Essays in Epistemic Game Theory

Akira Yokotani

Submitted in Partial Fulfillment
of the
Requirements for the Degree
Doctor of Philosophy

Supervised by
Professor Paulo Barelli

Department of Economics
Arts, Science and Engineering
School of Arts and Sciences

University of Rochester
Rochester, New York

2011
To Lori and Lorraine.
Curriculum Vitae

The author was born in Kyoto, Japan on 15 May, 1977. He entered the University of Tokyo in 1996, and was awarded a Bachelor of Economics degree in 2000 and a Master of Economics degree in 2002. After a one year exchange program at Yale University, he came to the University of Rochester in the Fall of 2006 and began graduate studies in Economics. He received a Fox International Fellowship in 2004 from Yale University, and from the University of Rochester, the Yasuo Uekawa Memorial Fellowship in 2006, the Akira Takayama Memorial Fellowship in 2007, and a Graduate Fellowship from 2007 to 2011. He pursued his research in Game Theory under the direction of Professor Paulo Barelli and received the Master of Arts degree from the University of Rochester in 2009.
Abstract

In this thesis, we make three achievements in the field of epistemic game theory: First, we show the existence of a universal type space where redundant types can be embedded. Type spaces with redundant types cannot be represented in the standard universal type space (Mertens and Zamir [39]). In Chapter 1, we extend the universal type space of Mertens-Zamir by introducing a payoff irrelevant parameter space $C$ as a missing source of uncertainty on the lines of Liu [33] so that redundant types also can be represented there. In contrast to Liu, we show that the parameter space $C$ can always be an exogenous space, and moreover $C = \{0, 1\}$ is always enough. In Chapter 2, we apply this idea of extended universal type spaces in order to generalize the existing results in robust implementation (Bergemann and Morris [7]). Adopting knowledge-belief spaces (Aumann [5] and [6]), we show that robust implementation is equivalent to Bayesian implementation on one particular belief structure. This result allows us to directly apply the results about Bayesian implementation, such as Jackson [29], to obtain a characterization result of robust implementation in a more general class of environments. In Chapter 3, we show the impossibility of robust implementation. We apply the idea of Saijo [49], which is about the impossibility of Nash implementation, and show that only constant social choice functions are robustly implementable in large domains of interdependent preferences.
## Contents

1. **The Belief Hierarchy Representation of Harsanyi Type Spaces with Redundancy**  
   1.1 Introduction ......................................................... 1  
   1.2 Preliminaries ....................................................... 9  
      1.2.1 Mathematical notations ........................................ 9  
      1.2.2 Harsanyi type space ......................................... 10  
   1.3 Universal type spaces ............................................. 11  
   1.4 An extended sequential belief space ............................. 12  
      1.4.1 Redundant types ............................................. 13  
      1.4.2 Extension with a payoff irrelevant parameter space ........ 13  
      1.4.3 S-isomorphic embedding onto $U(S \times C)$ ................ 16  
      1.4.4 Extension to $N > 2$ ...................................... 23  
   1.5 The characterization of the sequential belief in $U(S \times C)$ ........ 24  
      1.5.1 Symmetric types ............................................. 24  
      1.5.2 Characterization of symmetric types ......................... 27  
      1.5.3 Semantic interpretation of syntactic characterization ....... 32  
   1.6 Application to intrinsic correlation ............................. 32  
      1.6.1 Bayesian representation of correlated equilibrium ........... 32
1.6.2 Conditional independence and rationality and common certainty of rationality ........................................ 34
1.6.3 Conditional independence and redundancy ........................................... 35
1.7 Conclusion ........................................ 38
1.8 Appendix ........................................ 39
1.8.1 Bimeasurability of the function $\phi$ ........................................ 39

2 Knowledge-Belief Space Approach to Robust Implementation 46
2.1 Introduction ........................................ 46
2.2 Robust implementation ........................................ 51
  2.2.1 Bayesian implementation ........................................ 51
  2.2.2 Robust implementation ........................................ 53
2.3 The universal type space ........................................ 54
2.4 Extension property ........................................ 56
2.5 The universal knowledge-belief space ........................................ 59
  2.5.1 Knowledge-belief space ........................................ 59
  2.5.2 The universal knowledge-belief space ........................................ 61
  2.5.3 The universal knowledge-belief space with a payoff irrelevant parameter space ........................................ 71
2.6 Characterization of Implementation ........................................ 74
2.7 Conclusion ........................................ 82

3 Impossibility of Robust Implementation on Full Domain of Preferences 83
3.1 Introduction ........................................ 83
3.2 Model ........................................ 85
3.3 Necessary and sufficient condition for impossibility ........................................ 86
3.4 Comments on the impossibility in smaller domains ........................................ 90
3.4.1 Two person example ................................................. 91
3.5 Conclusion ................................................................. 93

Bibliography .................................................................. 94
Chapter 1

The Belief Hierarchy

Representation of Harsanyi Type

Spaces with Redundancy

1.1 Introduction

One difficulty in dealing with games of incomplete information is the infinite regress of uncertainty. Typically, an agent is uncertain about the payoff functions of the other agents.\footnote{The agents’ uncertainty about action spaces can be represented as the uncertainty about payoff functions. See Hu-Stuart \[28\] for the details} In order to analyze an agent’s decision under incomplete information, it is not enough to incorporate his belief over the basic uncertainty, i.e., the uncertainty about the agents’ payoffs. We have to incorporate what the agent believes about what his opponents believe about the basic uncertainty too. And next we have to consider the agent’s belief about what his opponents believe about what he believes about the basic uncertainty, and so on \textit{ad infinitum}. Therefore, in order to deal with games of incomplete information, we have to model this infinite regress of beliefs about beliefs. We call this hierarchy of beliefs a \textit{sequential belief}.\footnote{The agents’ uncertainty about action spaces can be represented as the uncertainty about payoff functions. See Hu-Stuart \[28\] for the details}
Since Harsanyi [25], we have been dealing with this difficulty by using the notion of \textit{type} and the associated Bayesian game. We postulate that all the informational attributes of agents, including sequential beliefs, can be reduced to one variable called the agent’s “type”. This postulation allows us to apply equilibrium concepts of games of complete information to games of incomplete information. In this paper, we say that the types defined by Harsanyi are \textit{Harsanyi types} in order to distinguish them from \textit{epistemic types} which we will define later.

Concerning individual informational attributes, we can conceive the information brought by private signals, predetermined personal conjectures (ex. personal characters, or habits in thinking), and so on. We can easily model these attributes with parameters. However, it is not clear that Harsanyi types correctly reflect the agents’ sequential beliefs. This suspicion is cleared by Mertens and Zamir [39] and Brandenburger and Dekel [11]. They showed that, under reasonable conditions, the space of the sequential beliefs over the basic uncertainty forms a Harsanyi type space, and we can embed arbitrary Harsanyi type spaces into the space of sequential beliefs. We say that this space of sequential beliefs is the \textit{universal type space} and sequential beliefs are \textit{epistemic types}.

Still we have another difficulty about the sequential beliefs and Harsanyi types. Indeed Mertens-Zamir and Brandenburger-Dekel verified that we can represent sequential beliefs as Harsanyi types, but only when there are no \textit{redundant types}, which are types that are associated with the same sequential belief. But redundant types ought to be considered, as the following example shows.

\textbf{Example 1 (Ely and Peski (2006))}: Consider the following two Harsanyi type spaces.

\textit{Type space A}: The payoff parameter space is $S = \{-1, 1\}$, the set of agents is $N = \{1, 2\}$, the
The set of types is $T_i = \{-1, 1\}$ for $i = 1, 2$, and the belief structure is characterized by a common prior $\mu \in \Delta(S \times T)$ such that

$$
\mu(s, t_i, t_{-i}) = \begin{cases} 
\frac{1}{4} & \text{if } s = t_i \cdot t_{-i} \\
0 & \text{otherwise}
\end{cases}
$$

Let $h^k_i(t_i)$ be the $k$th order belief of the agent $i$ associated with his type $t_i$. We can derive the sequential beliefs over $S$ in the above structure as follows:

$$
h^1_i(-1)[s] = \begin{cases} 
\frac{1}{2} & \text{if } s = -1 \\
\frac{1}{2} & \text{if } s = 1
\end{cases}
$$

$$
h^2_i(-1)[s] = \begin{cases} 
\frac{1}{2} h^1_j(-1)[-1] + \frac{1}{2} h^1_j(1)[-1] = \frac{1}{2} & \text{if } s = -1 \\
\frac{1}{2} h^1_j(-1)[1] + \frac{1}{2} h^1_j(1)[1] = \frac{1}{2} & \text{if } s = 1
\end{cases}
$$

$$
h^3_i(-1)[s] = \begin{cases} 
\frac{1}{2} h^2_j(-1)[-1] + \frac{1}{2} h^2_j(1)[-1] = \frac{1}{2} & \text{if } s = -1 \\
\frac{1}{2} h^2_j(-1)[1] + \frac{1}{2} h^2_j(1)[1] = \frac{1}{2} & \text{if } s = 1
\end{cases}
$$

$$
:$$

The resulting sequential belief of $t_i = -1$ is $\frac{1}{2}$ for all orders. In the same way, $h_i(1)$ is $\frac{1}{2}$ at each order for $i = 1, 2$.

*Type space B:* The payoff parameter is $S = \{-1, 1\}$, the set of agents is $N = \{1, 2\}$, the set of types is $T_i = \{0\}$ for $i = 1, 2$, and the belief structure is characterized by a common prior $\mu \in \Delta(S \times T)$ such that

$$
\mu(s, 0, 0) = \begin{cases} 
\frac{1}{2} & \text{if } s = -1 \\
\frac{1}{2} & \text{if } s = 1
\end{cases}
$$
In this case, both agents put probability \( \frac{1}{2} \) on each element of \( S \), and this is common knowledge between the agents. Therefore the resulting sequential belief of the type is \( \frac{1}{2}, \frac{1}{2}, \ldots \) for \( i = 1, 2 \). Type space A and type space B have different type structures, but they result in the same sequential beliefs. It means that the representation of a sequential belief using a Harsanyi type is not unique.

Clearly, the type space A and the type space B in the example have different informational structures.\(^2\) In the example, the types \( t_i = -1 \) and \( t'_i = 1 \) in the type space A are redundant types. The existence of redundant types shows the difficulty in modeling games of incomplete information. We can also interpret these examples in a different way, that is, when Harsanyi type spaces are given, sequential beliefs over the payoff parameter are not enough to characterize the belief structure of agents. The universal type space does not allow redundancy of types. However, without redundant types, we cannot deal with an interesting class of games such as the type space A. In the type space A, redundancy happens due to the strong correlation of the agents’ belief over the payoff parameter and their belief over the other agent’s types. Such correlation is common in applications. In Morris and Shin [41], for instance, the investors share the market information, such as the GDP report and personnel affairs in firms, with some private noises. In their model, the private signals are independent. But, if those private noises are correlated and every agent knows it, in order to model it as a Bayesian game, some types must be strongly correlated with each other and the basic uncertainty so that they result in the same sequential beliefs as in the following example.

**Example 2: Correlated public information with noise**

Let \( N = \{1, \cdots, n\} \). There are two states \( S = \{G, B\} \). The agents receive private signals \( X_i \) about the states from the government. The government tries to hide the state when the state is bad, but it cannot be completely hidden, because there is one agent that receives the true signal. Likewise, when

---

\(^2\)Ely-Peski showed that they have different sets of Bayesian equilibrium and rationalizable strategies.
the state is good, the government tries to make it public, but it cannot do so because one
agents receives a wrong signal. The distribution of the signals and the states are given by a
common prior \( \mu \) such that

\[
\forall i \in N, \mu(X_1 = G, \cdots, X_i = B, \cdots, X_n = G | s = G) = \frac{1}{n},
\]

Otherwise, \( \mu(\cdot | s = G) = 0 \).

\[
\forall i \in N, \mu(X_1 = G, \cdots, X_i = B, \cdots, X_n = G | s = B) = \frac{1}{n},
\]

Otherwise, \( \mu(\cdot | s = B) = 0 \).

Then, each type \( X_i \) assigns probability \( \frac{1}{2} \) to both states. Therefore, the resulting sequential
belief is \( \frac{1}{2}, \cdots \) at each type.

The universal type space has received considerable attention lately\(^3\). But type spaces with
redundant types cannot be represented in the universal type space.

In order to make redundant types tractable in the epistemic space, Ely and Peski ([20])
constructed a different kind of sequential beliefs called \( \Delta \)-hierarchies. That is, sequential be-
liefs over the space of probability distributions over the space of parameters. By using beliefs
over beliefs as the first order belief, we can deal with the correlation between beliefs over the
payoff parameter and beliefs over the other agents’ types. And, in particular, some types
that would be called redundant under standard sequential beliefs are mapped to different
\( \Delta \)-hierarchies. \( \Delta \)-hierarchies can represent richer information about the belief structure of
the agents than ordinary epistemic types, and give us a better foundation to work on the
epistemic analysis of games. In \( \Delta \)-hierarchies, however, we can only distinguish redundant

\(^3\)See Weinstein and Yildiz [51], Dekel, \textit{et al} [17, 18], and Bergemann and Morris [8].
types up to rationalizable actions. Harsanyi types which have different sets of rationalizable actions result in different $\Delta$-hierarchies, but Harsanyi types which share the same set of rationalizable actions result in the same $\Delta$-hierarchy. In the above example, the types $t_i = 1$ in the type space A and $t_i = 0$ in the type space B can be distinguished from each other in $\Delta$-hierarchies, but $t_i = 1$ and $t'_i = -1$ in the type space A cannot be distinguished there. Therefore we cannot always map Harsanyi type spaces into the space of the $\Delta$-hierarchies isomorphically.

Liu [33] took a different approach from Ely-Peski. He augmented the universal type space by adding an additional parameter space, which he called the payoff irrelevant parameter space. He showed that any Harsanyi type space, even if it has redundant types, has its isomorphic image in the space of the sequential beliefs over the payoff parameter $S$ and a payoff irrelevant parameter $C$. However, the payoff irrelevant parameter space that Liu used was the agents’ type space $T$. Therefore the resulting epistemic type space varies depending on Harsanyi type spaces to be studied. Since we cannot compare Harsanyi type spaces on one epistemic space, topological arguments such as Dekel, et al [17] and Ely and Peski [21] are not possible here. In this sense, the space that Liu constructed is different from the universal type space that Mertens-Zamir and Brandenburger-Dekel constructed. Besides, from the epistemic perspective, we cannot obtain any insight into what kind of information beyond the universal type space is needed to deal with the redundancy of types.

In this paper, we offer a solution by finding an exogenous payoff irrelevant parameter space. Moreover, for any Harsanyi type spaces to be mapped, the exogenous parameter space can be a two-valued set $\{0, 1\}$. To get an intuition of our argument, consider a two person game. Let us make on the agents’ Harsanyi type spaces a partition of equivalence classes whose elements have the same sequential belief over the payoff parameter. Equivalently, we sort Harsanyi types into classes of redundant types. In type spaces with redundant types, the beliefs of re-
dundant types have the same probability distribution over the equivalence classes of the other agent’s type space although they are different within each equivalence class. This means that even if redundant types have different conjectures over the payoff parameter and the other agents’s type, they are different just within each equivalence class of the other agent, not across equivalence classes. Since the members of each equivalence class of the other agent’s types cannot be distinguished by their sequential beliefs, the agent’s redundant types also result in the same sequential belief. Our method to deal with redundancy is to distinguish the members of each equivalence class by attaching to each type of an agent a different conjecture over a newly added payoff irrelevant parameter. As a result, those redundant types have different first order beliefs over the payoff parameter and the payoff irrelevant parameter. It enables us to distinguish the redundant types of the other agent by their second order beliefs because we can distinguish their different conjectures within each equivalence class. We have a further result when we assume that $S$ and $T_i$, for all $i \in N$, are uncountable Polish spaces, that is, complete, metrizable and separable spaces. Then it is sufficient to distinguish the redundant types if the payoff irrelevant parameter space has two elements. Here is an explanation. All spaces are infinite and Polish, and so the type space of the agent 1 is Borel equivalent to the closed interval $[0, 1]$. On the other hand, the space of probability measures over $\{0, 1\}$ is homeomorphic to $[0, 1]$. Thus we can assign Borel equivalent different first order beliefs over the set $\{0, 1\}$ to all the types of the player 1. By doing this, we can distinguish the members of each equivalence class of the player 1’s redundant types by their first order beliefs, and so the player 2’s redundant types are distinguishable by their second order beliefs whenever they have different conjectures over the payoff parameter and the player 1’s type space.

Now we can completely represent any Harsanyi type space as a subspace of the “univer-

---

4Even if $S$ and $T$ are countable sets equipped with discrete topology, we can still apply the following argument since we can embed $T$ to $[0, 1]$ Borel isomorphically.
sal type space" over $S \times \{0, 1\}$. It is beneficial for two reasons. First, it gives an epistemic foundation of Harsanyi type spaces. Any correlation of beliefs of agents which is not captured by the sequential beliefs over $S$ can be recovered just by introducing a coin flip as a moderator across agents. Alternatively, any hidden uncertainty in Harsanyi type spaces can be identified as the uncertainty about an agent’s personality. For example, whether or not he believes in God. The sequential conjecture over an agent’s personality generates the correlation of beliefs over the payoff parameter and agents’ types. Second, it allows us to deal with Bayesian games in a “universal” space. The payoff irrelevant parameter $\{0, 1\}$ is exogenous and we do not have to change the payoff irrelevant parameter space as in Liu’s construction. In fact, as we explain later, the points on $U(S \times C)$ characterizes Bayesian equilibrium strategies.

We have other contributions in this paper. One is to fill the gap between two methods in the epistemic game theory: syntactic one and semantic one. Concerning the syntactic method, Sadzik [48] adopted a first order epistemic language a la Aumann [6]. He distinguished Harsanyi types with the sets of the sentences which can be true at the types, and showed that this identification of types is essentially equivalent to identifying types with the set of possible Bayesian equilibrium strategies there. Compared to $\Delta$-hierarchies by Ely-Peski, which identifies with IIR, it is a finer epistemic characterization of Harsanyi types. And, this result also gives an epistemic characterization of Bayesian equilibrium, which is the most successful one so far. However his method is totally different from the existing literature, and we could not compare Sadzik’s syntactic characterization with the other universal type space approach such as Liu’s. In this paper, we show that Sadzik’s syntactic characterization of types is essentially equivalent to whether or not they are mapped to the same sequential beliefs on $U(S \times C)$.

Another contribution is in the Bayesian formulation of complete information games. Since Aumann [2, 3], mixed strategies and correlated strategies in complete information games
have been given a foundation by assuming a basic uncertainty not described in the game\textsuperscript{5} and reinterpreting them as incomplete information games. However, the resulting games often have redundant types. Brandenburger and Friedenberg [13] considered the set of correlated equilibria which can be achieved only thorough the correlation of sequential beliefs over the basic uncertainty(\textit{intrinsic correlation}). This is equivalent to consider the set of correlated equilibria achieved in Bayesian formulations without redundant types. Therefore, we can apply our result and show that every correlated equilibrium can be achieved through intrinsic correlation when we add a coin flip to the basic uncertainty. It is the same as the result in Brandenburger-Friedenberg derived in a different way.

This paper is organized as follows. In Section 1.3 and 1.4, we present the formal model and the proof of our main result: the elimination of redundancy by adding the payoff irrelevant parameter space $C = \{0, 1\}$. In Section 1.5, we characterize our result with Bayesian equilibrium, and interpret Sadzik’s syntactic approach on the universal type space. In Section 1.6, we discuss the intrinsic correlation in terms of redundant types. In the appendix, we provide detailed proofs about some measurability issues involved in our construction.

\section{1.2 Preliminaries}

We briefly explain mathematical symbols required in this thesis. We make it more specific as we go on depending on necessity.

\subsection{1.2.1 Mathematical notations}

Throughout this thesis, we assume that all spaces are endowed with some topology. Let $Y$ be an arbitrary set endowed with some topology. Then we use $\Sigma(Y)$ for the Borel sets on $Y$, and $\Delta(Y)$ for the set of probability measures on $(Y, \Sigma(Y))$ endowed with $w^*$-topology. $\mathcal{K}(Y)$

\textsuperscript{5}Without loss of generality, we can consider it to be the space of actions.
is the set of compact subspaces of $Y$ endowed with Hausdorff topology.\footnote{See details in Kechris \cite{32}.}

Let $N$ be a finite set, and $(Y_i)_{i \in N}$ be a family of sets. Then, for any $i \in N$, we use $Y_{-i}$ to denote the product space $\Pi_{j \in N \setminus \{i\}} Y_j$.

### 1.2.2 Harsanyi type space

We consider a finite set of agents $N = \{1, \ldots, n\}$. All the agents face the same basic uncertainty about their payoffs. It can be represented by a parameter space $S$.\footnote{See Mertens-Zamir \cite{39}, and Hu-Stuart \cite{28}.} We call this $S$ the payoff parameter space. A Harsanyi type space is a tuple $\langle S, (T_i)_{i \in N}, (\lambda_i)_{i \in N} \rangle$, where, for each $i \in N$, $\lambda_i$ is a function from $T_i$ to $\Delta(S \times T_{-i})$. We call each element $t_i \in T_i$ a Harsanyi type. By the function $\lambda_i$, each type stands for a belief over the payoff parameter and the other players’ types. Hereafter we make some assumptions on Harsanyi type spaces.

**Assumption 1:** The parameter space $S$ and the each agent’s type space $T_i$ are uncountable Polish spaces.

Let $T \equiv \Pi_{i \in N} T_i$. Then, as it is known, the product type space $T$ is also a Polish space.

In many works such as Mertens-Zamir \cite{39} *etc.*, the belief mapping $\lambda_i$ is assumed to be homeomorphism. Here we relax this usual assumption slightly.

**Definition 1.2.1.** A function $f : X \to Y$ is bimeasurable if $f$ is measurable and, for each measurable set $E \subset X$, $f(E)$ is also measurable.
Assumption 2: For each $i \in N$, the function $\lambda_i$ is a \textit{bimeasurable} injection.

This assumption precludes \textit{purely redundant} types, which are Harsanyi types $t_i, t'_i \in T_i$ such that $t_i \neq t'_i$ and $\lambda_i(t_i) = \lambda_i(t'_i)$.

1.3 Universal type spaces

The universal type space was introduced by Mertens-Zamir\cite{39}. It is the space of the sequential beliefs over $S$ which satisfy some coherency conditions. They showed that the space is also a Harsanyi type space and any Harsanyi type space without redundant types is embedded there. To define the universal type space, we have to define the space of the sequential beliefs first. Let a family of spaces $(Z^k)_{k \geq 1}$ be such that

$$Z^1 \equiv S$$

$$\text{For } k > 1, Z^k \equiv Z^{k-1} \times \Delta(Z^{k-1}).$$

The space $Z^k$ is the set of the $k$th order beliefs over $S$. We say that $\prod_{k=1}^{\infty} Z^k$ is the \textit{sequential belief space} and each element of it is the \textit{sequential belief}. Let a sequential belief be $z \equiv (z_1, \cdots)$ where, for all $k \in \mathbb{N}$, $z_k \in Z_k$. We say that $z$ satisfies \textit{coherency} if, for all $k \in \mathbb{N}$, the marginal distribution of $z_{k+1}$ over $Z_k$ is the same as $z_k$. Under coherency of beliefs, we can consider each element in $\prod_{k=1}^{\infty} Z^k$ as a projection limit. Let the set of the projection limits be $Z^\infty$. We say that each $e_i \in Z^\infty$ is an \textit{epistemic type}. The \textit{universal type space} is the set of all the sequential beliefs that satisfy coherency. We denote it as $U(S)$. Mertens-Zamir showed the following strong theorem about the universal type space.
Theorem 1.3.1. (Mertens-Zamir [39]) The universal type space $U(S)$ and its associated natural homeomorphism constitutes a Harsanyi type space.

