Essays on Game Theory

by

Taejun Oh

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2016

Doctoral Committee:

Professor Tilman M. Börgers, Chair
Professor Kfir Eliaz, Tel-Aviv University
Assistant Professor Heng Liu
Assistant Professor David A. Miller
Professor Scott E. Page

Acknowledgments

I am extremely grateful to my advisor, Tilman Börgers. I am also very grateful to my committee members Kfir Eliaz, Heng Liu, David Miller, and Scott E. Page.

This thesis has benefited from discussions with the faculty and students at University of Michigan and other institutions. My appreciation goes to Andreas Blume and Francoise Forges for their valuable comments. I also thank for seminar participants at University of Michigan.

I have been fortunate to receive generous financial support from University of Michigan with the Fred M. Taylor Award in Economic Theory and Rackham One-Term Graduate Fellowship.

Contents

List of Figures

Abstract

I study how to model various strategic interactions with incomplete information and how to properly analyze them. The first chapter of the dissertation suggests new solution concepts for incomplete information games that are preceded by communication opportunities. Second chapter, written with Tilman Börgers, studies type spaces that model incomplete information among players with a specific property: the independence property. Third chapter suggests a new behavioral model to study real world decision makers' behavior in specific games.

In the first chapter, I suggest new solution concepts for incomplete information games that are preceded by communication opportunities (communication games) in order to solve the following problem: traditional equilibrium analysis for communication games does not properly incorporate and explain uncertainties about players' communication strategies, and so it excludes plausible outcomes of communication games from the set of equilibrium outcomes. Thus, I define *correlated cheap talk equilibrium* for games with cheap talk opportunities, and *correlated communication equilibrium* for games with general communication opportunities. New definitions suggested in this paper are extensions of Aumann's (1974) definition of correlated equilibrium for complete information games to communication games. Following Aumann's (1987) epistemic justification of the correlated equilibrium, I provide epistemic justifications of new definitions. That is, new definitions represent

the common knowledge of the Bayesian rationality of players with a common prior for a given communication game. Then I compare new definitions with other correlated equilibrium definitions for incomplete information games.

In the second chapter, we study common prior type spaces in which for each agent the agent's payoff type and the agent's belief type are independent. Such type spaces deserve attention as the polar opposites of common prior type spaces in which agents' beliefs determine their preferences - a class of type spaces whose special properties are much studied. We find a necessary and sufficient condition for the independence of each agent's payoff type and belief type. Different agents' payoff types must be independent. Agents may hold payoff irrelevant information. The payoff irrelevant signals that agents receive may be correlated with each other, but they must be jointly independent of all agents' payoff types. We conclude that type spaces with independent payoff types, as commonly used in game theory and mechanism design, constitute, up to payoff irrelevant information, the class of *all* type spaces in which payoff types and belief types are independent for each agent.

In the third chapter, I suggest a modified cognitive hierarchy (CH) model which improves on the original CH model by Camerer, Ho and Chong (2004). Players are endowed with cognitive types which characterize their iterative reasoning abilities. Unlike the original CH model, the cognitive type space has a common prior, so players know the existence of equal or higher type players. In order to estimate equal or higher type players' choices, which are impossible for correct anticipation, players use other players' past choice data. Players utilize additional information and improve their estimation if all other players' individual choice data are given, compared to the case where only other players' aggregate choice data are given. Some experimental findings support the modified CH model. First, Sbriglia's (2009) experimental findings about repeated beauty contest game suggest that the modified CH model improves on the original CH model. Second, Duffy and Hopkins' (2005) experimental findings show that the modified CH model might provide better explanation about repeated market entry game compared to other learning models, the replicator dynamics and the fictitious play.

Chapter 1

Correlated Cheap Talk and Correlated Communication Equilibrium

1.1 Introduction

A communication game is a 2-stage game in which a simultaneous move incomplete information game is preceded by a communication opportunity among players. There are 2 sources of uncertainties associated with informative messages in communication games. For example, suppose the Fed announces that "the U.S. labor market conditions are nearly normal," which is an informative message about relevant uncertainty - in this case the U.S. unemployment rate. Consider a 2 stage game in which the Fed sends a message, then observers of the Fed's message choose actions. In that game, an observer cannot pinpoint what the Fed knows about the unemployment rate after observing the above message for two reasons. First, one might not know in which context the Fed uses the message "nearly normal": the Fed might say nearly normal when the unemployment rate is inside the range of 1-3 percent or inside the range of 2-4 percent. That is, an observer might be uncertain about the Fed's messaging strategy, so he might be uncertain about how to interpret a given message. Second, even if an observer knows the Fed's messaging strategy and the exact range of unemployment rates associated with "nearly normal," the exact unemployment rate inside a given range is unknown. In the 2-stage game between the Fed and observers, both types of uncertainties might be relevant when players choose optimal strategies.

In the literature, existing works analyze either the perfect Bayesian equilibrium or the sequential equilibrium for communication games.[1] For example, Crawford and Sobel (1982) analyzed the cheap talk equilibrium, which is the perfect Bayesian equilibrium of a cheap talk game: a 2-stage game in which a 2-player incomplete information game is preceded by a cheap talk communication opportunity. However, analyzing the perfect Bayesian equilibrium of communication games implies that there is no uncertainty about messaging strategies. For example, in the communication game just described between the Fed and observers, in any perfect Bayesian equilibrium, there is no uncertainty about the Fed's messaging strategy. That is, in any perfect Bayesian equilibrium, the exact range associated with the message "nearly normal" is known and fixed. Likewise, in the equilibrium analysis provided by Crawford and Sobel (1982), only the second source of uncertainties is incorporated and explained by the cheap talk equilibrium, which is the perfect Bayesian equilibrium of a given cheap talk game.

In this paper, I show that the first source of uncertainties, the uncertainties about messaging strategies, can be incorporated into formal equilibrium analysis by considering special forms of correlated equilibrium for

[1]The sequential equilibrium is a refinement of the perfect Bayesian equilibrium for general extensive form games, resulting in a smaller set of equilibrium outcomes. This paper is not about equilibrium refinement from perfect Bayesian equilibrium to sequential equilibrium; rather it is about introducing new equilibrium definitions that give larger sets of equilibrium outcomes than the perfect Bayesian equilibrium.

communication games. Accordingly, I suggest two new solution concepts for multi-player incomplete information games, *correlated cheap talk equilibrium* for incomplete information games extended by cheap talk opportunities[2] and *correlated communication equilibrium* for incomplete information games extended by general communication opportunities.[3] Also, analyzing such correlated equilibria of communication games will generate strictly larger equilibrium predictions compared to the equilibrium prediction based on the perfect Bayesian equilibrium. That is, introducing new solution concepts enables us to have new equilibrium outcomes, which were not analyzed and explained by the traditional equilibrium analysis of communication games.

In communication games, larger sets of equilibrium outcomes are provided by analyzing the correlated cheap talk equilibrium or the correlated communication equilibrium, rather than the perfect Baysian equilibrium of communication games. New definitions introduced in this paper are extensions of Aumann's (1974) correlated equilibrium for one-shot complete information games to 2-stage incomplete information games with communication opportunities. After introducing new equilibrium definitions, I give epistemic justifications for new definitions, thereby extending Aumann's (1987) epistemic justification for correlated equilibrium for complete information games to communication games. Detailed discussions are provided in sections 3 and 4.

I start the discussion of this paper by introducing a 2-player incomplete information game preceded by a cheap talk communication opportunity in which all perfect Bayesian equilibria entail no meaningful communication between players. Due to possible uncertainties about messaging strategies, a new equilibrium arises with meaningful communication when we analyze a

[2]By cheap talk, I refer to a communication protocol in which, prior to the given incomplete information game, all players simultaneously exchange one-shot public messages.

[3]The general communication opportunities represent every possible communication protocol that might precede a given incomplete information game.

correlated equilibrium of the given communication game.

A motivating example

Consider an incomplete information game as follows. There are two players, S, the speaker (male), and L, the listener (female). The speaker, who is the manager of a firm, privately observes $\theta \in \Theta \equiv \{nc, c\}$, where Θ is the set of relevant states of the world. Θ is associated with common prior $p = (0.6, 0.4)$, where 0.6 is the probability of nc. State c represents an imminent crisis for the business, and nc represents no such crisis. The listener is the owner of the firm without private information. She chooses an action from the set $A \equiv \{sq, f, supp\}$, where sq represents maintaining the status quo, f represents firing the manager, and $supp$ represents providing more support to the manager.

The speaker's utility does not depend on $\theta \in \Theta$. It is given by $u_S : A \to \mathbb{R}$ such that $u_S(sq) = 3,\ u_S(f) = 0,\ u_S(supp) = 7$. That is, the manager likes to have more support from the owner of the firm, dislikes being fired, and ranks the status quo between these two outcomes.

The listener's utility $u_L : \Theta \times A \to \mathbb{R}$ depends on both the state of the world and the chosen action as follows: $u_L(nc, sq) = 4,\ u_L(nc, f) = 0,\ u_L(nc, supp) = 3,\ u_L(c, sq) = 0,\ u_L(c, f) = 4,\ u_L(c, supp) = 2$. If there is no crisis, the owner prefers maintaining the status quo to give more support to the manager. Firing the manager will disrupt the business, and so it is least preferred. However, if there is an imminent crisis for the business, the owner prefers to fire the manager. If the owner does not fire the manager, she wants to give the manager more support in response to the crisis. Note that if the owner is uncertain about the state of the world - that is, if she has some intermediate beliefs - it may be optimal to provide more support to the manager.

Now let the above game be extended by a cheap talk opportunity à la

Crawford and Sobel (1982), where the manager speaks first and then the owner chooses an action. It is easy to see that in all perfect Bayesian equilibria of this game, the owner chooses an action strategy regardless of received messages. No communication is possible because the speaker's utility does not depend on the state of the world.[4] However, I suggest an equilibrium in which fully rational players use pure strategies with meaningful communication and players are possibly uncertain about meanings and interpretations of messages.

First, I give a formal description of such an equilibrium, and then I provide an intuitive explanation. I will consider a form of correlated equilibrium of the cheap talk game, in which a mediator gives out extra signals to players before players play the cheap talk game. After the original state of the world $\theta \in \Theta$ is realized, a mediator, knowing the state, sends extra signals to players as follows. When the state of the world is nc, the speaker observes a signal from the set $\{1, 2\}$, and the listener observes a signal from the set $\{\ell, r\}$. The signals are drawn from the common prior shown in Figure 1.

	ℓ	r
1	$\frac{2}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$

Figure 1.1: Distribution of extra signals conditional on state nc

When the state of the world is c, then the speaker always observes signal

[4]There could be equilibria in which different speaker types choose different mixed strategies, but this would require that the listener's updated conditional beliefs over messages induce same action strategy for all received messages. In this case, speaker types have different strategies and messages contain some information, but that information is not utilized in the equilibrium. I also describe such equilibria as equilibria that lack meaningful communication.

$\{3\}$, and the listener observes a signal from the set $\{\ell, r\}$ as in Figure 2.

	ℓ	r
3	$\frac{1}{2}$	$\frac{1}{2}$

Figure 1.2: Distribution of extra signals conditional on state c

The joint distribution of states of the world and additional signals can be represented by the extended type space depicted in Figure 3, where the speaker observes the row and the listener observes the column. Note that observations of additional signals do not change the beliefs of players about the underlying states of the world, Θ. For the speaker, this is so because he knows the state of the world. For the listener, the probability of receiving ℓ or r is the same in both underlying states of the world. Therefore, listener types ℓ and r share same beliefs about Θ, which is the same as the common prior p. In the literature this property of additional signals is called "belief invariance." Belief invariance of extra signals implies that in the extended type space, as in Figure 3, listener types ℓ and r, and speaker types $nc, 1$ and $nc, 2$, share the same belief hierarchies over Θ. I have thus introduced what the literature calls "redundant" types, which describe types that have the same belief hierarchies in a type space.[5]

I describe an equilibrium of the cheap talk game wherein players receive the information described in Figure 3. In contrast to any perfect Bayesian

[5] A player's belief over Θ is called the first order belief, and a player's belief over the set of possible beliefs over Θ by the other player together with his first order belief is called the second order belief. Likewise, we can define the nth order belief for any natural number n, and a combination of all possible nth order beliefs is called a belief hierarchy of a player with respect to Θ. In a type space, two types that share the same belief hierarchy are called redundant types. Types ℓ and r, and $(nc, 1)$ and $(nc, 2)$ are redundant types in the type space given by Figure 3.

	ℓ	r
$nc, 1$	0.2	0.1
$nc, 2$	0.1	0.2
$c, 3$	0.2	0.2

Figure 1.3: An extended type space with belief invariant extra signals to both players

equilibrium of the cheap talk game without extension to the type space, in the following equilibrium meaningful information transmission does occur.

An Equilibrium: *The speaker sends message nw when he observes that the state of the world is nc, and when he also receives the additional signal* 1*. Otherwise, the speaker sends the message* $w \neq nw$*. The listener, when observing message nw, understands that the state of the world is nc and therefore she maintains the status quo (action sq). If the listener observes message* w*, and receives the additional signal* ℓ*, she fires the manager (action* f*), whereas, if she observes message* w *and her own additional signal is* r*, then she will provide additional support to the manager (action supp).*

Define an equilibrium outcome of an incomplete information game as a function that maps each realization of the state of the world to a probability distribution on the set of actions. The above equilibrium induces an equilibrium outcome as described in Figure 4. In contrast, in any cheap talk equilibrium without extra signals for players, the listener's action is always be *supp* because there could be no communication.

Intuitively, in this equilibrium, diverse beliefs about the speaker's messaging strategy and the meaning of the message w induce meaningful information transmission. Interpret the message w as the warning, "there might be a cri-

	sq	f	$supp$
nc	1/2	1/6	1/3
c	0	1/2	1/2

Figure 1.4: An equilibrium outcome of given game: $\Theta \to \Delta A$

sis", because it is sent in both states nc (no crisis) or c (crisis). Interpret nw as no warning (i.e, "there is no crisis") because it is sent only in the state nc (no crises). Although there is no uncertainty about the meaning of nw (no warning), the message w (warning) is intended and interpreted differently by players with different extra signals. Listener type ℓ takes this warning w seriously with the updated belief on Θ as (1/3, 2/3) after observing message w, and she optimally responds to w by firing the manager. On the other hand, the listener type r with updated belief (1/2,1/2) after w, considers the warning more lightly, and she optimally chooses to provide more support to the manager. Note that listener types ℓ and r share the same belief hierarchy about Θ, so that they are redundant types. However, they choose different action strategies because they have different beliefs about the joint space of Θ and speaker's messaging strategies.

Then check speaker types' incentives on the suggested equilibrium. Speaker type $nc, 1$ says there is no crisis by nw because he believes that if he sends a warning with message w, the listener will be more likely to panic (being type ℓ) and fire him.[6] In contrast, speaker type $nc, 2$ sends a warning through message w because he anticipates that, given w, the listener is more likely to doubt the warning (being type r) and respond with moderate ac-

[6] Note that speaker type $nc, 1$ believes that the listener is more likely to be type ℓ. That is, $nc, 1$ type believes that it is more likely that the listener believes that the state of the world is probably c given message w, and so chooses f when she received the message w.

tion *supp*, which is the speaker's preference. Like speaker type $nc, 2$, speaker type $c, 3$ sends a warning by the message w with similar incentives. Note that speaker types $nc, 1$ and $nc, 2$ share the same belief hierarchy about Θ, so that they are redundant types. However, they choose different messaging strategies because they have different beliefs about the joint space of Θ and listener's action strategies. Also note that speaker types' beliefs about the joint space of Θ and listener's action strategies come from the speaker types' beliefs about the listener types' beliefs about the speaker types' messaging strategies.

In sum, I introduce a mediator who gives out extra signals that depend on the state $\theta \in \Theta$, but the extra signals do not give additional information about Θ to players. Introducing belief invariant extra signals create new equilibrium with meaningful communication. Those extra signals represent the players' uncertainties about Θ and each other's strategies (by representing listener's beliefs about speaker's messaging strategy and speaker's belief about listener action strategies given messages) on the suggested equilibrium. Therefore, the mediator and additional signals do not constitute a real mediator that send out actual signals that players observe; instead, they are modeling devices that represent players' beliefs about Θ and each other's strategies. Also, we can see that the extra signals, which represent players' beliefs about Θ and players' strategies, must not give additional information about Θ to players. Thus, the belief invariance condition seems to be a natural requirement for cheap talk games.

The belief invariant extra signal structure is introduced by Liu (2015) when he defined belief invariant Bayes correlated equilibrium for simultaneous-move incomplete information games,[7] which is one of many possible corre-

[7]Liu (2015) called his new equilibrium the correlated equilibrium for incomplete information games. To avoid confusion with other correlated equilibrium definitions, I follow Bergemann and Morris (2015) and call Liu's definition the belief invariant Bayes correlated

lated equilibrium definitions for simultaneous move incomplete information games.[8] My research applies extra signal structures with belief invariance to 2-stage games in which incomplete information games are preceded by communication opportunities: communication games. Detailed discussions and comparisons with Liu's (2015) equilibrium concept, as well as other equilibrium concepts, are provided in section 5.

Note that I could have applied extra signals with different characteristics to the suggested cheap talk game. That is, different assumptions regarding the correlation between the original and the extra signals could generate different types of extra signal structure for a cheap talk game. The intuitive reason of choosing extra signals that are possibly correlated to Θ with belief invariance for cheap talk games is explained above: extra signals represent players' beliefs about Θ and other players' strategies, so that they must not give more information about Θ. I provide formal justification of the belief invariant extra signal structure for cheap talk games in the epistemic analysis (section 3 and 4).

In the real world, uncertainties about the meanings of transmitted messages are prevalent and diverse beliefs about messaging strategies do matter in strategic situations with communication. Informative messages are coded to allow diverse interpretations, sometimes intentionally, or their meanings can be open to interpretation because of the limits of the language. Examples of informative messages with possibly diverse interpretations include firms' business plans and financial statements, rumors in the financial markets, credit ratings published by credit rating agencies, pledge made by political

equilibrium.

[8]In Aumann's (1987) definition of the correlated equilibrium for complete information games, players are allowed to observe additional signals before choosing actions. For incomplete information games, one can define various correlated equilibrium concepts that depend on the relationship between original uncertainty and extra signals. The original and extra signals could be independent, arbitrarily correlated, or have the belief invariance in the middle of the two extremes.

candidates, and oil reserve estimations by firms and governments.

Generalizing the example

In this paper, I generalize in several ways intuitions from the example described above. First, I define **correlated cheap talk equilibrium** for general multi-player incomplete information games. I define cheap talk as a simultaneous one-shot public message exchange that transpires before players choose actions. A correlated cheap talk equilibrium is a version of the correlated equilibrium of a cheap talk extension of an incomplete information game, wherein players observe belief invariant extra signals before exchanging cheap talk messages and choosing actions.

I provide epistemic justification for the correlated cheap talk equilibrium. The epistemic analysis shows that the correlated cheap talk equilibrium is the proper equilibrium concept for incomplete information games preceded by cheap talk communication opportunities. Some solution concepts have been justified by epistemic analysis. For example, Aumann (1987) provides a rationale for correlated equilibrium with complete information games. He shows that assumptions of the common knowledge of the Bayesian rationality of players with a common prior imply correlated equilibrium outcomes in a given complete information games. I conduct the same analysis for incomplete information games that are preceded by a cheap talk communication opportunity. Proposition 1 states that the correlated cheap talk equilibrium represents the common knowledge of the Bayesian rationality of players with a common prior in a given cheap talk game. Therefore, we should consider the correlated cheap talk equilibrium if we wish to understand the reasonable equilibrium expectations of a given incomplete information game that are preceded by a possible cheap talk communication opportunity.

The epistemic analysis for the correlated cheap talk equilibrium also tells us how to interpret extra signals and belief invariance in the definition of the correlated cheap talk equilibrium. Extra signals represent players' beliefs about the state of the world and players' strategies. The belief invariance of extra signals translates into an important informational assumption of games with cheap talk communication opportunities as follows: players' information about Θ comes only from their original information about Θ prior to communication and from transmitted messages regarding cheap talk communication.

Next, all previous intuitions are applied to general mediated communication. Moving away from one-shot simultaneous communication, Forges (1986, 1990, 2006) has defined communication equilibrium for incomplete information games. Her objective is to characterize equilibrium outcomes of incomplete information games that are preceded by arbitrary communication protocols. A communication game is a 2-stage extensive form game in which a mediator invites players to report their types, then gives out action recommendations to players according to predetermined rules. Forges (1990) shows that such mediated communication can represent any possible communication protocol before an incomplete information game, and she identifies a communication equilibrium of an incomplete information game as the perfect Bayesian equilibrium of a communication extension of a given game. Note that the communication equilibrium does not incorporate and explain players' uncertainties about each others' messaging strategies because it is defined to be the perfect Bayesian equilibrium of a given communication game. That is, the equilibrium suggested in the motivating example is not a communication equilibrium of suggested game. Therefore, I define **correlated communication equilibrium** for multi-player incomplete information games by adding belief invariant extra signals prior to the initiation of mediated communication. An epistemic justification of correlated communication equilibrium shows that the new equilibrium represents common knowledge of the Bayesian rationality of players with a common prior given a communication extension of an incomplete information game. Extra signals represent players' beliefs about Θ and all players' communication and action strategies. Belief invariance represents the fact that players' information about Θ comes only from original information acquired prior to communication and from communication outputs.

Related literature

Crawford and Sobel (1982) suggest the cheap talk equilibrium as proper solution concept of cheap talk games. Forges (1986, 1990, 2006) generalizes cheap talk communication to general communication protocols, and she suggests the communication equilibrium as a proper solution concept of communication games. New solution concepts suggested in this paper replace existing equilibrium definitions for cheap talk games, or communication games.

Aumann (1974, 1987) defines the correlated equilibrium for complete information games, and he provides epistemic justification of it. I extend Aumann (1974)'s logic to 2-stage games in which incomplete information games are preceded by communication stages, define correlated cheap talk and correlated communication equilibrium, and provide epistemic justifications of the new definitions.

Lipman (2009) asks how vague communication can be formally modeled so that the meanings of vague messages such as "tall" or "short" are not precisely defined and used. I answer Lipman's question by showing that the vagueness of communication can be a product of players' uncertainties about other players' messaging strategies, and I suggest forms of correlated equilibrium of communication games to represent such uncertainties.

In the motivating example, I considered a special form of 2-player cheap talk game, in which the speaker's utility does not depend on the state of the world. Glazer and Rubinstein (2012) call such games "persuasion games."

They assume that players are bounded rational in order to have equilibria with meaningful communication in persuasion games. In contrast, I assume full rationality of players. Kamenica and Gentzkow (2011) assume that the speaker can choose extra signals with commitment before his original signals are realized, and they study the speaker's problem of choosing an optimal signal structure. They examine an example of a persuasion game in which meaningful communication is possible under a given commitment assumption. In contrast, I suggest new equilibrium concepts without any commitment assumption, and this provides better equilibrium predictions than existing theories do.

Ishida and Shimizu (2012) and Barreda (2013) consider 2-player cheap talk games in which the listener has access to private information about Θ, and they show that providing more private information to the listener can reduce information transmission incentives in cheap talk equilibria. In contrast, I consider general multi-player incomplete information games that are extended by extra signals with belief invariance.

The new equilibrium definitions presented in this paper can also be interpreted as new correlated equilibrium definitions for simultaneous move incomplete information games. The literature on various correlated equilibrium definitions for incomplete information games is closely related to the work presented here. Forges (1986, 2006) suggests various correlated equilibrium definitions for incomplete information games, as well as Bergemann and Morris (2013, 2015) and Liu (2015). Detailed discussions and comparisons are provided in section 5.

Ely and Pesky (2006), Dekel, Fudenberg and Morris (2007), and Liu (2015) discuss meanings and roles of redundant types in type spaces for incomplete information games. I discuss meanings of redundant types for communication games.

Several authors, including Asheim and Perea (2005) and Battigalli and Siniscalchi (1999, 2002, 2007) provide various epistemic analyses of extensive games for rationalizability notions. In contrast, I conduct epistemic analysis for equilibrium definitions of limited class of extensive games with the common prior assumption, in which cheap talk or mediated communication opportunities precede simultaneous move incomplete information games.

Lambie-Hanson and Parameswaran (2015) study a model of vague communication, in which 2 players, a speaker and a listener, have an identical preference and there is an uncertainty about the listener's rationality. My model considers multi-player incomplete information games with full rationality of all players.

Outline of the paper

In section 2, I give the basic setting and notations. In section 3, I define the correlated cheap talk equilibrium with epistemic justification, and I discuss the meanings of extra signals and belief invariance. In section 4, I define the correlated communication equilibrium with epistemic justification. In section 5, I compare new definitions to existing correlated equilibrium definitions. Section 6 presents several properties of new definitions, and section 7 provides more examples. Section 8 concludes with a discussion of future research topics.

1.2 Setting and Notation

In this section, I describe the setting and notations used throughout the paper. An incomplete information game $G \equiv (I, \Theta, (T_i)_{i \in I}, p, f, (A_i)_{i \in I}, (u_i)_{i \in I})$ consists of the following elements:

- A finite set of players $I \equiv \{1, 2, \dots, I\}$.

