Interdependent Preferences and Strategic Distinguishability

Dirk Bergemann[*] Stephen Morris[†] Satoru Takahashi[‡]

October 6, 2009

This extremely preliminary and incomplete was prepared for a seminar at Northwestern. Please do not quote without permission. You are advised to wait for a better draft before reading.

1 Introduction

Economists often assume that agents' preferences are interdependent for informational or psychological reasons. Thus one agent's preferences depend on another agent's preferences. This introduces a circularity in the description of preferences. In this paper, we introduce a "universal" type space of interdependent, or higher order, preferences analogous to the universal type space of higher order beliefs introduced by Mertens and Zamir (1985). Our

[*]Yale University, dirk.bergemann@yale.edu
[†]Princeton University, smorris@princeton.edu
[‡]Princeton University, satorut@princeton.edu

construction maintains the assumption of common certainty that (i) agents are expected utility maximizers; and (ii) there are not indifferent between all outcomes. Our space encompasses only the agents' actual preferences conditional on events the agent thinks possible, and does not incorporate "counterfactual" information about what the agent's preferences would be conditional on events he considers impossible.

Economists' traditional view of preferences is that they are not directly observed but are best understood as being "revealed" by agents' choices in actual or hypothetical decision problems. There is a developed revealed preference theory of individual choice behavior. If preferences are interdependent, then preferences can only be revealed by choices in "interactive decision problems," i.e., games. In this paper, we carry out a "revealed strategic preference" exercise. Suppose that we fix (i) any type space description of agents' higher order preferences over fixed outcome space; (ii) a game specifying a mapping from arbitrary action sets to the outcome space; and (iii) a solution concept specifying the possible action choices for any given type and game. We say that two types are "strategically distinguished" if there exists a finite game where those types could not choose the same action in equilibrium. We show that universal type space of higher order preferences characterizes strategic distiguishability, i.e., two types are strategically distinguished if and only if they map to distinct types in our universal type space.

Our motivation for studying this problem is threefold. First, we think the strategic revealed preference question is of intrinsic interest. Second, we highlight the limits of what can be deduced about agents interdependent preferences in static games from observed behavior only. For example, there is no way of distinguishing informational or psychological reasons for independent preferences. Third, our canonical description of types gives rise for a natural language to discuss implicit common knowledge assumptions in economic models.

We also discuss analogues of our main results for rationalizability. We show that there is a family of versions of rationalizability for which the same characterization of strategic distinguishability holds.

Our work relates to and relates together a number of strands in the literature.

1. In the framework of Savage, Epstein and Wang (1996) extend the Mertens-Zamir hierarchies of beliefs to incorporate non-subjective expected utility preferences such as ambiguity aversion, but maintain a weak form of monotonicity as well as a regularity condition. Our hierarchical structure dispenses with monotonicity but imposes independence, and thus an expected utility representation. Di Tillio (2008) relaxes Epstein and Wang's monotonicity. Instead, he strengthens Epstein and Wang's regularity condition to finiteness of states and outcomes.

2. Abreu and Matsushima (1992) provides a characterization of strategic distinguishability, in the course of proving results about virtual implementation under incomplete information. The characterization of our main result is of the same type. An advantage of our characterization is that it is terms of a universal space of expected utility preferences,

whereas their characterization depending on the finite space within which types were embedded.

3. Bergemann and Morris (2009) defined and characterized a "belief free" version of strategic distinguishability (depending only on payoff types) that was relevant in proving results about robust virtual implementation.

4. Mertens and Zamir (1985) and Brandenburger and Dekel (1993) construct the universal type space for a game with incomplete information, consisting of infinite hierarchies of beliefs. In their Remark (p. 3), Mertens and Zamir provide two sources of uncertainty: players' von Neumann-Morgenstern utility indices over outcomes and other relevant characteristics of the game. Since then, one approach in the subsequent literature (e.g., Ely and Pęski, 2006; Dekel, Fudenberg, and Morris, 2006, 2007) is to analyze hierarchies of beliefs about exogenously given states of the world, assuming common certainty about players' utility indices. Given such an interpretation of Mertens and Zamir's work, our universal type space is *larger* than the Mertens-Zamir universal type space, relaxing the common certainty of players' utility indices. On the other hand, if players have beliefs (and higher-order beliefs) about their utility indices, then our hierarchy of preferences corresponds to an *equivalence class* of Mertens-Zamir hierarchies of beliefs, identifying hierarchies with the same preference relation over outcomes, and our universal type space is a *coarsening* of the Mertens-Zamir universal

type space.[1]

5. Gul and Pesendorfer (2007) construct of a universal space of behavioral types (without any explicit consideration of uncertainty). Their purpose was to identify a rich set of possible types which capture all distinctions that can be expressed in a natural language. These distinctions are going to be much finer than those that can be revealed by rational behavior in strategic settings. In particular, their types reflect much counterfactual information (what preferences would be conditional on other agents' types). But we focus on information that can be revealed in static games, we cannot identify counterfactual information.

6. Battigalli and Siniscalchi (2003) introduce Δ-rationalizability. We will work with a concept of interim Φ-rationalizability that is an adaptation of their concept in the framework of interdependent preferences.

7. Ledyard (1986) was an early paper identifying revealed information from rational behavior in strategic settings. See also Haile, Hortaçsu, and Kosenok (2008)

2 The Main Question

An outside observer will see a finite set of players, $\mathcal{I} = \{1, ..., I\}$, making choices in strategic situations, where there is a finite set of "outcomes" Z and a compact and metrizable set of "observable states" Θ. A strategic

[1] This is parallel to the fact that a state-dependent preference can be represented by many different pairs of beliefs about states and tastes about outcomes.

situation is modelled as a mechanism where each player i has a finite set of actions A_i and an outcome function $g : \Theta \times A \to \mathbb{R}$. Thus a mechanism $\mathcal{M} = \left((A_i)_{i\in\mathcal{I}}, g\right)$. We are interested in what the outside observer can infer about players' (perhaps interdependent) preferences by observing choices in this setting.

Consider standard expected type space models of players' perhaps interdependent preferences. A belief-utility type space will consist, for each player i, of a measurable space of types T_i, a bounded and measurable utility function $u_i : \Theta \times T \times Z \to \mathbb{R}$ and a measurable belief function $\nu_i : T_i \to \Delta(\Theta \times T_{-i})$. Thus a belief-utility type space $\mathcal{T}_{\Delta u} = (T_i, u_i, \nu_i)_{i\in\mathcal{I}}$.

The pair $(\mathcal{T}_{\Delta u}, \mathcal{M})$ induce a game of incomplete information. A strategy for player i in this game is a measurable function $\sigma_i : T_i \to \Delta(A)$. We can extend the domain of g to mixed strategies in the usual way. Now strategy profile $\sigma = (\sigma_i)_{i\in\mathcal{I}}$ is a (Bayesian Nash) equilibrium of the game $(\mathcal{T}_{\Delta u}, \mathcal{M})$ if, for every $i \in \mathcal{I}$ and $t_i \in T_i$, $\sigma_i(t_i)$ maximizes

$$\int_{\Theta\times T_{-i}} u_i(\theta, t_i, t_{-i}, g(\theta, a_i, \sigma_{-i}(t_{-i})))\nu_i(t_i)(d(\theta, t_{-i}))$$

We write $E_i(t_i, \mathcal{T}_{\Delta u}, \mathcal{M})$ for the set of all pure actions type t_i plays with positive probability in some Bayesian-Nash equilibrium of $(\mathcal{T}_{\Delta u}, \mathcal{M})$.

We say that two types of player i, t_i in $\mathcal{T}_{\Delta u}$ and t'_i in $\mathcal{T}'_{\Delta u}$, are *strategically indistinguishable* if, for every mechanism $\mathcal{M}$, there exists some action that can be chosen by both types, so that $E_i(t_i, \mathcal{T}_{\Delta u}, \mathcal{M}) \cap E_i(t'_i, \mathcal{T}'_{\Delta u}, \mathcal{M}) \neq \emptyset$ for every $\mathcal{M}$. Conversely, t_i and t'_i are *strategically distinguishable* if there exists a mechanism in which no action can be chosen by both types, so that $E_i(t_i, \mathcal{T}_{\Delta u}, \mathcal{M}^*) \cap E_i(t'_i, \mathcal{T}'_{\Delta u}, \mathcal{M}^*) = \emptyset$ for some $\mathcal{M}^*$.

Our main result will be a characterization of strategic distinguishability.

3 Examples

3.1 Example 1: A Characterization of Strategic Distinguishability

Consider the following simple situation. There are two agents 1 and 2 and two equally likely but unobservable states, $\Omega = \{L, H\}$. The common value of the object to the two agents is 0 in state L, 90 in state H. Each agent i observes conditional independent signal $s_i \in \{l, h\}$, with $\Pr(l|L) = \Pr(h|H) = \frac{2}{3}$.

How can we represent this routine situation with our standard expected utility type spaces? We will give five equivalent alternatives. In each of them, the set of types of each agent will correspond to their possible signals, so $T_1 = T_2 = \{l, h\}$.

3.1.1 The "True Value" Representation

First, consider a "natural" representation of this story, where we allow for the product space of uncertainty $T_1 \times T_2 \times \Omega$ with eight states. The common prior on the eight states is represented by the following table:

$\omega = L$:

$t_1 \backslash t_2$	l	h
l	$\frac{2}{9}$	$\frac{1}{9}$
h	$\frac{1}{9}$	$\frac{1}{18}$

$\omega = H$:

$t_1 \backslash t_2$	l	h
l	$\frac{1}{18}$	$\frac{1}{9}$
h	$\frac{1}{9}$	$\frac{2}{9}$

The common valuation of the two agents as a function of the eight states is represented in the following table.

$\omega = L$:

$t_1 \backslash t_2$	l	h
l	0	0
h	0	0

$\omega = H$:

$t_1 \backslash t_2$	l	h
l	90	90
h	90	90

3.1.2 The "Conditional Value" Representation

But if the states L and H are never observed by the analyst or either agent, we know that we can integrate them out. Thus we can work with the reduced state space $T_1 \times T_2$ with four states. The common prior is now:

$t_1 \backslash t_2$	l	h
l	$\frac{5}{18}$	$\frac{2}{9}$
h	$\frac{2}{9}$	$\frac{5}{18}$

and the common valuation of the object

$t_1 \backslash t_2$	l	h
l	18	45
h	45	72

3.1.3 The "Independent Beliefs" Representation

But also, there is an indeterminacy in state-dependent expected utility representations: we cannot distinguish beliefs from state-dependence in utilities. Thus all we could ever observe is the size of the product of probability and value, and not the two components separately. This means that we can always substitute any type space where the product of probability and value is unchanged. This means in particular that we can always substitute independent types. Thus with space, $T_1 \times T_2$, we can have common prior

$t_1 \backslash t_2$	l	h
l	$\frac{1}{4}$	$\frac{1}{4}$
h	$\frac{1}{4}$	$\frac{1}{4}$

and the common valuation of the object

$t_1 \backslash t_2$	l	h
l	20	40
h	40	80

3.1.4 The "True Value Mertens Zamir" Representation

Mertens and Zamir (1985) suggest one canonical way of representing types: let there be a set of states capturing possible variation in payoffs and consider a type space embedding beliefs and higher order beliefs about payoffs. We could do this in our example by letting the set of payoff types for each agent be their possible valuations, so $V_1 = V_2 = \{0, 90\}$. Now the relevant space of uncertainty is $T_1 \times V_1 \times T_2 \times V_2$, with 16 states. Now we do not need to define values (they are embedded in the description of payoffs). Rather than define the common prior over 16 states, we can look at the interim beliefs of a particular type, type h of agent 1:

$t_2 = l$	0	90
0	$\frac{2}{9}$	0
90	0	$\frac{2}{9}$

$t_2 = h$	0	90
0	$\frac{1}{9}$	0
90	0	$\frac{4}{9}$

3.1.5 The "Conditional Value Mertens Zamir" Representation

But what are the right values to use in a Mertens-Zamir construction? We could also focus on the expected values conditional on types, so that $V_1 = V_2 = \{18, 45, 72\}$. Now the relevant space of uncertainty is $T_1 \times V_1 \times T_2 \times V_2$ with 36 states. Type h of agent 1 has beliefs:

$t_2 = l$	18	45	72
18	0	0	0
45	0	$\frac{4}{9}$	0
72	0	0	0

$t_2 = h$	18	45	72
18	0	0	0
45	0	0	0
72	0	0	$\frac{5}{9}$

3.1.6 Observable Implications

A canonical way to represent types is to define a type by how how it would behave. Consider type h of agent 1. What can we say about how this type will behave in difference strategic situations? A "first level" observation is that this type will have an unconditional willingness to pay for the object of 60 ($= \frac{2}{3} \times 90$). This is all we could find out about this type's preferences in a single person choice setting. But in a richer strategic setting, we could identify that type's willingness to pay for the object conditional on agent 2's unconditional willingness? In particular, type h of agent 1 will not pay anything to receive the object conditional on agent 2's unconditional willingness to pay being anything other than 30 or 60. Conditional on agent 2's unconditional willingness to pay being 30, agent 1 would be prepared to pay 20 ($= \frac{4}{9} \times 45$) for the object. Conditional on agent 2's unconditional willingness to pay being 60, agent 1 would be prepared to pay 40 ($= \frac{5}{9} \times 72$)

for the object.

