Preprints of the Max Planck Institute for Research on Collective Goods Bonn 2011/28

Beliefs, Payoffs, Information: On the Robustness of the BDP Property in Models with Endogenous Beliefs

Alia Gizatulina
Martin Hellwig

MAX PLANCK SOCIETY

Preprints of the
Max Planck Institute
for Research on Collective Goods
Bonn 2011/28

Beliefs, Payoffs, Information: On the Robustness of the BDP Property in Models with Endogenous Beliefs

Alia Gizatulina / Martin Hellwig

First version December 2011
Revised Version March 2012

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.coll.mpg.de

Beliefs, Payoffs, Information: On the Robustness of the BDP Property in Models with Endogenous Beliefs*

Alia Gizatulina† and Martin Hellwig‡
Max Planck Institute for Research on Collective Goods
Kurt-Schumacher-Str. 10, D-53113 Bonn, Germany

First version: December 1, 2008
This version: March 23, 2012

Abstract

Neeman (2004) and Heifetz and Neeman (2006) have shown that, in auctions with incomplete information about payoffs, full surplus extraction is only possible if agents' beliefs about other agents are fully informative about their own payoff parameters. They argue that the set of incomplete-information models satisfying this so-called BDP property ("beliefs determine preferences") is negligible, in a geometric and a measure-theoretic sense. In contrast, we show that, in models with finite-dimensional type spaces, this property is topologically generic if the set of objects about which beliefs are formed is sufficiently rich and beliefs are derived by conditioning on the available information; for any agent, this information includes his own payoff parameters.

JEL Classification: D40, D44, D80, D82

Keywords: universal type space, surplus extraction, mechanism design, BDP, correlated information.

---

*For helpful comments and discussions, we would like to thank Beth Allen, Sophie Bade, Felix Bierbrauer, Yi-Chun Chen, Jacques Crémer, Christoph Engel, Amanda Friedenberg, Bruno Jullien, Aniol Llorente-Saguer, Benny Moldovanu, Stephen Morris, Zvika Neeman, Wojciech Olszewski, and Siyang Xiong.

†gizatulina@coll.mpg.de

‡hellwig@coll.mpg.de

1 Introduction

Neeman (2004) and Heifetz and Neeman (2006) have drawn attention to the so-called BDP property and the role this property plays in mechanism design models with correlated values. The label "BDP" - beliefs determine preferences - refers to the fact that, if this property holds and if one knows an agent's beliefs, then one also knows his preferences. They show that this property underlies the findings of Crémer and McLean (1988) or McAfee and Reny (1992) that correlations of agents types can be used to eliminate information rents. In Crémer and McLean (1988) or McAfee and Reny (1992), differences in an agent's beliefs about the other agents' characteristics induce differences in attitudes towards state-contingent payment schemes; such schemes are then used to screen agents in order to extract rents. Such rent extraction is necessarily incomplete if there are different states of the world in which an agent has different payoff parameters and the same beliefs.[1]

Heifetz and Neeman (2006) suggest that the set of incomplete-information models having the BDP property is a negligible subset of the set of all incomplete-information models that are consistent with common priors.[2] Their suggestion is based on the view, that, in a model with private values, the preferences of an agent can be specified independently of his beliefs about other agents (Heifetz and Neeman (2006), p. 215).

We want to take issue with this view. In a common-prior setting, beliefs are the result of agents conditioning their expectations on whatever information they have. This information includes their own preferences. The view that preferences and beliefs can be specified independently presumes that the information that an agent has about this preferences is not relevant for forming expectations about other agents' characteristics. This would be the case, for example, if, under the common prior, the agent's preferences and the other agents' characteristics were stochastically independent. If the agent's preferences and the other agent's characteristics were not independent, the agent's information

[1] Neeman (2004) applies this reasoning to a public-good provision problem with participation constraints. He shows that feasible provision levels are close to zero when there are many agents and certain violations of BDP are uniform over agents and states of the world, regardless of the number of participants. Gizatulina and Hellwig (2010) show that such uniform violations of BDP are incompatible with the notion that, in a large economy, individuals may be informationally small.

[2] Formally, they study the status of BDP models *within a fixed family of incomplete-information models*. Assuming that this family is what they call "closed under finite unions", they show that, if the family contains at least one BDP model, then within this family, failure of the BDP property is generic in a geometric and in a measure theoretic sense. We discuss their results in Section 5.

about his own preferences would still be irrelevant for expectations formation if this information was redundant in the sense that it is already contained in some other information variable that the agent also observes. Stochastic independence would be incompatible with correlated values, the very specification that Heifetz and Neeman (2006) as well as Crémer and McLean (1988) are concerned with. Complete redundancy would appear to be very special.[3]

Even if preferences and beliefs are not specified independently, however, the BDP property cannot be taken for granted. If beliefs are obtained by conditioning on information and if this information includes the agent's preferences, the question is whether the map from information to beliefs is invertible. If it is, then the agent's information, including his preferences, can be inferred from his beliefs. If the map from information to beliefs is not invertible, this inference is not possible everywhere.

There can be several reasons why the map from information to beliefs might not be invertible. First, if there are many information variables and few parameters characterizing beliefs, it may be impossible to tell which combination of information variables has given rise to a given belief. Suppose, for example, that in an auction with two participants, an agent observes not only the value he assigns to the object at auction but also a signal of the value that other agent assigns to the object. With correlated private values, a given belief that the other agent assigns a high value to the object may then be due to the fact that the agent's own value is high or to the fact that the agent has received an optimistic signal about the other agent. The confounding of influences of different information variables on beliefs makes it impossible to recover the value of any information variable from the beliefs.

Second, even if the agent observes only the value he assigns to the object, correlations might be such that the map from information variables to beliefs is not monotonic. In this case, a given belief about the other agent's valuation might be generated by different realizations of the agent's own valuation.

To assess whether the BDP property is negligible or not, one must consider whether a lack of invertibility of the map from information variables to beliefs should be considered

[3]Redundance of information about preferences might appear to be natural in a universal-type-space setting where the information on which an agent conditions is represented by his "type", a vector made up of his preferences and his beliefs. In this framework, conditioning of beliefs on information ("type") is trivial because the beliefs themselves are specified as part of the information. There is no account however, of what information gives rise to the beliefs. Moreover, the assumed redundancy of information about preferences imposes severe restrictions on the common prior.

to be the rule or the exception. Heifetz and Neeman (2006) do not actually study this question. When they impose a common-prior assumption, they require beliefs to be given by conditional distributions, but they do not investigate the implications of this requirement for their program.

This is where our paper steps in. In an abstract type space formulation, we explicitly consider the mappings from the information variables that agents observe to their beliefs, i.e., conditional distributions. Our main result shows that a confounding of influences of different information variables is unlikely to occur if the set of objects about which the agent forms his beliefs is sufficiently rich. In this case, the BDP property then is the rule, rather than the exception.

Throughout the paper, we consider an agent's "type" as reflecting not just his preference parameters but also additional information variables that he may observe. We do not limit ourselves to the "naive" type spaces considered by Crémer and McLean (1988), where agents' types are defined by their preference parameters only and differences in preference parameters are the only source of heterogeneity in beliefs.

If type sets are finite, however, the same logic as in Crémer and McLean (1988) implies that, if the cardinality of an agent's type set is less than or equal to the product of the cardinalities of the other agents' type sets, then, for this agent, the BDP property holds for an open and dense set of priors. If all agents' type sets have the same cardinality, then, for all agents, the BDP property holds for an open and dense set of common priors.[4]

Our main results concern models with a continuum of types. We treat the type t_i of agent i as a vector in $\mathbb{R}^{n_i}$ for some natural number n_i. The vector of all agents' types is a vector in $\mathbb{R}^N$ where $N = \sum_i n_i$. Priors are probability distributions on $\mathbb{R}^N$. We restrict the analysis to priors for which marginal distributions of agents' types have full supports and regular conditional probability distributions have continuous versions. For any agent i, such a regular conditional probability distribution is represented by a continuous function from the space $\mathbb{R}^{n_i}$ of agent i's types into the space $\mathcal{M}(\mathbb{R}^{N \setminus n_i})$ of probability distributions for the other agents' types. The topology on the space of priors is specified as the coarsest topology under which the maps from priors to marginal distributions of agents' types and to continuous regular conditional probability distributions over the other agents' types are continuous. With this topology, the BDP property holds for any agent i on a residual set of priors, i.e., on a countable intersection of open and dense sets of probability distributions on $\mathbb{R}^N$.

[4]This confirms a conjecture in Compte and Jehiel (2009), p. 188.

These results are based on an extension of the classical Embedding Theorem for continuous functions.[5] An embedding of a metric space X in a metric space Y is a one-to-one bicontinuous function from X into (a subset of) Y. The standard version of the Embedding Theorem asserts that, for any natural numbers n and m, if $m \geq 2n+1$, then the set of embeddings of a set $X \subset \mathbb{R}^n$ in $[0,1]^m$ contains a residual subset of the set of continuous functions from X into $[0,1]^m$. We use this theorem to show that, for any natural number n, the set of embeddings of a compact set $X \subset \mathbb{R}^{n_i}$ into the space $Y = \mathcal{M}(Z)$ of probability measures on a compact subset $Z \subset \mathbb{R}^{N \setminus n_i}$ that has infinitely many elements contains a residual subset of the set of continuous functions from X into $\mathcal{M}(Z)$. The proof of this result is based on the observation that, no matter what n_i and N may be, the dimension of $\mathcal{M}(Z)$ exceeds $2n_i + 1$.

In our setting, the Embedding Theorem implies that the set of embeddings contains a residual subset of the set of continuous regular conditional distributions for any agent i. Under the given topology on the set of priors, it follows that the set of priors giving rise to regular conditional distributions that are embeddings must contain a residual subset of the set of all priors that give rise to continuous regular conditional distributions. Because embeddings are injective, any prior in this set exhibits the BDP property, i.e., under any such prior, one can recover the agent's type from his beliefs about the other agents' types. For any agent i, the BDP property thus holds on a residual set of priors.

Because the set of priors that satisfy the BDP property for all agents is given by the (finite) intersection of the sets of priors that satisfy the BDP property for agent i, $i = 1, ...I$, the BDP property also holds for all agents on a residual set of common priors. On the set of common priors that have continuous density functions, the topology that yields residualness of the BDP property turns out to be the uniform topology for density functions.

In the following, Section 2 lays out the basic framework of our analysis. Section 3 introduces the BDP property and gives a few examples in order to build some intuition. Section 4 formulates and proves our genericity results. Section 4.1 gives the result for finite type spaces. Sections 4.2 - 4.4 give the results for continuous type spaces: Section 4.2 deals with the BDP property of a prior for a single agent, Section 4.3 with the BDP

[5] See Chapter V in Hurewicz and Wallman (1941). In economics, the literature on generic existence of completely revealing rational expectations equilibria has made extensive use of Embedding Theorems; see, in particular, Allen (1981). That literature, however, relied on Whitney's Embedding Theorem for C^r functions (Hirsch (1994), p. 35). Lacking the requisite differentiability properties, we use the Embedding Theorem for C^0 functions.

property of a common prior for all agents. Section 4.4 deals with the BDP property of common priors with continuous density functions.

The relation of our analysis to the literature is discussed in Section 5. There we show that the difference between our analysis and the analysis of Heifetz and Neeman (2006) is due to their considering genericity of the BDP property *within* a fixed family of models. If the family in question is the family of all common-prior models in the universal type space, this approach involves no loss of generality. In finite-dimensional abstract type spaces, however, the analysis of Heifetz and Neeman begs the question whether the families of models within which their genericity results hold are themselves robust. In a companion paper, Gizatulina and Hellwig (2011), we show that this is not the case. Using the results of this paper, we show there that the set of families within which the BDP property is robust is itself a residual set in the set of all families of incomplete-information models.

In contrast to this paper, Barelli (2009) and Chen and Xiong (2011) work with the universal type space. Barelli (2009) asserts that non-BDP models are topologically generic. Chen and Xiong (2011) point to a flaw in his analysis and show that the BDP property holds on a residual set of models in the universal type space. Their analysis uses the fact that, in the weak* topology on the universal type space, models with finite type sets are dense in the set of all models and BDP models are dense in the set of models with finite type sets. In contrast, we do not rely on finite approximations but on embedding theorems for continuous functions with finite-dimensional domains.

2 The Basic Framework

An abstract (Harsanyi) type space formulation of an incomplete-information model with $I \geq 2$ agents involves a collection

$$\mathcal{T} = \{T_i, \hat{\theta}_i, \hat{\pi}_i\}_{i=1}^{I} \tag{1}$$

such that, for any i, T_i is a set of abstract "types" for agent i, $\hat{\theta}_i$ is a mapping from T_i into a set Θ_i of payoff parameter vectors for agent i, and $\hat{\pi}_i$ is a mapping from T_i into the set $\mathcal{M}(T_{-i})$ of probability distributions on the space $T_{-i} := \prod_{j \neq i} T_j$ of the other agents' abstract types.[6] For any $t_i \in T_i$, $\hat{\theta}_i(t_i)$, the payoff type of agent i, indicates the agent's payoff

[6]See, e.g., Bergemann and Morris (2005).

parameters when his abstract type is t_i; $\hat{\pi}_i(t_i)$, the belief type of agent i, represents the agent's probabilistic beliefs about the other agents' types. Following Heifetz and Neeman (2006), we assume that the spaces T_i and Θ_i are complete separable metric spaces, and that the functions $\hat{\theta}_i : T_i \to \Theta_i$ and $\hat{\pi}_i : T_i \to \mathcal{M}(T_{-i})$ are continuous, where $\mathcal{M}(T_{-i})$ has the topology of weak convergence of probability measures, i.e. the weak* topology.

In this abstract type space formulation *a model exhibits the BDP property for agent i* if for any t_i and t'_i in T_i,

$$\hat{\pi}_i(t_i) = \hat{\pi}_i(t'_i) \text{ implies } \hat{\theta}_i(t_i) = \hat{\theta}_i(t'_i). \tag{2}$$

In the absence of restrictions on the mappings $\hat{\theta}_i$ and $\hat{\pi}_i$, there is no reason why this property should hold. One expects the set of models exhibiting this property to be in some sense negligible in the set of all models.

This conclusion cannot be taken for granted, however, if the belief mapping $\hat{\pi}_i$ must take the form

$$\hat{\pi}_i(\cdot) = b_i(\cdot, \nu_i), \tag{3}$$

where $\nu_i \in \mathcal{M}(T_i \times T_{-i})$ is a prior for agent i and $b_i(\cdot, \nu_i)$ is a regular conditional distribution for t_{-i} given t_i that is induced by ν_i, i.e. a function from T_i to $\mathcal{M}(T_{-i})$ such that for any bounded continuous function $f : T_{-i} \to \mathbb{R}$, $\int_{T_{-i}} f(t_{-i}) b_i(dt_{-i}|t_i, \nu_i)$ is the conditional expectation of $f(t_{-i})$ given t_i. Heifetz and Neeman impose this restriction with the additional requirement that ν_i be common to all agents.[7] We shall consider both, the case of agent-specific priors ν_i, $i = 1, ..., I$, and the case of a common prior ν such that $\nu_i = \nu$ for all i.

