Essays on Information and Incentives

by

Juan P. Xandri Antuña

Licenciado en Economia, Universidad de Montevideo (2005)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUN 06 2013
LIBRARIES

© 2013 Juan P. Xandri Antuña. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author ....
Department of Economics
May 15, 2013

Certified by ....
Robert Townsend
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Certified by. ....
Ivan Werning
Professor of Economics
Thesis Supervisor

Accepted by ....
Michael Greenstone
3M Professor of Environmental Economics
Chairman, Departmental Committee on Graduate Studies

Essays on Information and Incentives

by

Juan P. Xandri Antuña

Submitted to the Department of Economics on May 15, 2013, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

This thesis studies problems of belief and information formation of agents, and its effect on incentive provision in problems of experimental and mechanism design.

Chapter 1 is based on joint work with Arun Chandrasekhar and Horacio Larreguy. In this chapter we present the results of an experiment we conducted in rural Karnataka, India, to get evidence on how agents learn from each other's actions in the context of a social network. Theory has mostly focused on two leading models of social learning on networks: Bayesian updating and local averaging (DeGroot rules of thumb) which can yield greatly divergent behavior; individuals employing local averaging rules of thumb often double-count information and, in our context, may not exhibit convergent behavior in the long run. We study experiments in which seven individuals are placed into a network, each with full knowledge of its structure. The participants attempt to learn the underlying (binary) state of the world. Individuals receive independent, identically distributed signals about the state in the first period only; thereafter, individuals make guesses about the underlying state of the world and these guesses are transmitted to their neighbors at the beginning of the following round. We consider various environments including incomplete information Bayesian models and provide evidence that individuals are best described by DeGroot models wherein they either take simple majority of opinions in their neighborhood

Chapter 2 is based on joint work with Arun Chandrasekhar, and studies how researchers should design payment schemes when making experiments on repeated games, such as the game studied in Chapter 1. It is common for researchers studying repeated and dynamic games in a lab experiment to pay participants for all rounds or a randomly chosen round. We argue that these payment schemes typically implement different set of subgame perfect equilibria (SPE) outcomes than the target game. Specifically, paying a participant for a randomly chosen round (or for all rounds with even small amounts of curvature) makes the game such that early rounds matter more to the agent, by lowering discounted future payments. In addition, we characterize the mechanics of the problems induced by these payment methods. We are able to measure the extent and shape of the distortions. We also establish that a simple payment scheme, paying participants for the last (randomly occurring) round, implements the game. The result holds for any dynamic game with time separable utility and discounting. A partial converse holds: any payment scheme implementing the SPE should generically be history and time independent and only depend on the contemporaneous decision.

Chapter 3 studies a different but related problem, in which agents now have imperfect information not about some state of nature, but rather about the behavior of other players, and how this affects policy making when the planner does not know what agents expects her to do. Specifically, I study the problem of a government with low credibility, who decides to make a reform to remove ex-post time inconsistent incentives due to lack of commitment. The government has to take a policy action, but has the ability to commit to limiting its discretionary power. If the public believed the reform solved this time inconsistency problem, the policy maker could achieve complete discretion. However, if the public does not believe the reform to be successful some discretion must be sacrificed in order to induce public trust. With repeated interactions, the policy maker can build reputation about her reformed incentives. However, equilibrium reputation dynamics are extremely sensitive to assumptions about the publics beliefs, particularly after unexpected events. To overcome this limitation, I study the *optimal robust policy* that implements public trust for all beliefs that are consistent with common knowledge of rationality. I focus on robustness to all extensive-form rationalizable beliefs and provide a characterization. I show that the robust policy exhibits both partial and permanent reputation building along its path, as well as endogenous transitory reputation losses. In addition, I demonstrate that almost surely the policy maker eventually convinces the public she does not face a time consistency problem and she is able to do this with an exponential arrival rate. This implies that as we consider more patient policy makers, the payoff of robust policies converge to the complete information benchmark. I finally explore how further restrictions on beliefs alter optimal policy and accelerate reputation building.

Thesis Supervisor: Robert Townsend
Title: Elizabeth and James Killian Professor of Economics

Thesis Supervisor: Ivan Werning
Title: Professor of Economics

To my brother Fefo.
I miss you, man

Acknowledgments

I am deeply indebted to Robert Townsend and Ivan Werning for their generous support, advice, and mentoring; not just on academic research but also on my personal life. Through tough times they offered moral support and friendship, and I will be always grateful and honored for being able to have worked with them

I am extremely grateful towards Daron Acemoglu, George-Marios Angeletos, Abhijit Banerjee, Dave Donaldson, Esther Duflo, Glenn Elison and Muhamet Yildiz at MIT, who provided invaluable advice. I would also like to thank Marcelo Caffera, Andrés Neumeyer, Juan Pablo Nicolini, Marzia Raybaudi, and specially Juan Dubra, who mentored me as a young student, encouraged me every step of the way towards getting my doctorate, and have remained great friends ever since.

My experience at MIT, both academically and personally, was quite an amazing one, due to the formidable group of friends and collegues I had the fortune to share these last sixs years with. I'm very grateful to Ben Golub, Felipe Iachan, Anton Kolotilin, José Montiel Olea, Plamen Nenov, Alex Wolitsky and specially to Arun Chandrasekhar, Sebastián Di Tella, Pam Ginocchio, Pablo Kurlat, Horacio Larreguy, Juan Passadore, Mercedes Politi, Xiao Yu Wang and Luis Zermeño for their friendship and support over all these years.

Finally, nothing could have been possible without the undying and limitless support of my long time friends and family. I want to thank Ale, Alexa, Bermu, Fede, Fizu, Fran, Jor, the Juans, Leo, Martin, Maru, Mariana, Mer, Nacho, Noe, Pablo, Santi, Seba, Tonga, Vero and Yosi, for being not just friends, but rather my extended family. over all these years. I want to thank most specially to my parents, Fernanda and Jose Luis and my brother Federico: none of this could have been possible without you. This dissertation belongs to you as much as it does to me.

Cambridge, MA
May 15th, 2013

Contents

Chapter 1

Testing Models of Social Learning on Networks

1.1 Introduction

The social learning process is central to many economic environments. Information and opinions about new products, , political candidates, job opportunities, among others, are transmitted through word-of-mouth or observational learning. This is especially true in developing countries, where a lack of formal institutions and markets as well as information aggregating mechanisms forces agents in developing economies often rely on social connections for information and opportunities. However, the manner in which individuals acquire and process information – their learning mechanism – can greatly influence how society at large learns about the state of the world., Given that social learning is a fundamental component of many economic processes, before employing models of social learning to make policy recommendations we must first understand which models best describe features of empirical social learning. That is, the mechanics of the learning process are of policy interest. If the features of social learning are better described by certain models, those models should be the environment in which the relevant economic outcomes are studied.

There are two broad classes of models that describe learning on social networks. The first is Bayesian learning, wherein individuals process information using Bayes' rule (see, e.g., (Gale and Kariv 2003), (Mossel and Tamuz 2010), (Acemoglu, Dahleh, Lobel, and Ozdaglar 2010), among others.) The second class consists of DeGroot rule of thumb models ((DeGroot 1974)). In these models, agents are myopic and, after seeing the behavior of their network neighbors, individuals take a weighted average of the behaviors to construct their myopic belief going into the subsequent period. DeGroot models are local averaging models and, as such, when individuals freely communicate beliefs formed from repeatedly averaging continuous signals, they commonly converge to the truth ((Golub and Jackson 2010)). However, this convergence is inefficient and, as shown by (Golub and Jackson 2009), networks that exhibit significant homophily – wherein individuals tend to be connected more with those of their own "types" – will converge more slowly to the truth.

Many, if not most, environments of interest, however, have environments in which the actions observed by individuals are discrete. For instance, individuals may observe what sort of technology their neighbhor has adopted or whether or not their neighbor supports a particular candidate. The differences between the Bayesian and DeGroot models in these settings are particularly pronounced. While Bayesian learning will typically generate convergence to the truth in large societies, DeGroot learning may generate misinformation traps wherein pockets of individuals hang on to an incorrect opinion for all but finitely many periods.

Consider the case where the state of the world is either 0 or 1. In this binary environment, individuals employing DeGroot rules of thumb often double-count information and may not reach consensus in the long run even for extremely large graphs (as we show below). Meanwhile, in such an environment, Bayesian learning mechanisms will generically generate consensus in finite graphs and, moreover, in very large graphs the populations' limit opinion will coincide with the true state of the world ((Gale and Kariv 2003), (Mossel and Tamuz 2010)). That is, if the world was truly 0, all individuals would eventually come to believe this.

In this paper, we will study whether Bayesian learning or DeGroot rules of thumb models do a better job of describing empirical learning processes on networks. To study this question we conduct a unique lab experiment in the field across 19 villages in rural Karnataka, India. We ran our experiments directly in the villages so that we could study the relevant population of interest – namely those who could be potentially targeted by policy that depends on social learning (e.g., the introduction of a new fertilizer or credit opportunity, an informational campaign against clientelism). Our approach was to optimally design simple networks that provide statistical power to distinguish between the different learning models in the (Gale and Kariv 2003) environment. We then conducted a lab experiment in the field using these networks to address the proposed question.

We created networks of seven individuals and gave each individual a map of the entire graph so that the full informational structure was comprehended. The underlying state of the world was either 1 or 0 with equal probability. At $t = 0$ each individual received an independent identically distributed (iid) signal about the underlying state of the world and were informed that signals were correct with probability $5/7$. After receiving the signal each individual privately made a guess about the state of the world. These guesses were communicated to each individual's network neighbors at the start of the first period, $t = 1$.

Thereafter, in any given period, each individual knew the guesses of all of her network neighbors from all past periods. Using this information, she made a guess about the state of the world, which in turn was communicated to each of her network neighbors at the beginning of the following period. Every individual was paid for her guess in a randomly chosen round from the set of rounds that she played that day over the course of all the experiments. Consequently, participants had strong incentives to make their best guess in each round.[1]

[1]While it could be the case that players were extremely sophisticated and engaged in experimentation in early

We are able to analyze the data at two levels: the network level and the individual level. Network level analysis considers the entire network and sequence of actions as a single observation. That is, theory predicts a path of actions under a model of social learning, for each individual in each period given a network. At the network level, we address a question about how well social learning behaves; the observational unit in this approach is the social network itself. Meanwhile individual level analysis considers the action of an individual, given a history, as the observational unit.

Our core results are as follows. First, at the network level, we find evidence that a DeGroot rule of thumb model better explains the data than the Bayesian learning model. This is not to say, however, that the social learning process does not resemble the data generated by Baysian learning. In fact, the Bayesian learning model explains 62% of the actions taken by individuals while the best DeGroot rule of thumb explains over 76% of the actions taken by individuals.[2]

Second, at the individual level, we find that a DeGroot rule of thumb model of learning performs significantly better than Bayesian learning in explaining the actions of an individual given a history of play. In fact this model explains nearly 87% of the actions taken by individuals given a history.

Third, to address the limitation that we have designed and restricted our analysis to test whether Bayesian or DeGroot models of social learning better fit the experimental data, we extend the DeGroot model to allow for time variant weights that put a larger emphasis over the actions of the neighbors that are potentially more informative. This new model fits the data as goods as its original DeGroot model, and consequently, points out the need to conduct further social learning experiments with network topologies that allow us to separate between time variant and time invariant DeGroot models.

We also establish several supplementary results which may also be of independent interest.First, we develop a simple algorithm to simulate Bayesianlearning on networks which is computationally tight in the sense that asymptotically there can be no faster algorithm. Namely, the algorithm is $O(T)$ were T is the number of rounds played.[3] Second, we argue that there are problems in estimating models of Bayesian learning on networks with trembles or quantal response equilibrium (QRE). We demonstrate that networks that are small enough to avoid computational constraints are not large enough to tease out the differences between DeGroot and Bayesian learning with trembles. Meanwhile those that are large enough to separate the models become computationally infeasible to study using trembles or QRE. Third, we discuss why such a model selection exercise with network data must be done via a structural approach from the lab. We show that natural examples of reduced form analyses, wherein researches use the intuitions of Bayesian learning and

rounds, anecdotal evidence from participants suggests that this is not the case. In addition, the theoretical and experimental literature uses this assumption (see, e.g., (Choi, Gale, and Kariv 2009)).

[2]When we say a model explains $x\%$ of the actions, we are interested in $x := \frac{y-50}{50}$, where y is the percent of actions predicted correctly. This is the right normalization since we could always explain half the actions by flipping a fair coin.