Our main result will be a generalization of this example. In Section 4, we provide a formal description of a universal space of possible expected utility types, consisting of (i) unconditional (expected utility) preferences; (ii) preferences conditional on others' unconditional preferences; and so on. In Section 5, we confirm that two types are guaranteed to behave differently in equilibrium of some mechanism if and only if they correspond to different types in this universal space.

3.2 Example 2: Redundancy, Strategic Equivalence and Rationalizability

Our second example will illustrate a number of further issues that arise. Suppose we start with the situation described in example 1. But now assume agent i's valuation is the common value component (0 or 90) plus private value component x_i, where $x_i = 0$ if $s_i = h$ and $x_i = 30$ if $s_i = l$. Now we will have the same common prior probability distributions over type profiles (t_1, t_2):

$t_1 \backslash t_2$	l	h
l	$\frac{5}{18}$	$\frac{2}{9}$
h	$\frac{2}{9}$	$\frac{5}{18}$

but now we add 30 to the common conditional valuation of the low type only, giving valuation profiles:

$t_1 \backslash t_2$	l	h
l	$48, 48$	$75, 45$
h	$45, 75$	$72, 72$

Now the unconditional valuation of the low type of each agent is $\frac{4}{9}(48) + \frac{5}{9}(75) = 60$, while the unconditional valuation of the high type of each agent is $\frac{5}{9}(45) + \frac{4}{9}(72) = 60$, i.e., the same. This immediately implies that both types will map to the same type in the universal preference space, and will therefore be equilibrium strategically indistinguishable from each other and from any "complete information" type with common certainty that the unconditional valuation in 60. This example illustrates a form of "redundancy" analogous to (but different from) the redundancy of Mertens and Zamir (1985) and Dekel, Fudenberg and Morris (2007).

While types l, h and the complete information type with valuation 60 are strategically indistinguishable, it is easy to construct a game where equilibrium actions of one type are not equilibrium actions of the other type. Consider the two player game where each agent can (i) opt out; or (ii) opt in and pay 1 (for sure) and get the object and pay another 72 only if the other agent opts out.

	in	out
in	pay 1	pay 73 and get object
out	-	-

On the "reduced" complete information type space (without redundant types), each agent must opt out in equilibrium. But on the "rich" type space (with redundant types), there will be an strict equilibrium type h opts out and type l opts in.

This game illustrates the fact that two types may be strategically indistinguishable, i.e., have an equilibrium action in common in every game, but not be strategically equivalent, i.e., have the same set of equilibrium actions in all games.

In the example, opting out is also the unique interim correlated rationalizable (ICR) action (Dekel, Fudenburg and Morris (2007)) for type h and the complete information type, even though opting in is also an ICR action for type l. Thus the types are not ICR strategically equivalent.

In Section 6, we discuss a family of rationalizability notions and show that the characterization of equilibrium strategic distinguishability extends to this family. We also discuss why a clean characterization of strategic equivalence is elusive in this setting.

4 A Universal Space of Expected Utility Preferences

4.1 Preliminaries

4.1.1 Notation

The identity mapping on a set X is denoted id_X. For a collection of mappings $\varphi_\lambda\colon X_\lambda \to Y_\lambda$ indexed by $\lambda \in \Lambda$, we define the product mapping $\prod_\lambda \varphi_\lambda\colon X =$

$\prod_\lambda X_\lambda \to Y = \prod_\lambda Y_\lambda$ by

$$\forall x \in X, \quad \left(\prod_\lambda \varphi_\lambda\right)(x) = (\varphi_\lambda(x_\lambda))_{\lambda\in\Lambda}.$$

We also write $\varphi_1 \times \cdots \times \varphi_n$ for $\prod_\lambda \varphi_\lambda$ if $\Lambda = \{1, \ldots, n\}$. Similarly, for a collection of correspondences $\Gamma_\lambda \colon X_\lambda \rightrightarrows Y_\lambda$ indexed by $\lambda \in \Lambda$, we define the product correspondence $\prod_\lambda \Gamma_\lambda \colon X \rightrightarrows Y$ by

$$\forall x \in X, \quad \left(\prod_\lambda \Gamma_\lambda\right)(x) = \prod_\lambda \Gamma_\lambda(x_\lambda).$$

4.1.2 Signed Measures

For a measurable space (X, Σ) with set X and σ-algebra Σ, a set function $\mu \colon \Sigma \to \mathbb{R}$ is called a *signed measure* on (X, Σ) if it is σ-additive and $\mu(\emptyset) = 0$. For a signed measure μ on (X, Σ), define the *total variation of* μ by

$$||\mu|| = \sup\left\{\sum_{k=1}^n |\mu(E_k)| : \{E_1, \ldots, E_n\} \text{ is a partition of } X\right\}.$$

μ has a *bounded variation* if $||\mu|| < \infty$. A measurable set E is *μ-null* if $\mu(E') = 0$ for every $E' \in \Sigma$ with $E' \subseteq E$.

The set of all signed measures of bounded variation on (X, Σ) is denoted $ca(X, \Sigma)$ or $ca(X)$. We regard $ca(X, \Sigma)$ as a measurable space with the σ-algebra generated by a family of $\{\mu \in ca(X) : \mu(E) \geq p\}$, where E varies over Σ and p varies over $\mathbb{R}$. We also regard $ca(X, \Sigma)$ as a normed vector space with the variation norm.

A signed measure μ on (X, Σ) is a *probability* if it takes nonnegative values and $\mu(X) = 1$. The set of all probabilities on (X, Σ) is denoted $\Delta(X, \Sigma)$ or $\Delta(X)$.

Every finite set is endowed with the discrete topology. For a compact metrizable topological space (X, τ) with set X and a family τ of open sets, we endow X with the Borel σ-algebra $\mathcal{B}$ generated by τ and $ca(X, \mathcal{B})$ with the weak* topology generated by a family $\{\mu \in ca(X, \mathcal{B}) : p < \int f d\mu < q\}$, where f varies over continuous real-valued functions and p and q vary over $\mathbb{R}$. Note that the σ-algebra on $ca(X, \mathcal{B})$ coincides with the Borel σ-algebra generated by the weak* topology on $ca(X, \mathcal{B})$.[2]

For a sequence $\{X_n\}$ of measurable (compact metrizable, resp.) spaces, we endow the product set $\prod_n X_n$ with the product σ-algebra (the product topology, resp.). Note that the product space of compact metrizable spaces is also compact and metrizable, and its Borel σ-algebra coincides with the product of the Borel σ-algebras generated by the topologies on X_n.[3]

For a signed measure $\mu \in ca(X, \Sigma_X)$ and a measurable space (Y, Σ_Y), a measurable mapping $\varphi\colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ induces a signed measure $\varphi^{ca}(\mu) \in ca(Y, \Sigma_Y)$ defined by

$$\forall E \in \Sigma_Y, \quad (\varphi^{ca}(\mu))(E) = \mu(\varphi^{-1}(E)).$$

For a signed measure $\mu \in ca(X \times Y, \Sigma_X \otimes \Sigma_Y)$, the projection from $X \times Y$ to X induces the *marginal measure of* μ, $\operatorname{marg}_X \mu \in ca(X, \Sigma_X)$, which satisfies $(\operatorname{marg}_X \mu)(E) = \mu(E \times Y)$ for every $E \in \Sigma_X$.

[2] This claim is a simple extension of Kechris (Theorem 17.24) from probabilities to signed measures.

[3] See Kechris (p. 68).

4.1.3 State-Dependent Preferences

This subsection provides a definition of state-dependent preferences in the framework of Anscombe and Aumann (1963). We begin with a measurable space X of states and a finite set Z of outcomes with $|Z| \geq 2$. An (*Anscombe-Aumann*) *act* is a measurable mapping from X to $\Delta(Z)$. The set of all such acts is denoted by $F(X)$ and endowed with the sup norm. For $z, z' \in Z$ and measurable $E \subseteq X$, $z_E z'$ is the act that yields the point mass on z over E and the point mass on z' over $X \setminus E$. We consider the following conditions on binary relation $\succsim$ over $F(X)$.

1. *weak order*: for every $f, f' \in F(X)$, $f \succsim f'$ or $f' \succsim f$; for every $f, f', f'' \in F(X)$, if $f \succsim f'$ and $f' \succsim f''$, then $f \succsim f''$.

2. *independence*: for every $f, f', f'' \in F(X)$ and $\lambda \in (0, 1]$, $f \succsim f'$ if and only if $\lambda f + (1 - \lambda) f'' \succsim \lambda f' + (1 - \lambda) f''$.

3. *continuity*: for every $f, f', f'' \in F(X)$, if $f \succ f' \succ f''$, then there exists $\varepsilon \in (0, 1)$ such that $(1 - \varepsilon) f + \varepsilon f'' \succ f' \succ (1 - \varepsilon) f'' + \varepsilon f$.

4. *monotone continuity*: for every $z, z', z'' \in Z$ and decreasing sequence $\{E_n\}_{n \in \mathbb{N}}$ of measurable subsets of X with $\bigcap_n E_n = \emptyset$, if $z \succ z' \succ z''$, then there exists $n \in \mathbb{N}$ such that $z''_{E_n} z \succ z' \succ z_{E_n} z''$.

Let $P(X)$ be the set of all binary relations over $F(X)$ that satisfy weak order, independent, continuity, and monotone continuity. We have the following representation theorem.[4]

[4]We do not claim our originality for Proposition 1, for it is probably a well-known result. See Fishburn (1970, Theorem 13.1) for the case of finite X. Hill (2008) shows a

Proposition 1. *$\succsim \in P(X)$ if and only if there exists $\mu \in ca(X \times Z)$ that satisfies*

$$\forall f, f' \in F(X), \quad f \succsim f' \Leftrightarrow \int f(x)(z) d\mu(x,z) \geq \int f'(x)(z) d\mu(x,z).$$

Furthermore, μ and μ' represent the same preference in $P(X)$ if and only if there exist $a > 0$ and $\nu \in ca(X)$ such that $\mu'(E \times \{z\}) = a\mu(E \times \{z\}) + \nu(E)$ for every measurable $E \subseteq X$ and $z \in Z$.

Proof. It is easy to see that $\succsim$ represented by $\mu \in ca(X \times Z)$ satisfies weak order, independent, continuity, and monotone continuity. If $\succsim$ satisfies weak order, independent, and continuity, then, by the standard argument in the decision theory and the Riesz representation theorem (Aliprantis and Border, 1999, Theorem 13.4), $\succsim$ is represented by a signed charge (i.e., a finitely additive set function) μ of bounded variation on $X \times Z$. Since $\succsim$ satisfies monotone continuity, μ is σ-additive. The uniqueness of representations is standard. $\square$

$P(X)$ is endowed with the σ-algebra generated by a family of $\{\succsim \in P(X) : f \succsim f'\}$, where f and f' vary over $F(X)$. $\succsim \in P(X)$ is the *complete indifference*, denoted ind_X, if $f \sim f'$ for every $f, f' \in F(X)$. If X is a compact and metrizable topological space, then let $F_c(X) \subseteq F(X)$ be the set of all continuous acts over X, and endow $P(X)$ with the topology generated by $\{\text{ind}_X\}$ and a family of $\{\succsim \in P(X) : f \succ f'\}$, where f and f' vary over $F_c(X)$.

similar result in the Savage framework.

Fix arbitrary $z_0 \in Z$ and let $Z_0 = Z \setminus \{z_0\}$. For each $\succsim \in P(X)$, we normalize its representation $\mu \in ca(X \times Z)$ so that $X \times \{z_0\}$ is μ-null. In this case, we abuse notations and write $\mu \in ca(X \times Z_0)$. If $\succsim \neq \mathrm{ind}_X$, then we further normalize μ so that $||\mu|| = 1$.