The information on which agent i conditions his beliefs includes his payoff type $\hat{\theta}_i(t_i)$. To make this dependence explicit, we find it convenient to write (3) in the form

$$\hat{\pi}_i(\cdot) = \hat{b}_i(\hat{\theta}_i(\cdot), \hat{s}_i(\cdot), \nu_i), \tag{4}$$

where, for any $t_i \in T_i$, $\hat{s}_i(t_i)$ is a vector of payoff-irrelevant information variables that agent i observes in addition to his payoff type $\hat{\theta}_i(t_i)$, and the mapping $\hat{s}_i(\cdot)$ takes values in some set S_i.[8] In this formulation, the underlying type space T_i matters only to the extent

[7] Dekel, Fudenberg, and Morris (2006) suggest that, in a non-common-prior environment, the approach of Heifetz and Neeman would not even be well defined.

[8] This is without loss of generality. Any model $\{T_i, \theta_i, \pi_i\}_{i=1}^I$ can trivially be rewritten in the form $\{\hat{T}_i, \hat{\theta}_i, \hat{\pi}_i\}_{i=1}^I$ where $\hat{T}_i = \Theta_i \times T_i$, $\hat{\theta}_i$ is the projection from $\Theta_i \times T_i$ to Θ_i, and, with $\hat{s}_i$ given as the projection from $\Theta_i \times T_i$ to T_i, $\hat{\pi}_i$ satisfies (4). In this formulation, the original type space T_i itself is interpreted as a space of signals on which expectations are conditioned. For a discussion in the context of the universal type space, see Section 5 below.

that t_i affects the payoff type $\hat{\theta}_i(t_i)$ and the information vector $\hat{s}_i(t_i)$. Therefore there is no loss of generality in identifying abstract types with pairs of payoff and signal vectors and writing

$$T_i = \Theta_i \times S_i, \tag{5}$$

with the understanding that, for any $t_i \in T_i$,

$$t_i = (\theta_i, s_i) \text{ implies } \hat{\theta}_i(t_i) = \theta_i \text{ and } \hat{s}_i(t_i) = s_i, \tag{6}$$

i.e., the maps $\hat{\theta}_i(\cdot)$ and $\hat{s}_i(\cdot)$ are the projections from T_i to Θ_i and S_i. The representation (1) of the incomplete-information model then takes the form

$$\hat{\mathcal{T}} = \{T_i, \hat{\pi}_i\}_{i=1}^I = \{\Theta_i \times S_i, \hat{\pi}_i\}_{i=1}^I, \tag{7}$$

with the understanding that payoff and information mappings are the projections and that belief mappings satisfy (4) for some prior ν_i. In this formulation, the beliefs

$$\hat{\pi}_i(t_i) = \hat{b}_i(\hat{\theta}_i(t_i), \hat{s}_i(t_i), \nu_i) = b_i(t_i, \nu_i) \tag{8}$$

of agent i concern the pairs $t_j = (\hat{\theta}_j(t_j), \hat{s}_j(t_j))$, $j \neq i$, of the other agents' payoff and information vectors.[9]

As is standard for abstract type space formulations and is proved in the Appendix to Heifetz and Neeman (2006), the mapping $\hat{\pi}_i(\cdot) = b_i(\cdot, \nu_i)$ can be used to build an infinite hierarchy of beliefs of i about the distribution of θ_{-i}, the joint distribution of θ_{-i} and the other agents' beliefs about preferences, etc. The abstract type space can thus be mapped into the Θ-based universal type space of Mertens and Zamir (1985). If there is a common prior, i.e., if the priors ν_i are all the same, the universal type space image of the abstract type space model must have the prior assign all probability mass to a belief-closed subset of the universal type space with the additional property that the belief of any type must correspond to the value of a regular conditional distribution given the information that this type has.

[9] A formulation with conditioning on other information variables, in addition to payoff parameters, was previously proposed by Compte and Jehiel (2009), p. 188.

3 The BDP Property

3.1 Definition

Heifetz and Neeman (2006) study the question what can be said about the set of models having the BDP property as a question about agents' priors. We shall do the same but we shall use a slightly different formulation.

Given a prior ν on the set $T := \prod_{j=1}^{I} T_j$, we say that the marginal distribution for the type t_i of agent i has *full support* if the support of this marginal distribution is T_i.

DEFINITION 3.1 *For any i, let ν be a prior on T such that the marginal distribution for t_i has full support and, moreover, there exists a continuous regular conditional distribution $\hat{\pi}_i$ for t_{-i} given t_i. The prior ν has the BDP property for agent i if, for any $t_i = (\theta_i, s_i)$ and $t_i' = (\theta_i', s_i')$ in $T_i = \Theta_i \times S_i$, $\hat{\pi}_i(t_i) = \hat{\pi}_i(t_i')$ implies $\hat{\theta}_i(t_i) = \hat{\theta}_i(t_i')$.*

DEFINITION 3.2 *Let ν be a common prior on T such that, for each i, the marginal distribution for t_i has full support and, moreover, there is a continuous regular conditional distribution $\hat{\pi}_i$ for t_{-i} given t_i.The common prior ν has the BDP property if it has the BDP property for each agent $i = 1, ..., n$.*

In contrast to our approach, Heifetz and Neeman (2006) do not require marginal type distributions to have full support. Instead they use a definition of the BDP property which neglects null sets. In their analysis, a prior on the underlying type space satisfies the BDP property for agent i if there exists a subset $\hat{T}_i$ of the type space for agent i such that the prior assigns probability one to $\hat{T}_i$ and, moreover, the BDP condition (2) holds for all t_i and t_i' in $\hat{T}_i$.

To see why we use a slightly different formulation, with a more restricted set of distributions and a stronger version of the BDP property, consider the following example: For given i, let $\Theta_i = [0,1]$, $S_i = [0,1]$, hence $T_i = [0,1]^2$, and let $\hat{\pi}_i$ be a continuous function from $[0,1]^2$ into $\mathcal{M}(T_{-i})$ that satisfies the BDP condition (2) for all t_i and t_i' in T_i. Let $\nu_i \in \mathcal{M}(T_i)$ be such that $\nu_i(\Theta_i \times \{s_i > 0\}) = 0$, and let $\nu \in \mathcal{M}(T)$ be given by the composition of ν_i and $\hat{\pi}_i$. By construction, then, $\hat{\pi}_i(\cdot)$ is a regular conditional distribution for

t_{-i} given t_i under the measure ν. Define a new function $\bar{\pi}_i(\cdot)$ from $[0,1]^2$ into $\mathcal{M}(T_{-i})$ by setting

$$\bar{\pi}_i(\cdot|\theta_i, s_i) \equiv \pi_i(\cdot|(1-s_i)\theta_i, 0). \tag{9}$$

for all $(\theta_i, s_i) \in [0,1]^2$. Because $\bar{\pi}_i(\cdot|\theta_i, s_i) \equiv \pi_i(\cdot|(1-s_i)\theta_i, 0)$ whenever $s_i = 0$ and, moreover, $\nu((\Theta_i \times \{s_i > 0\}) \times T_{-i}) = 0$, the function $\bar{\pi}_i(\cdot)$ is also a regular conditional distribution for t_{-i} given t_i under the measure ν.

The measure ν and the belief function $\bar{\pi}_i(\cdot)$ satisfy the BDP property for agent i in the sense of Heifetz and Neeman but *not* in the sense of Definition 3.1. If we set $\hat{T}_i = [0,1] \times \{0\}$, then, for the given prior ν, we have $\nu(\hat{T}_i \times T_{-i}) = 1$ and $\bar{\pi}_i(\cdot|\theta_i, s_i) \equiv \pi_i(\cdot|\theta_i, s_i)$ for $(\theta_i, s_i) \in \hat{T}_i$. Because $\pi_i(\cdot)$ satisfies the BDP condition (2) for all t_i and t_i' in T_i, it follows that $\bar{\pi}_i(\cdot)$ satisfies the BDP condition (2) for all t_i and t_i' in $\hat{T}_i$, as required for the Heifetz-Neeman definition of the BDP property. However, for every type vector $(\theta_i, s_i) \in [0,1]^2$, we have $\bar{\pi}_i(\theta_i, s_i) = \bar{\pi}_i(\theta_i', s_i')$ for all (θ_i', s_i') for which $(1-s_i')\theta_i' = (1-s_i)\theta_i$; For every type $(\theta_i, s_i) \in T_i$, there exist types $(\theta_i', s_i') \in T_i$ such that Condition (2) is violated by (θ_i, s_i) and (θ_i', s_i').

From a mechanism design perspective, the neglect of null sets in the Heifetz-Neeman definition of the BDP property is unsatisfactory. A mechanism designer is constrained by the condition that, for any $t_i = (\theta_i, s_i)$ in T_i and any $t_i' = (\theta_i', s_i') \in T_i$, there should be no incentive for agent i with type t_i to dissemble and act as if the type was t_i'. If the set $T_i \backslash \hat{T}_i$ has probability zero under the given prior, the mechanism designer need not much care about this incentive compatibility condition for $t_i \in T_i \backslash \hat{T}_i$. However, he must care about the fact that the presence of types in $T_i \backslash \hat{T}_i$ affects incentive compatibility conditions for all types in $\hat{T}_i$. In the example, *all* types in $\hat{T}_i$ can actually mimic types in $T_i \backslash \hat{T}_i$ that have the same beliefs and different payoffs. In particular, for any type $t_i = (\theta_i, s_i)$ with $\theta_i > 0$, there is another type $t_i' = (\theta_i', s_i')$ with θ_i' arbitrarily close to zero so that the belief function $\bar{\pi}_i$ assigns the same beliefs to both t_i and t_i'. Whereas the set of mimicked types has prior measure zero, the set of types that can mimic has full measure. If the mechanism designer respects incentive compatibility only for $t_i \in \hat{T}_i$ and $t_i' \in \hat{T}_i$, he may end up with a mechanism that satisfies incentive compatibility for t_i and t_i' in $\hat{T}_i$ but violates incentive compatibility for $t_i \in \hat{T}_i$ and $t_i' \in T_i \backslash \hat{T}_i$, i.e. it may violate incentive compatibility for a full set of types.

Underlying this discussion, there is a deeper problem: The incomplete-information model (1) takes the belief function $\hat{\pi}_i$ as a primitive, and we initially defined the BDP property for belief functions. In focussing on priors rather than belief functions, one must come to terms with the fact that a given prior is usually compatible with many belief function. Specifically, if $\hat{\pi}_i(\cdot) = b_i(\cdot|\nu)$ is a regular conditional distribution for t_{-i} given t_i under the prior ν, then any function $\bar{\pi}_i(\cdot)$ such that $\bar{\pi}_i(t_i) = \hat{\pi}_i(t_i)$ for ν-almost all $t_i \in T_i$ is also a regular conditional distribution for t_{-i} given t_i under the prior ν. Our example shows that it is quite possible for Definition 3.1 to be satisfied when the belief function is taken to be $\hat{\pi}_i(\cdot)$ but *not* when it is $\bar{\pi}_i(\cdot)$.

The Heifetz-Neeman definition of the BDP property for priors deals with this problem by eliminating a null set of types from the analysis. This implies that, if the BDP property in their definition holds for one regular conditional distribution, then it must hold for all regular conditional distributions under the given prior.

Definition 3.1 deals with the problem by restricting the analysis to priors under which the marginal type distributions have full support. The following lemma shows that, for such priors, the requirement that belief functions must be continuous eliminates the multiplicity of belief functions that are compatible with a given prior.[10]

LEMMA 3.3 *For $j = 1, ..., I$, let T_j be a complete separable metric space, and let $T := \prod_{i=1}^{I} T_i$. For any i, let $\hat{\pi}_i : T_i \to \mathcal{M}(T_{-i})$ be a continuous function and let $\nu_i \in \mathcal{M}(T)$ be such that (3) holds, i.e., ν_i is a prior for $\hat{\pi}_i$. If the support of the marginal distribution on agent i's types that is induced by F_i is equal to T_i, then $\hat{\pi}_i$ is the unique continuous regular conditional distribution for t_{-i} given t_i that is induced by T_i.*

Proof. If the lemma is false, there exists a prior $\nu_i \in \mathcal{M}(T)$ such that the support of the marginal distribution on agent i's types is equal to T_i and there exist two belief functions

[10]Without continuity, the multiplicity of regular conditional distributions for a given prior is even more vexing. For an example, suppose again that $\Theta_i = [0,1]$, $S_i = [0,1]$, hence $T_i = [0,1]^2$. For a given prior ν_i, let $b_i(\cdot|\cdot, \nu_i)$ be a regular conditional distribution for t_{-i} given $t_i = (\theta_i, s_i)$ such that the BDP property in the sense of Definition 3.1 is satisfied. Construct another regular conditional distribution for t_{-i} given t_i by setting

$$\hat{b}_i(\cdot|\theta_i, s_i, \nu_i) = b_i(\cdot|\theta_i, s_i, \nu_i) \quad \text{if } \theta_i \in (0,1],\ s_i \in [0,1]$$

and

$$\hat{b}_i(\cdot|\theta_i, s_i, \nu_i) = b_i(\cdot|\psi(s_i), \nu_i) \quad \text{if } \theta_i = 0,\ s_i \in [0,1],$$

where $\psi : [0,1] \to [0,1]^2$ is Peano's space-filling function. If ν_i assigns measure zero to the event $\theta_i = 0$, we have $\hat{b}_i(t_i, \nu) = b_i(t_i, \nu)$ for ν_i-almost all (θ_i, s_i), implying that $\hat{b}_i(\cdot, \nu_i)$ is indeed another regular conditional probability distribution for t_{-i} given t_i under ν_i. However, for every type vector $(\theta_i, s_i) \in (0,1] \times [0,1]$ for agent i, there is $s_i' \in [0,1]$ such that $\hat{b}_i(\theta_i, s_i, \nu_i) = \hat{b}_i(0, s_i', \nu_i)$.

$b_i^1(\cdot,\nu_i)$,$b_i^2(\cdot,\nu_i)$ be such that $b_i^1(\cdot,\nu_i)$ and $b_i^2(\cdot,\nu_i)$ are both continuous and both regular conditional distributions for t_{-i} given t_i under ν_i. If $b_i^1(\cdot,\nu_i) \neq b_i^2(\cdot,\nu_i)$, there exists $t_i \in T_i$ such that $b_i^1(t_i,\nu_i) \neq b_i^2(t_i,\nu_i)$. Because $b_i^1(\cdot,\nu_i)$ and $b_i^2(\cdot,\nu_i)$ are both regular conditional distributions for t_{-i} given t_i under ν_i, it must be the case that $\nu_i(\{t_i\} \times T_{-i}) = 0$. Because the support of the marginal distribution on agent i's types is equal to T_i, it must be the case that, for any $\varepsilon > 0$, $\nu_i(B_\varepsilon(t_i) \times T_{-i}) > 0$, where $B_\varepsilon(t_i)$ is an ε-neighbourhood of t_i. However, because $b_i^1(\cdot,\nu_i)$ and $b_i^2(\cdot,\nu_i)$ are both continuous, $b_i^1(t_i,\nu_i) \neq b_i^2(t_i,\nu_i)$ implies $b_i^1(t_i',\nu_i) \neq b_i^2(t_i',\nu_i)$ for all $t_i' \in B_\varepsilon(t_i)$ if $\varepsilon > 0$ is sufficiently small. But then, $b_i^1(t_i',\nu_i) \neq b_i^2(t_i',\nu_i)$ for all t_i' in a set that has positive measure under $\nu_i(\cdot \times T_{-i})$. This is incompatible with the assumption that $b_i^1(\cdot,\nu_i)$ and $b_i^2(\cdot,\nu_i)$ are both regular conditional distributions for t_{-i} given t_i under ν_i. ■

On the domain of Definitions 3.1 there is a one-to-one relationship between priors and belief functions. Moreover, a prior has the BDP property for agent i in the sense of definition 3.1 if and only if the belief function has the BDP property in the sense of the original definition.