- A finite set of payoff-relevant states of the world Θ.
- Finite sets of types T_i for each $i \in I$, and a common prior $p \in \Delta T$, where $T \equiv \prod_{i \in I} T_i$.
- A function $f : T \to \Theta$.[9]
- Finite sets of actions A_i for each $i \in I$.
- Utility functions $u_i : \Theta \times A \to \mathbb{R}$ for each $i \in I$.

Thus, the basic setting of this paper describes a finite multi-player simultaneous move incomplete information game G. In this game all players have private information, and players' beliefs are derived from a common prior. I extend G by a cheap talk in section 3 and by general communication protocol in section 4 in order to define a 2-stage cheap talk extension, or a communication extension of G.

The following notations are used throughout the paper. For any player i-specific sets or functions $X_i, i \in I$, denote $X \equiv \prod_{i \in I} X_i$, and $X_{-i} \equiv \prod_{j \in I, j \neq i} X_j$. For any set X, ΔX denotes the set of probability measures over X. $prob(.)$ is the probability calculated by priors and Bayes' rule.

1.3 Correlated Cheap Talk Equilibrium

In this section, I define a version of a correlated equilibrium of cheap talk extensions of incomplete information games, and call it the *correlated cheap*

[9]Given that player types $t \in T$ completely determine the payoff-relevant state of the world $\theta \in \Theta$, the *join feasibility* condition as explained in Bergemann and Morris (2015) is satisfied in any equilibrium definition with this setting. That is, play can only depend on joint information from all players. This assumption simplifies my analysis without loss of generality because I only consider possible equilibria after communication. Note that any player's information about basic uncertainty, Θ, after any possible communication is limited to what is determined by all players' information, $t \in T$.

talk equilibrium. The new definition generalizes the intuition observed by the motivating example. That is, for cheap talk extensions of incomplete information games, studying the perfect Bayesian equilibrium, or the sequential equilibrium, excludes some plausible outcomes of cheap talk games from the set of equilibrium outcomes because the uncertainties about other players' strategies are not captured. Thus, we need to study a form of correlated equilibrium of cheap talk games. Given a multi-player incomplete information game G, players observe belief invariant extra signals and then they have an opportunity to simultaneously transmit publicly observable messages before choosing actions. After formally defining the correlated cheap talk equilibrium, I provide an epistemic justification for the new definition, and I explain the meanings and significance of extra signals and belief invariance.

Define a decision rule to be a function $\sigma : T \to \Delta A$. I define a correlated cheap talk equilibrium outcome of a game G as a decision rule σ which is a perfect Bayesian equilibrium outcome of an extended game of G. The game G is extended by belief invariant extra signals to players and a cheap talk opportunity. Extra signals to players are represented by a *correlating device*.

Definition 1. *A correlating device for G is defined to be $\Gamma \equiv ((C_i)_{i\in I}, (q^t)_{t\in T})$ such that*

- *C_i are finite sets for all $i \in I$.*
- *$q^t \in \Delta C$ for all $t \in T$*
- *$prob(\theta|t_i, c_i) = prob(\theta|t_i)$ for any $\theta \in \Theta, t_i \in T_i, c_i \in C_i$, and $i \in I$.*

A correlating device defines sets of extra signals $(C_i)_{i\in I}$ and a common prior q^t over C, for each realized $t \in T$. The third condition, which in the literature is called "belief invariance," says that players' beliefs on Θ do not change when they observe extra signals $(c_i \in C_i)_{i\in I}$.

A correlating device Γ and sets of finite messages $(M_i)_{i\in I}$ with an incomplete information game G describe the following 2-stage game. First, nature chooses $(t_i \in T_i)_{i\in I}$ according to p, and $\theta \in \Theta$ is determined by $t \in T$ and f. Then there is an impartial mediator who is represented by a correlating device Γ. The mediator knows the realized $t \in T$ and he sends out extra signals $(c_i \in C_i)_{i\in I}$ to players according to $q^t \in \Delta C$.[10] Note that the mediator sends out extra signals such that players do not update their beliefs about Θ when they observe extra signals. In stage 1, players simultaneously choose and publicly reveal messages $(m_i)_{i\in I}$ from finite sets $(M_i)_{i\in I}$. In stage 2, after observing all signals $m \in M$, players choose $(a_i \in A_i)_{i\in I}$ simultaneously. Then I define the correlated cheap talk equilibrium as follows.

Definition 2. *A decision rule $\sigma : T \to \Delta A$ is a correlated cheap talk equilibrium outcome of an incomplete information game G if*

1. *There exists a correlating device Γ and finite sets of messages $(M_i)_{i\in I}$.*
2. *There exists player i's messaging strategy $\bar{m}_i : T_i \times C_i \to M_i$, and action strategy $\bar{a}_i : T_i \times C_i \times M \to A_i$, for all $i \in I$.*
3. *Players' strategy profiles $(\bar{a}_i, \bar{m}_i)_{i\in I}$ and their beliefs constitute a perfect Bayesian equilibrium of described 2-stage game.*[11]

[10] When I define the correlated cheap talk equilibrium, I let players receive extra signals by an impartial mediator, which follow Aumann's (1974) definition of the correlated equilibrium of complete information games by extra signals. Later, with the epistemic justification of the correlated cheap talk equilibrium, I show that extra signals represent players' beliefs about the state of the world Θ and other players' strategies. The impartial mediator is just a modeling device to represent players' beliefs. The epistemic interpretations of mediators and extra signals closely follow and extend Aumann's (1987) work on the correlated equilibria of complete information games, in which he explains that extra signals in correlated equilibria represent players' beliefs about each other's chosen actions.

[11] As I don't restrict T or C to have full support, some of correlated cheap talk equilibria could be rendered invalid when there are probability zero events and when equilibrium entails irrational behavior on such events. In such cases, one might want to apply equilibrium refinement in light of the sequential equilibrium. However, the correlated cheap

4. σ is induced by p, Γ, and $(\bar{a}_i, \bar{m}_i)_{i \in I}$.

The set of correlated cheap talk equilibrium outcomes of G is the union of all possible correlated cheap talk equilibrium outcomes of G, with all possible Γ and $(M_i)_{i \in I}$.

Example revisited: A decision rule associated with the game given in the motivating example is provided in Figure 4, which cannot be realized by any perfect Bayesian equilibrium of the cheap talk game with the original type space, Θ. Also, the extra signal structure given by Figures 1 and 2 is a correlating device for the game in the motivating example. Note that the decision rule depicted in Figure 4 is a correlated cheap talk equilibrium outcome of the given game. For such correlated cheap talk equilibrium, the correlating device is given by Figures 1 and 2 and equilibrium strategies are described in the motivating example.

Note that a correlated cheap talk equilibrium is a version of a correlated equilibrium of a cheap talk extension of G because players observe extra signals before playing a 2-stage cheap talk game. On the other hand, if we interpret the cheap talk communication as a source of correlation for G, the correlated cheap talk equilibrium could also be interpreted as a version of a correlated equilibrium for given simultaneous move incomplete information game G.

Providing belief invariant extra signals to players prior to cheap talk games is not the only way to define a correlated equilibrium of cheap talk games. One can introduce different correlated equilibria by allowing different

talk equilibrium in this paper is defined as a perfect Bayesian equilibrium of a 2-stage cheap talk game with extra signals for the sake of simplicity and for similarity with widely understood cheap talk equilibrium. Note that Gerardi and Myerson (2007) defined the strong sequential communication equilibrium for communication games, in order to conduct equilibrium refinement of the communication equilibrium. By combining correlated version of cheap talk equilibrium and sequential equilibrium refinement, one might define "strong sequential correlated cheap talk equilibrium" for incomplete information games.

degrees of correlation between extra signals and original signals. I justify my choice of the specific form of correlated equilibrium with epistemic analysis.

Epistemic foundation of the correlated cheap talk equilibrium[12]

I extend Aumann's (1987) work to provide epistemic justification of correlated cheap talk equilibrium. The epistemic analysis of correlated cheap talk equilibrium shows that the new solution concept represents common knowledge of the Bayesian rationality of players with a common prior given a 2-stage cheap talk game. Therefore, if a modeler knows an incomplete information game G, and a modeler knows the fact that players can engage in cheap talk before G with a common prior, then the correlated cheap talk equilibrium must be regarded as the appropriate solution concept. Moreover, epistemic analysis shows that the extra signals in the definition of the correlated cheap talk equilibrium represent players' beliefs and information about the state of the world, Θ and all chosen messaging and action strategies by players. Therefore, the mediator who gives out extra signals to players is just a modeling device to represent players' relevant beliefs in cheap talk games. In addition, the belief invariance of extra signals represents the fact that in cheap talk games, players' information about Θ comes only from their original signals $((t_i \in T_i)_{i \in I})$ and from transmitted messages $(m \in M)$.

Aumann (1987) applied the Bayesian approach to game theory, so that players are assumed to maximize their expected utilities given their information and beliefs about exogenous uncertainties (uncertainties about Θ) and

[12] The epistemic analysis of the correlated cheap talk equilibrium is an adaptation of Auman's (1987) epistemic analysis of complete information simultaneous move games to 2 stage cheap talk extensions of incomplete information games. Note that Battigalli and Siniscalchi (1997) defined the conditional probability system and the epistemic type space for general extensive form games, and their construction could have applied to 2-stage cheap talk games for epistemic characterization of correlated cheap talk equilibrium. Instead, I choose to define an epistemic type space by following and extending Aumann's (1987) analysis, and my approach is simpler and more appropriate for simple 2-stage cheap talk games and provides a clearer explanation of the main points.

endogenous uncertainties (uncertainties about other players' chosen strategies). The Bayesian approach is different from the traditional approach to game theory, in which probabilities are assigned to exogenous uncertainties only. Applying the Bayesian approach to complete information simultaneous move games, Aumann (1987) defined common prior epistemic type spaces for complete information games. A state in an epistemic type space describes both exogenous uncertainties (realized $\theta \in \Theta$, which is trivial in the case of complete information games) and endogenous uncertainties (all players' chosen strategies).

Then Aumann (1987) showed that if in complete information games players are rational in all states of a given epistemic type space (which implies common knowledge of the Bayesian rationality of players), then the distribution of actions is a correlated equilibrium distribution. Note that a correlated equilibrium for a complete information game is defined by extra signals to players and players' equilibrium strategies that depend on extra signals. Aumann's (1987) epistemic analysis shows, first, that the correlated equilibrium represents the common knowledge of the Bayesian rationality of players with a common prior given a complete information game, and second, that extra signals in the definition of the correlated equilibrium represent players' beliefs about other players' chosen strategies (endogenous uncertainties).

I generalize Aumann's (1987) analysis of complete information simultaneous move games to 2-stage incomplete information cheap talk games. Consider a 2 stage game, which is a cheap talk extension G_C of an incomplete information game G. In the first stage of G_C, all players simultaneously send publicly observable messages from finite sets of messages $(M_i)_{i \in I}$. In the second stage of G_C, players simultaneously choose actions defined in G. Note that in the definition of G_C, there does not exist a mediator who gives out some extra signals. I define a common prior epistemic type space Ω of G_C. As in Aumann (1987), an epistemic state $\omega \in \Omega$ represents not only exogenous uncertainties ($\theta \in \Theta$ and $t \in T$) but also endogenous uncertainties (players' chosen messaging and action strategies). With an epistemic type space, players are assumed, first, to take both exogenous and endogenous uncertainties alike and, second, rationally choose optimal messages and actions based on their beliefs on all possible uncertainties. A common prior epistemic type space Ω for a cheap talk extension G_C of G consists of the following:

- A finite set of states of the world Ω and a common prior $\hat{p} \in \Delta\Omega$.
- Functions that specify exogenous uncertainties for states, $\hat{t} : \Omega \to T, \hat{\theta} : \Omega \to \Theta$. By definition, $\hat{\theta} \equiv f \circ \hat{t}$, and $\hat{p}$ together with $\hat{t}$ must be congruent to p.
- Functions that specify players' chosen messaging strategies for states, $\hat{m}_i : \Omega \to M_i$ for all $i \in I$.
- Functions that specify players' chosen action strategies for states, $\hat{A_i^M} : \Omega \to \mathcal{A}_i^{\mathcal{M}}$ for all $i \in I$. Here, $\mathcal{A}_i^{\mathcal{M}}$ is the space of all functions from M to A_i, with a typical element $A_i^M : M \to A_i$. [13]
- Each player i's information partition $\mathbb{P}_i^{A_i}$ on Ω when he chooses $A_i^M \in \mathcal{A}_i^{\mathcal{M}}$ for all $i \in I$, such that player i knows his type $t_i \in T_i$, his messaging strategy $m_i \in M_i$, his action strategy $A_i^M \in \mathcal{A}_i^{\mathcal{M}}$, and the transmitted public message $m \in M$. In other words, $\hat{t}_i, \hat{m}_i, \hat{A_i^M}$, and $\hat{m}$ are measurable with respect to $\mathbb{P}_i^{A_i}$ for all $i \in I$.
- Each player i's information partition $\mathbb{P}_i^{M_i}$ on Ω when he chooses $m_i \in M_i$ for all $i \in I$, such that player i knows his type $t_i \in T_i$, his messaging

[13] Player i's action strategy given $\omega \in \Omega$ is defined to be a function that maps the set of realized messages M to the set of player i's actions A_i, due to the structure of 2-stage game G_C. Therefore, a state $\omega \in \Omega$ contains information about player i's action choice in the hypothetical situation if player i received a different message combination other than what he actually receives in ω.

strategy $m_i \in M_i$, and his action strategy $A_i^M \in \mathcal{A}_i^{\mathcal{M}}$, but does not know the message combination that will be sent from other players, $m_{-i} \in M_{-i}$. In other words, $\mathbb{P}_i^{M_i}$ is a coarser partition than $\mathbb{P}_i^{A_i}$, and each element of $\mathbb{P}_i^{M_i}$ is obtained by merging elements of $\mathbb{P}_i^{A_i}$ which share same $\hat{t}_i, \hat{m}_i$, and $\hat{A_i^M}$.[14]

- Player i's belief about Θ is affected only by his information about t_i, and by information observed by public messages $m \in M$. In other words, information about θ is only updated by communication. Take any two $p_i, p_i' \in \mathbb{P}_i^{M_i}$ such that they share same t_i, then we have

$$\hat{p}(\theta|p_i) = \hat{p}(\theta|p_i') \tag{1.1}$$

  for any $\theta \in \Theta$.

The last condition in the above definition of the epistemic type space is the main informational assumption of G_C: players' information about Θ comes only from either their original signals or from communication outputs. In the epistemic type space, player i's partition $\mathbb{P}_i^{A_i}$ describes his knowledge about t_i, m_i, A_i^M, and m_{-i}, and his belief about Θ. His belief about Θ could possibly be different from what he believed without information m_{-i}. The last condition says that the updated belief about Θ comes solely from his knowledge about m_{-i}. With the coarser information partition $\mathbb{P}_i^{M_i}$, which is acquired by forgetting player i's knowledge about m_{-i} from $\mathbb{P}_i^{A_i}$, the player i's belief about Θ is determined by his knowledge of t_i.

Example revisited: An epistemic type space that describes realized $\theta \in \Theta$ and the players' chosen strategies for the motivating example is illustrated in Figure 5. For example, the top left state describes the state of the world (nc),

[14]Note that player i's action strategy is a function that maps the set of message combinations to his actions, and it is decided before he receives realized messages. Therefore, player i's action strategy is decided when his beliefs and information are given by $\mathbb{P}_i^{M_i}$.

together with the speaker's messaging strategy (nw) *and the listener's action strategy* (sq, f)*, so that the listener chooses* (sq) *when she received message* (w)*, and she chooses* (f) *when she receives message* (w)*. Note that the epistemic type space looks similar to the extended type space given by Figure 3, but there is no extra signal in the epistemic type space. The information partitions of players are congruent with the information that players hold when they choose messages or actions. Even if the listener does not send messages and the speaker does not choose actions in the example, we can naturally define 2 information partitions for both players. For example,* $\mathbb{P}_L^{M_L}$ *consists of two elements that are two columns of the epistemic type space, whereas* $\mathbb{P}_L^{A_L}$ *consists of four elements: the top left state; the top right state; the set of the middle left and the bottom left state; and the set of the middle right and the bottom right state. On the other hand,* $\mathbb{P}_S^{M_S}$ *and* $\mathbb{P}_S^{A_S}$ *consist of three elements that are three rows of the type space.*

	sq, f	$sq, supp$
nc, nw	0.2	0.1
nc, w	0.1	0.2
c, w	0.2	0.2

Figure 1.5: An epistemic type space Ω of the game in the example.

Then I suggest epistemic justification of the correlated cheap talk equilibrium as follows.

Proposition 1. *A decision rule* $\sigma : T \to \Delta A$ *is a correlated cheap talk equilibrium outcome of* G *if and only if there exists a common prior epistemic type space* Ω *of* G_C *with sets of messages* $(M_i)_{i \in I}$*, such that the followings conditions are satisfied.*

- *σ is induced by Ω.*

- *Each player i is Bayes rational at each state of the world in choosing $m_i \in M_i$ given information $\mathbb{P}_i^{M_i}$.*

- *Each player i is Bayes rational at each state of the world in choosing $A_i^M \in \mathcal{A}_i^{\mathcal{M}}$ given information $\mathbb{P}_i^{A_i}$.*[15]

Proof. For the "if" direction, assume that there is an epistemic type space Ω of G_C with described conditions. I construct a correlated cheap talk equilibrium of G by sets of messages, a correlating device, and the equilibrium strategies of players. Take $(M_i)_{i \in I}$ as given by G_C. Define

$$C_i \equiv \mathbb{P}_i^{M_i} \tag{1.2}$$

Therefore, an extra signal for player i, $c_i \in C_i$, determines player i's information $\hat{t}_i$, and strategies $\hat{m}_i$ and $\hat{A_i^M}$. Also, define

$$C \equiv \{c \subset \Omega | \exists c_i \in C_i, \forall i \in I, c = \bigcap_{i \in I} c_i\} \tag{1.3}$$

Denote for all $c \in C, c = \bigcap\limits_{i \in I} c_i \equiv (c_1, c_2, \ldots, c_I)$. Note that a fixed $c \in C$ exhausts all possible uncertainties in Ω, and so any $c \in C$ is a singleton set, and we can define an 1-1, onto function $\tilde{f} : C \to \Omega$. For all $t \in T$ and all $c \in C$, define

$$q^t(c) \equiv \frac{\hat{p}(\tilde{f}(c))}{\hat{p}(\omega \in \Omega | \hat{t}(\omega) = t)} \tag{1.4}$$

Now we know that for all $\theta \in \Theta$, and $c_i, c_i' \in C_i, t_i \in T_i$ such that $\hat{t}(\tilde{f}(c_i)) = \hat{t}(\tilde{f}(c_i')) = t_i$,

[15] As noted in Aumann (1987), if all players are rational in every states of the epistemic type space for choosing all strategies, then the rationality of players is common knowledge.

$$prob(\theta|t_i, c_i) = \hat{p}(\theta|\tilde{f}(c_i)) = \hat{p}(\theta|\tilde{f}(c_i')) = prob(\theta|t_i, c_i') \tag{1.5}$$

, by the definition of p, q, and the condition (1.1). Therefore, we know that defined $\Gamma \equiv ((C_i)_{i\in I}, (q^t)_{t\in T})$ is a correlating device for G. Now I define players' equilibrium strategies. For all $t_i \in T_i, c_i \in C_i$ with $\hat{t}(\tilde{f}(c_i)) = t_i$ and all $i \in I$,

$$\bar{m}_i(t_i, c_i) \equiv \hat{m}_i(\tilde{f}(c_i)) \tag{1.6}$$

Also, for all $t_i \in T_i, c_i \in C_i, m \in M$ with $\hat{t}(\tilde{f}(c_i)) = t_i$, and all $i \in I$, define

$$\bar{a}_i(t_i, c_i, m) \equiv \hat{A_i^M}(\tilde{f}(c_i))(m) \tag{1.7}$$

The above strategies $\bar{m}_i$ and $\bar{a}_i$ for all $i \in I$ constitute a perfect Bayesian equilibrium in the extended cheap talk game of G, which is extended by the correlating device Γ and sets of messages $(M_i)_{i\in I}$. First, $\bar{m}_i(t_i, c_i) = \hat{m}_i(\tilde{f}(c_i))$ maximizes player i's expected utility since he is Bayes rational in choosing $m_i \in M_i$ when he has information $\tilde{f}(c_i)$. Also, $\bar{a}_i(t_i, c_i, m) \equiv \hat{A_i^M}(\tilde{f}(c_i))(m)$ maximizes player i's expected utility because the knowledge of $\tilde{f}(c_i) \in \mathbb{P}_i^{M_i}$ and $m \in M$ determines a member $p_i^A \in \mathbb{P}_i^{A_i}$, and player i is Bayes rational in choosing $A_i^M \in \mathcal{A}_i^{\mathcal{M}}$ with information $p_i^A \in \mathbb{P}_i^{A_i}$.

For the "only if" direction, let σ be a correlated cheap talk equilibrium outcome with given Γ and $(M_i)_{i\in I}$. Define an epistemic type space Ω such that each $\omega \in \Omega$ describes $\hat{t}, \hat{m}_i, \hat{A_i^M}$ for all i as given by game G, correlating device Γ, and the equilibrium strategies of players. Assign $\hat{p}(\omega)$ for all $\omega \in \Omega$ as derived from p, q, and players' equilibrium strategies. Check that the properties listed in Proposition 1 are satisfied.

$\square$

Example revisited: Note that in the epistemic type space described by Figure 5, players are rational in choosing messaging and action strategies given their information partitions. Therefore, the decision rule induced by the epistemic type space is a correlated cheap talk equilibrium, according to Proposition 1. We can see that the described equilibrium in the motivating example with Γ *(Figures 1 and 2), generates the equilibrium outcome described in Figure 4. This, in turn, is induced by the epistemic type space described by Figure 5.*

Interpretations of extra signals and belief invariance

The epistemic analysis of the correlated cheap talk equilibrium sheds light on the interpretations and meanings of extra signals and the belief invariance condition in the definition of the correlated cheap talk equilibrium.

Consider how sets of extra signals are determined in the proof of Proposition 1. When a correlated cheap talk equilibrium is constructed given an epistemic type space Ω of G_C with rational players, each set C_i is defined to be the same as $\mathbb{P}_i^{M_i}$, player i's information partition when he chooses messaging strategy $m_i \in M_i$. That is, player i's extra signal $c_i \in C_i$ represents his belief and knowledge about both exogenous uncertainties (uncertainties about Θ), and endogenous uncertainties (uncertainties about the messaging and action strategies of other players) when he chooses messaging strategies without knowing the realized message combination of other players, $m_{-i} \in M_{-i}$.[16]

[16] In the proof of Proposition 1, player i's extra signal $c_i \in C_i$, which is constructed from information $p_i^M \in \mathbb{P}_i^{M_i}$, determines player i's knowledge about $t_i \in T_i$, as well as optimal messaging and action strategies. Note that in the definition of the correlated cheap talk equilibrium, an extra signal c_i of player i does not necessarily determine his original signal t_i and his optimal strategies. However, for every extra signal structure for a correlated cheap talk equilibrium, we can redefine an outcome equivalent extra signal structure that contains all information from original signals (outcome equivalence means that there is a 1-1, onto function that maps the original type space, which includes original and extra signals of players, to the new type space so that the function preserves all types' beliefs about Θ and all other players' strategies in a given equilibrium.) In the proof of Proposition

In the proof of Proposition 1, the belief invariance of extra signals in the definition of correlated cheap talk equilibrium is derived from condition (1.1) in the definition of epistemic type space. Condition (1.1) says that players' information about Θ comes only from the original signals $(T_i)_{i\in I}$ and from the communication outcome, $m \in M$. Therefore, the belief invariance of extra signals for a correlated cheap talk equilibrium comes from the main informational assumption of a cheap talk extension G_C of G.

Example revisited: The speaker's extra signals for the correlated cheap talk equilibrium in the motivating example characterize his belief and information when he chooses a messaging strategy. For example, the speaker's extra signal 1 *characterizes his information partition given as the top row in the epistemic type space given in Figure 5. That information partition determines his beliefs about the listener's action when she receives message* w*. Speaker types with extra signals* 1 *and* 2 *have different such beliefs about the joint space of* Θ *and the listener's action strategies. However, speaker types with extra signals* 1 *and* 2 *share same beliefs about* Θ *because they observe the same original signal* nc*. The listener's extra signals and belief invariance can be interpreted similarly.*

1.4 Correlated Communication Equilibrium

In this section, I generalize the analysis of cheap talk extensions of incomplete information games to general communication extensions of incomplete information games. Suppose we wish to know the set of possible equilibrium outcomes of an incomplete information game G when an unknown form of communication is possible. It is well known that a more complicated communication protocol might, in the context of incomplete information games,

1, I let players' extra signals to determine players original signals and strategies when I construct a correlated cheap talk equilibrium for a given epistemic type space.

generate larger set of equilibrium outcomes compared to cheap talk. For example, Krishna and Morgan (2004) show that in 2 player games, as studied by Crawford and Sobel (1982), multi-stage cheap talk might induce strictly larger set of equilibrium outcomes than one stage cheap talk. One can imagine various communication protocols for multi-player incomplete information games, such as a communication protocol in which only subset of players could communicate each other for an extended period. Then given an incomplete information game G, one might ask the following question: What is the set of possible equilibrium outcomes of G when G is extended by arbitrary communication protocols?