Lemma 1. *If X is compact and metrizable, then $P(X)$ is compact and metrizable.*

Proof. Note that $P(X)$ is homeomorphic to $\{\mu \in ca(X \times Z_0) : ||\mu|| = 1\} \cup \{\bar{0}\}$, where $\bar{0} \in ca(X \times Z_0)$ defined by $\bar{0}(E) = 0$ for every measurable $E \subseteq X \times Z_0$. By the Riesz representation theorem (Aliprantis and Border, 1999, Corollary 13.15) and Alaoglu's theorem (Aliprantis and Border, 1999, Theorem 6.25), $\{\mu \in ca(X \times Z_0) : ||\mu|| = 1\}$ is compact and metrizable. Thus $P(X)$ is compact and metrizable. $\square$

An event E is *$\succsim$-null* if $z'_{E'}z \sim z$ for every $z, z' \in Z$ and measurable $E' \subseteq E$. For $\succsim$ represented by $\mu \in ca(X \times Z_0)$, E is $\succsim$-null if and only if $E \times Z_0$ is μ-null. An event E is *$\succsim$-certain* if $X \setminus E$ is $\succsim$-null.

For a preference $\succsim \in P(X)$ and a measurable space Y, a measurable mapping $\varphi \colon X \to Y$ induces a preference $\varphi^P(\succsim) \in P(Y)$ given by

$$\forall f, f'^P(\succsim)\ f' \Leftrightarrow f \circ \varphi \succsim f' \circ \varphi.$$

If $\succsim \in P(X)$ is represented by $\mu \in ca(X \times Z)$, then $\varphi^P(\succsim) \in P(Y)$ is represented by $\varphi^{ca}(\mu) \in ca(Y \times Z)$. For a preference $\succsim \in P(X \times Y)$, the projection from $X \times Y$ to X induces the *marginal preference of* $\succsim$, $\mathrm{marg}_X \succsim \in P(X)$, which is the restriction of $\succsim$ to acts over $X \times Y$ that do not depend on the Y-coordinate.

4.2 Interdependent Preferences

This section formulates type spaces for interdependent preferences and constructs "hierarchies of preferences" à la Mertens and Zamir (1985) and Brandenburger and Dekel (1993).

4.2.1 Type Spaces

Fix a finite set $\mathcal{I} = \{1, \ldots, I\}$ of players with $I \geq 2$ and a compact and metrizable set Θ of states of nature.

Definition 1. A *(preference-)type space* $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$ consists of, for each $i \in \mathcal{I}$, a measurable space T_i of player i's types and a measurable mapping $\pi_i \colon T_i \to P(\Theta \times T_{-i})$ that maps his types to preferences over acts over states of nature and his opponents' types, where $T_{-i} = \prod_{j \in \mathcal{I} \setminus \{i\}} T_j$.

Clearly, every belief-utility type space $\mathcal{T}_{\Delta u} = (T_i, u_i, \nu_i)_{i \in \mathcal{I}}$ is equivalent corresponds to a preference-type space $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$, where each $\pi_i(t_i)$ is the preference represented by (u_i, ν_i).

A *belief-type space* $\mathcal{T}_\Delta = (T_i, \pi_{i,\Delta})_{i \in \mathcal{I}}$ consists of, for each $i \in \mathcal{I}$, a measurable space T_i of player i's types and a measurable mapping $\pi_{i,\Delta} \colon T_i \to \Delta(\Theta \times T_{-i})$ that maps his types to beliefs over states of nature and his opponents' types. Assume that outcomes are private goods, $Z = \prod_i Z_i$, and that it is common knowledge that each player i has a non-constant von Neumann-Morgenstern utility index $u_i \colon Z_i \to \mathbb{R}$ over his private goods. We say that a belief-type space combined with utility indices over private goods is a *classical environment*. A classical environment induces a preference-type space

$\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$, where each $\pi_i(t_i)$ is represented by $\mu_i(t_i) \in ca(\Theta \times T_{-i} \times Z)$ with $\mu_i(t_i)(E \times \{z\}) = u_i(z_i)\pi_{i,\Delta}(t_i)(E)$ for every measurable $E \subseteq \Theta \times T_{-i}$.

4.2.2 The Universal Type Space

Let $X_0 = \Theta$ and $X_n = X_{n-1} \times P(X_{n-1})^{I-1}$ for each $n \geq 1$. Note that $X_n = \Theta \times \prod_{k=0}^{n-1} P(X_k)^{I-1}$. Let $X_\infty = \Theta \times \prod_{n=0}^{\infty} P(X_n)^{I-1}$. By Lemma 1, each X_n is compact and metrizable, and thus X_∞ is compact and metrizable. Let $Y_0 = \prod_{n=0}^{\infty} ca(X_n \times Z_0)$ be the set of hierarchies of signed measures. A hierarchy of signed measures, $\{\mu_n\}_{n=1}^{\infty} \in Y_0$, is *uniformly bounded* if $\sup_n ||\mu_n|| < \infty$; *coherent* if $\operatorname{marg}_{X_{n-2} \times Z_0} \mu_n = \mu_{n-1}$ for every $n \geq 2$. Let $Y_1 \subset Y_0$ be the set of all uniformly bounded and coherent hierarchies of signed measures.

For each $\mu_n \in ca(X_{n-1} \times Z_0)$ with $n \geq 1$, let $\rho_n(\mu_n) \in P(X_{n-1})$ denote the preference represented by μ_n. Let $\rho = \prod_{n=1}^{\infty} \rho_n \colon Y_0 \to \prod_{n=0}^{\infty} P(X_n)$ be the product of mappings ρ_n. Similarly, for each $\mu_\infty \in ca(X_\infty \times Z_0)$, let $\rho_\infty(\mu_\infty) \in P(X_\infty)$ denote the preference represented by μ_∞. By coherency, for every $\{\mu_n\}, \{\mu'_n\} \in Y_1$, we have $\rho(\{\mu_n\}) = \rho(\{\mu'_n\})$ if and only if there exists $a > 0$ such that $\{\mu'_n\} = \{a\mu_n\}$, i.e., $\mu'_n = a\mu_n$ for every $n \geq 1$.

Lemma 2. *There is a homeomorphism $\psi_{ca} \colon Y_1 \to ca(X_\infty \times Z_0)$ that preserves scaler multiplication.*

Proof. Pick any $\{\mu_n\} \in Y_1$. By Kolmogorov's extension theorem for signed measures (Haagen, 1981, Theorem 2.4), there exists $\mu_\infty \in ca(X_\infty \times Z_0)$ such that $\operatorname{marg}_{X_{n-1} \times Z_0} \mu_\infty = \mu_n$ for every $n \geq 1$. Let $\psi_{ca}(\{\mu_n\}) = \mu_\infty$. Similarly to Brandenburger and Dekel (1993, Proposition 1), both ψ_{ca} and

ψ_{ca}^{-1} are continuous. We have $\psi_{ca}(\{a\mu_n\}) = a\psi_{ca}(\{\mu_n\})$ for every $a \in \mathbb{R}$ and $\{\mu_n\} \in Y_1$. $\square$

Let $T_1 = \rho(Y_1) \subset \prod_{n=0}^{\infty} P(X_n)$. Note that every $\{\succsim_n\}_{n=1}^{\infty} \in T_1$ not only satisfies coherency, i.e., $\operatorname{marg}_{X_{n-2}} \succsim_n = \succsim_{n-1}$ for every $n \geq 2$, but also inherits a certain regularity conditon from the uniform boundedness of the hierarchies of signed measures in Y_1.

Lemma 3. *There is a measurable isomorphism $\psi_P \colon T_1 \to P(X_\infty)$.*

Proof. Let $\psi_P = \rho_\infty \circ \psi_{ca} \circ \rho^{-1}$. This mapping is well defined and bijective since ψ_{ca} preserves scaler multiplication, and, for each $\succsim \in P(X_\infty)$ and $\{\succsim_n\} \in T_1$, $\rho_\infty^{-1}(\succsim)$ and $\rho^{-1}(\{\succsim_n\})$ are unique up to positive multiplication. Both ψ_P and ψ_P^{-1} are measurable since ψ_{ca} is a homeomorphism, and ρ_∞ and ρ are measurable isomorphisms up to positive multiplication. $\square$

Note that ψ_P is not continuous. For example, $\{\mathrm{ind}_{X_\infty}\}$ is open in $P(X_\infty)$, but its inverse image $\{\{\mathrm{ind}_{X_n}\}_{n=0}^{\infty}\}$ is not open in T_1.[5]

For $n \geq 2$, let

$$T_n = \{t \in T_1 : \Theta \times (T_{n-1})^{I-1} \text{ is } \psi_P(t)\text{-certain}\}$$

and $T^* = \bigcap_{n=1}^{\infty} T_n$. Note that T_n is compact for every $n \geq 1$, and hence T^* is also compact.

Proposition 2. *There is a measurable isomorphism $\pi^* \colon T^* \to P(\Theta \times (T^*)^{I-1})$.*

[5]If we changed the definition of the topology on $P(X)$ for compact metrizable X so that the singleton $\{\mathrm{ind}_X\}$ were not open, then ψ_P would be a homeomorphism. In this case, however, the topology on $P(X)$ would be non-Hausdorff, and hence non-metrizable.

Proof. By the monotone continuity of preferences, we have $T^* = \{t \in T_1 : \Theta \times (T^*)^{I-1}$ is $\psi_P(t)$-certain$\}$. Since ψ_P is onto, we have $\psi_P(T^*) = \{\succsim \in P(X_\infty) : \Theta \times (T^*)^{I-1}$ is $\succsim$-certain$\}$, which is measurable isomorphic to $P(\Theta \times (T^*)^{I-1})$. $\square$

Definition 2. The *universal type space* is a type space $\mathcal{T}^* = (T_i^*, \pi_i^*)_{i\in\mathcal{I}}$ with $T_i^* = T^*$ and $\pi_i^* = \pi^*$ for every $i \in \mathcal{I}$.

Definition 3. For two type spaces $\mathcal{T} = (T_i, \pi_i)_{i\in\mathcal{I}}$ and $\mathcal{T}' = (T_i', \pi_i')_{i\in\mathcal{I}}$, a profile $\varphi = (\varphi_i)_{i\in\mathcal{I}}$ of measurable mappings $\varphi_i \colon T_i \to T_i'$ *preserves preferences* if $\pi_i' \circ \varphi_i = (\mathrm{id}_\Theta \times \varphi_{-i})^P \circ \pi_i$ for every $i \in \mathcal{I}$, where $\varphi_{-i} = \prod_{j\in\mathcal{I}\setminus\{i\}} \varphi_j$.

Proposition 3. *Every type space $\mathcal{T} = (T_i, \pi_i)_{i\in\mathcal{I}}$ has a preference-preserving mapping $\hat{\pi} = (\hat{\pi}_i)_{i\in\mathcal{I}}$ from $\mathcal{T}$ to the universal type space $\mathcal{T}^*$.*

Proof. For every $i \in \mathcal{I}$ and $t_i \in T_i$, there exists $\mu_i(t_i) \in ca(\Theta \times T_{-i} \times Z_0)$ that represents $\pi_i(t_i) \in P(\Theta \times T_{-i})$. Define $\hat{\pi}_{i,1}(t_i) = \mathrm{marg}_\Theta\, \pi_i(t_i) \in P(X_0)$, $\hat{\mu}_{i,1}(t_i) = \mathrm{marg}_{\Theta\times Z_0}\, \mu_i(t_i) \in ca(X_0 \times Z_0)$, and, for each $n \geq 2$,

$$\begin{aligned}
\hat{\pi}_{i,n}(t_i) &= \left(\mathrm{id}_\Theta \times \hat{\pi}_{-i}^{n-1}\right)^P (\pi_i(t_i)) \in P(X_{n-1}),\\
\hat{\mu}_{i,n}(t_i) &= \left(\mathrm{id}_\Theta \times \hat{\pi}_{-i}^{n-1} \times \mathrm{id}_{Z_0}\right)^{ca} (\mu_i(t_i)) \in ca(X_{n-1} \times Z_0),
\end{aligned}$$

where $\hat{\pi}_{-i,n} = \prod_{j\in\mathcal{I}\setminus\{i\}} \hat{\pi}_{j,n}$ and $\hat{\pi}_{-i}^n(t_{-i}) = \{\hat{\pi}_{-i,k}(t_{-i})\}_{k=1}^n$ for each $t_{-i} \in T_{-i}$. Let $\hat{\pi}_i(t_i) = \{\hat{\pi}_{i,n}(t_i)\}_{n=1}^\infty$ and $\hat{\mu}_i(t_i) = \{\hat{\mu}_{i,n}(t_i)\}_{n=1}^\infty$. Since $\hat{\mu}_{i,n}(t_i)$ represents $\hat{\pi}_{i,n}(t_i)$ for every $n \geq 1$ and $\hat{\mu}_i(t_i)$ is uniformly bounded and coherent, we have $\hat{\pi}_i(t_i) \in T_1$. Since $\hat{\pi}_i(t_i) \in T_1$ for every $i \in \mathcal{I}$ and $t_i \in T_i$, we have $\hat{\pi}_i(t_i) \in T^*$ for every $i \in \mathcal{I}$ and $t_i \in T_i$.