3.2 Examples

We illustrate the BDP property by several examples.

EXAMPLE 3.4 *Let $I = 2$, $\Theta_1 = S_2 = \mathbb{R}$, $\Theta_2 = S_1 = \{0\}$, and suppose that ν_1 is a multivariate normal distribution. Then agent 1's conditional expectations satisfy*

$$E_1[s_2|\theta_1, s_1] = \frac{cov(s_2,\theta_1)}{var\theta_1}(\theta_1 - E_1\theta_1) + E_1 s_2. \tag{10}$$

If $cov(s_2,\theta_1) \neq 0$, i.e., if the signal s_2 contains any information about θ_1, one can infer θ_1 from the belief variable $E_1[s_2|\theta_1, s_1]$ and the parameters $E_1\theta_1$, $E_1 s_2$, $cov(s_2,\theta_1)$, $var\theta_1$ of the prior ν_1. Thus, the BDP property holds for agent 1 unless the signal s_2 is uncorrelated with the payoff type θ_1. Within the set of models covered by Example 3.4, BDP is generic.

In this example, agent 2 receives a signal about agent 1's payoff type. Knowing this, agent 1 treats his own payoff type as a signal about agent 2's signal. Therefore, his belief about agent 2's signal varies with his payoff type. The relation is monotonic, and his payoff type can be inferred from his belief type.

The following examples concern the possibility that BDP may fail because of the confounding influence of another information variable. In Example 3.5, confounding occurs, in Example 3.6, confounding is neutralized by the presence of another dimension of beliefs.

EXAMPLE 3.5 *Let $I = 2$, $\Theta_1 = S_1 = S_2 = \mathbb{R}$, $\Theta_2 = \{0\}$, and suppose that ν_1 is a multivariate normal distribution. Then agent 1's conditional expectations satisfy*

$$E_1[s_2|\theta_1, s_1] = \alpha_\theta(\theta_1 - E_1\theta_1) + \alpha_s(s_1 - E_1 s_1) + E_1 s_2, \tag{11}$$

where

$$\begin{pmatrix} \alpha_\theta & \alpha_s \end{pmatrix} = \begin{pmatrix} cov(s_2, \theta_1) & cov(s_2, s_1) \end{pmatrix} \begin{pmatrix} var\theta_1 & cov(\theta_1, s_1) \\ cov(\theta_1, s_1) & vars_1 \end{pmatrix}^{-1}. \tag{12}$$

In this case, as in Example 3.4, agent 1's belief about agent 2's signal is affected by agent 1's payoff type unless $cov(s_2, \theta_1) = 0$. However, if agent 1's own signal is also correlated with s_2, it is not possible to infer θ_1 from the belief variable $E_1[s_2|\theta_1, s_1]$ and the parameters of the prior ν_1. For such an inference, one would also have to know the realization of agent 1's own signal. In this setting, for any prior on T that is multivariate normal, the BDP property fails to hold for agent 1 except in the negligible case where $cov(s_2, s_1) = 0$.

EXAMPLE 3.6 *Let $I = 2$, $\Theta_1 = S_1 = \mathbb{R}$, $\Theta_2 = \{0\}$, $S_2 = \mathbb{R}^2$, and suppose that ν_1 is a multivariate normal distribution. Then agent 1's conditional expectations satisfy*

$$\begin{pmatrix} E_1[s_2^1|\theta_1, s_1] \\ E_1[s_2^2|\theta_1, s_1] \end{pmatrix} = A \begin{pmatrix} \theta_1 - E_1\theta_1 \\ s_1 - E_1 s_1 \end{pmatrix} + \begin{pmatrix} E_1 s_2^1 \\ E_1 s_2^2 \end{pmatrix}, \tag{13}$$

where $A = \Sigma_{21}\Sigma_{11}^{-1}$, with Σ_{21}, Σ_{11} as submatrices of the variance-covariance matrix Σ of $s_2^1, s_2^2, \theta_1, s_1$, partitioned so as to reflect the distinction between the variables that agent 1 observes and the variables that he does not observe. In this specification, both θ_1 and s_1 can be inferred from the conditional expectations $E_1[s_2^1|\theta_1, s_1]$ and $E_1[s_2^2|\theta_1, s_1]$ whenever the matrix A is invertible, i.e., whenever the matrix Σ_{21} is nonsingular. Because the set of nonsingular two-by-two matrices is open and dense in the set of all two-by-two matrices, the BDP property is generic for agent 1 within the set of models covered by this example.

Whereas Example 3.5 involves a failure of BDP due to a confounding of influences of different information variables, the following Example 3.7 shows that BDP will also fail if, for given parameters of the prior ν_1, the map from payoff types to conditional expectations is not monotonic (not one-to-one). Subsequently, Example 3.8 will show that this problem is likely to disappear if there are more variables about which to form expectations so that the vector of conditional expectations has a sufficiently high dimension.

EXAMPLE 3.7 *Let $I = 2, \Theta_1 = S_2 = \mathbb{R}$, $\Theta_2 = S_1 = \{0\}$, and suppose that ν_1 is the distribution that is generated when*

$$s_2 = \theta_1^2 + \varepsilon, \tag{14}$$

where θ_1 and ε are independent normal random variables. In this case, agent 1's conditional distribution for s_2 given θ_1 is normal with mean

$$E_1[s_2|\theta_1, s_1] = (\theta_1^2 - E_1\theta_1^2) + E_1 s_2 \tag{15}$$

and variance $Var\varepsilon$. From the belief $E[s_2|\theta_1, s_1]$, one can infer θ_1^2, but one cannot tell whether it is the positive or the negative solution of the equation

$$\theta_1^2 = E_1[s_2|\theta_1, s_1] - E_1 s_2 + E_1\theta_1^2. \tag{16}$$

The BDP property fails to hold.

EXAMPLE 3.8 *Let $I = 2$, $\Theta_1 = \mathbb{R}$, $\Theta_2 = S_1 = \{0\}$, $S_2 = \mathbb{R}^2$, and suppose that ν_1 is the distribution that is generated when*

$$s_2^1 = \theta_1^2 + \varepsilon, \tag{17}$$

$$s_2^2 = A\theta_1 + B\theta_1^2 + \eta \tag{18}$$

where θ_1,ε, and η are independent normal random variables and A and B are constants. In this case, agent 1's conditional distribution for s_2^1 and s_2^2 given θ_1 is normal with means

$$E_1[s_2^1|\theta_1, s_1] = (\theta_1^2 - E_1\theta_1^2) + E_1 s_2^1, \tag{19}$$

$$E_1[s_2^2|\theta_1, s_1] = A(\theta_1 - E_1\theta_1) + B(\theta_1^2 - E_1\theta_1^2) + E_1 s_2^2 \tag{20}$$

and variance-covariance matrix $\begin{pmatrix} Var\varepsilon & 0 \\ 0 & Var\eta \end{pmatrix}$. If $A \neq 0$, then from (19) and (20), one obtains

$$\theta_1 = E_1\theta_1 + \frac{1}{A}\left[E_1[s_2^2|\theta_1, s_1] - E_1 s_2^2 - B(E_1[s_2^1|\theta_1, s_1] - E_1 s_2^1\right], \tag{21}$$

which shows that θ_1 can be inferred by looking at $E_1[s_2^1|\theta_1, s_1]$ and $E_1[s_2^2|\theta_1, s_1]$ jointly. By looking at the two belief variables together, one overcomes the difficulty that neither belief variable alone is injective in θ_1. The BDP property holds for agent 1 except in the negligible case $A = 0$.

4 Genericity Results

4.1 The BDP Property with Finite Type Sets

Turning from these examples to a more general analysis, we first consider the case where the type sets T_i are finite. If each type set T_i is finite, with n_i elements, the space space $T := \prod_{i=1}^{I} T_i$ is also finite, with $N = \prod_{i=1}^{I} n_i$ elements, and the prior F is represented by a vector $\Pi \in \mathbb{R}^N$, such that $\sum_{k=1}^{N} \Pi_k = 1$. The set of such vectors is endowed with the usual (Euclidean) topology.

PROPOSITION 4.1 *Assume that, for each i, T_i is a finite set with n_i distinct elements. For any i, let*

$$N_{-i} := \prod_{\substack{j=1 \\ j \neq i}}^{I} n_j \tag{22}$$

be the cardinality of the set T_{-i}. If $n_i \leq N_{-i}$ for all i, then, for each i, the set P_i of priors that exhibit the BDP property for agent i is open and dense in the set of all priors on T.

Proof. Fixing i, we note that any N vector Π^N of probabilities on T can be written in matrix form as $\Pi(i) = (\pi_{t_i t_{-i}})$ where the different rows refer to different types t_i of agent i and different columns refer to the different elements t_{-i} of T_{-i}. The belief $b_i(\cdot|t_i)$ of agent i, conditional on t_i, about the other agents' types is represented by a vector of conditional probabilities on T_{-i}. Bayes' Law implies that, under the prior Π, this vector is proportional to the vector $(\pi_{t_i t_{-i}})$; one can write:

$$b_i(t_{-i}|t_i) = \lambda(t_i) \times \pi_{t_i t_{-i}} \tag{23}$$

for all $t_{-i} \in T_{-i}$, where

$$\lambda(t_i) = \frac{1}{\sum_{t_{-i} \in T_{-i}} \pi_{t_i t_{-i}}} \tag{24}$$

is chosen to ensure that the entries in (23) sum to one. Equations (23) and (24) imply that, if the rows of the matrix $\Pi(i)$ are linearly independent, as well as strictly positive, then so are the belief vectors $b_i(\cdot|t_i), t_i \in T_i$. This implies, in particular, that the belief vectors $b_i(\cdot|t_i), t_i \in T_i$, are all distinct and the function $t_i \rightarrow b_i(\cdot|t_i)$ is invertible, i.e. one can infer the type t_i of agent i from his belief vector. Given that $t_i = (\theta_i, s_i)$, this means, in particular, that one can infer θ_i from $b_i(\cdot|t_i)$. The proposition follows because, by standard arguments, for $n_i \leq N_{-i}$, the set of $n_i \times N_{-i}$ matrices with linearly independent rows is an open and dense subset P_i of $\mathbb{R}^{n_i} \times \mathbb{R}^{N_{-i}} = \mathbb{R}^N$. ■

COROLLARY 4.2 *Under the assumptions of Proposition 4.1, the set of common priors exhibiting the BDP property is open and dense in the set of all priors on T.*

Proof. Proposition 4.1 implies that, for each i, the set P_i of priors that satisfy BDP for agent i is open and dense in the set of all priors on T. Because I is finite, the intersection $P = \cap_{i=1}^{I} P_i$ of these open and dense sets for the different agents i is still open and dense. The corollary follows by observing that a common prior satisfies BDP for all agents i if and only if it belongs to P. ■

The assumption that $n_i \leq N_{-i}$ ensures that the set of things about which agent i forms his beliefs is larger than the set of things on which he conditions. When both sets are finite, it follows that, generically, the map from types to beliefs is one-to-one. Indeed, the belief vectors of different types are linearly independent, generically. This latter property is sufficient for surplus extraction.[11]

4.2 The BDP Property for Agent i With a Continuum of Types

We next allow for a continuum of types of each agent. We assume that, for each i, there is a positive integer n_i such that the type set T_i of agent i is a subset of $\mathbb{R}^{n_i}$. The space $T := \prod_{i=1}^{I} T_i$ of vectors of all agents' types is a subset of $\mathbb{R}^N$, where $N := \sum_{j=1}^{I} n_j$. The space T_{-i} of vectors of types of agents other than i is a subset of $\mathbb{R}^{N_{-i}}$, where $N_{-i} = N - n_i$. The sets T_i, $i = 1, ..., I$, and therefore also T and $T_{-i}, i = 1, ..., I$, are assumed to be compact.

In this setting, a prior is a probability measure on T. Given a prior $\nu \in \mathcal{M}(T)$, we write $\bar{\nu}_i(\nu)$ for the marginal distribution on T_i that is induced by ν. A regular conditional distribution for t_{-i} given t_i, is a function from T_i into $\mathcal{M}(T_{-i})$. We endow the range of this function with the topology of weak convergence of probability measures, i.e. the weak* topology.

For reasons discussed in Section 3.1, we restrict our attention to the subset $\mathcal{N}_i^c(T)$ of priors on T such that, for $\nu \in \mathcal{N}_i^c(T)$, the support of the marginal distribution $\bar{\nu}_i(\nu)$ is T_i and, moreover, there exists a continuous function $b_i(\cdot, \nu)$ from T_i into $\mathcal{M}(T_{-i})$ that is a

[11] For a discussion of the role of linear independence, see Crémer and McLean (1988), McAfee and Reny (1992).

regular conditional distribution for t_{-i} given t_i. By Lemma 3.3, this restriction eliminates the indeterminacy in the relation between priors and belief functions that is due to the fact that regular conditional distributions can be arbitrarily modified on null sets. For $\nu \in \mathcal{N}_i^c(T)$, there is therefore no ambiguity in Definition 3.1 saying that ν has the BDP property for agent i if and only if the associated continuous belief function $b_i(\cdot, \nu)$ has the BDP property.

The assessment of genericity, robustness, or negligibility of the BDP property in the space $\mathcal{N}_i^c(T)$ depends on the topology that is imposed on this space. Because we are interested in the behaviour of conditional distributions, it would be inappropriate to simply specify the space of priors as a subspace of $\mathcal{M}(\mathbb{R}^N)$ endowed with the topology of weak convergence of probability measures, i.e. the weak* topology. The topology of weak convergence fails to control for the behaviour of conditional distributions. If a sequence of measures ν^k converges weakly to a limit ν, the regular conditional distributions can exhibit a discontinuity in the limit.[12] Such a discontinuity in the information that is available to the agent is likely to induce a discontinuity in his behaviour.

To avoid such discontinuities, we want to specify the topology on $\mathcal{N}_i^c(T)$ so that the map from priors to joint distributions of types and beliefs is continuous.[13] If payoff functions are also continuous, this condition ensures that behaviour correspondences are upper hemi-continuous in types. Continuity of the map from priors to joint distributions of types and beliefs ensures that, if payoff functions are continuous, then the map from priors to joint distributions of types and actions is upper hemi-continuous.

Because belief functions are continuous, continuity of the map from priors to joint distributions of types and beliefs is obtained if the maps $\nu \to \bar{\nu}_i(\nu)$ and $\nu \to b_i(\cdot, \nu)$ from priors to type distributions and from priors to belief functions are continuous. For any $\nu \in \mathcal{N}_i^c(T)$, the marginal distribution $\bar{\nu}_i(\nu)$ is an element of the space $\mathcal{M}(T_i)$ of probability measures on T_i and the induced belief function $b_i(\cdot, \nu)$ is an element of the space $C(T_i, \mathcal{M}(T_{-i}))$ of continuous functions from T_i into $\mathcal{M}(T_{-i})$. Endowing $\mathcal{M}(T_i)$ with the

[12] Examples are given by Jordan (1977) and Hellwig (1996). In the present setting, let $I = 2$, $n_1 = n_2 = 1$. For $k = 1, 2, ...$, let ν^k be such that t_1 is uniformly distributed on $[0, 1]$ and $t_2 = \frac{1}{2}[1 + \sin(kt_1)]$. The sequence $\{\nu^k\}$ converges weakly to the uniform distribution on $[0, 1]^2$. For any k, agent 1 is able to perfectly infer the type of agent 2 from his own type, but, in the limit, observation of his own type provides him with no information at all about the type of agent 2.