To answer that question, Forges (1986, 1990, 2009) defines a communication equilibrium of an incomplete information game G as a perfect Bayesian equilibrium of an extended game in which G is preceded by a communication device. This communication device asks players to report their types in the interim stage of the game, and then it send out recommendations of actions to players according to predetermined functions. According to the revelation principle, a perfect Bayesian equilibrium outcome of any arbitrary communication extension of G can be replicated by a communication equilibrium of G when an appropriate communication device is chosen. Whereas messages from all players are revealed to all other players with cheap talk communication protocol, a communication device enables information transmissions in which players could be kept from some portion of information which comes from other players. In contrast to a situation where all information are revealed to all players, in this case players' incentive constraints are relaxed, which could induce a larger set of equilibrium outcomes.

Note that the intuition from the motivating example applies also for general communication extensions of incomplete information games. That is, the equilibrium outcome in Figure 4 is what we can expect from a given game extended by a communication opportunity. However, it is not a communication equilibrium outcome of a given game without the extra signals in the motivating example.

Therefore, to define *correlated communication equilibrium* for an incomplete information game G, I allow players to observe belief invariant extra signals before they report to a communication device. By epistemic analysis, I show that the correlated communication equilibrium represents the common knowledge of the Bayesian rationality of players with a common prior given a communication protocol. The epistemic analysis shows also that extra signals represent players' beliefs, and belief invariance indicates that players' information comes only from the original signal and from communication outcomes.

Consider a game with incomplete information $G \equiv (I, \Theta, (T_i)_{i\in I}, p, f, (A_i, u_i)_{i\in I})$ as defined in section 2. For G, I define a correlated communication device.

Definition 3. *A correlated communication device is defined to be* $\Gamma' \equiv ((C_i)_{i\in N}, (q^t)_{t\in T}, \gamma)$ *such that*

1. C_i *is a finite set for all* $i \in I$*, with* $C \equiv \prod_{i\in I} C_i$
2. $q^t \in \Delta C$ *for all* $t \in T$*.*
3. $prob(\theta|t_i) = prob(\theta|t_i, c_i)$*, for any* $\theta \in \Theta, i \in I, t_i \in T_i, c_i \in C_i$*.*
4. $\gamma : T \times C \to \Delta A$*.*

A correlated communication device Γ' can be thought of as consisting of two different types of mediators. First, after the original signals $t \in T$ and the state of the world $\theta \in \Theta$ is realized, there is a mediator who gives out belief invariant extra signals $(c_i \in C_i)_{i\in I}$ to players. As noted before, the first type of mediator is a modeling device that models players' beliefs about Θ and other players' strategies in communication games. Second, after players receive original and extra signals, there is a mediator who invites players to report their original and extra signals, and he provides recommendations of actions to players according to predetermined rules. As suggested by Forge's (1986, 1990, 2006) analysis of communication equilibrium, the second type of mediator is a modeling device that represents some communication protocol, which is a rule of communication with a given incomplete information game.

A correlated communication device Γ' together with an incomplete information game G describes a 2-stage communication extension of G as the following. Nature draws $t \in T$ and $c \in C$ according to p and $(q^t)_{t\in T}$, and each player $i \in I$ observes $t_i \in T_i$ and $c_i \in C_i$. Then there is a communication device that invites players to report their original and additional signals $((t_i, c_i) \in T_i \times C_i)_{i\in I}$, and it gives out recommendations to players according to $\gamma : T \times C \to \Delta A$.[17] I define the correlated communication equilibrium as follows.

Definition 4. *A decision rule $\sigma : T \to \Delta A$ is a correlated communication equilibrium outcome of an incomplete information game G if*

1. *There exists a correlated communication device Γ'.*
2. *Players cannot gain by unilaterally lying about their types or by deviating from recommended actions in the extended game G with Γ'.*[18]
3. *σ is induced by p and Γ'.*

[17]The definition of the correlated communication device and the communication extension of G implicitly assumes the revelation principle. That is, I could have defined the correlated communication device and a communication extension of G with sets of messages for players and sets of signals to players from the mediator.

[18]As in the case of the correlated cheap talk equilibrium, the correlated communication equilibrium is defined as a perfect Bayesian equilibrium, not a sequential equilibrium, of an extended game, for the sake of simplicity and for similarity with the communication equilibrium.

The set of correlated communication equilibrium outcomes of G is the union of all possible correlated communication equilibrium outcomes of G that are generated by all possible correlated communication devices which represent all possible belief invariant extra signal structures and all possible communication protocols.

Note that the correlated communication equilibrium is a generalization of the correlated cheap talk equilibrium.

Observation 1. *Given an incomplete information game G, a correlated cheap talk equilibrium outcome σ of G is also a correlated communication equilibrium outcome of G.*

Proof. Given a correlating device Γ and equilibrium strategies for correlated cheap talk equilibrium, I construct a correlated communication device Γ'. Take $((C_i)_{i\in I}, (q^t)_{t\in T})$ as the same ones provided by Γ; γ is induced by equilibrium strategies of players in the correlated cheap talk equilibrium. Players' incentives for messaging and action strategies are satisfied with given correlated cheap talk equilibrium. $\square$

The converse of the Observation 1 is not necessarily true: for a given incomplete information game, there could be correlated communication equilibrium outcomes which are not correlated cheap talk equilibrium outcomes.[19] In section 7, I discuss several examples in which the converse of the Observation 1 does not hold.

Epistemic foundation of the correlated communication equilibrium

In this section, I show that the set of correlated communication equilibrium outcomes describe the common knowledge of the Bayesian rationality

[19] For the game described in the motivating example, the set of correlated cheap talk equilibrium outcomes is the same as the set of correlated communication equilibrium outcomes. That is because of the simplicity of the example.

of players with a common prior given a communication game. First, define a communication protocol $\mathcal{C}$ for an incomplete information game G.

Definition 5. *A communication protocol* $\mathcal{C} \equiv \{(M_i)_{i\in I}, (O_i)_{i\in I}, \beta\}$ *of an incomplete information game* G *consists of the following.*

- *Finite set* M_i*, the set of messages for player* i *for each* $i \in I$*.*
- *Finite set* O_i*, the set of outputs for player* i *for each* $i \in I$*.*
- *A function* $\beta : M \to \Delta O$*, and derived* $\beta_i : M \to \Delta O_i$ *for each* $i \in I$*.*

With a communication protocol $\mathcal{C}$, each player i sends a message $m_i \in M_i$ to a communication device, and receives an outputs $o_i \in O_i$ from the communication device according to β_i. Naturally, I define a 2-stage communication extension $G_{\mathcal{C}}$ of an incomplete information game G so that G is preceded by $\mathcal{C}$. A communication protocol represents a specific communication opportunity that might enable more information transmission than cheap talk.

Now define a common prior epistemic type space Ω of a communication extension $G_{\mathcal{C}}$ of a game G that has the given communication protocol $\mathcal{C}$. A state $\omega \in \Omega$ describes exogenous uncertainties ($\theta \in \Theta, t \in T$) and endogenous uncertainties (players' chosen messaging and action strategies.) A common prior epistemic type space Ω of $G_{\mathcal{C}}$ describes the following:

- A finite set of states of the world Ω and a common prior $\hat{p} \in \Delta\Omega$.
- Functions that specify exogenous uncertainties for states, $\hat{t} : \Omega \to T$, and $\hat{\theta} : \Omega \to \Theta$. By definition, $\hat{\theta} \equiv f \circ \hat{t}$ and $\hat{p}$ together with $\hat{t}$ are congruent to p.
- Functions that specify players' chosen messaging strategies for states, $\hat{m}_i : \Omega \to M_i$ for all $i \in I$.

- Functions that specify players' received communication outputs for states, $\hat{\beta}_i : \Omega \to O_i$ for all $i \in I$. Note that $\hat{\beta}_i \equiv \beta_i \circ \hat{m}$ by definition.

- Functions that specify players' chosen action strategies for states, $\hat{A_i^{O_i}} : \Omega \to \mathcal{A}_i^{\mathcal{O}_i}$ for all $i \in I$. Note that $\mathcal{A}_i^{\mathcal{O}_i}$ is the space of all functions from the set of communication outputs O_i to the chosen action A_i, with a typical element $A_i^{O_i} : O_i \to A_i$.[20]

- Each player i's information partition $\mathbb{P}_i^{A_i}$ on Ω when he chooses $A_i^{O_i} \in \mathcal{A}_i^{\mathcal{O}_i}$ for all $i \in I$, such that player i knows his type $t_i \in T_i$, his messaging strategy $m_i \in M_i$, action strategy $A_i^{O_i} \in \mathcal{A}_i^{\mathcal{O}_i}$, and output from communication $\beta_i(m)$. In other words, $\hat{t}_i, \hat{m}_i, \hat{A_i^{O_i}}$, and $\hat{\beta}_i$ are measurable with respect to $\mathbb{P}_i^{A_i}$.

- Each player i's information partition $\mathbb{P}_i^{M_i}$ on Ω when he chooses $m_i \in M_i$ for all $i \in I$, such that player i knows his type $t_i \in T_i$, his messaging strategy $m_i \in M_i$, and action strategy $A_i^{O_i} \in \mathcal{A}_i^{\mathcal{O}_i}$, but does not know output from communication $\beta_i(m)$. In other words, $\mathbb{P}_i^{M_i}$ is a coarser partition than $\mathbb{P}_i^{A_i}$, and each element of $\mathbb{P}_i^{M_i}$ is obtained by merging elements of $\mathbb{P}_i^{A_i}$ which share the same $\hat{t}_i, \hat{m}_i$, and $\hat{A_i^{O_i}}$.

- Player i's belief about Θ is affected only by his information about t_i and by information from the communication. Take any two $p_i, p_i' \in \mathbb{P}_i^{M_i}$ such that they share the same t_i. Then we have

$$\hat{p}(\theta|p_i) = \hat{p}(\theta|p_i') \tag{1.8}$$

  for any $\theta \in \Theta$.

[20] Like the epistemic type space for cheap talk games, players' action strategies are functions that map the set of communication outputs to the set of actions due to the structure of 2-stage game $G_{\mathcal{C}}$.

Then I suggest epistemic justification of the correlated communication equilibrium as follows.

Proposition 2. *A decision rule $\sigma : T \to \Delta A$ is a correlated communication equilibrium outcome of an incomplete information game G if and only if there exists an epistemic type space Ω of a communication extension $G_{\mathcal{C}}$ of G with a communication protocol $\mathcal{C}$, such that the followings conditions are satisfied.*

- *σ is induced by Ω.*
- *Each player i is Bayes rational at each state of the world in choosing $m_i \in M_i$ with information $\mathbb{P}_i^{M_i}$.*
- *Each player i is Bayes rational at each state of the world in choosing $A_i^{O_i} \in \mathcal{A}_i^{\mathcal{O}_i}$ with information $\mathbb{P}_i^{A_i}$.*

Proof. Omitted. (Similar to the proof of Proposition 1). $\square$

Proposition 2 says that the the set of correlated communication equilibrium outcomes of a given incomplete information game G represents the common knowledge of the Bayesian rationality of players with a common prior given $G_{\mathcal{C}}$ with a fixed communication protocol $\mathcal{C}$. Therefore, when a modeler knows (a) the underlying game G and (b) that players might engage in some form of communication without knowing the exact communication protocol, the modeler expects the set of correlated communication equilibrium outcomes as plausible equilibrium outcomes with a common prior.

Moreover, the definition of the epistemic type space Ω of $G_{\mathcal{C}}$ and the proof of Proposition 2 indicate that the extra signals that players receive in the definition of correlated communication device represent their beliefs regarding both exogenous and endogenous uncertainties when they choose messaging strategies. Additionally, the belief invariance of extra signals represents the main informational assumption of communication extensions of incomplete information games - i.e, condition (1.8). Players' information about Θ comes only from their information about $(t_i \in T_i)_{i \in I}$ and from communication outputs. I do not discuss this issue in further detail here because it resembles my discussion about extra signals and belief invariance in section 3.

1.5 Comparisons with Other Equilibrium Definitions

So far, I defined new solution concepts, *correlated cheap talk equilibrium* and *correlated communication equilibrium* for incomplete information games that are preceded by communication opportunities. New definitions represent reasonable predictions of a given incomplete information game G with a common prior, preceded by a cheap talk opportunity or by a general communication opportunity. Note that new definitions can be interpreted as correlated equilibrium definitions of incomplete information games, wherein the correlation between players' actions come from players' correlated beliefs and communications. In this section, I compare new definitions with existing correlated equilibrium definitions.

Aumann (1974, 1987) defines correlated equilibrium for complete information games. New definitions introduced in this paper are generalizations of Aumann's correlated equilibrium as follows:

Observation 2. *With complete information, the set of correlated cheap talk equilibrium outcomes and the set of correlated communication equilibrium outcomes are the same as the set of correlated equilibrium outcomes.*

Proof. Established by definitions. □

The new solution concepts presented in this paper generalize Aumann's (1974, 1987) correlated equilibrium in two respects. First, they examine equilibria in incomplete information games. Second, they allow communication between players. With regarding to equilibrium concepts of incomplete information games extended by communication opportunity, Crawford and Sobel (1982) analyzed the cheap talk equilibrium of incomplete information game with cheap talk communication opportunities, and Forges (1986, 1990, 2006) analyzed the communication equilibrium of incomplete information with general mediated communication opportunities. The cheap talk equilibrium and communication equilibrium are perfect Bayesian equilibrium definitions for 2-stage games in which incomplete information games are preceded by communication stages. The correlated cheap talk equilibrium and the correlated communication equilibrium are specific forms of correlated equilibria for 2-stage incomplete information games with communication, and so they generalize existing definitions as follows.

Observation 3. *For a multi-player incomplete information game G, define a cheap talk equilibrium as a perfect Bayesian equilibrium of a 2-stage game in which a simultaneous public message exchange opportunity precedes G. Then any cheap talk equilibrium of G is a correlated cheap talk equilibrium. Also, a communication equilibrium of G is a correlated communication equilibrium.*

Proof. Established by definitions. □

The correlated cheap talk equilibrium and the correlated communication equilibrium generalize cheap talk equilibrium and communication equilibrium by letting players observe belief invariant extra signals before 2-stage communication extensions of incomplete information games. By letting players observe belief invariant extra signals before simultaneous move incomplete information games, Liu (2015) defined the belief invariant Bayes correlated equilibrium. The following observation follows from definitions.

Observation 4. *For any incomplete information game G, a belief invariant*

Bayes correlated equilibrium, as defined in Liu (2015), is a correlated cheap talk equilibrium and also a correlated communication equilibrium.

Proof. Established by definitions. □

Note that the set of belief invariant Bayes correlated equilibrium outcomes describes all possible equilibrium outcomes in which players' belief hierarchies over Θ are fixed by their original types, $(t_i \in T_i)_{i \in I}$ in the description of G. Also, in the belief invariant Bayes correlated equilibrium, belief invariance is introduced in order to study possible equilibrium outcomes in which players' beliefs are not changed from what are given by their original types. In correlated cheap talk equilibrium and the correlated communication equilibrium, players' beliefs about Θ are updated as a result of communications, and the belief invariance represents the fact that in games with communication, players' information about Θ is acquired only from either their original information, $(t_i \in T_i)_{i \in I}$ or from communication outputs.

	No communication	Communication
Complete info	Aumann (1987)	Aumann (1987)
Incomplete info	Liu (2015)	Current Paper

Figure 1.6: Correlated equilibrium definitions with belief invariant extra signals

Figure 6 summarizes my discussion up to this point. It categorizes correlated equilibrium definitions, which are defined by belief invariant extra signals. With complete information games, all possible extra signals are belief invariant because players do not have private information. In such cases, allowing a communication opportunity does not affect the set of correlated equilibrium outcomes. Any extra signals or communication opportunity only work to correlated players' chosen actions, and communication does not transmit any information about exogenous uncertainties. Therefore, Aumann's (1974, 1987) correlated equilibrium is the relevant equilibrium concept of complete information games with and without communication opportunities. On the other hand, with incomplete information games, belief invariant extra signals work differently in 2-stage games with communication than they do in simultaneous move games. Whereas belief invariant extra signals work only to correlate players' action strategies in simultaneous move incomplete information games, they work to correlate players' communication and action strategies in 2-stage communication games. Therefore, extra signals help players to transmit more information about exogenous uncertainties in 2-stage communication games.

Bergemann and Morris (2013, 2015) define Bayes correlated equilibrium for incomplete information games by letting players, prior to a given game, observe arbitrary extra signals given by an omniscient mediator. When considering the set of the Bayes correlated equilibrium outcomes for a given incomplete information game G, a modeler imagines that players could have information about Θ beyond what is given by players' types $(t_i \in T_i)_{i \in I}$ in the description of G. Therefore, in the Bayes correlated equilibrium, players' beliefs about Θ could be arbitrarily updated compared to what are given by $(t_i \in T_i)_{i \in I}$. Extra signals represent additional information structure known to players, but not to the modeler. In the correlated cheap talk equilibrium and correlated communication equilibrium outcomes, players' beliefs about Θ also can be updated from what is given by $(t_i \in T_i)_{i \in I}$. However, in correlated cheap talk and correlated communication equilibrium, a modeler knows that players' updated beliefs come from communication. In the new definitions proposed in this paper, extra signals are not the actual information structure; instead, they represent players' beliefs about Θ and other players' strategies in communication games. Also, in this paper, players' information about Θ is updated only when the information change is approved by other players who pass that information through communication.

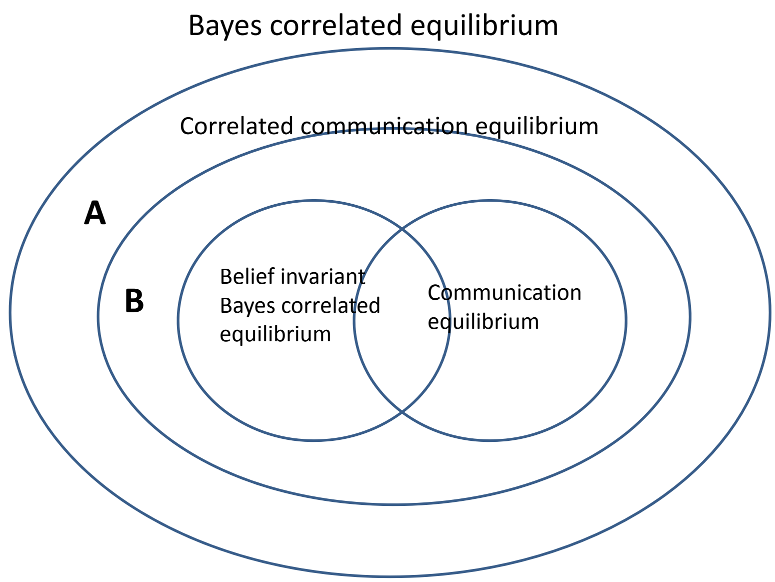
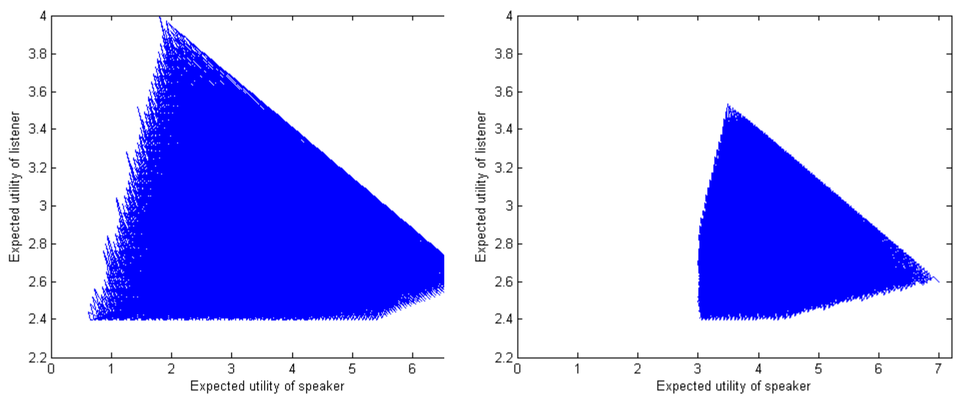


Figure 1.7: Correlated communication equilibrium and other definitions

Figure 7 shows set relations between several different correlated equilibrium definitions for incomplete information games. Note that in area **A**, there are Bayes correlated equilibrium outcomes such that the omniscient mediator gives out information to players, but some of the information transmission is not approved by other players. Consequently, such equilibrium outcomes cannot be realized by games that are extended by belief invariant extra signals and communication opportunities. In area **B**, there are correlated communication equilibrium outcomes as described by the motivating example, so that such outcomes cannot be a communication equilibrium outcome nor a belief invariant Bayes correlated equilibrium outcome. The information transmitted to players can only be incentivized by correlated communication: players need a communication opportunity and they should have uncertainties about other players' strategies, which are represented by belief invariant extra signals.

Example revisited: Figure 8 shows the set of Bayes correlated equilibrium (BCE) outcome payoffs and the set of correlated communication equilibrium (CCE) outcome payoffs of the game G given by the motivating example. Note that the set of correlated communication equilibrium outcome payoffs is a proper subset of the set of Bayes correlated equilibrium outcome payoffs. Also, the set of communication equilibrium outcome payoffs and the set of belief invariant Bayes correlated outcome payoffs consist of a single point, $(7, 2.6)$. *The equilibrium outcome payoffs described in the motivating example and Figure 4 is the point* $(3.7, 3)$ *in Figure 8.*:[21]



(a) Set of BCE outcome payoffs (b) Set of CCE outcome payoffs

Figure 1.8: Comparison of BCE and CCE outcome payoffs of the example

1.6 Several Properties of New Definitions

Several meaningful properties of new definitions can be easily verified.

[21]The set of BCE and CCE outcomes are numerically calculated and plotted using MATLAB.

First, correlated cheap talk and correlated communication equilibrium outcomes are invariant to adding redundant types to the given type space in an incomplete information game G. That is, the set of correlated cheap talk and correlated communication equilibrium outcomes depend only on players' belief hierarchies regarding Θ, which are determined by players' types $(T_i)_{i\in I}$ in given G.[22] This property is shared with Liu's (2015) belief invariant Bayes correlated equilibrium and Dekel, Fudenberg and Morris' (2007) interim correlated rationalizability.

Observation 5. *Given an incomplete information game G, the set of correlated cheap talk equilibrium outcomes and the correlated communication equilibrium outcomes are invariant to the addition of redundant types to the type structure T in G.*

Proof. Established by definitions. □

Next, we know from Forges (1986) that the set of communication equilibrium outcomes is a convex polyhedron because it can be described by linear inequalities. For a given G, if we fix the correlating device or the correlated communication device, then the set of correlated cheap talk equilibrium outcomes and the set of correlated communication equilibrium outcomes are convex polyhedron because they are described by sets of linear inequalities. The sets of equilibrium outcomes given G with varying correlating devices or correlated communication devices are unions of convex polyhedra. I strongly suspect that the set of correlated cheap talk equilibrium outcomes and the set of correlated communication equilibrium outcomes are also convex polyhedra, as we can observe from Figure 8-(b) for the example. However, I do not know the general proof of the claim.

[22]A type space T in an incomplete information game G models players' belief hierarchies about Θ, as shown by Mertens and Zamir (1985) and Brandenburger and Dekel (1993).

Forges (1990) justifies her definition of communication equilibrium by showing that any equilibrium of an extension of G with some form of pre-play communication (which does not explain the correlated communication examined in this paper) can be replicated by a communication equilibrium, in which a communication device receives reports from players about their original types $(t_i \in T_i)_{i \in I}$, and it sends out action recommendations to players. The correlated communication device introduced in this paper combines two types of mediators: a mediator which represent players' beliefs in communication games and a mediator which represents a communication protocol as in Forges (1990). Therefore, a correlated communication device can represent any pre-play correlated communication (a communication with uncertainties about players' strategies) given an incomplete information game as follows.

Observation 6. *Let us extend an incomplete information game G with the following arbitrary communication $\bar{C}$, thereby define $G_{\bar{C}}$. Before and after observing $t_i \in T_i$, players are engaged in finite communication stages. At each stage, players receive extra signals (these do not change their beliefs about Θ), send messages to a communication device, and receive communication outcomes from the communication device according to a predetermined rule. The communication protocol described here includes both ex-ante and interim communication, and it is the most general form of finite pre-play communication for a given G. Then any perfect Bayesian equilibrium outcome of some communication extension $G_{\bar{C}}$ of G as described above can be obtained by a correlated communication equilibrium outcome of G.*

Proof. Given an incomplete information game G, I show that any equilibrium outcome of an arbitrary finite communication extension $G_{\bar{C}}$ as described above can be replicated in an epistemic type space of $G_{\mathcal{C}}$, which is a 2-stage game of G extended by some communication protocol $\mathcal{C}$ with properties described in the Proposition 2. Define Ω as the set of final histories of $G_{\bar{C}}$.

For all $\omega \in \Omega$, define $\hat{t}(\omega)$ and $\hat{\theta}(\omega)$ given by G, and define $\hat{m}_i(\omega)$ to be player i's reporting strategy in the history ω. Let player i's information partition $\mathbb{P}_i^{M_i}$ consist of sets of final histories which share same reporting strategies of player i. Let information partition $\mathbb{P}_i^{A_i}$ to be player i's information partition at the final histories. Also, define $\hat{\beta}_i(\omega)$ so that it coincides with $\mathbb{P}_i^{A_i}$. Players' action strategies $\hat{A}_i^{O_i}$ are defined to coincide with action strategies with given information $\hat{\beta}_i(\omega)$. Check that rationality conditions of players are satisfied in the constructed epistemic type space.