Then we can define the function which maps Harsanyi types onto the sequential belief space. Let the first order mapping $h_1^1 : T_i \to \Delta(S)$ be such that

$$h_1^1(t_i) = \text{Marg}_i(\lambda_i(t_i)).$$

For $k > 1$, let the $k$th order mapping $h_k^k : T_i \to \Delta(Z^k)$ be such that

$$h_k^k(t_i) = \lambda_i(t_i) \circ [\text{Id}_S, (h_j^{k-1})_{j \in N \setminus \{i\}}]^{-1},$$

where $\text{Id}_S$ is an identical function from $S$ to $S$.

We say that the function $(h_k^k)_{k=1}^\infty : T_i \to \Pi_{k=1}^\infty Z_k$ is the hierarchy mapping. Let $h \equiv (h_i)_{i \in N}$. Then, this $h$ enables us to map any Harsanyi type space to the sequential belief space. Also you can see that sequential beliefs derived in this way satisfy the coherency condition.

1.4 An extended sequential belief space

In this section, we extend the universal type space by adding a payoff-irrelevant parameter space $C$. And we show that we can isomorphically embed Harsanyi type spaces there even if they have redundant types.
1.4.1 Redundant types

Let $\Lambda = \langle S, T, (\lambda)_{i \in N} \rangle$ be a Harsanyi type space. Mertens-Zamir showed that Harsanyi type spaces can be embedded as a subspace of $U(S)$ homeomorphically only if they have no redundant types. To discuss the matter, we have to define redundant types first.

**Definition 1.4.1.** In a Harsanyi type space $\Lambda$, two Harsanyi types $t_i$ and $t'_i \in T_i$ are redundant if $h_i(t_i) = h_i(t'_i)$.

We say that the Harsanyi types which are not redundant are non-redundant types. Then we can formally state what Mertens-Zamir showed.

**Proposition 1.4.2.** (Mertens and Zamir [39]) Any Harsanyi type space without redundant types can be embedded onto $U(S)$ homeomorphically. And the hierarchy mapping $h$ is the unique embedding.

1.4.2 Extension with a payoff irrelevant parameter space

Now we construct an extended space of sequential beliefs so that we can embed Harsanyi type spaces there even if they have redundant types. We introduce a parameter space $C = \{0, 1\}$ and consider the sequential belief space over $S \times C$ instead of $S$. In the rest of this section, we assume that $N = \{1, 2\}$.

Let $C \equiv \{0, 1\}$. We assume that any element does not affect the payoffs of the agents.
Therefore we call $C$ the *payoff irrelevant parameter space*. We define sequential beliefs over $S \times C$ and construct the coherent sequential belief space over $U(S \times C)$ in the same way as we did over $S$.

Let

$$Z_1 \equiv S \times C,$$

$$\forall k \geq 2, \ Z_k \equiv \Delta(\Pi_{n=1}^{n=k-1}Z_n).$$

And let

$$H^k(S \times C) \equiv \Delta(\Pi_{n=1}^{n=k}Z_k)$$

$$= Z_{k+1}.$$ 

and

$$H(S \times C) \equiv \Pi_{k=1}^{k=\infty}H^k(S \times C)$$

$$= \Pi_{k=1}^{k=\infty}\Delta(Z_k).$$

For each $k$, $H^k(S \times C)$ is the set of the $k$th order belief over $S \times C$. Let $U(S \times C) \subset \Pi_{i \in N}H(S \times C)$ be the product space of the coherent sequential beliefs.

We also define Harsanyi type spaces based on $S \times C$ by the sequence $\Phi = (S \times C, V, (\phi_i)_{i \in N})$ where $\phi_i$ is a bimeasurable injection from $V_i$ to $\Delta(S \times C \times V_{-i})$.

Before we embed a Harsanyi type space onto $U(S \times C)$, we extend it to a Harsanyi type space on $S \times C$. To do that, we should clarify what is “isomorphism” between Harsanyi type spaces.
Definition 1.4.3. (Liu [33]) Let $X = \langle S, T, \lambda \rangle$ and $Y = \langle S \times C, V, \phi \rangle$ be Harsanyi type spaces on $S$ and $S \times C$ respectively. Then, $X$ and $Y$ are $S$-isomorphic to each other if there exists a $g = (g_i)_{i \in \{0\} \cup N}$ such that (1) $g_0 : S \to S$ is an identity function, (2) $g_i : T_i \to V_i$ is Borel equivalence for all $i \in N$, and (3) $\text{Marg}_{S \times V} \phi_i(v_i) = \lambda_i(t_i) \circ g^{-1} \circ \text{Proj}_{S \times V}$.

Hereafter, when Harsanyi type spaces $X$ and $Y$ are $S$-isomorphic, we use $X \sim_S Y$. And when both spaces are defined on $S$, we use $X \sim Y$.

Next, we want to construct a Harsanyi type space on $S \times C$ which is $S$-isomorphic to the original type space on $S$. For the construction, we need the next well-known theorem.\(^8\)

Theorem 1.4.4. Let $X$ be an uncountable Polish space. Then $X$ is Borel equivalent to the closed interval $[0, 1]$.

Let $\Lambda \equiv \langle S, (T_i)_{i \in \{1,2\}}, (\lambda_i)_{i \in \{1,2\}} \rangle$ be a Harsanyi type space. Since $T_i$ is an uncountable Polish space, there exists a Borel equivalence from $T_i$ to $[0, 1]$. Let this equivalence be $p_i : T_i \to [0, 1]$. Using $p_i$, we define a Harsanyi type space $\Phi = \langle S \times C, (V_i)_{i \in \{1,2\}}, (\phi_i)_{i \in N} \rangle$ so that

1. For all $i \in \{1, 2\}$, $V_i = [0, 1]$.

2. For all $i \in \{1, 2\}$, $\phi_i : V_i \to \Delta(S \times C \times V_{-i})$ satisfies the next property;

For the agent 1,

\(^{8}\text{See Royden [47] for the detailed argument.}\)
\( \forall v_1 \in V_1, \quad \text{Marg}_{S \times V_2} \phi_1(v_1) = \lambda_1(p_1^{-1}(v_1)) \circ [\text{Id}_S, p_2]^{-1}, \)

\( \forall E \in \Sigma(S \times V_2), \quad \phi_1(v_1)[E \times \{0\}] = v_1 \lambda_1(p_1^{-1}(v_1)) \circ [\text{Id}_S, p_2]^{-1}[E]. \)

For the agent 2,

\( \forall v_2 \in V_2, \quad \text{Marg}_{S \times V_1} \phi_2[v_2] = \lambda_2[p_2^{-1}(v_2)] \circ [\text{Id}_S, \ p_1]^{-1}, \)

\( \text{Marg}_{(C)} \phi_2(\{0\}) = 1. \)

The bimeasurability of \((\phi_i)_{i \in \{1,2\}}\) is proven in the appendix. Then, you can see that \(\Phi\) is a well defined Harsanyi type space. Concerning this Harsanyi type space \(\Phi\), we have the next fundamental lemma.

**Lemma 1.4.5.** The above type space \(\Phi\) is S-isomorphic to \(\Lambda\).

*Proof.* Let \(\text{Id}_S : S \rightarrow S\) be identity function. Then, \((\text{Id}_S, p_1, p_2)\) is S-isomorphism from \(\Lambda\) to \(\Phi\) by construction. \(\square\)

### 1.4.3 S-isomorphic embedding onto \(U(S \times C)\)

We go to the main part of this paper. We show that, in the Harsanyi type space \(\Phi\) defined above, all elements of \(V_i\) correspond to different sequential beliefs over \(S \times C\).
Theorem 1.4.6. Let $\Lambda$ and $\Phi$ be Harsanyi type spaces defined above. Then, for each $i \in \{1, 2\}$, the agent $i$’s hierarchy mapping induced by $\Phi$, $h_i : V_i \to H(S \times C)$, is an injection.

Proof. Let $h^k_i : V_i \to H^k_i(S \times C)$ be the agent $i$’s $k$th order belief mapping on $S \times C$ induced by $\Phi$, and let $g^k_i : T_i \to H^k_i(S)$ be the agent $i$’s $k$th order belief mapping onto $S$ induced by $\Lambda$.

(Step 1: For the agent 1)

Let $v_1, v'_1 \in V_1$ be such that $v_1 \neq v'_1$. His first order belief of $v_1$ is

$$
\forall E \in \Sigma(S), \ h^1_1(v_1)[E \times \{0\}] = \phi_1(v_1)[E \times \{0\} \times V_2] = v_1 \lambda_1(p^{-1}_1(v_1)) \circ \text{Id}_{S, p_2}^{-1}[E \times V_2] = v_1 g^1_1(p^{-1}_1(v_1))[E].
$$

By the symmetric argument,

$$
\forall E \in \Sigma(S), \ h^1_1(v'_1)[E \times \{0\}] = v'_1 g^1_1(p^{-1}_1(v'_1))[E].
$$

(Case 1:) Suppose that $v_1 g^1_1[p^{-1}_1(v_1)](E) = v'_1 g^1_1[p^{-1}_1(v'_1)](E)$. Then, since $v_1 \neq v'_1$, $g^1_1[p^{-1}_1(v_1)](E) \neq g^1_1[p^{-1}_1(v'_1)](E)$. 

On the other hand,

\[ h_1(v_1)[E \times C] = \phi_1(v_1)[E \times C \times V_2] \]
\[ = \operatorname{Marg}_{(S \times V_2)} \phi_1(v_1)[E \times V_2] \]
\[ = \lambda_1(p_1^{-1}(v_1)) \circ [\operatorname{Id}_S, p_2]^{-1}[E \times V_2] \]
\[ = g_1(p_1^{-1}(v_1))[E]. \]

From these results, we have

\[ h_1(v_1)[E \times \{1\}] = h_1(v_1)[E \times C] - h_1(v_1)[E \times \{0\}] \]
\[ = g_1(p_1^{-1}(v_1))[E] - v_1 g_1(p_1^{-1}(v_1))[E] \]
\[ = g_1(p_1^{-1}(v_1))[E] - v_1' g_1(p_1^{-1}(v_1'))[E] \]
\[ \neq g_1(p_1^{-1}(v_1'))[E] - v_1' g_1(p_1^{-1}(v_1'))[E] \]
\[ = h_1(v_1')[E \times \{1\}] . \]

Thus \( h_1 \) is injective.

(Case 2:) Suppose that \( v_1 g_1(p_1^{-1}(v_1))(E) \neq v_1' g_1(p_1^{-1}(v_1'))(E) \). It means that \( h_1(v_1)[E \times \{0\}] \neq h_1(v_1')[E \times \{0\}] \). Thus \( h_1 \) is injection.

(Step 2: For the agent 2)
Let $v_2, v'_2 \in V_2$ be such that $v_2 \neq v'_2$. Concerning his first order belief, by construction,

$$
\forall E \in \Sigma(S), \\
h_2^1(v_2)(E \times \{0\}) = g_2^1(p_2^{-1}(v_2))[E]. \\
h_2^1(v_2)(E \times \{1\}) = 0.
$$

Let, for each $\mu_1 \in \Delta(S \times C)$, $h_1^{-1}(\mu_1) \equiv \{v_1 \in V_1 : h_1^1(v_1) = \mu_1\}$. As we have shown, the function $h_1^1 : V_1 \rightarrow \Delta(S \times C)$ is injective. Therefore $h_1^{-1} : h_1^1(V_1) \rightarrow V_1$ is the well defined inverse bijection.

Then we can derive the agent 2’s second order belief over $S \times C$.\(^9\)

Note that

$$
\forall E \in \Sigma(S), \forall Q \in \Sigma(\Delta(S \times C))
\\h_2^2(v_2)(E \times \{0\} \times Q) = \phi_2(v_2)(E \times \{0\} \times h_1^{-1}(Q))
\\= \lambda_2(p_2^{-1}(v_2)) \circ [\text{Id}_S, p_1]^{-1}[E \times h_1^{-1}(Q)].
$$

Since $\lambda_2 : T_2 \rightarrow \Delta(S \times T_1)$ is a bimeasurable injection, $\lambda_2(p_2^{-1}(v_2)) \neq \lambda_2(p_2^{-1}(v'_2))$. By Dynkin’s lemma\(^{10}\), there exists a rectangle $F \equiv \hat{S} \times \hat{T_1}$ such that $\hat{S} \in \Sigma(S)$, $\hat{T_1} \in \Sigma(T_1)$, and $\lambda_2(p_2^{-1}(v_2))[F] \neq \lambda_2(p_2^{-1}(v'_2))[F]$. Let $\hat{V}_1 \equiv p_1(\hat{T_1})$. Then, $\hat{V}_1 \in \Sigma(V_1)$ and $h_1^1(\hat{V}_1) \in$

\(^9\)Concerning the bimeasurability of $h_1^1$, see appendix.

\(^{10}\)See Theorem 10-10 in [1]
\[ \Sigma(\Delta(S \times C)). \] Therefore

\[
h_2^2(v_2)(\hat{S} \times \{0\} \times h_1^1(\hat{V}_1)) = \phi_2(v_2)[\hat{S} \times \{0\} \times h_1^{-1}(h_1^1(\hat{V}_1))] \\
= \phi_2(v_2)[\hat{S} \times \{0\} \times \hat{V}_1] \\
= \lambda_2[p_2^{-1}(v_2)] \circ [\text{Id}_S, p_1^{-1}](\hat{S} \times \hat{V}_1) \\
= \lambda_2(p_2^{-1}(v_2))(\hat{S} \times \hat{T}_1) = \lambda_2(p_2^{-1}(v_2))[F] \\
\neq \lambda_2(p_2^{-1}(v_2'))[F] \\
= h_2^2(v_2')(\hat{S} \times \{0\} \times h_1^1(\hat{V}_1)).
\]

It means that \( h_2(v_2) \neq h_2(v_2') \). Therefore, \( h_2 \) is injection. \( \square \)

So far we did not consider topological structures of Harsanyi type spaces except that they are Polish. As it plays a crucial role in Weistein and Yildiz \cite{51} and others, it is important to show that each agent’s type space \( V_i \) is homeomorphic to the belief space \( \Delta(S \times C \times V_{-i}) \).

**Definition 1.4.7.** A Harsanyi type space \( \mathcal{X} = (X, T, (x_i)_{i \in N}) \) is a continuous type space if, for all \( i \in N \), \( x_i : T_i \to \Delta(X \times T_{-i}) \) is homeomorphic embedding.

Next we show that the embedded image on \( U(S \times C) \) of each Harsanyi type by the hierarchy mapping is a continuous Harsanyi type space.

**Lemma 1.4.8.** Let \( \Phi \) be a type space and \( H(S \times C) \) be the space of sequential belief over \( S \times C \) as we defined before. The function \( h_i : V_i \to H(S \times C) \) is the full hierarchy mapping.
Now let \( f : h_i(V_i) \to \Delta(S \times C \times h_i(V_{-i})) \) be such that \( f(h_i(v_i)) \equiv \phi(v_i) \circ [\text{Id}_{S \times C}, h_{-i}]^{-1} \). Then, \( f \) is homeomorphism.

**Proof.** Since \( S \times C \) is a Polish space, there exists a unique homeomorphism \( \psi : H(S \times C) \to \Delta(S \times C \times H(S \times C)) \) such that, for each \( m \in H(S \times C) \), \( \psi(m) \) is the Kolmogorov extension of \( m \).\(^{11}\) So it is enough to show that, for all \( i \in \{1, 2\} \) and \( v_i \in V_i \), \( f_i(h_i(v_i)) \) is the Kolmogorov extension of \( h_i(v_i) \)

Let \( m_i \in h_i(V_i) \). First, for all \( E \in \Sigma(S \times C \times H(S \times C)) \), by letting \( f_i(m_i)(E) \equiv f_i(m_i)[E \cap (S \times C \times h_{-i}(V_{-i}))] \), we can extend \( f_i(m_i) \) so that \( f_i(m_i) \in \Delta(S \times C \times H(S \times C)) \). And as we defined before,

\[
Z_1 \equiv S \times C,
\]
\[
\forall k \geq 2, \quad Z_k \equiv \Delta(\Pi_{n=1}^{n=k-1} Z_n),
\]

\[
H^k(S \times C) = \Delta(\Pi_{n=1}^{n=k} Z_k)
= Z_{k+1},
\]

and

\[
H(S \times C) = \Pi_{k=1}^{k=\infty} H^k(S \times C)
= \Pi_{k=1}^{k=\infty} \Delta(Z_k).
\]

The equations above also imply that

\[
H(S \times C) = \Pi_{n=2}^{n=\infty} Z_n.
\]

\(^{11}\)See Prop 1 and Prop 2 in Brandenburger-Dekel [11]
Therefore,
\[ S \times C \times H(S \times C) = \prod_{n=1}^{\infty} Z_n. \]

From these equalities, we have \( f_i(m_i) \in \Delta(\prod_{n=1}^{\infty} Z_n) \), and \( m_i \in H(S \times C) = \prod_{k=1}^{\infty} \Delta(Z_k) \).

To show that \( f_i(m_i) \) is the Kolmogorov extension of \( m_i \), it is enough to show that the next property holds:
\[
\forall k, \ Marg(\prod_{n=1}^{k} Z_n) f_i(m_i) = \text{Proj}_k m_i.
\]

Let \( E \in \Sigma(\prod_{n=1}^{k} Z_n) \) and \( \hat{E} = E \times \prod_{n=k+1}^{\infty} Z_n \). Then,
\[
f(m_i)(\hat{E}) = \phi[v_i] \circ [\text{Id}_{S \times C}, h_{-i}]^{-1}(\hat{E} \cap h_{-i}(V_{-i}))
= \phi(v_i)(E_1 \times \hat{V}_{-i}^k),
\]
where \( \hat{V}_{-i}^k = \{v_{-i} \in V_{-i} : h_{-i}^k(v_{-i}) \in \prod_{n=2}^{k} E_n\} \).

On the other hand, from the \( k \)th order belief of the agent \( i \),
\[
\exists v_i \in V_i, \text{Proj}_k m_i[E] = h_{-i}^k[v_i](E)
= \phi_i[v_i](E_1 \times \hat{V}_{-i}).
\]

This equation means that \( f_i(m_i)[\hat{E}] = \text{Proj}_k(m_i)(E) \). Consequently, \( f_i(m_i) \) is the Kolmogorov extension of \( m_i \).

As a consequence, we have the next theorem.
**Theorem 1.4.9.** For any Harsanyi type space $\Lambda = \langle S, T, (\lambda_i)_{i \in \{1,2\}} \rangle$, there exists a continuous BL-subspace in $U(S \times C)$ which is $S$-isomorphic to $\Lambda$.

*Proof.* Let a Harsanyi type space $\Phi = \langle S \times C, V, \phi \rangle$ be an $S$-isomorphic extension of $\Lambda$, and let $E_i = h_i(V_i)$ for all $i \in N$. Let $\mathcal{E} = \langle S \times C, E, (f_i)_{i \in N} \rangle$, where $f_i$ is defined as in the lemma. Since $h_i$ is bimeasurable injection, $\mathcal{E}$ is $S$-isomorphic to $\Phi$ by construction. As a direct result of the lemma, $\mathcal{E}$ is a continuous Harsanyi type space. \qed

1.4.4 Extension to $N > 2$

We can extend the above theorems to $N$-person game. Let $N$ be the finite set of the agents and $|N| = n$. Consider an $N$-person Harsanyi type space $\Lambda = \langle S, T, (\lambda_i)_{i \in N} \rangle$ as before. We maintain the same assumptions on $S$, $T$, $C$ and $\lambda_i$.

We can define an extension of $\Lambda$ on $S \times C$, $\Phi = \langle S \times C, V, (\phi_i)_{i \in N} \rangle$, as follows. For all $i \in N$, let $p_i : T_i \to [0, 1]$ be a Borel equivalence. Let $\Phi$ be such that

\[
\forall i \in N, \ V_i = [0, 1].
\]
\[
\forall i \in N \setminus \{1\}, \ \forall v_i \in V_i, \ \forall E \in \Sigma(S \times V_{-i}),
\phi_i(v_i)[E \times \{0\}] = v_i\{\lambda_i(p_i^{-1}(v_i)) \circ [\text{Id}_s, p_{-i}^{-1}](E)\}.
\]
\[
\text{Marg}_{(S \times V_{-i})}\phi_i(v_i) = \lambda_i(p_i^{-1}(v_i)) \circ [\text{Id}_s, p_{-i}^{-1}].
\]
And,

∀v_1 ∈ V_1, ∀E ∈ Σ(S × V_{-1}),

\[ \phi_1(v_1)[E × \{0\}] = \lambda_1(p_{-1}^{-1}(v_1)) ◦ [\text{Id}_s, p_{-1}^{-1}](E). \]

\[ \text{Marg}_{(S×V_{-1})}\phi_1(v_1) = \lambda_1(p_{-1}^{-1}(v_1)) ◦ [\text{Id}_s, p_{-1}^{-1}]. \]

In the same way as we did above, we can show that Φ is S-isomorphic to Λ and the resulting hierarchy mapping is injection.

1.5 The characterization of the sequential belief in \( U(S × C) \)

In this section, we show that the sequential belief in \( U(S × C) \) and Bayesian equilibrium characterize each other. In this process, a new notion, symmetric types, gives us a new insight and help us to establish the characterization.

1.5.1 Symmetric types

Throughout this section, we assume that \( S \) and \( T \) are finite for technical convenience. Unless otherwise stated, the results below are valid in the general case of infinite \( S \) and \( T \).

**Definition 1.5.1.** Let \( t_i, t_i' \in T_i \) in \( \Lambda \). The Harsanyi types \( t_i \) and \( t_i' \) are one sided symmetric if there exists a permutation \( \pi_{-i} : T_{-i} → T_{-i} \) such that, for all \( E \in Σ(S) \),

\[ \forall t_{-i} ∈ T_{-i}, \; \lambda_i(t_i)(s, t_{-i}) = \lambda_i(t_i')(s, \{\pi_{-i}(t_{-i})\}). \]

**Definition 1.5.2.** The Harsanyi types \( t_i \) and \( t_i' \) in \( \Lambda \) are symmetric if (1) \( t_i \) and \( t_i' \) are one side symmetric, (2) for each \( t_{-i} ∈ T_{-i} \), \( t_{-i} \) and \( \pi_{-i}(t_{-i}) \) are one sided symmetric with regard
to a permutation \( \pi_i : T_i \rightarrow T_i \), and (3) \( t_i' = \pi_i(t_i) \).

The permutation \( \pi \equiv (\pi_i, \pi_{-i}) \) is just a renaming of types in \( \Lambda \). The above definition states that we can exchange the roles of symmetric types without changing the entire structure of \( \Lambda \). As a result we can say that symmetric types have the same set of Bayesian equilibria in any game.

**Definition 1.5.3.** A game \( \Gamma \) on \( S \) is a tuple of \( ((u_i)_{i \in N}, A) \), where \( u_i : A \times S \rightarrow \mathbb{R} \).

Let \( \beta_i : T_i \rightarrow A \) be the agent \( i \)'s (pure) strategy. Bayesian equilibrium is defined as follows;

**Definition 1.5.4.** A tuple of strategies \( \beta \equiv (\beta_i)_{i \in N} \) is Bayesian equilibrium if, for all \( i \in N \), \( t_i \in T_i \), and \( a_i \in A_i \),

\[
\int_{S \times T_{-i}} u_i(\beta_i(t_i), \beta_{-i}, s) d\lambda_i(t_i) \geq \int_{S \times T_{-i}} u_i(a_i, \beta_{-i}, s) d\lambda_i(t_i).
\]

**Definition 1.5.5.** For each \( t \in T \) in \( \Lambda \) and \( \Gamma \),

\[
BE(t, \Gamma) \equiv \{ a \in A : \exists \beta^* \text{ s.t. } \beta^* \text{ is Bayesian equilibrium in } \Gamma, \text{ and } \beta^*(t) = a \}.
\]

**Proposition 1.5.6.** Let \( t_i, t_i' \in T_i \) be symmetric types. Then, for any game \( \Gamma \), \( BE(t_i, \Gamma) = BE(t_i', \Gamma) \).