[13] This is the continuity property required in Jordan (1977). As mentioned above, with continuous payoff functions, this continuity property implies that the map from priors to joint distributions of types and actions is upper hemi-continuous. For details, see Jordan (1977), Hellwig (1996).

topology of weak convergence and $C(T_i, \mathcal{M}(T_{-i}))$ with the topology of uniform convergence, we specify the topology on $\mathcal{N}_i^c(T)$ to be the coarsest topology under which the mapping φ_i from $\mathcal{N}_i^c(T)$ to $\mathcal{M}(T_i) \times C(T_i, \mathcal{M}(T_{-i}))$ that is defined by the formula

$$\varphi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu)) \tag{25}$$

is continuous.

Given this topology, the following results establish the residualness of the BDP property in $\mathcal{N}_i^c(T)$. The first result is an extension of the standard embedding theorem for continuous functions to the case where the range of the functions is a space of measure.

PROPOSITION 4.3 *Let T_i, T_{-i} be compact subsets of $\mathbb{R}^{n_i}$ and $\mathbb{R}^{N_{-i}}$. If T_{-i} has infinitely many elements, the set $\mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ of embeddings of T_i in $\mathcal{M}(T_{-i})$ is a countable intersection of open and dense subsets of $C(T_i, \mathcal{M}(T_{-i}))$.*

Proposition 4.3 contains the main insight of our analysis: If T_{-i} has infinitely many elements, the space $\mathcal{M}(T_{-i})$ of agent i's beliefs is an infinite-dimensional space. If agent i's type set is finite-dimensional, the set of belief functions that are embeddings is a residual set. Because embeddings are injective, any belief function that is an embedding necessarily has the BDP property: for any such function $b_i(\cdot, \nu)$ and any $\bar{b}_i \in \mathcal{M}(T_{-i})$ there is at most one $t_i = (\theta_i, s_i)$ such that $b_i(t_i, \nu) = \bar{b}_i$. For the proof of Proposition 4.3, the reader is referred to Appendix A.1

PROPOSITION 4.4 *Let T_i, T_{-i} be compact subsets of $\mathbb{R}^{n_i}$ and $\mathbb{R}^{N_{-i}}$ and let $T := T_i \times T_{-i}$. Let $\mathcal{N}_i^c(T)$ be endowed with the coarsest topology under which the mapping $\nu \to \varphi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu))$ is continuous. If $T_{-i} \subset \mathbb{R}^{N_{-i}}$ has infinitely many elements, then the set $\mathcal{N}_i^*(T)$ of priors in $\mathcal{N}_i^c(T)$ that have the BDP property for agent i is a residual subset, i.e., $\mathcal{N}_i^*(T)$ contains a countable intersection of open and dense subsets of $\mathcal{N}_i^c(T)$.*

Proof. Let $\mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ be the set of embeddings of T_i in $\mathcal{M}(T_{-i})$ and note that any measure ν with $b_i(\cdot, \nu) \in \mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ has the BDP property. It therefore suffices to show that the set $\varphi_i^{-1}(\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i})))$ is a residual subset of $\mathcal{N}_i^c(T)$.

By Proposition 4.3, $\mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ is a residual subset of $C(T_i, \mathcal{M}(T_{-i}))$. Trivially then, $\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ is a residual subset of the space $\mathcal{M}(T_i) \times C(T_i, \mathcal{M}(T_{-i}))$. By Lemma A.2 in the Appendix, residualness of $\varphi_i^{-1}(\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i})))$ in $\mathcal{N}_i^c(T)$ follows if the mapping φ_i is open as well as continuous.

Because the topology on $\mathcal{N}_i^c(T)$ is the coarsest topology under which φ_i is continuous, the open sets in $\mathcal{N}_i^c(T)$ all take the form $V = \varphi_i^{-1}(U)$ where U is an open subset of $\mathcal{M}(T_i) \times C(T_i, \mathcal{M}(T_{-i}))$. For any open set $V \subset \mathcal{N}_i^c(T)$, therefore, the set $\varphi_i(V)$ is open, i.e., φ_i maps open sets into open sets. Thus, φ_i is open as well as continuous. By Proposition 4.3 and by Lemma A.2 in the Appendix, therefore, $\varphi_i^{-1}(\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i})))$ is a residual subset of $\mathcal{N}_i^c(T)$. ■

4.3 The BDP Property for a Common Prior With a Continuum of Types

With finite type sets, the genericity of the BDP property for common priors was a simple corollary of the genericity of the BDP property for any single agent. With a continuum of types, the matter is slightly more complicated. The reason is that we now require continuity of belief functions for all agents, so the relevant set of measures is the set $\mathcal{N}^c(T) := \cap_{k=1}^I \mathcal{N}_k^c(T)$ of priors on T that admit continuous regular conditional distributions for t_{-i} given t_i for *all* agents i. An assessment of the set of priors on T that have the BDP property for all agents simultaneously requires a consideration of the intersections $\cap_{i=1}^I \mathcal{N}_i^*(T)$, as a subset of $\mathcal{N}^c(T)$. Proposition 4.4 establishes residualness of the set $\mathcal{N}_i^*(T)$ of priors having the BDP property for agent i in the set $\mathcal{N}_i^c(T)$ of priors on T that admit continuous regular conditional distributions for t_{-i} given t_i. This does not, however, permit any inference about the relation between $\mathcal{N}_i^*(T)$ and the intersection $\mathcal{N}_i^c(T) = \cap_{k=1}^I \mathcal{N}_k^c(T)$. There is no guarantee that the perturbations of belief functions that serve to establish denseness of $\mathcal{N}_i^*(T)$ in $\mathcal{N}_i^c(T)$ can be chosen to lie in the intersection $\cap_{k=1}^I \mathcal{N}_k^c(T)$, as would be required to establish denseness of $\mathcal{N}_i^*(T)$ in $\cap_{k=1}^I \mathcal{N}_k^c(T)$.

To avoid this difficulty, we go back to the residualness of BDP belief functions that was established in Proposition 4.3. From this result, we derive the residualness of common priors with the BDP property in $\mathcal{N}^c(T)$ by the same argument that we used in the proof of Proposition 4.4. In the following proposition, as before, the spaces $\mathcal{M}(T_i)$ and $\mathcal{M}(T_{-i}))$ have the topology of weak convergence, and $C(T_i, \mathcal{M}(T_{-i}))$ has the topology of uniform convergence.

PROPOSITION 4.5 *For $i = 1, ..., I$, let T_i be a compact subset of $\mathbb{R}^{n_i}$, and assume that $\mathcal{N}^c(T) := \cap_{k=1}^I \mathcal{N}_k^c(T)$ is nonempty, where, as before, $T = \prod_{j=1}^I T_j$. Let $\mathcal{N}^c(T)$ be endowed with the coarsest* *topology under which the mapping*

$$\nu \to \varphi(\nu) := (\varphi_1(\nu), ..., \varphi_I(\nu)), \tag{26}$$

with $\varphi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu))$, $i = 1, ..., I$, *from* $\mathcal{N}^c(T)$ *to* $\prod_{i=1}^{I}[\mathcal{M}(T_i) \times C(T_i, \mathcal{M}(T_{-i}))]$ *is continuous. If the sets* T_i, $i = 1, ...I$, *have infinitely many elements, then the set* $\mathcal{N}^*(T)$ *of elements of* $\mathcal{N}^c(T)$ *that have the BDP property for all agents is a residual subset, i.e.,* $\mathcal{N}^*(T)$ *contains a countable intersection of open and dense subsets of* $\mathcal{N}^c(T)$.

Proof. Because the set $\mathcal{N}^*(T)$ of priors in $\mathcal{N}^c(T)$ that have the BDP property for all agents is equal to the intersection $\cap_{i=1}^{I}\mathcal{N}_i^*(T)$ of the sets $\mathcal{N}_i^*(T)$ of priors in $\mathcal{N}^c(T)$ that have the BDP property for agents $i = 1, ..., I$, it suffices to show that each of the sets $N_i^*(T)$ contains a countable intersection of open and dense subsets of $\mathcal{N}^c(T)$.

The proof of this latter claim is similar to the proof of Proposition 4.4. Because any measure ν with $b_i(\cdot|\cdot, \nu) \in \mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ has the BDP property, it suffices to show that the set $\varphi_i^{-1}(\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i})))$ is a residual subset of $\mathcal{N}^c(T)$. By Proposition 4.3, $\mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ is a residual subset of $C(T_i, \mathcal{M}(T_{-i}))$. Therefore, $\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i}))$ is a residual subset of $\mathcal{M}(T_i) \times C(T_i, \mathcal{M}(T_{-i}))$. Residualness of $\varphi_i^{-1}(\mathcal{M}(T_i) \times \mathcal{E}(T_i, \mathcal{M}(T_{-i})))$ in $\mathcal{N}^c(T)$ follows if the mapping φ_i is open as well as continuous.

By the same argument that was given for the mapping φ_i in the proof of Proposition 4.4, the mapping $\varphi = (\varphi_1, ..., \varphi_I)$ is a continuous and open mapping from $\mathcal{N}^c(T)$ to $\prod_{i=1}^{I}[\mathcal{M}(T_i) \times C(T_i, \mathcal{M}(T_{-i}))]$. Because the projection is also continuous and open, it follows that φ_i is also open and continuous. Residualness of $\mathcal{N}_i^*(T)$ in $\mathcal{N}^c(T)$ follows immediately. ■

4.4 The BDP Property for Common Priors with Continuous Densities

In the formulation of Proposition 4.5, it is somewhat unsatisfactory that nonemptyness of $\mathcal{N}^c(T)$ is assumed and that the topology on $\mathcal{N}^c(T)$ is not specified explicitly. Both these weaknesses are eliminated if the analysis is restricted to probability measures with continuous, bounded densities. We now assume that, for each i, the type set T_i for agent i is compact and is the closure of an open set $\mathring{T}_i$. As before, we assume that T_i is a subset of $\mathbb{R}^{n_i}$, and we write T and T_{-i} for the products $\prod_{j=1}^{I} T_j$ and $\prod_{j=1, j\neq i}^{I} T_j$. Given the set T, we consider the set $\mathcal{M}^d(T)$ of measures ν such that, for some continuous and bounded function f^ν : $T \to \mathbb{R}_+$, we have

$$\nu(B) = \int_{B\cap T} f^\nu(t)dt \tag{27}$$

for any measurable $B \subset T$. For any $\nu \in \mathcal{M}^d(T)$ and any $t_i \in T_i$, set

$$\bar{f}_i^\nu(t_i) = \int_{T_{-i}} f^\nu(t_i, t_{-i})dt_{-i}. \tag{28}$$

Because f^ν is continuous and bounded on T and because the set T_{-i} is bounded, the integral in (28) is well defined for all $t_i \in T_i$. Moreover, $\bar{f}_i^\nu(t_i)$ depends continuously on t_i. The function $\bar{f}_i^\nu(\cdot)$ from T_i to $\mathbb{R}_+$ that is defined by (28) is the density function of the *marginal* distribution $\bar{\nu}_i(\nu)$ that is induced by ν. We write $\mathcal{M}_+^d(T)$ for the set of measures $\nu \in \mathcal{M}^d(T)$ such that $\bar{f}_i^\nu(t_i) > 0$ for all i and all $t_i \in T_i$.

For $\nu \in \mathcal{M}_+^d(T)$ and $t_i \in T_i$, we can then define a density function $\beta_i(\cdot|t_i,\nu)$ on T_{-i} by setting

$$\beta_i(t_{-i}|t_i,\nu) := \frac{f^\nu(t_i,t_{-i})}{\bar{f}_i^\nu(t_i)} \tag{29}$$

for $t_{-i} \in T_{-i}$. With this density function, we associate the probability distribution $b_i(\cdot|t_i,\nu)$ such that

$$b_i(B_{-i}|t_i,\nu) := \int_{B_{-i}\cap \mathring{T}_{-i}} \beta_i(t_{-i}|t_i,\nu)dt_{-i} \tag{30}$$

for any measurable set $B_{-i} \subset \mathbb{R}^{N_{-i}}$. The function $b_i(\cdot|\cdot,\nu)$ is obviously a regular conditional probability distribution for t_{-i} given t_i.

By inspection of (29), the density $\beta_i(t_{-i}|t_i,\nu)$ depends continuously on $t_i \in T_i$ and $t_{-i} \in T_{-i}$. By standard arguments, again using Lebesgue's bounded-convergence theorem, it follows that the function $t_i \to b_i(\cdot|t_i,\nu)$ that is given by (30) maps the domain T_i of the marginal density $\bar{f}_i^\nu$ continuously into $\mathcal{M}(T_{-i})$. By the same argument as in Lemma 3.3, this is the *only* regular conditional distribution for t_{-i} given t_i that maps T_i continuously into $\mathcal{M}(T_{-i})$.

By construction, we have $\mathcal{M}_+^d(T) \subset \mathcal{N}^c(T)$. The continuity property of the regular conditional distribution here is actually stronger than in the preceding analysis. For $\nu \in \mathcal{M}_+^d(T)$, the belief function $b_i(\cdot|\cdot,\nu)$ actually takes values in the subspace $\mathcal{M}^d(T_{-i})$ of $\mathcal{M}(T_{-i})$ consisting of those measures that have densities that are continuous on T_{-i}. If we endow $\mathcal{M}^d(T_{-i})$ with the topology that is induced by the uniform topology for density functions, we find that the function $b_i(\cdot|\cdot,\nu)$ maps the set T_i continuously into $\mathcal{M}^d(T_{-i})$, i.e., that $b_i(\cdot|\cdot,\nu)$ is an element of the space $C(T_i, \mathcal{M}^d(T_{-i}))$ of continuous functions from T_i into $\mathcal{M}^d(T_{-i})$. If we endow $C(T_i, \mathcal{M}^d(T_{-i}))$ with the uniform topology, we obtain the following analogue of Proposition 4.3.

PROPOSITION 4.6 *For $i = 1, ..., I$, let T_i be a compact subset of $\mathbb{R}^{n_i}$ and let $T_{-i} := \prod_{\substack{j=1 \\ j\neq i}}^{I} T_j$ and $T = \prod_{j=1}^{I} T_j$. Assume that the sets T_i have nonempty interiors $\mathring{T}_i$. Then for any i, the set $\mathcal{E}(T_i, \mathcal{M}^d(T_{-i}))$ of embeddings of T_i in $\mathcal{M}^d(T_{-i})$ is a residual subset, i.e., it contains a countable intersection of open and dense subsets of $C(T_i, \mathcal{M}^d(T_{-i}))$.*

For a proof of Proposition 4.6, the reader is referred to Appendix A.2.