Let $\hat{\pi}_{-i} = \prod_{j\in\mathcal{I}\setminus\{i\}} \hat{\pi}_j$. Then we have $\pi^* \circ \hat{\pi}_i = \rho_\infty \circ \psi_{ca} \circ \rho^{-1} \circ \hat{\pi}_i = \rho_\infty \circ \psi_{ca} \circ \hat{\mu}_i = \rho_\infty \circ (\mathrm{id}_\Theta \times \hat{\pi}_{-i} \times \mathrm{id}_{Z_0})^{ca} \circ \mu_i = (\mathrm{id}_\Theta \times \hat{\pi}_{-i})^P \circ \pi_i$, thus $\hat{\pi} = (\hat{\pi}_i)_{i\in\mathcal{I}}$ preserves preferences. $\square$

We say that $\hat{\pi}_{i,n}(t_i)$ in the proof is the *n-th order preference of t_i*, and $\hat{\pi}_i(t_i)$ the *hierarchy of preferences of t_i*.

5 Equilibrium Strategic Distinguishability Result

To give a characterization of equilibrium strategic distinguishability, we must impose two restrictions on types.

First, in order to have existence of equilibrium, we will restrict attention to countable type spaces. Thus a belief-utility type space $\mathcal{T}_{\Delta u} = (T_i, u_i, \nu_i)_{i\in\mathcal{I}}$. is countable if each T_i is countable. A type t_i is countable if it belongs to countable type space $\mathcal{T}_{\Delta u} = (T_i, u_i, \nu_i)_{i\in\mathcal{I}}$.

Second, we will require a "boundedness" restriction on belief-utility types. This will require there is a uniform bound on how much utility conditional on others' types can vary. To give this bound bite, we must define the bound relative to differences in unconditional utility. Thus if we fix a belief-utility type space $\mathcal{T}_{\Delta u} = (T_i, u_i, \nu_i)_{i\in\mathcal{I}}$, write $\overline{u}_i : \Theta \times T_i \times Z \to \mathbb{R}$ for agent i's payoff unconditional on other agents' types:

$$\overline{u}_i(\theta, t_i, z) = \int_{T_{-i}} u_i(\theta, t_i, t_{-i}, z)\, \nu_i(t_i)(\{\theta\} \times dt_{-i}).$$

Now say that belief-utility type space $\mathcal{T}_{\Delta u}= (T_i, u_i, \nu_i)_{i\in\mathcal{I}}$ is bounded by $K < \infty$ if, for every $i \in \mathcal{I}$ and $t_i \in T_i$, there exist $\overline{\theta} \in \Theta$ and $\overline{z}, \overline{z}' \in Z$ such that

$$\overline{u}_i\left(\overline{\theta}, t_i, \overline{z}\right) \neq \overline{u}_i\left(\overline{\theta}, t_i, \overline{z}'\right) \text{ and } \nu_i\left(t_i\right)\left(\left\{\overline{\theta}\right\} \times T_{-i}\right) > 0$$

and

$$\left|u_i\left(\theta, t_i, t_{-i}, z\right) - u_i\left(\theta, t_i, t'_{-i}, z'\right)\right| \leq K\left|\overline{u}_i\left(\overline{\theta}, t_i, \overline{z}\right) - \overline{u}_i\left(\overline{\theta}, t_i, \overline{z}'\right)\right|$$

for every $\theta \in \Theta$, t_{-i}, $t'_{-i} \in T_{-i}$, and $z, z' \in Z$. A type t_i is bounded if it belongs to a bounded type space $\mathcal{T}_{\Delta u}= (T_i, u_i, \nu_i)_{i\in\mathcal{I}}$.

Now we have:

THEOREM. Suppose types t_i (in $\mathcal{T}$) and t'_i (in $\mathcal{T}'$) are countable and bounded. Then they are strategically indistinguishable if and only if they map to the same type in the universal space of expected utility preferences (i.e., $\hat{\pi}_i(t_i) = \hat{\pi}_i(t_i)$).

This theorem follows from Propositions 4 and 6 in the next section.

6 Rationalizability

In this section, we introduce a family of rationalizability solution concepts and then simultaneously prove results for various notions of rationalizability and equilibrium.

6.1 Solution Concepts

6.1.1 Local Bayesian-Nash Equilibrium

(Bayesian-Nash) equilibria do not necessarily on large type spaces, including, in particular, the universal type space. However, even when equilibria do not exist on large type spaces, equilibria may exist on closed subsets of the large type space. Since such "local" equilibria will be useful to work with in our arguments, we formally define them.

A product set $\tilde{T} = \prod_i \tilde{T}_i$ of types is a *preference-closed subspace of* $\mathcal{T}$ if, for every $i \in \mathcal{I}$ and $t_i \in \tilde{T}_i$, $\tilde{T}_i \subseteq T_i$ and $\Theta \times \tilde{T}_{-i}$ is $\pi_i(t_i)$-certain.

Definition 4. A profile of behavioral strategies, $\sigma = (\sigma_i)_{i \in \mathcal{I}}$ with measurable $\sigma_i \colon T_i \to \Delta(A_i)$, is a *local equilibrium* if there exists a preference-closed subspace $\tilde{T}$ of $\mathcal{T}$ such that

$$\forall i \in \mathcal{I},\ t_i \in \tilde{T}_i,\ a_i \in A_i, \quad g(\cdot, \sigma_i(t_i), \cdot) \circ (\mathrm{id}_\Theta \times \sigma_{-i})\ \pi_i(t_i)\ g(\cdot, a_i, \cdot) \circ (\mathrm{id}_\Theta \times \sigma_{-i}),$$

where $\sigma_{-i} = \prod_{j \in \mathcal{I} \setminus \{i\}} \sigma_j$.

Let $LE_i(t_i)$ be the set of all local equilibrium plays of type t_i. By Kakutani's fixed-point theorem, $LE_i(t_i) \neq \emptyset$ if t_i is a countable type, i.e., there exists a finite preference-closed subspace $\tilde{T} = \prod_j \tilde{T}_j$ with $T_i \ni t_i$.

6.1.2 Interim Rationalizability

Let $\Phi = (\Phi_i)_{i \in \mathcal{I}}$ be a profile of correspondences $\Phi_i \colon T_i \rightrightarrows P(\Theta \times T_{-i} \times A_{-i})$, where $\Phi_i(t_i)$ is the set of all possible interim preferences of type t_i over acts over states of nature as well as opponents' types and actions.

Definition 5. An action $a_i \in A_i$ is a Φ_i-*best reply for type* $t_i \in T_i$ *against* $\Gamma_{-i} = \prod_{j \in \mathcal{I} \setminus \{i\}} \Gamma_j$ *with* $\Gamma_j \colon T_j \rightrightarrows A_j$ if there exists $\succsim_i \in \Phi_i(t_i)$ such that $\Theta \times \mathrm{graph}(\Gamma_{-i})$ is $\succsim_i$-certain and

$$\forall a_i' \in A_i, \quad g(\cdot, a_i, \cdot) \ (\mathrm{marg}_{\Theta \times A_{-i}} \succsim_i) \ g(\cdot, a_i', \cdot).$$

$\Gamma = (\Gamma_i)_{i \in \mathcal{I}}$ is a Φ-*best reply correspondence* if, for every $i \in \mathcal{I}$, $t_i \in T_i$, and $a_i \in \Gamma_i(t_i)$, a_i is a Φ_i-best reply for type t_i against Γ_{-i}. An action a_i is *interim* Φ-*rationalizable for type* t_i if there exists a Φ-best reply correspondence Γ with $\Gamma_i(t_i) \ni a_i$.

Let $R_i^{\Phi}(t_i)$ be the set of interim Φ-rationalizable actions for type t_i. Let $R_{i,0}^{\Phi}(t_i) = A_i$ for every $i \in \mathcal{I}$ and $t_i \in T_i$, and, for every $n \geq 1$, let $R_{i,n}^{\Phi}(t_i)$ be the set of preference-correlated best replies for type t_i against $R_{-i,n-1}^{\Phi}$. We have $R_i^{\Phi}(t_i) \subseteq \bigcap_{n \geq 0} R_{i,n}^{\Phi}(t_i)$.[6]

Φ is *compatible with type space* $\mathcal{T}$ if, for every $i \in I$, $t_i \in T_i$, and $\succsim_i \in \Phi_i(t_i)$, we have $\mathrm{marg}_{\Theta \times T_{-i}} \succsim_i = \pi_i(t_i)$. For every $i \in \mathcal{I}$ and $t_i \in T_i$, let $\Phi_i^{\mathrm{PC}}(t_i)$ be the set of preferences $\succsim_i \in P(\Theta \times T_{-i} \times A_{-i})$ such that $\mathrm{marg}_{\Theta \times T_{-i}} \succsim_i = \pi_i(t_i)$. Note that $\Phi^{\mathrm{PC}} = (\Phi_i^{\mathrm{PC}})_{i \in \mathcal{I}}$ is the weakest restriction compatible with $\mathcal{T}$. We call interim Φ^{PC}-rationalizability *interim preference-correlated rationalizability.* In the setup of games with complete information, where each player has only one type, Morris and Takahashi (2009) argue that preference-correlated rationalizability is the implication of common certainty of rationality.

[6] Probably we also have $R_i^{\Phi}(t_i) \supseteq \bigcap_{n \geq 0} R_{i,n}^{\Phi}(t_i)$, but we will not use this direction of set inclusion in this paper.

Consider a classical environment with $Z = \prod_i Z_i$ and each type t_i's preference $\pi_i(t_i)$ is represented by her von Neumann-Morgenstern utility index $u_i\colon Z_i \to \mathbb{R}$ and belief $\pi_{i,\Delta}(t_i) \in \Delta(\Theta \times T_{-i})$. For every $i \in \mathcal{I}$ and $t_i \in T_i$, let $\Phi_i^{\mathrm{BC}}(t_i)$ be the set of preferences $\succsim_i \in P(\Theta \times T_{-i} \times A_{-i})$ represented by u_i and some $\mu_{i,\Delta} \in \Delta(\Theta \times T_{-i} \times A_{-i})$ with $\operatorname{marg}_{\Theta \times T_{-i}} \mu_{i,\Delta}(t_i) = \pi_{i,\Delta}(t_i)$. Interim Φ^{BC}-rationalizability corresponds to interim belief-correlated rationalizability proposed by Dekel, Fudenberg, and Morris (2007). Define $\Phi_i^{\mathrm{BI}}(t_i) \subseteq \Phi_i^{\mathrm{BC}}(t_i)$ by further imposing stochastic independence on $\mu_{i,\Delta}(t_i)$ between nature and player i's opponents: there exists $\sigma_{-i} = \prod_{j \in \mathcal{I} \setminus \{i\}} \sigma_j$ with measurable mixed strategies $\sigma_j\colon T_j \to \Delta(A_j)$ such that

$$\mu_{i,\Delta}(t_i)(E \times \{a_{-i}\}) = \int_{(\theta, t_{-i}) \in E} \sigma_{-i}(t_{-i})(a_{-i}) d\pi_{i,\Delta}(t_i)(\theta, t_{-i})$$

for every measurable $E \subseteq \Theta \times T_{-i}$ and $a_{-i} \in A_{-i}$. Interim Φ^{BI}-rationalizability corresponds to interim belief-independent rationalizability (Ely and Pęski, 2006).

6.2 Strategic Distinguishability

In this section, we introduce the social planner's viewpoint and discuss whether and how the social planner can construct a mechanism to distinguish types from others in terms of their behavior. We show that, under a mild richness condition and a boundedness condition on Φ, strategic distinguishability under interim Φ-rationalizability is essentially equivalent to having different hierarchies of preferences. This result gives an operational definition to hierarchies of preferences.

