and the same range of beliefs,

$$\hat{b}_i(A_i) = \hat{b}_i(A_i') \subseteq \Delta(\Theta_{-i}),$$

but different valuations. That is, if $\theta_i \in A_i$ and $\theta_i' \in A_i'$ are such that

$$\hat{b}_i(\theta_i') = \hat{b}_i(\theta_i),$$

then

$$\hat{v}_i(\theta_i') \neq \hat{v}_i(\theta_i).$$

It therefore follows that

$$p_{|\Theta_i}(A_i) = \alpha p_{1|\Theta_i}(A_i) + (1 - \alpha)p_{2|\Theta_i}(A_i) > 0,$$
$$p_{|\Theta_i}(A_i') = \alpha p_{1|\Theta_i}(A_i') + (1 - \alpha)p_{2|\Theta_i}(A_i') > 0,$$

which implies that p cannot be a BDP prior.

We now show that p is BDP if both p_1 and p_2 are BDP. If p_1 and p_2 are BDP priors for player i, then by Definition 1 there exist subsets $\Theta_i^{p_k} \subseteq \Theta_i$, $k \in \{1, 2\}$, such that $p_{|\Theta_i}(\Theta_i^{p_k}) = 1$ and $\Theta_i^{p_k}$ is the graph of a function $\Phi_i^{p_k} : B_i^{p_k} \to V_i$, where $B_i^{p_k}$ is the projection of $\Theta_i^{p_k}$ on $\Delta(\Theta_{-i})$.

We show that the graph of the function $\Phi_i : B_i^{p_1} \cup B_i^{p_2} \to V_i$ defined by

$$\Phi_i(b_i) = \begin{cases} \Phi_i^{p_1}(b_i), & b_i \in B_i^{p_1}, \\ \Phi_i^{p_2}(b_i), & \text{otherwise,} \end{cases}$$

is assigned probability 1 by $p_{|\Theta_i}$. This implies that p is BDP for bidder i.

This is obvious if $B_i^{p_1} \cap B_i^{p_2} = \emptyset$, because in this case the graph of Φ_i is simply the union of the graphs of $\Phi_i^{p_1}$ and $\Phi_i^{p_2}$. We show that the graph of Φ_i is assigned probability 1 by $p_{|\Theta_i}$ (and hence p is BDP for bidder i) also if $B_i^{p_1} \cap B_i^{p_2} \neq \emptyset$. The proof consists of showing that the graphs of $\Phi_i^{p_1}$ and $\Phi_i^{p_2}$ coincide almost surely on $B_i^{p_1} \cap B_i^{p_2}$ according to $p_{|\Theta_i}$.

Pick some bidder $j \neq i$ and, for $k = 1, 2$, denote by

$$\bar{\Theta}_j^{p_k} = \{\theta_j \in \Theta_j : \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_k}) = 1\}$$

the set of j's types that assign probability 1 to $\Theta_i^{p_k}$. Because p_k is a common prior,

$$(3) \qquad 1 = p_{k|\Theta_i}(\Theta_i^{p_k})$$

$$= \int_{\Theta_j} \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_k}) \, dp_{k|\Theta_j}(\theta_j)$$

$$= \int_{\bar{\Theta}_j^{p_k}} \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_k}) \, dp_{k|\Theta_j}(\theta_j) + \int_{\Theta_j \setminus \bar{\Theta}_j^{p_k}} \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_k}) \, dp_{k|\Theta_j}(\theta_j)$$

$$= \int_{\bar{\Theta}_j^{p_k}} 1 \cdot dp_{k|\Theta_j}(\theta_j) + \int_{\Theta_j \setminus \bar{\Theta}_j^{p_k}} \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_k}) \, dp_{k|\Theta_j}(\theta_j).$$

Since $\hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_k}) < 1$ on $\Theta_j \setminus \bar{\Theta}_j^{p_k}$, we conclude that $p_{k|\Theta_j}(\bar{\Theta}_j^{p_k}) = 1$ (and $p_{k|\Theta_j}(\Theta_j \setminus \bar{\Theta}_j^{p_k}) = 0$), because otherwise the sum of the two integrals in (3) would be smaller than 1.