$\square$

A large literature[23] discusses how to emulate equilibrium outcomes with mediated communication through some combination of direct communications. Direct communication, or cheap talk, is the most familiar form of communication, but the existence of a communication device as a mediator can seem to be artificial and unnatural. As noted in the literature, various extended protocols of cheap talk are powerful enough to mimic mediated talk in most cases: Given that the correlated communication equilibrium is identical to the communication equilibrium after belief invariant signals are added, observations from the literature regarding relationships between mediated and direct communication are applicable to relationships between correlated communication equilibrium and some extended protocols of correlated cheap talk.

Observation 7. *In most cases, a correlated communication equilibrium outcome of game G can be achieved by players who observe belief invariant sig-*

[23]See Forges (1990), Bárány (1992), Ben-Porath (2003), Gerardi (2004), Krishna (2007), and Vida and Forges (2013). Note that Vida and Forges (2013) defines a "correlated cheap talk" that is different from what is defined in this paper. Vida and Forges (2013) allow extra signals that are independent to original signals before repeated cheap talk game, in order to emulate a mediated communication by a repeated cheap talk preceded by independent extra signals. In contrast, the correlated cheap talk suggested in this paper allows players to observe extra signals that can be correlated to the original signals with belief invariance, and extra signals represent players' beliefs.

nals C and then engage in some extended direct communication protocol.

1.7 More Examples

In this section, I consider some additional examples.

General Mediated Communication Compared to Cheap Talk

As explained before, general mediated communication might lead to more meaningful communication (compared to cheap talk) if some information were hidden from players, because this would relax the incentive constraints for truthful reporting.

First, consider a persuasion game similar to the one described in the motivating example. There are two players, a speaker with private information about Θ and a listener who chooses an action and has no access to private information. Suppose that an epistemic type space corresponding to a correlated communication equilibrium is given in Figure 9. The row represents speaker's information partition $\mathbb{P}_S^{M_S} \equiv \{1, 2, 3, 4\}$. The column represents listener's information partition $\mathbb{P}_L^{M_L}$. The small five boxes represent listener's information partition $\mathbb{P}_L^{A_L}$. The letters a, b, c, d, e are recommended actions by the correlated communication device. In Figure 9, some information about Ω are omitted for the sake of simplicity.

Note that the decision rule induced by suggested correlated communication equilibrium cannot be implemented by a correlated cheap talk equilibrium because the amount of information received by the listener depends on the combination of inputs received from both players. When the listener is recommended to play b, she cannot differentiate speaker type 3 from 2, whereas that information is revealed if the listener is recommended to play c, d, or e. In other words, some information sent by the speaker is selectively revealed to listener types. In the case of cheap talk, the amount of

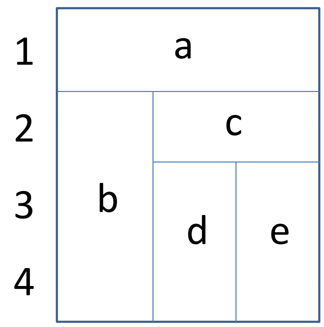


Figure 1.9: An epistemic type space corresponding to a correlated communication equilibrium

information received by all listener types should be the same. Therefore, in Figure 9, a decision rule implied by a given epistemic type space cannot be implemented by a correlated cheap talk equilibrium.

Next, consider a 3 player incomplete information game with $I \equiv \{1, 2, 3\}$ and a correlated communication equilibrium as follows. In an epistemic type space that corresponds to a given correlated communication equilibrium, player 1's information partition $\mathbb{P}_1^{A_1}$ contains information about the elements of $\mathbb{P}_2^{M_2}$, whereas $\mathbb{P}_3^{A_3}$ does not contain that information. The epistemic type space models a communication protocol where player 2 gives information to player 1 but not to player 3. A decision rule induced by a communication protocol of this type cannot be realized by cheap talk, which is a public message exchange.

In sum, compared to cheap talk, mediated communication might enable more communication and more equilibrium outcomes because it does not reveal all information to all players at the same time.

Correlated Cheap Talk with Persuasion Games

Figure 8-(b) shows that for the given game in the motivating example,

the correlated communication equilibrium gives strictly larger set of equilibrium outcomes compared to the communication equilibrium. Then one might be curious for which incomplete information games G, the correlated communication makes a difference in equilibrium predictions compared to the communication equilibrium. Currently I am unaware of general conditions on incomplete information games so that the correlated communication equilibrium gives strictly larger set of equilibrium outcomes compared to the communication equilibrium. Instead, I generalize a little from the motivating example to find sufficient conditions of meaningful communication with 2-player persuasion games.[24] Such an exercise might be illustrative and give us some ideas about when correlated communication makes a difference in equilibrium predictions.

Consider a persuasion game with 2 players, a speaker, S, and a listener, L. The speaker learns about the state of the world from a finite set Θ which is associated with a common prior p; the listener chooses an action from a finite set A. Assume that the speaker's utility depends only on A, so that $u_S : A \to \mathbb{R}$, and the listener's utility is given by $u_L : \Theta \times A \to \mathbb{R}$. A generalization of the equilibrium given in the motivating example gives us sufficient conditions for the existence of an equilibrium with meaningful communication.

Observation 8. *A 2 player persuasion game has a correlated cheap talk equilibrium with meaningful communication if the following conditions are satisfied.*

- *There are 3 actions* $x_1, x_2, x_3 \in A$

[24]Note that a 2-player persuasion game is an incomplete information game with a speaker and a listener. In a persuasion game, the speaker has private information about Θ without choosing an action and the listener chooses an action without private information. Moreover, the speaker's utility does not depend on Θ. In any 2-player persuasion game, all communication equilibrium imply that there is no meaningful communication.

- *There exist $b_1, b_2, b_3 \in \Delta\Theta$ such that given belief b_j, the listener's optimal action is x_j for $j = 1, 2, 3$.*

- *There exist $w_1, w_2 \in [0, 1]$ such that $w_1 b_1 + w_2 b_2 + (1 - w_1 - w_2) b_3 = p$.*

- *b_1 and b_2 share same support on Θ.*

- *The speaker ranks x_1, x_2, x_3 as $u_S(x_1) > u_S(x_3), u_S(x_2)$, without loss of generality.*

- *For every $\theta \in \Theta$ which is in the support of b_1 and b_2,*

$$w_1 b_1(\theta) u_S(x_1) + w_2 b_2(\theta) u_S(x_2) \geq u_S(x_3) \tag{1.9}$$

Proof. One can construct a correlated cheap talk equilibrium of a given persuasion game, such as in the epistemic type space described in Figure 10, with the conditions given in the Observation 8. In the given epistemic type space, the probability distributions of Θ of the boxes with associated actions x_1 and x_2 are given by $w_1 b_1$ and $w_2 b_2$ respectively. The distribution of the left side of the box with action x_3 is given by $(1 - \epsilon)(1 - w_1 - w_2) b_3$ with small ϵ; the right side of the box with x_3 is given by $\epsilon(1 - w_1 - w_2) b_3$.

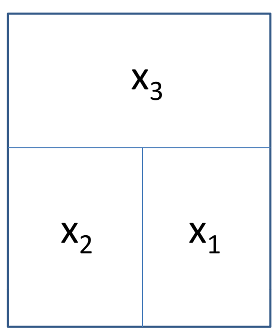


Figure 1.10: An epistemic type space of a correlated cheap talk equilibrium for a given persuasion game

□

Given a 2 player persuasion game, meaningful communication is possible if the following conditions are satisfied. First, there are 3 actions such that all 3 actions are listener's best responses depending on the listener's belief about Θ. Second, the listener's belief about Θ and about the speaker's messaging strategies is such that all 3 actions are optimal depending on received messages. Third, the speaker's belief about the listener's action strategy is such that all speaker types have incentives to send equilibrium messages. In a sense, the sufficient conditions suggested in Observation 8 show that players' interests and beliefs are aligned to enable meaningful communication in a given persuasion game.

1.8 Conclusion

In this paper, I suggest two new equilibrium concepts to represent reasonable outcomes of 2-stage games, in which simultaneous move incomplete information games are preceded by communication opportunities. I suggest the *correlated cheap talk equilibrium* for incomplete information games that are preceded by one-shot public message exchange, and I suggest the *correlated communication equilibrium* for incomplete information games that are preceded by general communication opportunities. New equilibrium concepts suggested in this paper are generalizations of Aumann's (1974) correlated equilibrium for complete information games to 2-stage communication extensions of incomplete information games.

Following Aumann (1987), I provide epistemic justifications of new definitions. The correlated cheap talk equilibrium represents the common knowledge of the Bayesian rationality of players with a common prior given a cheap talk game. The correlated communication equilibrium represents the common knowledge of the Bayesian rationality of players with a common prior given a game with a communication protocol. The epistemic analyses show that the extra signals given in the new definitions are not actual signals observed by players; extra signals are modeling devices that represent players' beliefs given a communication game. Also, the epistemic analyses show that belief invariance of extra signals represent the fact that players' information comes only from their original information or from communication outputs.

New definitions adopt Liu's (2015) belief invariant correlating device to 2-stage communication games, whereas Liu (2015) applied correlating devices to simultaneous move games to define belief invariant Bayes correlated equilibrium. In simultaneous move games, belief invariant extra signals work to correlate players' action strategies, and players' beliefs about the underlying states of the world do not change in an equilibrium. In 2-stage communication games, belief invariant extra signals work to correlate players' messaging and action strategies, and so extra signals enable more information transmission about the underlying states of the world in an equilibrium. Also, new definitions are different from Bergemann and Morris' (2013, 2015) Bayes correlated equilibrium so that belief updates of players come only from communication.

I conclude with several future research possibilities. First, the new equilibrium definitions suggested in this paper could be used to analyze various games with communication opportunities. For example, one might apply equilibrium definitions suggested in this paper to auction theory and to market design problems. Suppose that a market designer wants to know the set of possible equilibrium outcomes of certain forms of auctions - such as first price or second price auctions - knowing that players might communicate before the auction. The market designer would need to consider the set of correlated cheap talk equilibrium or the set of correlated communication equilibrium outcomes depending on communication protocols, instead of considering the set of cheap talk equilibrium or communication equilibrium outcomes.

Second, my idea regarding the uncertainties about message meanings might be applied to other forms of extensive games in addition to communication games. For example, one might apply the idea to contract theory. Consider contract games that are 2-stage games in which players make a contract in the first period and execute what is agreed on the contract in the second period. Contract games are different from communication games because, in contrast to non-binding communication, what is agreed to in the first stage is binding. In contact games, players' interpretations of what is written in a contract can be diverse, which can produce uncertainties and new equilibria, as in communication games. One might interpret the uncertainty about the interpretation of clauses in contracts as a source of contract incompleteness. I hope that the intuitions described in this paper improve understandings of various games and real world phenomena.

Bibliography

[1] Asheim, G.B., and Perea, A. (2005): "Sequential and Quasi-perfect Rationalizability in Extensive Games," *Games and Economic Behavior* 53(1), 15-42.

[2] Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics* 1, 67-96.

[3] Aumann,R. J. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica 55*, 1-18.

[4] Bárány, I. (1992): "Fair distribution protocols or how players replace fortune," *Mathematics of Operations Research* 17, 327-340.

[5] Barreda, I. M. (2013): "Cheap Talk with Two-Sided Private Information," working paper.

[6] Battigalli, P., and Siniscalchi, M. (1999): "Hierarchies of conditional beliefs and interactive epistemology in dynamic games," *Journal of Economic Theory* 88, 188-230

[7] Battigalli, P. and Siniscalchi, M. (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory* 106, 356-391

[8] Battigalli, P., and Siniscalchi, M. (2007): "Interactive Epistemology in Games with Payoff Uncertainty," *Research in Economics* 61, 165-184

[9] Ben-Porath, E. (1998): "Correlation without Mediation: Expanding the Set of Equilibrium Outcomes by Cheap Talk Pre-play Procedures," *Journal of Economic Theory* 80, 108-122

[10] Ben-Porath, E. (2003): "Cheap Talk in Games with Incomplete Information," *Journal of Economic Theory*108, 45-71

[11] Bergemann, D., and Morris. S. (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81-4, 1251-1308.

[12] Bergemann, D., and Morris. S. (2015): "Bayes Correlated Equilibrium and the Comparison of Information Structure in Games," *Theoretical Economics*, forthcoming.

[13] Blume, A., Board, O., and Kawamura, K. (2007): "Noisy Talk," *Theoretical Economics* 2, 395-440.

[14] Brandenburger, A., and Dekel, E. (1993): "Hierarchies of Belief and Common Knowledge," *Journal of Economic Theory*59, 189-198.

[15] Crawford, V.P., and Sobel, J. (1982): "Strategin Information Transmission," *Econometrica*, 50-6, 1431-1451.

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

[17] Forges, F. (1986): "An Approach to Communication Equilibrium," *Econometrica*, 54-6 1375-1385.

[18] Forges, F. (1990): "Universal Mechanism," *Econometrica*, 58-6 1341-1364.

[19] Forges, F. (2006): "Correlated Equilibrium in Games with Incomplete Information Revisited," *Theory and Decision* 61, 329-344.

[20] Gerardi, D. (2004): "Unmediated Communication in Games with Complete and Incomplete Information," *Journal of Economic Theory*, 114, 104-131

[21] Gerardi, D., and Myerson, R. B. (2007): "Sequential Equilibria in Bayesian Games with Communication," *Games and Economic Behavior*, 60, 104-134

[22] Glazer, J., and Rubinstein, A. (2012): "A Model of Persuasion with Boundedly Rational Agent," *Journal of Political Economy* 120-6 1057-1082.

[23] Ishida, J., and Shimizu, T. (2012): "Cheap Talk with an Informed Receiver," working paper.

[24] Kamenica, E., and Gentzkow, M. (2011): "Bayesian Persuasion," *American Economic Review* 101, 2590 - 2615.

[25] Krishna, V., and Morgan, J. (2004): "The Art of Conversation: Eliciting Information from Experts through Multi-stage Communication," *Journal of Economic Theory* 117, 147-179.

[26] Lambie-Hanson. T., and Parameswaran, G. (2015): "A Model of Vague Communication," *working paper.*

[27] Lipman, B. L. (2009): "Why is Language Vague?," working paper.

[28] Liu, Q. (2015): "Correlation and Common Priors in Games with Incomplete Information," *Journal of Economic Theory* 157, 49-75.

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

[30] Vida, P., and Forges, F. (2013): "Implementation of Communication Equilibria by Correlated Cheap Talk: The Two Player Case," *Theoretical Economics* 8, 95-123.

Chapter 2

Common Prior Type Spaces In Which Payoff Types and Belief Types Are Independent

2.1 Introduction

The notion of a type space is central to the analysis of games with incomplete information (Harsanyi, 1967-68) and to mechanism design (e.g. Myerson, 1981, Bergemann and Morris, 2005). Types describe agents' payoff relevant as well as other, payoff irrelevant information, and also agents' beliefs about other agents' types, and agents' beliefs about other agents' beliefs, etc. Bayesian Nash equilibria, or, for example, correlated equilibria of games are defined with respect to a given type space. Type spaces are flexible modeling devices that can describe complex belief structures.

Applied game theory often focuses on common prior type spaces in which all information that an agent receives is payoff relevant. If we call the payoff relevant information of an agent that agent's "payoff type," [1] then types and payoff types are the same for each agent if all information is payoff relevant. Two special classes with such property have received special attention. One such class consists of type spaces in which different agents' types are independent (e.g. Myerson 1981). An assumption embedded in this construction is that agents' first order beliefs about other agents' types are the same, irrespective of their own type. This implies that agents' first order beliefs about other agents' types are common knowledge among the agents.

A second special class of type spaces that are frequently studied in the literature are type spaces in which no agent has two distinct types with identical hierarchies of beliefs. Referring to an agent's hierarchy of beliefs about another agents' types as the agent's "belief type," these type spaces are characterized by the property that "belief types determine payoff types." Implicit in this construction is the assumption that the function mapping belief types into payoff types is common knowledge among agents. In mechanism design these type spaces often allow the construction of mechanisms that elicit agents' beliefs about other agents, and by doing so also elicit agents' payoff types. Agents then earn no information rents, and the mechanism designer can "extract the full surplus" (Crémer and McLean, 1985, 1988, Neeman, 2004). A recent line of work has examined whether the sets of type spaces that have the "belief types determine payoff types" property, or that allow "full surplus extraction," are generic (Heifetz and Neeman, 2006, Chen and Xiong, 2011a, 2011b, Gizatulina and Hellwig, 2011).

The polar opposite of the condition that belief types determine payoff types is the condition that belief types and payoff types are stochastically independent for every agent, so that knowing the belief type of an agent

[1] We borrow the expression "payoff type" from Bergemann and Morris (2005).

does not allow any inferences at all about that agent's payoff type.[2] In this paper we investigate the class of type spaces for which this opposite condition, to which we shall refer as the "independence property," is true. We are interested in type spaces satisfying this strong condition because an analysis of games or mechanism design problems for such type spaces allows the modeler to exclude all effects due to correlation between payoff and belief types. Moreover, it will turn out that large portions of the existing game theoretic and mechanism design literature can be re-interpreted as being concerned with exactly the class of type spaces that satisfy the independence condition.

We restrict attention to type spaces in which agents' beliefs are derived from a common prior. We allow type spaces in which an agent's type includes payoff-irrelevant information. Type spaces in which types are independent obviously have the independence property because in such type spaces all types of a given agent have the same belief types so that belief types are constant, and constant random variables are stochastically independent of any other random variable. Our interest is in the question whether there are other type spaces with the independence property. We answer this question positively, and we characterize all type spaces with the independence property. All such type spaces can be interpreted as follows: Agents have independent payoff types. They also receive further information that is potentially not independent among agents, but that is independent of all agents' payoff types. Therefore, all types of a given agent have the same belief about other agents' payoff types, as is the case in type spaces with independent types,

[2]Note that the condition that we investigate is in an informal sense the *opposite*, but importantly by no means the *negation* of the "beliefs determine preferences" condition. The negation encompasses the condition that we study in this paper, but is far more general. "Environment 2""in Neeman (2004) is an example of a common prior type space in which one agent's belief types don't determine that agent's payoff types, but in which this agent's belief and payoff types are not stochastically independent either.

but different types of the same agent may hold different beliefs about other agents' payoff irrelevant information. Thus, the class of type spaces with the independence property is a generalization of the class of type spaces with independent types. A simple, and not surprising, implication of our result is that common priors for which belief and payoff types are *not* independent are generic in the senses considered in the literature on the genericity of the "beliefs determine preferences" property which we mentioned earlier.[3]

What is remarkable about our characterization is that we begin with an independence assumption that refers to each agent separately: each agent's payoff type and belief type are independent, and we show that this is equivalent to a form of independence across agents: different agents' payoff types are independent, and payoff irrelevant information is independent of all agents' payoff types. Figuratively speaking, independence propagates from each agent separately to the group of agents as a whole.

Using the language of the recent literature, type spaces with the independence property differ from type spaces with independent types only through the introduction of "redundant types," that is, multiple types that have the same payoff types, and the same hierarchies of beliefs regarding the underlying payoff relevant uncertainty. There is thus a connection between our main result and Theorems 1 and 2 in Liu (2011), who characterizes for general common prior type spaces the connection between type spaces with redundant types, and the same type spaces without redundant types. He shows for common prior type spaces that the type space with redundant type is obtained from the corresponding type space without redundant types by adding a common prior correlation device where the correlation is conditional on the vector of agents' payoff types. Our result shows in a common prior context that the independence property holds if and only if different players' payoff

[3]As this is straightforward to see, but tedious to state formally, we have not included this observation in the main body of the paper.

types are independent of each other, and the payoff irrelevant information is independent of all players' payoff types.

Our analysis is subtly related to Aumann and Brandenburger (1995). Seeking an epistemic foundation for Nash equilibrium, they infer in their Theorem B from the assumption that beliefs are common knowledge that beliefs must be product measures. Although their model and their motivation are entirely different from ours, the proof of our main result includes an important step that is also included in Aumann and Brandenburger's proof of their Theorem B. At the end of Section 3, we shall comment further on the relation between Aumann and Brandenburger's result and ours.[4]

In the last two sections of the paper we describe the implications of our analysis for game theory and mechanism design. In game theory an exploration of the Bayesian equilibria of a strategic form game using a type space with the independence property is equivalent to the exploration of the "strategic form correlated equilibria" (Cotter, 1991, Forges, 1993) of the game with the type space in which the payoff irrelevant information is omitted. This result is closely related to Lemma 2 in Liu (2011). However, Liu studies general type spaces, and therefore his result refers to a more general version of correlated equilibrium than ours. In his version of correlated equilibrium, before suggesting strategies to agents, the "mediator" observes the agents' types. By contrast, in "strategic form correlated equilibrium" the "mediator" does not observe agents' types before recommending strategies.

In mechanism design we show for a wide variety of possible objectives of the mechanism designer, that mechanisms that are optimal for a type space with the independence property are essentially the same as the mechanisms that are optimal for the corresponding type space in which no payoff irrelevant information is provided.

[4]We are very grateful to Qingmin Liu for pointing out the relation between our result and Aumann and Brandenburger's result.

2.2 Framework

There are $n \geq 2$ agents. We write N for the set of agents. For each agent $i \in N$ there is a finite set Θ_i of possible "payoff types" θ_i of agent i. We borrow the expression "payoff type" from Bergemann and Morris (2005), where payoff types are the possible realizations of a signal that agent i observes, and whose realizations potentially affect i's own or other agents' payoffs in a game. The payoff type is the only signal that i observes that may affect payoffs. Agent i may make other observations, but these don't affect payoffs. In this and the next section, payoff types are in fact completely abstract. In these sections it is irrelevant whether there is an underlying game. In Sections 4 and 5, the interpretation of the elements of Θ_i as payoff types will, by contrast, be important. For concreteness, we shall even in Sections 2 and 3 occasionally interpret payoff types as payoff relevant information, and the reader may have this interpretation in mind throughout.

Throughout the paper, we use notations such as $\theta \in \Theta \equiv \prod_{i \in N} \Theta_i$, and $\theta_{-i} \in \Theta_{-i} \equiv \prod_{j \neq i} \Theta_j$. Also, for any non-empty, finite set X, we denote by $\Delta(X)$ the set of all probability distributions on X.

We use type spaces to describe the agents' beliefs about their own and others' payoff types, their beliefs about these beliefs, etc. The modeling device of type spaces is due to Harsanyi (1967-68). The focus of this paper is on type spaces with a common prior. The analysis does not apply to type spaces with subjective priors.To keep our analysis straightforward, we restrict attention to finite type spaces where the common prior has full support.

Definition 6. *A* type space *is a list* $((T_i)_{i \in N}, (\hat{\theta}_i)_{i \in N}, \mu)$ *such that:*

1. *for every* $i \in N$, T_i *is a non-empty, finite set;*

2. *for every* $i \in N$, $\hat{\theta}_i$ *is a function of the form:* $\hat{\theta}_i : T_i \to \Theta_i$*;*

3. $\mu \in \Delta(T)$ *where* $T \equiv \prod_{i \in N} T_i$*;*

4. $\mu(t) > 0$ *for all* $t \in T$*.*

Here, a (standard) implicit assumption is that the type space is common knowledge, and that each agent i observes her own type t_i, but not other agents' types t_{-i}.

Without loss of generality, we assume that the range of $\hat{\theta}_i$ is Θ_i. Writing $\mu(t_{-i}|t_i)$ for the conditional probability of t_{-i} where we condition on i's type being t_i, we define next:

Definition 7. *For a given type space* $((T_i)_{i \in N}, (\hat{\theta}_i)_{i \in N}, \mu)$*, agent* i*'s belief function*

$$\hat{b}_i : T_i \to \Delta(T_{-i}) \tag{2.1}$$

is defined by:

$$\hat{b}_i(t_i)(t_{-i}) = \mu(t_{-i}|t_i) \tag{2.2}$$

for every $t_i \in T_i, t_{-i} \in T_{-i}, i \in N$*.*

Thus, $\hat{b}_i(t_i)$ is the belief about other players' types that agent i holds if her type is t_i. This belief is derived from the prior μ by conditioning on t_i. We shall refer to $\hat{b}_i(t_i)$ also as agent i's "belief type." We write B_i for the range of $\hat{b}_i$. B_i is thus the set of all belief types. We shall write $\mu(\theta_i, b_i)$ for the probability assigned by μ to the set of all type vectors such that agent i's preference is θ_i and agent i's belief is b_i, and similarly use notation such as $\mu(\theta_i)$, $\mu(b_i)$, etc.

We make throughout the following assumption which says in words that there are no "duplicate types:"

Assumption 1. *For every* $i \in N$*, if* $t_i, t'_i \in T_i$ *and* $t_i \neq t'_i$*, then* $\hat{\theta}_i(t_i) \neq \hat{\theta}_i(t'_i)$ *or* $\hat{b}_i(t_i) \neq \hat{b}_i(t'_i)$*.*

Duplicate types, that we rule out, are thus types with identical payoff types and with identical beliefs. To apply our main result to type spaces in which duplicate types exist, one has to successively "merge" duplicate types into a single type. Assumption 1 means that every type t_i is uniquely identified by t_i's payoff type $\hat{\theta}_i(t_i)$ and t_i's belief type $\hat{b}_i(t_i)$.