[3]An algorithm is $O(T)$ if the number of computations as a function of T, $f(T)$, is such that $\frac{f(T)}{T} \to M$ for some constant M. In particular, this is true if $f(T) = MT$, as it is in our algorithm

DeGroot learning to test for correlations in regression analysis of social learning data, may be problematic. Namely, the data generated even by Bayesian learning models do not conform to the intuition motivating the regressions. We maintain that, in turn, the researcher ought to proceed by a structural analysis. Given the computational constraints for structural estimation of learning models in large networks, this suggests that separating models of social learning are best addressed in a lab setting.

There is little empirical evidence comparing Bayesian learning with rules of thumb learning in non-experimental contexts. Without experimental data it is difficult to control priors of agents in the network and the signal quality. Moreover, even in field experiments separating between Bayesian and DeGroot models may be difficult. First, structural approaches are computationally infeasible even with moderately sized networks (10 nodes), as it will become clear below in our discussion of computational complexity. Second, reduced form tests may not suffice for separating between these models. Third, empirical network data may not be precisely measured, affecting the conclusions of a researcher who is trying to select between these models of learning. There may be problems with estimating a structural model on a sampled network and the survey-obtained social network may not be precisely the communication channels used in practice, both of which would induce biases ((Chandrasekhar and Lewis 2010)). Since network-based estimation (which is inherently structural even when using reduced form regressions) is often sensitive to misspecification of the network, it is difficult to cleanly identify which model best describes the data in a non-laboratory context. Fourth, we are unable to know exactly what information is being transmitted in empirical data. Without knowing whether the information transmitted in this context is beliefs, actions or something else all together, one may mistakenly select the wrong model because of not properly specifying the information that is communicated.

Meanwhile, we believe that for our purposes, conducting a lab experiment outside the field of interest is insufficient because we desire to describe the social learning process for our population of interest. We are precisely interested in studying the social behavior of rural populations in a developing country as this is the relevant population in the aforementioned literature.

(Acemoglu, Dahleh, Lobel, and Ozdaglar 2010) and (Jackson 2008a) provide extensive reviews of the social learning on networks literature. The literature is partitioned by whether the learning is Bayesian or myopic (following some rule of thumb). On top of this, the literature layers a myriad of questions such as whether individuals learn from the communication of exact signals (or beliefs or payoffs of other agents) or by observing others' actions, whether the information arrives once or enters over time, whether the interaction is simultaneous or sequential, etc.[4]

[4]The benchmark models of Bayesian learning come from (Banerjee 1992) and (Bikhchandani, Hirshleifer, and Welch 1992). They examine sequential decisions by Bayesian agents who observe past actions. These papers point out that consensuses are formed and thereafter agents end up making the same decision, which may in fact be the wrong decision. (Smith and Sorensen 2000) and (Celen and Kariv 2005) look at Bayesian-rational sequential decision making and explore the conditions under which asymptotic learning is attainable. (Acemoglu, Dahleh, Lobel, and Ozdaglar 2010) extend the framework to consider a social network environment in which individuals

(Gale and Kariv 2003) study the Bayesian learning environment that is closest to ours. They only focus on Bayesian learning and extend the learning model to a finite social network with multiple periods. At time t each agent makes a decision given her information set, which includes the history of actions of each of their neighbors in the network. Via the martingale convergence theorem, they point out that connected networks with Bayesian agents will yield uniform actions in finite time with probability one. (Choi, Gale, and Kariv 2005; Choi, Gale, and Kariv 2009) make a seminal contribution to the empirical literature of social learning by testing the predictions derived by (Gale and Kariv 2003) in a laboratory experiment. They are able to show that features of the Bayesian social learning model fit the data well for networks of three individuals. However, they do not allow for statistical power under the DeGroot alternatives. In extremely simple networks, such as the ones studied in their paper, there are few (if any) differences in the predicted individual learning behavior by the Bayesian and the rule of thumb learning models.[5]

Turning to DeGroot learning, (DeGroot 1974) provides the most influential non-Bayesian framework. Agents observe signals just once and communicate with each other and update their beliefs via a weighted and possibly directed trust matrix. (Golub and Jackson 2010) characterize the asymptotic learning for a sequence of growing networks. They argue that crowds are wise, provided that there are not agents that are too influential.[6]

The work most closely related to ours is (Möbius, Phan, and Szeidl 2011), who study how information decays as it spreads through a network, and (Mueller-Frank and Neri 2012), who conduct a similar experiment to ours. (Möbius, Phan, and Szeidl 2011) test between DeGroot models and a streams model that they develop in which individuals "tag" information by describing where it comes from. Their experiment uses Facebook network data from Harvard undergraduates in conjunction with a field experiment and finds evidence in favor of the streams model. In our experiment, we shut down this ability for individuals to "tag" information to be able to compare the Bayesian model to DeGroot alternatives.

The rest of the paper is organized as follows. Section 1.2 develops the theoretical framework. Section 1.3 contains the experimental setup. Section 1.4 describes the structural estimation proce-

have stochastic neighborhoods. Their main result is that asymptotic learning occurs even with bounded beliefs for stochastic topologies such that there is an infinitely growing subset of agents who are probabilistically "well informed" (i.e. with some probability observe the entire history of actions) with respect to whom the rest of the agents have expanding observations.

[5]The literature on social learning experiments begins with (Anderson and Holt 1997), (Hung and Plott 2001), and (Kubler and Weizsacker 2004). Explicit network structure are considered in a series of papers by (Gale and Kariv 2003) , (Choi, Gale, and Kariv 2005; Choi, Gale, and Kariv 2009), and Çelen, Kariv, and Schotter (2010).

[6](DeMarzo, Vayanos, and Zwiebel 2003) also consider a DeGroot style model and show that as agents fail to account for the repetition of information propagating through the network, persuasion bias may be a serious concern affecting the long run process of social opinion formation. Moreover, they show that even multidimensional beliefs converge to a single line prior to obtaining a consensus, thereby demonstrating how a multidimensional learning process can be characterized by a uni-dimensional convergence. (Chatterjee and Seneta 1977), (Berger 1981), (Friedkin and Johnsen 1997), and (Krackhardt 1987) are among other papers that examine variations on the DeGroot models.

dure and the main results of the estimation. Section 1.5 presents the discussion of the difficulties of reduced form approaches. Section 1.7 concludes.

1.2 Framework

1.2.1 Notation

Let $G = (V, E)$ be a graph with a set V of vertices and E of edges and put $n = |V|$ as the number of vertices. We denote by $A = A(G)$ the adjacency matrix of G and assume that the network is an undirected, unweighted graph, with $A_{ij} = \mathbf{1}\{ij \in E\}$. Individuals in the network are attempting to learn about the underlying state of the world, $\theta \in \Theta = \{0, 1\}$. Time is discrete with an infinite horizon, so $t \in \mathbb{N}$.

At $t = 0$, and only at $t = 0$, agents receive iid signals $s_i|\theta$, with $\mathrm{P}(s_i = \theta|\theta) = p$ and $\mathrm{P}(s_i = 1 - \theta|\theta) = 1 - p$. The signal correctly reflects the state of the world with probability p. In every subsequent period, the agent takes action $a_{i,t} \in \{0, 1\}$ which is her best guess of the underlying state of the world. Figure 1-1 provides a graphical illustration of the timeline.

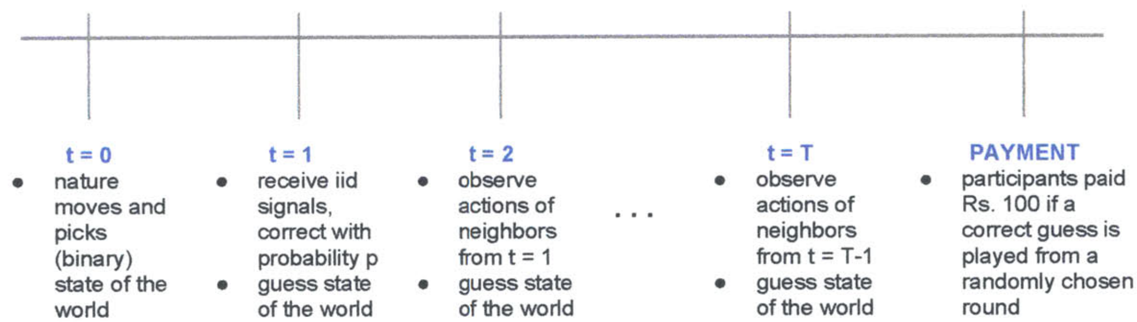


Figure 1-1: Timeline

In addition, we denote by W the set of all possible combinations of signals among agents, which we will refer to as "worlds". Therefore $s \in S$ is an element $s = (s_1, ..., s_n)$ with $s_i \in \{0, 1\}$. Note that $|W| = 2^n$. We will use $d_i = \sum_j A_{ij}$ to refer to the vector of degrees for $i \in \{1, ..., n\}$ and ξ for the eigenvector corresponding to the maximal eigenvalue of A.

1.2.2 Bayesian Learning

In our analysis we consider a model of Bayesian learning with incomplete information. Individuals will have common priors over the relevant state spaces (described below) and update according to Bayes' rule in each period. We formalize the model in Appendix 1.9. Each agent is drawn from a population which has π share Bayesian agents and $1 - \pi$ share DeGroot agents and this fact is common knowledge, as is the structure of the entire network. However, there is incomplete information about the types of the other agents in the network, and the Bayesian individuals will attempt to learn about the types of the other agents in the network along the path while attempting to learn about the underlying state of the world. The incomplete information setup is a useful step beyond the fully Bayesian environment, restricting $\pi = 1$. For instance, if an individual believes that her neighbor does not act in a Bayesian manner, she will process the information about observed decisions accordingly; as outside observers, the econometricians might think that she is not acting as a Bayesian. This is a very common problem when testing Bayesian learning, because we need to make very strong assumptions about common knowledge. A model in which there is incomplete information about how other players behave attempts to address this issue while only minimally adding parameters to be estimated in an already complicated system.

1.2.3 DeGroot Learning

We begin with a classical model of rule of thumb learning on networks and discuss three specific and natural parametrizations. (Jackson 2008b) contains an extensive review of DeGroot learning models. In our experiment, we consider DeGroot action models as opposed to communication models. In action models individuals observe the historical actions of their network neighbors, while in communication models individuals will be able to communicate their beliefs to their neighbors. One might also call these (weighted) *majority* models; individuals choose the action that is supported by a weighted majority of their neighborhood.

We are interested in action models for several reasons. First, the models of Bayesian learning on networks are action models, so it is the appropriate comparison. Second, it is extremely difficult to get reliable, measurable, and believable data of beliefs in a communication model for a lab experiment conducted in the field in rural villages. Third, as it is difficult to control and map into data exactly what is (or is not) communicated by various agents in a more general communication model, we are able to focus on the mechanics of the learning process by restricting communication to observable actions. Fourth, this also fits with the motivating literature wherein individuals may observe the actions, such as technology or microfinance adoption decisions, of their neighbors.

Let $T = T(A)$ be a weighted matrix which parametrizes the weight person i gives to the action of person j. We study three natural parametrizations of the DeGroot model. The first is *uniform weighting* wherein each individual weights each of her neighbors exactly the same. The weight matrix $T^u(A)$ is given by

$$T^u_{ij} = \frac{A_{ij}}{d_i + 1} \text{ and } T^u_{ii} = \frac{1}{d_i + 1}$$

meaning that each individual puts $(d_i + 1)^{-1}$ weight on each of her d_i neighbors as well as on herself.

The second model we consider is *degree weighting.* Each individual weights her neighbors by their relative popularity, given by degree. $T^d(A)$ is given by

$$T^d_{ij} = \frac{d_j}{\sum_{j \in N_i} d_j + d_i} \text{ and } T^d_{ii} = \frac{d_i}{\sum_{j \in N_i} d_j + d_i}.$$

The third model is *eigenvector weighting.* An individual places weight on her neighbor proportional to the neighbor's relative importance, given by eigenvector centrality. $T^e(A)$ is given by

$$T^e_{ij} = \frac{\xi_j}{\sum_{j \in N_i} \xi_j + \xi_i} \text{ and } T^e_{ii} = \frac{\xi_i}{\sum_{j \in N_i} \xi_j + \xi_i}$$

where ξ is the eigenvector corresponding to the maximal eigenvalue of A. This is motivated by the idea that an individual may put greater weight on more information-central neighbors, which eigenvector centrality captures.