Since p_k is a common prior (and $\Theta_i^{p_k} \subset B_i^{p_k} \times V_i$), it follows that

$$\int_{B_i^{p_k} \times V_i} \hat{b}_i(\theta_i)_{|\Theta_j}(\bar{\Theta}_j^{p_k}) \, dp_{k|\Theta_i}(\theta_i) = p_{k|\Theta_j}(\bar{\Theta}_j^{p_k}) = 1$$

and hence that $\hat{b}_i(\theta_i)_{|\Theta_j}(\bar{\Theta}_j^{p_k}) = 1$ for $p_{k|\Theta_i}$-almost every $\theta_i \in B_i^{p_k} \times V_i$. This implies that for $p_{|\Theta_i}$-almost every $\theta_i \in (B_i^{p_1} \cap B_i^{p_2}) \times V_i$ we have $\hat{b}_i(\theta_i)_{|\Theta_j}(\bar{\Theta}_j^{p_1} \cap \bar{\Theta}_j^{p_2}) = 1$.

Furthermore, the belief of every $\theta_j \in \bar{\Theta}_j^{p_1} \cap \bar{\Theta}_j^{p_2}$ is concentrated on $\Theta_i^{p_1} \cap \Theta_i^{p_2}$ or

$$\hat{b}_j(\theta_j)_{|\Theta_i}\big((B_i^{p_1} \cap B_i^{p_2}) \times V_i\big) = \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{p_1} \cap \Theta_i^{p_2}) = 1.$$

In other words, if $B^0 = \{b_i \in B_i^{P_1} \cap B_i^{P_2} : \Phi_i^{P_1}(b_i) \neq \Phi_i^{P_2}(b_i)\}$ is the subset of i's belief types in $B_i^{P_1} \cap B_i^{P_2}$ on which the graphs $\Theta_i^{P_1}$ and $\Theta_i^{P_2}$ are disjoint, then $\hat{b}_j(\theta_j)_{|\Theta_i}(B^0 \times V_i) = 0$. Since p is a common prior, we therefore have

$$p_{|\Theta_i}(\Theta_i^{P_1} \cap \Theta_i^{P_2}) = \int_{\bar{\Theta}_j^{P_1} \cap \bar{\Theta}_j^{P_2}} \hat{b}_j(\theta_j)_{|\Theta_i}(\Theta_i^{P_1} \cap \Theta_i^{P_2}) \, dp_{|\Theta_j}(\theta_j)$$

$$= \int_{\bar{\Theta}_j^{P_1} \cap \bar{\Theta}_j^{P_2}} \hat{b}_j(\theta_j)_{|\Theta_i}\big((B_i^{P_1} \cap B_i^{P_2}) \times V_i\big) \, dp_{|\Theta_j}(\theta_j)$$

$$= p\big(((B_i^{P_1} \cap B_i^{P_2}) \times V_i) \times (\bar{\Theta}_j^{P_1} \cap \bar{\Theta}_j^{P_2}) \times \Theta_{-i-j}\big)$$

$$= \int_{(B_i^{P_1} \cap B_i^{P_2}) \times V_i} \hat{b}_i(\theta_i)_{|\Theta_j}(\bar{\Theta}_j^{P_1} \cap \bar{\Theta}_j^{P_2}) \, dp_{|\Theta_i}(\theta_i)$$

$$= \int_{(B_i^{P_1} \cap B_i^{P_2}) \times V_i} 1 \cdot dp(\theta_i)$$

$$= p_{|\Theta_i}\big((B_i^{P_1} \cap B_i^{P_2}) \times V_i\big).$$

This equality implies that on $B_i^{P_1} \cap B_i^{P_2}$ the graphs of $\Phi_i^{P_1}$ and $\Phi_i^{P_2}$ coincide p-almost surely, as required, because if this were not the case on a subset of $B_i^{P_1} \cap B_i^{P_2}$ with a positive measure according to p, then it would follow that $p_{|\Theta_i}(\Theta_i^{P_1} \cap \Theta_i^{P_2}) < p_{|\Theta_i}((B_i^{P_1} \cap B_i^{P_2}) \times V_i)$.   *Q.E.D.*