To translate this result into a proposition about priors, we endow the space $\mathcal{M}^d_+(T)$ with the coarsest topology under which the maps

$$\nu \to \psi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu)), \tag{31}$$

from distributions on T into marginal distributions and belief functions, are continuous. In this context, the range $\mathcal{M}^d_+(T_i)$ of the function $\nu \to \bar{\nu}_i(\nu)$ is taken to have the topology that is induced by the uniform topology for density functions, the range $C(T_i, \mathcal{M}^d(T_{-i}))$ of the function $\nu \to b_i(\cdot, \nu)$ the topology of uniform convergence.

PROPOSITION 4.7 *For $i = 1, ..., I$, let T_i be a compact subset of $\mathbb{R}^{n_i}$ and let $T_{-i} := \prod_{\substack{j=1 \\ j\neq i}}^{I} T_j$ and $T = \prod_{j=1}^{I} T_j$. Assume that the sets T_i have nonempty interiors $\mathring{T}_i$. Let $\mathcal{M}^d_+(T)$ is endowed with the coarsest topology under which the mapping*

$$\nu \to \psi(\nu) := (\psi_1(\nu), ..., \psi_I(\nu)), \tag{32}$$

with $\psi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu))$, from $\mathcal{M}^d_+(T)$ to $\prod_{i=1}^{I}[\mathcal{M}^d_+(T_i) \times C(T_i, \mathcal{M}^d(T_{-i}))]$ is continuous. Then the set $\mathcal{N}^{**}(T)$ of priors in $\mathcal{M}^d_+(T)$ that have the BDP property for all agents is a residual subset of $\mathcal{M}^d_+(T)$.

Proposition 4.7 follows from Proposition 4.6 by the same argument by which Proposition 4.5 was derived from Proposition 4.3. The details are left to the reader.

COROLLARY 4.8 *For $i = 1, ..., I$, let T_i be a compact subset of $\mathbb{R}^{n_i}$ and let $T = \prod_{j=1}^{I} T_j$. Assume that the sets T_i have nonempty interiors $\mathring{T}_i$. If $\mathcal{M}_+^d(T)$ is endowed with the topology that is induced by the uniform topology for density functions, then the set $\mathcal{N}^{**}(T)$ of priors on T that have the BDP property for all agents is a residual subset of $\mathcal{M}_+^d(T)$.*

Corollary 4.8 follows immediately from Proposition 4.7 and the following lemma, the proof of which is given in Appendix A.3.

LEMMA 4.9 *The topology that is induced by the topology of uniform convergence of density functions is the coarsest topology under which the mapping $\nu \to \psi(\nu)$ in Proposition 4.7 is continuous.*

5 Relation to the Literature

The thrust of our results runs counter to that of Heifetz and Neeman (2006) and Barelli (2009), and parallels Chen and Xiong (2011). It is therefore appropriate to discuss the relation of our analysis to theirs. Heifetz and Neeman (2006) consider families $\{\mathcal{T}^k\}_{k \in K}$ of incomplete-information models of the form

$$\mathcal{T}^k = \{T_i^k, \hat{\theta}_i^k, \hat{\pi}_i^k\}_{i=1}^I \tag{33}$$

where, for any k in some index set K, for any i, T_i^k is a set of abstract "types" for agent i, $\hat{\theta}_i^k$ is a mapping from T_i^k into a set Θ_i^k of payoff parameter vectors for agent i, and $\hat{\pi}_i^k$ is a continuous mapping from T_i^k into the set $\mathcal{M}\left(T_{-i}^k\right)$ of probability distributions on the space T_{-i}^k of the other agents' abstract types. They restrict their attention to incomplete-information models that are consistent with common priors and study the genericity of the BDP property in the set

$$\mathcal{P} \subset \mathcal{M}\left(\prod_{i=1}^{I} \bigcup_{k \in K} T_i^k\right) \tag{34}$$

such that $F \in \mathcal{P}$ if and only if F is a common prior for one model T^k in the family that is being considered. Under the assumption that the family $\{\mathcal{T}^k\}_{k \in K}$ is "closed under finite unions", they show that the set $\mathcal{P}$ is convex: If F^{k^1} and F^{k^2} are common priors for the incomplete-information models $\mathcal{T}^{k^1}, \mathcal{T}^{k^2}$, then, for any $\alpha \in [0, 1]$, the measure

$$\alpha F^{k^1} + (1 - \alpha) F^{k^2} \in \mathcal{M}\left(\prod_{i=1}^{I} (T_i^{k^1} \cup T_i^{k^2})\right) \tag{35}$$

is a common prior for the model $\mathcal{T}^{\hat{k}} = \{(T_i^{k^1} \cup T_i^{k^2}), \hat{\theta}_i^{\hat{k}}, \hat{\pi}_i^{\hat{k}}\}_{i=1}^{I}$, where $\mathcal{T}^{\hat{k}}$ is the model that corresponds to the "union" of $\mathcal{T}^{k^1}$ and $\mathcal{T}^{k^2}$, with $\hat{\theta}_i^{\hat{k}}, \hat{\pi}_i^{\hat{k}}$ specified so that

$$(\hat{\theta}_i^{\hat{k}}(t_i), \hat{\pi}_i^{\hat{k}}(t_i)) = (\hat{\theta}_i^{k^1}(t_i), \hat{\pi}_i^{k^1}(t_i)) \text{ if } t_i \in T_i^{k^1} \tag{36}$$

and

$$(\hat{\theta}_i^{\hat{k}}(t_i), \hat{\pi}_i^{\hat{k}}(t_i)) = (\hat{\theta}_i^{k^2}(t_i), \hat{\pi}_i^{k^2}(t_i)) \text{ if } t_i \in T_i^{k^2}. \tag{37}$$

Given this finding, they go on to show that any prior $F^{\hat{k}} \in \mathcal{P}$ that can be represented in the form

$$F^{\hat{k}} = \sum_{j=1}^{J} \alpha_j F^{k^j}, \tag{38}$$

with $\alpha_j > 0$ for all j, has the BDP property *if and only if* every one of the distributions F^{k^j} has the BDP property. This leads them to conclude that, unless the incomplete-information models $\mathcal{T}^k$, $k \in K$, admit *only* BDP priors, the set of non-BDP priors will be geometrically and measure-theoretically generic in $\mathcal{P}$. Specifically, if the incomplete-information models $\mathcal{T}^k$, $k \in K$, admit one or more non-BDP priors, the set of BDP priors will be a proper face of the convex set $\mathcal{P}$. Moreover, under certain additional regularity conditions, the set of BDP priors will be finitely shy in $\mathcal{P}$.

Chen and Xiong (2011) contrast the geometric and measure theoretic approaches of Heifetz and Neeman (2006) with their own topological approach, which yields genericity of the BDP property in a universal type space setting. Our approach is also topological. From our perspective, however, with finite-dimensional type spaces, the difference between our results and those of Heifetz and Neeman (2006) is not one of topological versus geometric or measure-theoretic genericity but one of genericity in the full space versus genericity in a specially chosen subspace. Heifetz and Neeman (2006) are concerned with the genericity of non-BDP priors relative to the set of priors that is associated with a given family $\{T^k\}_{k\in K}$ of incomplete-information models that is closed under finite unions. If the family $\{\mathcal{T}^k\}_{k\in K}$ is the set of common-prior models in the universal type space, this approach involves no loss of generality. With finite-dimensional abstract type spaces, however, the requirement that the family $\{\mathcal{T}^k\}_{k\in K}$ be closed under finite unions is quite restrictive. In a companion paper, Gizatulina and Hellwig (2011), we show that, if the priors have continuous densities, then a family that satisfies this requirement is at most countable; more precisely, such a family consists of unions of members of a countable family of models with type sets that do not intersect each other. As an implication of Proposition 4.7, we then find that the set of such families for which all models satisfy the

BDP property contains a residual set. The set of families of models for which Heifetz and Neeman (2006) obtain geometric and measure-theoretic genericity on non-BDP priors is itself a sparse set.

Topological genericity of the BDP property is also discussed by Barelli (2009) and Chen and Xiong (2011). Barelli suggests that the measure-theoretic approach of Heifetz and Neeman (2006) is problematic and states a topological genericity result for non-BDP priors. Chen and Xiong (2011) point to an error in Barelli's argument and prove a topological genericity result for BDP priors.

In contrast to this paper, Barelli (2009) and Chen and Xiong (2011) specify the type space $T = \prod_{i=1}^{I} T_i$ as the Θ-based universal type space, i.e., the space of payoff parameters and belief hierarchies that is generated by the payoff type space $\Theta = \prod_{i=1}^{I} \Theta_i$.[14] Chen and Xiong rely on the fact that, if the space of common priors on this space is endowed with the topology of weak convergence of probability measures, i.e., the weak* topology, the set of common priors with finite supports is dense and the set of BDP priors is generic in the set of finite priors.

Because Chen and Xiong work with the universal type space and with finite approximations and we work with finite-dimensional abstract type spaces and embedding theorems, their results and ours reflect different aspects of the underlying structure. In a universal type space setting, one can rely on the weak* topology because there is no need to worry about informational discontinuities. The dependence of beliefs on information plays no role because agents' beliefs themselves are encoded in their types.

To understand this point, go back to Example 3.7, where agent 2 has a noisy signal $s_2 = (\theta_1)^2 + \varepsilon$ and agent 1's belief about the signal s_2 reveals $(\theta_1)^2$ but not θ_1. If the support of agent 1's type distribution was the finite set

$$\{-\frac{2n}{\sqrt{n}}, -\frac{2(n-1)}{\sqrt{n}}, ..., -\frac{2}{\sqrt{n}}, 0, \frac{1}{\sqrt{n}}, \frac{1+2}{\sqrt{n}}, ..., \frac{1+2(n-1)}{\sqrt{n}}, \frac{1+2n}{\sqrt{n}}\} \tag{39}$$

and the support of the noise term ε contains only even multiples of $\frac{1}{\sqrt{n}}$, the signal s_2 would

[14] Barelli (2009) presents his analysis as if he was considering arbitrary type spaces. However, like Heifetz and Neeman (2006), he works with fixed belief functions. By the arguments in Gizatulina and Hellwig (2011), this implies that distinct indecomposable type sets cannot intersect. As he goes on to talk about convergence of sequences of models in terms of Hausdorff convergence of the corresponding sequences of type sets, he cannot be working with abstract finite-dimensional type spaces but must implicitly be working with the universal type space.

actually reveal θ_1 and so would agent 1's belief about s_2. In this case, if s_2 is an even multiple of $\frac{1}{\sqrt{n}}$, agent 2 will know that θ_1 is negative; if s_2 is an odd multiple of $\frac{1}{\sqrt{n}}$, agent 2 will know that θ_1 is positive. If we consider a sequence of type distributions with supports (39) approximating the continuous type distribution, there is an informational discontinuity because, in the continuous limit, agent 2 cannot distinguish whether θ_1 is negative or positive. This informational discontinuity, however, appears only in the abstract type space approach of Example 3.7. In a universal type space approach, the informational discontinuity is avoided. With beliefs defined by types, weak convergence of type distributions implies that the equations $E_2[\theta_1|s_2] = \theta_1$ and $E_1[s_2|\theta_1] = \theta_1$ holds in the limit as well as along the sequence.

The universal type space approach is not well suited to dealing with the endogeneity of beliefs. When, in Section 2 above, we introduced the role of beliefs as conditional distributions given the information that is available to agents, we argued that, without loss of generality, any incomplete-information model $\mathcal{T} = \{T_i, \hat{\theta}_i, \hat{\pi}_i\}_{i=1}^I$ could be rewritten in the form

$$\hat{\mathcal{T}} = \{\Theta_i \times S_i, \hat{\pi}_i\}_{i=1}^I \tag{40}$$

with the understanding that, for any $t_i = (\theta_i, s_i) \in \Theta_i \times S_i$, $\hat{\theta}_i(t_i) = \theta_i$ is the payoff type of agent i and $\hat{s}_i(t_i) = s_i$ is an additional signal observed by agent i. In principle, this reformulation is also available when $T = \prod_{i=1}^I T_i$ is the universal type space. In this case, however, the type of any agent i is a pair $t_i = (\theta_i, h_i)$ where θ_i is the agent's payoff type and h_i indicates the hierarchy of the agent's beliefs about the other agents' payoffs, the other agents payoffs and first-order beliefs, etc. Reformulating a universal type space model in the form (40) is trivial, but requires that the space S_i of signals available to agent i be identified with the space H_i of belief hierarchies for this agent. The question what information can be inferred from the observation of the type $t_i = (\theta_i, s_i) = (\theta_i, h_i)$ is moot because, for each belief hierarchy $h_i \in H_i$, there is a unique measure $\hat{\pi}_i(h_i)$ on T_{-i} that is compatible with the belief hierarchy h_i. By construction, all relevant information is encoded in the belief hierarchy $s_i = h_i$, and any other information, e.g., about payoff types, is redundant.

In the universal type space approach, the question whether beliefs properly reflect the information that is available to agents becomes a question about type sets and priors. For a single agent i, the belief function $\hat{\pi}_i$ must satisfy the equation $\hat{\pi}_i(h_i) \equiv b_i(\theta_i, h_i|\nu_i)$ for some measure ν_i and for all (θ_i, h_i) in the support of the measure ν_i. This is possible *only* if if the BDP property holds so that all information about θ_i is already encoded in h_i or if θ_i is to some extent independent of the other agents' types so that beliefs about the other agents' types do not depend on θ_i. The latter alternative is very special. The analysis of Chen and Xiong (2011) exploits this structure of the universal type space without paying attention to the fact that beliefs should be treated as the result of conditioning on available information.

6 Concluding Remarks

We conclude this paper with several additional remarks. First, the genericity properties established in Propositions 4.6 and 4.7 are still obtained if we replace the space $\mathcal{M}_+^d(T)$ by the space $\mathcal{M}_+^d(\mathring{T})$ of measures with densities that are defined on the interior of T and if we endow the spaces $\mathcal{M}_+^d(\mathring{T}_i)$ and $\mathcal{M}^d(\mathring{T}_{-i})$ of marginal distributions on the interiors of the type spaces of agent i and agent other than i with the topologies that is induced by the compact open topologies, rather than the uniform topologies for the associated density functions. In this case, however, an analogue of Corollary 4.8 is only obtained if an additional condition of uniform boundedness is imposed on the density functions of the priors under consideration.

Second, in Gizatulina and Hellwig (2011), we show that the genericity properties established in Propositions 4.6 and 4.7 hold also for an infinite-dimensional type space of the form $T = \cup_{\ell=1}^{\infty} T^\ell$, where, for each ℓ, $T^\ell = \prod_{i=1}^{I} T_i^\ell$ is a finite-dimensional set. This corresponds to a specification with *ex ante* uncertainty about the incomplete-information model that is going to be relevant. If the set of such models is countable and each model has a finite-dimensional type space, the BDP property is still generic.