Note that Assumption 1 does not rule out what the literature refers to as "redundant types," that is, multiple types with identical payoff types and hierarchies of beliefs about other players' payoff types. This is because a players' type may encode more information than just the player's payoff type and the players' beliefs about other players' payoff types. This point is crucial for our paper. The potential importance of redundant types for the analysis of incomplete information games has been emphasized by Forges (1993, pp. 284/5). The following discussion of the role of redundant types is taken from Liu (2009, p. 2117):

> "..., if the analyst knows only the payoff structures - he is unaware of (or unable to specify) some other variables that the players know, ... , but he is aware of his unawareness (or misspecification) - then a redundant type structure is a "safe" modeling choice: the players "reason" within a redundant structure *as if* they were reasoning about some parameters unknown to the analyst. In other words, the analyst should not make use of a redundant structure unless he is not sure of the players' space of basic uncertainties."

Liu (2009, 2011) provides formal results that support this interpretation of redundant types, and that apply to our model. When allowing redundant types in our model it is Liu's interpretation that we have in mind, and thus we allow that the type space is constructed by an analyst who is aware that he is unaware of some variables that players may have beliefs about.

The property of type spaces in which we are interested in this paper is the following:

Definition 8. *A type space* $((T_i)_{i \in N}, (\hat{\theta}_i)_{i \in N}, \mu)$ *has the* independence property *if for every* $i \in N$ *the random variables* $\hat{\theta}_i$ *and* $\hat{b}_i$ *are independent.*

As explained in the Introduction, we view this property as the polar opposite of the "beliefs determine payoff types" property. In type spaces with the independence property, knowing an agent's beliefs provides no information about that agent's preferences. If a type space has the independence property with no duplicate types as in Assumption 1, we can relabel the type space as follows:

$$T_i = \Theta_i \times B_i \text{ for all } i \in N. \tag{2.3}$$

The Assumption 1 and the relabeling of type spaces will make our representations clearer and easier.

2.3 Result

Before stating our result, we give an example that illustrates the result. We observed already in the Introduction that every naive type space with independent types has the independence property trivially because each agent's beliefs are constants. Type spaces with independent types, however, embed a very restrictive common knowledge assumption: each agent's first order beliefs are common knowledge. We therefore give an example in which agents beliefs about the other agents' types are not constant, and the agents' first order beliefs are not common knowledge.

Example 1. $N = \{1, 2\}$. *For every* $i \in N$*, the set of payoff types is* $\Theta_i = \{\theta_i^1, \theta_i^2\}$*, and the set of types is* $T_i = \{t_i^k : k = 1, 2, 3, 4\}$*. Payoff types are given by* $\hat{\theta}_i(t_i^1) = \hat{\theta}_i(t_i^2) = \theta_i^1$ *and* $\hat{\theta}_i(t_i^3) = \hat{\theta}_i(t_i^4) = \theta_i^2$ *for* $i = 1, 2$*. The* *common prior μ is described in Figure 2. Conditional on agent 1's payoff type being θ_1^k, his beliefs about agent 2's types are $(1/6, 2/6, 1/6, 2/6)$ with probability 0.5, and $(2/6, 1/6, 2/6, 1/6)$ with probability 0.5. This probability does not depend on k. Therefore, for agent 1, beliefs and payoff types are independent. A similar calculation shows that also for agent 2 beliefs and payoff types are independent.*

μ	t_2^1	t_2^2	t_2^3	t_2^4
t_1^1	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{1}{24}$	$\frac{2}{24}$
t_1^2	$\frac{2}{24}$	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{1}{24}$
t_1^3	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{1}{24}$	$\frac{2}{24}$
t_1^4	$\frac{2}{24}$	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{1}{24}$

Figure 2.1: The common prior μ in Example 1

There is an equivalent representation of the type space in Example 1. Note that in Example 1 the pair of the agents' payoff types, $(\hat{\theta}_1, \hat{\theta}_2)$, is independent of the pair of the agents' belief types, $(\hat{b}_1, \hat{b}_2)$. This is a stronger property than the independence property which only requires independence of payoff types and belief types agent by agent. In Example 1 one can then imagine types being determined by two independent draws: one draw determines (θ_1, θ_2), and another draw determines (b_1, b_2). We describe these draws in Figure 2, where the left square represents the common prior for the draw of (θ_1, θ_2), and the right square represents the common prior for the pair (b_1, b_2). We denote the common prior distribution of payoff types by μ', the two possible belief types of each agent by b_i^1 and b_i^2 (in the order that they were listed in the description of Example 1), and the common prior distribution of belief types by μ''.

μ'	θ_2^1	θ_2^2
θ_1^1	$\frac{1}{4}$	$\frac{1}{4}$
θ_1^2	$\frac{1}{4}$	$\frac{1}{4}$

μ''	b_2^1	b_2^2
b_1^1	$\frac{1}{6}$	$\frac{2}{6}$
b_1^2	$\frac{2}{6}$	$\frac{1}{6}$

Figure 2.2: An equivalent representation of the common prior in Example 1: $\mu = \mu' \times \mu''$.

Now note a further independence: the distribution of payoff types is a product distribution, that is, payoff types are independent across agents. This implies that agents' beliefs about other agents' payoff types are in fact constant in the model, and therefore common knowledge, as they are when types are drawn independently. The variation in agents' beliefs stems from the variation in their beliefs about other variables, that are not payoff related. These are captured by the belief distribution in Figure 2. Note that this distribution is not a product distribution.

The main result of this paper is that a similar representation as the one in Figure 2 can be given for any type space with the independence property.

Proposition 3. *A type space has the independence property if and only if*

$$\mu(t_1, t_2, \ldots t_N) = \mu(\theta_1)\mu(\theta_2)\ldots\mu(\theta_n)\mu(b_1, b_2, \ldots, b_n) \tag{2.4}$$

for all $(t_1, t_2, \ldots, t_n) \in T$.

Proof. It is immediate that (2.4) implies that $\hat{\theta}_i$ and $\hat{b}_i$ are independent for each agent. We prove that (2.4) is necessary for the independence property in three claims.

Claim 1: For all $i \in N, \theta_i \in \Theta_i, b_i \in B_i, \theta_{-i} \in \Theta_{-i}, b_i \in B_{-i}$:

$$\mu(\theta_{-i}, b_{-i}|\theta_i, b_i) = \mu(\theta_{-i}, b_{-i}|b_i) \tag{2.5}$$

Proof.

$$\begin{aligned}
\mu(\theta_{-i}, b_{-i}|b_i) &= \frac{\mu(\theta_{-i}, b_i, b_{-i})}{\mu(b_i)} \\
&= \sum_{\theta_i' \in \Theta_i} \frac{\mu(\theta_i', \theta_{-i}, b_i, b_{-i})}{\mu(b_i)} \\
&= \sum_{\theta_i' \in \Theta_i} \frac{\mu(\theta_i', b_i)}{\mu(b_i)} \frac{\mu(\theta_i', \theta_{-i}, b_i, b_{-i})}{\mu(\theta_i', b_i)} \\
&= \sum_{\theta_i' \in \Theta_i} \mu(\theta_i'|b_i)\mu(\theta_{-i}, b_{-i}|\theta_i', b_i) \\
&= \sum_{\theta_i' \in \Theta_i} \mu(\theta_i'|b_i) b_i(\theta_{-i}, b_{-i}) \\
&= \sum_{\theta_i' \in \Theta_i} \mu(\theta_i'|b_i)\mu(\theta_{-i}, b_{-i}|\theta_i, b_i) \\
&= \mu(\theta_{-i}, b_{-i}|\theta_i, b_i) \qquad (2.6)
\end{aligned}$$

Here, the fifth and sixth line follow from the definition of belief types.

Claim 2: For all $i \in N, \theta \in \Theta, b \in B$:

$$\mu(\theta, b) = \mu(\theta_i)\mu(\theta_{-i}, b). \qquad (2.7)$$

Proof.

$$\begin{aligned}
\mu(\theta, b) &= \mu(\theta_i, b_i)\mu(\theta_{-i}, b_{-i}|\theta_i, b_i) \\
&= \mu(\theta_i)\mu(b_i)\mu(\theta_{-i}, b_{-i}|\theta_i, b_i) \\
&= \mu(\theta_i)\mu(b_i)\mu(\theta_{-i}, b_{-i}|b_i) \\
&= \mu(\theta_i)\mu(\theta_{-i}, b) \qquad (2.8)
\end{aligned}$$

The second line follows from the independence property, and the third line follows from Claim 1.

CLAIM 3:[5] If for all $i \in N, \theta \in \Theta, b \in B$:

$$\mu(\theta, b) = \mu(\theta_i)\mu(\theta_{-i}, b), \tag{2.9}$$

then for all $\theta \in \Theta, b \in B$:

$$\mu(\theta, b) = \mu(\theta_1)\mu(\theta_2)\dots\mu(\theta_N)\mu(b) \tag{2.10}$$

Proof. We prove this by induction over n, beginning with the case $n = 2$. By assumption:

$$\mu(\theta_1, \theta_2, b) = \mu(\theta_2)\mu(\theta_1, b). \tag{2.11}$$

Therefore, we can complete the proof by showing:

$$\mu(\theta_1, b) = \mu(\theta_1)\mu(b). \tag{2.12}$$

By assumption:

$$\mu(\theta_1, \theta_2, b) = \mu(\theta_1)\mu(\theta_2, b). \tag{2.13}$$

Summing (2.13) over all θ_2, we obtain (2.12).

Now suppose the claim had been shown for all numbers of agents up to some number $n \geq 2$. We prove the claim for $n + 1$. By assumption:

$$\mu(\theta, b) = \mu(\theta_{n+1})\mu(\theta_{-(n+1)}, b) \tag{2.14}$$

Therefore, we can complete the proof by showing:

$$\mu(\theta_{-(n+1)}, b) = \mu(\theta_1)\mu(\theta_2)\dots\mu(\theta_n)\mu(b) \tag{2.15}$$

[5] CLAIM 3 and its proof are identical to Lemma 4.6 and its proof in Aumann and Brandenburger (1995), except that the type space in Aumann and Brandenburger's model does not include a component that is analogous to the component "b" in our type space. We comment further on the relation between our work and that of Aumann and Brandenburger in the paragraph following the proof of Proposition 1.

We prove this using the inductive assumption. For this, it is sufficient to show that the "if-condition" of CLAIM 3 holds for $(\theta_{-(n+1)}, b)$:

$$\mu(\theta_{-(n+1)}, b) = \mu(\theta_i)\mu(\theta_{-(i,n+1)}, b) \tag{2.16}$$

for all $i \neq n+1$. This is implied by the "if-condition" of CLAIM 3 for (θ, b):

$$\mu(\theta, b) = \mu(\theta_i)\mu(\theta_{-i}, b) \tag{2.17}$$

if we sum over all θ_{n+1}. □

Proposition 1 is subtly related to Theorem B in Aumann and Brandenburger (1995). In Aumann and Brandenburger's model a type space describes hierarchies of beliefs over strategies, not over payoff types. However, one can reinterpret their model, replacing strategies by payoff types. Aumann and Brandenburger then investigate the assumption that beliefs about other players' payoff types are common knowledge. They infer that beliefs have to be product measures.[6] Their assumption is stronger than ours, as the assumption that beliefs are common knowledge implies that they are independent of payoff types, but it is in another sense weaker, because it only refers to beliefs about payoff types, not to beliefs about types per se. Their conclusion is similar to ours, except that their conclusion does not address the possible existence of redundant types.

We can translate Aumann and Brandenburger's result into our setting. Suppose we say that a fact is common knowledge in our model if it is true for every $t \in T$.[7] In particular, let us say that agent i's beliefs are common

[6]Combining this with the assumption of mutual knowledge of rationality, they obtain that beliefs form a Nash equilibrium.

[7]Our assumption of full support beliefs for every type implies that the standard definition of a fact being common knowledge reduces in our model to the condition that the fact is true for all t.

knowledge if there is some $b_i \in \Delta(T_{-i})$ such that $\hat{b}_i(t_i) = b_i$ for all $t_i \in T_i$. Then Aumann and Brandenburger's proof of their Theorem B shows:

Remark 1. *All agents' beliefs are common knowledge if and only if for all $t \in T$:*

$$\mu(t) = \mu(t_1)\mu(t_2)\ldots\mu(t_n). \tag{2.18}$$

The proof of this remark is essentially the same as the proof of Proposition 1. In particular, to show that (2.18) is necessary for beliefs to be common knowledge, one begins with the observation that the constancy of belief types implies: $\mu(t_{-i}|t_i)\mu(t_{-i})$ for all $t \in T$, which is the analog of Claim 1. The proof continues with analogs of Claims 2 and 3, omitting, as in the analog of Claim 1, the conditioning on belief types, as belief types are the same everywhere in the type space.

We mentioned already in the Introduction and at the beginning of this section that type spaces with independent types are important in the literature, yet extremely special. We noted at the beginning of this section that independent types imply that beliefs about others' types are common knowledge. Remark 1 adds to this the observation that the reverse is also true: common knowledge of beliefs implies that types are independent. Remark 1 thus characterizes the most prominent special case of type spaces with the independence property.[8]

2.4 Implications for Game Theory

Now we introduce a game played by the n agents whom we have also considered so far. The (finite) sets of pure actions in this game are: $S_1, S_2, \ldots, S_n$. Also, for each player a utility function $u_i : S_1 \times S_2 \times \ldots \times S_n \times \Theta \to \mathbb{R}$ is

[8] A small caveat is that Remark 1, unlike our earlier comments, covers type spaces in which types and payoff types are not the same.

given. If we combine a type space with the action sets and utility functions, then we obtain a game of incomplete information. We shall refer to this game as "the incomplete information game generated by the type space." A pure strategy of player i is a mapping: $\sigma_i : \Theta_i \to S_i$. Denote the set of all pure strategies of player i by Σ_i. We define: $\Sigma = \prod_{i \in N} \Sigma_i$.

Our goal is to find a connection between the Bayesian equilibria of a game generated by a type space that has the independence property and the equilibria, for an appropriate equilibrium concept, of the game in which we have dropped the payoff irrelevant component from the type space. If we find such a relation, it will be possible to analyze games with independent payoff and belief types without taking account of the possibility of payoff irrelevant information, and yet at the same time capture the results that an analysis of the Bayesian equilibria of all incomplete information games generated by a type space with independent payoff and belief types would yield.

It turns out that for our purposes the relevant equilibrium concept for the analysis of the game without payoff irrelevant information is a version of correlated equilibrium. Care is needed regarding the precise definition of a correlated equilibrium. Cotters (1991, 1994), Forges (1993, 2006), Liu (2011), and others distinguish different notions of correlated equilibria of incomplete information games. In this paper the appropriate notion is what Forges refers to as "strategic form correlated equilibrium" (Cotters, 1991, and Forges, 1993). A "strategic form correlated equilibrium" is a probability distribution γ on Σ that is a correlated equilibrium in the sense of Aumann (1974, 1987) of the strategic form of the incomplete information game. A Bayesian equilibrium is a strategic form correlated equilibrium γ that is the product of its marginals on the n pure strategy sets Σ_i.[9]

To conduct our analysis formally, we next need to be precise about what

[9]To simplify our notation, we use Milgrom and Weber's (1985) distributional approach to the representation of mixed strategies.

it means to drop the sets B_i from a type space, and what it means to re-introduce them. This is done in the following definition.

Definition 9. *(i) For given type space with the independence property* $((T_i)_{i\in N},$ $(\hat{\theta}_i)_{i\in N},\ \mu)$ *such that:* $T_i = \Theta_i \times B_i$*, the* corresponding reduced type space $((T_i')_{i\in N},\ (\hat{\theta}_i')_{i\in N}, \mu')$ *is:* $((T_i')_{i\in N},\ (\hat{\theta}_i')_{i\in N}, \mu')$ *where:* $T_i' = \Theta_i$ *for all* $i \in N$, $\hat{\theta}_i'(\theta_i) = \theta_i$ *for all* $i \in N$ *and* $\theta_i \in \Theta_i$*, and* $\mu'(\theta_1, \theta_2, \ldots, \theta_n) = \mu(\theta_1)\mu(\theta_2)\ldots\mu(\theta_n)$ *for all* $(\theta_1, \theta_2, \ldots, \theta_n) \in \Theta$.

(ii) For given type space with the independence property $((T_i)_{i\in N}, (\hat{\theta}_i)_{i\in N}, \mu)$ *where* $T_i = \Theta_i$ *for all* $i \in N$*, a* corresponding augmented type space $((T_i')_{i\in N},$ $(\hat{\theta}_i')_{i\in N}, \mu')$ *is a type space with the independence property such that the corresponding reduced space is* $((T_i)_{i\in N}, (\hat{\theta}_i)_{i\in N}, \mu)$.

Next, we introduce a correspondence between a vector of mixed strategies for an incomplete information game generated by a type space with the independence property and a correlated strategy for the incomplete information game generated by the type space for which the payoff irrelevant information is dropped.

Definition 10. *Let* γ *be a product distribution on the set of pure strategy combinations* Σ *in the incomplete information game generated by a type space with the independence property. Then the* equivalent probability distribution γ' on the set of pure strategy combinations Σ' in the game generated by the corresponding reduced type space *is defined by:*

$$\gamma'(s') = \sum_{s\in\Sigma} \left(\gamma(s)\mu\left(\{(b_1, \ldots, b_N) | s_i(\theta_i, b_i) = s_i'(\theta_i) \text{ for all } i \in N \text{ and } \theta_i \in \Theta_i\}\right)\right) \tag{2.19}$$

for all $s' \in \Sigma'$.

Our result is:

Proposition 4. *(i) Let γ be a Bayesian equilibrium of the incomplete information game generated by a type space with the independence property. Then the equivalent probability distribution γ' on the set of pure strategy combinations Σ' in the incomplete information game generated by the corresponding reduced type space is a strategic form correlated equilibrium of that incomplete information game.*

(ii) Let γ' be a strategic form correlated equilibrium of the incomplete information game generated by a type space with the independence property in which $T_i = \Theta_i$ for all $i \in N$. Then there is a corresponding augmented type space, and a product distribution γ on the space of pure strategies Σ' in the incomplete information game generated by the augmented type space such that γ' is equivalent to γ, and such that γ is a Bayesian equilibrium of that incomplete information game.

Proposition 2 is a re-statement of the revelation principle for our model. We therefore omit a formal proof. Cotter (1991, p. 54) and Forges (1993, p. 289) observed that the revelation principle applies to the strategic form correlated equilibrium. An appropriately phrased version of part (i) of Proposition 2 remains true if one replaces strategic form correlated equilibrium by agent normal form correlated equilibrium, because, roughly speaking, every strategic form correlated equilibrium is also an agent normal form correlated equilibrium (Forgers, p. 290). It is not true, however, that every agent normal form correlated equilibrium is a strategic form correlated equilibrium (see Example 3 in Forges (1993)), and thus part (ii) of Proposition 2 does not hold for agent normal form correlated equilibria.

The question answered by Proposition 2 for Bayesian equilibria can also be asked for other game theoretic solution concepts. An alternative to Bayesian equilibria is in particular the concept of rationalizability. Several notions of rationalizability for incomplete information games have been proposed in the literature. If we employ the concept of "interim correlated rationalizability" as defined by Dekel, Fudenberg and Morris (2007), then the result is simple. According to Proposition 1 in Dekel, Fudenberg and Morris (2007), the set of interim correlated rationalizable strategies of a player only depends on that player's hierarchy of beliefs about payoff relevant information. It is not affected by payoff irrelevant information included in the type space. Therefore, it is without loss of generality in our context, in which we postulate the independence property, to analyze the set of interim correlated rationalizable strategies using the reduced type space in which only payoff types are included.

2.5 Implications for Mechanism Design

Next, we examine the implications of our analysis for mechanism design. We consider the same n agents as in the previous sections, as well as a mechanism designer. There are a (finite) set of possible outcomes Y, and for every agent i a utility function $u_i : \Theta \times Y \to \mathbb{R}$. The mechanism designer supposes that the agents' information is described by a type space with the independence property. The mechanism designer chooses a game form, consisting of strategy sets for each agent, a mapping of strategies into outcomes, and a Bayesian equilibrium of the incomplete information game defined by the game form, the utility functions, and the type space. We leave the mechanism designer's objective function unspecified except that we assume that it only depends on the implied mapping between agents' payoff types and probability distributions over outcomes. By the revelation principle we can restrict attention to direct game forms $q : T \to \Delta(Y)$ such that truth telling is a Bayesian equilibrium in the corresponding incomplete information game. We refer to such direct game forms as "incentive compatible."

Our objective is to find a correspondence between the direct and incentive compatible mechanisms for a type space with the independence property and the direct and incentive compatible mechanisms for the corresponding reduced type space. Here, we use the terminology for type spaces introduced in the previous section. We shall find such a correspondence if we focus on the mapping between payoff types and probability distribution over outcomes. As we have postulated that the mechanism designer's objective depends only on that mapping, our result implies that mechanisms that are optimal for a type space with the independence property and mechanisms that are optimal for the the corresponding type space in which all payoff irrelevant information has been removed can achieve exactly the same values of the mechanism designer's objective function. It is therefore without loss of generality to study the mechanism designer's maximization problem only for the reduced type space, as the literature has mostly done.

We first define how we relate direct mechanisms for a type space with the independence property to direct mechanisms for the same type space but without payoff irrelevant information.

Definition 11. *(i) Consider a direct mechanism $q : T \to \Delta(Y)$ for a type space with the independence property. The* equivalent direct mechanism for the corresponding reduced type space *is the mechanism $q' : T' \to \Delta(Y)$ where for every $\theta \in \Theta$ and $y \in Y$ we have:*

$$q'(\theta, y) = \sum_{b \in B} \left(q((\theta, b), y)\mu(b) \right). \tag{2.20}$$

Here $q(t, y)$ denotes the probability that a direct mechanism assigns to outcome y when the vector of types is t.

(ii) Consider a type space with the independence property where $T'_i = \Theta_i$ for all $i \in N$, and a corresponding augmented type space*. Let $q : T' \to \Delta(Y)$ be a direct mechanism for the first type space. Then the* equivalent direct mechanism for the augmented type space *is the mechanism $q : T \to \Delta(Y)$*

where for every $t = (\theta_i, b_i)_{i \in N}$ and $y \in Y$ we have:

$$q(t, y) = q'(\theta, y). \tag{2.21}$$

Our result is:

Proposition 5. *(i) If a direct mechanism $q : T \to \Delta(Y)$ for a type space with the independence property is incentive compatible, then the equivalent direct mechanism for the corresponding reduced type space is incentive compatible.*

(ii) If a direct mechanism $q : T \to \Delta(Y)$ for a type space with the independence property and with $T_i' = \Theta_i$ for all $i \in N$ is incentive compatible, then the equivalent direct mechanism for a corresponding augmented type space is incentive compatible.

Part (ii) is immediate, as in the augmented type space agents simply ignore the payoff irrelevant information B which then is strategically irrelevant as well. Like Proposition 2, part (i) of Proposition 3 is a version of the revelation principle. In particular, suppose the true type space were the reduced type space, but the mechanism designer provided the payoff irrelevant information B to agents as part of an extensive form mechanism. By the standard revelation principle, the mechanism could collapse such an extensive form mechanism into a direct mechanism in which truth-telling is an equilibrium. This is essentially what part (i) of Proposition 3 says. We omit the proof of Proposition 3.

Propositions 2 and 3 together indicate that a mechanism designer's range of possibilities does not expand if the mechanism designer is allowed to suggest a strategic form correlated equilibrium to agents rather than a Bayesian equilibrium.

Bibliography

[1] Aumann, Robert, (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics* 1, $67 - 96$.

[2] ——————- (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica* 55, $1 - 18$.

[3] Aumann, Robert and Brandenburger, Adam, (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica* 63, $1161 - 1180$.

[4] Bergemann, Dirk and Morris, Stephen (2005): "Robust Mechanism Design," *Econometrica* 73, $1771 - 1813$.

[5] Chen, Yi-Chun and Xiong, Siyang (2011a): "The Genericity of Beliefs-Determine-Preferences Models Revisited," *Journal of Economic Theory* 146, $751 - 761$.

[6] ——————- (2011b): "Genericity and Robustness of Full Surplus Extraction," discussion paper.

[7] Cotter, Kevin D. (1991): "Correlated Equilibrium in Games with Type-Dependent Strategies," *Journal of Economic Theory* 54, $48 - 68$.

[8] ——————- (1994): "Type Correlated Equilibria for Games with Payoff Uncertainty," *Economic Theory* 4, $617 - 627$.

[9] Crémer, Jacques, and McLean, Richard (1985): "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands Are Interdependent," *Econometrica* 53, $345 - 361$.

[10] —————- (1988): "Full Extraction of Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica* 56, $1247 - 1257$.

[11] Dekel, Eddie, Fudenberg, Drew, and Morris, Stephen (2007): "Interim Correlated Rationalizability," *Theoretical Economics* 2, 15-40.

[12] Forges, Françoise (1993): "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information," *Theory and Decision* 35, $277 - 310$.

[13] —————- (2006): "Correlated Equilibrium in Games with Incomplete Information Revisited," *Theory and Decision* 61, $329 - 344$.

[14] Gizatulina, Alia and Hellwig, Martin (2011): "Beliefs, Payoffs, Information: On the Robustness of the BDP Property in Models with Endogenous Beliefs," discussion paper.