Note that we consider various type spaces $\mathcal{T}$ and mechanisms $\mathcal{M}$. To emphasize their dependency on type spaces and/or mechanisms, we write $\hat{\pi}_i(t_i, \mathcal{T}) = \{\hat{\pi}_{i,n}(t_i, \mathcal{T})\}$ for the hierarchy of type t_i's preferences in type space $\mathcal{T}$, $\Phi_i(t_i, \mathcal{T}, \mathcal{M})$ for restrictions on type t_i's preferences in game $(\mathcal{T}, \mathcal{M})$, $LE_i(t_i, \mathcal{T}, \mathcal{M})$ for the set of local Bayesian-Nash equilibrium plays of type t_i in $(\mathcal{T}, \mathcal{M})$, and $R_i^{\Phi}(t_i, \mathcal{T}, \mathcal{M})$ for the set of interim Φ-rationalizable actions of type t_i in $(\mathcal{T}, \mathcal{M})$.

Definition 6. A mechanism $\mathcal{M}$ *distinguishes types* $t_i \in T_i$ *in* $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$ *and* $t_i' \in T_i'$ *in* $\mathcal{T}' = (T_i', \pi_i')_{i \in \mathcal{I}}$ *under interim* Φ*-rationalizability* if $R_i^{\Phi}(t_i, \mathcal{T}, \mathcal{M}) \cap R_i^{\Phi}(t_i', \mathcal{T}', \mathcal{M}) = \emptyset$. Types t_i and t_i' are *distinguishable under interim* Φ*-rationalizability* if there exists a mechanism that distinguishes t_i and t_i'.

Note that our exercises of strategic distinguishability make sense only under the implicit assumption that type spaces are defined independently of mechanisms. If players made inferences and changed their preferences based on the social planner's choice of a mechanism, then the social planner could only solicit the information about players' preferences *conditional on* the particular mechanism, and would not be able to extrapolate it to other mechanisms. Note also that using strategic distinguishability under interim Φ-rationalizability implies that the social planner *a priori* assumes common certainty of rationality as well as restrictions Φ on players' preferences.

Our definition of distinguishability is in a strong form: the social planner can construct a mechanism that *always* reveals differences of two different types. A weaker form of distinguishability is to construct a mechanism that *possibly* reveals their differences: there exists a mechanism $\mathcal{M}$ such that

$R_i^\Phi(t_i, \mathcal{T}, \mathcal{M}) \neq R_i^\Phi(t'_i, \mathcal{T}', \mathcal{M})$.[7]

6.2.1 Necessary Conditions

Here we show that, in order for two types to be strategically distinguishable, they need to have different hierarchies of preferences.

Lemma 4. *For every pair of type spaces $\mathcal{T}$ and $\mathcal{T}'$ and mechanism $\mathcal{M}$, if $\varphi = (\varphi_i)_{i\in\mathcal{I}}$ is a preference-preserving mapping from $\mathcal{T}$ to $\mathcal{T}'$, then*

$$LE_i(t_i, \mathcal{T}, \mathcal{M}) \supseteq LE_i(\varphi_i(t_i), \mathcal{T}', \mathcal{M})$$

for every $i \in \mathcal{I}$, $t_i \in T_i$.

Proof. Pick any local Bayesian-Nash equilibrium σ' in $(\mathcal{T}', \mathcal{M})$ associated with preference-closed subspace $\tilde{T}' = \prod_i \tilde{T}'_i$ of $\mathcal{T}'$. Let $\tilde{T} = \prod_i \varphi_i^{-1}(\tilde{T}'_i)$ and $\sigma = (\sigma'_i \circ \varphi_i)_{i\in I}$. Since φ preserve preferences, $\tilde{T}$ is a preference-closed sub-

[7] There are two open questions concerning the weak form of distinguishability.

1. Is the weak form of distinguishability equivalent to the strong form of distinguishability?
2. If not, what is the "right" definition of hierarchies of preferences?

In classical environments, Dekel, Fudenberg, and Morris (2006, 2007) show the equivalence between the strong and weak forms of distinguishability under interim belief-correlated rationalizability. In general, the weak form of distinguishability leads to a finer description of hierarchies than the standard universal type space. Ely and Pęski (2006) and Sadzik (2008) provide definitions of "extended" hierarchies of beliefs corresponding to the weak form of distinguishability under interm belief-independent rationalizability and local Bayesian-Nash equilibrium, respectively.

space of $\mathcal{T}$ and σ is a local Bayesian-Nash equilibrium in $(\mathcal{T}, \mathcal{M})$ associated with $\tilde{T}$. $\square$

Proposition 4. *For every pair of finite type spaces $\mathcal{T}$ and $\mathcal{T}'$ and mechanism $\mathcal{M}$, we have*

$$\hat{\pi}_i(t_i, \mathcal{T}) = \hat{\pi}_i(t'_i, \mathcal{T}') \Rightarrow LE_i(t_i, \mathcal{T}, \mathcal{M}) \cap LE_i(t'_i, \mathcal{T}', \mathcal{M}) \neq \emptyset$$

for every $i \in \mathcal{I}$, $t_i \in T_i$, and $t'_i \in T'_i$.

Proof. By Proposition 3, $\hat{\pi}(\cdot, \mathcal{T})$ and $\hat{\pi}(\cdot, \mathcal{T}')$ are preference-preserving mappings from $\mathcal{T}$ and $\mathcal{T}'$ to the universal type space $\mathcal{T}^*$, respectively. By Lemma **??**, $LE_i(t_i, \mathcal{T}, \mathcal{M}) \cap LE_i(t'_i, \mathcal{T}', \mathcal{M}) \supseteq LE_i(t^*_i, \mathcal{T}^*, \mathcal{M})$, where $t^*_i = \hat{\pi}_i(t_i, \mathcal{T}) = \hat{\pi}_i(t'_i, \mathcal{T}')$. Since $\hat{\pi}(T, \mathcal{T}) \cap \hat{\pi}(T', \mathcal{T}')$ is a finite preference-closed subspace of $\mathcal{T}^*$, we have $LE_i(t^*_i, \mathcal{T}^*, \mathcal{M}) \neq \emptyset$. $\square$

In order to relate local Bayesian-Nash equilibrium with interim Φ-rationalizability, we impose a mild "richness" condition on Φ as follows.

Definition 7. Φ *allows randomization in* $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$ if, for every $i \in \mathcal{I}$, $t_i \in T_i$, $\mathcal{M}$, and $\sigma_{-i} = (\sigma_j)_{j \in \mathcal{I} \setminus \{i\}}$ with measurable $\sigma_j \colon T_j \to \Delta(A_j)$, there exists $\succsim_i \in \Phi_i(t_i)$ such that $\Theta \times \operatorname{graph}\left(\prod_{j \in \mathcal{I} \setminus \{i\}} \operatorname{supp} \sigma_j(\cdot)\right)$ is $\succsim_i$-certain and

$$\begin{aligned} \forall a_i,\, a'_i \in A_i, \quad & g(\cdot, a_i, \cdot) \circ (\mathrm{id}_\Theta \times \sigma_{-i}) \ \pi_i(t_i) \ g(\cdot, a'_i, \cdot) \circ (\mathrm{id}_\Theta \times \sigma_{-i}) \\ & \Leftrightarrow g(\cdot, a_i, \cdot) \ \operatorname{marg}_{\Theta \times A_{-i}} \succsim_i g(\cdot, a'_i, \cdot). \end{aligned}$$

That is, if Φ allows randomization in $\mathcal{T}$, then every type in $\mathcal{T}$ can have a belief that the opponents can choose mixed actions independently across players. In a classical environment, allowing randomization is equivalent to having $\Phi(\cdot, \mathcal{T}, \cdot) \supseteq \Phi^{\mathrm{BI}}(\cdot, \mathcal{T}, \cdot)$.

Proposition 5. *For every pair of finite type spaces $\mathcal{T}$ and $\mathcal{T}'$ and mechanism $\mathcal{M}$, if restriction Φ that is rich in both $\mathcal{T}$ and $\mathcal{T}'$, then we have*

$$\hat{\pi}_i(t_i, \mathcal{T}) = \hat{\pi}_i(t'_i, \mathcal{T}') \Rightarrow R_i^{\Phi}(t_i, \mathcal{T}, \mathcal{M}) \cap R_i^{\Phi}(t'_i, \mathcal{T}', \mathcal{M}) \neq \emptyset$$

for every $i \in I$, $t_i \in T_i$, and $t'_i \in T'_i$.

Proof. Since Φ is rich in $\mathcal{T}$ and $\mathcal{T}'$, by Proposition **??**, we have $R_i^{\Phi}(t_i, \mathcal{T}, \mathcal{M}) \cap R_i^{\Phi}(t'_i, \mathcal{T}', \mathcal{M}) \supseteq LBNE_i(t_i, \mathcal{T}, \mathcal{M}) \cap LBNE_i(t'_i, \mathcal{T}', \mathcal{M}) \neq \emptyset$. □

6.2.2 Sufficient Conditions

Φ is *compatible with type space* $\mathcal{T}$ if, for every mechanism $\mathcal{M}$, $\Phi(\cdot, \mathcal{T}, \mathcal{M})$ is compatible with $\mathcal{T}$. We impose the following boundedness condition on Φ.

Definition 8. Φ is *uniformly bounded by $K < \infty$ in $\mathcal{T}$* if, for every $i \in I$, $t_i \in T_i$, $\mathcal{M}$, and $\succsim_i \in \Phi_i(t_i, \mathcal{T}, \mathcal{M})$, there exists $\mu_i \in ca(\Theta \times T_{-i} \times A_{-i} \times Z_0)$ that represents $\succsim_i$ and $||\mu_i|| \leq K|| \operatorname{marg}_{\Theta \times Z_0} \mu_i||$.

If Φ is compatible with type space $\mathcal{T}$, then the uniform boundedness requires that every type have a preference $\succsim_i \in P(\Theta \times T_{-i} \times A_{-i} \times Z_0)$ whose preferences conditional on her opponents' types and actions do not differ "too much" from her first-order preference (her preference over acts over the state of nature). Note that the bound K is imposed uniformly over all mechanisms. Φ^{PC} is unbounded. In a classical environment, Φ^{BC} is uniformly bounded by $K = 1$.

In addition to the compatibility and the uniform boundedness, we assume that no type in a type space of interest has completely indifferent first-order preferences.

Let d^* be a metric on the set T^* of hierarchies of preferences compatible with its product topology.

Proposition 6. *For every $\varepsilon > 0$ and $K < \infty$, there exists a mechanism $\mathcal{M}$ such that, for every pair of type spaces $\mathcal{T}$ and $\mathcal{T}'$ without completely indifferent first-order preferences, if restriction Φ is uniformly bounded by K in both $\mathcal{T}$ and $\mathcal{T}'$ and compatible with both $\mathcal{T}$ and $\mathcal{T}'$, then we have*

$$d^*(\hat{\pi}_i(t_i, \mathcal{T}), \hat{\pi}_i(t'_i, \mathcal{T}')) > \varepsilon \Rightarrow R_i^\Phi(t_i, \mathcal{T}, \mathcal{M}) \cap R_i^\Phi(t'_i, \mathcal{T}', \mathcal{M}) = \emptyset$$

for every $i \in I$, $t_i \in T_i$, and $t'_i \in T'_i$.

Note that, in Proposition 6, the construction of $\mathcal{M}$ depends on ε and K, but is independent of the details of Φ, $\mathcal{T}$, and $\mathcal{T}'$.

In the universal belief-type space (the space of Mertens-Zamir hierarchies), Dekel, Fudenberg, and Morris (2006, Lemma 4) construct a "discretized" direct mechanism in which only actions close to truth telling are interim belief-correlated rationalizable. Their result corresponds to Proposition 6 in classical environments under interim belief-correlated rationalizability, which is uniformly bounded by $K = 1$. Our proof uses a similar mechanism, but needs to take care of the following two issues that potentially destroy players' incentives for truth telling. (i) Outcomes are not necessarily private goods, so the social planner cannot necessarily give a reward to a player without affecting other players' incentives. Especially, a player's incentives to report her lower-order preferences are affected by how the social planner uses her reports to solicit other players' higher-order preferences. (ii) As a player sends less accurate reports about her lower-order preferences, other players become less willing to report their higher-order preferences accurately. (i) originates the issue, whereas (ii) "multiplies" it.[8] The uniform boundedness plays an important role to isolate these issues from the original truth-telling mechanism. The next two subsections are devoted for the proof of Proposition 6.