While a very natural parametrization of learning, the DeGroot model misses strategic and inferential features of learning. Behavior is as follows. At time $t = 0$, individuals receive signals $s = (s_1, s_2, ..., s_n)$, and accordingly, actions $a_{i,0} = \mathbf{1}\{s_i = 1\}$ are taken. Let $\mathbf{a}_0 = (a_{1,0}, a_{2,0}, ..., a_{n,0})$ At the beginning of $t = 1$, individual i observe $a_t^{(j)}$ for all $j \in N_i$ and aggregates information to form beliefs to $b_1 = T\mathbf{a}_0$. In turn, actions are chosen $\mathbf{a}_1$=$\mathbf{1}\{b_1 > 1/2\}$. Now consider time $t = k$ with action profile $\mathbf{a}_k$. Then beliefs for stage $k+1$ are formed $b_{k+1} = T\mathbf{a_k}$ and accordingly actions are chosen a_{k+1}=$\mathbf{1}\{b_{k+1} > 1/2\}$. In turn, we have if the limit exists,

$$\begin{aligned} a_\infty &= \lim_{k \to \infty} 1\{Ta_{k+1} > 1/2\} \\ &= \lim_{k \to \infty} \mathbf{1}\{T \cdot \mathbf{1}\{Ta_k > 1/2\} > 1/2\}, \ a_k = \mathbf{1}\{Ta_{k-1} > 1/2\}. \end{aligned}$$

While we cannot easily analyze this using the theory of linear operators (due to nested indicator functions), we will discuss its implications in section 1.2.4.

1.2.4 An Illustrative Example: Concentric Social Quilts

We present a simple setup which yields asymptotic learning under communication DeGroot models and consensus under action Bayesian models, but fails asymptotic learning and violate consensus with action DeGroot models. Namely, a number of nodes will become "stuck" in an information trap in a local neighborhood of the network. This demonstrates a wedge between DeGroot and Bayesian learning in models with discrete actions.

We argue that there is an asymmetry in the literature; the right abstraction to think about social learning ought to be parallel across the competing Bayesian and DeGroot models. For motivation we examine a key example of a sequence of networks which satisfy the properties to have asymptotic learning under both Bayesian (action) and DeGroot (communication) learning models, but fail to have asymptotic learning with DeGroot (action) learning models.[7]

The motivation for the graph structure comes from (Jackson, Barraquer, and Tan 2010) who study network architecture that arise as equilibria in favor exchange games. They show that these

[7] The wisdom of the DeGroot learning hinges on the fact that an extensive amount of information is passed along such a model relative to the action model of Bayesian learning. For a parallel, in Bayesian learning if we introduced a communication model, then the filtering problem would be somewhat simpler since an agent would know her neighbors' posteriors exactly.

networks will be social quilts; a social quilt is a patchwork of substructures (e.g., triangles) pasted together in specific ways. We take a very simple example of this style of a network. From the applied perspective, the intuition is that if graphs are constructed as equilibria of risk-sharing or favor-exchange games, then they may have such quilt-like substructures. While these quilt-like structures enable network members to maintain favor-exchange relationships in equilibrium through local punishment of misbehavior, the same networks are also the surface on which information passes among members. We note that, if individuals are indeed DeGroot in a discrete learning process, it may be the case that information does not transmit efficiently through social quilts.

To illustrate this we define a social quilt tree (SQT) as a graph that consists of triangles quilted together around a central triangle such that every triangle (in the interior of the structure) is connected to exactly three triangles in the following way. Consider a sequence of SQTs which can be constructed following a recursive process as the number of nodes goes to infinity. We index this sequence by $r \in \mathbb{N}$.

1. Take SQT_{r-1} and let T_r be the set of terminal nodes of SQT_{r-1}.

2. To each terminal node $i \in T_r$, attach a triangle with two new nodes added.

Figure 1-2 shows such a network and the model is developed in detail in Appendix 1.11.

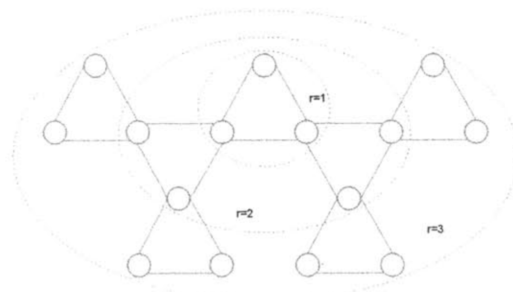
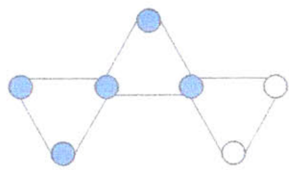


Figure 1-2: A social quilt tree

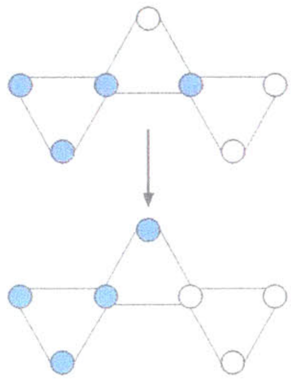
Definition 1. We say that node $i \in V_r$ is *stuck* if there exists a $t_i \in \mathbb{N}$ such that for all $t \geq t_i$, $a_{i,t} = 1 - \theta$.

A node is stuck if the node for all but finitely many periods takes the same (wrong) action. Figure 1-3 provides two examples of nodes that get stuck despite the majority of nodes in the network receiving the right signal.

Panel A of Figure 1-3 illustrates the problem. Assume that for some subtree of the SQT, which connects to the rest of the network through the top-most node, we have the initial signal endowment shown. Any information from the rest of the graph will come via top-most node, which we will call parent node. To get a lower bound on the number of nodes that get stuck in the wrong action, we can simply assume that the parent node of the subtree always chooses the right action for all



Panel A:



Panel B:

Figure 1-3: In Panel A two nodes are stuck for all periods $t \in \mathbb{N}$, even though 5 of the 7 nodes have received the true signal. In Panel B in the first period 4 nodes receive the true signal, and after one node switches, 3 are asymptotically stuck.

rounds. However, even in this case the nodes in the lower right triangle act in the same (wrong) manner for all but finitely many periods. As the sequence of networks grow, $r \to \infty$, there will be a non-vanishing fraction of subtrees with this initial configuration of signals. These subtrees will have at least $3/7$ nodes which become stuck. This example has demonstrated the following result.

Proposition 2. *For a sequence of concentric social quilts with iid signals with probability p, with probability approaching one*

1. under the Bayesian action model there is consensus and asymptotic learning,
2. under the DeGroot communication model with uniform weighting the network is wise,[8]
3. but under the DeGroot action model with uniform weighting a non-vanishing fraction of nodes get stuck.

Proof. All proofs are contained in Appendix 1.12. □

[8] Let $\mu = \theta p + (1-p)(1-\theta)$ and T_n is a sequence of convergent row-normalized matrices. As defined in (Golub and Jackson 2010), the sequence is wise if $plim_{n\to\infty} \sup_{i\leq n} \left|\lim_{t\to\infty} T_n^t s_n - \mu\right| = 0$. In our context wisdom corresponds to asymptotic learning since in the limit a share of nodes that have belief μ goes to one and therefore the nodes can distinguish $\mu > 0$ or $\mu < 0$, as p is known.

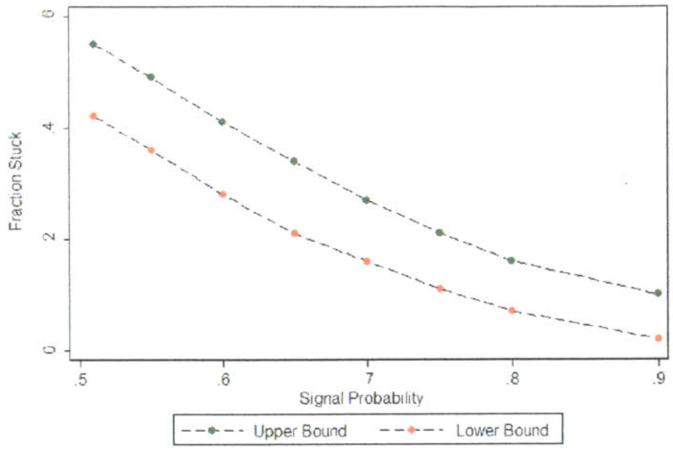


Figure 1-4: Bounds for fraction stuck

çThat asymptotic learning occurs with the Bayesian action model follows from (Mossel, Sly, and Tamuz Forthcoming) and that it occurs with DeGroot communication for this model follows from Corollary 1 of (Golub and Jackson 2010). However, the result for the DeGroot action model is apparent from the previous example. To illustrate the severity of Proposition 2, in Figure 1-4 we show lower bounds on the expected fraction of nodes that are stuck. Even with high quality signals ($p = 0.7$) at least 16% of nodes become stuck and do not asymptotically learn. In particular, recall that the benchmark for mistakes is 50%, since a node can always randomly guess. Therefore, relative to the expected fraction of nodes that should have learned, at least 25% actually get stuck with the wrong information.

In addition to motivating the study of DeGroot action models in our experiment, this example is of independent interest as it raises the question about whether certain network structures are better for social learning, given that asymptotic learning may not occur due to this stuck property. We conjecture that graphs with sufficiently good expansion properties will generate asymptotic learning even with action DeGroot models.

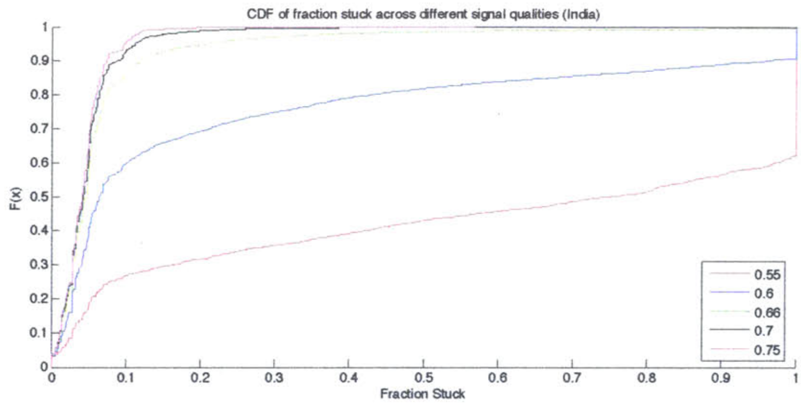


Figure 1-5: Fraction of nodes stuck across 100 simulations per each of 75 Indian village networks with varying probability of the signal being correct.

DISCUSSION.

1.3 Experiment

1.3.1 Setting

Our experiment was conducted at 19 villages in Karnataka, India for a total of 95 experimental sessions for each of three chosen networks. The villages range from 1.5 to 3.5 hours' drive from Bangalore. A village setting was chosen because social learning through networks such as by word-of-mouth communication is of the utmost importance in rural environments; information about new technology ((Conley and Udry 2010)), microfinance ((Banerjee, Chandrasekhar, Duflo, and Jackson 2010)), political candidates ((Cruz 2012)) among other things propagates through the social network.

1.3.2 Overall Game Structure

In each village, individuals played the social learning game three times, each time with a different network structure. The three networks (see Figures 1-6) were played with a random order in each village. Every network consisted of seven individuals and each participant was shown the entire network structure as well as her own location in the network.

At the beginning of each game, every individual was shown two identical bags, one with five yellow balls and two blues ball and the other which had five blue balls and two yellow balls. One of the two bags was chosen at random to represent the state of the world and the goal of the game was that the participant had to independently guess whether the blue bag or the yellow bag had been selected. Since there was an equal probability that either bag could be chosen, we induced priors of 1/2. As the selected bag contained three balls reflecting the state of the world, individuals anticipated receiving independent signals that were correct with probability 5/7.

After receiving their signals in round zero, all individuals simultaneously and independently made their best guesses about the underlying state of the world (which bag had been selected). The game continued to the next round randomly and on average lasted 6 rounds. If the game continued to the second round, at the beginning of the second round each individual was shown the round one guesses of the other individuals in her neighborhood, N_i. Agents updated their beliefs about the state of the world and then again made their best guesses about the state of the world. The game again continued to the following round randomly. This process repeated until the game came to a close. Notice that after the time zero set of signals, no more signals were drawn during the course of the game. Individuals would only observe the historical decisions of their neighbors and update their own beliefs accordingly.

Individuals were paid for a randomly chosen round from a randomly chosen game and therefore faced non-trivial incentives to submit a guess which reflected their belief about the underlying state of the world. Figure 1-1 depicts the timing.