*Proof.* Let \( \tilde{t}_i, \hat{t}_i \in T_i \) be symmetric types. Suppose that \( a_i^* \in BE(\tilde{t}_i, \Gamma) \). Then, there exists a B.E. \( \tilde{\beta} \) such that \( \tilde{\beta}_i(\tilde{t}_i) = a_i^* \).
Now \( \hat{t}_i \) and \( \hat{\hat{t}}_i \) are symmetric. Therefore, there exists \((\pi_i)_{i \in N}\) such that

\[
\forall i \in N, \forall t_i \in T_i, \lambda_i(t_i) = \lambda_i(\pi_i(t_i)) \circ [\text{Id}_S, \pi_i]. \tag{1.1}
\]

Let \( \hat{\beta} \) be a pair of strategies such that, for each \( i \in N \) and \( t_i \in T_i \), \( \hat{\beta}_i(t_i) = \hat{\beta}_i(\pi_i(t_i)) \). Under the strategy \( \hat{\beta} \), for each \( i \in N \) and \( t_i \in T_i \), the expected payoff to the agent \( i \) by taking an action \( a_i \) at his type \( \pi_i(t_i) \) is:

\[
U_i(a_i, \hat{\beta}_i, \pi_i(t_i)) = \int u_i(s, a_i, \hat{\beta}(t_i))d\lambda_i(\pi_i(t_i)) \tag{1.2}
\]

\[
= \int u_i(s, a_i, \hat{\beta}(\pi_i(t_i)))d\lambda_i(\pi_i(t_i)) \tag{1.3}
\]

\[
= \int u_i(s, a_i, \hat{\beta}(t_i))d\lambda_i(t_i) \tag{1.4}
\]

\[
= U_i(a_i, \hat{\beta}_i, t_i). \tag{1.5}
\]

Since \( \hat{\beta}_i(t_i) \in \arg\max_{a_i \in A_i} U_i(a_i, \hat{\beta}_i, t_i) \), we have, for each \( i \in N \) and \( t_i \in T_i \), \( \hat{\beta}_i(\pi_i(t_i)) \in \arg\max_{a_i \in A_i} U_i(a_i, \hat{\beta}_i, \pi_i(t_i)) \). Therefore, \( \hat{\beta} \) is also BNE of the game \( \Gamma \). Since \( \hat{t}_i = \pi_i(\hat{t}_i) \), \( \hat{\beta}_i(\hat{t}_i) = \hat{\beta}_i(\hat{t}_i) \). Therefore, \( BE(\hat{t}_i, \Gamma) \subset BE(\hat{\hat{t}}_i, \Gamma) \). By the symmetric argument, we also have \( BE(\hat{t}_i, \Gamma) \supset BE(\hat{t}_i, \Gamma) \). Thus \( BE(\hat{t}_i, \Gamma) = BE(\hat{\hat{t}}_i, \Gamma) \). \( \square \)

We show that symmetric types characterize the sequential belief in \( U(S \times C) \).

**Proposition 1.5.7.** Symmetric types are \( S \)-isomorphically mapped to the same points on \( U(S \times C) \).

**Proof.** Let \( \hat{t}_i, \tilde{t}_i \in T_i \) be symmetric types. For \( i = 1, 2 \), let \( \pi \equiv (\pi_i)_{i \in N} \) be a permutation defined in the definition of symmetric types.

We want to show that \( \pi \) is a \( S \)-isomorphism from \( \Lambda \) to itself. Since \( T \) is countable, \( \pi_i \) is
Borel isomorphism from $T_i$ to $T_i$. By the definition of symmetry, we have that:

$$\forall i \in N, \forall t_i \in T_i, \lambda_i(t_i) = \lambda_i(\pi_i(t_i)) \circ [\text{Id}_S, \pi^{-1}_i].$$

It means that $\pi$ is S-isomorphism from $\Lambda$ to $\Lambda$. Also, the definition of symmetry implies that $\hat{t}_i = \pi_i(\tilde{t}_i)$. Both of the types are S-isomorphically mapped to each other. Thus they are S-isomorphically mapped to the same points on $U(S \times C)$.

Next we have to show that the symmetry characterizes the sequential belief in $U(S \times C)$.

**Proposition 1.5.8.** When we embed $\Lambda$ to $M \subset U(S \times C)$, if two Harsanyi types in $T_i$ can be mapped to the same point in $M$ by some S-isomorphisms, then they are symmetric.

**Proof.** Let $t_i, t'_i \in T$ and $h$ be an S-isomorphic embedding from $T$ to $M$. Suppose that there exists another S-isomorphism $h' : T \to M$ such that $h'(t') = h(t)$. Then $g \equiv h^{-1} \circ h'$ is an S-isomorphism from $T$ to $T$ and $t = g(t')$. By the definition, $t$ and $t'$ are symmetric with regard to the permutation $g$. \[\square\]

1.5.2 Characterization of symmetric types

Next we show that symmetric types are characterized by Bayesian equilibrium. We will make use of a result of Sadzik [48], who adopted the syntactic approach to obtain an epistemic characterization of Bayesian equilibrium. We show that Sadzik’s syntactic condition is characterized by the symmetry of types.

We do not deal with the syntactic details here. We introduce only the part of the paper that we need here. Sadzik added a signal to the Harsanyi type.
Definition 1.5.9. For each \(i \in N\), let \(X_i \equiv [0, 1]^N\). A signal from \(T_i\) to \(X_i\) is a function \(z_i : T_i \rightarrow X_i\).

Let \(z \equiv (z_i)_{i \in N}\). We denote \(Z\) as the set of all possible \(z\). We assume \(z\) is common knowledge among the agents. Then, given a Harsanyi type space \(\Lambda\) and the realized private information \(z\), we can derive a hierarchy mapping \(\delta^z_i : T_i \rightarrow U(S \times X)\) in the same way we derived sequential beliefs over \(S\). Let \(\delta^z \equiv (\delta^z_i)_{i \in N}\). Notice that \(z\) does not have to be a bijection. Therefore, the image of \(\Lambda\) by \(\delta^z\) is no longer \(S\)-isomorphic to \(\Lambda\) generally.

We assume that \(A\) is Polish. Since \(X_i\) is the Hilbert cube, we can embed \(A_i\) to \(X_i\). Therefore we can interpret the set of signals \(Z\) as the set of potential strategies. For the characterization, we need the next notation.

Definition 1.5.10. For each \(t \in T\) in \(\Lambda\) and \(\Gamma\),

\[
LBE(t, \Gamma) \equiv \{a \in A : \exists \beta^* \text{ s.t. } \beta^* \text{ is B.E. in } \Lambda^t \text{ and } \Gamma, \text{ and } \beta^*(t) = a\}.
\]

Here \(\Lambda^t\) is the smallest sub type space of \(\Lambda\) which includes \(t\).

Theorem 1.5.11. (Sadzik [48]) For \(t, t' \in T\) in \(\Lambda\), if \(\{\delta^z(t) : z \in Z\} = \{\delta^z(t') : z \in Z\}\), then, \(LBE(t, \Gamma) = LBE(t', \Gamma)\) for any \(\Gamma\).

Theorem 1.5.12. (Sadzik [48]) For \(t, t' \in T\) in \(\Lambda\), if \(BE(t, \Gamma) = BE(t', \Gamma)\) for any \(\Gamma\), then \(\{\delta^z(t) : z \in Z\} = \{\delta^z(t') : z \in Z\}\).\(^{12}\)

Next we show how to interpret Sadzik’s characterization by using the universal type space argument.

\(^{12}\)When \(T\) is infinite, the latter part is relaxed to the equality of the closure of \(\{\delta^z(t) : z \in Z\}\) and \(\{\delta^z(t') : z \in Z\}\).
Proposition 1.5.13. For \( t, t' \in T \) in \( \Lambda \), if \( t \) and \( t' \) are symmetric, then \( \{ \delta^z(t) : z \in Z \} = \{ \delta^z(t') : z \in Z \} \).

Proof. Suppose that \( t, t' \in T \) are symmetric and \( \pi \) is the associated permutation on \( T \) such that \( t = \pi(t') \). Let \( z \in Z \). Then \( z \circ \pi^{-1} \in Z \). All we have to show is \( \delta^z(t') = \delta^{z \circ \pi^{-1}}(t) \).

Let \( \delta^z_k \) be the \( k \)th order belief mapping induced by a signal \( z \), and let \( H^k_X \) be the space of the \( k \)th order sequential beliefs over \((S \times X)\). Then

\[
\forall i \in N, \forall l \in T, \\
\forall s \in S, \forall x \in X, \delta^z_{1,i}(l_i)(s, x) = \lambda_i(l_i)[\text{Id}_S, z_i]^{-1}(s, x).
\]

\[
\forall E \in \Sigma(H^k_X), \delta^z_{k+1,i}(l_i)[E] = \lambda_i(l_i)[\text{Id}_S, \delta^z_{k,i}]^{-1}[E].
\]

We also have

\[
\forall i \in N, \forall l \in T, \\
\forall s \in S, \forall x \in X, \delta^{z \circ \pi^{-1}}_{1,i}(l_i)(s, x) = \lambda_i(l_i)[\text{Id}_S, z_i \circ \pi^{-1}_i]^{-1}(s, x).
\]

\[
\forall E \in \Sigma(H^k_X), \delta^{z \circ \pi^{-1}}_{k+1,i}(l_i)[E] = \lambda_i(l_i)[\text{Id}_S, \delta^{z \circ \pi^{-1}}_{k,i}]^{-1}[E].
\]

we show that \( \delta^z(t') = \delta^{z \circ \pi^{-1}}(t) \) by mathematical induction with regard to \( k \). For \( k = 1 \), we have , for all \( i \in N \), and any \( t, t' \in T \) such that \( t = \pi(t') \), the equations below.

\[
\forall s \in S, \forall x \in X, \delta^z_{1,i}(t'_i)(s, x) = \lambda_i(t'_i)[\text{Id}_S, z_{-i}]^{-1}(s, x)
\]

\[
= \lambda_i(t'_i)[\{(s, l'_i) : z_{-i}(l'_i) = x_{-i}\}]
\]

\[
= \lambda_i(\pi_i(t'_i))[\{(s, \pi_{-i}(l'_i)) : z_{-i}(l'_i) = x_{-i}\}]
\]

\[
= \lambda_i(t_i)[\{(s, l_{-i}) : z_{-i} \circ \pi^{-1}_{-i}(l_{-i}) = x_{-i}\}]
\]

\[
= \delta^{z \circ \pi^{-1}}_{1,i}(t_i)(s, x).
\]
For the higher order belief, for all \( i \in \mathbb{N} \),

\[
\forall E \in \Sigma(H^k_X), \quad \delta^z_{k+1,i}(t'_i)[E] = \lambda_i(t'_i)[\text{Id}_S, \delta^z_{k-1,i}]^{-1}[E]
\]

\[
= \lambda_i(t'_i)[\{(s, l'_i) : (s, \delta^z_{k-1,i}(l'_i)) \in E\}]
\]

\[
= \lambda_i(\pi_i(t'_i))[\{(s, \pi_{-i}(l'_i)) : (s, \delta^z_{k-1,i}(l'_i)) \in E\}]
\]

The induction hypothesis is that, for all \( i \in \mathbb{N} \), and any \( t, t' \in T \) such that \( t = \pi(t') \), \( \delta^z_{k,i}(t'_i)(s, x) = \delta^z_{\boxplus}^{-1}(t_i)(s, x) \). Therefore,

\[
\delta^z_{k+1,i}(t'_i)[E] = \lambda_i(\pi_i(t'_i))[\{(s, \pi_{-i}(l'_i)) : (s, \delta^z_{k-1,i}(l'_i)) \in E\}]
\]

\[
= \lambda_i(t_i)[\{(s, l_i) : (s, \delta^z_{\boxplus}^{-1}(l_i)) \in E\}]
\]

\[
= \delta^z_{\boxplus}^{-1}(t_i)(s, x).
\]

As a result, \( \delta^z(t') = \delta^z_{\boxplus}^{-1}(t) \). Thus any image of \( t' \) induced by a signal \( z \) is always attained by its symmetric type \( t \) with a signal \( z \circ \pi^{-1} \).

**Proposition 1.5.14.** For \( t, t' \in T \) in \( \Lambda \), if \( \{\delta^z(t) : z \in Z\} = \{\delta^z(t') : z \in Z\} \), then the smallest sub type spaces of \( \Lambda, \Lambda^t \) and \( \Lambda^{t'} \), can be S-isomorphically embedded to the same space in \( U(S \times C) \) where \( t \) and \( t' \) fall onto the same point.

**Proof.** Suppose that \( \{\delta^z(t) : z \in Z\} = \{\delta^z(t') : z \in Z\} \). Let \( \Phi \) be the smallest sub type space of \( T \) which includes \( t \), and let \( \Phi' \) be the smallest sub type space of \( T \) which includes \( t' \). We can pick the identity function on \( T \) for \( z \). Then \( \delta^z \) becomes the same as S-isomorphism to \( U(S \times T) \), the universal type space on \( S \times T \) in Liu [33]. Let \( M(m) \) be the smallest sub type space on \( U(S \times T) \). Then \( M(\delta^z(t)) \) is S-isomorphic to \( \Phi \). If \( |\Phi| > |\Phi'| \), then \( |\Phi'| \) can never mapped to \( M(\delta^z(t)) \) by any belief mapping. Therefore, we have \( |\Phi'| = |\Phi| \).
By the assumption, there exists $z' : T \to T$ such that $z'(t') = \delta z(t)$. Then we can consider $\Phi'$ as a Harsanyi type space on the payoff parameter $S \times T$. Since $|\Phi| = |\Phi'|$, $\delta z'$ must be bijection from $\Phi'$ to $M(\delta z(t))$. Therefore, $\Phi'$ does not have redundant types concerning $S \times T$. According to Mertens-Zamir, it implies that $\Phi'$ is $S \times T$-isomorphic to $M(\delta z(t))$. Let $\lambda z' \in \Delta(S \times T \times \Phi)$ which induced by $\lambda$ and $z'$, and $\mu_i \in \Delta(S \times T \times M_{-i}(\delta z(t)))$ be a natural belief mapping induced by Kolmogorov extension. Then,

$$\forall i \in N, \forall l_i \in \Phi_i, \quad \lambda z'(l_i) = \mu_i(\delta z'_i(l_i)) \circ [\text{Id}_{S \times T}, \delta z'_i]^{-1}. \tag{1}$$

Therefore,

$$\forall i \in N, \forall l_i \in \Phi_i, \quad \text{Marg}_{(S \times \Phi') \lambda z'}(l_i) = \text{Marg}_{(S \times M_{-i}(\delta z(t))) \mu_i(\delta z'_i(l_i)) \circ [\text{Id}_{S}, \delta z'_i]^{-1}}. \tag{2}$$

By construction, $\text{Marg}_{(S \times \Phi') \lambda z'_i} = \lambda_i$. Therefore, $\Phi'$ is $S$-isomorphic to $M_{-i}(\delta z(t))$. Therefore, $\Phi'$ is $S$-isomorphic to $\Phi$, and $t'$ and $t$ can be mapped to the same point $S$-isomorphically. □

Under a plausible condition, we obtain the next result.

**Lemma 1.5.15.** Let $t, t' \in \Lambda$ in $\Lambda$, and $\Lambda$ is the smallest sub type space that includes $t$. Then, if $\{\delta z(t) : z \in Z\} = \{\delta z(t') : z \in Z\}$, $t$ and $t'$ are symmetric.

**Proof.** By the above proposition, the smallest type space $\Lambda_{t'}$ is $S$-isomorphic to $\Lambda$. Therefore, by Proposition 4.6., they are symmetric. □

We get the next theorem as a corollary.
Theorem 1.5.16. Let \( t, t' \in T \) in \( \Lambda \), and \( \Lambda \) is the smallest sub type space includes \( t \). Then, \( BE(t, \Gamma) = BE(t', \Gamma) \) for any \( \Gamma \) if and only if they are symmetric to each other.

1.5.3 Semantic interpretation of syntactic characterization

Sadzik adopted a first order language which can describe the modal logic of epistemology. He showed that \( \{ \delta^z(t) : z \in Z \} = \{ \delta^z(t') : z \in Z \} \) if and only if the set of sentences which can be true by appropriate values of signals at the types are the same. The results in this section show that this syntactic characterization of types are equivalent to symmetry of types, and whether or not they can be mapped to the same sequential beliefs on \( U(S \times C) \).

1.6 Application to intrinsic correlation

We have shown that any Harsanyi type space can be mapped isomorphically to a sub-space of \( U(S \times C) \). One application of this theorem is the intrinsic correlation of beliefs proposed by Brandenburger and Friedenberg [13]. They showed that, in some complete information games, we cannot achieve all correlated rationalizable actions without any external mediator. They also showed that we can achieve all correlated rationalizable actions as intrinsic ones by adding a coin-flip to the basic uncertainty. In fact, their results are closely related to redundant types. In this section, we show the results of Brandenburger-Friedenberg in a different way; using redundant types and our theorems above.

1.6.1 Bayesian representation of correlated equilibrium

Consider a complete information game. Let \( G \equiv \langle (A_i)_{i \in N}, (\pi_i)_{i \in N} \rangle \) be a game, where \( A_i \) is the strategy space of the agent \( i \) and \( \pi_i : A \rightarrow \mathbb{R}_+ \) is a payoff function. We assume that, for all \( i \in N \), \( A_i \) is finite.\(^{13}\) To define the correlated equilibrium of the game \( G \), we introduce the Bayesian framework \( a la \) Aumann. Let the basic uncertainty space be \( \Omega \), the information

\(^{13}\)This is the same assumption as Brandenburger-Friedenberg.
partition of the agent be \( \mathcal{H}_i \), and the interim belief systems be \( P(.|\mathcal{H}_i) \in \Delta(\Omega) \). Since \( A \) is finite, \( \Omega \) can be chosen to be finite in order to represent correlated equilibria.\(^{14}\)

**Definition 1.6.1.** (Aumann [2]): For all \( i \in N \), let \( f_i : \Omega \to A_i \) be measurable with regard to \( \mathcal{H}_i \). Then \( f \equiv (f_i)_{i \in N} \) is an a posteriori equilibrium iff

\[
\forall i \in N, \forall \omega \in \Omega, \forall a_i \in A_i, \quad \sum_{\omega \in \Omega} \pi_i(f_i(\omega), f_{-i}(\omega)) \cdot P(\omega|\mathcal{H}_i(\omega)) \geq \sum_{\omega \in \Omega} \pi_i(a_i, f_{-i}(\omega)) \cdot P(\omega|\mathcal{H}_i(\omega)).
\]

**Definition 1.6.2.** (Bernheim [9], Pearce [43]): A set of strategies \( R^\infty \subset \Pi_{i \in N} A_i \) is the set of the correlated rationalizable actions if (1) for each \( i \in N \) and each \( a_i \in R_i^\infty \), there exists \( \mu \in \Delta(R_i^\infty) \) such that \( a_i \) is a best response to \( \mu \), and (2) there is no set \( F \subset \Pi_{i \in N} A_i \) such that it satisfies (1) and \( F \supseteq R^\infty \).

Concerning a posteriori equilibria and correlated rationalizable actions, we have the next equivalence result.

**Proposition 1.6.3.** (Epstein [22]\(^{15}\)) For any \( a^* \in R^\infty \), there exists a posteriori equilibrium \( \langle A, (\mathcal{H}_i)_{i \in N}, (P(.|\mathcal{H}_i))_{i \in N}, f \rangle \) such that, for all \( i \in N \), \( \mathcal{H}_i = A_i \), for all \( a \in A \), \( f(a) = a \), and \( f(a^*) = a^* \).

From \( \langle A, (\mathcal{H}_i)_{i \in N}, (P(.|\mathcal{H}_i))_{i \in N}, f \rangle \), where \( \mathcal{H}_i = A_i \), we can construct a Harsanyi type space

\(^{14}\)As in the following argument, we can set \( \Omega \) to be \( A \).

\(^{15}\)Aumann [3] and Brandenburger and Dekel [10] showed the same result.
on $A$. For all $i \in N$, let $T_i \equiv \mathcal{R}_i$. and $\lambda_i : T_i \to \Delta(A \times T)$ be as follows;

$$
\lambda_i(a_i)[(a_{-i}, a_{-i})] = \begin{cases} 
P(a_{-i}|a_i) & \text{if } a_{-i} = f_{-i}(a_{-i}) \\ 0 & \text{otherwise.} \end{cases}
$$

Let $\Lambda \equiv \langle A, T, \lambda \rangle$. We can easily confirm that $\Lambda$ is a Harsanyi type space on $A$.

Let $G' \equiv \langle \pi, \Lambda \rangle$ be a Bayesian game. For $i \in N$, let a strategy $\beta_i : T_i \to A_i$ be such that $\beta_i(t_i) = f_i(\omega)$ where $\omega \in t_i$. Then $\beta \equiv (\beta)_{i \in N}$ becomes a Bayesian Nash equilibrium of the game $G'$. The a posteriori equilibrium of the original game $G$ is a Bayesian Nash equilibrium of $G'$.

### 1.6.2 Conditional independence and rationality and common certainty of rationality

Brandenburger-Friedenberg characterized intrinsic correlation by two conditions on Harsanyi types.

**Definition 1.6.4.** A Harsanyi type $t_i \in T_i$ satisfies conditional independence if $\lambda_i(t_i)[a_{-i}|h(t_{-i})] = \Pi_{j \in N \setminus \{i\}} \lambda_i(t_i)[a_j|h(t_{-i})]$, where $h$ is the hierarchy mapping from $T \to U(A)$.

For the definition of another condition, *rationality and common certainty of rationality*, we need some preliminary definitions.

**Definition 1.6.5.** For each $i \in N$, a pair $(a_i, t_i) \in A_i \times T_i$ satisfies rationality if $a_i$ is a best response to $\text{Marg}(A_{-i})\lambda_i(t_i)$. 
We use $R_i$ to denote the set of the pairs that satisfies rationality.

**Definition 1.6.6.** For any $E \subset A_\ast - i \times T_\ast - i$, $t_i \in K_i(E)$ if $\lambda_i(t_i)[E] = 1$.

**Definition 1.6.7.** For each $i \in N$, $t_i \in T_i$ satisfies rationality and common certainty of rationality if $t_i \in R_i \cap \bigcap_{k=1}^\infty K^k(R)$, where $K^k$ is the $k$th iteration of the operator $K$.

Since $\beta$ is a BNE, it is almost clear that, for all $t_i \in T_i$, $(t_i, \beta_i(t_i))_{i \in N}$ satisfies RCBR. Polak [44] showed that RCBR is not sufficient condition for Nash equilibrium as shown in Aumann-Brandenburger [4], but Nash equilibrium satisfies RCBR under complete information about the payoffs. And he showed that the same thing applies to BNE. Here is a brief sketch of the proof. By construction, it is clear that, for all $i \in N$, $(t_i, \beta_i(t_i)) \in R^1_i$. Suppose that, for all $i \in N$ and $t_i \in T_i$, $(t_i, \beta_i(t_i)) \in R^k_i$. Then, since $\lambda_i(t_i)[\{(t_{\ast - i}, a_{\ast - i}) : a_{\ast - i} = \beta_{\ast - i}(t_{\ast - i})\}] = 1$ and $\{(t_{\ast - i}, a_{\ast - i}) : a_{\ast - i} = \beta_{\ast - i}(t_{\ast - i})\} \subset R^k_i$, we have $t_i \in B(R^k_{\ast - i})$. By the induction hypothesis, $(t_i, \beta_i(t_i)) \in R^k_i \cap [A_i \times B(R^k_{\ast - i})]$. Thus, $(t_i, \beta_i(t_i)) \in R^\infty_i$.