Third, the assumption that the type sets T_i in Proposition 4.7 are taken as fixed and given is not essential. In principle, the set T in Proposition 4.7 could be any compact subset of $\mathbb{R}^N$ that has a nonempty interior $\mathring{T}$. Given a measure ν with a density that is continuous on $\mathring{T}$, the belief function $b_i(\cdot|\cdot,\nu)$ is defined on the set D_i^ν of types of agent i for which the marginal density $\bar{f}_i^\nu(t_i)$ is strictly positive. D_i^ν is an element of the set D_i^o of open subsets of $\mathbb{R}^{n_i}$ that have compact closures. Thus $b_i(\cdot|\cdot,\nu)$ can be treated as an element of the space $\cup_{D\in\mathcal{D}_i^o} C(D, \mathcal{M}^d(T_{\text{-}i}))$. If the topology on $\cup_{D\in\mathcal{D}^o} C(D, \mathcal{M}^d(T_{-i}))$ is specified so that a sequence $\{b_i^k\}$ of functions in $\cup_{D\in\mathcal{D}^o} C(D, \mathcal{M}^d(T_{-i}))$ converges to a limit $b_i \in \cup_{D\in\mathcal{D}^o} C(D, \mathcal{M}^d(T_{-i}))$ if and only if, every compact subset K of the domain

D_{b_i} of b_i is also a subset of the domains $D_{b_i^k}$ of the functions b_i^k for any sufficiently large k, and, moreover, $\lim_{k\to\infty} b_i^k(t_i) = b_i(t_i)$, uniformly on K, an extension of the argument given to prove Proposition 4.6 can be used to show that the set $\cup_{D\in\mathcal{D}^o}\mathcal{E}(D, \mathcal{M}^d(T_{-i}))$ of embeddings is a residual subset of $\cup_{D\in\mathcal{D}^o} C(D, \mathcal{M}^d(T_{-i}))$. A similar generalization of Proposition 4.3 is also available.

Forth, in a previous version of this paper, we had applied the Embedding Theorem to the conditional-expectations functions $\bar{t}^{\nu}_{-i}(\cdot)$, which are given by the formula

$$\bar{t}^{\nu}_{-i}(t_i) := \int_{T^{\nu}_{-i}} t_{-i}\beta_i(t_{-i}|t_i,\nu)dt_{-i} = \frac{\int_{T^{\nu}_{-i}} t_{-i} f^{\nu}(t_i, t_{-i})dt_{-i}}{\bar{f}^{\nu}_i(t_i)}.$$

Under the dimensionality assumption that $2n_i + 1 \leq N_{-i}$, the Embedding Theorem for continuous functions implies that, for any $\varepsilon > 0$ there exists an embedding $t_i \to \hat{t}^{\varepsilon}_{-i}(t_i)$ such that $\left\|\hat{t}^{\varepsilon}_{-i}(t_i) - \bar{t}^{\nu}_{-i}(t_i)\right\| < \varepsilon$ for all $t_i \in D^{\nu}_i$. Given this embedding, approximating densities can be defined by writing

$$f^{\nu^{\varepsilon}}(t_i, t_{-i}) = \beta(t_{-i} + \hat{t}^{\varepsilon}_{-i}(t_i) - \bar{t}^{\nu}_{-i}(t_i)|t_i)\bar{f}^{\nu}_i(t_i) \text{ if } t_i \in D^{\nu}_i, \tag{41}$$

and

$$f^{\nu^{\varepsilon}}(t_i, t_{-i}) = 0 \text{ if } t_i \notin D^{\nu}_i. \tag{42}$$

The BDP property of approximating priors is then directly embodied in the conditional expectations, and there is no need to appeal to the fact that probability distributions themselves are infinite-dimensional objects.

Finally, if we strengthen the assumptions on the set of priors under consideration so that we can use a differentiable approach, then, under the dimensionality assumption that $2n_i + 1 \leq N_{-i}$, we can strengthen the claim that BDP priors are generic from residual to open and dense. Specifically, if $\bar{N}^{cd}$ is a subset of $\bar{N}^{c}$ so that, for any $\nu \in \bar{N}^{cd}$, the support T^{ν} of ν is a compact manifold and the conditional-expectations functions $\bar{t}^{\nu}_{-i}$ are continuously differentiable on T^{ν}_i, and if the topology on $\bar{N}^{cd}$ is such that, for all i, the map from priors into agent i's conditional-expectations functions is continuous when the range of this map is given the strong C^1 topology, then, under the dimensionality assumption in Proposition 4.7, the set of common priors exhibiting the BDP property is open and dense in $\bar{N}^{cd}$. The proof is basically the same as the proof of Proposition 4.7, except that the Embedding Theorem for continuous functions with compact domains must be replaced by Whitney's Embedding Theorem for C^1 functions, see, e.g., Hirsch (1994), p. 35.

A Appendix

In this appendix, we state and prove the embedding theorems that are used in Section 4. For any separable metric space X, a subset of X is said to be *residual* in X if it contains a countable intersection of open and dense subsets of X. For any two separable metric spaces X and Y, $C(X,Y)$ will denote the space of continuous functions from X to Y. The space $C(X,Y)$ will be endowed with the uniform topology. The set of embeddings of X into Y will be denoted as $\mathcal{E}(X,Y)$.

A.1 Proof of Proposition 4.3

Proposition 4.3 in the text is an instance of the following result, with $X = T_i$ and $Z = T_{-i}$.

PROPOSITION A.1 *Let $X \subset \mathbb{R}^n$ be compact, and let Z be a compact metric space with infinitely many elements. Let $Y = \mathcal{M}(Z)$ be the space of probability measures on Z, endowed with the topology of weak convergence of probability measures. Then the set $\mathcal{E}(X,Y)$ of embeddings of X in Y is a residual subset of $C(X,Y)$.*

In proving Proposition A.1, we will repeatedly use the following lemma from topology.

LEMMA A.2 *Let X, Y be any two topological spaces and let g be a continuous and open function from X to Y. Then, for any open and dense set $U \subset Y$, the inverse image $g^{-1}(U)$ is an open and dense subset of X. For any residual set $Y^r \subset Y$, the inverse image $g^{-1}(Y^r)$ is a residual subset of X.*

Proof. Continuity of $g : X \to Y$ implies that $g^{-1}(U)$ is open in X whenever U is open in Y. Further, openness of g implies that $g(V)$ is open in Y whenever V is open in X. Thus, if V is open in X and U is dense in Y, we must have $g(V) \cap U \neq \emptyset$ and, therefore, $V \cap g^{-1}(U) = g^{-1}(g(V) \cap U) \neq \emptyset$. Thus, $g^{-1}(U)$ is dense in X whenever U is dense in Y. The first statement of the lemma is proved.

To prove the second statement, it suffices to observe that, if $Y^r \subset Y$ satisfies $Y^r \supset \cap_{k=1}^{\infty} U_k$ for some sequence $\{U_k\}$ of open and dense subsets of Y, then

$$g^{-1}(Y^r) \supset g^{-1}(\cap_{k=1}^{\infty} U_k) = \cap_{k=1}^{\infty} g^{-1}(U_k),$$

and, by the first statement of the lemma, the sets $g^{-1}(U_k)$, $k = 1, 2, ...$, are open and dense in X. ■

LEMMA A.3 *Let X and Y be as specified in Proposition A.1. If there exists a separable metric space Q that is homeomorphic to $[0,1]^{2n+1}$ and if there exists a mapping Φ from $C(X,Y)$ to $C(X,Q)$ that is continuous and open, then the set $\mathcal{E}(X,Y)$ of embeddings of X in Y is a residual subset of $C(X,Y)$.*

Proof. Let h be a homeomorphism from Q to $[0,1]^{2n+1}$. We claim that the formula

$$H(f) := h \circ f$$

defines a homeomorphism H from $C(X,Q)$ to $C(X,[0,1]^{2n+1})$. To establish this claim, we note that, because h is continuous, H takes values in $C(X,[0,1]^{2n+1})$. Continuity of h also implies that H is continuous. Because h is invertible, the inverse of H is well defined, with

$$H^{-1}(\hat{f}) = h^{-1} \circ \hat{f}$$

for any $\hat{f} \in [0,1]^{2n+1}$. Continuity of h^{-1} implies that H^{-1} takes values in $C(X,Q)$ and that H^{-1} is continuous.

By the classical Embedding Theorem,[15] the set $\mathcal{E}(X,[0,1]^{2n+1})$ of embeddings of X in $[0,1]^{2n+1}$ is a residual subset of $C(X,[0,1]^{2n+1})$. By Lemma A.2 therefore, $H^{-1}(\mathcal{E}(X,[0,1]^{2n+1}))$ is a residual subset of $C(X,Q)$. If the mapping Φ from $C(X,Y)$ to $C(X,Q)$ is continuous and open, then, again by Lemma A.2, it follows that the set $\Phi^{-1}(H^{-1}(\mathcal{E}(X,[0,1]^{2n+1})))$ is a residual subset of $C(X,Y)$.

To prove the lemma, it therefore suffices to show that

$$\Phi^{-1}(H^{-1}(\mathcal{E}(X,[0,1]^{2n+1})) \subset \mathcal{E}(X,Y). \tag{43}$$

For this purpose, suppose that $f \in \Phi^{-1}(H^{-1}(\mathcal{E}(X,[0,1]^{2n+1}))\backslash\mathcal{E}(X,Y)$. Then there exist $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$ and $h(\Phi(x_1|f)) \neq h(\Phi((x_2|f))$. Because h is a homeomorphism, $h(\Phi(x_1|f)) \neq h(\Phi((x_2|f))$ implies $\Phi(x_1|f) \neq \Phi(x_2|f)$, which in turn implies $f(x_1) \neq f(x_2)$. The assumption that $\Phi^{-1}(H^{-1}(\mathcal{E}(X,[0,1]^{2n+1}))\backslash\mathcal{E}(X,Y) \neq \emptyset$ thus leads to a contradiction, which proves (43). The lemma follows immediately. ∎

To prove Proposition A.1, it thus suffices to show that there exist Q and Φ as specified in the lemma. We define Q as the unit simplex in $\mathbb{R}_+^{2n+1}$, i.e.,

$$Q := \left\{ \mathbf{q} \in \mathbb{R}_+^{2n+1} \middle| \sum_{i=1}^{2n+1} q_i \leq 1 \right\}. \tag{44}$$

[15] See Theorem V.2, p. 56, in Hurewicz and Wallman (1941).

LEMMA A.4 *The set Q that is defined by (44) is homeomorphic to $[0,1]^{2n+1}$.*

Proof. Because both Q and $[0,1]^{2n+1}$ are compact convex subset of $\mathbb{R}_+^{2n+1}$ and have non-empty interiors, they are both homeomorphic to the closed unit ball in $\mathbb{R}_+^{2n+1}$ and therefore to each other.[16] ■

To construct the mapping Φ, we proceed as follows: Exploiting the fact that Z has infinitely many elements, we fix $2n+1$ distinct elements $z_1, ..., z_{2n+1}$ of Z. For this purpose, we note that, as compact metric space, Z is separable. Therefore, there exists some $\varepsilon > 0$ so that the open ε-balls $B_\varepsilon(z_1), ..., B_\varepsilon(z_{2n+1})$ around $z_1, ..., z_{2n+1}$ do not intersect each other, i.e., $B_\varepsilon(z_i) \cap B_\varepsilon(z_j) = \emptyset$ for all i and all $j \neq i$. Moreover, by Urysohn's lemma, there exist continuous functions $g_1, ...g_{2n+1}$ from Z into $[0,1]$ such that, for any i, $g_i(z_i) = 1$ and $g_i(z) = 0$ for $z \in Z \backslash B_\varepsilon(z_i)$. For any $z \in Z$, therefore, $\sum_{i=1}^{2n+1} g_i(z) \in [0,1]$, and, for any $\mu \in \mathcal{M}(Z)$, the vector

$$\varphi(\mu) = \left(\int_Z g_1(z) d\mu(z), ..., \int_Z g_{2n+1}(z) d\mu(z) \right) \tag{45}$$

is an element of the set Q that is defined by (44). Given the mapping $\varphi : \mathcal{M}(Z) \to Q$, the formula

$$\Phi(b) = \varphi \circ b \tag{46}$$

defines a mapping from $C(X, \mathcal{M}(Z))$ into the space of functions from X to Q. We need to show that $\Phi(b) \in C(X, Q)$ for all $b \in C(X, \mathcal{M}(Z))$ and that the mapping Φ is continuous and open. We begin by showing that the mapping φ has these properties.

LEMMA A.5 *The mapping $\varphi : \mathcal{M}(Z) \to Q$ that is defined by formula (45) is continuous and open.*

Proof. Continuity of φ is immediate from (45) and the definition of the topology of weak convergence.

Note that, if Z is a compact metric space and if the space $Y = \mathcal{M}(Z)$ of probability measures on Z is endowed with the topology of weak convergence of probability measures, i.e., the weak* topology, then Y is a compact metric space.[17] If $\{h_1, h_2, ...\}$ is

[16] See Proposition 4.26 in Lee (2000).

[17] See Theorem 6.4, p. 45, in Parthasarathy (1967)

a countable dense set of continuous functions from Z into $[0,1]$, then, by an argument Parthasarathy (1967), p. 43, the mapping

$$\mu \to h(\mu) := \left(\int_Z h_1(z) d\mu(z), \int_Z h_2(z) d\mu(z), ... \right)$$

defines a homeomorphism between $\mathcal{M}(Z)$ and a subspace S of the infinite product $[0,1]^\infty$. Without loss of generality, we may assume that $h_i = g_i$ for $i = 1, ..., 2n+1$. Let $\pi_{2n+1} : [0,1]^\infty \to [0,1]^{2n+1}$ be the natural projection from $[0,1]^\infty$ to the first $2n+1$ factors of the infinite product. Then

$$\varphi = \pi_{2n+1} \circ h.$$

Because the homeomorphism $h : \mathcal{M}(Z) \to S$ is open and the composition of two open mappings is also open, openness of the mapping $\varphi : \mathcal{M}(Z) \to Q$ will follow if the restriction π^S_{2n+1} of π_{2n+1} to S is shown to be an open mapping from S to Q.

For this purpose, we note that φ maps $\mathcal{M}(Z)$ *onto* Q, i.e., that $\varphi(\mathcal{M}(Z)) = Q$: For any $\mathbf{q} = (q_1, ..., q_{2n+1}) \in Q$, any $\mu \in \mathcal{M}(Z)$ such that $\mu(\{z_i\}) = q_i$ for $i = 1, 2, ...2n+1$, and $\mu(Z \setminus \cup_i B_\varepsilon(z_i)) = 1 - \sum_{i=1}^{2n+1} q_i$ satisfies $\varphi(\mu) = \mathbf{q}$. Thus, we obtain

$$\pi_{2n+1}(S) = \pi_{2n+1} \circ h(\mathcal{M}(Z)) = \varphi(\mathcal{M}(Z)) = Q.$$

Now let V be any open subset of S. By definition of the subspace topology for S, there exists an open set $U \subset [0,1]^\infty$ such that $V = S \cap U$. Thus,

$$\pi^S_{2n+1}(V) = \pi^S_{2n+1}(S \cap U) = \pi_{2n+1}(S \cap U) = \pi_{2n+1}(S) \cap \pi_{2n+1}(U) = Q \cap \pi_{2n+1}(U).$$

Because the projection π_{2n+1} is an open mapping from $[0,1]^\infty$ to $[0,1]^{2n+1}$, the set $\pi_{2n+1}(U)$ is an open subset of $[0,1]^{2n+1}$. Therefore, the set $\pi^S_{2n+1}(V) = Q \cap \pi_{2n+1}(U)$ is open in the subspace topology for Q as a subset of $[0,1]^{2n+1}$. Thus, π^S_{2n+1}maps any open subset of $S = h(\mathcal{M}(Z))$ into an open subset of Q. This proves that the mapping $\pi^S_{2n+1} : S \to Q$ is open. The mapping $\varphi = \pi^S_{2n+1} \circ h : \mathcal{M}(Z) \to Q$ is therefore also open. ■

LEMMA A.6 *The mapping Φ that is defined by formula (46) maps $C(X, \mathcal{M}(Z))$ continuously into $C(X, Q)$.*

Proof. Because φ is continuous, obviously, $\Phi(b) \in C(X, Q)$ if $b \in C(X, \mathcal{M}(Z))$. To prove that Φ is continuous, we note that, as mentioned above, if Z is a compact metric space, then $\mathcal{M}(Z)$ is a compact metric space. A metric ρ for $\mathcal{M}(Z)$ is given by the formula

$$\rho(\mu,\nu) = \sup_k \left| \int h_k(zy) d\mu(z) - \int h_k(z) d\nu(z) \right|, \tag{47}$$

where $\{h_1, h_2, ...\}$ is a countable dense subset of the set of continuous functions from Z into $[0,1]$. Again there is no of generality in assuming that $h_k = g_k$ for $k = 1, ..., 2n+1$.