1.3.3 Network Choice

We selected networks specifically so that we could separate between various DeGroot and Bayesian models considered in the paper. The previous experimental literature on Bayesian learning on networks ((Choi, Gale, and Kariv 2005; Choi, Gale, and Kariv 2009)) make use of several interesting and parsimonious three-person networks. However, we are unable to borrow these networks for our study as they were not designed for the purpose of separating between DeGroot and Bayesian learning. In fact, the networks utilized in (Choi, Gale, and Kariv 2005; Choi, Gale, and Kariv 2009) lack power to pit Bayesian learning against the DeGroot alternatives posited above. Panel A of Table 2 shows the fraction of observations that differ across complete information Bayesian learning and the DeGroot alternatives for each of the three networks used in (Choi, Gale, and Kariv 2005) and (Choi, Gale, and Kariv 2009). In two of the networks, there are no differences between the equilibrium paths of Bayesian learning and each of the DeGroot alternatives and in the third network the differences are on the order of 15% of the observations.

Given our goal of separating between Bayesian and DeGroot alternatives, we move to an environment with seven agents as opposed to three agents, so that we obtain more power to distinguish between these models while still maintaining computational tractability.[9]

We considered all connected, undirected networks with seven nodes. Next, we established a model selection criterion function. This criterion function depended on power to detect a DeGroot alternative against a complete information Bayesian null, using our pilot data to generate an estimate of the noise, as well as a divergence function. The divergence function measures the share of node-time observations for which the Bayesian model (with $\pi = 1$) and a DeGroot model pick different

[9]Moving to eight agents, for instance, would be exponentially more difficult for our structural estimation.

actions,

$$D(G) := \frac{1}{nT} \sum_{w \in W} \sum_{t=1}^{T} \sum_{i=1}^{n} \left| a_{i,t}^{B}(w \mid G) - a_{i,t}^{m}(w \mid G) \right| \cdot \mathrm{P}(w \mid \theta = 1),$$

where $a_{i,t}^{B}(w|G)$ is the action predicted under the Bayesian model and $a_{i,t}^{m}(w|G)$ is the action predicted under DeGroot with m-weighting, where m is uniform, degree, or eigenvector weighting. Figure ?? depicts the Pareto frontier between power and divergence and shows one of the networks that we have selected. The procedure yields the networks shown in Figure 1-6.

1.4 Testing the Theory

In order to test how well a model m fits the data in village r, we will use the fraction of discrepancies between the actions taken by individuals in the data and those predicted by the model. This is given by

$$D(m,r) := \frac{1}{n(T_r - 1)} \sum_{i=1}^{n} \sum_{t=2}^{T_r} D_{i,t,r}^{m}$$

where $D_{i,t,r}^{m} = |a_{i,t,r}^{obs} - a_{i,t,r}^{m}|$ which computes the share of actions taken by players that are not predicted by the model m.[10] To examine how poorly model m predicts behavior over the entirety of the data set, we define the *divergence function* as

$$\mathbf{D}(m) := \frac{1}{R} \sum_{r=1}^{R} \frac{1}{(T_r - 1)n} \sum_{i=1}^{n} \sum_{t=2}^{T_r} D_{i,t,r}^{m}.$$

This is simply the average discrepancy taken over all villages. Model selection will be based on the minimization of this divergence measure.

While the divergence is the deviation of the observed data from the theory, we may define the action prescribed by theory in one of two ways. First, we may look at the network level which considers the entire social learning process as the unit of observation and, second, we may study the individual level wherein the unit of observation is an individual's action at an information set.

When studying *network level divergence*, we consider the entire learning process as a single observation. Theory predicts a path of actions under the true model for each individual in each period given a network and a set of initial signals. This equilibrium path that model m predicts is gives the theoretical action, $a_{i,t,v}^{m}$. When using this approach, we try to assess how the social learning process as a whole is explained by a model. This method maintains that the predicted action under m is not path-dependent and is fully determined by the network structure and the set of initial signals. When we consider the *individual level divergence*, the observational unit is the

[10]Since all models and all empirical data have a fixed first action (given by the signal endowment), the first round should not enter into a divergence metric. In turn, we restrict attention to $t \geq 2$.

individual. The action prescribed by theory is conditional on the information set available to i at $t-1$: $a^m_{i,t,v}$ is the action predicted for agent i at time t in treatment r given information set $I_{i,r,t}$.

1.4.1 Learning at the Network Level

We begin by treating the observational unit as the social network itself and take the whole path of actions as a single observation. Table 3 presents the network level divergence for each of the three DeGroot models as well as the complete information Bayesian model ($\pi = 1$). Across all the networks, uniform weighting fails to explain 12% of the data, degree weighting fails to explain 14% of the data, complete information Bayesian learning fails to explain 19% of the data, and eigenvector centrality fails to explain 27% of the data. This suggests that the degree and uniform DeGroot models as well as the Bayesian learning models each explain between 60 to 80% of the observations.

Figure 1-7 presents the data in a graphical manner. Uniform weighting dominates each of the other models, including the complete information Bayesian model, in terms of fit. Meanwhile, eigenvector centrality weighting performs uniformly worse. The data shows that these models outperform the eigenvector weighting model considerably: they explain approximately 84% of all agent-round observations, while the latter explains less than 67% of these observations.

We then compare, pairwise, the complete information Bayesian learning model to each of the DeGroot alternatives. To be able to statistically test the difference of fits across these different models, we conduct a non-nested hypothesis test (e.g. Rivers and Vuong, 2002). Specifically, we perform a nonparametric bootstrap at the village-game level wherein we draw, with replacement, 75 village-game blocks of observations, and compute the network level divergence.[11] This procedure is analogous to clustering and, therefore, is conservative exploiting only variation at the block level. We then create the appropriate test statistic, which is a normalized difference of the divergence functions from the two competing models.

Our key hypothesis of interest is a one-sided test with the null of Bayesian learning against the alternative of the DeGroot model. Table 4 presents the p-value results of the inference procedure. Note that most of the values are extremely close to the boundary – highly significant at the 0.5% level at a one-sided test (let alone a 1% test). First, looking across all topologies, we find evidence to reject the Bayesian model in favor of the uniform weighting alternative and the degree weighting alternative. Second, we find that uniform weighting dominates every alternative across every topology both separately and jointly. Third, we note that eigenvector centrality weighting is dominated; looking across all networks, it is summarily rejected in favor of any alternative model. Ultimately, the bootstrap provides strong evidence that the uniform weighting DeGroot model best describes the data generating process when analyzed at the network level.

Next, we study the incomplete information Bayesian learning model with DeGroot alternatives. We estimate the parameter that minimizes the network level divergence: the best-fitting value of π,

[11] We have 95 village-game blocks in network 1 and 75 for each of networks 2 and 3. We redraw with replacement the same number that we have in our empirical data.

given by $\widehat{\pi} = \text{argmin } \mathbf{D}(m, \pi)$. Figures 1-8, 1-9, 1-10 show the divergence against π and our optimal estimate is given by minimizing the expected diveregence. We find that the optimal $\widehat{\pi} = 0$ is a corner solution where any Bayesian agent would believe that almost no other agents are Bayesian and the population share of Bayesian agents is approximately zero. The divergence under the incomplete information Bayesian learning models with DeGroot alternatives with $\pi \approx 0$ are approximately close to the divergence under.the DeGroot alternative. Thus, we omit these results to avoid redundancy.[12]

1.4.2 Learning at the Individual Level

Having looked at the network level divergence, we turn our attention to individual level divergence. While this does not purely address the mechanics of the social learning process as a whole, it does allow us to look at individual learning patterns. Understanding the mechanics of the individual behavior may help us microfound the social learning process.[13]

1.4.2.1 DeGroot Models

We begin by calculating the individual level divergence for the DeGroot models.[14] Panel A of Table 5 contains the results of the exercise. First, uniform weighting systematically outperforms eigenvector weighting (by a large margin) and degree weighting (by a smaller margin). It is worth noting how well the DeGroot models perform in terms of predicted individual behavior. Across all three networks, the uniform weighting model explains approximately 87% of all individual observations. Degree and eigenvector centrality weighting models predict 73% and 79% of all individual observations, respectively.

1.4.2.2 Bayesian Learning with Incomplete Information

We now turn our attention to the Bayesian learning model. Unlike the myopic models, when considering the empirical divergence and the subsequent predicted action $a^m_{i,t,v}$, we need to consider the whole path of observed actions for all agents.

A potential problem arises: since our model of Bayesian learning implies that actions taken by individuals are deterministic functions of the underlying environment, this implies that the support of the set of potential paths that individuals could have observed is rather limited. Therefore, there is a possibility that empirically, Bayesian agents may arrive to an information set that has zero probability of occurrence. This is problematic for identification, since the Bayesian learning model is mute when agents have to condition their inference on zero probability events; any observed action from then on would be admissible for a Bayesian learning agent.

[12]Details are provided in a supplementary appendix, available upon request from the authors.

[13]It may be the case that agents themselves are do not each behave in according to a particular model while the aggregate social group may best be described by such a model

[14]When an agent faces a tie, they stay with their previous action. We considered a random tie-breaking alternative as well, which does not change the results much.

Table 6 shows that zero probability information sets are hit quite frequently. With any DeGroot alternative, 100% of the treatments in networks 1 and 2 have at least one agent hitting a zero-probability information set. Moreover, at least 62% of players in these networks have hit a zero-probability information set at some point. Though one may be tempted to interpret the lack of support itself as evidence against a Bayesian model, this is a delicate issue requiring a more careful treatment.

To compute the divergence across all observations, we need to make the support of possible paths extend over our entire data set. The usual way to deal with this problem is to introduce disturbances. In the following subsection we explore the possibility of estimating a trembling hand or quantal response equilibrium (QRE) style version of Bayesian learning in which we introduce the possibility of making mistakes by all agents. In such a model, individuals can make mistakes with some probabilities, and Bayesian agents, knowing the distribution of these disturbances, integrate over this possibility when updating.

1.4.2.3 Bayesian Learning with Disturbances and Complexity Problems

To simplify exposition, we restrict attention to the case of a complete information Bayesian model where each agent is Bayesian. In this environment, each agent makes a mistake with probability ε and chooses the opposite action that a Bayesian agent would chose. This guarantees full support; any agent can take any action given any history with positive probability.[15]

Introducing disturbances comes at great computational cost in an environment where agents learn on networks. The only sufficient statistic for the information set that each agent sees is the information set itself, as there is no deterministic function between signal endowments and information sets. This means that through time, the relevant state space (the histories each agents could have seen) grows exponentially. We show that this makes the problem intractable for any practical purpose.

First, we note that the algorithm that we use to simulate the Bayesian learning model without trembles is computationally "tight" in the sense that, asymptotically, there is no faster algorithm.[16] Because any algorithm would have to take order T steps to print output for each of the T periods, an algorithm that is $O(T)$ is asymptotically tight.

Proposition 3. *The algorithm for computing Bayesian learning with no disturbances is $O(T)$. Moreover, it is asymptotically tight; i.e. any algorithm implementing Bayesian learning running time must be at least $O(T)$*

Specifically, the algorithm is $\Theta(n4^nT)$.[17] Notice that if n was growing this algorithm would be

[15](Haile, Hortaçsu, and Kosenok 2008) show that QRE imposes no falsifiable restrictions and can rationalize any distribution of behavior in normal form games. Relating this intuition to our context, one may be able to pick a distribution of ε such that it rationalizes the incomplete information Bayesian model as describing the data well.

[16]Our environment consists of finite graphs where n does not grow in T.

[17]Recall that we say $f_1(n) \in \Theta(f_2(n))$ if f_1 is asymptotically bounded above and below by f_2, upto a constant

exponential time. Second, we show that the extension of this algorithm to an environment with disturbances is computationally intractable.

Proposition 4. *Implementing the Bayesian learning algorithm with disturbances has computational time complexity of* $\Theta(4^{nT})$.

To see how computationally intractable is this algorithm, take as an example our experimental design. Assume that the original code takes one second to run. With $n = 6$ and $T = 5$ the computational time is on the order of 6.9175×10^{18} seconds, which is 8.0064×10^{13} days or 2.1935×10^{11} years. To get some perspective, let us compare the number of calculations with this very simplistic algorithm using the (Choi, Gale, and Kariv 2005; Choi, Gale, and Kariv 2009) enviroment in which $n = 3$. In this setup, the expected time would be 1.2288×10^5 seconds or 1.42 days which is entirely reasonable for structural econometrics.