### 1.6.3 Conditional independence and redundancy

Note that conditional independence defined above is conditional on the sequential beliefs of the other agents’ types. Therefore, when there are redundant types in $\Lambda$, it is hard for CI to be satisfied. However, the results that we have shown allows us to get rid of redundant types without affecting resulting equilibria.

In this section, we show first that, for any $a^* \in R^\infty$, there exists a Harsanyi type space $\Phi$ such that $a^*$ is a realization of a BNE, and $\Phi$ has no purely redundant types except for one agent. As a result, we get the result that, for any $a^* \in R^\infty$, there exists a Bayesian
formulation where $a^*$ satisfies RCBR at a type which satisfies CI.

**Proposition 1.6.8.** For any $a^* \in R^\infty$, there exists a posterior equilibrium such that, for some $\omega \in \Omega$, $f(\omega) = a^*$, and, for all $i \neq 1$, if $H_i \neq H'_i$, $P([a_j]_{j \neq i}|H_i) \neq P([a_j]_{j \neq i}|H'_i)$ for some $a_{-i}$.

**Proof.** By the proposition above, there exists a posterior equilibrium such that $\Omega = A_i$ for all $i \in N$, $\mathcal{A}_i = \{a_i \times A_{-i} : a_i \in A_i\}$ and $f_i(a) = a_i$. Let this a posteriori equilibrium be $F$ and $[a_i] \equiv a_i \times A_{-i}$. For notational convenience, we denote each class in the agent $i$'s information partition as $[a_i]$. Now it is possible that there exists $[a_i] \neq [a'_i]$ such that, for all $H_{-i}$, $P(H_{-i}|[a_i]) = P(H_{-i}|[a'_i])$. Then we can duplicate the agent 1’s information partition.

Suppose that, for $b_1 \in A_1$, $P([b_1]| [a_i]) > 0$. We add another set of states so that the states of the world $\hat{\Omega} = (A_1 \cup \{a_2^i\}) \times A_{-1}$ and associate another information partition $[a_2^i]$ to the agent 1. We define a new a posterior equilibrium $\hat{F} \equiv \langle A, \hat{\Omega}, \hat{P}, \hat{f} \rangle$ such that

For $j = 1$, \[ \hat{f}_1([a_1^2]) = b_1 \]
\[ \forall a_1 \in A_1, \hat{f}_1([a_1]) = f_1([a_1]). \]
\[ \forall j \neq 1, \forall a_j \in A_j, \hat{f}_j([a_j]) = f_j([a_j]). \]
The new sequence of conditional beliefs is defined in the following way;

For $j = 1,$
\[
\hat{P}(\cdot|a_1^2) = P(\cdot|a_1) \\
\hat{P}(\cdot|a_1) = P(\cdot|a_1) \text{ otherwise.}
\]

\[\forall j \neq 1, \forall a_j \neq [a_j'], \]
\[
\hat{P}(\cdot|a_j) = P(\cdot|a_j) \\
\hat{P}(\cdot|a_j) = P(\cdot|a_j) \text{ otherwise.}
\]

For $j = i$ and $a_i', \forall a_{-1,i} \in A_{-1,i},$
\[
\hat{P}((a_1^2, a_{-1,i})|[a_i']) = P((b_1, a_{-1,i})|[a_i']) \\
\hat{P}(\cdot|[a_i']) = P(\cdot|[a_i']) \text{ otherwise.}
\]

It is easy to show that $\hat{\Omega} \equiv \langle A, \hat{\Omega}, \hat{P}, \hat{f} \rangle$ is an a posteriori equilibrium, and $\hat{f}(a^*) = a^*.$ Note that, in this a posterior equilibrium, $\hat{P}(\cdot|a_i)$ and $\hat{P}(\cdot|[a_i'])$ are distinguishable at the event $[b_i].$

Since $N$ and $A$ are finite, we can iterating this process until every pair of each agent’s, except for the agent 1, information states $[a_j] \neq [a_j']$ have different conditional beliefs over the other players’ information states. \qed

**Corollary 1.6.9.** For any $a^* \in R^\infty$, there exists a Harsanyi type space $\Lambda = \langle A, T, \lambda \rangle$ and a pair of Bayesian equilibrium strategy $\beta = (\beta_i)_{i \in N}$ such that $a^* = \beta(t)$, and, for all $i \neq 1$, $T_i$ has no purely redundant types.

Then we can apply the theorem to find an S-isomorphic Harsanyi type space $\Phi$ on $A \times \{0, 1\}$ which has no redundant types. And, in $\Phi,$ no types result in the same sequential belief. Therefore, each type and its action associated by the equilibrium strategy $\beta$ satisfy CI. Therefore we have the next result, which is the same result shown in a different way by Brandenburger-Friedenberg.
**Theorem 1.6.10.** For any $a^* \in R^\infty$, there exists a Harsanyi type space $\Phi = \langle A \times \{0,1\}, V, \phi \rangle$ such that $a^*$ satisfies RCBR at some state $v \in V$ which satisfies CI.

### 1.7 Conclusion

In this paper, we showed that it is possible to embed Harsanyi type spaces isomorphically onto the space of sequential beliefs over an augmented uncertainty, even if they have redundant types. The technique to introduce a payoff irrelevant parameter is an extension of Liu. However we have the following distinctions: (1) our payoff irrelevant parameter space is exogenous, and (2) it is enough that the parameter space has only two values. That is, any correlation of types in Bayesian frameworks which cannot be explained by the basic uncertainty is resolved by adding a coin flip to the uncertainty. Concerning the first finding, the exogeneity of the parameter allowed us to show the existence of the universal type space where the vast majority of Harsanyi type spaces are uniquely embedded.

We showed that our results can be applied to provide a characterization of Bayesian Equilibrium and an interpretation of intrinsic correlation in games. Because we presented a universal type space that includes redundant type spaces, the recent research on strategic topologies on Mertens-Zamir’s universal type space can potentially be extended to our space. Moreover, we can use Bayesian Equilibrium as the solution concept, differently from Dekel et al who used ICR and Ely-Peski who used IIR.
1.8 Appendix

1.8.1 Bimeasurability of the function $\phi$

Let $\Phi = \langle S \times C, V_1 \times V_2, (\phi_i)_{i \in \{1, 2\}} \rangle$ be a Harsanyi type space, $V_1 = V_2 = [0, 1]$, and $C = \{0, 1\}$ as defined in the section 3. First we show that $\phi_1 : V_1 \to \Delta(S \times V_2 \times C)$ is bimeasurable. It is worth while to notice that $\phi_1$ maps each element in $V_1$ to a product measure on the measurable space $(S \times V_2 \times C, \Sigma(S \times V_2 \times C))$.

We define the following functions.

$$f_1 : V_1 \to \Delta(S \times V_2) \text{ such that } f_1(v_1) = \lambda_1(p_1^{-1}) \circ [Id_S, p_2]^{-1}.$$  

$$g_1 : V_1 \to \Delta(C) \text{ such that } g_1(v_1)(0) = v_1.$$  

You can see that both $f_1$ and $g_1$ are bimeasurable functions. Then we have that $\phi_1(v_1) = f_1(v_1) \times g_1(v_1)$, where $f_1(v_1) \times g_1(v_1)$ is the product measure on the Borel measure space $(S \times V_2 \times C, \Sigma(S \times V_2 \times C))$. Since $S \times V_2$ and $C$ are both second countable, $\Sigma(S \times V_2) \times \Sigma(C) = \Sigma(S \times V_2 \times C)$.

Lemma 1.8.1. The Borel $\sigma$-algebra $\Sigma(S \times V_2 \times C) = \{E : \exists A \in \Sigma(S \times V_2), \exists B \in \Sigma(C), E = A \times B\}$.

Proof. Let $\hat{\Sigma} \equiv \{E : \exists A \in \Sigma(S \times V_2), \exists B \in \Sigma(C), E = A \times B\}$. We only have to show that $\hat{\Sigma}$ is a $\sigma$-algebra. It is clear that $\emptyset, S \times V_2 \times C \in \hat{\Sigma}$. Let $E \in \hat{\Sigma}$. Then there exists $A \in \Sigma(S \times V_2)$ and $B \in \Sigma(C)$ such that $E = A \times B$. Therefore $E^c = A^c \times C \cup A \times B^c$.

\footnote{By Caratheodory’s extension theorem, the product measure is uniquely determined.}
Let $\Delta^P(S \times V_2 \times C) \subset \Delta(S \times V_2 \times C)$ be the set of the product measures over $S \times V_2$ and $C$.

**Lemma 1.8.2.** The subspace $\Delta^P(S \times V_2 \times C)$ is homeomorphic to the product space $\Delta(S \times V_2) \times \Delta(C)$.

**Proof.** By Caratheodory’s extension theorem, the function $d : \Delta(S \times V_2) \times \Delta(C) \rightarrow \Delta^P(S \times V_2 \times C)$ such that $(\eta, \mu) \mapsto \eta \times \mu$ is bijection.

First we want to show that $d$ is a continuous function. The topological base of $S \times V_2 \times C$ is $t = \{G \times a : G \text{ is an open subset of } S \times V_2, \text{ and } a \in 2^C\}$. Therefore any open set $G' \subset S \times V_2 \times C$ takes the form of

$$G' = \tilde{G}_1 \times \{0\} \cup \tilde{G}_2 \times \{1\} \cup \tilde{G}_3 \times \{0, 1\},$$

where, for $i = 1, 2, 3$, $\tilde{G}_i$ is an open set in $S \times V_2$. It is reduced to

$$G' = G_1 \times \{0\} \cup G_2 \times \{1\},$$

where, for $i = 1, 2$, $G_i$ is an open set in $S \times V_2$.

Let $\{\eta_\alpha\}$ be a net in $\Delta(S \times V_2)$ such that $\eta_\alpha \rightarrow \eta$. And let $\{\mu_\alpha\}$ be a net in $\Delta(C)$ such that $\mu_\alpha \rightarrow \mu$. Then,

$$\forall G : \text{ open, } \liminf \eta_\alpha(G) \geq \eta(G),$$

$$\forall a \in 2^C, \liminf \mu_\alpha(a) \geq \mu(a).$$
Let \( \nu_\alpha \equiv \eta_\alpha \times \mu_\alpha \), and \( \nu = \eta \times \mu \). Then, for each open set \( G' \subset S \times V_2 \times C \),

\[
\nu_\alpha(G') = \nu_\alpha(G_1 \times \{0\}) + \nu_\alpha(G_2 \times \{1\}) \\
= \eta_\alpha(G_1)\mu_\alpha(\{0\}) + \eta_\alpha(G_2)\mu_\alpha(\{1\}).
\]

In the same way,

\[
\nu(G') = \eta(G_1)\mu(\{0\}) + \eta(G_2)\mu(\{1\}).
\]

Since \( \eta_\alpha \to \eta \) and \( \mu_\alpha \to \mu \),

\[
\liminf \eta_\alpha(G_1)\mu_\alpha(\{0\}) \geq \eta(G_1)\mu(\{0\}).
\]

\[
\liminf \eta_\alpha(G_2)\mu_\alpha(\{1\}) \geq \eta(G_2)\mu(\{1\}).
\]

And,

\[
\liminf \nu_\alpha(G') = \liminf \{ \eta_\alpha(G_1)\mu_\alpha(\{0\}) + \eta_\alpha(G_2)\mu_\alpha(\{1\}) \} \\
\geq \liminf \eta_\alpha(G_1)\mu_\alpha(\{0\}) + \liminf \eta_\alpha(G_2)\mu_\alpha(\{1\}) \\
\geq \eta(G_1)\mu(\{0\}) + \eta(G_2)\mu(\{1\}) \\
= \nu(G').
\]

Therefore, \( \nu_\alpha \to \nu \). Therefore \( d \) is a continuous function.

Next, we show that \( d^{-1} \) is a continuous function. Let \( \{ \nu_\alpha \} \equiv \{ \eta_\alpha \times \mu_\alpha \} \) be a net of product measures such that \( \nu_\alpha \to \nu = \eta \times \mu \). Then, \( \nu_\alpha(S \times V_2 \times a) = \mu_\alpha(a) \), and \( \nu(S \times V_2 \times a) = \mu(a) \).

Since \( \liminf \nu_\alpha(S \times V_2 \times a) \geq \nu(S \times V_2 \times a) \), \( \liminf \mu_\alpha(a) \geq \mu(a) \). In the symmetric
way, \( \liminf \eta_\alpha(G) \geq \eta(G) \). It means that \((\eta_\alpha, \mu_\alpha) \to (\eta, \mu)\). Therefore \(d^{-1}\) is a continuous function. \(\square\)

**Corollary 1.8.3.** The subspace \(\Delta^F(S \times V_2 \times C)\) is closed.

Since \(\Delta(S \times V_2) \times \Delta(C)\) is second countable, \(\Sigma(\Delta(S \times V_2) \times \Delta(C)) = \Sigma(\Delta(S \times V_2)) \times \Sigma(\Delta(C))\)

**Proposition 1.8.4.** The function \(\phi_1 : V_1 \to \Delta(S \times V_2) \times \Delta(C)\) is a bimeasurable function.

*Proof. (Inverse measurability)* Let \(\phi_1 = (f_1, g_1)\). The space of the probability measures \(\Delta(C)\) is homeomorphic to \(V_1\), and \(g_1\) is its homeomorphism. We consider that \(\phi_1 : V_1 \to \Delta(S \times V_2) \times V_1\) and \(\phi_1 = (f_1, Id)\). It allows us to consider that \(\phi_1(V_1) \subset \Delta(S \times V_2) \times V_1\) is a graph of the function \(f_1^{-1}\). Since \(f_1^{-1}\) is a measurable function, the graph \(\phi_1(V_1)\) is a Borel set in the product measure space \(\Delta(S \times V_2) \times V_1\). For each \(E \in \Sigma(V_1)\), \(\phi_1(E) = f_1(E) \times E \cap \phi(V_1)\). We know that both \(f_1(E) \times E\) and \(\phi(V_1)\) are measurable. Therefore, \(\phi_1(E)\) is measurable.

*Measurability* Let \(E \subset \Delta(S \times V_2) \times V_1\) be a rectangle. Let \(\pi_1\) and \(\pi_2\) be the projection onto \(V_1\) and \(\Delta(S \times V_2)\) respectively. Let \(F_2 \equiv f_1 \circ \pi_1(E) \subset \Delta(S \times V_2)\). Since \(f_1\) is bimeasurable, \(F_2\) is also Borel. For each \(y \in \pi_2(E)\), \(f_1^{-1}(y) \in \pi_1(E)\) if and only if \(y \in \pi_2(E) \cap F_2\). Let Therefore, the intersection of the rectangle \(E\) and the entire graph \(\phi_1(V_1) \equiv \{(x, f_1(x)) : x \in V_1\}\) becomes \(G \equiv \{(f_1^{-1}(y), y) : y \in \pi_2(E) \cap F_2\}\). Since \(\pi_2(E)\) and \(F_2\) are both Borel, \(\pi_2(E) \cap F_2\) is also Borel. We can see that \(\phi_1^{-1}(E) = \pi_1(G)\). Since \(f_1\) is measurable, \(\pi_1(G)\) is also Borel.

\(^{17}\)See Halmos [26] pp143.
Therefore $\phi_1^{-1}(E)$ is Borel.

\[ \phi_1^{-1}(E) \text{ is Borel.} \]

**Proposition 1.8.5.** The function $\phi_2 : V_2 \to \Delta(S \times V_1 \times \Delta(C))$ is a bimeasurable function.

**Proof.** Let

\[ \Delta_0 \equiv \{ \mu \in \Delta(S \times V_1 \times C) : \forall E \in \Sigma(S \times V_1), \mu(E \times \{1\}) = 0 \}. \]

Let $f : S \times V_1 \times C \to \mathbb{R}$ such that

\[
\forall e \in S \times V_1, f(e, 0) = a, \\
    f(e, 1) = b.
\]

Then, $f \in C_b(S \times V_1 \times C)$. Therefore, when a net $\{\mu_\alpha\}$ converges to some probability measure, it must be in $\Delta_0$. Therefore, $\Delta_0$ is a closed set. Since $\lambda_2$ is bimeasurable between $V_2$ and $\Delta(S \times V_1)$ and $\Delta(S \times V_1)$ is homeomorphic to $\Delta_0$, $\phi_2$ is bimeasurable between $V_2$ and $\Delta_0$.

I use the term “bimeasurable” in a slightly different way.

**Definition:** A function $f : X \to Y$ is bimeasurable if $f$ is measurable and, for each measurable set $E \subset X$, $f(E)$ is also measurable.
Lemma 1.8.6. Let $X$ and $Y$ be Polish, and $f_1 : X \to \Delta(X)$ and $g_1 : X \to Y$ be both bimeasurable. Let $f_2 : X \to \Delta(Y)$ be such that, for each $x \in X$, $f_2(x) = f_1(x) \circ g_1^{-1}$. Then, the function $f_2$ is bimeasurable.

Proof. The measurability of $f_2$ is shown by Liu. We only show that, for all $E \in \Sigma(X)$, $f_2(E) \in \Sigma(\Delta(Y))$.

Let $g_2 : \Delta(X) \to \Delta(Y)$ such that, for all $\mu \in \Delta(X)$, $g_2 : \mu \mapsto \mu \circ g_1^{-1}$. Let $A \equiv \{\mu \in \Delta(X) : \mu(E) \geq p\}$, where $E \in \Sigma(X)$ and $p \in [0,1]$. Then, $g_2(A) = \{\nu \in \Delta(Y) : \nu(g_1(E)) \geq p\} \cap \{\nu \in \Delta(Y) : \nu(g_1(X)) = 1\}$. Notice that $g_1(X) \in \Sigma(Y)$. The both sets in the right hand side of the equation are measurable. Therefore $g_2(A) \in \Sigma(Y)$.

Since $f_2 = g_2 \circ f_1$, we have that, for each $E \in \Sigma(X)$, $f_2(E) \in \Sigma(\Delta(Y))$. \qed

Lemma 1.8.7. For each $k \geq 1$, the $k$th order hierarchy mapping $h_k^i$ is bimeasurable.

Proof. Without loss of generality, we only have to show that $h_1^1$ is bimeasurable.

For $k = 1$, let $X \equiv S \times V \times C$, $Y \equiv S \times C$, $f_1 \equiv \tilde{\phi}_i$, and $g_1 \equiv \text{proj}_{(S \times C)}$, where, for all $(s,c,v_1,v_2) \in S \times V \times C$, $\tilde{\phi}_i(s,c,v_1,v_2) \equiv \phi_1(v_1)$. It is easy to see that $f_1$ and $g_1$ are bimeasurable. By the lemma, $h_1^1 : S \times V \times C \to \Delta(S \times C)$ is bimeasurable. We can just restrict the domain from $S \times V \times C$ to $V_1$ to get the $h_1^1$ which is measurable.

\footnote{See Lemma 5 in Liu [33]}
For \( k \geq 2 \), we assume that, for \( i = 1, 2 \), \( h_i^{k-1} \) is bimeasurable as the induction hypothesis. Let \( X \equiv S \times V \times C \), \( Y \equiv S \times C \times H^k(S \times C) \), \( f_1 \equiv \tilde{\phi}_1 \), and \( g_1 \equiv \tilde{h}_2^{k-1} \), where 
\[
\tilde{h}_2^{k-1}(s, c, v_1, v_2) \equiv h_2^{k-1}(v_2). \]
By the lemma, \( \tilde{h}_1^k : S \times V \times C \rightarrow \Delta(S \times C) \) is bimeasurable. We can just restrict the domain from \( S \times V \times C \) to \( V \) to get the \( h_1^k \) which is measurable. □

**Proposition 1.8.8.** The full hierarchy mapping \( h_i \) is bimeasurable.

**Proof.** First, we show that \( h_i \) is measurable. The \( \sigma \)-algebra on \( \prod_{k=1}^{\infty} H^k \) is the \( \sigma \)-algebra generated by 
\[
\mathcal{F} \equiv \{ F = \Pi_{k \in I} E_k \times \Pi_{k \in I} H^k : I \subset \mathbb{N} \text{ is finite, and } E_k \in \Sigma(H^k) \}. 
\]
Since \( h_i \equiv (h_1^i, \ldots, ) \), for each \( F \in \mathcal{F} \), \( h_i^{-1}(F) \in \Sigma(V_i) \). By Theorem 4-1-6 in Dudley [19], \( h_i \) is measurable.

Next, we show that, for each \( E \in \Sigma(V_i) \), \( h_i(E) \) is measurable. Let \( E \in \Sigma(V_i) \). Since \( h_1^i \) and \( h_2^i \) are bimeasurable injection, \( h_2^i \circ (h_1^i)^{-1} \) is bimeasurable bijection from \( h_1^i(E) \) to \( h_2^i(E) \). Let the image of \( E \) by \( (h_1^i, h_2^i) \) be \( \Gamma_2(E) \). It means \( \Gamma_2(E) \equiv \{ (h_1^i(v_i), h_2^i(v_i)) \in H^1 \times H^2 : v_i \in E \} \). We can see that it is the graph of \( h_2^i \circ (h_1^i)^{-1} \). Therefore, \( \Gamma_2(E) \) is measurable in the product measurable space \( H^1 \times H^2 \). By the mathematical induction, for each \( k \geq 1 \), \( \Gamma_k(E) \subset \prod_{l=1}^{k} H^l \), the image of \( E \) by \( (h_1^1, \ldots, h_k^k) \), is measurable. The image of the full hierarchy \( h_i(E) \) is the projective limit of \( (\Gamma_k(E))_{k \in \mathbb{N}} \), and as we saw, each \( \Gamma_k(E) \) is measurable. Therefore \( h_i(E) \) is measurable. □
Chapter 2

Knowledge-Belief Space Approach to Robust Implementation

2.1 Introduction

Mechanism design is the field of economics that studies the following question: when agents have private information, is it possible to design a game (mechanism) that has an equilibrium implementing a desirable allocation, i.e., a social choice function (SCF)? The literature is vast: it started with Arrow’s work in social choice theory in the 1950’s and it remains an active field of current research. However there are at least two practical issues. First, there may be other equilibria of the designed game that do not implement the SCF. Second, it is required that the designer knows the beliefs of the agents so that she can compute the set of Bayesian (Nash) equilibria of the induced game.

The first problem is addressed with what is known as “full implementation”: every equilibrium of the designed game has to implement the desirable allocation. As such, it worsens the second problem. The designer must know the belief structure, which is known as the type space, so as to compute every Bayesian equilibrium of the game. This is a formidable task.
One way to address this problem is to consider “belief-free” mechanism design.

In this paper, we study belief-free mechanism design, which was first defined by Bergemann and Morris, robust implementation. Robust implementation is to find a game that implements a SCF as the unique Bayesian equilibrium outcome no matter what the agents believe. In other words, a game robustly implements a SCF if, on every possible type space of the agents, the SCF is obtained as the unique Bayesian equilibrium outcome of the resulting Bayesian game. If there is such a robust game, the planner does not have to worry about the agents’ type spaces. However, according to the definition of robust implementation, the planner has to check all the possible type spaces of the agents in order to see if a game robustly implements a SCF. It is also a formidable task. The agents can form not only any beliefs about the other agents’ private information, but also any beliefs about the other agents’ beliefs and so on. That is, in principle, infinitely many type spaces are possible, so it is a difficult task to check if a game can work well on all the type spaces.

The main result of this paper is that there exists a type space such that a game robustly implements a SCF if and only if it implements the SCF as the unique Bayesian equilibrium outcome on that type space. It implies that the planner only has to check implementation in that one particular type space. It makes arguments both conceptually and technically simpler, and enables us to study robust implementation in more general environments than the existing literature.