6.2.3 Single-Player Revelation Mechanism

As a preliminary step of the proof of Proposition 6, we first analyze a single-player mechanism that reveals her preferences. In this subsection, fix a compact metric space X of states with metric d. Let d_P be a metric compatible with the topology on $P(X)$. For each $\succsim \in P(X)$, we define the indicator function of $\succsim$, $\chi_{\succsim}$, that maps pairs of acts over X to 0, 1/2, or 1 according to $\succsim$ as follows:

$$\forall f, f' \in F(X), \quad \chi_{\succsim}(f, f') = \begin{cases} 1 & \text{if } f \succ f', \\ 1/2 & \text{if } f \sim f', \\ 0 & \text{if } f \prec f'. \end{cases}$$

Recall that $F_c(X)$ is the set of continuous acts over X. Since X is a compact metric space, by the Stone-Weierstrass Approximation theorem, there exists a countable dense subset $F = \{f_1, f_2, \dots\} \subset F_c(X)$ in the sup norm. Fix such an F.

We consider the following direct mechanism $\mathcal{M}^0 = (A^0, g^0)$ for a single player with action set $A^0 = P(X) \setminus \{\mathrm{ind}_X\}$ and outcome function $g^0 \colon X \times$

[8] Inaccurate reports may occur in Dekel, Fudenberg, and Morris (2006), but they come purely from discretization.

$A^0 \to \Delta(Z)$ given by

$$a \in A^0, \quad g^0(\cdot, a) = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} 2^{-k-l+1} \chi_a(f_k, f_l) f_k.$$

Under the mechanism $\mathcal{M}^0$, the player reports her preference. Then the social planner randomly draws a pair of acts from F and assigns the player with her preferred act according to her reported preference.[9]

In Lemma 5 below, we show that truth telling is a dominant strategy in $\mathcal{M}^0$ for every type. Indeed, by invoking the compactness of X, we show a "robust" version of strategy proofness: in every mechanism close to $\mathcal{M}^0$, every type strictly prefers reporting almost true preferences to reporting others according to almost true preferences.

For each $\delta > 0$ and measurable space W, let $G^{\delta,W}$ be the set of functions $g\colon X \times A^0 \times W \to \Delta(Z)$ with $|g(\cdot, w) - g^0| \leq \delta$ for every $w \in W$. We interpret W as the set of "noises" that may make the outcome slightly different from g^0. For each $\delta > 0$, let D^δ be the δ-neighborhood of the diagonal of $X \times X$, $\{(x, y) \in X \times X : d(x, y) \leq \delta\}$. For each $\delta > 0$, $K < \infty$, $\succsim \in A^0$, and measurable space W, let

$$P^{\delta,K,W}(\succsim) = \left\{ \succsim' \in P(X \times W) : \begin{array}{l} \exists \mu \in ca(X \times X \times W \times Z_0) \text{ s.t.} \\ (1)\ \operatorname{marg}_{1,4} \mu \text{ represents } \succsim, \\ (2)\ \operatorname{marg}_{2,3,4} \mu \text{ represents } \succsim', \\ (3)\ (X \times X \setminus D^\delta) \times W \times Z_0 \text{ is } \mu\text{-null}, \\ (4)\ ||\mu|| \leq K ||\operatorname{marg}_{1,4} \mu|| \end{array} \right\},$$

[9] Strictly speaking, $\mathcal{M}^0$ is not a mechanism according to Definition **??**, for its action set is infinite. The mechanism we will construct in the next subsection to prove Proposition 6, however, has finite action sets.

where $\operatorname{marg}_\Lambda \mu$ with $\Lambda \subset \{1,2,3,4\}$ denotes the marginal of μ with respect to the coordinates in Λ. $P^{\delta,K,W}(\succsim)$ is the set of preferences over "noise"-dependent acts that are "close" to $\succsim$.

Lemma 5. *For every $\varepsilon > 0$ and $K < \infty$, there exists $\delta > 0$ such that, if $\succsim, a, b \in A^0$ satisfy $d_P(\succsim, a) \leq \delta$ and $d_P(\succsim, b) > \varepsilon$, then, for every measurable space W, $g \in G^{\delta,W}$, and $\succsim'^{\delta,K,W}(\succsim)$, we have $g(\cdot, a) \succ' g(\cdot, b)$.*

Proof. See Appendix. □

6.2.4 Proof of Proposition 6

Let d_Θ be a metric compatible with the topology on Θ. For each $n \geq 1$, let $d_{P,n}$ be a metric compatible with the topology on the set of all n-th order preferences, $P(X_{n-1})$. Let

$$d_n((\theta, t_{-i,1}, \ldots, t_{-i,n}), (\theta', t'_{-i,1}, \ldots, t'_{-i,n})) = \max\left\{ d_\Theta(\theta, \theta'), \max_{k \leq n, j \in \mathcal{I} \setminus \{i\}} d_{P,k}(t_{j,k}, t'_{j,k}) \right\},$$

which is compatible with the product topology on $X_n = \Theta \times \prod_{k=0}^{n-1} P(X_k)^{I-1}$.

Fix any $\varepsilon > 0$ and $K < \infty$. Recall that d^* is a metric compatible with the product topology on $T^* \subset \prod_{n=0}^{\infty} P(X_n)$. By the definition of the product topology, there exist $\bar{\varepsilon} > 0$ and $N \in \mathbb{N}$ such that, for every $t = \{t_n\}_{n=1}^{\infty}, t' = \{t'_n\}_{n=1}^{\infty} \in T^*$, if $d^*(t, t') > \varepsilon$, then there exists $n \leq N$ such that $d_{P,n}(t_n, t'_n) > \bar{\varepsilon}$. Pick such $\bar{\varepsilon}$ and N.

For each $n \leq N$, substitute $X = X_{n-1}$, $d = d_{n-1}$, and $d_P = d_{P,n}$ in the previous subsection, and define $A^0_n = P(X_{n-1}) \setminus \{\operatorname{ind}_{X_{n-1}}\}$, $g^0_n \colon X_{n-1} \times A^0_n \to \Delta(Z)$, $G^{\delta,W}_n$, D^δ_n, and $P^{\delta,K,W}_n(\succsim)$ for $\delta > 0$, $\succsim \in A^0_n$, and measurable space W. By Lemma 5, there exist $0 < \varepsilon_0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_{N-1} \leq \varepsilon_N \leq \bar{\varepsilon}/2$ such that, for every $n \le N$, if $\succsim, a, b \in A_n^0$ satisfy $d_{P,n}(\succsim, a) \le \varepsilon_{n-1}$ and $d_{P,n}(\succsim, b) > \varepsilon_n$, then, for every measurable space W, $g \in G_n^{\varepsilon_{n-1},W}$, and $\succsim' \in P_n^{\varepsilon_{n-1},K,W}(\succsim)$, we have $g(\cdot, a) \succ' g(\cdot, b)$.

We define a mechanism $\mathcal{M}^* = ((A_i^*)_{i\in\mathcal{I}}, g^*)$ as follows. For each $i \in \mathcal{I}$ and $n \le N$, let $A_{i,n}^*$ be any ε_{n-1}-dense finite subset of A_n^0 with respect to $d_{P,n}$, $A_i^* = \prod_{n=1}^N A_{i,n}^*$, and $A^* = \prod_{i\in\mathcal{I}} A_i^*$. Define $g^*\colon \Theta \times A^* \to \Delta(Z)$ by

$$\forall \theta \in \Theta,\ a \in A^*, \quad g^*(\theta, a) = \frac{1-\delta}{I(1-\delta^N)} \sum_{i\in\mathcal{I}} \sum_{n=1}^N \delta^{n-1} g_n^0(\theta, a_{-i,1}, \dots, a_{-i,n-1}, a_{i,n}),$$

where $\delta > 0$ is small enough to satisfy $(1-\delta)/\delta \ge (I-1)(1-\varepsilon_0)/\varepsilon_0$.

Claim 1. *For every type space $\mathcal{T} = (T_i, \pi_i)_{i\in\mathcal{I}}$ without completely indifferent first-order preferences and $n \le N$, if restriction Φ is uniformly bounded by K in $\mathcal{T}$ and compatible with $\mathcal{T}$, then we have*

$$a_i \in R_{i,n}^\Phi(t_i, \mathcal{T}, \mathcal{M}^*) \Rightarrow d_{P,n}(\hat\pi_{i,n}(t_i, \mathcal{T}), a_{i,n}) \le \varepsilon_n$$

for every $i \in \mathcal{I}$ and $t_i \in T_i$.

Proof. The proof is by induction on n. Suppose that, for every $k \le n-1$, $a_i \in R_{i,n-1}^\Phi(t_i, \mathcal{T}, \mathcal{M}^*)$ implies $d_{P,k}(\hat\pi_{i,k}(t_i, \mathcal{T}), a_{i,k}) \le \varepsilon_k \le \varepsilon_{n-1}$. Suppose that there exists $a_i^* \in R_{i,n}^\Phi(t_i, \mathcal{T}, \mathcal{M}^*)$ such that $d_{P,n}(\hat\pi_{i,n}(t_i, \mathcal{T}), a_{i,n}^*) > \varepsilon_n$. Then there exist $\succsim_i \in P(\Theta \times T_{-i} \times A_{-i}^*)$ and $\mu_i \in ca(\Theta \times T_{-i} \times A_{-i}^* \times Z_0)$ such that $\Theta \times \mathrm{graph}(R_{-i,n-1}^\Phi(\cdot, \mathcal{T}, \mathcal{M}^*))$ is $\succsim_i$-certain, $g(\cdot, a_i^*, \cdot)\ (\mathrm{marg}_{\Theta\times A_{-i}^*} \succsim_i)\ g(\cdot, a_i, \cdot)$ for every $a_i \in A_i^*$, $\mathrm{marg}_{\Theta\times T_{-i}} \succsim_i = \pi_i(t_i)$, μ_i represents $\succsim_i$, and $||\mu_i|| \le K||\mathrm{marg}_{\Theta\times Z_0} \mu_i||$.

Let $W = \prod_{k=n}^N A_{-i,k}^*$ and $\varphi_{-i}\colon \Theta \times T_{-i} \times A_{-i}^* \to X_{n-1} \times X_{n-1} \times W$ such that $\varphi_{-i}(\theta, t_{-i}, a_{-i}) = (\theta, \hat\pi_{-i,1}(t_{-i}, \mathcal{T}), \dots, \hat\pi_{-i,n-1}(t_{-i}, \mathcal{T}), \theta, a_{-i})$. Collect all the terms in g^* that depend on $a_{i,n}$ and define $g^*_{i,n}\colon X_{n-1} \times A^*_{i,n} \times W$ by

$$\begin{aligned}
&g^*_{i,n}(\theta, a_{-i,1}, \ldots, a_{-i,n-1}, a_{i,n}, a_{-i,n}, \ldots, a_{-i,N}) \\
&= C\left(g^0_n(\theta, a_{-i,1}, \ldots, a_{-i,n-1}, a_{i,n}) + \sum_{j\in\mathcal{I}\setminus\{i\}} \sum_{k=n+1}^{N} \delta^{k-n} g^0_k(\theta, a_{-j,1}, \ldots, a_{-j,k-1}, a_{j,k})\right),
\end{aligned}$$

where $a_{i,-n} = a^*_{i,-n}$ and C is a positive normalization constant. We have $g^*_{i,n} \in G^{\varepsilon_0,W}_n \subseteq G^{\varepsilon_{n-1},W}_n$. Let $\succsim'_i = (\varphi_{-i})^P(\succsim_i) \in P(X_{n-1} \times X_{n-1} \times W)$ and $\mu'_i = (\varphi_{-i} \times \mathrm{id}_{Z_0})^{ca}(\mu_i) \in ca(X_{n-1} \times X_{n-1} \times W \times Z_0)$. Note that we have $\operatorname{marg}_{1,4} \mu'_i$ represents $\hat{\pi}_{i,n}(t_i, \mathcal{T})$, $\operatorname{marg}_{2,3,4} \mu'_i$ represents $\operatorname{marg}_{\Theta\times A^*_{-i}} \succsim_i$, and $||\mu'_i|| \leq K|| \operatorname{marg}_{\Theta\times Z_0} \mu_i|| = K|| \operatorname{marg}_{1,4} \mu'_i||$. Note also that, since $\varphi_{-i}(\Theta \times \operatorname{graph}(R^\Phi_{-i,n-1}(\cdot, \mathcal{T}, \mathcal{M}^*))) \subseteq D^{\varepsilon_{n-1}}_n \times W$ is $\succsim'_i$-certain, $(X_{n-1} \times X_{n-1} \setminus D^{\varepsilon_{n-1}}_n) \times W \times Z_0$ is μ'_i-null. Thus, we have $\operatorname{marg}_{\Theta\times A^*_{-i}} \succsim_i \in P^{\varepsilon_{n-1},K,W}(\hat{\pi}_{i,n}(t_i, \mathcal{T}))$. Since $A^*_{i,n}$ is ε_{n-1}-dense in A^0_n, there exists $a'_{i,n} \in A^*_{i,n}$ such that $d_{P,n}(\hat{\pi}_{i,n}(t_i, \mathcal{T}), a'_{i,n}) \leq \varepsilon_{n-1}$. By Lemma 5, $\operatorname{marg}_{\Theta\times A^*_{-i}} \succsim_i$ strictly prefers $g^*_{i,n}(\cdot, a'_{i,n}, \cdot)$ to $g^*_{i,n}(\cdot, a^*_{i,n}, \cdot)$, thus, by the independence of $\operatorname{marg}_{\Theta\times A^*_{-i}} \succsim_i$, $\operatorname{marg}_{\Theta\times A^*_{-i}} \succsim_i$ strictly prefers $g^*(\cdot, a'_{i,n}, a^*_{i,-n}, \cdot)$ to $g^*(\cdot, a^*_i, \cdot)$. This is a contradiction. $\square$