In the above exercise, we used the most natural algorithm and one that was efficient for the case without disturbances; an objection may be made that there could perhaps be a more efficient algorithm. The decision problem we are interested in is determining whether an agent i in time period t given a history always picks the same action under a proposed algorithm as under the Bayesian model with trembles. We conjecture that the problem is NP-hard, which we are investigating in ongoing work. This means that the computational problem is at least as hard as NP-complete problems.[18] Whether there may or may not be polynomial time solutions for NP-hard problems is open; if P $\neq$ NP, then none would exist. The computer science literature studying Bayesian learning networks shows that obtaining the probabilities is NP-hard (Cooper, 1990) in any given network of events. In this context the networks are networks of events. Translating our framework into this setup involves constructing a network of belief states for each individual in the network and each time period, so a node in the Bayesian learning network would be a pair (i, t), so the size of it would be NT. Our ongoing work seeks to extend their argument to our decision problem which involves checking that the action taking by each person in each time period is identical when comparing a proposed algorithm with the true Bayesian learning model. The intuition is that the learning network is growing linearly in the number of periods and individuals and therefore for any algorithm there can be some action sequence such that to be able to decide whether individual i at time t, given the history, needs to decide whether to guess 0 or 1, one needs all the probabilities. Based on Cooper (1990), which applies to a broader class of networks (and therefore will have weakly worse complexity), we conjecture that the argument for our sub-class of networks will also be NP-hard.

1.4.2.4 Results

We have argued that estimating the Bayesian model with trembles has computational complexity constraints. In turn, we now turn to studying which model best fits the data, taking these constraints

scale. Formally, if $\exists c_1, c_2 > 0, \underline{n}$ such that $\forall n > \underline{n}$, $c_1 \cdot |f_2(n)| < |f_1(n)| < c_2 \cdot |f_2(n)|$.

[18] A problem is said to be NP-complete if (a) it is NP which is to say that a given solution can be verified in polynomial time and (b) it is NP-hard so that any NP problem can be converted to this problem in polynomial time.

into account. We look at the deviation of each agent's action, given the history that the agent has observed at that time, from the predicted action by the model for that agent given the history. The formalities are developed in Appendix B.

Since guaranteeing full support in this model by reintroducing trembles induces computational problems, we make the following arguments regarding the relative performance of the Bayesian model. First, the fact that we repeatedly observe agents facing zero probability events, even when there is positive probability that agents may be behaving in another manner, may be taken as prima facie evidence supporting the idea that this model of Bayesian learning with incomplete information on networks fails to explain the experimental data.

Second, one could make the objection that the considered incomplete information Bayesian model is not sufficiently rich to capture the characteristics of the data and that, perhaps, one needs a more nuanced model. This could indeed be the case, but as demonstrated in Proposition 4, it would be computationally infeasible to estimate a model generating full support.

Third, it might be the case that we have the right model of incomplete information Bayesian model but we lack a theory of what individuals do once they hit zero probability events. This implies that we assume the existence of a correct set of beliefs when encountering zero probability events that rationalizes individuals' actions. If this is the case we may take two different approaches. First, we could be agnostic about the correct off equilibrium beliefs. Second, we could consider the case of a richer Bayesian model that rationalizes the actions taken after an agent hits a zero probability event and precisely matches the supposed off equilibrium behavior. Such a model, of course, has the degree-of-freedom problem.

We begin with the first approach, by being agnostic about the off-equilibrium behavior and instead restrict attention to observations for which agents were in an information set that had a positive probability of occurrence. This is the only feasible comparison we can do given our assumption and agnosticism about the off-equilibrium beliefs. In this subset of observations, we can calculate the individual level divergence, since Bayes' rules applies and the Bayesian learning models gives us a concrete prediction. Of course, we have not eliminated observations at random, but rather we have eliminated those that were not in accordance to the Bayesian learning model equilibrium (i.e. those that happened with zero probability). This is an admittedly ad hoc approach, requiring the assumption that the DeGroot model does not perform sufficiently worse off-equilibrium (where the Bayesian model in principle could rationalize anything), to which we will return below. Under such an assumption, if it turns out that even in this subset of observations, Bayesian performs worse than the alternative myopic models considered, then this would be further evidence that, at the individual level, the Bayesian learning model would not seem to fit the experimental data well.

Based on this idea, we present the calculation of the individual level divergence measure for observations that were in the support of the model. As for the case of the network level divergence, to compute the individual level divergence for the incomplete information Bayesian learning model, we need an estimate of π. Again the optimal $\widehat{\pi} = 0$.[19] Table 7 shows the individual level divergence for the incomplete information Bayesian learning model with DeGroot alternatives. It shows that across all networks, as well as for each network, all DeGroot models have a lower divergence than the Bayesian model.

To be able to perform inference on the null hypothesis of the Bayesian learning model against the alternative DeGroot models, again we perform a nonparametric bootstrap. Our main hypothesis of interest is a one-sided test of the Bayesian learning null against the DeGroot alternatives. Table 8 shows us that, while both the complete and incomplete information Bayesian models are strongly rejected against the alternatives of uniform and eigenvector centrality weighting DeGroot models, we cannot reject them against the alternative of degree weighting. Moreover, as in the case of the network level analysis, we find that uniform weighting beats all alternative DeGroot models.

We now return to the second approach. If we assume that indeed we have the right model of incomplete information Bayesian model but we are simply missing the correct off equilibrium behavior, we could consider the case that a richer Bayesian model could be the one that rationalizes the actions taken after an agent hits a zero probability event and precisely matches the supposed off equilibrium behavior. Notice that even if, for short T, the Bayesian model might be under-performing, with probability 1 this will be the opposite in the long run, inverting the present results. This follows because, if we consider the Bayesian model as rationalizing anything off-equilibrium, once we are off equilibrium, as $t \to \infty$, Bayesian would never be penalized while DeGroot will be penalized infinitely often (assuming the behavior does not identically match the DeGroot model for all but finitely many periods).

To summarize this section's results, first we have presented evidence that the considered model of Bayesian learning result arrives at zero probability information sets extremely often. This can be taken as evidence against these particular models. Second, we provide computational theory that shows that models with trembles, which would smooth out the zero probability information set problem, are of little practical use to structurally evaluate empirical data. In turn, methodologically, structural approaches must restrict themselves to models which allow for zero probability information sets. Third, we take a pass at the data by ignoring the off-equilibrium information sets. This evidence suggests that, when restricting the analysis to positive probability information sets, the divergence minimizing models have in the limit no Bayesian agents. Finally, we point out that this approach, while ad hoc, may be inappropriate for a model of incomplete information behavior wherein the off-equilibrium behavior is well-matched. However, assuming the researcher is interested in incomplete information models (because of the computational infeasibility of QRE models), the argument in favor of the ad hoc approach is rejected only if the researcher believes in an untestable model which performs well off equilibrium (since we know on equilibrium it performs poorly). But such a model is unlikely to be the empirically relevant object.

[19]Details are provided in a supplementary appendix, available upon request from the authors.

1.5 Why a Lab Experiment with Structural Estimation

In this section, we discuss two reduced form approaches to study the data. Our motivation for this exercise is twofold. First, given the computational limits of the structural approach, we are interested in seeing whether reduced form patterns of Bayesian learning (as opposed to DeGroot learning) may be obtained from the data. Second, since larger networks, such as those found in empirical data sets, do not lend themselves to structural approaches for computational reasons, it is worth looking into the effectiveness of reduced form approaches to address these questions.

The central intuition we focus on has to do with double counting information. Under any of the aforementioned Bayesian models, Bayesian agents should not double-count information. However, DeGroot agents do double-count information, and it is on this intuition that we build the exercise.

We provide two examples of regressions which researchers may run. The first set of regressions explores whether individuals overweight the same information if they receive it through multiple channels. The second set of regressions explore whether individuals treat old information that cycles back to them as if it is new, additional information. The null in these regressions is Bayesian model, since one would assume that the relevant parameters ought to be zero. Thus, a rejection of a zero may provide evidence in the direction of the DeGroot rules of thumb. The empirical data shows that both these reduced form analyses seem to provide support in favor of the DeGroot alternatives. However, because we are able to simulate out the data under the null, we show that these intuitions are wrong. Specifically, when we simulate social learning data under the Bayesian null, the coefficients are not as one may have expected.

1.5.1 Multiplicity

We define a variable which is a dummy for whether individual i makes a guess of 1 in the final period T, $y_i := 1\{a_{i,T} = 1\}$. As before, d_i is the degree of individual i and N_i is the set of (direct) neighbors $N_i = \{j \in V : ij \in E\}$. Note that $d_i = |N_i|$. Moreover, N_{2i} is the set of second-neighbors of person i; that is, $j \in N_{2i}$ means that there is at least one path of length two between i and j, but no path of length one. Finally, we define N_{2i}^l to be the set of second neighbors to whom she has exactly l paths.

The first regression we run is of the form

$$y_i = \beta_0 + \beta_1 s_i + \beta_2 \mathbb{E}_{N_i}[s_j | j \in N_i] + \sum_l \beta_{3l} \mathbb{E}_{N_{2i}^l}[s_j | j \in N_{2i}^l] + \epsilon_i. \tag{1.1}$$

This is a regression of whether or not individual i ultimately makes a guess of 1 on whether the individual's signal is 1 (s_i) the share of ones ($\mathbb{E}_{N_i}[s_j|j \in N_i]$) in individual i's neighborhood, and the share of ones given to each subset of second neighbors to whom i has exactly l paths ($\mathbb{E}_{N_{2i}^l}[s_j|j \in N_{2i}^l]$).

The interpretation is as follows. β_2 measures the impact of her neighborhood receiving a greater share of ones on an individual's guess. We expect $\beta_2 > 0$. Moreover, β_{3l} measures the impact of the subset of her second-neighborhood with multiplicity l. The intuition is that as the signals of individuals with greater multiplicity ought not be double-counted under a Bayesian frame, $\beta_{3l+1} > \beta_{3l}$ would be evidence of overweighting redundant information that has arrived via multiple channels, while $\beta_{3l+1} = \beta_{3l}$ would provide evidence in favor of the Bayesian hypothesis.

Table 8 provides the simulation and empirical results. Looking at the empirical results, as expected, an individual's own signal being one and the share of individuals in one's neighborhood with signals of one increase the probability of the final guess being one. However, we can reject that $\beta_{3l+1} > \beta_{3l}$. While this seems to be inconsistent with the intuition that agents engage in double-counting, the simulation exercise shows that these patterns cannot be interpreted in that manner.

Given the learning model, the network structure, and signal endowment, we simulated out the learning path and then ran the relevant regressions. We present results when simulating the learning process from the complete information Bayesian model (every agent is Bayesian) as well as each of the three DeGroot alternatives. The table shows that the Bayesian null does not have coefficients that are near identical across multiplicities 1 and 2. Furthermore, the increasing correlation with indirect friends of higher multiplicities is also not uniformly found across the DeGroot models. Ultimately, the regressions suggest that the linear projection of this learning process is complex and may depend crucially on the network structure and set of initial signals.

1.5.2 Historical Information

Another reduced form that one may look at is whether individuals re-incorporate historical information that they have previously observed. Consider an individual at period 3. They have observed both their own signals and the signals of their direct neighbors (insofar as the first period guesses of their neighbors will be identical to their signals). In period three, therefore, a Bayesian individual's guess should not re-incorporate this information. Instead, it should only update using information about second-neighbors and the like, about whom they have yet to receive information.

To examine this formally, we perform the following regression. We regress the period three guess of individual i on her own signal (s_i) and the average signal of her neigbhorhood ($\mathbb{E}_{N_i}[s_j|j \in N_i]$) which she would have seen in period three. We also include as regressors the average signal of second neighbors ($\mathbb{E}_{N_{2i}}[s_k|k \in N_{2i}]$) which should be new information in period three. Lastly, we include the average signal of direct neighbors whose signals can cycle back via a path of length two back to individual i. Of course, we also include the agent herself in this set. (Formally, we use $\mathbb{E}_{C_i}[s_j|j \in C_i]$, where $C_i = \{j \in V - \{i\} : A_{ij}^2 A_{ij} > 0\} \cup \{i\}$.) The regression is as follows.

$$y_i = \alpha_0 + \alpha_1 s_i + \alpha_2 \mathbb{E}_{N_i}[s_j] + \alpha_3 \mathbb{E}_{N_{2i}}[s_k|k \in N_{2i}] + \alpha_4 \mathbb{E}_{C_i}[s_j|j \in C_i] + \epsilon_i. \tag{1.2}$$

We test the hypothesis of whether $\alpha_4 = 0$, which is our Bayesian null. Notice that $\alpha_4 > 0$ provides evidence that individuals reincorporate information that they already knew as it cycles through the network.