The intuition behind this result is Bayesian implementation on the universal type space. The universal type space is the space of all possible hierarchies of beliefs, first constructed by Mertens and Zamir [39]. It is the type space such that any type space can be considered as one of its sub spaces. Intuitively, the universal type space is the largest possible belief structure of agents and any possible belief structure can be realized inside it. This univer-
sality tempts us to think in the following way: If a game implements a SCF as the unique Bayesian equilibrium outcome on the universal type space, then it also implements the SCF as the unique Bayesian equilibrium outcome on every possible type space because they are sub spaces of the universal space. However this conjecture is not true due to two problems: the failure of Extension property and the existence of redundant types. We obtain our result by resolving these two problems and making the above intuition valid.

The failure of Extension property means that, given a game, the set of Bayesian equilibria on the universal type space is not the same as those on its sub type spaces, and this failure prevents us from studying Bayesian implementation on the universal type space. It is worthwhile to emphasize that, according to the definition c.f. Myerson [40], a Bayesian game consists of a game and a type space where the game is defined. Therefore a game on the universal type space is a different Bayesian game from the same game on its sub type space. However one may think that the Bayesian equilibrium strategies taken on the sub type space are the same in both Bayesian games because the sub type space is included in both Bayesian games.

However this conjecture is not necessarily true. Friedenberg and Meier [23] gave an example that, for some equilibrium strategies of the Bayesian game on a sub type space, there is no equilibrium of the Bayesian game on the universal type space which coincides with the equilibrium strategies of the Bayesian game on the sub type space. When Extension property is not satisfied, Bayesian implementation on the universal type space is not robust implementation. Even if a game implements a SCF as the unique Bayesian equilibrium on the universal type space, there may exist other Bayesian equilibrium outcomes of the Bayesian game on a sub type space and there is no guarantee that those equilibrium outcomes coincide with the SCF.
Another problem preventing us from taking this “universal type space approach” is the existence of redundant types. Although we referred to the universal type space as the type space where every type space can be considered as its sub space, there are some exceptions, as suggested by Ely and Peski [20]. They showed that some type spaces have elements which are “degenerated” when they are mapped to the universal type space. We say that those type spaces are *type spaces with redundancy*. This argument means that Bayesian implementation on the universal type space does not implies Bayesian implementation on the type space with redundancy. However, we can apply the result in Yokotani [50] to deal with this problem. Yokotani showed that we can construct the universal type space where there is no problem of redundancy by introducing a payoff irrelevant parameter space on the lines of Liu [33]. Liu suggested that redundancy occurs due to some source of uncertainty missing in the model and can be resolved by introducing a payoff irrelevant parameter space as the missing source of uncertainty.

Having dealt with redundancies, it remains to deal with the failure of Extension property. We solve this problem by using a *knowledge-belief space* (KB space), which was first defined by Aumann [5]. A KB space is a type space with a knowledge operator. Using KB spaces means that we explicitly deal with what the agents know about their environments. Friedenberg-Meier suggested that the failure of Extension property occurs because the agents do not know *the context of the game*. We refer to the context of the game as the action space, the payoff function, and the type space where the game is defined. Note that, in any Bayesian game analysis, the context of the game is assumed to be a common knowledge among the agents. So, in order to study Bayesian equilibria of different type spaces, we have to clarify what the agents know at the type spaces. One way to deal with it is to adopt KB spaces instead of type spaces. We show that Extension property is always satisfied on KB spaces and there exists *the universal KB space*. As a consequence, we can show that there exists a universal KB space which was augmented by a payoff irrelevant parameter space, and Bayesian implementation
on that universal space implies Bayesian implementation on every possible type space, i.e. it is robust implementation.

The fact that robust implementation is Bayesian implementation on a particular space enables us to study robust implementation in a more general environments. One application is the robust implementation of social choice correspondences (SCC). By now, we have just assumed that the planner’s target is uniquely pinned down, i.e. a SCF. But, in many economic applications, this requirement is too strong. It often happens that there is no game that implements the SCF. Therefore it is more realistic that the planner has a target range of allocations, i.e. a SCC, instead of a SCF. Since our main result is also true for SCCs, we can directly apply the existing results about Bayesian implementation, such as Jackson [29], to obtain results about robust implementation of SCCs. However, robust implementation of SCCs can not be handled on the lines of the existing literature such as Bergemann and Morris [7]. Bergemann-Morris obtains robust implementation through another implementation concept, rationalizable implementation, which uses rationalizability for the solution concept instead of Bayesian equilibrium. In order for rationalizable implementation to be equivalent to Bayesian implementation, it is required that the rationalizable outcome be unique, i.e., a single valued function. In the case of a SCC, the equivalence breaks down, and their method cannot be applied.

The second contribution of our result is to give foundations for robust implementation when agents know some features about the environment. In economic applications, it is often the case that some agents are informed about fundamentals such as economic statistics, project profitabilities, and so on. Although there are some studies considering what agents know in the Bayesian implementation literature\(^1\), it was difficult to consider robustness regarding to knowledge. However our result makes it possible to take the existing approach on the

---

\(^1\)For example, Palfrey and Srivastava [42]
universal KB space to handle robust implementation when agents have some pre-specified knowledge about the environments.

The rest of this paper is organized as follows. We survey the existing results of robust implementation, the universal type space, and Extension property in Sections 2.2, 2.3 and 2.4. We show the main result of this paper in Section 2.5, and give a characterization result as its application in Section 2.6.

2.2 Robust implementation

2.2.1 Bayesian implementation

Let $N$ be the set of agents and $X$ be the space of alternatives. Let $S_i$ be the (ex post) payoff parameter space of the agent $i$. We define $S \equiv \Pi_{i \in N} S_i$. The optimal or targeted alternatives for the planner is represented as a social choice correspondence (SCC) $F : S \rightharpoonup X$. If $F$ is function, we call it a social choice function (SCF). The agents’ (ex post) utility depends on alternatives and agents’ payoff types: $u_i : X \times S \rightarrow \mathbb{R}_+$. Note that we allow interdependent types. Then, at the interim stage, the agents form conjectures over $S$, the other agents’ conjecture over $S$, and so on. Such hierarchies of beliefs about beliefs and some other uncertainty make the interim states of the world, i.e. the interim type space or the Harsanyi type space. The definition is as follows:

**Definition 2.2.1.** A space $T \equiv \langle T, (\lambda_i)_{i \in N}, (\pi_i)_{i \in N} \rangle$ is a Harsanyi type space if

\[
T \equiv \Pi_{i \in N} T_i,
\]

\[
\lambda_i : T_i \rightarrow \Delta(T_{-i}),
\]

\[
\pi_i : T_i \rightarrow S_i.
\]
Here $\pi_i$ means that the agent $i$ at the type $t_i$ knows his payoff type $\pi_i(t_i)$, and $\lambda_i$ represents the agent $i$’s conjecture over the other agents’ types. The planner has to think about how to implement a SCC $F$ at the interim stage. Palfrey-Srivastava [42] and Jackson [29] studied implementation at the interim stage, which is known as Bayesian implementation.

At the interim stage, we have to think about not only payoff types but also epistemic types. So SCC must be slightly modified as in Palfrey-Schmeidler [45]. The planner’s target is, instead of SCC, a social choice set (hereafter SCS). A SCS is given by a set of functions $F \subset \mathcal{F} \equiv \{q \mid q : T \rightarrow X\}$. Bayesian implementability is about whether we can achieve $F$ as an equilibrium with some mechanism. Formally, mechanism is a tuple $\mathcal{M} \equiv \langle M, g \rangle$, where $M \equiv \Pi_{i \in N} M_i$ is the message space and $g : M \rightarrow X$ is the payoff function. The pair of a mechanism and a Harsanyi type space constitutes a Bayesian game form $\Gamma(\mathcal{M}, T)$. Here we assume that each agent $i$ has vNM preference $R_i$, that is

$$\forall q, r \in \mathcal{F}, \quad q R_i(t_i) r \quad \text{if and only if} \quad \int_T u_i(q(t)) d\lambda_i(t_i) \geq \int_T u_i(r(t)) d\lambda_i(t_i).$$

The equilibrium concept is as usual. Let $G = \langle A, g \rangle$ be an arbitrary game, and $\sigma_i : T_i \rightarrow A_i$ be the agent $i$’s strategy.

**Definition 2.2.2.** (Bayesian equilibrium) The family of strategies $(\sigma_i)_{i \in N}$ is a Bayesian equilibrium if, for each $i \in N$, $t_i \in T_i$ and $a_i \in A_i$, $\int_T u_i(g \circ \sigma) d\lambda_i(t_i) \geq \int_T u_i(g \circ (a_i, \sigma_{-i})) d\lambda_i(t_i)$.

In an uncountable type space, we focus only on measurable equilibria. The next definition is from Jackson.

**Definition 2.2.3.** (Bayesian implementation) A mechanism $\mathcal{M}$ Bayesian implements a SCS $F$ if the mapping $\pi$ assigns a unique $s$ to each $t_i$, so all uncertainty about $S$ is included in the one about $T_{-i}$.
2.2.2 Robust implementation

The practical difficulty in the Bayesian implementation argument is that the planner must know the Harsanyi type space of the agents in order to design a mechanism. It is not likely to be met in applications. Therefore we are required to think about some kind of “belief-free” implementation concept. Bergemann and Morris first formally defined such robust implementation and gave a characterization. For the rest of this section, we only deal with SCF. We discuss why it is later.

**Definition 2.2.4.** (Bergemann-Morris [8]) A mechanism $M$ robustly implements a SCF $f$ if $M$ Bayesian implements $f$ in every Harsanyi type space $T$.

We say that a SCF $f$ is robustly implementable if there exists a mechanism that robustly implements $f$.

We introduce Bergemann-Morris’s characterization result without proof. They use the following concepts.

**Definition 2.2.5.** (EPIC) A social choice function $f$ satisfies ex-post incentive compatibility (EPIC) if, for each $i \in N$, $s_i, s'_i \in S_i$, and $s_{-i} \in S_{-i}$, $u_i(f(s_i, s_{-i}), s_i, s_{-i}) \geq u_i(f(s'_i, s_{-i}), s_i, s_{-i})$.

**Definition 2.2.6.** (Robust monotonicity)

Let $Y_i(s_{-i}) \equiv \{y \in X \mid \forall s'_i \in S_i , u(y, s'_i, s_{-i}) \leq u(f(s'_i, s_{-i}), s'_i, s_{-i})\}$. Then, a SCF $f$ satisfies
robust monotonicity if, for each unacceptable deception $\beta$,

$\exists i \in N, \ s_i \in S_i, \ and \ s'_i \in \beta_i(s_i) \ s.t.

\forall s'_{-i} \in S_{-i}, \ \exists y \in Y_i(s'_{-i}) \ s.t.

\forall s_{-i} \in S_{-i} \ with \ s'_{-i} \in \beta_{-i}(s'_{-i}),

u_i(y, s_i, s'_{-i}) > u_i(f(s'_i, s'_{-i}), s_i, s'_{-i}).$

**Definition 2.2.7. (Bad outcome condition)** For each $i \in N$, there exists $y^*_i \in X$ such that

$\forall s_i \in S_i, \ \forall s_{-i} \in S_{-i}, \ and \ \forall \psi_i \in \Delta(S_{-i}),$

there exists $y_i \in Y_i(s_{-i}) \ s.t.$

$\int_{S_{-i}} u_i(y_i, s_i, s'_{-i})d\psi > \int_{S_{-i}} u_i(y^*_i, s_i, s'_{-i})d\psi.$

Here is Bergemann-Morris result:

**Theorem 2.2.8. (Bergemann and Morris [7])** Under the bad outcome condition, a SCF $f$ is robustly implementable if and only if it satisfies EPIC and robust monotonicity.

### 2.3 The universal type space

Bergemann-Morris gave a necessary condition and a sufficient condition for robust implementation of a SCF. However they derived it through another implementation concept using rationalizable actions. Therefore their canonical mechanism that implements a SCF is different from that in the existing implementation literature which uses some equilibrium concept such as Nash equilibrium or Bayesian equilibrium. One problem here is that we cannot apply the mechanism to SCC. In this paper, we take a different approach from Bergemann-Morris. We try to characterize robust implementation by applying the existing results about Bayesian
implementation to *the universal type space* founded by Mertens and Zamir [39]. In this section, we introduce the construction of the universal type space by Brandenburger and Dekel [11]. We assume that $S$ is a compact metrizable space and $N = \{1, 2\}$. It is straightforward to extend to more than two agents.

The universal type space is the space of coherent hierarchies of beliefs over beliefs over $S$. A hierarchy of beliefs is an infinite sequence of beliefs over $S$ and the other agents’ beliefs over $S$ and so on. Formally, let

$$
Z^0 \equiv S,
$$

$$
Z^1 \equiv Z^0 \times \Delta(S),
$$

$$
\vdots
$$

$$
Z^k \equiv Z^{k-1} \times \Delta(Z^{k-1}),
$$

$$
\vdots
$$

Let $Z^\infty \equiv \prod_{k \geq 0} \Delta(Z^k)$. Then $Z^\infty$ is the entire set of hierarchies of beliefs over $S$. But we require each order of belief to be coherent with the previous order of belief. Let $T^* \subset Z^\infty$ be the set of coherent belief hierarchies where coherency is commonly known. Brandenburger-Dekel showed the next theorem.

**Theorem 2.3.1.** (*Brandenburger and Dekel [11]*) There exists a homeomorphism between $T^*$ and $\Delta(S \times T^*)$.

From this theorem, we can define the universal type space as $U(S) \equiv \langle S, T^* \times T^*, h_1^*, h_2^* \rangle$ where $h_i^*$ is homeomorphism. And we have the theorem by Mertens-Zamir.
Theorem 2.3.2. (Mertens and Zamir [39]) Every Harsanyi type space without redundant
types can be homeomorphically embedded to $U(S)$.

Redundancy will be discussed in the section 6.3. For now, note that one could conceive the
following approach: because of Theorem 4.2, if a SCC $F$ if Bayesian implementable on $U(S)$
then it is robustly implementable, and conversely.

2.4 Extension property

Friedenberg and Meier [23] showed that a part of any Bayesian equilibrium on $U(S)$ focusing
on its sub type space becomes a Bayesian equilibrium on that sub space, but not all Bayesian
equilibria of a sub type space can be realized as such a part of an equilibrium on $U(S)$.
Friedenberg-Meier named the former property of Bayesian equilibrium Pull-Back property
and the latter Extension property. Failure of Extension property is a problem here because
there is no guarantee that a mechanism implementing a SCF on $U(S)$ implements the func-
tion on each sub type space of $U(S)$.

We give the formal definition of Extension property and an example of its failure by Friedenberg-
Meier. Let a game $\Gamma$ be a family of an action space $A \equiv \Pi_{i \in N} A_i$ and a payoff function
$g : A \to X$, i.e. $\Gamma \equiv \langle A, g \rangle$.

Definition 2.4.1. (Friedenberg and Meier [23]) Let $T$ and $\tilde{T}$ be arbitrary Harsanyi type
spaces such that $\tilde{T} \subset T$. Bayesian equilibrium satisfies the extension property if, for each
Bayesian equilibrium $\tilde{\sigma}$ of the game $\langle \tilde{T}, \Gamma \rangle$, there exists a Bayesian equilibrium $\sigma$ of $\langle T, \Gamma \rangle$
such that $\sigma$ is an extension of $\tilde{\sigma}$ to $\langle T, \Gamma \rangle$. 
However this property does not necessarily hold as in the example below.

**Example (Friedenberg-Meier):** Let $N = \{1, 2\}$ and consider the following two type spaces $T_0$ and $T_1$: $T_0 \equiv (\{s^*\}, \{t_i^*\}_{i \in N}, (\lambda_i)_{i \in N})$, where for each $i \in N$, $\lambda_i(t_i^*)[(s^*, t_{i-1}^*)] = 1$. To define $T_1$, we need another preliminary type space $T' = (S_1 \times S_2, T_1 \times T_2, (\lambda'_i)_{i \in N})$. Here, for each $i \in N$, we assume $s_i^* \notin S_i$ and $t_i^* \notin T_i$. $T_1$ is a kind of compound type space of $T_0$ and $T'$:

$$T_1 \equiv (S \cup \{s^*\}, (T_i \cup \{t_i^*\})_{i \in N}, (\lambda_i)_{i \in N}) \text{ such that}$$

$$\forall i \in N, t_i \in T_i, \lambda_i(t_i)[(s^*, t_{i-1}^*)] = p,$$

$$\lambda_i(t_i)[E] = (1 - p)\lambda'(t_i)[E] \text{ for each } E \subset [S \times T_{i-1}].$$

This belief structure of $T_1$ is shown in the picture below.

From this picture, it is clear that $T_0$ is represented as a point $(s^*, t_1^*, t_2^*)$ in $T_1$. 
Next we define a symmetric game on these type spaces. Let $\Gamma = (\langle A_i \cup \{a_i^*\} \rangle_{i \in N}, g)$, where $g : S \cup \{s^*\} \times A \cup \{a^*\} \to \mathbb{R}^2$. The detailed structure of $g$ is given in the next picture.

Here $x, y > 0$, and for each $(a_1, a_2) \in A_1 \times A_2$, $g_1(a_1, a_2) \in [1, 2]$. Agent 2’s payoff is similarly defined. And we have another assumption about $g$.

**Assumption 2.4.2.** The Bayesian game $G' \equiv \langle T', \Gamma' = (A, g) \rangle$ does not have Bayesian equilibrium.

Now two games $G_0 \equiv \langle T_0, \Gamma \rangle$ and $G_1 \equiv \langle T_1, \Gamma \rangle$ are defined. Although $G_0$ shares the same action space and payoff functions and $T_0$ is included in $T_1$, $G_0$ is a technically different game from $G_1$. Extension property does not hold for these two games. The intuition is following.

There is a Bayesian equilibrium $\sigma^0$ of $G_0$ such that, for each $i \in N$, $\sigma^0_i(t^*_i) = a_i \in A_i$. If Extension property holds, there exists a Bayesian equilibrium $\sigma^1$ of $G_1$ such that, for each

---

3By choosing proper $g$, we can satisfy this condition. See details in Friedenberg-Meier and Sion-Wolf
\[ i \in N, \quad \sigma_1^i(t_i^*) = a_i \in A_i. \] Then, it is clear that \( \sigma_2^j(t_j^*) \in A_2. \) Then, from the figure, we know that \( a_1^* \) is strictly dominated at each \( t_1 \in T_1 \), which means that \( \sigma_1^i(t_i) \in A_1 \) for all \( t_1 \in T_1 \cup \{ t_1^* \} \). By the same logic, \( a_2^* \) is strictly dominated at each \( t_2 \in T_2 \). Thus the equilibrium strategy \( \sigma^1 \) takes actions only from \( A \), so by focusing on the subset \( (S_1 \times S_2, T_1 \times T_2) \), \( \sigma^1 \) is also strategy on the game \( G' \). From the belief structure in the figure, we know that \( \sigma^1 \) is a Bayesian equilibrium on the game \( G' \). But this violates Assumption 2.4.2. \[ \square \]

2.5 The universal knowledge-belief space

2.5.1 Knowledge-belief space

We discussed the failure of the Extension property on the universal type space. It happens because the agents do not know the “context” of the game. Here “context” means the true type structure they belong to. The Extension property does not fail if the agents know the context at each epistemic type. For this purpose, we introduce knowledge in addition to subjective belief. We adopt a knowledge-belief (hereafter KB space) space introduced by Aumann [6]. A KB space is a Harsanyi type space with knowledge operator.

**Definition 2.5.1.** A tuple \( V = (V \subset S \times T, (\lambda_i)_{i \in N}, (P_i)_{i \in N}) \) is a knowledge-belief space if it
satisfies

(1) $\lambda_i : T_i \to \Delta(S \times T_{-i})$, 

(2) $P_i : T_i \to \mathcal{K}(S \times T_{-i})$, 

(3) $\forall i \in N, \ t_i \in T_i, \ \text{Supp}(\lambda_i(t_i)) \subset P_i(t_i)$, 

(4) $\forall i \in N, \ \forall t_i, \ \exists s_i \in S_i \ s.t. \exists s_{-i} \in S_{-i} \ and \ s \in P_i(t_i)$, 

(5) $\forall i \in N, \ t_i \in T_i$, 

$\quad \text{if } (s, t_i, t_{-i}) \in P_i(t_i), \text{ then, } \forall j \in N, \ (s, t_i, t_{-i}) \in P_j(t_j)$, 

(6) $V \subset S \times T \ s.t. \ V = \{(s, t) \in S \times T \mid \forall i \in N, (s, t) \in P_i(t_i)\}$.

It is worth while to emphasize that it is without loss of generality to study KB spaces instead of type spaces. In any Bayesian game analysis, it is implicitly assumed that the type space which the agents belong to is a common knowledge. From this perspective, any Harsanyi type space should be interpreted as a KB space where the agents know that they are in the type space.

From (2) in Definition 2.5.1, it follows that every agent knows the context. It leads to the next theorem which states that the Extension property is satisfied on KB spaces. To show this theorem, we first give the lemma that every KB space is decomposable into sub KB spaces.

**Lemma 2.5.2.** Let $\mathcal{V} \equiv \langle V, (\lambda_i)_{i \in N}, (P_i)_{i \in N} \rangle$ be a knowledge-belief space, and $\mathcal{V}_0 \equiv \langle V_0, (\lambda_i)_{i \in N}, (P_i)_{i \in N} \rangle$ be a sub knowledge-belief space of $\mathcal{V}$. Then, the tuple $\langle V_0^c, (\lambda_i)_{i \in N}, (P_i)_{i \in N} \rangle$ is also a knowledge-belief space.

**Proof.** Let $x \in V_0^c$ be an arbitrary point in the complement of $V$. Suppose that, for some
$i \in N$, $P_i(x_i) \cap V_0 \neq \emptyset$. Let $y \in P_i(y_i) \cap V_0$. By the definition, for each $i \in N$, $y \in P_i(y_i)$ and $P_i(y_i) \subset V_0$. However, the condition (5) in the definition 2.5.1 implies that $x \in P_i(y_i)$. It means that $P_i(x_i) \cap V_0^c \neq \emptyset$. It contradicts the fact that $P_i(y_i) \subset V_0$. Therefore $P_i(x_i) \subset V_0^c$.

Since $\mathcal{V}$ is a KB space, the conditions (1)-(6) in 2.5.1 are satisfied at $x$. So the fact $P_i(x_i) \subset V_0^c$ implies that the conditions (1)-(6) are satisfied for $(V_0^c, (\lambda_i)_{i \in N}, (P_i)_{i \in N})$. \hfill \Box

As a corollary of Lemma 2.5.2, we have the next theorem.

**Theorem 2.5.3.** Let $\Gamma = (A, (g_i)_{i \in N})$ be a game. For each KB space $\mathcal{V}$ and each sub KB space $\mathcal{V}_0$ s.t. $\mathcal{V}_0 \subset \mathcal{V}$, Extension property for Bayesian equilibrium holds as long as there exists at least one Bayesian equilibrium of the game $G = (\mathcal{V}, \Gamma)$.