For any $\gamma > 0$, define $\eta(\gamma) := \frac{\gamma}{2n+1}$. Then, for any μ and ν in $\mathcal{M}(Z)$, $\rho(\mu,\nu) < \eta(\gamma)$ implies

$$\left| \int h_k(z) d\mu(z) - \int h_k(z) d\nu(z) \right| < \frac{\gamma}{2n+1}$$

for all k. Thus, $\rho(\mu,\nu) < \eta(\gamma)$ implies $\|\varphi(\mu) - \varphi(\nu)\| \leq \gamma$, where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^{2n+1}$. If b and $\hat{b}$ in $C(X, \mathcal{M}(Z))$ are such that

$$\sup_{x \in X} \rho(b(x), \hat{b}(x)) < \eta(\gamma) = \frac{\gamma}{2n+1},$$

it follows that

$$\sup_{x \in X} \left\| \varphi(b(x)) - \varphi(\hat{b}(x)) \right\| \leq \gamma,$$

i.e., if b and $\hat{b}$ are $\eta(\gamma)$-close, then $\Phi(b)$ and $\Phi(\hat{b})$ are γ-close. This proves that Φ is continuous. ■

LEMMA A.7 *The mapping $\Phi : C(X, \mathcal{M}(Z)) \to C(X, Q)$ that is defined by (46) is open.*

Proof. By Proposition 2.2.b, p. 269, in Clausing (1978), the lemma follows from the fact that $Y = \mathcal{M}(Z)$ and Q are compact convex sets and that the mapping $\varphi : Y \to Q$ is affine, as well as continuous and open. ■

Proposition A.1 now follows from Lemmas A.3, A.6, and A.7.

A.2 Proof of Proposition 4.6

Proposition 4.6 in the text is an instance of the following result, with $X = T_i$ and $Z = T_{-i}$.

PROPOSITION A.8 *Let $X \subset \mathbb{R}^n$, and let $Z \subset \mathbb{R}^{N-n}$ be compact sets with nonempty interiors. Let $Y = \mathcal{M}^d(Z)$ be the space of probability measures on Z that have continuous density functions and let the topology on Y be induced by the uniform topology for density functions. Then the set $\mathcal{E}(X,Y)$ of embeddings of X in Y is a residual subset of $C(X,Y)$.*

The proof proceeds along the same lines as the proof of Proposition A.1. We note that $\mathcal{M}^d(Z)$ is a separable metric space. A metric ρ for $\mathcal{M}^d(Z)$ is given by the formula

$$\rho(\mu,\nu) = \sup_{z\in Z} \left| f_\mu(z) - f_\nu(z) \right|, \tag{48}$$

where f_μ and f_ν are the continuous densities associated with μ and ν.

The conclusion of Lemma A.3 remains valid with $X \subset \mathbb{R}^n$ and $Y = \mathcal{M}^d(Z)$ having the topology that is induced by uniform convergence of density functions. To prove Proposition A.8, it therefore suffices to specify a mapping $\hat{\Phi} : C(X, \mathcal{M}^d(Z)) \to C(X, [0,1]^{2n+1})$ that is continuous and open. We do so by setting

$$\hat{\Phi}(b) = \hat{\varphi} \circ b \tag{49}$$

for $b \in C(X, \mathcal{M}^d(Z))$, where $\hat{\varphi} : \mathcal{M}^d(Z) \to [0,1]^{2n+1}$ is given by the formula:

$$\hat{\varphi}(\mu) = \left(\frac{f_\mu(z_1)}{1+f_\mu(z_1)}, ..., \frac{f_\mu(z_{2n+1})}{1+f_\mu(z_{2n+1})} \right), \tag{50}$$

where $z_1, ..., z_{2n+1}$ is an arbitrary but fixed collection of distinct elements of the interior of Z.

LEMMA A.9 *The mapping $\hat{\varphi} : \mathcal{M}^d(Z) \to [0,1]^{2n+1}$ is continuous and open.*

Proof. Continuity is immediate from (50). Openness follows from observing that $\hat{\varphi}$ is the composition of the homeomorphism $\mu \to f_\mu$ between $\mathcal{M}^*(Z)$ and $C(Z, \mathbb{R}_+)$, the projection $f \to (f(z_1), ...f(z_{2n+1}))$ from $C(Z, \mathbb{R}_+)$ to $\mathbb{R}_+^{2n+1}$, and the homeomorphism

$$(f_1, ...f_{2n+1}) \to \left(\frac{f_1}{(1+f_1)}, ..., \frac{f_{2n+1}}{(1+f_{2n+1})} \right)$$

from $\mathbb{R}_+^{2n+1}$ to the product $[0,1)^{2n+1}$. ■

LEMMA A.10 *The function $\hat{\Phi}$ that is defined by (49) and (50) maps $C(X, \mathcal{M}^d(Z))$ continuously into $C(X, [0,1]^{2n+1})$.*

Proof. By routine calculations, (50) implies

$$\|\hat{\varphi}(\mu) - \hat{\varphi}(\bar{\mu})\| \leq (2n+1) \max_i \left| f_\mu(z_i) - f_{\bar{\mu}}(z_i) \right|$$

for any μ and $\bar{\mu}$ in $\mathcal{M}^d(Z)$, where $\|\cdot\|$ is again the Euclidean norm on $\mathbb{R}^{2n+1}$. For any μ and $\bar{\mu}$ in $\mathcal{M}^d(Z)$, therefore,

$$\|\hat{\varphi}(\mu) - \hat{\varphi}(\bar{\mu})\| \leq (2n+1)\rho(\mu, \bar{\mu}).$$

The function $\hat{\varphi}$ thus is uniformly continuous. By the same argument as in the proof of Lemma A.6, it follows that $\hat{\Phi}$ takes values in $C(X, [0,1]^{2n+1})$ and that $\hat{\Phi}$ is continuous. ■

The proof that $\hat{\Phi}$ is also open is more involved. The range $\mathcal{M}^d(Z)$ of belief functions now consists of measures with continuous density functions and is therefore not compact.[18] Therefore we cannot rely on the result of Clausing (1978) to infer that $\hat{\Phi}$ is open if $\hat{\varphi}$ is open. Instead we need a new argument.

LEMMA A.11 *For any* $(\mathbf{q}, \mu) \in [0,1)^{2n+1} \times \mathcal{M}^d(Z)$*, the infimum* $\rho^*(\mathbf{q}, \mu)$ *of the distance* $\rho(\mu, \nu)$ *over the set of measures* $\nu \in \mathcal{M}^d(Z)$ *that satisfy*

$$\hat{\varphi}(\nu) = \mathbf{q}. \tag{51}$$

is well defined and satisfies

$$\rho^*(\mathbf{q}, \mu) = \max_i \left| \frac{q_i}{1-q_i} - f_\mu(z_i) \right|. \tag{52}$$

Proof. Fix $\mathbf{q} \in [0,1)^{2n+1}$ and $\mu \in \mathcal{M}^d(Z)$ and let f_μ be the density of μ. For any $\varepsilon > 0$ and any i, let $B_\varepsilon(z_i)$ be the open ε-ball around z_i and let $g_i^\varepsilon \in C(Z, [0,1])$ be such that $g_i^\varepsilon(z_i) = 1$ and $g_i^\varepsilon(z) = 0$ for all $z \notin B_\varepsilon(z_i)$. For any $\varepsilon > 0$ and any $\alpha > 0$, consider the function $\hat{f}_{\alpha\varepsilon}$ that is given by the formula

$$\hat{f}_{\alpha\varepsilon}(z) = \max\left[0, \alpha f_\mu(z) + \sum_{i=1}^{2n+1} \left(\frac{q_i}{1-q_i} - \alpha f_\mu(z_i) \right) g_i^\varepsilon(z) \right]. \tag{53}$$

By construction, the function $\hat{f}_{\alpha\varepsilon}(\cdot)$ from Z into $\mathbb{R}_+$ is continuous. Thus, if ε and α are such that

$$\int_Z \hat{f}_{\alpha\varepsilon}(z)dz = 1, \tag{54}$$

then $\hat{f}_{\alpha\varepsilon}(\cdot)$ is the density of a measure $\nu_{\alpha\varepsilon} \in \mathcal{M}^d(Z)$.

[18] Compactness would require that the density functions belong to an equicontinuous set of functions.

We claim that, if $\varepsilon > 0$ is sufficiently small, there exists $\alpha(\varepsilon)$ such that equation (54) holds for $\alpha = \alpha(\varepsilon)$ and ε. Moreover, if ε is close to zero, $\alpha(\varepsilon)$ is close to one. To prove this claim, we note that, because $\int_Z f_\mu(z)dz = 1$, (53) implies

$$\int_Z \hat{f}_{\alpha\varepsilon}(z)dz \leq \alpha + \sum_{i=1}^{2n+1} \frac{q_i}{1-q_i} \int_Z g_i^\varepsilon(z)dz \leq \alpha + H(\varepsilon) \tag{55}$$

and

$$\int_Z \hat{f}_{\alpha\varepsilon}(z)dz \geq \alpha + \sum_{i=1}^{2n+1} \left[\frac{q_i}{1-q_i} - f_\mu(z_i)\right] \int_Z g_i^\varepsilon(z)dz \geq \alpha - H(\varepsilon), \tag{56}$$

where

$$\begin{aligned} H(\varepsilon) &:= \sum_{i=1}^{2n+1} \max\left[\frac{q_i}{1-q_i}, f_\mu(z_i)\right] \int_Z g_i^\varepsilon(z)dz \\ &\leq \sum_{i=1}^{2n+1} \max\left[\frac{q_i}{1-q_i}, f_\mu(z_i)\right] \int_{B_\varepsilon(z_i)} dz. \end{aligned} \tag{57}$$

Then, for $\alpha \leq 1 - H(\varepsilon)$, (55) and (57) imply $\int_Z \hat{f}_{\alpha\varepsilon}(z)dz \leq 1$ and, for $\alpha \leq 1 + H(\varepsilon)$, (56) and (57) imply $\int_Z \hat{f}_{\alpha\varepsilon}(z)dz \geq 1$. By the intermediate value theorem, there exists $\alpha(\varepsilon) \in [1-\eta, 1+\eta]$ such that $\int_Z \hat{f}_{\alpha\varepsilon}(z)dz = 1$. Moreover, by (57), $\lim_{\varepsilon\to 0} H(\varepsilon) = 0$ and, therefore, $\lim_{\varepsilon\to 0} \alpha(\varepsilon) = 1$.

Consider the measures $\nu_{\alpha(\varepsilon)\varepsilon} \in \mathcal{M}^d(Z)$ that are induced by the density functions $\hat{f}_{\alpha(\varepsilon),\varepsilon}$ for $\varepsilon > 0$ sufficiently small and $\alpha(\varepsilon)$ chosen so that (54) holds. Because the points $z_1, ..., z_{2n+1}$ in Z are distinct, there exists $\bar{\varepsilon} > 0$ such that, for $\varepsilon < \bar{\varepsilon}$, no two of the open balls $B_\varepsilon(z_i)$ intersect each other. In this case, (53) implies

$$\hat{f}_{\alpha\varepsilon}(z_i) = \frac{q_i}{1-q_i} \tag{58}$$

for all i. For $\alpha = \alpha(\varepsilon)$, it follows that the measure $\nu_{\alpha(\varepsilon)\varepsilon} \in \mathcal{M}^d(Z)$ satisfies $\hat{\varphi}(\nu_{\alpha\varepsilon}) = \mathbf{q}$. This implies, in particular, that the set of measures satisfying (51) is nonempty so that the infimum $\rho^*(\mathbf{q}, \mu)$ is well defined.

For any one of the measures $\nu_{\alpha(\varepsilon)\varepsilon}$ with $\varepsilon < \bar{\varepsilon}$, we compute

$$\begin{aligned} \rho(\mu, \nu_{\alpha(\varepsilon)\varepsilon}) &= \sup_{z\in Z} \left| f_\mu(z) - \hat{f}_{\alpha(\varepsilon),\varepsilon}(z) \right| \\ &\leq \sup_{z\in Z} \left| (1-\alpha(\varepsilon)) f_\mu(z) - \sum_{i=1}^{2n+1} \left(\frac{q_i}{1-q_i} - \alpha(\varepsilon) f_\mu(z_i) \right) g_i^\varepsilon(z) \right| \\ &\leq (1-\alpha(\varepsilon)) \sup_{z\in Z} f_\mu(z) + \max_i \left| \frac{q_i}{1-q_i} - f_\mu(z_i) \right| + (1-\alpha(\varepsilon)) f_\mu(z_i). \end{aligned}$$

If ε and $\alpha(\varepsilon)$ go to zero, the first term and the third term on the right-hand side vanish and only the middle term remains. Therefore,

$$\rho^*(\mathbf{q}, \mu) \leq \max_i \left| \frac{q_i}{1 - q_i} - f_\mu(z_i) \right|.$$

Equation (52) follows because, for ν satisfying (51), we also have

$$\begin{aligned}\rho(\mu, \nu) &= \sup_{z \in Z} |f_\mu(z) - f_\nu(z)| \\ &\geq \max_i |f_\mu(z_i) - f_\nu(z_i)| \\ &= \max_i \left| f_\mu(z_i) - \frac{q_i}{1 - q_i} \right|,\end{aligned}$$

and hence

$$\rho^*(\mathbf{q}, \mu) \geq \max_i \left| \frac{q_i}{1 - q_i} - f_\mu(z_i) \right|.$$

■

LEMMA A.12 *For any $\varepsilon > 0$, there exists a continuous function ν_ε from $[0, 1)^{2n+1} \times \mathcal{M}^d(Z)$ to $\mathcal{M}^d(Z)$ such that, for any $(\mathbf{q}, \mu) \in [0, 1)^{2n+1} \times \mathcal{M}^d(Z)$, $\hat{\varphi}(\nu_\varepsilon(\mathbf{q}, \mu)) = \mathbf{q}$ and*

$$\rho(\mu, \nu_\varepsilon(\mathbf{q}, \mu)) \leq \rho^*(\mathbf{q}, \mu) + \varepsilon, \tag{59}$$

where $\rho^*(\mathbf{q}, \mu)$ is again the infimum of the distance $\rho(\mu, \nu)$ over the set of measures $\nu \in \mathcal{M}^d(Z)$ that satisfy (51).