[15] Harsanyi, John C. ($1967 - 8$): "Games with Incomplete Information Played by Bayesian Players," *Management Science* 14, $159 - 182, 320 - 334, 486 - 502$.

[16] Heifetz, Aviad and Neeman, Zvika (2006): "On the Generic (Im)possibility of Full Surplus Extraction in Mechanism Design," *Econometrica* 74, $213 - 233$.

[17] Liu, Qingmin (2009): "On Redundant Types and Bayesian Formulation of Incomplete Information," *Journal of Economic Theory* 144, $2115 - 2145$.

[18] ——————- (2011): "Correlation and Common Priors in Games with Incomplete Information," discussion paper.

[19] Mertens, Jean François and Zamir, Shmuel (1985): "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory* 14, $1-29$.

[20] Milgrom, Paul, and Weber, Robert J. (1985): "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research* 10, 619-632.

[21] Myerson, Roger B. (1981): "Optimal Auction Design," *Mathematics of Operational Research* 6, $58-73$.

[22] Neeman, Zvika (2004): "The Relevance of Private Information in Mechanism Design," *Journal of Economic Theory* 117, $55-77$.

Chapter 3

A Modified Cognitive Hierarchy Model

3.1 Introduction

Standard equilibrium concepts in game theory assume that players act rationally given their beliefs of other players' actions, and players' beliefs about other players' actions are consistent with what other players actually do. That is, equilibrium concepts are defined by rationality and mutual consistency. However, in many real world games, standard equilibrium predictions are often far from actual choices players make, and sometimes players seem to have mistaken beliefs about other players' choices. For example, in a beauty contest game, players are asked to choose a number between 0 and 100, and the player whose number is closest of 2/3 of the average wins a prize. The assumptions of players' rationality and consistency of players' beliefs imply that the unique Nash equilibrium is all players choosing zero. However, when the beauty contest is played in experimental settings, the average choice is typically between 20 and 35, as first pointed out by Nagel (1995).

In response to the implausibility of standard equilibrium predictions of the beauty contest game, Camerer, Ho, and Chung (2004) suggest a cognitive hierarchy model (a CH model). In the CH model, there is a cognitive type space $\{0, 1, 2, \ldots, K\}$, and each player draws a type. A 0 type player does not assume anything about other players and chooses according to a fixed probability distribution on [0,100]. A k type player assumes that their opponents are distributed from step 0 to step $k-1$, so he ignores the possibility that some players may be doing as many or more steps of reasoning compared to himself. A k type player can correctly anticipate what lower type players will choose at each period. All players with types greater than zero maximize utilities based on their beliefs of other players' choices. Then the model implies that one-shot beauty contest game will result in players choosing numbers significantly greater than 0, thus explaining real world players' behavior.

In the CH model, players' reasoning abilities are restricted to have k limits on thinking steps, and also players are overconfident and not aware of other players who use as many thinking steps as them. Therefore, Camerer, Ho and Chong (2004) relax two assumptions in standard equilibrium theory to define the cognitive hierarchy model. They assume that players' iterative reasoning abilities are limited, and players have incorrect beliefs about the cognitive type space. I suggest 2 reasons why relaxing two above assumptions simultaneously could be implausible, and suggest a modified cognitive hierarchy model in which players' assessments of the cognitive type space are correct.

First, the original CH model performs poorly in explaining the experimental results of repeated beauty contest game. As Camerer, Ho and Chong (2004) and Sbriglia (2009) observe, real world players' choices in repeated beauty contest game approach quickly to the unique Nash equilibrium. Then the CH model implies that all players' cognitive limits are rapidly increased when a beauty contest game is repeated. That explanation seems to be implausible given Camerer, Ho, and Chung (2004)'s interpretation on players' cognitive limits, because it not likely that players' limits on steps of thinking could increase dramatically in short intervals of the repeated game.

Second, it is not clear why 2 assumptions in the standard equilibrium theory should be simultaneously relaxed. The CH model first assumes a cognitive type space with players' limited thinking abilities, and then assumes heterogenous prior on the cognitive type space. Note that a player's limit on iterative thinking ability is different from his inability to correctly specify the true prior on the cognitive type space. For the latter assumption, Camerer, Ho, and Chung (2004) explain that people are frequently overconfident, and a brain has limits and does not always understand its own limits. However, people are often aware of other people who are smarter than themselves, and even if a brain might not understand its own limits, it might know the existence of things which are incomprehensible. It might be better to relax only the assumption of players' unlimited reasoning abilities so that we can distill only the consequences of the introduction of the cognitive type space.

In response, I suggest a modified cognitive hierarchy model, in which I assume the existence of cognitive type space and players' limits on reasoning abilities as in the original CH model. However, I maintain the common prior assumption on the cognitive type space, and players know the existence of equal or higher type players. Note that there might be discrepancy between the realized distribution of cognitive types from the prior, especially if the number of players is not big. When a k type player chooses an action, he estimates of all other players' actions, including equal or higher type players' choices. Whereas a k type player always correctly predicts what other $1, 2, \ldots, k-1$ type players will choose, he can only guess equal or higher type players' choices due to his limit on the thinking ability, using past data to anticipate equal or higher type players' choices. Players also use past data to improve their estimations on the realized distribution of cognitive types. Therefore, players use deductive reasoning to predict lower types' actions, and learning, or inductive reasoning, to predict equal or higher players' actions.

The modified CH model combines players' utility maximization and learning, so it is meaningful in explaining real-world behavior of players in repeated interactions. Proposition 1 shows that if the modified CH model dynamics converge for a given repeated game, it converges to a Nash equilibrium. In the modified CH model, players only use learning, or inductive reasoning, to assist their deductive reasoning. Thus, the model provides explanations why in real world, players sometimes appear to use only deductive reasoning and sometimes try to learn other players' behavior by their past actions. In that sense, the model suggests a new direction for both behavioral models with bounded rationality, and learning models in game theory. Therefore, it is natural to compare the modified CH model to existing learning models.

There are two broad categories of learning models, the replicator dynamics and the fictitious play. First, the replicator dynamics assume that players are programmed to play fixed pure strategies. As the game is repeatedly played, the players with strategies that result in higher payoffs reproduce faster than other players with strategies that result in lower payoffs. Börgers and Sarin (1997) showed that a learning model converges to a replicator dynamic in continuous time limit, so that they provided an interpretation of the replicator dynamics as a learning model. In contrast with the replicator dynamics, the modified CH model assumes that players are utility maximizers, and use learning only partially to update their estimations of equal or higher type players' choices. Second, the fictitious play models assume that players believe that there is a fixed distribution of all other players' choices, (which is wrong because all other players also follow the fictitious play) so players use only past data to form estimation of other players' choices. The modified CH model assumes that players have better understandings about other players' behavior, and they use past data only to assist deductive reasoning. Also, the modified CH model assumes that players know that other players have different cognitive types, and more information about other players' individual choices are given, players will utilize such data. On the other hand, with fictitious play, individual choice data would not be useful. In addition, in both existing learning models, the initial conditions and first period choices are exogenously given, but in the modified CH model, first period choices are endogenously decided by utility maximization. Detailed theoretical comparisons with other models will be given in section 3.

As a behavioral model, the modified CH model should provide explanations for empirical findings. First, the modified CH model predicts that in repeated beauty contest game, players' choices will converge to the unique Nash equilibrium outcome by the learning process. That is exactly what Camerer, Ho and Chong (2004) and Sbriglia (2009) observed in experiments of the repeated beauty contest game. Therefore, the modified CH model improves the original CH model in explaining the experimental results of the repeated beauty contest game. Detailed discussions are found in section 4.

The modified CH model also provides explanations for the experimental findings of the repeated market entry game with multiple players. In the market entry game, players are asked to choose to either enter or stay out of a given market, and players will enjoy higher utility by entering if and only if there are less players choosing to enter than the market size. Duffy and Hopkins (2005) conduct experiments on the repeated market entry game, and compare experimental findings with predictions of the replicator dynamics and the fictitious play. They find that players' choices tend to converge to pure strategy Nash equilibria as the market entry game is repeated, which might be explained by either of the existing learning models. However, they couldn't explain the following result by the replicator dynamics or the fictitious play: when players are given the information about other players' individual choices, the speed of convergence to the Nash equilibrium is much faster than the case when players' are only given the information about other players' aggregate choices. Also, their experiments show a large variance for the speed of convergence even with same information treatment. The modified CH model provides better explanations for Duffy and Hopkins' (2005) experimental results compared to other learning models. Detailed comparisons with other learning models in the experimental repeated market entry game as in Duffy and Hopkins (2005) are found in section 5.

Related Literature

There is a large literature on "level-k" models, such as Stahl and Wilson (1995), Nagel (1995), and Costa-Gomes, Crawford and Broseta (2001). However, in the level-k literature, players' cognitive types do not necessarily mean their k level iterative reasoning abilities. Rather, Stahl and Wilson (1995) defined a level k type player as a player who has a wrong model of all other players so that he thinks all other players have 0 to $k-1$ types.

It is Camerer, Ho and Chong (2004)'s cognitive hierarchy model that interpreted the level-k type as the cognitive limit on the iterative reasoning of players, while maintaining the assumption in the level-k literature so that players have wrong models about other players. In this paper, players have the correct model about other players so that there is a common prior on the cognitive type space. Players acknowledge the existence of equal or higher type players, and use deductive reasoning to estimate and learn about those incomprehensible choices.

There is a large literature about learning in game theory, as reviewed by Fudenberg and Levine (1998). There are two broad categories of existing learning theory: the replicator dynamics and the fictitious play. I provide comparisons of the modified CH model and existing learning theories in section 3.

Sbriglia (2009) observed that in experiments of the repeated beauty contest game, players' choices converge to Nash equilibrium. Moreover, he observed that the speed of convergence depend on the level of information that the players receive after each period. This paper discusses the relationship between Sbriglia's (2009) experimental findings and the modified CH model in section 4.

Erev and Rapoport (1998) and Duffy and Hopkins (2005) conducts experiments on the repeated market entry game. Especially, Duffy and Hopkins (2005) try to use learning theories (the replicator dynamics and the fictitious play) to explain their empirical findings. I provides explanation of their experimental findings by the modified CH model, and compare the new model with existing learning models in section 5.

Plan of the Paper

In section 2, I will define the modified cognitive hierarchy model for repeated strategic form games. Section 3 defines convergence and stability of the model and observe some basic consequences. Section 4 provides theoretical comparisons with other models. Section 5 studies the repeated beauty contest game and provides theoretical and empirical consequences of the modified CH model in comparison with the original CH model. Section 6 studies the repeated market entry game and provides theoretical and empirical consequences of the modified CH model in comparison with the replicator dynamics and the fictitious play. Section 5 concludes.

3.2 The Modified Cognitive Hierarchy Model

Throughout this paper, I consider complete information strategic form games $G \equiv \{N, A, u\}$, where $N \equiv \{1, 2, \ldots, N\}$ is the set of players, $A \subset R$ is the finite set of actions shared by all players, and $u : A \times \bar{A} \to R$ is the utility function shared by all players. A player i's utility depends on his own choice $a \in A$, and the average of the choice of all players, $\bar{a} \in \bar{A}$, where $\bar{A}$ is the set of all real numbers between the minimum and the maximum of A.[1] Therefore, this paper only considers symmetric games where all players share same set of actions and same utility function.[2] Game G is repeated for periods $t \equiv \{1, 2, \ldots, \}$. In the model, the amount of information each player gets at the end of each stage is important for players' decision making processes. I consider two possible information structures for repeated game as follows:

- **(Aggregate Information Structure)** : After each period t, players are informed about $a^t \in \bar{A}$, a number representing the average choice of all players at period t.
- **(Full Information Structure)** After each period t, players are informed about all players' individual choices, $a_i^t \in A$, for all $i \in \{1, 2, \ldots, N\}$.

Before the game, players independently draw types from a cognitive type space, $k \in \{1, 2, \ldots, K\}$, which is associated with a common prior $p \in \Delta\{1, 2, \ldots, K\}$. A type k player is assumed to be able to perform k times of iterative optimizations. Note that the common prior assumption implies that players know the existence of equal or higher type players. Note that the realized cognitive type distribution might be different from the common prior p, especially when the number of players, N, is small.[3]

[1] Note that $A \in R$, so the average number $\bar{a} \in \bar{A}$ of all players' choices can be calculated to decide utility given all possible $a \in A$ chosen.

[2] Current setting is a simplification and the modified cognitive hierarchy model can be defined for games with heterogenous sets of actions and utility functions for players, incomplete information games, and(or) extensive form games.

[3] In this paper, the different assumptions on the information structure affects players' update processes of estimations of the realized distribution of the cognitive types. Detailed explanation will be provided later.

At each period t, players maximize per-period payoffs based on their estimations of all other players' choices at that period. There are several characteristics of the modified CH model that I wish to highlight before describing actual estimation processes. The following assumptions are made largely for the simplicity of the model. First, players are assumed to be myopic and do not think in advance, and they don't care about reputation effects. Second, at each period t, players make point estimations for 2 unknowns: the realized cognitive type distribution and equal or higher type players' choices at that period.[4] Therefore, players don't have distributional beliefs about those unknowns at each period. Third, when players use past choice data to estimate the realized cognitive type distribution and equal or higher type players' choices, they use lexicographic order in the estimation process with conservative update rule for the realized cognitive type space. At the start of each period t, players compare the newly updated choice data for other players with theoretical predictions for lower type players' choices and change the old estimation for the realized cognitive distribution only when the old estimation is incongruent with new data. After that, players use the previous choice data of other players to make estimations for equal or higher type players' choices at period t. Therefore, it is assumed that players treat two different unknowns differently. Those points will be made clear with the detailed explanation of the model.

Players' decision rules for repeated game depend on the information they receive after each period. First, I recursively define players' decision rules with the aggregate information structure . At period 1, a type 1 player knows that all other players are of equal or higher types compared to himself. Because he is not able to deduce other players' reasonings and optimize

[4] As will be explained shortly, players know exactly what other players with lower types will choose at period t, so that they don't need to make estimations for lower type players' choices.

against them, a type 1 player makes a fixed estimation (which is known to all players with types higher than 1) for all other players' choices.[5]

Then consider a type k player's estimation about other players' choices in period 1. Note that a k type player's assessment of the realized distribution of types about other players might be wrong because he only knows the prior p on the type space. That possible difference between the prior and the realized distribution is more important if the number of players, N, is small. However, in period 1, a k type player doesn't have any information to infer the realized distribution on the type space, other than the prior. Denote a type k player's estimation of the number of type ℓ players at period t for $\ell < k$ as $(q^t_{k,\ell})_{\ell=1}^{k-1}$.[6] Then $q^1_{k,\ell} \equiv p(\ell)N$ for $\ell < k$.[7] That is, at period 1, a k type player assumes that the realized number of a lower type ℓ is the same as given from the common prior. At period 1, a k type player knows what lower type players think, and follow their reasoning to correctly calculate all lower type players' decisions for each lower type $\ell < k$. Then he use $q^1_{k,\ell}$ for $\ell < k$ to calculate lower type players' aggregate choices. In contrast, he uses a fixed estimation (which is known to all players with types higher than k) for equal or higher type players' choices, and use $N - \sum\limits_{\ell=1}^{k-1} q^1_{k,\ell}$ for the number of equal or higher type players, and calculate equal or higher type players' aggregate choices .

[5] It is natural that type 1 player cannot correctly predict other type 1 players' choices. If it were possible, then a type 1 player would first simulate other type 1 players' choices by an optimization, and then optimize to the result. Therefore, type 1 player would be able to perform optimizations twice, contradicting the assumption that he can perform only 1 optimization per period.

[6] A k type player might also have estimations of the number of equal or higher type players. However, such estimation is not used by a k type player because he cannot deduce any meaningful conclusion about equal or higher type players' choices.

[7] When $p(\ell)N$ is not a natural number, set $q^1_{k,\ell}$ as the closest natural number to $p(\ell)N$. For notational convenience, I denote it as just $q^1_{k,\ell} \equiv p(\ell)N$, and such notational abuse is maintained throughout the paper.

Consider a type k player's decision making process at period $t > 1$. First, at the start of period t, a k type player updates his estimation of $(q^t_{k,\ell})^{k-1}_{\ell=1}$ as follows. If observed data from period 1 to period $t-1$ is incongruent to the previous estimation $(q^{t-1}_{k,\ell})^{k-1}_{\ell=1}$, then choose $(q^t_{k,\ell})^{k-1}_{\ell=1}$ so that the difference of total number of per-type players with the previous estimation is the smallest.[8] When there are multiple possible choices for estimation change according to the above rule, then choose a new estimation that always changes the number of lower type players. For example, assume observed data by a k type player is such that either estimated number of type 1 players should be decreased by 1 or type 2 players should be decreased by 1. Then a type k player would choose new estimation that decreases the number of type 1 players by 1. A k type player combines updated $(q^t_{k,\ell})^{k-1}_{\ell=1}$ with the accurate predictions for lower type players' choices to calculate aggregate estimations of lower type players' choices. In sum, in this model, players are assumed to make point estimations for realized cognitive types of other players, and they change their previous estimations only when there is enough evidence to do so.

Then consider a k type players' estimation of equal or higher type players' choices at period t. Because a type k player cannot deduce what equal or higher type will do, he estimates equal and higher type players' choices by past data. In this paper, I assume that all players use fixed weighted moving average of equal and higher type players' past choices to estimate their current choices. Denote a type k player's estimation of the average number of all equal or higher type players' choices at period t as $e^t_k \in \bar{A}$. Also, denote a type k player's chosen action at period t as $a^t_k \in A$. Then we

[8] Note that with aggregate information structure, a k type player can detect the difference between $(q^{t-1}_{k,\ell})^{k-1}_{\ell=1}$ and realized distribution only in the case when there is a lower type whose realized number of players is strictly less than the current estimation given by $(q^{t-1}_{k,\ell})^{k-1}_{\ell=1}$. That is, if the realized number of lower type players is larger than the current estimation of that type, a k type player cannot detect the abnormality because from the k type player's point of view, there might be some equal or higher type players who happen to choose same actions with that lower type.

have:

Assumption 2. ***(Estimation rule with aggregate information structure)*** *With aggregate information structure, a k type player estimates the aggregate choice of all equal or higher type players as follows. There exists a constant I and constants $\{A_i\}_{i=1}^{I}$ with $A_i \geq 0$ for all $i \in \{1, 2, \ldots, I\}$, $\sum_{i=1}^{I} A_i = 1$, such that*

$$e_k^t \equiv \sum_{i=1}^{I} \left(A_i \left(\frac{N}{1 - \sum\limits_{\ell=1}^{k-1} q_{k,\ell}^t} \right) \left(a^{\max(t-i,1)} - \left(\sum_{\ell=1}^{k-1} \frac{q_{k,\ell}^t}{N} a_\ell^{\max(t-i,1)} \right) \right) \right) \tag{3.1}$$

for all $k \in \{1, 2, \ldots, K\}$, and $t \in \{2, 3, \ldots\}$

All players use the past weighted moving average of derived average choices of all equal or higher type players, with weights given by $(A_i)_{i=1}^{I}$ for past periods $(t-1, t-2, \ldots, t-I)$. The average choices of equal or higher type players are calculated by subtracting the estimated average choices of lower type players from the average choice of all players. In sum, at period t, players conservatively update their estimations on the number of lower type players, $(q_{k,\ell}^t)_{\ell=1}^{k-1}$, and they discard the oldest aggregate data c^{t-I-1} and add newest one c^{t-1} to make estimations about equal or higher type players as given in Assumption 1. The weights, $\{A_i\}_{i=1}^{I}$, are only restricted to be non-negative, so Assumption 1 encompasses various possible estimation rules.

Based on their estimations of all other players' choices at that period, all type players choose actions to maximize their per-period expected utilities. In the modified CH model, players are always assumed to choose pure strategies, and higher type players are assumed to know decision rules of lower type players with the following assumptions. First, each player type uses a number to estimate equal or higher type players' choices at period 1, and that number is known to higher type players. Second, when there are multiple actions that maximize a k type player's expected utility, he always chooses an action with lowest label in the set A. Again, that is also known to higher type players. The above assumptions are simplifications. While it is natural that higher type players can correctly follow lower type players' reasoning, it might not be always true that higher type players always have correct knowledge about lower type players' actual choices. In the real world, higher type players' anticipations about lower type players' choices might also have some error terms, which include random choices at period 1 and possible random choices with indifference.

Next, I define players' decision making processes with full information structure, when players are informed about individual choices of all other players at the end of each period. With more information, players are able to infer more information about the realized distribution on the cognitive type space, and also players are able to detect possible correlation between equal or higher type players' past choices. Players' period 1 decision making processes are given as the same with the aggregate information structure, with $q^1_{k,\ell} \equiv p(\ell)N$ for $\ell < k$ and for all $k \in \{2, 3, \ldots, K\}$.

Then consider what happens in period k. At the start of period t, a k type player first determines $(q^t_{k,\ell} \equiv p(\ell))^{k-1}_{\ell=1}$ as the following. A k type player starts with previous period's type realization estimation $(q^{t-1}_{k,\ell} \equiv p(\ell))^{k-1}_{\ell=1}$, and he changes it only when there is enough evidence to do that. First, a k type player compares all other players' previous play records with his prediction of all lower type players' previous play. Starting from player $n' = 1$,(according to the given player labeling $N \equiv \{1, 2, \ldots, N\}$), if player 1's previous play record is congruent to k type player's calculation of player type 1's play, then player 1 is assigned to be a tentative type 1 player for the period t. If not, move to compare player 1's previous play record with a type 2 player's play, up to type $k-1$. A k type player stops comparing n' players' previous play record with type 1's play when $q_{k,1}^{t-1}$ number of players are assigned to type 1. In this manner, a k type player goes over players $n' = 1, 2, \ldots, N$ (excluding himself) to assign other players to tentative lower types $\ell = 1, 2, \ldots, k-1$. If the number of players assigned to ℓ type after going over all other players is less than $q_{k,\ell}^{t-1}$, then $q_{k,\ell}^{t}$ is decreased accordingly. Therefore, a k type player can adjust $q_{k,\ell}^{t}$ when there is evidence that the number of realized ℓ type player is less than $q_{k,\ell}^{t-1}$. For the other direction in adjusting $q_{k,\ell}^{t}$, assume that there is a large enough $\bar{t}$ such that at period $\bar{t}$, a player type $k \in \{2, 3, \ldots, K\}$ possibly increases $q_{k,\ell}^{t}$ for $\ell = \{1, 2, \ldots, k-1\}$ if there are more players whose previous record is the same as k type's play estimation for $\ell \in \{1, 2, \ldots, k-1\}$.[9] Note that there needs to be only one such $\bar{t}$ that players might increase their estimations of the realized numbers of lower type players. After $\bar{t}$, the estimations for lower type players can only decrease because players identify another player as a lower type only when all previous choice records are congruent to a lower type's theoretical past choices.

After fixing $(q_{k,\ell}^{t})_{\ell=1}^{k-1}$ and assigned tentative lower types to other players, a k type player uses past play data of perceived equal or higher types to estimate their choices at period t. Note that with full information structure, a k type player can detect possible correlation between other players' previous play records. Most importantly, he can detect that some other players always choose the same strategy, which results from the fact that they are of the same higher type. Denote a k type's perceived number of equal or higher type players at period t as $N - \sum_{\ell=1}^{k-1} q_{k,\ell}^{t} \equiv \mathbf{k}$. A k type player subtracts the

[9]This is a simplification. A k type player might have a threshold probability so that at each period, if the probability of the case where a higher type player's previous play record happens to be the same as ℓ type's play becomes lower than the threshold, then the k type player would increase $q_{k,\ell}^{t}$. Then the model becomes too complicated.

number of all players with perceived lower types at period t from N. Then the k type player relabel all perceived equal or higher type players as the same order in the original labeling N, so that the remaining players has new label $\mathbf{k} \equiv \{1, 2, \ldots, \mathbf{k}\}$. For each period $t' < t$, a k type player has data for all perceived equal or higher type player's play, which is recorded in a **k**th order tensor, $\mathbf{A_k^{t'}}$ with size $|A| \times |A| \times \ldots, |A|$, where each element (which is 0 or 1) indicates the actions chosen by each player in $\mathbf{k}$. For example, let $\mathbf{k} = 3$, at period t for a type k. It means that a k type player thinks there are 3 equal or higher type players at period t. Then the k type player has data represented by $\mathbf{A_k^{t'}}$ for all previous periods $t' = 1, 2, \ldots, t-1$, and each of $\mathbf{A_k^{t'}}$ is a 3rd order tensor. If the tensor $\mathbf{A_k^1}$ contains the element $a_{2,1,4} = 1$, it means that in period 1, (supposedly equal or higher type) player 1 (as in the new labeling $\mathbf{k} \equiv \{1, 2, \ldots, \mathbf{k}\}$) chose the second action in A, player 2 chose the first action, and player 3 chose the fourth action. Denote a k type player's probabilistic estimation of equal or higher type players' choices at period t as E_k^t. Note that E_k^t is also a **k**th order tensor.