### 2.5.2 The universal knowledge-belief space

Since Extension property of Bayesian equilibrium holds in KB spaces, if we find a “universal” KB space, Bayesian implementation on that space implies Bayesian implementation on every KB space, apart from the existence of redundant types. Meier [38] showed the existence of the universal KB space in the case without topology by using a syntactical method. However in order to deal with the problem of redundant types (see subsection 6.3), we need to work with a topological KB space. Thus, in this subsection, we construct the topological universal KB space. Formally, a topological KB space is a continuous KB space, defined as follows:

**Definition 2.5.4.** A KB space $\mathcal{V}$ is a continuous knowledge-belief space if, for each $i \in N$, the spaces $S_i$, $T_i$ and $V$ are compact metrizable and the mappings $\lambda_i$ and $P_i$ are continuous.
We can basically apply the standard belief hierarchy technique to construct the universal continuous KB space. But we also have to think about knowledge as well as belief. The space of knowledge and belief hierarchies over $S$ is as follows;

\[
\tilde{Z}_0 \equiv S,
\]

\[
\tilde{Z}_1 \equiv \tilde{Z}^0 \times \Delta(\tilde{Z}^0) \times K(\tilde{Z}^0),
\]

\[
\vdots
\]

\[
\tilde{Z}_k \equiv \tilde{Z}_{k-1} \times \Delta(\tilde{Z}_{k-1}) \times K(\tilde{Z}_{k-1}),
\]

\[
\vdots
\]

Let the entire space of knowledge-belief hierarchies be $T_0 \equiv \Pi_{k \geq 0}\{\Delta(\tilde{Z}_k) \times K(\tilde{Z}_k)\}$.

When we assume the agents’ rationality, among the entire set of the knowledge-belief hierarchies, we should focus on the hierarchies which are coherent with their previous orders. Let $\tau_n \equiv (\mu_n, \kappa_n) \in \Delta(\tilde{Z}_{n-1}) \times K(\tilde{Z}_{n-1})$. Then, the set of the coherent knowledge-and belief hierarchies are given as follows;

\[
T_1 \equiv \{(\tau_n)_{n \geq 1} \in T_0 \mid \forall n \geq 1, \text{ Proj}_{(\Delta(\tilde{Z}_n) \times K(\tilde{Z}_n))}\mu_{n+1} = \mu_n, \text{ and Proj}_{(\Delta(\tilde{Z}_n) \times K(\tilde{Z}_n))}\kappa_{n+1} = \kappa_n}\}.
\]

Concerning the space of coherent belief-knowledge hierarchies, we can show a version of Kolmogorov extension theorem on the lines of arguments by Brandenburger and Dekel [11] and Mariotti et al. [35].
Lemma 2.5.5. Let \( \{Z_k\}_{k \geq 0} \) be a countable family of compact metrizable spaces. Let \( Z^n \equiv \Pi_{0 \leq k \leq n} Z_k \), \( Z^\infty \equiv \Pi_{k \geq 0} Z_k \), and

\[
T \equiv \{(\tau_n)_{n \geq 1} \mid \forall n, \ \tau_n = (\mu_n, \kappa_n) \text{ s.t. } (1) \ \mu_n \in \Delta(Z^{n-1}), \ \kappa_n \in K(Z^{n-1}), \text{ and } \\
(2) \ \text{Proj}_{(Z^n-1)} \mu_{n+1} = \mu_n, \ \text{Proj}_{(Z^n-1)} \kappa_{n+1} = \kappa_n \}.
\]

Then there exists homeomorphism \( h : T \to \Delta(Z^\infty) \times K(Z^\infty) \) such that, for each \( n \geq 1 \),

\[
\text{Proj}_{(Z^n)} h((\tau_k)_{k \geq 1}) = (\mu_n, \kappa_n).
\]

Proof. Given \( (\tau_n)_{n \geq 1} \in T \), let \( \tilde{\mu} \) be the Kolmogorov extension of \( (\mu_n)_{n \geq 1} \) to \( Z^\infty \). We can also define compact spaces such that \( K_n \equiv \kappa_n \times \Pi_{k \geq n} Z_k \) and \( K \equiv \bigcap_{n \geq 1} K_n \).

Let \( h : T \to \Delta(Z^\infty) \times K(Z^\infty) \) be such that \( (\tau_n)_{n \geq 1} \mapsto (\tilde{\mu}, K) \). By construction, we can say that \( h \) is injection. And for each \( (\mu, \kappa) \in \Delta(Z^\infty) \times K(Z^\infty) \), we can find a sequence \( (\tau_n)_{n \geq 1} = (\mu_n, \kappa_n)_{n \geq 1} \) such that \( \text{Proj}_{(Z^n)} (\mu, \kappa) = (\mu_n, \kappa_n) \) by taking projection over \( Z^n \) of \( \mu \) and \( \kappa \) respectively. From these results, we have that \( h \) is bijection. Since, for each \( k \), \( Z_k \) is compact metrizable, \( h_\Delta : (\mu_n)_{n \geq 1} \mapsto \mu \) and \( h_K : (\kappa_n)_{n \geq 1} \mapsto \kappa \) are continuous mappings.\(^4\)

Therefore \( h \) is a continuous bijection from \( T \) to \( \Delta(Z^\infty) \times K(Z^\infty) \). Both \( T \) and \( \Delta(Z^\infty) \times K(Z^\infty) \) are compact metrizable spaces. Thus \( h \) is a homeomorphism. \( \square \)

We have the next proposition as a corollary of Lemma 2.5.5.

Proposition 2.5.6. The set of coherent knowledge-belief hierarchies \( T_1 \) is homeomorphic to \( \Delta(S \times T_0) \times K(S \times T_0) \).

To exploit agents’ rationality completely, we have to not only require agents’ coherency but also require it to be common knowledge across the agents. So we have to study the following

\[^4\text{See Brandenburger and Dekel [11] and Mariotti \textit{et al} [35].}\]
$T^*$ instead of $T_1$. The space $T^*$ is the set of coherent KB hierarchies that satisfy common knowledge of coherency, i.e.,

$$
T_k \equiv \{ \tau \in T_{k-1} \mid h_1(\tau)[S \times T_{k-1}] = 1, \text{and} \ h_2(\tau) \in K(S \times T_{k-1}) \},
$$

$$
\vdots
$$

$$
T^* \equiv \bigcap_{k \geq 1} T_k.
$$

**Proposition 2.5.7.** The set $T^*$, which is defined above, is non-empty, and homeomorphic to the space $\Delta(S \times T^*) \times K(S \times T^*)$.

**Proof.** Let $h$ be a homeomorphism from $T_1 \to \Delta(S \times T_0) \times K(S \times T_0)$. We define

$$
\tilde{T}^* \equiv \{ \tau \in T_1 \mid h_1(\tau)[S \times T^*] = 1, \text{and} \ h_2(\tau) \in K(S \times T^*) \}.
$$

By construction, for each $\tau \in T^*$, $h_1(\tau)[S \times T^*] = 1$, and $h_2(\tau) \in K(S \times T^*)$. Therefore, $T^* \subset \tilde{T}^*$. On the other hand, for each $\tau \in \tilde{T}^*$, $h_1(\tau)[S \times T_1] = 1$, and $h_2(\tau) \in K(S \times T_1)$. It means that $\tilde{T}^* \subset T_1$. Recursively, we have that, for each $n \geq 1$, $\tilde{T}^* \subset T_n$. Therefore $\tilde{T}^* \subset T^*$, which concludes that $\tilde{T}^* = T^*$.

From the above results,

$$
h(T^*) \equiv \{ \tau = (\mu, \kappa) \in \Delta(S \times T_0) \times K(S \times T_0) \mid \mu[S \times T^*] = 1, \text{and} \ \kappa \in K(S \times T^*) \}.
$$

Since $T^*$ is a compact subset of $T_0$, it follows that $\Delta(S \times T^*)$ is homeomorphic to the space $\{ \mu \in \Delta(S \times T_0) \mid \mu[S \times T^*] = 1 \}$ and $K(S \times T^*)$ is homeomorphic to $\{ \kappa \in K(S \times T_0) \mid \kappa \in K(S \times T^*) \}$. Let $g_1$ and $g_2$ be the homeomorphism respectively. Then $g \equiv (g_1, g_2)$ is homeomorphism from $\Delta(S \times T^*) \times K(S \times T^*)$ to $h(T^*)$. It follows that the mapping $g^{-1} \circ h : T^* \to \Delta(S \times T^*) \times K(S \times T^*)$ is homeomorphism. $\square$
Note that $U^*(S) \equiv S \times \Pi_{i \in N} T^*$ is not the universal space (c.f. Mertens-Zamir) because we imposed the axioms (3)-(6) in Definition 2.5.1 on knowledge-belief hierarchies. To construct the universal KB space, we have to pick up the largest subset of $U^*(S)$ satisfying the axioms. Note that we also use the term $U^*(S)$ to mean the space associated with the homeomorphism $\lambda^*$ and $P$ hereafter.

Even if $U^*(S)$ is not the universal KB space we seek, we can show that any continuous KB space without redundant types, can be embedded there. For the proceeding argument, we should clarify what it means for a KB space to be embedded in another and what redundant types are.

**Definition 2.5.8.** (S-homeomorphism) Let $V$ and $W$ be continuous KB spaces. A mapping $h: V \rightarrow W$ is S-homeomorphism if

1. $h$ is homeomorphism,
2. $\forall i \in N, \forall t_i \in T_i, \lambda_i(t_i) \circ h^{-i} = \phi_i(h(t_i))$,
3. $\forall i \in N, \forall t_i \in T_i, h[P_i(t_i)] = \bar{P}_i(h(t_i))$,

where $v \equiv (t_i, t_{-i}, s) \in V$, and $h(t_i) \equiv \text{Proj}_{T_i} h(v)$.

We need the next lemma to show that every KB space without redundant types can be embedded in $U^*(S)$.

**Lemma 2.5.9.** Let $V$ be a continuous KB space. Then there exists a continuous mapping $\gamma: V \rightarrow U^*(S)$.

**Proof.** Given types in continuous KB spaces, we can derive the hierarchy mapping $\gamma$ from
those types to their knowledge-belief hierarchies over $S$ as follows;

\[
\forall i \in N, \quad \gamma_1^i : t_i \mapsto (\text{Proj}_S \lambda_i(t_i), \text{Proj}_S P_i(t_i)),
\]

\[
\gamma_2^i : t_i \mapsto [(\text{Proj}_S \lambda_i(t_i), \text{Proj}_S P_i(t_i))], \quad (\lambda_i(t_i) \circ (\gamma_1^i)^{-1}, \gamma_1^i[P_i(t_i)]),
\]

\[
\vdots
\]

\[
\gamma_k^i : t_i \mapsto [(\lambda_i(t_i) \circ (\gamma_{k-1}^i)^{-1}, \gamma_{k-1}^i[P_i(t_i)]), \ldots, (\lambda_i(t_i) \circ (\gamma_{k-1}^1)^{-1}, \gamma_{k-1}^1[P_i(t_i)]),
\]

\[
\vdots
\]

Let $\gamma_i^\infty \equiv \lim_{k \to \infty}(\gamma_i^k)$. We show that this inverse limit $\gamma_i^\infty$ is a continuous mapping from $V$ to $\Omega$ endowed with the product topology. Let $(t_i^l)_{l \in \mathbb{N}}$ be such that, for each $l \in \mathbb{N}$, $t_i^l \in T_i$ and there exists $t_i^* \in T_i$ with $t_i^* = \lim_{l \to \infty} t_i^l$. We want to show that, for each level $k \in \mathbb{N}$, $\lim_{l \to \infty} \gamma_i^k(t_i^l) = \gamma_i^k(t_i^*)$.

We use induction. For each $i \in N$, at $k = 1$, we have $\gamma_i^1(t_i^*) = \lim_{l \to \infty} \gamma_i^1(t_i^l)$ because $\lambda_i$ and $P_i$ are both continuous. Suppose that, for each $i \in N$, $\gamma_i^k$ is continuous. Then, $\lambda_i(t_i^l) \circ (\gamma_{k-1}^i)^{-1} \to \lambda_i(t_i^*) \circ (\gamma_{k-1}^i)^{-1}$ as $l \to \infty$ because $\lambda_i$ is continuous and $\gamma_{k-1}^i$ is fixed. Also since $\gamma_{k-1}^i$ is continuous, we have $\gamma_{k-1}^i[P_i(t_i^l)] \to \gamma_{k-1}^i[P_i(t_i^*)]$ as $l \to \infty$. It means that $\gamma_i^{k+1}(t_i^*) = \lim_{l \to \infty} \gamma_i^{k+1}(t_i^l)$.

Therefore, for each $k \in \mathbb{N}$, $\lim_{l \to \infty} \gamma_i^k(t_i^l) = \gamma_i^k(t_i^*)$, which means that $\gamma_i^\infty$ is continuous.

\[\square\]

Based on this hierarchy mapping, we can define redundant types.

**Definition 2.5.10.** Let $t_i$ and $t_i'$ be types from the same continuous KB space $V$. The types $t_i$ and $t_i'$ are redundant types if they are mapped to the same hierarchy by the hierarchy mapping $\gamma$. 

The next proposition directly follows from Lemma 2.5.9.

**Proposition 2.5.11.** As long as a continuous KB space $V$ does not have redundant types, the hierarchy mapping $\gamma : V \to U^*(S)$ is $S$-isomorphism.

**Proof.** Since $V$ does not have redundant types, the hierarchy mapping $\gamma^\infty$ is a bijection. Therefore, by Theorem 2.36 in Aliprantis and Border [1], $\gamma^\infty$ is (homeomorphic) embedding from $V$ to $\Omega$. □

Although we know $U^*(S)$ is large enough to embed any KB space without redundancy, there is no guarantee that $U^*(S)$ is a continuous KB space as we mentioned above. The space $U^*(S)$ is just the set of coherent knowledge-belief hierarchies, and does not have to satisfy the axioms in Definition 2.5.1. One way to find the universal KB space is to pick up the largest subset of $U^*(S)$ which satisfies the axioms. That is given by the following space.

$$\Omega(S) \equiv \{ \omega \in U^*(S) \mid \text{There exists a continuous KB space } V \text{ with } \omega \in V \}$$

Our plan is to show that $\Omega(S)$ is the *universal continuous KB space*, in the sense that it is a continuous KB space in which we can embed any continuous KB space. Therefore we only have to show that $\Omega(S)$ is a continuous KB space.

For this purpose, we define another subspace of $U^*(S)$ similar to $\Omega(S)$:

$$\check{\Omega}(S) \equiv \{ \omega \in U^*(S) \mid \text{There exists a KB space } L \text{ with } \omega \in L \}$$

Note that $\Omega(S) \subset \check{\Omega}(S)$ by construction.
A KB space is called semi-continuous if it is a continuous KB space without the compactness of the spaces $S$ and $T$

**Lemma 2.5.12.** The space $\tilde{\Omega}(S)$ is a semi-continuous KB space.

**Proof.** As shown above, there exist the homeomorphism $\lambda^*_i : T^* \to \Delta(S \times T^*)$ and $P_i : T^* \to K(S \times T^*)$. By restricting the domains, we obtain the continuous functions $\lambda^*_i : \text{Proj}_T, \tilde{\Omega}(S) \to \Delta(\text{Proj}_{S \times T^*}, \tilde{\Omega}(S))$ and $P_i : \text{Proj}_T, \tilde{\Omega}(S) \to K(\text{Proj}_{S \times T^*}, \tilde{\Omega}(S))$. Then, from (3) in Definition 2.5.1, we have, for each $\omega \in \tilde{\Omega}(S)$, $\text{Supp}(\lambda^*_i(\omega)) \subset P^*_i(\omega)$. And each $\omega$ is in some KB subspace $L$, so (4) and (5) in Definition 2.5.1 are satisfied. For each $\omega \in \tilde{\Omega}(S)$, there exists a KB space $L$ with $\omega \in L$. Therefore, for each $i \in N$, $\omega \in P^*_i(\omega_i)$. It means $(\tilde{\Omega}(S), \lambda^*, P^*)$ is a semi-continuous KB space. □

Next we show that $\tilde{\Omega}(S)$ is a continuous KB space by showing that $\tilde{\Omega}(S)$ is compact.

**Proposition 2.5.13.** The space $\tilde{\Omega}(S)$ is a compact space.

**Proof.** Since $\tilde{\Omega}(S) \subset U^*(S)$ and $U^*(S)$ is compact, we only have to show that $\tilde{\Omega}(S)$ is closed. Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence such that, for each $k \in \mathbb{N}$, $\omega^k \in \tilde{\Omega}(S)$ and there exists a limit $\omega^* \in U^*(S)$ i.e. $\lim_{k \to \infty} \omega^k = \omega^*$. From the lemma above, $\tilde{\Omega}(S)$ is a semi-continuous KB space. Therefore, for each $k \in \mathbb{N}$ and $i \in N$, it holds that $\text{Supp}(\lambda^*_i(\omega^k)) \subset P_i(\omega^k)$. And it also implies that, for each $i \in N$ and $k \in \mathbb{N}$, $\lambda^*(\omega^k_i)[P_i(\omega^k_i)] = 1$. Remind that $\lim_{k \to \infty} \omega^k = \omega^*$ by the assumption. By the continuity of $P^*_i$, we have, for each $i \in N$, $P_i(\omega^k_i) \to P_i(\omega^*_i)$ in the Hausdorff topology.

So, for each $\epsilon_1 > 0$, there exists $k_1$ such that, for each $k \geq k_1$, $H_d(P_i(\omega^k_i), P_i(\omega^*_i)) \leq \epsilon_1$. 

Therefore, for each \( k \geq k_1 \), \( P_i(\omega^k) \subset B_{cl}(P_i(\omega^*), \epsilon_1) \). Here \( B_{cl}(X, \epsilon) \) stands for the \( \epsilon \)-closed ball of the set \( X \). It means that \( \lambda^*_i(\omega^k)[B_{cl}(P_i(\omega^*), \epsilon_1)] = 1 \). By the construction, \( \lambda^*_i(\omega^k) \rightarrow \lambda^*_i(\omega^*) \) as \( k \rightarrow \infty \). So we have \( \lim_{k \rightarrow \infty} \lambda^*_i(\omega^k)[B_{cl}(P_i(\omega^*), \epsilon_1)] \leq \lambda^*_i(\omega^*)[B_{cl}(P_i(\omega^*), \epsilon_1)] \), which means that \( \lambda^*_i(\omega^*)[B_{cl}(P_i(\omega^*), \epsilon_1)] = 1 \). In the same way, for each \( 0 < \epsilon \leq \epsilon_1 \), \( \lambda^*_i(\omega^*)[B_{cl}(P_i(\omega^*), \epsilon_1)] = 1 \). We can pick up a sequence \( (\epsilon_i)_{i=1}^{\infty} \) such that \( \lim_{i \rightarrow \infty} \epsilon_i = 0 \). Then, since \( (B_{cl}(P_i(\omega^*), \epsilon_i))_{i=1}^{\infty} \) is monotonically decreasing, \( \lim_{i \rightarrow \infty} B_{cl}(P_i(\omega^*), \epsilon_i) = P_i(\omega^*) \).

Since, for each \( l \), \( \lambda^*_i(\omega^*)[B_{cl}(P_i(\omega^*), \epsilon_l)] = 1 \) and \( \lambda^*_i(\omega^*)[P_i(\omega^*)] = 1 \). So \( \text{Supp}(\lambda^*_i(\omega^*)) \subset P_i(\omega^*) \). The convergence of the knowledge sets \( (P^k)_{k \in \mathbb{N}} \) also implies that, for each \( i \in \mathbb{N} \) and \( \epsilon > 0 \), \( P_i(t^*_i) \subset B(\tilde{\Omega}(S), \epsilon) \).

Before going further, we consider the condition (4) in the Definition 2.5.1. We can easily check that there exists \( s^*_i \in S_i \) such that \( \text{Proj}_{(S_i)} P_i(t^*_i) = \{ s^*_i \} \) because \( s^*_i = \lim_{k \rightarrow \infty} s^k_i \) with \( \text{Proj}_{(S_i)} P_i(t^k_i) = \{ s^k_i \} \).

Let \( \tilde{\Omega}(S) \) be the closure of \( \tilde{\Omega}(S) \). For each \( \omega \in \tilde{\Omega}(S) \setminus \tilde{\Omega}(S) \), it is the limit of a convergent sequence in \( \tilde{\Omega}(S) \). Therefore, from the preceding results, for each \( i \in \mathbb{N} \), \( \text{Supp}(\lambda^*_i(\omega_i)) \subset P_i(\omega_i) \).

From the above argument, we have, for any \( \epsilon > 0 \), \( P_i(\omega_i) \subset B(\tilde{\Omega}(S), \epsilon) \). So, for any \( \epsilon > 0 \), it holds that \( P_i(\omega_i) \subset B(\tilde{\Omega}(S), \epsilon) \). Since the set \( \tilde{\Omega}(S) \) is closed, \( P_i(\omega_i) \subset \tilde{\Omega}(S) \). As a result, \( \tilde{\Omega}(S) \) satisfies Definition 2.5.1 except for (5) and (6).

Finally we show that (5) and (6) in Definition 2.5.1 hold at \( \omega \in \tilde{\Omega}(S) \). Suppose that \( \omega \in \tilde{\Omega}(S) \setminus \tilde{\Omega}(S) \). Otherwise it is clear that (5) and (6) are satisfied. Let \( \tilde{\omega} \in P_i(\omega) \). All we have to show is (i) For all \( j \in \mathbb{N} \), \( \omega \in P_j(\omega) \), and (ii) For all \( j \neq i \), \( \tilde{\omega} \in P_j(\tilde{\omega}) \).\(^5\) Now there exists \( (\omega^k)_{k \in \mathbb{N}} \) such that, for each \( k \) and \( j \in \mathbb{N} \), \( \omega^k \in P_j(\omega^k) \) and \( \lim_{k \rightarrow \infty} \omega^k = \omega \). By the continuity of \( P_j \), for each \( j \in \mathbb{N} \), \( \lim_{k \rightarrow \infty} \omega^k \in \lim_{k \rightarrow \infty} P_j(\omega^k) \), that is, \( \omega \in P_j(\omega) \).

\(^5\)For \( i \), it is clear that \( \tilde{\omega} \in P_i(\tilde{\omega}) \) from the construction of \( U^*(S) \).
Next we show that there exists a subsequence \((\omega_{k_n})_{k_n \in \mathbb{N}}\) such that there exists \((\tilde{\omega}_n)_{n \in \mathbb{N}}\) with, for each \(n\), \(\tilde{\omega}_n \in P_i(\omega_{k_n})\) and \(\tilde{\omega}_n \to \tilde{\omega}\) as \(n \to \infty\). We can construct such a sequence in the following way. By the definition of the Hausdorff topology, for each \(\delta_n > 0\), we have \(d(\tilde{\omega}, P_i(\omega_{k_n})) < \delta_n\) for any sufficiently large \(n\). Since \(P_i(\omega_{k_n})\) is compact and the distance function is continuous, there exists \(\tilde{\omega}_n \in P_i(\omega_{k_n})\) such that \(d(\tilde{\omega}, \tilde{\omega}_n) < \delta_n\). We can find such \(\tilde{\omega}_n\) for each \(\delta_n > 0\). Clearly, the sequence \((\tilde{\omega}_n)_{n \in \mathbb{N}}\) converges to \(\tilde{\omega}\). Since, for each \(k_n, \omega_{k_n} \in \tilde{\Omega}(S)\), for each \(j \in \mathbb{N}\), \(\tilde{\omega}_n \in P_j(\tilde{\omega}_n)\). So the continuity of \(P_j\) implies that \(\lim_{n \to \infty} \tilde{\omega}_n \in \lim_{n \to \infty} P_j(\tilde{\omega}_n)\), that is, \(\tilde{\omega} \in P_j(\tilde{\omega})\).

The above argument shows that \(\tilde{\Omega}(S)\) is a KB space.\(^6\) By the construction, \(\tilde{\Omega}(S) \subset \bar{\Omega}(S)\), and so \(\bar{\Omega}(S) = \tilde{\Omega}(S)\). It means that \(\bar{\Omega}(S)\) is closed.\(^7\)

As a corollary of this result, we have the next theorem.

**Theorem 2.5.14.** Except for KB spaces with redundant types, the space \(\Omega(S)\) is the universal continuous knowledge belief space.

**Proof.** From the construction and Proposition 2.5.11, it is enough to show that \(\Omega(S)\) is a continuous knowledge belief space. By the definition, we know that \(\Omega(S) \subset \bar{\Omega}(S)\). The above lemmas show that \(\bar{\Omega}(S)\) is a continuous KB space. Therefore, \(\Omega(S) \supset \bar{\Omega}(S)\), which means that \(\Omega(S) = \bar{\Omega}(S)\).