Proof. Fix $\varepsilon > 0$. For any $\mathbf{q} \in Q$ and any $\mu \in \mathcal{M}^d(Z)$, let

$$\psi_\varepsilon(\mathbf{q}, \mu) := \{\nu \in \mathcal{M}^d(Z) | \varphi(\nu) = \mathbf{q} \text{ and } \rho(\mu, \nu) < \rho^*(\mathbf{q}, \mu) + \varepsilon\}, \tag{60}$$

and let $\bar{\psi}_\varepsilon(\mathbf{q}, \mu)$ be the closure of $\psi_\varepsilon(\mathbf{q}, \mu)$. To prove the lemma, it suffices to show that the correspondence $\bar{\psi}_\varepsilon$ from $Q \times \mathcal{M}^d(Z)$ into $\mathcal{M}^d(Z)$ has a continuous selection.

By Theorem 1.2 of Michael (1964), existence of a continuous selection of the closed-valued correspondence $\bar{\psi}_\varepsilon$ from $Q \times \mathcal{M}^d(Z)$ into $\mathcal{M}^d(Z)$ is guaranteed if $\bar{\psi}_\varepsilon$ is convex-valued and lower hemi-continuous. To show that $\bar{\psi}_\varepsilon$ has these properties, it suffices that to show that ψ_ε has them: Convex-valuedness of $\bar{\psi}_\varepsilon$ then follows from the convex-valuedness of ψ_ε and the observation that the closure of a convex set is convex; similarly, lower hemi-continuity of $\bar{\psi}_\varepsilon$ follows from the lower hemi-continuity of ψ_ε.[19]

[19] Hildenbrand (1974), pp. 36, 26.

To establish convex-valuedness of ψ_ε, fix $(\mathbf{q}, \mu) \in Q \times \mathcal{M}^d(Z)$. Let ν_1, ν_2 be any two elements of $\psi_\varepsilon(\mathbf{q}, \mu)$, and, for some $\lambda \in (0,1)$, consider the convex combination $\lambda\nu_1 + (1-\lambda)\nu_2$ of ν_1 and ν_2. Because ν_1 and ν_2 belong to $\psi_\varepsilon(\mathbf{q}, \mu)$, the densities $f_{\nu_1} f_{\nu_2}$ of ν_1, ν_2 satisfy

$$f_{\nu_1}(z_i) = f_{\nu_2}(z_i) = \frac{q_i}{1-q_i} \tag{61}$$

for all i and

$$\sup_{z \in Z} \rho(\mu, \nu_j) < \rho^*(\mathbf{q}, \mu) + \varepsilon \tag{62}$$

for $j = 1, 2$. From (61), we immediately obtain

$$f_{\lambda\nu_1+(1-\lambda)\nu_2}(z_i) = \lambda f_{\nu_1}(z_i) + (1-\lambda) f_{\nu_2}(z_i) = \frac{q_i}{1-q_i}$$

for all i, hence, $\hat{\varphi}(\lambda\nu_1 + (1-\lambda)\nu_2) = \mathbf{q}$. Moreover, (62) implies

$$\begin{aligned}
&\rho(\mu, \lambda\nu_1 + (1-\lambda)\nu_2) \\
&= \sup_{z \in Z} \left| f_\mu(z) - \lambda f_{\nu_1}(z) - (1-\lambda) f_{\nu_2}(z) \right| \\
&\leq \lambda \sup_{z \in Z} \left| f_\mu(z) - f_{\nu_1}(z) \right| + (1-\lambda) \sup_{z \in Z} \left| f_\mu(z) - \lambda f_{\nu_2}(z) \right| \\
&< \rho^*(\mathbf{q}, \mu) + \varepsilon.
\end{aligned}$$

Thus, $\lambda\nu_1 + (1-\lambda)\nu_2 \in \psi_\varepsilon(\mathbf{q}, \mu)$.

To prove lower hemi-continuity of ψ_ε, consider any pair $(\mathbf{q}, \mu) \in Q \times \mathcal{M}^d(Z)$, any measure $\nu \in \psi_\varepsilon(\mathbf{q}, \mu)$, and any sequence $\{(\mathbf{q}^r, \mu^r)\}$ that converges to $(\mathbf{q}, \mu)$. Because $\hat{\varphi}$ is open, there exists a sequence $\{\nu^r\}$ converging to ν such that $\varphi(\nu^r) = \mathbf{q}^r$ for all r. For any r, the triangle inequality implies

$$\rho(\mu^r, \nu^r) \leq \rho(\mu^r, \mu) + \rho(\mu, \nu) + \rho(\nu, \nu^r).$$

Since $\nu \in \psi_\varepsilon(\mathbf{q}, \mu)$ satisfies (62), it follows that, for some $\eta > 0$, we have

$$\rho(\mu^r, \nu^r) \leq \rho(\mu^r, \mu) + \rho^*(\mathbf{q}, \mu) + \varepsilon + \rho(\nu, \nu^r) - \eta$$

for all r. If r is large enough so that $\rho(\mu^r, \mu)$, $\rho(\nu, \nu^r)$, and $|\rho^*(\mathbf{q}, \mu) - \rho^*(\mathbf{q}^r, \mu^r)|$ are all less than $\frac{\eta}{3}$, it follows that

$$\rho(\mu^r, \nu^r) < \rho^*(\mathbf{q}^r, \mu^r) + \varepsilon$$

and, hence, that $\nu^r \in \psi_\varepsilon(\mathbf{q}^r, \mu^r)$. This proves that ψ_ε is lower hemi-continuous. ■

LEMMA A.13 *The mapping* $\hat{\Phi} : C(X, \mathcal{M}^d(Z)) \to C(X, [0,1]^{2n+1})$ *is open.*

Proof. The lemma is equivalent to the statement that $\hat{\Phi}^{-1}$ is a lower hemi-continuous correspondence. To verify lower hemi-continuity, consider any functions $b \in C(X, \mathcal{M}^d(Z))$ and $\beta \in C(X, Q)$ such that $\beta = \hat{\Phi}(b)$. Consider any sequence $\{\beta^r\}$ in $C(X, Q)$ that converges to β. For any r, define a function b^r by setting

$$b^r(x) = \nu_{1/r}(\beta^r(x), b(x))$$

for any $x \in X$, where $\nu_{1/r}$ is the function given by Lemma A.12 with $\varepsilon = \frac{1}{r}$ Because $\nu_{1/r}$, β^r, and b, are continuous, $b^r \in C(X, \mathcal{M}^d(Z))$. Lemma A.12 implies that , for any $\gamma > 0$, there exists $\eta(\gamma) > 0$ such that, if $\|\hat{\phi}(b(x)) - \beta^r(x)\| < \eta(\gamma)$ for all x, then

$$\rho(b(x), b^r(x)) \leq \rho^*(\beta^r(x), b(x)) + \frac{1}{r}$$

for all x. By Lemma A.11, it follows that

$$\begin{aligned}\rho(b(x), b^r(x)) &\leq \max_i \left| \frac{\beta_i^r(x)}{1 - \beta_i^r(x)} - f_{b(x)}(z_i) \right| + \frac{1}{r} \\ &= \max_i \left| \frac{\beta_i^r(x)}{1 - \beta_i^r(x)} - \frac{\beta_i(x)}{1 - \beta_i(x)} \right| + \frac{1}{r}\end{aligned}$$

for all x, where $f_{b(x)}$ is the density of the measure $b(x)$. Now the uniform convergence of the sequence $\{\beta^r\}$ to β implies that, for any $\eta > 0$, there exists $R(\eta)$ such that for all $r \geq R(\eta)$, $\rho(b(x), b^r(x)) < \eta$ for all x. This proves that the sequence $\{b^r\}$ converges uniformly to b. Lower hemi-continuity of $\hat{\Phi}$ follows immediately. ■

A.3 Proof of Lemma 4.9

Proof. We first show that, if $\mathcal{M}_+^d(T)$ is endowed with the topology that is induced by the uniform topology for density functions, then the maps $\nu \to \psi_i(\nu) = (\bar{v}_i(\nu), b_i(\cdot, \nu))$ are continuous and open. We begin by proving continuity. If $\{\nu^k\}$ is a sequence in $\mathcal{M}_+^d(T)$ that converges to a limit $\nu \in \mathcal{M}_+^d(T)$, the associated densities satisfy

$$\lim_{k \to \infty} f^{\nu^k}(t_i, t_{-i}) = f^\nu(t_i, t_{-i}), \tag{63}$$

uniformly on T. Because the continuous function f^ν is bounded on the compact set T, it follows that the densities f^{ν^k} are uniformly bounded. For any i, therefore, Lebesgue's bounded convergence theorem implies that

$$\lim_{k\to\infty} \bar{f}_i^{\nu^k}(t_i) = \lim_{k\to\infty} \int_{T_{-i}} f^{\nu^k}(t_i, t_{-i}) dt_{-i} = \int_{T_{-i}} f^{\nu}(t_i, t_{-i}) dt_{-i} = \bar{f}_i^{\nu}(t_i), \tag{64}$$

uniformly on T_i, which proves that the sequence $\{\bar{\nu}_i(\nu^k)\}$ converges to $\bar{\nu}_i(\nu)$.

Because $\nu \in \mathcal{M}_+^d(T)$, we have $\bar{f}_i^{\nu}(t_i) > 0$ for all $t_i \in T_i$; indeed, because $\bar{f}_i^{\nu}(\cdot)$ is continuous, $\bar{f}_i^{\nu}(t_i)$ is bounded away from zero on T_i. Because the marginal densities $f_i^{\nu^k}$ converge uniformly to $\bar{f}_i^{\nu}$, it follows that they are uniformly bounded away from zero. If we combine (63) and (64) with (29) in the text, we obtain

$$\lim_{k\to\infty} \beta_i(t_{-i}|t_i, \nu^k) = \beta_i(t_{-i}|t_i, \nu), \tag{65}$$

uniformly on T. For any $t_i \in T_i$, therefore, the sequence $\{b_i(\cdot|t_i, \nu^k)\}$ converges to $b_i(\cdot|t_i, \nu)$, uniformly on T_i. Continuity of the map $\psi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu))$ is thus proved.

To show that ψ_i is open, we note that, for any ψ_i has an inverse. For any $\bar{\nu}_i \in \mathcal{M}_+^d(T_i)$ and any $b_i \in C(T_i, \mathcal{M}^d(T_{-i}))$, let $\bar{f}^{\bar{\nu}_i}$ and $\beta_i(\cdot|\cdot)$ be the associated marginal and conditional density functions and consider the measure $\psi_i^{-1}(\bar{\nu}_i, b_i)$ that is defined by the formula

$$\psi_i^{-1}(B_i \times B_{-i}|\bar{\nu}_i, b_i) = \int_{B_i} \int_{B_{-i}} \beta_i(t_{-i}|t_i) \bar{f}^{\bar{\nu}_i}(t_i) dt_{-i} dt_i. \tag{66}$$

By another application of Lebesgue's bounded convergence theorem, one sees that $\psi_i^{-1}(\bar{\nu}_i, b_i) \in \mathcal{M}_+^d(T)$ and that ψ_i^{-1} is continuous. Therefore ψ_i is open.

Next, we show that the topology that is induced by uniform convergence of densities is in fact the coarsest topology under which the maps $\nu \to \psi_i(\nu) = (\bar{\nu}_i(\nu), b_i(\cdot, \nu))$ Let B be any open subset of $\mathcal{M}_+^d(T)$ in the topology that is induced by the uniform topology for density functions and let $\mathcal{T}$ be any other topology on $\mathcal{N}_+^{db}(T)$ such that the maps ψ_i are continuous. Because, in the topology that is induced by the uniform topology for density functions, ψ_i is an open mapping, $\psi_i(B)$ is an open subset of $\mathcal{M}_+^d(T_i) \times C(T_i, \mathcal{M}^d(T_{-i}))$. Because ψ_i is continuous when $\mathcal{M}_+^d(T)$ has the topology $\mathcal{T}$, it follows that $\psi_i^{-1}(\psi_i(B)) = B \in \mathcal{T}$. Thus, the topology on $\mathcal{M}_+^d(T)$ that is induced by the uniform topology for density functions is at least as coarse as $\mathcal{T}$. ∎

References

ALLEN, B. (1981): "Generic Existence of Completely Revealing Equilibria for Economies with Uncertainty when Prices Convey Information," *Econometrica*, 49, 1173 – 1199. 5

BARELLI, P. (2009): "On the genericity of full surplus extraction in mechanism design," *J. of Econ. Theory*, 144, 1320 – 1333. 6, 23, 25

BERGEMANN, D., AND S. MORRIS (2005): "Robust mechanism design," *Econometrica*, 73, 1521–1534. 6

CHEN, Y., AND S. XIONG (2011): "The Genericity of Beliefs-Determine-Preferences Models Revisited," *Journal of Economic Theory*, 146, 751– 761. 6, 23, 24, 25, 27

CLAUSING, A. (1978): "Retractions and Open Mappings Between Convex Sets," *Mathematische Zeitschrift*, 160, 263 – 274. 33, 35

COMPTE, O., AND P. JEHIEL (2009): "Veto Constraint in Mechanism Design: Inefficiency with Correlated Types," *American Economic Journal: Microeconomics*, 1, 182 – 206. 4, 8

CRÉMER, J., AND R. MCLEAN (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica*, 56, 1247–1257. 2, 3, 4, 16

DEKEL, E., D. FUDENBERG, AND S. MORRIS (2006): "Topologies on Types," *Theoretical Economics*, 1, 275 – 309. 7

GIZATULINA, A., AND M. HELLWIG (2010): "Informational Smallness and the Scope for Limiting Informational Rents," *Journal of Economic Theory*, 145, 2260 – 2281. 2

——— (2011): "On the Robustness of the BDP Property for Families of Incomplete-Information Models," Working paper, Max Planck Institute for Research on Collective Goods. 6, 24, 25, 27

HEIFETZ, A., AND Z. NEEMAN (2006): "On the generic (im)possibility of full surplus extraction in mechanism design," *Econometrica*, 117, 213–233. 1, 2, 3, 4, 6, 8, 9, 23, 24, 25

HELLWIG, M. (1996): "Sequential Decisions Under Uncertainty and the Maximum Theorem," *Journal of Mathematical Economics*, 25, 443 – 464. 17

HILDENBRAND, W. (1974): *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton, NJ. 37

HIRSCH, M. (1994): *Differential Topology*. Springer Verlag, New-York. 5, 28

HUREWICZ, W., AND H. WALLMAN (1941): *Dimension Theory*. Princeton Mathematical Series, no. 4., Princeton, NJ. 5, 30

JORDAN, J. (1977): "The Continuity of Optimal Dynamic Decision Rules," *Econometrica*, 45, 1365 – 1376. 17

LEE, J. (2000): *Introduction to Topological Manifolds*. Springer, New York. 31

MCAFEE, P., AND P. RENY (1992): "Correlated information and mechanism design," *Econometrica*, 60, 395–421. 2, 16

MERTENS, J., AND S. ZAMIR (1985): "Formulation of Bayesian analysis for games with incomplete information," *International Journal of Game Theory*, 10, 619–632. 8

MICHAEL, E. (1964): "Three Mapping Theorems," *Proceedings of the American Mathematical Society*, 15, 410 – 415. 37

NEEMAN, Z. (2004): "The relevance of private information in mechanism design," *J. Econ. Theory*, 117, 55–77. 1, 2

PARTHASARATHY, K. (1967): *Probability Measures on Metric Spaces*. Academic Press, New York and London. 31, 32