Table 9 presents the simulation and empirical results. Looking at the empirical results, as expected, an individual's own signal being one and the share of direct and new indirect neighbors with signals of one increase the probability of the final guess being one. Also, the empirical results show that the share of repeated indirect neighbors with signals of one increase the probability of the final guess being one, that is$\alpha_4 > 0$. While this seems to provide suggestive evidence for the intuition that DeGroot weighting reincorporates old information, the simulation results provide evidence that for our environment$\alpha_4 > 0$ even for the Bayesian model.

1.5.3 Reflection on Reduced Forms

Taken together, Tables 9 and 10 have shown that natural reduced form approaches to test between these models may be misguided without first checking whether the patterns by the learning processes actually match the intuitions. We are able to study the reduced form projections of the Bayesian model using our simulation algorithm and find evidence that, when projected onto a regression for these networks with this environment, the Bayesian data suggests that the coefficients can deviate greatly from our intuitions. This, we argue, provides a strong motivation for the structural approach to studying the models.

1.6 DeGroot-Like Extensions

We have shown that when we pit the various Bayesian models (complete and incomplete information) against DeGroot alternatives, DeGroot models better fit the data. Moreover, we have found that uniform weighting majority rule does the best out of our three ex ante hypotheses. These findings may be taken as prima facie evidence supporting the idea that social learning processes may be sub-optimal, with information often getting stuck in pockets of the network.

We note an important caveat regarding this statement. We have designed and restricted our analysis to test whether Bayesian or DeGroot models of social learning better fit the data. We have consequently left out of our analysis potentially countless alternative models that might better fit the social learning process. Such a restriction might be important for concluding that social learning processes may be sub-optimal. DeGroot action models present two main shortcomings that contribute to the sub-optimality of the social learning process. First, in DeGroot models individuals do not "tag" information by describing where it comes from ((Möbius, Phan, and Szeidl 2011)). Second, DeGroot models are memory-less in the sense that individuals do a weighted majority rule of t-1 neighbors' actions without taking into account whether they represent a change from t-2 neighbors' actions, and therefore, whether they incorporate potentially new information. In other words, individuals put equal weight in those neighbors that potentially have acquired new information and those that do not.

Since in our experiment communication among participants is shut down, the first caveat is less of a problem. However, omitted regression analysis shows partial correlations between actions and changes in neighbors' actions between t-2 and t-1 after controlling for neighbors ' actions in t-1, and thereby, suggest that the selected model might not best represent the data.

As a first step to address this potential problem, we extend our chosen DeGroot model of learning wherein individuals uniformly weight the actions of each of their network neighbors as follows,

$$a_{i,t} = \left\{ \frac{a_{i,t-1} + \sum_{j\in N_i} a_{j,t-1} + \sum_{j\in N_i} a_{j,t-1} \cdot \mathbf{1}\left\{a_{j,t-1} \neq a_{j,t-2}\right\}}{1 + d_i + \sum_{j\in N_i} \cdot \mathbf{1}\left\{a_{j,t-1} \neq a_{j,t-2}\right\}} > 1/2 \right\}.$$

We denote this model time variant uniform DeGroot model as opposed to the time invariant uniform DeGroot model. In this model extension, individuals put double weight over the actions of the neighbors that change theirs, who probably acquired more information than in the previous round.

Resulst indicate that the network and individual level divergences for the time variant uniform DeGroot model (0.1207 and 0.0663, respectively) are indistinguishable from the corresponding ones for the invariant uniform DeGroot model (0.1203 and 0.0648, respectively). Additionally, they are not statistically different from each other.

While our experiment design has no power to separate between DeGroot models, the fact that the time variant uniform DeGroot model fits the data as good as the and time invariant model suggests an avenue for future research. In this regard, in future work we plan to conduct further social learning experiments with network topologies that allow us to separate between time variant and time invariant DeGroot models.

1.7 Conclusions

In this paper we have investigated whether social learning patterns on small networks resemble those which would emerge if agents behaved in a Bayesian manner or if they are better modeled with DeGroot alternatives which are myopic and more simplistic models. To do so, we developed a simple experiment on networks that were designed to distinguish between these models, large enough to give us power on this dimension, but small enough to ensure that simulating a Bayesian learning on networks model was not computationally intractable. Given our experimental data, we were able to study the social learning process as a whole by taking the network as the unit of observation and study the behavior of an individual, which addresses whether an agent acts in a Bayesian manner.

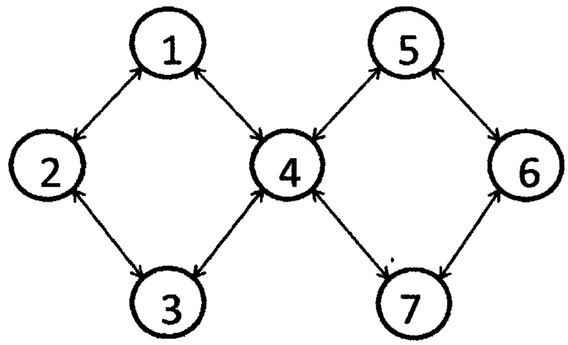
At the network level we find evidence that the uniform weighting DeGroot model best explains the data. The Bayesian learning null is rejected in favor of this alternative model. However, we maintain that Bayesian learning did an adequate job of describing the data process, explaining (beyond pure random guessing) 62% of the actions taken as compared to 76% by the DeGroot alternative.

At the individual level we find that uniform weighting DeGroot performs the best outperforming the Bayesian model. However, we show that the Bayesian model encounters the problem that many individuals come across zero probability information set. First, this provides suggestive evidence of the lack of fit of this incomplete information Bayesian model. Second, we demonstrate that introducing disturbances to smooth out the distribution cannot be a solution in this environment. The computational complexity of the problem is damaging to the very approach of applying QRE or trembles to the Bayesian learning on networks environment. As such, we recommend that researchers focus on computationally tractable models which will be easier to falsify.

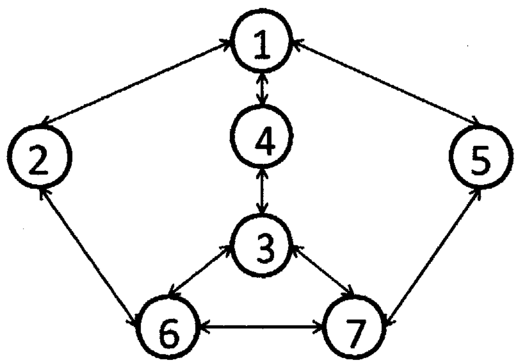
We also show that reduced form approaches may be problematic. We provide two natural examples of regressions which build on intuitions separating DeGroot and Bayesian learning patterns. Equipped with our Bayesian learning algorithm, we simulate learning data from the Bayesian model as well as from DeGroot models and show that the reduced form regression outcomes do not conform to the intuitions.

Ultimately, the findings suggest that agents and the learning process as a whole may better be thought of as coming form DeGroot action models where individuals myopically weight their neighbors' actions when updating their own beliefs rather from a Bayesian model. This may imply that social learning processes empirically may be sub-optimal, with information often getting stuck in pockets of the network. Having constructed an example of a network which satisfies asymptotic learning for DeGroot communication models, but where asymptotic learning fails for DeGroot action models, we argue that in action-learning environments DeGroot processes may be more damaging to the wisdom of society than previously anticipated.

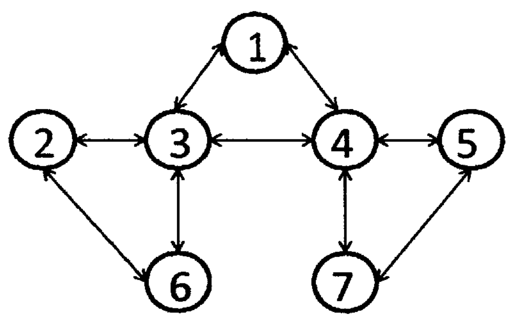
Figures and Tables



Panel A: Network 1



Panel B: Network 2



Panel C: Network 3

Figure 1-6: Network structures chosen for the experiment.

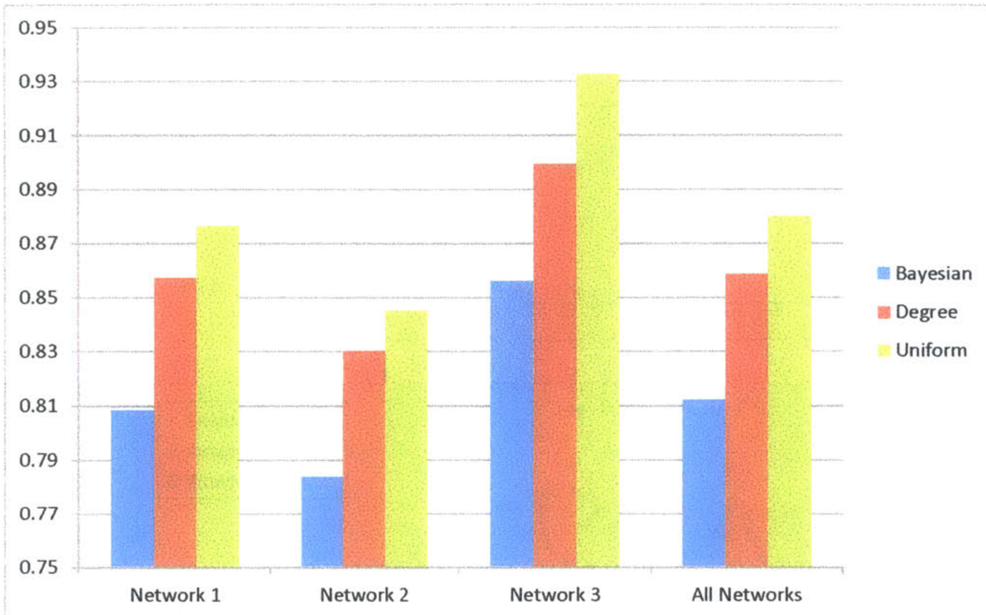


Figure 1-7: Fraction of actions explained at network level

Note: The fraction of actions explianed x is $x := \frac{y-50}{50}$, where y is the share of actions predicted correctly. This is the right normalization since we could always explain half the actions by flipping a fair coin.