\(^6\)Note that \(\bar{\Omega}(S)\) is also a continuous KB space.

\(^7\)Note that \(\tilde{\Omega}(S)\) is a continuous KB space.
2.5.3 The universal knowledge-belief space with a payoff irrelevant parameter space

Now we have to face the existence of redundant types. Even if we find a mechanism which Bayesian implements a social choice function on the universal KB space constructed above, it does not necessarily implements the function on KB spaces with redundant types. One way to deal with such redundant types is to introduce a payoff irrelevant parameter space $C$ as an additional source of uncertainty and construct the space of knowledge-belief hierarchies over $S \times C$ on the lines of Liu’s argument [33]. Liu considered that redundant types arise because some source of uncertainty, which does not affect the payoffs, is not incorporated in the model. As a result, the belief hierarchies over $S$ are not enough to explain the type structure of the agents. Although, in Liu’s argument, the choice of such a payoff irrelevant parameter space depends on the type space to be mapped, we showed in another paper [50] that when Harsanyi type spaces are Polish, the two-valued parameter space $C = \{0, 1\}$ is large enough to embed any type space into the universal type space over $S \times C$. We apply this method to construct the universal KB space that we want.

First we introduce the next mathematical fact.

**Theorem 2.5.15.** (Kechris [32], Theorem 4.14) Every Polish space is homeomorphic to a subspace of the Hilbert cube $[0, 1]^\mathbb{N}$.

Here we study KB space instead of Harsanyi type space, and we need a larger payoff irrelevant parameter space than $\{0, 1\}$. Note that all spaces we are studying are compact metrizable. Since compact metrizable spaces are Polish, they are homeomorphically embedded to $[0, 1]^\mathbb{N}$. This means that the Hibert cube is rich enough to represent the information structure of any KB space. For this reason, we set $C \equiv [0, 1]^\mathbb{N}$ as an exogenous parameter space.
We can define the KB space and all the necessary concepts over $S \times C$ in the same way as over $S$. Let $W \equiv ( W \subset S \times C \times \tilde{T}, (\phi_i)_{i\in\mathbb{N}}, (\bar{P}_i)_{i\in\mathbb{N}} )$ be a continuous KB space. Note that we can apply the same definition of S-homeomorphism to compare KB spaces over $S$ and those over $S \times C$.

The next proposition tells that, for any continuous KB space over $S$, we can always find its S-isomorphic continuous KB space over $S \times C$ which does not have redundant types.

**Proposition 2.5.16.** For any continuous KB space $V$ over $S$, there exists $W$ over $S \times C$ that is S-homeomorphic to $V$ and has no redundant types.

**Proof.** Let $V \equiv ( V \subset S \times T, (\lambda_i)_{i\in\mathbb{N}}, (P_i)_{i\in\mathbb{N}} )$ be a continuous KB space over $S$. Since $V$ is compact metrizable, there exists a homeomorphism to a compact subspace of $C$. Let $\pi : V \to C$ be the homeomorphism. We construct $W$ as follows. Let $W \equiv ( W \subset S \times C \times \tilde{T}, (\phi_i)_{i\in\mathbb{N}}, (\bar{P}_i)_{i\in\mathbb{N}} )$ be such that

1. $W \equiv \{ (v, \pi(v)) \in S \times T \times C \mid v \in V \}$,
2. $\forall i \in \mathbb{N}, \phi_i(t_i) \circ (\text{Id}_{S \times T}, \pi) \equiv \lambda_i(t_i)$,
3. $\forall i \in \mathbb{N}, \bar{P}_i(t_i) \equiv (\text{Id}_{S \times T}, \pi)[P_i(t_i)]$.

Then, $W$ is a continuous KB space over $S \times C$ and S-homeomorphic to $V$ because $\pi$ is homeomorphism.

And since $\pi$ is bijection, any $t_i \neq t'_i \in \text{Proj}_{T_i}W$ knows different $c$ and $c' \in C$. It means that their first order hierarchies over $S \times C$ are different, i.e. $\gamma^1_i(t_i) \neq \gamma^1_i(t'_i)$. Thus there exists no redundant types in $W$. \qed
As a corollary of Theorem 2.5.14 and Proposition 2.5.16, we have the next theorem;

**Theorem 2.5.17.** Any continuous KB space $V$ over $S$ is $S$-homeomorphic to a continuous sub KB space of $\Omega(S \times C)$.

**Theorem 2.5.18.** The continuous KB space $\Omega(S \times C)$ is also a continuous KB space over $S$.

**Proof.** We show that there exists a continuous KB space over $S$ in which the set of Bayesian equilibria is the same as that of $\Omega(S \times C)$ for every game $g : S \times A \to X$. We construct a continuous KB space $\hat{V} \equiv \langle \hat{V} \subset S \times T^*, (\lambda_i)_{i \in N}, (P_i)_{i \in N} \rangle$ as follows:

1. $\hat{V} = \text{Proj}_{(S \times T^*)} \Omega(S \times C)$,
2. $\forall i \in N, \forall t_i \in T^*_i, \lambda_i(t_i) = \text{Marg}_{(S \times T^*_i)} \phi_i(t_i)$,
3. $\forall i \in N, \forall t_i \in T^*_i, P_i(t_i) = \text{Proj}_{(S \times T^*_i)} \bar{P}_i(t_i)$.

Then, $\hat{V}$ is a continuous KB space over $S \times C$ and share the same equilibrium as $\Omega(S \times C)$ for every $g$. 

Now we want to state our first main result. Before that, we have to slightly modify SCC because SCC is defined on the ex post type space, but we are now working on the epistemic type space. We define $\hat{F}$ is a social choice set (hereafter SCS) if $\hat{F} \subset \hat{F} \equiv \{ f \mid f : \Omega(S \times C) \to X \}$.

Without loss of generality, we can identify a SCC $F$ to be a SCS $\hat{F}$ such that $\hat{F} \equiv \{ q \in \hat{F} \mid \forall v \in \Omega(S \times C), q(v) \in F(s(v)) \}$. Based on this, it is straightforward to extend the definition of robust implementation to the SCC case.
Definition 2.5.19. A mechanism $\mathcal{M}$ robustly implements a SCC $F$ if $\mathcal{M}$ Bayesian implements $\hat{F}$ in every Harsanyi type space $T$.

Then we have our first main result.

Theorem 2.5.20. A SCC $F$ is robustly implementable if and only if $\hat{F}$ is Bayesian implementable on $\Omega(S \times C)$.

Proof.

$(\rightarrow)$ Since $\Omega(S \times C)$ is a continuous KB space by Theorem 2.5.18, it is also a Harsanyi type space over $S$. Therefore the robust mechanism Bayesian implements $F$ on $\Omega(S \times C)$ by the definition of robust implementation.

$(\leftarrow)$ From Theorem 2.5.17, any Harsanyi type space $\Lambda$ can be embedded to $\Omega(S \times C)$ as a KB space $\Lambda^{KB}$ whose knowledge set is the entire type space $\Lambda$. Now $\hat{F}$ is Bayesian implementable on $\Omega(S \times C)$, so there exists at least one Bayesian equilibrium. By Theorem 2.5.3, if $\hat{F}$ is Bayesian implementable on $\Omega(S \times C)$, it is Bayesian implemented on $\Lambda^{KB}$, i.e. $\Lambda$. Thus $F$ is robustly implementable. \qed

2.6 Characterization of Implementation

Finally we give a necessary and sufficient condition about when a SCS is Bayesian implementable on $\Omega(S \times C)$. For notational convenience, we just use $\Omega$ for $\Omega(S \times C)$. Note that once we get all agents’ epistemic types $t$, we can deduce what $s$ is. Let $T_i$ be the space of the $i$’s epistemic types on $\Omega$. Then we can consider a SCS to be a set of functions defined on $\Pi_{i \in N}T_i$ instead of $\Omega$, i.e., a subset of $\mathcal{F} \equiv \{f \mid f : \Pi_{i \in N}T_i \to X\}$.\footnote{To be mathematically precise, we have to eliminate the combinations of “null” states. However, as explained later, it does not bring significant changes.} In this section, we use
the mathematical symbol \( F \) to represent a SCS, and assume that \( F \) is countable for technical simplicity.\(^9\)

First we have to define preliminary concepts for the characterization. A \textit{deception} is a way of lying.

\textbf{Definition 2.6.1.} A \textit{deception} is a measurable mapping \( \alpha_i : T_i \to T_i \). The family of the agents’ deception is denoted as \( \alpha \equiv (\alpha_i)_{i \in N} \).

For each \( z \in F, \tilde{t}_i \in T_i \), and \( t \in T \), we use the notation \( z_{\tilde{t}_i}(t) \equiv z(\tilde{t}_i, t-i) \).

The agent \( i \)'s interim preference is given by Von-Neumann-Morgenstern preference.

\textbf{Definition 2.6.2.} For each \( i \in N \), each \( z \) and \( z' \in F \),

\[
z \; R_i(t_i) \; z' \quad \text{if} \quad \int_{t-i \in T-i} zd\lambda_i^*(t_i) \geq \int_{t-i \in T-i} z'd\lambda_i^*(t_i).
\]

\[\text{Definition 2.6.3.} \quad \text{A social choice set } F \text{ satisfies Bayesian incentive compatibility (hereafter, BIC) if, for all } q \in F, \; i \in N, \; \text{and } \tilde{t}_i \in T_i, \]

\[
\forall t_i \in T_i, \quad q \; R_i(t_i) \; q_{\tilde{t}_i}
\]

The next definition is an uncountable version of the \textit{closure} concept used in Palfrey and

\(^9\)Even if the cardinality of a SCC is finite for each \( s \in S \), its corresponding SCS becomes uncountable. One interpretation here is that \( F \) is a subset of the corresponding SCS picked up by the planner for some practical feasibility.
Srivastava [42] and Jackson [29].

**Definition 2.6.4.** *(Closure)* Let \( \{ \Psi_k \}_{k \in \mathbb{N}} \) be a countable family of belief closed measurable spaces such that \( \Omega = \bigcup_{k \in \mathbb{N}} \Psi_k \). A social choice set \( F \) satisfies closure (hereafter C) if, for all \( \{ q^k \}_{k \in \mathbb{N}} \subset F \), there exists \( z \in F \) such that, for each \( k \in \mathbb{N} \) and \( t \in \Psi_k \), \( z(t) = q^k(t) \).

We consider the following environment which is an interim version of the bad outcome assumption by Bergemann-Morris.

**Assumption 2.6.5.** *(Bad outcome assumption)* For each \( i \in \mathbb{N} \), there exists \( z^*_i \in \mathcal{F} \) such that, for each \( q \in \mathcal{F} \), each \( t_i \in T_i \), and each \( \alpha \) there exists \( r \in \mathcal{F} \) such that, (1) for each \( t'_i \in T_i \), \( q R_i(t'_i) r \), and (2) \( (r \circ \alpha) P_i(t_i) (z^*_i \circ \alpha) \).

We adopt Jackson’s monotonicity condition.

**Definition 2.6.6.** *(Jackson [29]*) A social choice set \( F \) satisfies Bayesian monotonicity (hereafter BM) if, for each \( q \in \mathcal{F} \) and deception \( \alpha \),

\[
\text{if } \forall i \in \mathbb{N}, \forall t_i \in T_i, \forall r \in \mathcal{F}, \\
[\forall t'_i \in T_i, q R_i(t'_i) r_{P_i(t_i)}] \Rightarrow q \circ \alpha R_i(t_i) r \circ \alpha \]

\[
\text{then } q \circ \alpha \in \mathcal{F}
\]

As we discuss later, the bad outcome assumption is required only in order to obtain Proposition 2.6.8. It is not needed for Proposition 2.6.11. However, it is not enough for SCCs. We need another assumption.
Assumption 2.6.7. (Economic environment) For each \( q \in \mathcal{F} \), and each \( t \in T \), there exists \( i \in N \) such that

\[
\exists r^i \in \mathcal{F} \quad r^i \succ_P^i(t_i) \preceq q
\]

Under Assumption 2.6.5 and 2.6.7, we have the following theorem.

**Proposition 2.6.8.** Suppose \( |N| > 3 \). If a social choice set \( \mathcal{F} \) satisfies Bayesian incentive compatibility (BIC), closure (C), and Bayesian monotonicity (BM) on \( \Omega \), then \( \mathcal{F} \) is Bayesian implementable on \( \Omega \setminus \mathcal{C}_{\text{null}} \). Here \( \mathcal{C}_{\text{null}} \) is a measurable set such that, for each \( i \) and \( t_i \in \mathcal{C}_{\text{null}} \), \( \lambda_i^*(t_i)[\mathcal{C}_{\text{null}}] = 0 \).

**Proof.** We construct a mechanism \( \mathcal{M} \) as follows.

Let the agent \( i \)'s message space \( M_i \equiv T_i \times \mathcal{F} \times \mathbb{Z}_+ \times \mathcal{F}^F \), and \( M \equiv \Pi_{i \in N} M_i \) where, for each \( q \in \mathcal{F} \), \( m^i_1(q) \in \mathcal{F} \) such that, for each \( t'_i \in T_i \), \( q \succ R_i(t'_i) \succ m^i_1(q) \). We make a partition on \( M \) as follows:

\[
D_0 \equiv \{ m \in M \mid \forall i \in N, \exists q \in \mathcal{F}, \quad m_i = (\ldots, q, 1, \ldots) \},
\]

\[
D^i_1 \equiv \{ m \in M \mid \exists j \in N, \forall j \neq i, \exists q \in \mathcal{F}, \quad m_j = (\ldots, q, 1, \ldots), \quad \text{and} \quad m^j_1 \neq q \quad \text{or} \quad m^j_2 > 1 \},
\]

\[
D_1 \equiv \bigcup_{i \in N} D^i_1,
\]

\[
D_2 \equiv \{ m \in M \mid m \notin D_0 \cup D_1 \}
\]
The payoff function $g : M \rightarrow A$ is given as:

$$
g(m) = q(m^1) \quad \text{if } m \in D_0,$$

$$
g(m) = m_i^2(m^1)(1 - \frac{1}{1 + m_i^2}) + z^i(\frac{1}{1 + m_i^2}) \quad \text{if } m \in D_1 \text{ and } \forall t_i' \in T_i, \ q R_i(t_i') m_i^2 q R_i^* m_i^1,
$$

$$
g(m) = m_i^4(q)(m^1)(1 - \frac{1}{1 + m_i^2}) + z^i(\frac{1}{1 + m_i^2}) \quad \text{if } m \in D_1 \text{ and } \exists t_i' \in T_i, \ m_i^2 q R_i(t_i') q R_i^* P_i(t_i') q,
$$

$$
g(m) = m_k(m^1)(1 - \frac{1}{1 + m_k^2}) + z^k(\frac{1}{1 + m_k^2}) \quad \text{if } m \in D_2,
$$

where $k$ is the agent such that $\forall j \in N, \ m_j^3 \geq m_j^3$.

We show that this mechanism implements $F$ by showing the following two lemmas.

**Lemma 2.6.9.** If $F$ satisfies (BIC), then, for each $q \in F$, there exists a Bayesian equilibrium $\sigma$ of $(\Omega, M, g)$ such that $q = g(\sigma)$.

**Proof.** Fix $q \in F$. Suppose that $F$ satisfies (BIC). Then, for each $i \in N$, $t_i \in T_i$, and $t_i^* \in T_i$, we have that $q R_i(t_i) q R_i^*$. Now, for each $i \in N$, let $\sigma_i$ be the agent $i$'s strategy such that, for each $t_i \in T_i$, $\sigma_i(t_i) = (t_i, q, 1, \ldots)$. Then $g(\sigma) = q$. We show that $\sigma$ is a Bayesian equilibrium.

(Case 1) Let $\tilde{\sigma}_i(t_i) = (\alpha_i(t_i), q, 1, \ldots)$, where $\alpha_i$ is a deception. Then, $g(\tilde{\sigma}, \sigma_{-i}) = q_{\alpha_i}$. Since, for each $i \in N$, $t_i \in T_i$, and $t_i^* \in T_i$, $q R_i(t_i) q R_i^*$, it holds that $q R_i(t_i) q R_i(t_i) q R_i^*$. It means that $\sigma_i$ is better than $\tilde{\sigma}_i$ given $\sigma_{-i}$.

(Case 2) Let $\tilde{\sigma}_i(t_i) = (\alpha_i(t_i), r, 1, \ldots)$ and $r \neq q$. If there exists $t_i^* \in T_i$ such that $r_{\alpha_i(t_i)} P_i(t_i') q$, then, by the construction of $g$, $g(\tilde{\sigma}, \sigma_{-i}) = q_{\alpha_i}$. Therefore, the above argument implies that $\sigma_i$ is better than $\tilde{\sigma}_i$ given $\sigma_{-i}$. Next we assume that there exists $\hat{t}_i \in T_i$ such that, for each $t_i^* \in T_i$, $q R_i(t_i') r_{\alpha_i(t_i)}$. Then, for each $t_{-i} \in T_{-i}$, $g(\tilde{\sigma}(\hat{t}_i), \sigma_{-i}(t_{-i})) = \ldots$
\[ r(\alpha_i(\hat{t}_i), t_{-i})(1 - \frac{1}{1+n_i}) + z^i(\frac{1}{1+n_i}). \] So \( g(\hat{\sigma}, \sigma_{-i}) \) and \( r(\alpha_i(\hat{t}_i), t_{-i})(1 - \frac{1}{1+n_i}) + z^i(\frac{1}{1+n_i}) \) are equivalent in expected utility at \( \hat{t}_i \). On the other hand, \( q R_i(\hat{t}_i) r_{\alpha_i(t_i)} \) by the assumption. Therefore

\[
q R_i(\hat{t}_i) r_{\alpha_i(t_i)}
\]

\[
r_{\alpha_i(t_i)} R_i(\hat{t}_i) r(\alpha_i(\hat{t}_i), t_{-i})(1 - \frac{1}{1+n_i}) + z^i(\frac{1}{1+n_i}).
\]

It means that \( q R_i(\hat{t}_i) g(\hat{\sigma}, \sigma_{-i}) \).

The above arguments means that, for each \( i \) and \( t_i \), there is no incentive to deviate from \( \sigma_i \). Thus \( \sigma \) is a Bayesian equilibrium. \( \square \)

The following lemma completes the proof.

**Lemma 2.6.10.** If \( F \) satisfies \( (C) \) and \( (BM) \), then, for each Bayesian equilibrium \( \sigma \), there exists \( z \in F \) such that, for each \( t \in \Omega \setminus C^{null} \), \( z = g(\sigma) \).

**Proof.** Let \( \sigma \) be a Bayesian equilibrium for the game \((\Omega, g, M)\). Let \( \alpha_i : T_i \to M_i^1 \) be the deception by the agent \( i \) such that, for each \( t_i \), \( \alpha_i(t_i) = \sigma_i^1(t_i) \).

Let \( B_i^1 \equiv \{ t_i \in T_i \mid \sigma_i(t_i) = (\alpha_i(t_i), r, 1, . ) \} \) and \( B_r \equiv \{ t \in T \mid \forall i \in N, \; \sigma_i(t_i) = (\alpha_i(t_i), r, 1) \} \).

(Case 1) Suppose that, given the other agents’ strategy profiles \( \sigma_{-i} \), for some \( t_i \in T_i \), \( \sigma_i^3(t_i) > 1 \) is the best strategy. Then, for each \( t_{-i} \in T_{-i} \), the resulting message \( \sigma(t_i, t_{-i}) \in D_1 \cup D_2 \). Therefore, under Assumption 2.6.5, the agent \( i \) gets better by increasing \( m_i^3 \) to infinity. It means that \( \sigma_i^3(t_i) > 1 \) cannot be the best response. Thus at equilibrium, \( \sigma_i^3(m_i) = 1 \).
(Case 2) Suppose that, at an equilibrium strategy profile $\sigma$, for some $t_i \in T_i$, $\lambda^*(t_i)[\{t_{-i} \mid \sigma_{-i}(t_{-i}) \neq \sigma_i^2(t_i)\}] > 0$. By the above argument, $\lambda^*(t_i)[\{t_{-i} \mid \sigma_{-i}(t_{-i}) = 1\}] = 1$. In this case, if $\sigma_i^3 = 1$, then $m_i^3 \to \infty$ is better if $\sigma(t) \notin D_0$. And if $\sigma(t) \in D_0$, $i$ gets $\sigma_i^2(t_i)$. Therefore, by the message $\tilde{m}_i(t_i) = (\sigma_i^1(t_i), \sigma_i^2(t_i), \infty, \tilde{m}_i^4)$ such that $\tilde{m}_i^4(\sigma_i^2(t_i)) = \sigma_i^2(t_i)$, he gets better. It cannot be an equilibrium.

As a result, in the equilibrium, it must be that

$$\text{For each } i \in N, \text{ and } t_i \in T_i,$$

$$\exists q \in F, \quad \sigma_i(t_i) = (\alpha_i(t_i), q, 1, \ldots), \text{ and}$$

$$\lambda^*(t_i)[\{t_{-i} \mid \sigma_{-i}(t_{-i}) = (\alpha_{-i}(t_{-i}), q, 1, \ldots)\}] = 1$$

Therefore $\Omega = \bigcup_{r \in F} B_r \cup C_{null}$. Since $F$ is countable, $\bigcup_{r \in F} B_r$ is a countable union of measurable sets, which means both $\bigcup_{r \in F} B_r$ and its complement $t \in C_{null}$ are measurable. By Closure, there exists $\tilde{q} \in F$ such that, for $t \in \Omega \setminus C_{null}$, (1) $\tilde{q}(t) = r(t)$ if $t \in B_r$ and (2) $g(\sigma) = \tilde{q} \circ \alpha$.

We want to show that $\tilde{q} \circ \alpha \in F$. To do this, we use contradiction. Suppose that there is no $z \in F$ such that $z = \tilde{q} \circ \alpha$. Then we apply Bayesian monotonicity to obtain that

$$\exists i \in N, \exists t_i \in T_i, \exists r \in F,$$

$$\forall t'_i \in T_i, \tilde{q} R_i(t'_i) r_{\alpha_i(t_i)} \text{ and } r \circ \alpha P_i(t_i) \tilde{q} \circ \alpha.$$
Therefore,

\[ \tilde{\sigma}_i \equiv \begin{cases} 
\sigma(t_i) & \text{for } \forall t_i \neq \tilde{t}_i \\
(\alpha_i(\tilde{t}_i), r, v_i) & \text{for } \tilde{t}_i 
\end{cases} \]

where \( v_i \) is a sufficiently large number.

Then, this new strategy \( \tilde{\sigma}_i \) is better than \( \sigma_i \) for the agent \( i \) given \( \sigma_{-i} \). It is a contradiction to the assumption that \( \sigma \) is a Bayesian equilibrium. \qed

Next we show that the inverse direction of the theorem.

**Proposition 2.6.11.** If \( F \) is Bayesian implementable on \( \Omega \), it satisfies (C), (BIC), and (BM).

**Proof.** Suppose that a mechanism \( \mathcal{M} \equiv (M, g) \) implements a SCS \( F \) on \( \Omega \). Let

\[ \tilde{F} \equiv \{ z \in F \mid \exists \sigma \text{ s.t. } \sigma \text{ is a Bayesian equilibrium on } \mathcal{M}, \text{ and } z = g(\sigma) \}. \]

Since \( F \) is implemented by \( \mathcal{M} \), it follows that \( F = \tilde{F} \).