LEMMA 2: *The set $\mathcal{B}$ of BDP priors for bidder i is a Borel subset of the space of priors $\mathcal{P}_u$ on the universal space T, when $\mathcal{P}_u$ is endowed with a topology at least as strong as the topology of weak convergence.*

PROOF: If the lemma obtains when $\mathcal{P}_u$ is equipped with the topology of weak convergence, it also obtains for any stronger topology. It is therefore enough to proceed by assuming that $\mathcal{P}_u$ is equipped with the topology of weak convergence.

A prior $p \in \mathcal{P}_u$ is a BDP prior if and only if the marginal of p on $T_i = V_i \times \Delta(T_{-i})$ is concentrated on a measurable graph of a function $\Phi_i^p : B_i^p \to V_i$. This is expressible by countably many conditions, in the following way.

Since V_i is separable, there is a countable collection $\{A_i^n\}_{n \geq 1}$ of subsets of V_i that is closed under complements and finite unions, and generates the Borel sigma-field of V_i. Hence there are also countably many partitions $\{\Gamma_i^m\}_{m \geq 1}$ of V_i to finitely many disjoint subsets $\{A_i^{n_k^m}\}_{k=1}^{N_i^m} \subseteq \{A_i^n\}_{n \geq 1}$. Similarly, since $\Delta(T_{-i})$ is separable, there exists a countable collection $\{Y_i^\ell\}_{\ell \geq 1}$ of subsets of $\Delta(T_{-i})$ that is closed under complements and finite unions, and generates the Borel sigma-field of $\Delta(T_{-i})$. Hence, there are also countably many partitions $\{\Lambda_i^r\}_{r \geq 1}$ of $\Delta(T_{-i})$ to finitely many disjoint subsets in $\{Y_i^{\ell_k^r}\}_{k=1}^{L_i^r} \subseteq \{Y_i^\ell\}_{\ell \geq 1}$.

The marginal of p on $T_i = V_i \times \Delta(T_{-i})$ is concentrated on the graph of Φ_i^p if, and only if for every partition $\Gamma_i^m = \{A_i^{n_k^m}\}_{k=1}^{N_i^m}$ of S_i,

$$p\left(\bigcup_{k=1}^{N_i^m} \left(A_i^{n_k^m} \times (\Phi_i^p)^{-1}\left(A_i^{n_k^m}\right) \times T_{-i}\right)\right) = 1.$$

Intuitively, as the partitions $(\Gamma_i^m)_{m \geq 1}$ of V_i get finer, the union of the rectangles $A_i^{n_k^m} \times (\Phi_i^p)^{-1}(A_i^{n_k^m})$ approximates the graph of Φ_i^p increasingly well.

Now, for each partition $\Gamma_i^m = \{A_i^{n_k^m}\}_{k=1}^{N_i^m}$ of V_i, $\{(\Phi_i^p)^{-1}(A_i^{n_k^m})\}_{k=1}^{N_i^m}$ is a partition of $\Delta(T_{-i})$ that can be approximated arbitrarily well (in terms of the probabilities assigned to the partition members by the marginal of p on $\Delta(T_{-i})$) by partitions in $\{\Lambda_i^r\}_{r \geq 1}$. Hence, the marginal of p on $T_i = V_i \times \Delta(T_{-i})$ is concentrated on a measurable graph from $\Delta(T_{-i})$ to V_i if and only if, for every natural number $q \geq 1$ and for each partition $\Gamma_i^m = \{A_i^{n_k^m}\}_{k=1}^{N_i^m}$ of V_i, there exists a partition $\Lambda_i^r = \{Y_i^{\ell_k^r}\}_{k=1}^{L_i^r}$ of $\Delta(T_{-i})$ with $L_i^r = N_i^m$ and