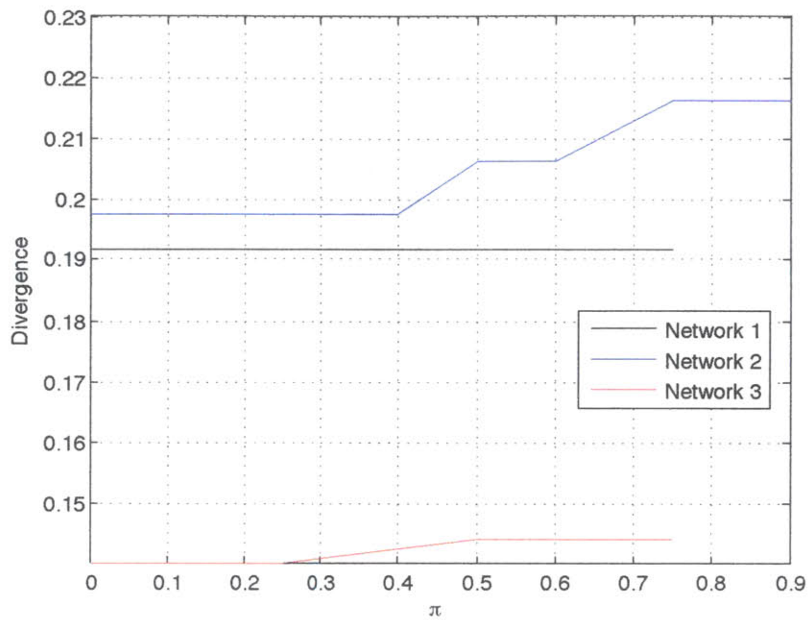


Figure 1-8: Divergence with Uniform Weighting

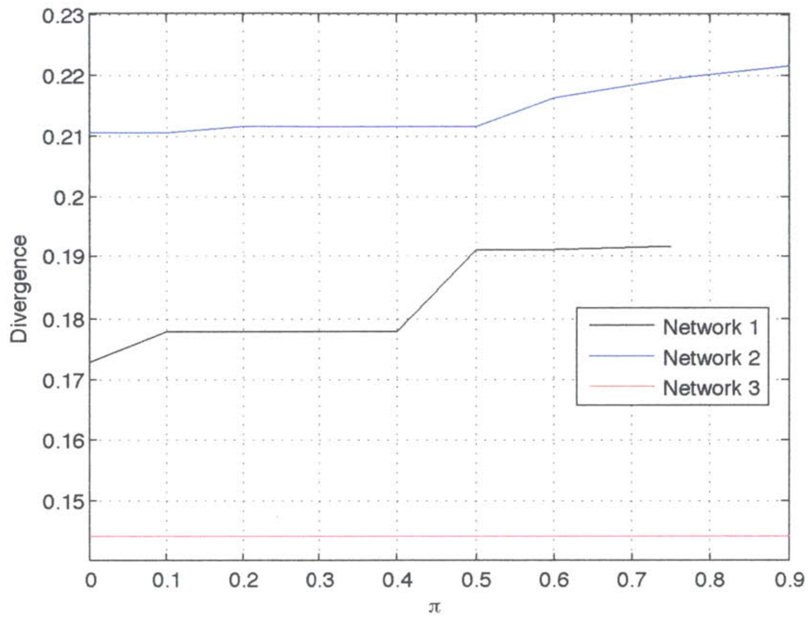


Figure 1-9: Divergence with Degree Weighting

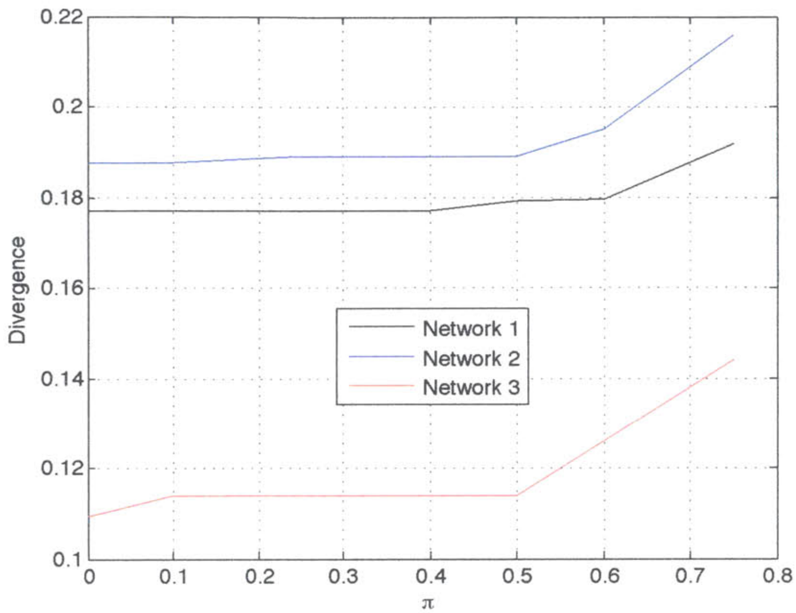


Figure 1-10: Divergence with Eigenvector Weighting

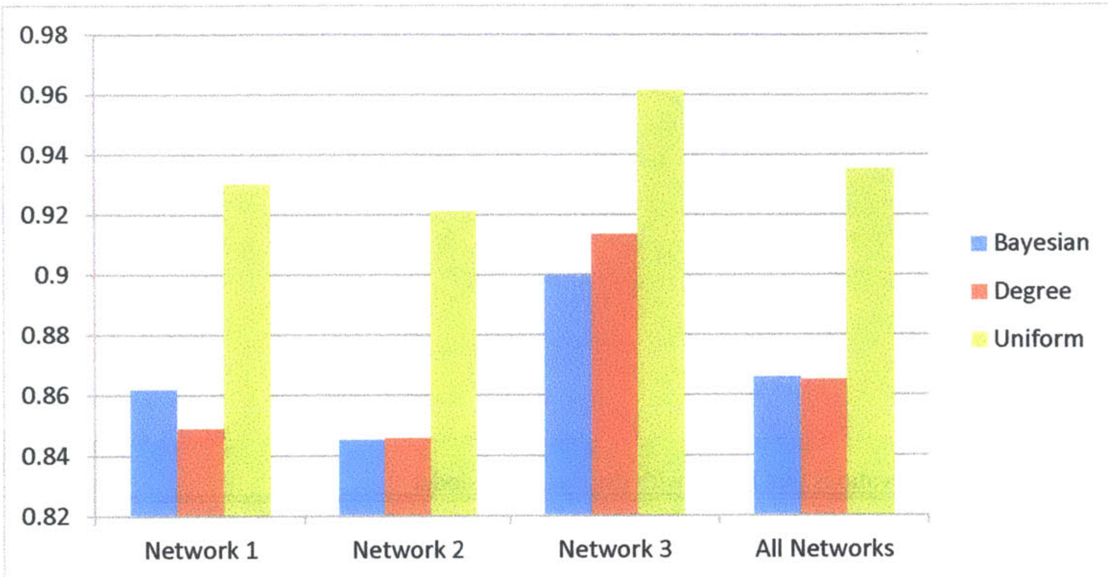


Figure 1-11: Fraction of actions explained at individual level
Note: The fraction of actions explianed x is $x := \frac{y-50}{50}$, where y is the share of actions predicted correctly. This is the right normalization since we could always explain half the actions by flipping a fair coin.

Table 1: Fraction of Observations that Differ with the Bayesian Model

Panel A: Networks from Choi et al. (2005, 2009)

	Total Divergence			Divergence in Final Period		
Network	*Uniform*	*Degree*	*Eigenvector*	*Uniform*	*Degree*	*Eigenvector*
1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
3	9.37%	21.87%	8.98%	12.67%	18.67%	7.67%

Panel B: Networks Selected in This Paper

	Total Divergence			Divergence in Final Period		

Table 3: Network Level Divergence

Network	*Total Obs*	*Bayesian*	*Uniform*	*Degree*	*Eigenvector*
All Networks	9,205	0.1878	0.1198	0.1413	0.2703
1	3,045	0.1917	0.1236	0.1428	0.2229
2	3,031	0.2161	0.1548	0.1698	0.3026
3	3,129	0.1440	0.0673	0.1006	0.2909

Notes: Network level divergence for the complete information Bayesian model, uniform DeGroot weighting, degree DeGroot weighting, and eigenvector DeGroot weighting.

Table 4: Significance Tests for Network Level Divergence

H_0	H_a	All Networks	Network 1	Network 2	Network 3
Bayesian	Degree	0.0001	0.0041	0.0019	0.0133
Bayesian	Uniform	0.0001	0.0001	0.0001	0.0001
Bayesian	Eigenvector	0.9999	0.9397	0.9999	0.9999
Degree	Uniform	0.0006	0.0615	0.0696	0.0001
Degree	Eigenvector	0.9999	0.9999	0.9999	0.9999
Uniform	Eigenvector	0.9999	0.9999	0.9999	0.9999

Notes: The test statistic is the normalized difference in the divergence functions of the null and the alternative model. We show the probability that the test statistic is less than 0, estimated via bootstrap with replacement.

Table 5a: Individual Divergence for DeGroot Models

Panel A: Divergence Across Networks

Network	*Observations*	Divergence		
		Uniform	*Degree*	*Eigenvector*
All	*9,205*	*0.0648*	*0.135*	*0.1083*
1	3,045	*0.0699*	*0.1513*	*0.0961*
2	3,031	*0.0788*	*0.1544*	*0.1562*
3	3,129	*0.0386*	*0.0866*	*0.0598*

Panel B: Share of Best-Fitting Sessions

H_0	H_1	Informative	H0 beats H1	Best model
Uniform	*Degree*	*0.65*	*0.8225*	*Uniform*
Uniform	*Eigenvector*	*0.4538*	*0.7712*	*Uniform*
Degree	*Eigenvector*	*0.3577*	*0.172*	*Eigenvector*

Notes: Panel A shows the individual level divergence across the three networks for each of the three DeGroot models.

In Panel B, the first column indicates the share of informative sessions over all sesssions, and the second column indicates the share of informative sessions where the H0 model beats the H1 model.

Table 5b: Lack of Bayesian Individual Behavior

H_1	Observations	Share Bayesian
Degree	46	17.39%
Uniform	98	17.35%
Eigenvector	98	17.35%

Note: "Observations" are the number of cases where there are discrepancies between the parent node action (which from $t>3$ Bayesian prescribes peripheral nodes should follow) and the action that the H1 model prescribes peripheral nodes. "Share Bayesian" is the share of observations where peripheral nodes indeed follow the parent node action.

Table 6: Zero Probability Information Sets Reached

Panel A: Complete Information Model

Network	% Individuals	% Treatments	% Observations
1	98.95%	98.95%	32.35%
2	100.00%	100.00%	33.18%
3	100.00%	100.00%	36.96%

Panel B: Incomplete Information Model

Degree Weighting Alternative

Network	% Individuals	% Treatments	% Observations
1	98.95%	98.95%	23.46%
2	100.00%	100.00%	19.78%
3	100.00%	100.00%	23.02%

Uniform Weighting Alternative

Network	% Individuals	% Treatments	% Observations
1	98.95%	98.95%	22.40%
2	100.00%	100.00%	29.29%
3	100.00%	100.00%	26.30%

Eigenvector Weighting Alternative

Network	% Individuals	% Treatments	% Observations
1	98.95%	98.95%	23.28%
2	100.00%	100.00%	25.91%
3	66.67%	100.00%	20.29%

Notes: Panel A presents results for the complete information Bayesian model. Panel B presents results for the incomplete information Bayesian model against DeGroot alternatives. % Individuals refers to the fraction of individuals who reach a zero probability information set. % Treatments refers to the fraction of treatments (network x village) that reaches a zero probability information set. % Observations refers to the fraction of individual x time units that reach a zero probability information set.

Table 7: Individual Level Divergence For Bayesian Model

Alternative	*Network*	*No. Obs.*	*At optimal π*
Degree	1	1858	0.173
	2	1898	0.210
	3	1358	0.144
Uniform	1	1847	0.192
	2	1675	0.198
	3	1300	0.140
Eigenvector Centrality	1	1822	0.177
	2	1750	0.188
	3	1383	0.109

Note: We present the individual level divergence for the Bayesian model on information sets in the support. No. of observations denotes the number of triples (individual, village, treatment) that were taken at non-zero probability information sets. Divergence is calculated conditional on all agents being Bayesian, with other potential types being Alternative. Optimal $\pi = 0$ for all networks and alternatives.

Table 8: Significance Tests for Individual Level Divergence

H_0	H_a	All Networks	Network 1	Network 2	Network 3
Bayesian	Degree	0.6438	0.794	0.357	0.0768
Bayesian	Uniform	0.0001	0.0001	0.0001	0.0001
Bayesian	Eigenvector	0.0001	0.0001	0.4267	0.0001
Degree	Uniform	0.0001	0.0001	0.0001	0.0001
Degree	Eigenvector	0.0001	0.0001	0.9999	0.0001
Uniform	Eigenvector	0.9999	0.9992	0.9999	0.9999

Notes: The test statistic is the normalized difference in the divergence functions of the null and the alternative. We show the probability that the test statistic is less than 0, estimated via bootstrap with replacement.