Now we prove Proposition 6. Pick any pair of type spaces $\mathcal{T}$ and $\mathcal{T}'$ without completely indifferent first-order preferences, restriction Φ uniformly bounded by K in both $\mathcal{T}$ and $\mathcal{T}'$ and compatible with both $\mathcal{T}$ and $\mathcal{T}'$, $i \in \mathcal{I}$, $t_i \in T_i$, and $t'_i \in T'_i$. Suppose that there exists $a_i = (a_{i,1}, \ldots, a_{i,N}) \in R^\Phi_i(t_i, \mathcal{T}, \mathcal{M}^*) \cap R^\Phi_i(t'_i, \mathcal{T}', \mathcal{M}^*)$. For every $n \leq N$, since $a_i \in R^\Phi_{i,n}(t_i, \mathcal{T}, \mathcal{M}^*) \cap R^\Phi_{i,n}(t'_i, \mathcal{T}', \mathcal{M}^*)$, we have

$$\begin{aligned}
&d_{P,n}(\hat{\pi}_{i,n}(t_i, \mathcal{T}), \hat{\pi}_{i,n}(t'_i, \mathcal{T}')) \\
&\leq d_{P,n}(\hat{\pi}_{i,n}(t_i, \mathcal{T}), a_{i,n}) + d_{P,n}(\hat{\pi}_{i,n}(t'_i, \mathcal{T}'), a_{i,n}) \leq 2\varepsilon_n \leq \bar{\varepsilon}
\end{aligned}$$

by Claim 1. Thus $d^*(\hat{\pi}_i(t_i, \mathcal{T}), \hat{\pi}_i(t'_i, \mathcal{T}')) \leq \varepsilon$.

6.2.5 Impossibility of Strategic Distinguishability

In Proposition 6, we assumed that Φ is uniformly bounded. In this subsection, we show that, if we use interim preference-correlated rationalizability and impose no restriction on players' preferences other than compatibility with underlying type spaces, then the social planner cannot infer any interdependency of players' preferences from their behavior.

Fix a type space $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$ and a mechanism $\mathcal{M} = ((A_i)_{i \in \mathcal{I}}, g)$. A mixed action $\alpha_i \in \Delta(A_i)$ *strongly dominates* $a_i \in A_i$ *for type* $t_i \in T_i$ if the first-order preference of type t_i, $\mathrm{marg}_\Theta\, \pi_i(t_i) = \hat{\pi}_i(t_i)$, strictly prefers $g(\cdot, \alpha_i, a_{-i})$ to $g(\cdot, a_i, a_{-i})$, and $g(\cdot, \alpha_i, a_{-i}) - g(\cdot, a_i, a_{-i})$ is independent of a_{-i}. a_i is *strongly undominated for type* t_i if no mixed action strongly dominates a_i for type t_i.

The next proposition extends Morris and Takahashi (2009, Propositions 2 and 3) to games with incomplete information, and shows that interim preference-correlated rationalizability is characterized by one-step elimination of strongly dominated actions.

Proposition 7. *Suppose that Θ is finite. Then a_i is interim preference-correlated rationalizable for type t_i if and only if a_i is strongly undominated for t_i.*

Proof. See Appendix. □

Since one-step elimination of strongly dominated actions depends only on the first-order preference, Proposition 7 implies that two types are distinguishable under interim preference-correlated rationalizability only if they have different first-order preferences.

Proposition 8. *Suppose that Θ is finite. For every pair of type spaces $\mathcal{T}$ and $\mathcal{T}'$ and mechanism $\mathcal{M}$, we have*

$$\hat{\pi}_{i,1}(t_i, \mathcal{T}) = \hat{\pi}_{i,1}(t_i', \mathcal{T}') \Rightarrow R_i^{\Phi^{\mathrm{PC}}}(t_i, \mathcal{T}, \mathcal{M}) = R_i^{\Phi^{\mathrm{PC}}}(t_i', \mathcal{T}', \mathcal{M})$$

for every $i \in \mathcal{I}$, $t_i \in T_i$, and $t_i' \in T_i'$.

A Appendix

A.1 Proof of Lemma 5

Suppose not. Then, there exist $\varepsilon > 0$ and $K < \infty$ such that, for every $n \in \mathbb{N}$, there exist $\succsim_n, a_n, b_n \in A^0$, measurable space W_n, $g_n \in G^{1/n, W_n}$, $\succsim_n' \in P^{1/n,K,W_n}(\succsim_n)$ such that $d_P(\succsim_n, a_n) \leq 1/n$, $d_P(\succsim_n, b_n) \geq \varepsilon$, and $g_n(\cdot, a_n) \precsim_n' g_n(\cdot, b_n)$. For each n, there exist signed measures $\mu_n \in ca(X \times Z_0)$ and $\nu_n \in ca(X \times X \times W_n \times Z_0)$ such that $\operatorname{marg}_{1,4} \nu_n = \mu_n$ represents $\succsim_n$, $\operatorname{marg}_{2,3,4} \nu_n$ represents $\succsim_n'$, $(X \times X \setminus D^{1/n}) \times W \times Z_0$ is ν_n-null, $||\mu_n|| = 1$, and $||\nu_n|| \leq K$. Since A^0 and $\{\mu \in ca(X \times Z_0) : ||\mu|| = 1\}$ are compact and metrizable by Lemma 1, by taking subsequences if necessary, we can find $\succsim^*, b^* \in A^0$ and $\mu^* \in ca(X \times Z_0)$ with $||\mu^*|| = 1$ such that $\succsim_n \to \succsim^*$, $b_n \to b^*$, and $\mu_n \to \mu^*$ as $n \to \infty$. Note that $a_n \to \succsim^*$ as $n \to \infty$, $\succsim^* \neq b^*$, and μ^* represents $\succsim^*$.

Claim 2. *There exist $k^*, l^* \in \mathbb{N}$ such that $\succsim^*$ strictly prefers f_{k^*} to f_{l^*} while b^* strictly prefers f_{l^*} to f_{k^*}.*

Proof of Claim 2. Since $\succsim^* \neq b^*$, there exist $f, f' \in F_c(X)$ such that $\succsim^*$ and b^* have different preferences between f and f'. Since $\succsim^*$ and b^* satisfy the continuity and neither of them is complete indifference, we can assume without loss of generality that $\succsim^*$ strictly prefers f to f' and b^* strictly prefers f' to f. To see this, suppose, for example, that $\succsim^*$ is indifferent between f and f' while b^* strictly prefers f' to f. Then, replace f by $(1-\lambda)f + \lambda f''$ and f' by $(1-\lambda)f' + \lambda f'''$ such that $\succsim^*$ strictly prefers f'' to f''' and $\lambda > 0$ is sufficiently small. A similar trick works when $\succsim^*$ strictly prefers f to f' while b^* is indifferent between f to f'. Since F is dense in $F_c(X)$ in the sup norm, by the continuity of $\succsim^*$ and b^*, we can assume $f, f' \in F$ without loss of generality. □

Claim 3. *There exists $n_0 \in \mathbb{N}$ such that, for every $n \geq n_0$, b_n strictly prefers f_{l^*} to f_{k^*}.*

Proof of Claim 3. Follows from $b_n \to b^*$ as $n \to \infty$. □

It follows from Claim 2 that there exists $\eta > 0$ such that

$$(5K+2)\eta < 2^{-k^*-l^*+1} \int (f_{k^*} - f_{l^*}) d\mu^*.$$

Pick $k_0 \geq \max\{k^*, l^*\}$ such that

$$\sum_{\max\{k,l\} > k_0} 2^{-k-l+1} < \eta.$$

Claim 4. *There exists $n_1 \in \mathbb{N}$ such that, for every $n \geq n_1$ and $\max\{k, l\} \leq k_0$, if $\succsim^*$ strictly prefers f_k to f_l, then a_n also strictly prefers f_k to f_l.*

Proof of Claim 4. Follows from $a_n \to \succsim^*$ as $n \to \infty$. □

Note that

$$(\chi_{a_n}(f_k, f_l) - \chi_{b_n}(f_k, f_l)) \int f_k d\mu^* + (\chi_{a_n}(f_l, f_k) - \chi_{b_n}(f_l, f_k)) \int f_l d\mu^*$$
$$= (\chi_{a_n}(f_k, f_l) - \chi_{b_n}(f_k, f_l)) \int (f_k - f_l) d\mu^*$$

since $\chi_{a_n}(f_l, f_k) = 1 - \chi_{a_n}(f_k, f_l)$ and $\chi_{b_n}(f_k, f_l) = 1 - \chi_{b_n}(f_l, f_k)$.

Claim 5. *For every $n \geq \max\{n_0, n_1\}$, we have*

$$(\chi_{a_n}(f_k, f_l) - \chi_{b_n}(f_k, f_l)) \int (f_k - f_l) d\mu^* \begin{cases} = \int (f_{k^*} - f_{l^*}) d\mu^* & \textit{if } (k, l) = (k^*, l^*), \\ \geq 0 & \textit{if } \max\{k, l\} \leq k_0. \end{cases}$$

Proof of Claim 5. By Claims 3 and 4, $\chi_{a_n}(f_{k^*}, f_{l^*}) = 1$ and $\chi_{b_n}(f_{k^*}, f_{l^*}) = 0$; $\chi_{a_n}(f_k, f_l) = 1 \geq \chi_{b_n}(f_k, f_l)$ and $\int (f_k - f_l) d\mu^* > 0$ if $\succsim^*$ strictly prefers f_k to f_l; $\chi_{a_n}(f_k, f_l) = 0 \leq \chi_{b_n}(f_k, f_l)$ and $\int (f_k - f_l) d\mu^* < 0$ if $\succsim^*$ strictly prefers f_l to f_k; $\int (f_k - f_l) d\mu^* = 0$ if $\succsim^*$ is indifferent between f_k and f_l. □

Claim 6. *There exists $n_2 \in \mathbb{N}$ such that, for every $n \geq n_2$ and $k \leq k_0$, we have*

$$\left| \int f_k d \operatorname{marg}_{2,4} \nu_n - \int f_k d\mu_n \right| < K\eta.$$

Proof of Claim 6. Since X is a compact metric space, every continuous function is uniformly continuous. Therefore, there exists $n_2 \in \mathbb{N}$ such that $|f_k(x) - f_k(y)| < \eta$ for every $k \leq k_0$ and $(x, y) \in D^{1/n_2}$. For every $n \geq n_2$, we have

$$
\begin{aligned}
&\left|\int f_k d\operatorname{marg}_{2,4}\nu_n - \int f_k d\mu_n\right| \\
&= \left|\int f_k(y)(z) d\operatorname{marg}_{1,2,4}\nu_n(x,y,z) - \int f_k(x)(z) d\operatorname{marg}_{1,2,4}\nu_n(x,y,z)\right| \\
&= \left|\int (f_k(x)(z) - f_k(y)(z)) d\operatorname{marg}_{1,2,4}\nu_n(x,y,z)\right| \\
&\leq \int |f_k(x)(z) - f_k(y)(z)| d\operatorname{marg}_{1,2,4}\nu_n(x,y,z) < K\eta
\end{aligned}
$$

since $|f_k(x)(z) - f_k(y)(z)| < \eta$ for every $(x,y,z) \in D^{1/n} \times Z_0 \subseteq D^{1/n_2} \times Z_0$, $(X \times X \setminus D^{1/n}) \times Z_0$ is $\operatorname{marg}_{1,2,4}\nu_n$-null, and $||\operatorname{marg}_{1,2,4}\nu_n|| \leq ||\nu_n|| \leq K$. $\square$