Since \( \tilde{F} \) is the set of Bayesian equilibrium, the condition (C) is satisfied. We can show that the condition (IC) is satisfied as follows. Let \( i \in N, q \in F \) and \( \sigma \) be a Bayesian equilibrium such that \( g(\sigma) = q \). Let us consider the agent \( i \)'s strategy \( \hat{\sigma}_i \) such that there exists \( t_i^* \in T_i \) and , for each \( t_i \in T_i, \hat{\sigma}_i(t_i) = \sigma_i(t_i^*) \). Since \( \sigma_i \) is the equilibrium strategy,

\[ \forall t_i \in T_i, g(\sigma) R_i(t_i) g(\hat{\sigma}, \sigma_{-i}). \]
We have $g(\sigma) = q$ and $g(\hat{\sigma}, \sigma_{-i}) = q_{i^*}$. Thus

$$\forall t_i \in T_i, \ q R_i(t_i) q_{i^*}.$$

The condition (BIC) is satisfied.

Next we show (BM). Let $q \in F$ and $\sigma$ be defined as before. Suppose that, for some deception $\alpha$, there is no $z \in F$ such that $z = q \circ \alpha$. Since $\bar{F} = F$, $q \circ \alpha \notin \bar{F}$. Therefore, there exist $t^* \in T$ where, for some agent $i$, $\sigma_i(t^*_i)$ is not the best response. It means that there exists a better response $\bar{m}_i \in M_i$ for the agent $i$ at $t^*_i$ given $\sigma_{-i}$. Let $\bar{\sigma}_i$ be such that, for each $t_i \in T_i$, $\bar{\sigma}_i(t_i) = \bar{m}_i$. And let $r \equiv g(\bar{\sigma}_i, \sigma_{-i})$. Then we have that $g[(\bar{\sigma}_i, \sigma_{-i}) \circ \alpha] P_i(t^*_i) g(\sigma \circ \alpha)$. It means that $r \circ \alpha P_i(t^*_i) q \circ \alpha$. On the other hand, for each $t_i \in T_i$, $g(\sigma) R_i(t_i) g(\bar{m}_i, \sigma_{-i})$ because $\sigma$ is a Bayesian equilibrium. Therefore, for each $t_i \in T_i$, $q R_i(t_i) r$. Note that, for each $t_i \in T_i$, $\bar{\sigma}_i(t_i) = \bar{m}_i$. Therefore, $r_{\alpha_i} = r$. From the above equations, we have that, for each $t_i \in T_i$, $q R_i(t_i) r_{\alpha_i(t^*_i)}$. Thus (BM) is satisfied. $\square$

### 2.7 Conclusion

In this paper, it is shown that robust implementation is equivalent to Bayesian implementation on a particular type space. This result makes the argument much simpler. First we can apply the existing results of Bayesian implementation to robust implementation. We generalize Bergemann-Morris’s result to social choice correspondences by applying Jackson’s result. Second, the space we construct is a knowledge-belief space. It allows us to consider what agents know when designing mechanisms. It is important especially in making regulations in financial markets, firms’ investment activities and so on. Our result gives a foundation for robust implementation in those environments.
Chapter 3

Impossibility of Robust Implementation on Full Domain of Preferences

3.1 Introduction

In this paper, we show the impossibility of robust implementation on the full domain of interdependent preferences. Robust implementation is a belief-free notion in mechanism design, first defined by Bergemann and Morris [7]. Its idea is to find a mechanism which implements a social choice function (SCF) as the unique Bayesian equilibrium outcome\(^1\) no matter what kind of beliefs the agents have. The reason why we have to think about robust implementation is a practical limitation of Bayesian implementation. In order to implement a SCF as the unique Bayesian equilibrium outcome, the planner must identify the belief structures of the agents. Otherwise, he cannot expect what outcome would be achieved as a Bayesian equilibrium of the game resulting from his mechanism. The existing literature suggests that

\(^1\)We refer to the term ‘implementation’ in this complete sense, which is to be contrasted with partial implementation. When a mechanism implements a SCF as one equilibrium outcome among many other equilibrium outcomes, we say that the mechanism \textit{partially implements} the SCF.
a slight misspecification of the agents’ beliefs will bring a significant difference in equilibrium outcomes.\textsuperscript{2} However this requirement seems quite strong. It is usually impossible to observe correctly what kind of beliefs the agents have in their minds. To overcome this limitation of Bayesian implementation, we have to find a mechanism which implements a SCF for every belief structure of the agents. Although among belief-free notions of mechanism design is \textit{strategic proofness}, it has been criticized for being too strong when the agents types (preferences) are interdependent. Bergemann-Morris defined robust implementation so that it is the mildest belief-free mechanism design. Moreover, under interdependent types, robust implementation is milder than the alternative belief-free concept, dominant strategy implementation.\textsuperscript{3}

Bergemann-Morris not only defined robust implementation but also provided a necessary and a sufficient condition for robust implementability. They showed that if a SCF $f$ is robustly implementable, then $f$ satisfies \textit{ex post incentive compatibility} (EPIC) and \textit{robust monotonicity}, and these conditions are sufficient in a wide class of economies. EPIC is the ex post version of (Bayesian) incentive compatibility, stating that it is beneficial for each agent to report his type truthfully as long as the other agents report truthfully, no matter what the others’ types are. Robust monotonicity is a belief-free version of Bayesian monotonicity, which is a necessary condition for Bayesian implementation.\textsuperscript{4} In this paper, we show that, on the full domain of interdependent preferences, the SCFs satisfying robust monotonicity must be only constant functions. This means that, on this domain, only constant SCFs are robustly implementable.

The intuition behind this impossibility theorem is to use a sub property of SCF, \textit{dual dominance} (Saijo [49]). Dual dominance requires the preference domain to be sufficiently rich

\textsuperscript{2}See, among others, [31].
\textsuperscript{3}See [30] and [7].
\textsuperscript{4}See [29].
relative to the range of the SCF. Saijo showed that Maskin monotonic\(^5\) functions satisfying
dual dominance are only constant functions. This result leads to a impossibility theorem of
Nash implementation on the full domain of (independent) preferences. We extend this idea
of dual dominance in order to obtain \textit{robust dual dominance}, and show that only constant
functions satisfy robust monotonicity on the full domain of interdependent preferences.

Among other impossibility results about robust implementation is Jehiel \textit{et al} [30]. Jehiel \textit{et
al} showed that only constant functions satisfies EPIC on a generic class of smooth quasi-linear
interdependent preference domains. Compared to their result, our result does not need any
numeraire goods, and so it can be applied to more general class of economies.

### 3.2 Model

Let \( N \) be the set of agents and \( A \) be the set of deterministic outcomes. We take the lottery
space \( X = \Delta(A) \) as the award space. A social choice function is a function \( f : \Theta \rightarrow X \). The
agents preferences are given by utility functions \( \{u_i\}_{i \in N} \) such that \( u_i : X \times \Theta \rightarrow \mathbb{R} \), where
\( \Theta \equiv \Pi_{i \in N} \Theta_i \) is the product space of agents’ \textit{types}. We assume that \( \Theta \) is a compact metric
space.

For the following arguments, we define a subspace of \( X \) as follows:

\[
\forall i \in N, \forall \theta_{-i} \in \Theta_{-i}, \quad Y_i(\theta_{-i}) \equiv \{y \in X : \forall \theta_i' \in \Theta_i, \ u_i(y, \theta_i', \theta_{-i}) \leq u_i(f(\theta_i', \theta_{-i}), \theta_i', \theta_{-i})\}.
\]

This is a robust version of the agent \( i \)'s lower counter set of alternatives for the allocation
assigned by the SCF \( f \).

\(^5\)Maskin monotonicity is a necessary condition for Nash implementation.
Bergemann-Morris extended Maskin monotonicity to characterize robust implementation as a part of the necessary and sufficient condition of robust implementability. Before going to the definition, we first give preliminary concepts a la Jackson.

**Definition 3.2.1.** A deception is a profile $\beta = (\beta_i)_{i \in N}$, where, for all $i \in N$ and $\theta_i \in \Theta_i$, $\beta_i : \Theta_i \rightarrow \Theta_i$ is non-empty valued and $\theta_i \in \beta_i(\theta_i)$.

**Definition 3.2.2.** A deception is acceptable if $\theta' \in \beta(\theta)$ implies $f(\theta') = f(\theta)$. A deception is unacceptable if it is not acceptable.

**Definition 3.2.3.** (Robust (dual) monotonicity [7], hereafter (RM)) A social choice function $f$ satisfies robust (dual) monotonicity if, for any unacceptable deception $\beta$,

$$
\exists i \in N, \ \theta_i \in \Theta_i, \ \text{and} \ \theta_i' \in \beta_i(\theta_i) \ \ s.t.
$$

$$
\forall \theta'_{-i} \in \Theta_{-i}, \ \exists y \in Y_i(\theta'_{-i}),
$$

$$
\forall \theta_{-i} \in \beta_{-i}^{-1}(\theta'_{-i}), \ u_i(y, \theta_i, \theta_{-i}) > u_i(f(\theta_i', \theta'_{-i}), \theta_i, \theta_{-i}). \quad (3.1)
$$

3.3 Necessary and sufficient condition for impossibility

Next we show a necessary and sufficient condition about when only constant functions satisfy robust monotonicity. For this purpose, we introduce a new property of social choice functions;

**Definition 3.3.1.** (Robust dual dominance, hereafter (RDD)) A social choice function $f$
satisfies robust dual dominance if the following condition is satisfied:

For $\forall i \in N$, $\forall \theta'_i, \theta''_i \in \Theta_i$,

\[ \exists \theta_i \in \Theta_i \text{ s.t.} \]
\[ \forall \theta_{-i} \in \Theta_{-i}, \ u_i(f(\theta'_i, \theta_{-i}), \theta_i, \theta_{-i}) \geq u_i(f(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}), \quad (3.2) \]
\[ u_i(f(\theta''_i, \theta_{-i}), \theta_i, \theta_{-i}) \geq u_i(f(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}). \quad (3.3) \]

RDD implies that the domain of preferences is sufficiently large compared to the range of allocations.

Then we have the next theorem.

**Theorem 3.3.2.** A social choice function $f$ is constant if and only if $f$ is robust monotonic and robust dual dominant.

**Proof.** Since any deception $\beta$ is acceptable under constant social choice functions, the ‘only if’ part is clear. So we prove the ‘if’ part.

Fix $i \in N$ and $\theta_i, \theta'_i, \theta''_i \in \Theta_i$ so that (3.2) and (3.3) are satisfied. Let $\beta$ be a deception as follows:

For the agent $i$,

\[ \beta_i(\theta_i) = \{\theta_i, \theta'_i\} \]
\[ \forall \tilde{\theta}_i \neq \theta_i, \ \beta_i(\tilde{\theta}_i) = \tilde{\theta}_i, \]

For each agent $j \neq i$,

\[ \forall \tilde{\theta}_j \in \Theta_j, \ \beta_j(\tilde{\theta}_j) = \{\tilde{\theta}_j\}, \]

This means that all agents except for $i$ always tell the truth, but the agent $i$ may pretend to
be at $\theta'_i$ when his true type is $\theta_i$. In the same way, we define another deception $\beta'$:

For the agent $i$,

$$\beta_i(\theta_i) = \{\theta_i, \theta''_i\} \quad \forall \tilde{\theta}_i \neq \theta_i, \beta_i(\tilde{\theta}_i) = \tilde{\theta}_i,$$

For each agent $j \neq i$,

$$\forall \tilde{\theta}_j \in \Theta_j, \beta_j(\tilde{\theta}_j) = \{\tilde{\theta}_j\}.$$

Here the agent $i$ may pretend to be at $\theta''_i$ when his true type is $\theta_i$.

Next we see if the equality condition in (3.1) is satisfied in $\beta$. First consider the agent $i$. Since, for each agent $j \neq i$ and each $\tilde{\theta}_j \in \Theta_j$, $\beta_j(\tilde{\theta}_j) = \{\tilde{\theta}_j\}$, we have

$$\forall \tilde{\theta}_i \in \Theta_{-i}, \beta_{-1}^{-1}(\tilde{\theta}_{-i}) = \tilde{\theta}_{-i}.$$

And by construction of $Y$,

$$\forall \tilde{\theta}_{-i} \in \Theta_{-i}, \forall y \in Y(\tilde{\theta}_{-i}),$$

$$u_i(y, \theta_i, \tilde{\theta}_{-i}) \leq u_i(f(\theta_i, \tilde{\theta}_{-i}), \theta_i, \tilde{\theta}_{-i}),$$

$$\leq u_i(f(\theta'_i, \tilde{\theta}_{-i}), \theta_i, \tilde{\theta}_{-i}).$$

The last inequality is derived from (3.2). The above result implies that the inequality in (3.1) is not satisfied for the agent $i$.

Next consider an agent $j \neq i$. So, for all $\tilde{\theta}_j \in \Theta_j$, $\beta_j(\tilde{\theta}_j) = \{\tilde{\theta}_j\}$. And, for all $\tilde{\theta}_i \neq \theta_i$.
and all $\tilde{\theta}_{-(i,j)} \in \Theta_{-(i,j)}$, $\beta_{-j}(\tilde{\theta}_i, \tilde{\theta}_{-(i,j)}) = \{(\tilde{\theta}_i, \tilde{\theta}_{-(i,j)})\}$. Therefore we have

$$\forall \tilde{\theta}_j \in \Theta_j, \forall \tilde{\theta}'_j \in \beta_j(\tilde{\theta}_j),$$

$$\forall y \in Y_j(\tilde{\theta}_i, \tilde{\theta}_{-(i,j)}),$$

$$u_j(y, \tilde{\theta}_j, \tilde{\theta}_i, \tilde{\theta}_{-(i,j)}) \leq u_j(f(\tilde{\theta}_j, \tilde{\theta}_i, \tilde{\theta}_{-(i,j)}), \tilde{\theta}_i, \tilde{\theta}_{-(i,j)})$$

$$\leq u_j(f(\tilde{\theta}'_j, \tilde{\theta}_i, \tilde{\theta}_{-(i,j)}), \tilde{\theta}_i, \tilde{\theta}_{-(i,j)}).$$

Therefore, the inequality in (3.1) is not satisfied for the agent $j$.

The above result means that the inequality condition in (3.1) is not satisfied for $\beta$. Thus $\beta$ is acceptable, i.e., for all $\tilde{\theta}_{-i} \in \Theta_{-i}$, $f(\theta_i, \tilde{\theta}_{-i}) = f(\theta'_i, \tilde{\theta}_{-i})$. We can obtain by the same logic that, for all $\tilde{\theta}_{-i} \in \Theta_{-i}$, $f(\theta_i, \tilde{\theta}_{-i}) = f(\theta''_i, \tilde{\theta}_{-i})$. It means that, for all $\tilde{\theta}_{-i} \in \Theta_{-i}$, $f(\theta'_i, \tilde{\theta}_{-i}) = f(\theta''_i, \tilde{\theta}_{-i})$.

Since $i$, $\theta'_i$, and $\theta''_i$ are arbitrary, the above results implies that $f$ is constant. \hfill $\square$

We say that a type $\theta_i \in \Theta_i$ is the flat preference if, for all $\tilde{\theta}_{-i} \in \Theta_{-i}$ and all $x \in X$, $u_i(x, \theta_i, \tilde{\theta}_{-i}) = c$ where $c$ is a constant number. Theorem 3.3.2 implies that only constant functions are robustly implementable if each agent has the flat preference.

Next we define the full domain of interdependent preferences as a natural extension of the full domain of independent preferences. However, in contrast to independent preferences, it is not so straightforward to construct the full domain of interdependent preferences because the agents’ preferences are determined in the entire structure of the agents’ types. Therefore we define a sufficiently large type structure where any combination of the agents’ preferences are represented. It corresponds to the full domain of preferences in the independent preference case.
First let $\mathcal{R}$ be the space of complete and transitive binary relationships over $X$. And let $\mathcal{C}(Y, Z)$ be the space of continuous functions from $Y$ to $Z$. Suppose that $\Theta_i = \Theta_j$ for all $i, j \in N$. We say that $\{u_i\}_{i \in N}$ is symmetric if, for each $i$ and $j$, and each $\theta_i$, and $\theta_j$, $u_i(., \theta_i, \theta_j, \theta_{-i,j}) = u_j(., \theta_j, \theta_i, \theta_{-i,j})$. We use the notation $u_{\theta_i}$ for the function $u_{\theta_i} : \Theta_{-i} \to \times_{i \in N} \mathcal{R}$ such that $u_{\theta_i}(\theta_{-i})$ is the representation of $(u_i(., \theta_i, \theta_{-i}), u_{-i}(., \theta_i, \theta_{-i}))$.

**Definition 3.3.3.** Let $\Theta^* \equiv \Pi_{i \in N} \Theta_i^*$, where, for all $i, j \in N$, $\Theta_i^* = \Theta_j^*$, and $\{u_i\}_{i \in N}$ be continuous and symmetric. The type space $\Theta^*$ is a large domain of interdependent preferences over $X$ if, for each $i \in N$ and each $g \in \mathcal{C}(\Theta_{-i}^*, \times_{i \in N} \mathcal{R})$, there exists $\theta_i$ such that $u_{\theta_i} = g$.

Then, $\Theta^*$ includes the flat preference. Thus we obtain the next result as a corollary.

**Corollary 3.3.4.** In large domains of interdependent preferences $\Theta^*$, any robust implementable social choice function is constant.

### 3.4 Comments on the impossibility in smaller domains

Jehiel, et al [30] showed that if the preference domain is quasi-linear and satisfies a certain differentiability condition, then generically only constant functions satisfy *ex post incentive compatibility* (EPIC), which is a necessary condition of robust implementation.

It corresponds to the impossibility theorem in a sub-domain. However, it does not necessarily implies the impossibility of robust implementation in larger domains.

---

7 Some components of the canonical interdependent preference space in Gul and Pesendorfer [24] satisfy this condition.

8 See [7]
To clarify the argument, we first define EPIC and what is a sub domain of an interdependent preference domain.9

**Definition 3.4.1.** A SCF $f$ satisfies EPIC if, for each $i \in N$, each pair $\theta_i, \theta'_i \in \Theta_i$, and each $\theta_{-i} \in \Theta_{-i}$, $u_i(f(\theta_1, \theta_{-i}), \theta_i, \theta_{-i}) \geq u_i(f'(\theta'_1, \theta_{-i}), \theta_i, \theta_{-i})$.

**Definition 3.4.2.** Let $\langle \Theta, \{u_i\}_{i \in N} \rangle$ be an interdependent preference domain over $X$. An interdependent preference domain $\langle \Theta', \{u'_i\}_{i \in N} \rangle$ is a sub-domain of $\Theta$ if there exists a homeomorphic embedding $h : \Theta' \rightarrow \Theta$ such that, for each $i \in N$ and each $\theta' \in \Theta'$, $u'_i(., \theta')$ and $u_i(., h(\theta'))$ represents the same preference over $X$.

The fact that only constant functions satisfy EPIC in a sub-domain does not necessarily imply the impossibility in a larger domain. We provide an example below.

### 3.4.1 Two person example

Let $N = \{1, 2\}$. Consider two interdependent preference domains, Domain 1 = $\langle \Theta_1, \Theta_2, (u_i)_{i=1,2} \rangle$ and Domain 2 = $\langle \Theta_1, \Theta'_2, (u_i)_{i=1,2} \rangle$. We can combine Domain 1 and Domain 2 to obtain a larger domain, Domain 3 = $\langle \Theta_1, (\Theta_2 \cup \Theta'_2), (u_i)_{i=1,2} \rangle$, where Domain 1 and 2 are sub-domains. Suppose that only constant functions satisfy EPIC in Domain 1 but there exists a non-constant SCF $f$ satisfying EPIC in Domain 2. We show when there exists a non-constant SCF satisfying EPIC in Domain 3.

---

9Concerning the preference domains $\Theta'$ and $\Theta$ in Definition 3.4.2, Gul and Pesendorfer [24] consider that these domains belong to different interdependent preference systems, and they do not consider $\Theta'$ to be a part of $\Theta$. 
Lemma 3.4.3. If there exists $\tilde{x} \in X$ such that,

$$
\forall y \in X,
\begin{aligned}
& u_2(\tilde{x}, \theta) \geq u_2(y, \theta) \text{ if } \theta \in \Theta_1 \times \Theta_2, \\
& u_2(\tilde{x}, \theta) \leq u_2(y, \theta) \text{ if } \theta \in \Theta_1 \times \Theta'_2,
\end{aligned}
$$

then there exists a non-constant SCF satisfying EPIC in Domain 3.

Proof. Let $g : \Theta_1 \times (\Theta_2 \cup \Theta'_2) \rightarrow X$ be such that

$$
g(\theta) = \begin{cases} 
\tilde{x} & \text{if } \theta \in \Theta_1 \times \Theta_2 \\
f(\theta) & \text{if } \theta \in \Theta_1 \times \Theta'_2
\end{cases}
$$

Since $f$ satisfies EPIC, for each $\theta_1, \theta'_1 \in \Theta_1$ and each $\theta_2 \in \Theta'_2$, we have $u_1(g(\theta_1, \theta_2), \theta_1, \theta_2) \geq u_1(g(\theta'_1, \theta_2), \theta_1, \theta_2)$. And, in $\Theta_2$, $g$ is constant. Therefore, we have

$$
\forall \theta_1, \theta'_1 \in \Theta_1, \forall \theta_2 \in (\Theta_2 \cup \Theta'_2),
\begin{aligned}
& u_1(g(\theta_1, \theta_2), \theta_1, \theta_2) \geq u_1(g(\theta'_1, \theta_2), \theta_1, \theta_2).
\end{aligned}
$$

For the agent 2, we have to consider three cases: (1) $\theta_2 \in \Theta_2$ and $\theta'_2 \in \Theta'_2$, (2) $\theta_2, \theta'_2 \in \Theta_2$, and (3) $\theta_2, \theta'_2 \in \Theta'_2$.

In case (1), by the assumption, $\tilde{x}$ is the best allocation for the agent 2 at $\theta_2$. Therefore, for each $\theta_1 \in \Theta_1$, $u_2(g(\theta_1, \theta_2), \theta_1, \theta_2) \geq u_2(g(\theta_1, \theta'_2), \theta_1, \theta_2)$. On the other hand, $\tilde{x}$ is the worst allocation for $\theta'_2$, so we have, for each $\theta_1 \in \Theta_1$, $u_2(g(\theta_1, \theta'_2), \theta_1, \theta'_2) \geq u_2(g(\theta_1, \theta_2), \theta_1, \theta'_2)$.

In case (2), $g$ is constant, so, for each $\theta_1 \in \Theta_1$, $u_2(g(\theta_1, \theta_2), \theta_1, \theta_2) = u_2(g(\theta_1, \theta'_2), \theta_1, \theta_2)$. And in case (3), $g = f$ on that sub-domain, and $f$ satisfies EPIC. Therefore, for each $\theta_1 \in \Theta_1$, 

\[ u_2(g(\theta_1, \theta_2), \theta_1, \theta_2) \geq u_2(g(\theta_1, \theta'_2), \theta_1, \theta_2). \]

To summarize the above arguments, we obtain

\[ \forall \theta_1 \in \Theta_1, \forall \theta_2, \theta'_2 \in (\Theta_2 \cup \Theta'_2), \]
\[ u_2(g(\theta_1, \theta_2), \theta_1, \theta_2) \geq u_2(g(\theta_1, \theta'_2), \theta_1, \theta_2). \]

Thus \( g \) satisfies EPIC on Domain 3. \( \square \)

### 3.5 Conclusion

In this paper, we provided a necessary and sufficient condition for constancy of a SCF, using a novel concept, robust dual dominance. As a direct implication of this result, we showed the impossibility of robust implementation in large domains of interdependent preferences. With independent preferences, our result corresponds to the well known impossibility of dominant strategy implementation in the full domain of preferences. Although there are other results about the impossibility of robust implementation in restricted domains of preferences, including Jehiel et al [30], our result is not logically implied by them. Moreover our result is more general because, in contrast to [30], numeraire goods are not required.
Bibliography