Assumption 3. ***(Estimation rule with full information structure)*** *With full information structure, a k type player estimates choices of perceived equal or higher type players as follows. There exists a constant I and constants $\{A_i\}_{i=1}^I$ with $A_i \geq 0$ for all $i \in \{1, 2, \ldots, I\}$, $\sum_{i=1}^{I} A_i = 1$, such that*

$$E_k^t \equiv \sum_{i=1}^{I} A_i \mathbf{A_k}^{\max(\mathbf{t-i,1})} \tag{3.2}$$

for all $k \in \{1, 2, \ldots, K\}$, and $t \in \{2, 3, \ldots\}$

Therefore, with full information structure, a k type player uses the weighted moving average of perceived equal or higher type players' past choice data to estimate their choices, taking into account possible correlation between them.

Then we have the following observation:

Observation 9. *The full information structure implies the followings with players' decision making process in the modified CH model compared to the aggregate information structure:*

- *Players' assessments for realized distribution of the cognitive type space are corrected faster, with higher possibility to identify the correct realized distribution if the game is repeated long enough without reaching a steady state.*

- *Players' estimations for equal or higher type players' choices are benefitted from observing equal or higher type players' possible correlation, especially the perfect correlation between same type players.*

First, with full information structure, players can track all other players' all previous play records, so they can detect the correct number of lower type players much more quickly than with the aggregate information structure. Players are even able to infer identities of lower type players. Also, unlike the aggregate information structure, players can correct their wrong guess on the realized cognitive types when the realized number of players for some lower type is greater than what is predicted by p. Second, with full information structure, players can use information about the correlation between equal or higher type players' past choices. Therefore, players utilize additional information with the full information structure in meaningful ways.

In sum, players use both deductive and inductive reasoning to anticipate other players' choices and maximize current period's expected utility. Players use deductive reasoning to correctly anticipate lower type players' choices, and they use inductive reasoning to utilize past data for equal or higher type players' choices. As the base game G is repeated, players update their data to utilize the most recent data and discard the oldest one. Also, players' estimations of the realized cognitive type distribution are corrected as the game is repeated. Note that with aggregate information structure, only with the case when the realized number of a lower type is strictly less than what is predicted by the common prior p, a k type player might be able to detect the discrepancy and fix $q^t_{k,\ell}$. On the other hand, with full information structure, if the game is repeated enough without reaching steady state, players might be able to reach the correct realization of cognitive types by observing enough past data. Also note that players are assumed to behave in a way that ignore reputation effects in the repeated game. Players are in a sense myopic so that they maximize per period utilities at each period.

Then consider the convergence of the dynamic. The modified CH model is different from traditional treatments of fictitious play or replicator dynamics so that it is a discrete time dynamic defined on finite games. I define convergence of the modified CH model dynamic as follows:

Definition 12. *A modified CH model dynamic converges if there exists a natural number T such that a k type player's choice a^t_k stays the same after all periods $t \in \{T, T+1, T+2, \dots,\}$ for all $k \in K$.*

Therefore, a modified CH model might converge even when players have wrong beliefs about the realized cognitive types. Note that similar to other learning dynamics, the modified CH dynamic may not converge. Then I observe the following:

Proposition 6. *When the modified CH model dynamic converges, it converges to a pure strategy Nash equilibrium.*

Proof. The following proof works for both the aggregate information structure and the full information structure. Assume that a modified CH model dynamics converges, so all type players' choices converge. Therefore, for every vectors $\epsilon > 0$, $\exists$ a natural number M such that

$$c_k^t - \epsilon \leq c_k^{t+1} \leq c_k^t + \epsilon \tag{3.3}$$

for all $t > M$, and for all $k \in \{1, 2, \ldots, K\}$. Note that from the definition of e_k^t and E_k^t for both aggregate and full information structure, players' estimation for perceived equal or higher types must be in the close range of their actual choices after period M, by (3.3). Therefore, players' anticipations about all other players' choices converge to their actual choices after long enough periods. That does not depend on whether players have correct estimation of the realized cognitive types or not. Even if a player i misunderstands another lower type player j as a higher type, player i's estimation of player j's choices is congruent to what player j actually chooses. Then we can conclude that when the modified CH model dynamics converges, at the convergence point, all players maximize their expected payoffs and their expectations about all other players are congruent to the actual choices made by them. $\square$

3.3 Comparison with Other Models

In this section, I compare the modified CH model with the original cognitive hierarchy model and existing learning models. The modified cognitive hierarchy model follows the original cognitive hierarchy model so that players are endowed with cognitive types, which represent their iterative reasoning abilities. But there are several important theoretical differences between the original and modified CH models.

First, unlike the original CH model, the cognitive type space has a common prior. As explained in the introduction, there are little theoretical reasons to discard the common prior assumption when we try to explain real world players' behavior with limited cognitive hierarchy. Second, in the modified CH model, there is no 0 type player. In the original CH model, level 0 players are not utility maximizers, but their choices are exogenously given. In the modified cognitive hierarchy model, all players maximize expected utilities given their expectations of other players, and players' expectation forming processes depend on their cognitive types, which represent their abilities to perform iterative optimizations. Thus, in the modified CH model, players are differentiated only by their cognitive abilities, not by the fact that whether they are utility maximizers or not. Third, the modified CH model suggests how players might utilize information about past play in repeated games with the limited cognitive hierarchy assumption. In modified CH model, players use past choice data to estimate other higher type players' choices, and also estimate realized distribution on the cognitive type space. Also, we can see that with more information, (full information structure compared to the aggregate information structure) players' estimations are corrected faster. On the other hand, with the original CH model, the past play data has no meaning and players don't utilize it, and that conclusion is rejected by empirical evidences as shown in section 5.

Then I compare the modified CH model with existing learning models in the economics literature. In the modified CH model, players are assumed to use past data to predict equal or higher type players' choices. Therefore in the repeated game, players use partial learning for other players' choices.[10] Note that there is a large literature about learning in game theory, with two broad categories: replicator dynamics (evolutionary models) and fictitious play.

The idea of replicator dynamics is adapted to the game theory from the theory of evolution in biology. In replicator dynamics, players are endowed

[10]Players only use learning partially in the sense that they utilize past data only for equal or higher type players. Players are assumed to be able to correctly deduce lower type players' choices, so learning is not needed for anticipations of lower type players' actions.

with propensities to play certain actions. Therefore, players' types are defined to be their strategy choices, and players do not maximize utilities: they simply choose actions according to their given propensity. As the game is played repeatedly, player types that enjoy higher utilities grow more rapidly (reinforced) than types with low utilities. Thus, realized distribution of types respond to payoffs given by G and evolve according to them. Börgers and Sarin (1997) showed that a learning model converges to a replicator dynamic in continuous time limit, thereby interpreting the replicator dynamics as a learning model.

The modified cognitive hierarchy model is different from the replicator dynamics in several important ways. First, the modified CH model assumes that players know the game G and maximize utilities, whereas replicator dynamic assumes either players don't know G including the payoff structure, or players are not utility maximizers but they are programmed to play certain actions. Thus, the modified CH model is closer to the standard assumptions in game theory and the model explains what happens when players know the environment and the game they are playing, and maximize expected utilities.

Second, the modified CH model puts meaningful restrictions to the initial condition of players' choices, providing rationale for chosen initial conditions. In the replicator dynamic, players' choices at period 1 is randomly assigned without much reason why certain players are assigned to certain actions at period 1. Note that the evolutionary dynamic and the convergence point (if the dynamic converges) may depend on the initial conditions. Therefore, with replicator dynamics, if there are multiple convergence points depending on different initial condition choices, then it is hard to justify one over the other. On the other hand, in the modified CH model, players' choices at period 1 are explained to be consequences of utility maximization and their cognitive limits. When there are multiple convergence points of a modified CH model depending on different initial conditions, different initial conditions can be explained to have different cognitive type distributions or different period 1 estimations. Thus, with the modified CH model, it is easier to find rationale for preferring one initial condition or one convergence point over the other. Related to that, there are restrictions on the set of possible choices at period 1 (that initial choices must be justified by a utility maximization and a cognitive type distribution), so that the modified CH model would provide a sharper prediction of given repeated game compared to the replicator dynamic. Note that with a fixed initial conditions and learning trajectory, players' overall utilities depend on their initial choices and their adjustment processes in both learning models. The modified CH model provides explanations about players' initial choices, their adjustment processes, and their overall utilities for given game.

There is a different class of learning models: fictitious play. Fudenberg and Levine (1998) summarize literature on fictitious play. In fictitious play, players maximize per period expected utilities, anticipating other players' choices solely by their previous play records. Therefore, in fictitious play, players only use inductive reasoning to form expectations about other players, whereas in the modified CH model, players use both inductive and deductive reasoning. An important underlying assumption of the fictitious play is that players assume that other players' choices are given by a fixed distribution, so that they try to estimate that unknown distribution by accumulating past data. In forming such expectation, players assume that all other players are of the same nature: there is no "cognitive types" that differentiates players. Also, in the fictitious play, players do not recognize that all other players use past data for estimation and maximize utility. In a sense, fictitious play assumes that players have the "wrong model" about other players' choice procedures. On the other hand, in the modified CH model, players have a better understanding of other players' choices. In the modified CH model, players correctly know that other players use both inductive and deductive reasoning for estimation and maximize utility. Players utilize past data because of their cognitive limits. Another difference between the fictitious play and the modified CH model is the same with the second reason in the previous paragraph: the modified CH model provides reasonable restrictions on players' initial choices, thus providing explanations for initial choices.

As a behavioral model, the modified CH model should be justified by empirical evidences. Camerer, Ho and Chong (2004) explain that the original CH model provides better explanation for the one-shot beauty contest game compared to the Nash equilibrium theory. The modified CH model improves the original CH model in that it provides better explanation for repeated beauty contest game. The explanation for repeated beauty contest game and comparison with the original CH model is provided in section 5. Duffy and Hopkins (2005) use existing learning models (the replicator dynamics and the fictitious play) to explain the empirical evidences of the repeated market entry game. Even if they justify some of experimental results with existing learning models, they acknowledge that there are some results that might not be explained. In section 6, I explain that the modified CH model might be able to provide answers for their questions.

3.4 The Repeated Beauty Contest Game

In this section, I describe the repeated beauty contest game, the prediction of the modified CH model, and the comparisons of the modified CH model and other models.

I present the beauty contest game as in Camerer, Ho and Chong (2004) and Sbriglia (2009). N players play the beauty contest game repeatedly as follows. Sbriglia (2009) chooses relatively large N, $N = 38$ or $N = 125$ in the experiments. Thus in this section, I assume that the realized distribution of the cognitive type space is congruent to what is given by p. In each period $t = \{1, 2, \ldots\}$, each player simultaneously chooses a multiple of $1/1000$ between 0 and 100.[11] After players choose their numbers, then the average of the players' numbers are announced,[12] and a player who is closest to 2/3 of the average wins a fixed prize. If there are multiple winners, then the prize is evenly shared. Then the game is moved to $t + 1$ period and repeated. Note that in both Camerer, Ho and Chong (2004) and Sbriglia (2009), players are only informed about the average choice of all players, thus excluding the possibility of the modified CH model analysis with full information structure.

I will introduce distributional assumptions that makes analysis in this section simpler. In the original CH model, players' types are drawn from a Poisson distribution, and a t type player mistakenly believes that all other players' types are distributed according to a normalized Poisson distribution, from type 0 to type $t - 1$.[13] In this section for the repeated beauty contest game, I assume that players' cognitive types are uniformly distributed.[14]

Assumption 4.

$$p(k) = \frac{1}{K} \tag{3.4}$$

for all $k \in \{1, 2, \ldots, K\}$.

Also I fix players' estimations for equal or higher type players' choices at period 1 as follows:

[11] Note that due to the given restriction, A is a finite set.

[12] Note that the information given for repeated beauty contest game provides the aggregate information structure, but not the full information structure.

[13] Camerer, Ho and Chong (2004) explain that they choose Poisson distribution for two reasons. First, Poisson distribution has only one parameter, λ, so it is easy to handle. Second, they think it is plausible that as k rises, fewer and fewer players do the next step of thinking beyond k.

[14] I choose the uniform distribution mainly because it is easy to handle with one parameter K.

Assumption 5. *All players estimate equal or higher type players' choices to be equal to 50 at period 1.*

The original CH model is first introduced to explain the difference between the Nash equilibrium theory prediction and real world choices of players with a one-shot beauty contest game. That is, in experiments, real world players do not choose 0, which is the unique Nash equilibrium outcome, but they choose numbers substantially greater than 0. The CH model explains that real world players are not perfectly rational, but they have limited iterative reasoning abilities. In the CH model, the actual choices players make are explained by proper choice of the parameter, λ. Note that with modified CH model, players will also choose numbers substantially greater than 0 in an one-shot beauty contest game. Similar to the original CH model, the actual choices players make can be explained by proper choice of parameters, K.

The modified CH model improves over the original CH model in explaining the repeated beauty contest game. In the real world experiment, as the beauty contest game is repeated, players' choices rapidly approach the unique Nash equilibrium outcome, zero, as Camerer, Ho and Chong (2004) and Sbriglia (2009) observe. With the assumptions of the original CH model, the only possible explanation is that players' cognitive types are rapidly increasing as the beauty contest game is repeated. But such an explanation is implausible because it is unlikely that a player's iterative reasoning ability increases substantially during the short amount of time with which the beauty contest game is repeated in the experimental setting. With the modified CH model, players use deductive reasoning to estimate lower type players' choices, and use inductive reasoning to estimate equal or higher type players' choices. Note that with Sbriglia's (2009) choice of $N = 38$ or $N = 125$, it might be safe to assume that the realized distribution of the cognitive types is the same with the common prior. Then as a result of learning by the past data on choices of equal or higher type players, the model predicts that the average of players' choices will approach to zero. Denote the average of all players' choices at period t as m^t.

Proposition 7. $\lim_{t\to\infty} m^t = 0$, *so that all players will eventually choose 0 in finite time.*

Proof. First, I use induction to prove that $e^t_{k+1} \leq e^t_k$ and $a^t_{k+1} < a^t_k$ for all $t \in \{1, 2, \ldots\}$, and for all $k \in \{1, 2, \ldots, K-1\}$, and also $a^t_k < e^t_k$ for all $k \in \{1, 2, \ldots, K\}$. Note that $e^1_{k+1} = e^1_k = 50$ and $a^1_{k+1} < a^1_k$ for all $k \in \{1, 2, \ldots, K-1\}$. Assume that $e^{j-1}_{k+1} \leq e^{j-1}_k$ and $a^t_{k+1} < a^t_k$ for all $t \in \{1, 2, \ldots, j-1\}$. Then we have

$$e^j_{k+1} = \sum_{i=1}^{I} A_i(K-k)\left(\sum_{m=k+1}^{K} \frac{1}{K} a_m^{\max(t-i,1)}\right) \tag{3.5}$$

$$\leq \sum_{i=1}^{I} A_i(K-k+1)\left(\sum_{m=k}^{K} \frac{1}{K} a_m^{\max(t-i,1)}\right) = e^j_k, \tag{3.6}$$

by inductive hypothesis. Note that $a^t_1 = \frac{2}{3}e^t_1 < e^t_1$. Suppose that $a^t_m < e^t_m$ for all $m = 1, 2, \ldots, k-1$. Then we have

$$a^t_k = \frac{2}{3}\left(\sum_{m=1}^{k-1}\left(\frac{1}{K}a^t_m\right) + \frac{K-k+1}{K}e^t_k\right) < \frac{2}{3}\left(\sum_{m=1}^{k-1}\left(\frac{1}{K}e^t_m\right) + \frac{K-k+1}{K}e^t_k\right) \leq e^t_k, \tag{3.7}$$

first inequality by inductive hypothesis and second inequality by (3.6). Thus we know that $a^t_k < e^t_k$. Now, denote the average choice of types $1, 2, \ldots, k-1$ for period j as B. Then we have:

$$a^j_{k+1} = \frac{2}{3}\left(\frac{k-1}{K}B + \frac{1}{K}a^j_k + \frac{K-k}{K}e^j_{k+1}\right) < \frac{2}{3}\left(\frac{k-1}{K}B + \frac{K-k+1}{K}e^j_k\right) = a^j_k, \tag{3.8}$$

by (3.6) and (3.7). Thus we have $e_{k+1}^t \leq e_k^t$ and $a_{k+1}^t < a_k^t$ for all $t \in \{1, 2, \ldots\}$, and for all $k \in \{1, 2, \ldots, K-1\}$ and also $a_k^t < e_k^t$ for all $k \in \{1, 2, \ldots, K\}$.

Therefore, we know that the average of all players' choices at period t, m^t, satisfies the following:

$$m^t < \frac{2}{3} e_1^t, \tag{3.9}$$

because for all types k, $a_k^t < a_1^t = \frac{2}{3} e_1^t$. Then for a large $t > nI$ with a natural number n, we have

$$\begin{aligned}
m^t &< \frac{2}{3} e_1^t = \frac{2}{3} \left(\sum_{i=1}^{I} A_i m^{t-i} \right) && (3.10)\\
&< \frac{2}{3} \left(\sum_{i=1}^{I} A_i \frac{2}{3} e_1^{t-i} \right) = \left(\frac{2}{3} \right)^2 \left(\sum_{i=1}^{I} A_i e_1^{t-i} \right) && (3.11)\\
&< \left(\frac{2}{3} \right)^3 \left(\sum_{i=1}^{I} \sum_{i_2=1}^{I} A_i A_{i_2} e_1^{t-i-i_2} \right) && (3.12)\\
&< \left(\frac{2}{3} \right)^n \left(\sum_{i=1}^{I} \sum_{i_2=1}^{I}, \ldots, \sum_{i_n=1}^{I} A_i A_{i_2} \ldots A_{i_n} e_1^{t-i-i_2, \ldots, -i_n} \right) && (3.13)
\end{aligned}$$

As $t \to \infty$, $m^t \to 0$ because $\left(\sum_{i=1}^{I} \sum_{i_2=1}^{I}, \ldots, \sum_{i_n=1}^{I} A_i A_{i_2} \ldots A_{i_n} e_1^{t-i-i_2 \ldots, -i_n} \right)$ is bounded.

□

There are several notable points in the proof of Proposition 2. First, it is important that players completely replace old data to new data every I periods. If players continue to utilize part of old data, then we cannot guarantee that m^t approaches 0.[15] Related to that, note that m^t is decreased less than $\frac{2}{3}$ every I periods. That is natural since players use I block of past data when they try to estimate equal or higher type players' choices, and response to them. In a sense, players' observations of other players' past choices replace one round of deductive reasoning in the iterative reasoning process. Therefore, when the game is repeated, utilization of new data works so that players' choices approach to the unique Nash equilibrium outcome. Also, we can see that by Proposition 2, the unique convergent point of the repeated beauty contest game is globally stable.

Note that for the repeated beauty contest game, players' choices will approach to the Nash equilibrium outcome faster when players use more weights on the most recent data. In the extreme, the convergence will be fastest when $I = 1$, so players estimate equal or higher type players' choices only by their choices of just one period before. That is because of the special feature of the beauty contest game: as the reasoning process is repeated, players' optimal choices are monotone decreasing. However, the same might not be said with different types of games and some other rules of estimation might induce faster convergence compared to the estimation rule that only use the latest data. For example, in the financial market, moving averages and old data are often used by investors who perform technical analysis.

Sbriglia's (2009) experimental results on the repeated beauty contest game could be interpreted as supporting the theory of the modified CH model. Sbriglia conducts repeated beauty contest game experiments with two different treatments. In the first treatment, only the average number of all players' choices is announced at the end of each period. That is congruent to the aggregate information structure of this paper. In the second

[15]For example, assume that players continue to partly use the data $m^1 > 0$ when making estimations of equal or higher type players' choices, and the weight players give to m^1 does not converge to zero. Then $m^t \not\to 0$.

"information" treatment, the winners are asked to provide explanation of their reasoning to choose winning number, and all other players are informed about the winner's explanation. In the second treatment, players are given more information compared to the aggregate information structure, but not the same as in the full information structure in this paper. Then Sbriglia finds that in both treatments, players' choices converge to the Nash equilibrium outcome. Note that the modified CH model predicts convergence in both treatments, thus the modified CH model's prediction is supported by the experimental evidences. Also, Sbriglia (2009) observes that in the second treatment with more information, the convergence is faster. The explanation of the winner's iterative reasoning might help other players' reasoning process so that other players' cognitive types might increase. Without such explanation and help from the winner, players' cognitive types are not likely to increase in short amount of time as I explained earlier about the reason why original CH model is insufficient to explain the convergence to Nash equilibrium in repeated beauty contest model.

Then I observe the theoretical predictions of other learning models for the repeated beauty contest game, and compare them with the predictions of the modified CH model. It is easy to see that both the replicator dynamics and the fictitious play will predict that all players' choices will converge to the unique Nash equilibrium outcome as the beauty contest game is repeated. However, the replicator dynamics or the fictitious play does not provide proper explanation of players' choices at the first period. As pointed out before, in other learning models, players' first period choices are exogenously decided. Recall that the motivation of the original CH model is to explain players' first period choices of the beauty contest game. The modified CH model provides not only the explanation of players' first period choices by cognitive types, but also the convergence to the unique Nash equilibrium when the game is repeated. In a sense, the modified CH model combines the original CH model and other learning models in explaining the repeated beauty contest game.

3.5 The Repeated Market Entry Game

In this section, I describe the repeated market entry game, the theoretical prediction by the modified CH model, and the experimental findings in the literature. I explain that the modified CH model provides good explanation of empirical results presented by Duffy and Hopkins (2005), comparing with explanations by existing learning theories, the replicator dynamics and the fictitious play.

I present the market entry game as in Erev and Rapport (1998) and Duffy and Hopkins (2005). N players simultaneously decide whether to enter a market, (e), or to stay out, (o). Player i's payoff at period t is given by:

$$u_i^t(o) \quad = v \tag{3.14}$$

$$u_i^t(e) \quad = v + r(c - m), \tag{3.15}$$

where v, c, r are positive constants, $1 \leq c < N$ and $0 \leq m \leq N$ is the number of players who choose e at period t. Here, the constant c has the interpretation as the capacity of the market. Therefore, the return to entry exceeds the return to staying out if and only if $m < c$. Note that in all pure strategy Nash equilibria of this game, between c and $c-1$ agents enter the market, and there is a symmetric mixed strategy equilibrium. Note that we can relabel e as 1, and o as 0, then the given game is an example of the symmetric games described in section 2 with $A \equiv \{1, 0\} \in R$, and a player's utility depends on his own choices and the average of all players' choices.

First, consider what will happen with full information structure with the modified CH model. I am interested in whether and when the modified CH model dynamic converges to pure strategy Nash equilibria. In the modified CH dynamic process, the discrepancy between the number of perceived lower type ℓ by higher type k, $q^t_{k,\ell}$, and the realized number of lower types ℓ are important in players' estimation and maximization processes. With the full information structure, if the dynamic is repeated long enough without convergence, then players' perceived lower types is likely to converge to the realized types. That is because players will be able to compare more and more repeated play records to the calculated lower types' play. Note that the only possible scenario that there exists a player whose estimation of the cognitive type space keeps to be wrong when the dynamic does not reach a convergence point is that when there are two different types $k \neq k'$, such that $a^t_k = a^t_{k'}$ for $t = 1, 2, \ldots$, and we can expect such cases would rarely happen.

Excluding such scenarios, with full information structure, one of the followings are realized as the game is repeated: either the dynamic converges before all players' estimations of the realized cognitive types coincide with the actual realization, all players' estimations of the realized cognitive type space are adjusted to be correct without reaching a convergent point. The purpose of the current analysis is to find out necessary conditions so that the modified CH dynamics does not converge. Thus, I find such necessary conditions when all players have come to the correct estimation of the realized cognitive type space. The detailed discussions and explanations are given in the Appendix. From the analysis, we can observe that the convergence of the modified CH dynamics depend much on the realization of the cognitive types of players. The modified CH dynamic might oscillate among multiple states without converging, or the modified CH dynamic might converge fast. Whether the dynamic converges or not, and also the speed of convergence vary greatly depending on the realized cognitive type distribution.

Now, consider how the modified CH dynamic evolves differently with the aggregate information structure. With the aggregate information structure, players' estimations about the realized cognitive types are corrected much slower compared to the full information structure. Note that whether the dynamic converges of not, and also the speed of convergence depend on whether players have the correct estimation about the realized cognitive type space. Therefore, the modified CH model predicts that the speed of convergence would be slower with the aggregate information structure compared to the full information structure, if the dynamic converges.

The theoretical predictions of modified CH model is that the repeated market entry game might converge to pure strategy Nash equilibria where some players permanently choose to enter, and other players stay out. Also, the full information structure induces convergence more than the aggregate information structure, and players utilize more information given by the full information structure. As I will explain, the modified CH model provides explanations for Duffy and Hopkins' (2005) experimental results.

For the repeated market entry game, Duffy and Hopkins (2005) repeat the market entry game 100 times for each session. The experimental question for the repeated market entry game is, whether the game converges or not, which equilibrium if it converges, and which model best explains real-world players' behavior.

Duffy and Hopkins (2005) conduct experiments with three different treatments. In the first "limited information" treatment, subjects are repeatedly asked to choose between e or o, without knowing the payoff function. The subjects even didn't know that they were playing games with other subjects. In the second "aggregate information" treatment, subjects were informed of the game and the payoff function for all players. Also, subjects were informed the hypothetical payoff of choosing e even if they choose o in that period. In the final "full information" treatment, additional to all information for the aggregate treatment, subjects were also informed of the individual actions chosen by each of the other 5 players in the session; this information was not available in other treatments.