$$p\left(\bigcup_{k=1}^{N_i^m}\left(A_i^{n_k^m} \times Y_i^{\ell_k^r} \times T_{-i}\right)\right) \geq 1 - \frac{1}{q}.$$

Formally, therefore, the set $\mathcal{B}$ of BDP priors is

$$\bigcap_{i \in N}\bigcap_{m \geq 1}\bigcap_{q \geq 1}\bigcup_{r \geq 1}\left\{p \in \mathcal{P}_u : p\left(\bigcup_{k=1}^{N_i^m}\left(A_i^{n_k^m} \times Y_i^{\ell_k^r} \times T_{-i}\right)\right) \geq 1 - \frac{1}{q}\right\},$$

which is a Borel subset of the space of priors $\mathcal{P}_u$.   *Q.E.D.*

REFERENCES

ABREU, D., AND H. MATSUSHIMA (1992): "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica*, 60, 993–1108.

ANDERSON, R. M., AND W. R. ZAME (2001): "Genericity with Infinitely Many Parameters," *Advances in Theoretical Economics*, 1, 1–62.

BRANDENBURGER, A., AND E. DEKEL (1993): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 59, 189–198.

CHRISTENSEN, J. P. R. (1974): *Topology and Borel Structure*, Mathematical Studies, Vol. 10. Amsterdam: North-Holland.

CRÉMER, J., AND R. MCLEAN (1985): "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands Are Interdependent," *Econometrica*, 53, 345–361.

——— (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1257.

ELY, J. C., AND M. PESKI (2006): "Hierarchies of Belief and Interim Rationalizability," *Theoretical Economics*, 1, to appear.

HEIFETZ, A. (1993): "The Bayesian Formulation of Incomplete Information—The Non-Compact Case," *International Journal of Game Theory*, 21, 329–338.

HEIFETZ, A., AND Z. NEEMAN (2005): "On the Generic Implementation of Arbitrary Decision Rules," Unpublished Manuscript.

HUNT, B. R., T. SAUER, AND J. A. YORKE (1992): "Prevalence: A Translation-Invariant 'Almost Every' on Infinite-Dimensional Spaces," *Bulletin (New Series) of the American Mathematical Society*, 27, 217–238.

MCAFEE, P. R., AND P. J. RENY (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395–421.

MERTENS, J. F., S. SORIN, AND S. ZAMIR (1994): "Repeated Games, Part A: Background Material," Discussion Paper 9420, CORE.

MERTENS, J. F., AND S. ZAMIR (1985): "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 10, 619–632.

NEEMAN, Z. (2004): "The Relevance of Private Information in Mechanism Design," *Journal of Economic Theory*, 117(1), 55–77.

OTT, W., AND J. A. YORKE (2005): "Prevalence," *Bulletin (New Series) of the American Mathematical Society*, 42, 263–290.

PERRY, M., AND P. RENY (2003): "Toward a Strategic Foundation for Rational Expectations Equilibrium," Mimeo, University of Chicago.

ROCKAFELLAR, T. R. (1970): *Convex Analysis*. Princeton, NJ: Princeton University Press.

WILSON, R. (1987): "Game-Theoretic Analyses of Trading Processes," in *Advances in Economic Theory: Fifth World Congress*, ed. by T. Bewley. Cambridge, U.K.: Cambridge University Press, 33–70.

YAMASAKI, Y. (1985): *Measures on Infinite Dimensional Spaces*. Singapore: World Scientific.