Since $\mu_n \to \mu^*$ as $n \to \infty$, there exists $n \geq \max\{n_0, n_1, n_2, 1/\eta\}$ such that, for every $k \leq k_0$, $|\int f_k d\mu_n - \int f_k d\mu^*| < \eta$. We decompose $\int (g_n(\cdot, a_n) - g_n(\cdot, b_n)) d\operatorname{marg}_{2,3,4}\nu_n$ into the following four terms:

$$
\begin{aligned}
&\int (g_n(\cdot, a_n) - g_n(\cdot, b_n)) d\operatorname{marg}_{2,3,4}\nu_n \\
&= \sum_{\max\{k,l\} \leq k_0} 2^{-k-l+1}(\chi_{a_n}(f_k, f_l) - \chi_{b_n}(f_k, f_l)) \int f_k d\mu^* \\
&\quad + \sum_{\max\{k,l\} \leq k_0} 2^{-k-l+1}(\chi_{a_n}(f_k, f_l) - \chi_{b_n}(f_k, f_l)) \left(\int f_k d\operatorname{marg}_{2,4} d\nu_n - \int f_k d\mu^*\right) \\
&\quad + \sum_{\max\{k,l\} > k_0} 2^{-k-l+1}(\chi_{a_n}(f_k, f_l) - \chi_{b_n}(f_k, f_l)) \int f_k d\operatorname{marg}_{2,4}\nu_n \\
&\quad + \int [(g_n(\cdot, a_n) - g^0(\cdot, a_n)) - (g_n(\cdot, b_n) - g^0(\cdot, b_n))] d\operatorname{marg}_{2,3,4}\nu_n.
\end{aligned}
$$

The first term is larger than $(5K+2)\eta$ by Claim 5. The other terms are at least -2η, $-2(K+1)\eta$, and $-2K\eta$, respectively, since $\sum_{\max\{k,l\} \leq k_0} 2^{-k-l+1} <$

2, $|\chi_{a_n} - \chi_{b_n}| \leq 1$,

$$
\begin{aligned}
&\left|\int f_k d \operatorname{marg}_{2,4} d\nu_n - \int f_k d\mu^*\right| \\
&\leq \left|\int f_k d \operatorname{marg}_{2,4} d\nu_n - \int f_k d\mu_n\right| + \left|\int f_k d\mu_n - \int f_k d\mu^*\right| \\
&< (K+1)\eta
\end{aligned}
$$

by Claim 6, $\sum_{\max\{k,l\}>k_0} 2^{-k-l+1} < \eta$, $|f_k| \leq 1$, $|g_n - g^0| \leq 1/n \leq \eta$, and $||\operatorname{marg}_{2,4}\nu_n|| \leq ||\operatorname{marg}_{2,3,4}\nu_n|| \leq ||\nu_n|| \leq K$. Thus $\succsim_n'$ strictly prefers $g_n(\cdot, a_n)$ to $g_n(\cdot, b_n)$, which is a contradiction.

A.2 Proof of Proposition 7

Lemma 6. *Suppose that Θ is finite. Then $a_i \in R_{i,1}^{\Phi^{\mathrm{PC}}}(t_i)$ if and only if a_i is strongly undominated for t_i.*

Proof. The "only if" part is obvious. To show the "if" part, suppose that $a_i \notin R_{i,1}(t_i)$. Let $\mu_i \in ca(\Theta \times Z)$ represent $\operatorname{marg}_\Theta \pi_i(t_i)$. Then there is no $\nu_i \in ca(\Theta \times A_{-i} \times Z)$ that satisfies the following system of linear (in)equalities:

$$
\begin{gathered}
\operatorname{marg}_{\Theta\times Z} \nu_i = \mu_i, \\
\forall a_i' \in A_i, \quad \int_{\Theta\times A_{-i}\times Z} (g(\cdot, a_i, \cdot) - g(\cdot, a_i', \cdot)) d\nu_i \geq 0.
\end{gathered}
$$

Since this system is essentially finite dimensional, by Farkas' lemma, there exist $p\colon \Theta \times Z \to \mathbb{R}$ and $q\colon A_i \to \mathbb{R}$ such that

$$
\forall \theta \in \Theta,\ a_{-i}' \in A_{-i},\ z \in Z, \quad p(\theta, z) + \sum_{a_i' \in A_i} q(a_i')(g(\theta, a_i, a_{-i})(z) - g(\theta, a_i', a_{-i})(z)) = 0,
$$

$$
\begin{gathered}
\forall a_i' \in A_i, \quad q(a_i') \leq 0, \\
\int_{\Theta\times Z} p d\mu_i < 0.
\end{gathered}
$$

Since every $q(a_i')$ is non-positive, and at least one is strictly negative, let $r = 1/\sum_{a_i'} q(a_i') < 0$ and $\alpha_i = rq \in \Delta(A_i)$. Then we have $g(\theta, \alpha_i, a_{-i}) - g(\theta, a_i, a_{-i}) = -rp(\theta, z)$ is independent of a_{-i} and

$$\int_{\Theta \times Z} (g(\cdot, \alpha_i, a_{-i}) - g(\cdot, a_i, a_{-i})) d\mu_i = -r \int_{\Theta \times Z} p d\mu_i > 0$$

for every $a_{-i} \in A_{-i}$. Thus α_i strongly dominates a_i. $\square$

Lemma 7. *Suppose that Θ is finite. Then $R_1 = (R_{i,1})_{i \in \mathcal{I}}$ is a preference-correlated best reply correspondence.*

Proof. Fix any player $i \in \mathcal{I}$. For each $j \in \mathcal{I} \setminus \{i\}$, $t_j \in T_j$, and $a_j \in A_j$, if $a_j \in R_{j,1}(t_j)$, then let $\alpha_j(t_j, a_j)$ be the point mass on a_j. If $a_j \notin R_{j,1}(t_j)$, then, by Lemma 6, let $\alpha_j(t_j, a_j)$ be a mixed action that strongly dominates a_j for type t_j. Without loss of generality, we can assume that $\alpha_j(t_j, a_j) \in \Delta(R_{j,1}(t_j))$. By the Kuratowski-Ryll-Nardzewski selection theorem (Aliprantis and Border, 1999, Theorem 17.13), we can assume that α_j is measurable. For each $t_{-i} \in T_{-i}$ and $a_{-i} \in A_{-i}$, define $\alpha_{-i}(t_{-i}, a_{-i}) \in \Delta(R_{-i,1}(t_{-i}))$ by $\alpha_{-i}(t_{-i}, a_{-i})(b_{-i}) = \prod_{j \neq i} \alpha_j(t_j, a_j)(b_j)$ for each $b_{-i} \in R_{-i,1}(t_{-i})$.

Pick any $t_i \in T_i$. Let $\mu_i \in ca(\Theta \times T_{-i} \times Z)$ be a signed measure that represents $\pi_i(t_i)$. Pick any $b_i \in R_{i,1}(t_i)$, which is a best reply under $\nu_i \in ca(\Theta \times T_{-i} \times A_{-i} \times Z)$ with $\mathrm{marg}_{\Theta \times T_{-i} \times Z}\, \nu_i = \mu_i$. Define $\nu_i' \in ca(\Theta \times T_{-i} \times A_{-i} \times Z)$ by

$$\nu_i'(E \times \{b_{-i}\}) = \int_{E \times A_{-i}} \alpha_{-i}(\cdot)(b_{-i}) d\nu_i$$

for every measurable $E \subseteq \Theta \times T_{-i} \times Z$ and $b_{-i} \in R_{-i,1}(t_{-i})$. We show that b_i is a preference-correlated best reply for type t_i. First,

$$\mathrm{marg}_{\Theta \times T_{-i} \times Z}\, \nu_i' = \mathrm{marg}_{\Theta \times T_{-i} \times Z}\, \nu_i = \mu_i.$$

Second, for every $a_i \in A_i$,

$$\begin{aligned}
&\int_{\Theta \times A_{-i} \times Z} g(\cdot, a_i, \cdot) d \operatorname{marg}_{\Theta \times A_{-i} \times Z} \nu_i' \\
&= \int_{\Theta \times T_{-i} \times A_{-i} \times Z} g(\cdot, a_i, \cdot) d\nu_i' \\
&= \int_{\Theta \times T_{-i} \times A_{-i} \times Z} g(\cdot, a_i, \alpha_{-i}(\cdot)) d\nu_i \\
&= \int_{\Theta \times T_{-i} \times A_{-i} \times Z} (g(\cdot, a_i, \cdot) + p) d\nu_i \\
&= \int_{\Theta \times A_{-i} \times Z} g(\cdot, a_i, \cdot) d \operatorname{marg}_{\Theta \times A_{-i} \times Z} \nu_i + \int_{\Theta \times T_{-i} \times A_{-i} \times Z} p d\nu_i,
\end{aligned}$$

where $p(\theta, t_{-i}, a_{-i}) = g(\theta, a_i, \alpha_{-i}(t_{-i}, a_{-i})) - g(\theta, a_i, a_{-i})$ is independent of a_i by the definition of strong dominance. Since b_i maximizes g with respect to $\operatorname{marg}_{\Theta \times A_{-i} \times Z} \nu_i$, b_i maximizes g with respect to $\operatorname{marg}_{\Theta \times A_{-i} \times Z} \nu_i'$. $\square$

Proposition 7 follows from Lemmas 6 and 7.

References

[1] D. Abreu and H. Matsushima (1992), "Virtual Implementation in Iteratively Undominated Actions," at http://www.princeton.edu/~dabreu/index_files/virtual%20implementation-incomplete.pdf

[2] C. D. Aliprantis and K. C. Border (1999), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin.

[3] F. J. Anscombe and R. J. Aumann (1963), "A Definition of Subjective Probability," *Annals of Mathematical Statistics* 34, 199–205.

[4] P. Battigalli and M. Siniscalchi (2003) "Rationalization and Incomplete Information," *Advances in Theoretical Economics* 3 (1), Article 3.

[5] D. Bergemann and S. Morris (2009). "Virtual Robust Implementation," *Theoretical Economics* 4, 45–88.

[6] T. Börgers (1993), "Pure Strategy Dominance," *Econometrica* 61 (2), 423–430.

[7] A. Brandenburger and E. Dekel (1993), "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory* 59, 189–198.

[8] E. Dekel, D. Fudenberg, and S. Morris (2006), "Topologies on Types," *Theoretical Economics* 1, 275–309.

[9] E. Dekel, D. Fudenberg, and S. Morris (2007), "Interim Correlated Rationalizability," *Theoretical Economics* 2, 15–40.

[10] A. Di Tillio (2008), "Subjective Expected Utility in Games," *Theoretical Economics* 3, 287–323.

[11] J. C. Ely and M. Pęski (2006), "Hierarchies of Belief and Interim Rationalizability," *Theoretical Economics* 1, 19–65.

[12] L. G. Epstein (1997), "Preference, Rationalizability and Equilibrium," *Journal of Economic Theory* 73, 1–29.

[13] L. G. Epstein and T. Wang (1996), ""Beliefs about Beliefs" without Probabilities," *Econometrica* 64, 1343–1373.

[14] P. A. Haile, A. Hortaçsu, and G. Kosenok (2008), "On the Empirical Content of Quantal Response Equilibrium," *American Economic Review* 98 (1), 180–200.

[15] B. Hill (2008), "An Additively Separable Representation in the Savage Framework," mimeo.

[16] F. Gul and W. Pesendorfer (2007), "The Canonical Space for Behavioral Types," mimeo.

[17] A. Kechris (1995), *Classical Descriptive Set Theory*, Springer, Berlin.

[18] J. O. Ledyard (1986), "The Scope of the Hypothesis of Bayesian Equilibrium," *Journal of Economic Theory* 39 (1), 59–82.

[19] J.-F. Mertens and S. Zamir (1985), "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory* 14 (1), 1–29.

[20] S. Morris and S. Takahashi (2009), "Preference-Correlated Rationalizability," mimeo.

[21] T. Sadzik (2008), "Beliefs Revealed in Bayesian-Nash Equilibrium," mimeo.

[22] A. J. Van Haagen (1981), "Finite Signed Measures on Function Spaces," *Pacific Journal of Mathematics* 95 (2), 467–482.