Duffy and Hopkins (2005) analyze the repeated market entry game with existing learning models and provides several theoretical predictions regarding to their three different treatments. First, both the replicator dynamics and the fictitious play predict that players' choices converges to pure strategy Nash equilibria. In addition, if players follow replicator dynamics, then all three treatments should show no difference in the speed of convergence because all information that players need to know in the replicator dynamics is provided in the limited information treatment. On the other hand, if players follow fictitious play, then the aggregate information treatment and the full information treatment should show same speed of convergence because all information that players need to know in the fictitious play is provided in the aggregate information treatment.[16]

Duffy and Hopkins (2005) draw several conclusions from experiments. First, they observe that in the long run, players' choices might converge to a pure strategy Nash equilibrium with some agents permanently in the market, and some permanently out. Both of the replicator dynamics and the fictitious play explain the tendency to the convergence to the pure strategy Nash equilibrium. However, Duffy and Hopkins' (2005) experimental results show that the process might not converge after 100 periods of the repeated market entry game, and that might be seen as incongruent with existing learning models.

Second, there are significant differences in the speed of convergence across different treatments. In the limited information treatment, only after close to 100 periods do subjects begin to approach equilibrium. Players' choices

[16]Because players know the game in the aggregate information treatment and they are informed about the hypothetical payoffs if they chose differently, they can deduce the aggregate choice of other players in each period.

converge faster in the aggregate information treatment, and even faster in the full information treatment. Because of the differences of the speed of convergence, the empirical results says that players do not entirely follow the replicator dynamics, though they could partly follow it because the limited information treatment does converge in the end. Also, because the convergence speeds are significantly different between the aggregate information and full information treatments, players are not considered as entirely following the fictitious play. In fictitious play, all players have the same expectation forming process such that they perceive all other players are of the same nature, so the only meaningful information for the players in fictitious play is the aggregate choice of all other players. There is no way to explain that players might utilize other players' individual choice data in the fictitious play.

The modified CH model provides explanations for Duffy and Hopkins' (2005) experimental results. First, the modified CH model predicts that the repeated market entry game might converge to pure strategy Nash equilibrium, but the speed of convergence depends much on the realized cognitive type distribution, the size of I, and shape of $(A_i)_{i\in\{1,2,\ldots,I\}}$. Therefore, the modified CH model predicts large variance of the speed of convergence for the repeated market entry game. That is exactly what Duffy and Hopkins' (2005) experimental results show. Even with aggregate or full information treatment, players' choices might not converge after 100 periods, or even when they converge, the speed of convergence varies greatly. On the other hand, it is difficult to explain different speed of convergence with same information structure by the fictitious play or the replicator dynamics.

Second, the modified CH model explains the empirical evidences that the aggregate information treatment shows slower convergence speed compared to the full information treatment. As explained before, in the modified CH model, players do utilize more information about other players' past individual choices with the full information structure in two ways, compared to the aggregate information structure. With individual choice data, players can identify lower type players faster. Also, players can utilize the correlation between equal or higher type players' choices, especially the perfect correlation between same type players. Thus, in one hand players can increase the proportion of deductive reasoning (which is more accurate than the inductive reasoning) more by correctly specifying the proportion of lower type players, and on the other hand players can use inductive reasoning more efficiently by having more information about equal or higher type players.

In addition, note that the modified CH model provides a possible explanation for players' initial choices. In other learning models, initial choices of players are exogenously given. In contrast, the modified CH model says that players' initial choices are results of their utility maximization based on some expectations. The details of players' adjustment processes depend on the realized cognitive type distributions and players' first period estimations of equal or higher type players.[17]

3.6 Conclusion

I presented a modified cognitive hierarchy model, which adds features of learning to the original cognitive hierarchy model by Camerer, Ho and Chong (2004). Players are endowed with cognitive types, which characterize their cognitive limits. A k type player can only perform iterative optimization k times. The cognitive type space has a common prior, so players knows the existence of equal or higher type players, even the realized distribution of cognitive types might be different from the prior. Players can follow lower type

[17] Duffy and Hopkins (2005) does not provide full data of players' individual choice history. In the future, more suitable experiments of the repeated market entry game might be conducted to compare the theoretical predictions of the modified CH model to real world players' choice dynamics.

players' reasonings and correctly anticipate their choices, but they cannot follow equal or higher type players' reasonings and perform maximization. Therefore, players are assumed to use past choice data of equal or higher type players' choices to estimate their current period choices. It is shown that if the modified CH model dynamics converge, it converges to a Nash equilibrium. Also, it is not likely that the modified CH model dynamics converge to an equilibrium that involves indifference. When the modified CH model dynamics converge to a strict Nash equilibrium, it is likely to be locally stable if the number of players is large or players use wide moving average of past data to estimate equal or higher type players' choices.

The new model is motivated by the fact that whereas the original CH model explains well about the discrepancy between the empirical findings of the one-shot beauty contest model and the Nash equilibrium prediction, the original CH model's explanation for the empirical results of the repeated beauty contest model is not satisfactory. The modified CH model can explain the experimental findings that with repeated beauty contest game, players' choices converge to the unique Nash equilibrium. That finding cannot be explained by the original CH model.

Because the modified CH model explains what happens in the repeated game play when players are involved with learning, it is important that how is the model different from existing learning models: the replicator dynamics and the fictitious play. The modified CH model have different assumptions about players' knowledge about the game and behavior compared to other learning models, so that the new model might be used for explaining real-world phenomena that are not fit to be explained by other learning models. Also, the new model have different interpretations and explanations about initial conditions/choices of players and how players utilize different level of information given to them.

I present some experimental results from the literature, and discuss how to interpret those results with the modified CH model. Sbriglia (2009) reports that in experiments of repeated beauty contest game, players' choices converge to the Nash equilibrium over time, and the speed of convergence is faster when players have additional information about the winner's rationale for choosing the winning number. The modified CH model explains the convergence and provides plausible reasons for the difference in the speed of convergence, thus improving the original CH model. Duffy and Hopkins (2009) conduct experiments on the repeated market entry game, with three different information treatments. The modified CH model provides better explanation for some of their experimental findings compared to existing learning models.

The modified CH model makes an innovation about how to model players' cognitive limits, and how to analyze some real phenomena. Even with cognitive limits, real world players might acknowledge that there are other players who have higher types than themselves. In the new model, learning and inductive reasoning is only used to assist the imperfect deductive reasoning in the presence of the cognitive limits. Real world players might use learning to estimate other players' behavior that are incomprehensible due to cognitive limits. Also, when more information is given about other players' choices, real world players might be able to utilize it to improve their estimations. The modified CH model might provide better explanation of real-world players' behavior than both the original CH model, and existing learning models for some situations.

In the future, the modified CH model might be extended for repeated incomplete information games and extensive form games. Also, it might be interesting to see the consequences of different expectation forming rules to the prediction of the model. In this paper, I used experimental results found in literature for the repeated beauty contest game and the market entry game. In the future, more suitable experiments might be designed to identify in which cases the modified CH model provides good explanations, and in which other cases it does not. Also, for given repeated game, individual players' play records might be tracked to estimate his or her cognitive type in the modified CH model. With such experimental findings, the modified CH model would have much more support.

Bibliography

[1] T. Börgers and R. Sarin (1997): "Learning Through Reinforcement and Replicator Dynamics," *Journal of Economic Theory* 77, 1-14.

[2] C. Camerer, T. Ho, and J. Chong (2004): "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics* 119(3), 861-898.

[3] M. Costa-Gomes, V. Crawford, and B. Broseta (2001): "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica* 39, 1193-1235.

[4] J. Duffy and E. Hopkins (2005): "Learning, Information and Sorting in Market Entry Games: Theory and Evidence," *Games and Economic Behavior*.

[5] I. Erev and A. Rapoport (1998): "Coordination, "magic", and reinforcement learning in a market entry game," *Games and Economic Behavior,* 23, 146-175.

[6] D. Fudenberg and D.K.Levine (1998): "The Theory of Learning in Games," *M.I.T. Press*.

[7] J. Maynard-Smith (1974): "The Theory of Games and the Evolution of Animal Conflicts," *Journal of Theoretical Biology* 47, 209-221.

[8] R. Nagel (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review* 65, 1313-1326.

[9] M. Osborne and A. Rubinstein (1994): "A Course in Game Theory," *MIT Press*

[10] P. Sbriglia (2008): "Revealing the depth of reasoning in p-beauty-contest games," *Experimental Economics* 11, 107-121.

[11] D. Stahl and P. Wilson (1995): "On Players' Model of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior* 10, 213-254.

Appendix A

Analysis of the Repeated Market Entry Game by the Modified CH Model

In the appendix, I provide some explanations for the theoretical predictions of the repeated market entry game.

Consider the choice of parameter values for the repeated market entry game. Duffy and Hopkins (2005) conduct experiments of repeated market entry game with different levels of information about previous play, which will be explained later. In their experiments, they used parameters $N = 6, v = 8, r = 2, c = 2.1$. Thus, in all pure strategy Nash equilibria, the equilibrium is strict and there are 2 players choosing e and 4 players choosing o. In the following analysis of the repeated market entry game, I maintain parametric assumptions of Duffy and Hopkins (2005), so that $N = 6, v = 8, r = 2, c = 2.1$.

First, assume the full information structure. As explained in the main text, I will only consider the cases where all players have correct estimations of the realized cognitive type distribution. Then the convergence depends on the realized distribution of the cognitive types. For example, assume that there are 2 realized cognitive types with 3 players each. Then the modified CH dynamic does not converge because same type players always choose the same action, so that the dynamic can't converge to a state where only 2 players choose e and 4 players choose o with given realized type distribution. Then we have the following observation.

Observation 10. *The modified CH dynamic for given repeated market entry game will converge to pure strategy Nash equilibria only if one of the following conditions are satisfied:*

- *There is a type k such that there are 2 players with k as the realized cognitive type.*
- *There are two types k and k' such that one player is assigned for each of the types.*

Note that the observation suggests only necessary conditions for the convergence. As we will see, the modified CH dynamic might not converge even with one of the listed necessary conditions satisfied.

We want to point out conditions for the case where the modified CH model dynamic does not converge. Due to the length concern, I will only consider when the first condition in Observation 4 is satisfied so that there is a type $\bar{k}$ such that there are 2 players with $\bar{k}$ type, and the $\bar{k}$ type is the highest type among all players. The following analysis shows that the modified CH dynamic is likely to converge with such an assumption, and the dynamic rarely ends up with a cycle. For the other possible realized type distributions when conditions in Observation 4 are satisfied, the analysis is similar to conduct.

Consider the earliest period t such that all player types' estimations of number of lower type players are congruent to the realization of such types.

With the assumption that there is the highest type $\bar{k}$ with exactly 2 players, I will contemplate the dynamics of the modified CH model with every possible states in terms of realized cognitive types and each types' choices at t.

For all possible realization of cognitive types, note that if there are exactly 2 players choosing e and 4 players choosing o at period t, then the dynamic came to a convergence point at period t regardless of the identity of types who choose e. If a lowest type player chose e in period t, then he will choose e in period $t+1$ because when he choose e at period t, the previous play records of all other players supported the choice of e. Then at period $t+1$, the previous play record of period t is added to the belief updating procedure, and the record of period t supports the choice of e because in period t, only one other player chose e. On the other hand, if a lowest type player chose o in period t, then he will choose o in period $t+1$ by the similar reasoning. All players will know what lower type players will choose in period $t+1$, and use previous records for equal or higher type players' choices. In period $t+1$, all type players will choose the same action as in period t due to the similar reasoning. Thus I will exclude the cases where exactly 2 players choosing e at period t in the following observations.

Also note that it is not possible that all players choose o. If all lower type players choose o, then the $\bar{k}$ type players will choose e because for a $\bar{k}$ type player, the maximum number of choice e by all other players is 1. Then all remaining possible states regarding to choices of players at period t with $\bar{k}$ as the highest type are given as follows.

- 3 players with types lower than $\bar{k}$ choose e and $\bar{k}$ type players choose o. (state 1)
- 4 players with types lower than $\bar{k}$ choose e and $\bar{k}$ type players choose o. (state 2)

- 1 player with types lower than $\bar{k}$ chooses e and $\bar{k}$ type players choose e. (state 3)
- 1 player with types lower than $\bar{k}$ chooses e and $\bar{k}$ type players choose o. (state 4)

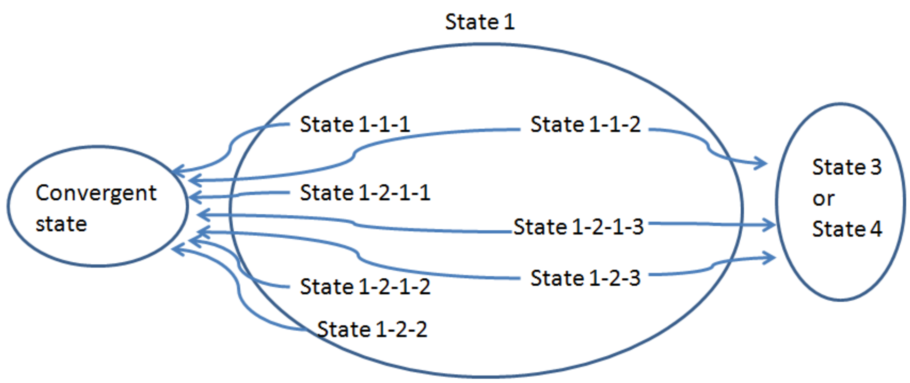


Figure 3.1: Possible dynamics from (state 1).

Start with (state 1). Figure 1 summarizes what happens from (state 1). There are 2 possibilities for 3 players with lower types who chose e: either all 3 players are of the same type (state 1-1), or there is 1 lowest type player and 2 players with a higher type (state 1-2).[1] In (state 1-1), there are two possible type distributions: either there is a player with lower type than the 3 players who choose e (state 1-1-1), or there is a player whose type is between the 3 players who choose e and $\bar{k}$ type players (state 1-1-2). In state (1-1-1), in periods $t+1$, if the 3 players with e update data at period t so that they switch to o, then the type $\bar{k}$ will choose e and the dynamic converges. The 1

[1] If there is a lowest type with 2 players and a higher type with 1 player, then the higher type won't choose e after correctly anticipating that the lowest type players would choose e. Therefore, it is not possible 3 players choosing e with 2 lowest type players and 1 higher type player. Also, the logic is similar if all 3 players are of different types.

player with lowest type will still choose o in period $t+1$ because after period t's observation, his choice of o is reinforced by observing 3 players choosing e in period t. If in periods $t+1$, 3 players with e will still choose e, then all players' choices of period $t+1$ will be the same as period t. As the game is repeated and the data is accumulated with 3 players choosing e, 3 players will switch to o before period $t+I$ because their data will be completely replaced every I periods. Then at that period when those 3 players switch, $\bar{k}$ type players will choose e and the dynamic converges. In state (1-1-2), again as the game is repeated, 3 players will switch to o before period $t+I$. At the period when the 3 players switch to o, the 1 player whose type is just below $\bar{k}$ type might choose e or o. If he chooses o, then $\bar{k}$ type chooses e and the dynamic converges. If the 1 player chooses e when 3 players switch to o, then the state goes to the same as in (state 3) or (state 4), and possible scenarios for (state 3) and (state 4) will be discussed shortly.

Consider the case (state 1-2). Similar to (state 1-1), 3 players who chose e in period t will eventually wish to switch to o before period $t+I$, but different types' incentives might be different. There are 3 possibilities: either all 3 players will simultaneously switch to o at some period $t+j$ (state 1-2-1), or only the 1 player with lowest type will switch to o and 2 players will stay at e at $t+j$ (state 1-2-2), or the 1 player with lowest type stay at e and 2 players switch to o at $t+j$ (state 1-2-3).[2] First consider (state 1-2-1). There are 3 possible type distributions: either the 1 other player is of the lowest type (state 1-2-1-1), the 1 other player's type is between those who chose e (state 1-2-1-2), or the 1 other player is just below type $\bar{k}$ (state 1-2-1-3). In (state 1-2-1-1) and (state 1-2-1-2), when 3 players switch to o at period $t+j$, the 1 other player will still choose o because he uses data of periods up to $t+j-1$ that reinforces his choice of o. Then the $\bar{k}$ type players choose e and the dynamic converges. In (state 1-2-1-3), at period $t+j$, either the 1 other

[2]Note that $j \leq I$ due to players' belief updating process given as in Assumption 3.

player chooses o and the dynamic converges, or the 1 other player chooses e and the state goes to the same as in (state 3) or (state 4). Now consider (state 1-2-2). At period $t + j$ where only 1 player switches from e to o, the 1 other player who chose o will still choose o.[3] $\bar{k}$ type players will also choose o and the dynamic converges. In (state 1-2-3), the dynamic might converge if the 1 other player who chose o at period $t + j - 1$ is of the type just below $\bar{k}$ and chooses e at period $t + j$. Then players with $\bar{k}$ type chooses o and the dynamic converges. Otherwise, the state goes to the same as in (state 3) or (state 4).

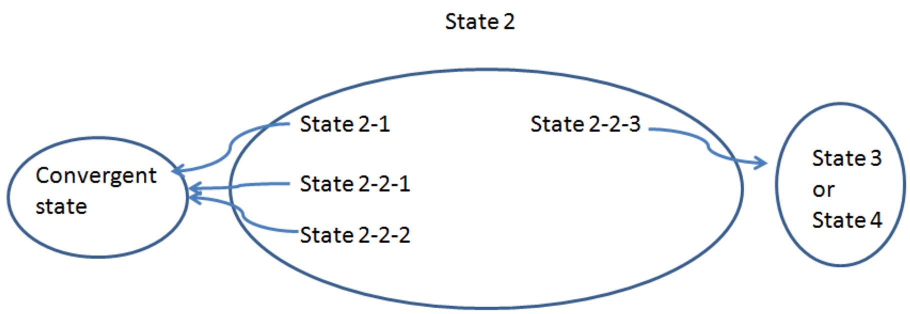


Figure 3.2: Possible dynamics from (state 2).

Then move to (state 2). Figure 2 summarizes what happens from (state 2). There are 2 possibilities: either all 4 players are of the same type (state 2-1), or there is 1 player with lowest type who chooses e and 3 higher type players who choose e (state 2-2).[4] In (state 2-1), the lowest type players will switch to o before period $t+I$, and when they switch, the dynamic converges

[3] If the 1 other player who chose o in period $t + j - 1$ is of the lowest type or the type in between the players who chose e in period $t+j-1$, then he will still choose o at period $t + j$ because the data up to period $t + j - 1$ reinforces the choice of o. If the 1 other player is of type just below $\bar{k}$, then he will choose o at period $t + j$ because he correctly anticipates 2 lower type players choose e at that period.

[4] Again, if the cognitive type distribution is such that if there is 2 or 3 players with lowest type who choose e, then the remaining higher type players will not choose e.

with $\bar{k}$ type players choosing e. In (state 2-2), at some later period $t+j$, either all players switch to o so that the dynamic converges (state 2-2-1), or only the lowest type player switches to o and the dynamic goes to state as same in (state 1-1-1) and eventually converges (state 2-2-2), or only 3 players switches to o and the dynamic goes to states as same in (state 3) or (state 4) - (state 2-2-3).

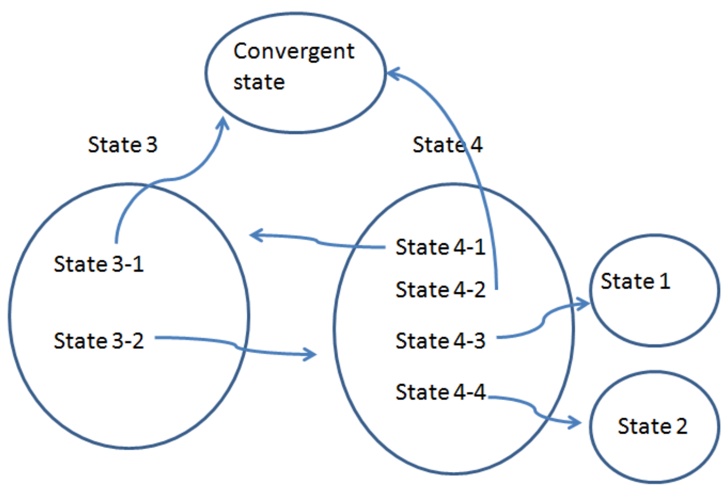


Figure 3.3: Possible dynamics from (state 3) and (state 4).

Figure 3 summarizes what happens from (state 3) and (state 4). Consider (state 3). Note that as long as 3 players who chose e keep choosing e, the 3 players who chose o will continue to choose o in later periods. Then there are 2 possibilities: at some period $t+j$, either only the lower type player switches to o and $\bar{k}$ type players choose e so that the dynamic converges - (state 3-1), or only $\bar{k}$ type players switches to o so that the dynamic goes to the state as same in (state 4) - (state 3-2).

Consider (state 4). As the game is repeated, at some period $t+j$, either the $\bar{k}$ type players switch to e (state 4-1) so that the dynamic goes to the state as same in (state 3), or some number of lower type players switch to e. There are 3 possibilities for possible future switch of lower type players who chose o in period t: at some later period $t+j$, either 1 more lower type player switch to e so the dynamic converges - (state 4-2), or 2 more lower type players switch to e so the state goes as same with (state 1) - (state 4-3), or 3 more lower type players switch to e so the state goes as same with (state 2) - (state 4-4).

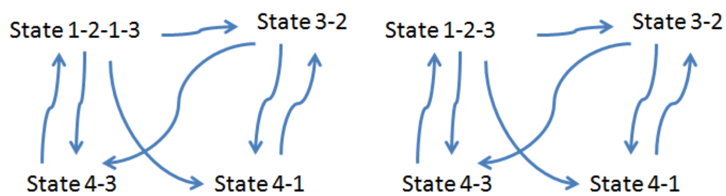


Figure 3.4: All possible cycles.

From the observations of Figure 1, Figure 2 and Figure 3, we can detect possible cycles. I will explain that some cycle scenarios are impossible so that they must be excluded from consideration. First, note that the cycle between (state 3-2) and (state 4-1) is impossible. The change from (state 3-2) to (state 4-1) and the change from (state 4-1) to (state 3-2) are exactly the opposite: in the former case, there is 1 lower type who keeps playing e and $\bar{k}$ type players switch from e to o, and in the latter case, there is 1 lower type who keeps playing o and $\bar{k}$ type players switch from o to e. If the former case happens, then it means $\bar{k}$ type players have accumulated more past data to switch from e to o compared to the 1 lower type player, so if the game is repeated, the latter case cannot happen. Second, the cycle initiating from (state 1-1-2) is also impossible. The only way that (state 1-1-2) ends up with cycling is to reach (state 4-3) so that the dynamic goes back to (state 1), but the distributional assumptions of (state 1-1-2) and (state 4-3) are different.[5] Third, the cycle initiating from (state 2-2-3) is also impossible. The only way that (state 2-2-3) ends up cycling is the dynamic oscillates between (state 2-2-3) and (state 4-4), with possible deviations to (state 3-2). However, the switch from (state 2-2-3) to (state 4) and the switch from (state 4-4) to (state 2-2-3) is exactly the opposite. Note that the distribution must be there is 1 player with lowest type and 3 player with higher type. When (state 2-2-3) moves to (state 4), 1 lowest type player keeps choosing e, and 3 higher type players switch from e to o, and when (state 4-4) moves to (state 2), 1 lowest type player keeps choosing e and 3 higher type players switch from o to e. Players' past choice data cannot support oscillating between those 2 states, even with occasional deviation to (state 3-2) because $\bar{k}$ type players' choice data are utilized in the same way by 1 lowest type player and 3 higher type players.

Therefore, there are only 2 possible cognitive type distributions and scenarios that the modified CH model ends up oscillating among 4 different states with $\bar{k}$ type being the highest as follows:

- The realized cognitive type space is such that among 4 players with lower types than $\bar{k}$, there is 1 player with a lowest type, 2 players with middle type, and 1 player with higher type. The dynamic oscillates among (state 1-2-1-3), (state 3-2), (state 4-1), and (state 4-3).

- The realized cognitive type space is such that among 4 players with lower types than $\bar{k}$, there is 1 player with lowest type, 1 player with middle type, and 2 players with higher type. The dynamic oscillates among (state 1-2-3), (state 3-2), (state 4-1) and (state 4-3).

[5] At (state 1-1-2), all 3 players who choose e are of the same type, and at (state 4-3), the dynamic moves back to (state 1) only if just 2 of 3 players who chose o switches to e.

Figure 4 summarizes all possible cycles that are described above. Note that compared to all possible cognitive type distributions and all possible trajectories depending on initial conditions given $\bar{k}$ highest type with 2 players, the above 2 cases are rare. For other possibilities that satisfy one of the necessary conditions given in Observation 4, such as when the $\bar{k}$ type with 2 players is the lowest type or in the middle, the dynamic analysis is similar. I expect that cycles also rarely happen.

While in the repeated beauty contest model, players with highest cognitive types always perform better than lower type players, in the repeated market entry game, it might not be the case.[6]

[6] For example, consider the dynamic from (state 1-2-2) in the previous explanation. The dynamic ends up with the convergent point where $\bar{k}$ type players ends up choosing o, and 2 lower type players choosing e, enjoying higher per-period payoffs than $\bar{k}$ types. If players are patient enough, the aggregate expected utilities for those lower type players exceed the utilities for $\bar{k}$ type players, who have the highest type.