Table 8: Weight on indirect neighbors according to the number of multiple direct neighbors

	Data		Bayesian		Degree		Uniform	
	All	Restricted	All	Restricted	All	Restricted	All	Restricted
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Signal	.4974***	.5351***	.3965***	.4429***	.5725***	.6204***	.6017***	.677***
	[.0218]	[.0279]	[.017]	[.0223]	[.0169]	[.0211]	[.0159]	[.0201]
Direct	.7735***	.7485***	.775***	.8048***	.9125***	.8913***	.8954***	.9078***
	[.0311]	[.0438]	[.0323]	[.0384]	[.0335]	[.0425]	[.0381]	[.0475]
One Way	.3157***	.2609***	.7056***	.8189***	0.0461	-0.0614	.1306***	.077*
	[.0405]	[.0542]	[.0363]	[.0486]	[.035]	[.0421]	[.0368]	[.0447]
Two Ways	.2177***	.1751***	.3582***	.4096***	.0797*	0.024	.1985***	.1841***
	[.0355]	[.0424]	[.0337]	[.0338]	[.0421]	[.0449]	[.0459]	[.0519]
TwoWays > OneWay								
t-statistic	-1.97	-1.557	-8.653	-8.136	0.7675	1.925	1.481	2.215
Probability	0.9741	0.9386	1	1	0.2224	0.0287	0.071	0.0146
N	1750	1271	1750	1271	1750	1271	1750	1271
R -squared	0.4724	0.373	0.5322	0.4632	0.588	0.5453	0.5974	0.5571

Note: Robust standard errors, clustered at the village by game level, in brackets. Ouctome variable is action in round 3. "Direct" is the average signal of direct neighbors, "One Way" is the average signal of indirect neighbors only thorugh one direct neighbor, and "Two Ways is the average signal of indirect neighbors thorugh two direct neighbors. Column (1) is the regression with all data. Column (2) is the regression restricting to treatments that are informative for the comparisons Bayesian - Degree and Bayesian - Uniform. * p<.1, ** p<.05, *** p<.01

Table 9: Weight on indirect neighbors according to whether they provide new information

	Data		Bayesian		Degree		Uniform	
	All	Restricted	All	Restricted	All	Restricted	All	Restricted
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Signal	.4284***	.4573***	.4729***	.5264***	.4507***	.482***	.5694***	.6468***
	[.0347]	[.0425]	[.0286]	[.0356]	[.032]	[.0359]	[.0307]	[.0357]
Direct	.7508***	.7011***	.8095***	.8007***	.8436***	.8167***	.7576***	.7275***
	[.0363]	[.0471]	[.0285]	[.0346]	[.0402]	[.0489]	[.0391]	[.0492]
Indirect New	.2455***	.1803***	.3768***	.3489***	.1558***	.0883**	.2548***	.2063***
	[.0331]	[.0377]	[.0261]	[.0317]	[.0333]	[.0384]	[.0283]	[.0338]
Indirect Repeated	.1715***	.1669***	.1029***	.0801*	.2463***	.2619***	.1724***	.1554***
	[.0416]	[.048]	[.0354]	[.0427]	[.0398]	[.0447]	[.0417]	[.0492]
N	1587	1250	1587	1250	1587	1250	1587	1250
R -squared	0.4628	0.3958	0.4953	0.4135	0.5819	0.5475	0.5945	0.5687

Note: Robust standard errors, clustered at the village by game level, in brackets. Outcome variable is action in round 3. "Direct" is the average signal of direct neighbors, "Indirect New" is the average signal of indirect neighbors that provide new information, and "Indirect Repeated" is the average signal of indirect neighbors that do not provide new information. Column (1) is the regression with all data. Column (2) is the regression restricting to treatments that are informative for the comparisons Bayesian - Degree and Bayesian - Uniform. * p<.1, ** p<.05, *** p<.01

1.8 Appendix A - Complete Information Bayesian Algorithm

In this appendix we describe the algorithm for computing the actions under the assumption of complete information Bayesian agents.

1.8.1 Setup

Suppose that all agents learn about the state of the world using Bayes' rule and assume that this is common knowledge. Note that the signal endowment w is a sufficient statistic over the actions that agents take, since θ is never observed. In turn, the inference that agents need to do from the other agents' play is only with respect to the signal endowment. To proceed we must establish some notation.

We define $\mu_0^{\theta}(w)$ as the probability of signal endowment w when the true state of the world is θ. Then

$$\begin{aligned}
\mu_0^{\theta=1}(w) &:= \Pr(w \mid \theta = 1) \qquad (1.3)\\
&= p^{\#\{i:w_i=1\}}(1-p)^{N-\#\{i:w_i=1\}} = p^{\left(\sum_{i=1}^{n} w_i\right)}(1-p)^{n\left(1-\sum_{i=1}^{n} w_i\right)}\\
\mu_0^{\theta=0}(w) &:= \Pr(w \mid \theta = 1) \qquad (1.4)\\
&= p^{\left(\sum_{i=1}^{i=N} \frac{w_i}{n}\right)}(1-p)^{\left(N-\sum_{i=1}^{i=N} \frac{w_i}{n}\right)}
\end{aligned}$$

Following the same reasoning, define $\mu_{i,t}^{\theta}(w, \overline{w})$ as the belief probability distribution that agent i has at period t of the game over states $w \in W$, given that the true signal endowment is $\overline{w}$. Observe that for different signal endowments, the information sets that each agent will observe are clearly going to be different. In turn, the belief over signal endowments that each agent has at each state depends on what was the true signal endowment.

Define $p_{i,t}(\overline{w})$ as the probabilitythat agent i at stage t puts on the event $\theta = 1$, if the information set she observes comes from the true signal endowment being $\overline{w}$. Moreover, put $\mathbf{a}_{i,t}(\overline{w}) \in \{0,1\}$ as the action that agent i takes at stage t of the game if the information set reached comes from the true signal endowment $\overline{w}$.

We will iteratively find $\mu^{\theta}_{i,t}(w,\overline{w})$, $\mathbf{a}_{i,t}(\overline{w})$ and $p^{\theta}_{i,t}(\overline{w})$. We begin by finding these objects at $t = 1$. These are given by

$$p^{\theta}_{i,1}(\overline{w}) = \begin{cases} p \text{ if } \overline{w}_i = s_i = 1 \\ 1 - p \text{ if } \overline{w}_i = s_i = 0, \end{cases} \tag{1.5}$$

$$\mathbf{a}_{i,1}(\overline{w}) = s_i, \tag{1.6}$$

$$\mu^{\theta=1}_{i,1}(w,\overline{w}) = \begin{cases} 0 \ \forall w \in W : \ w_i \neq s_i \\ \frac{\mu^{\theta=1}_0(w)}{\Pr(s_i|\theta=1)} \text{ if } w_i = s_i, \end{cases} \tag{1.7}$$

and

$$\mu^{\theta=0}_{i,1}(w,\overline{w}) = \begin{cases} 0 \ \forall w \in W : \ w_i \neq s_i \\ \frac{\mu^{\theta=0}_0(w)}{\Pr(s_i|\theta=0)} \text{ if } w_i = s_i. \end{cases} \tag{1.8}$$

Next, in order to model what exactly each agent observes, which will influence how beliefs are updated, we introduce some network-based notation. In particular, $N(i) := N_i \cup \{i\}$, the set of neighbors of agent i, including i herself. Next, define

$$\mathbf{a}^{(i)}_t(\overline{w}) := \left(\underbrace{\mathbf{a}_{j_1,t-1}(\overline{w}), \mathbf{a}_{j_2,t-1}(\overline{w}), ..., \mathbf{a}_{j_{d(i)},t-1}(\overline{w})}_{\text{actions taken by neighbors}}, \underbrace{\mathbf{a}_{i,t-1}(\overline{w})}_{\text{own past action}} \right) \tag{1.9}$$

to be the action profile that agent i sees at the beginning of state t, when the true state of the world is $\overline{w}$. If we just write $a^{(i)}_t$ we refer to a particular observed action profile.

1.8.2 Time $t+1$ iteration

At time t we have

$$\begin{aligned} \mu^{(\theta=1)}_{i,t}(w,\overline{w}) &= \Pr(w \mid \theta = 1, I_{i,t}(\overline{w})), \\ \mu^{(\theta=0)}_{i,t}(w,\overline{w}) &= \Pr(w \mid \theta = 0, I_{i,t}(\overline{w})), \\ p^{(\theta=1)}_{i,t}(\overline{w}) &= \Pr(\theta = 1 \mid I_{i,t}(\overline{w})), \\ p^{(\theta=0)}_{i,t}(\overline{w}) &= \Pr(\theta = 0 \mid I_{i,t}(\overline{w})), \end{aligned}$$

where $I_{i,t}(\overline{w})$ is the information set agent i at stage t of the game given that the true signal endowment is $\overline{w}$. Of course, agent i does not know $\overline{w}$, but only the information set $I_{i,t}$. Suppose now the agents receive new information, namely, $w \in I_{t+1}(\overline{w}) \subseteq I_t(\overline{w})$. Then

$$\mu_{i,t+1}^{(\theta=1)}(w,\overline{w}) := \Pr(w \mid \theta = 1, I_{i,t+1}(\overline{w})) = \begin{cases} 0 \text{ if } w \notin I_{i,t+1}(\overline{w}) \\ \frac{\mu_{i,t}^{(\theta=1)}(w,\overline{w})}{\sum_{w' \in I_{t+1}} \mu_{i,t}^{(\theta=1)}(w,\overline{w})} \text{ if } w \in I_{i,t+1}(\overline{w}) \end{cases}. \tag{1.10}$$

Likewise,

$$\mu_{i,t+1}^{(\theta=0)}(w,\overline{w}) := \Pr(w \mid \theta = 0, I_{i,t+1}(\overline{w})) = \begin{cases} 0 \text{ if } w \notin I_{i,t+1}(\overline{w}) \\ \frac{\Pr(w|\theta=0, I_t(\overline{w}))}{\sum_{w' \in I_{t+1}} \Pr(w'|\theta=0, I_{i,t}(\overline{w}))} \text{ if } w \in I_{i,t+1}(\overline{w}) \end{cases}. \tag{1.11}$$

Based on the new probability distribution over signal endowments, we can get the probability over θ as

$$\begin{aligned} p_{i,t+1}^{(\theta=1)}(\overline{w}) &= \Pr(\theta = 1 \mid I_{i,t+1}(\overline{w})) = \Pr(\theta = 1 \mid I_{i,t}(\overline{w}) \cap I_{i,t+1}(\overline{w})) \\ &= \frac{p_{i,t}^{(\theta=1)}(\overline{w}) \sum_{w \in I_{i,t+1}(\overline{w})} \mu_{i,t}^{(\theta=1)}(w,\overline{w})}{p_{i,t}^{(\theta=1)}(\overline{w}) \sum_{w \in I_{i,t+1}} \mu_{i,t}^{(\theta=1)}(w,\overline{w}) + \left(1 - p_{i,t}^{(\theta=1)}(\overline{w})\right) \sum_{w \in I_{i,t+1}} \mu_{i,t}^{(\theta=0)}(w,\overline{w})} \end{aligned} \tag{1.12}$$

Therefore, we need to compute the relevant information sets. Let $\mathbf{a}_{i,t}(w)$ be the action that agent i takes at time t if configuration of signals is w. Then we can consider the set of worlds that have positive probability at time t, given by

$$W_{t+1}^{(i)}\left(\mathbf{a}_t^{(i)}\right) = \{w \in W : \mathbf{a}_{j,t}(w) = a_{j,t} \text{ for all } j \in N(i)\} \tag{1.13}$$

and hence,

$$I_{i,t+1}(\overline{w}) = W_{t+1}^{(i)}\left(\mathbf{a}_t^{(i)}(\overline{w})\right). \tag{1.14}$$

Note that in deriving the information set, we were able to eliminate the path of actions observed by agent i so far by realizing that the actions taken by an agent are *deterministic functions of what they observe.* Thus, once we have conditioned on the signal endowment w, the actions observed $a_t^{(i)}(w)$ are completely determined. However, in environments in which we allow for random actions, as opposed to deterministic actions, this fact is no longer true. Though this would define a perfectly sensible and reasonable model, it astronomically complicates things in an astronomical manner.

1.8.3 Actions

The algorithm described so far gives us how a Bayesian agent i would update her beliefs if she knew $\mathbf{a}_{j,t}(\overline{w})$ for all j and has prior beliefs on signal endowments and states of the world given by

$$\left\{\mu_{i,t}^{(\theta=1)}(w,\overline{w}),\mu_{i,t}^{(\theta=0)}(w,\overline{w}),p_{i,t}^{(\theta=1)}(\overline{w})\right\}. \tag{1.15}$$

If agent i is Bayesian, then the decision at $t+1$ is given by (when there are no ties)

$$\mathbf{a}_{i,t+1}(\overline{w})=\begin{cases} 1 \text{ if } p_{i,t+1}^{(\theta=1)}(\overline{w})>\frac{1}{2} \\ 0 \text{ if } p_{i,t+1}^{(\theta=1)}(\overline{w})<\frac{1}{2} \\ a\in\{0,1\} \text{ if } p_{i,t+1}^{(\theta=1)}(\overline{w})=\frac{1}{2}. \end{cases} \tag{1.16}$$

Note that when $p_{i,t+1}^{(\theta=1)}(\overline{w})=\frac{1}{2}$, we need to use some tie breaking rule. We will use the "past action" rule. That is, when faced with a tie, an individual will play the action she played in the previous round, $p_{i,t+1}^{(\theta=1)}(\overline{w})=\frac{1}{2}\implies \mathbf{a}_{i,t+1}(\overline{w})=\mathbf{a}_{i,t}(\overline{w})$. Of course, one could think of many other tie breaking rules, including random tie breaking rule, where the agent plays each action with the same probability. However, as we will see, this such a model will be computationally intractable in our framework.

Observe that the above framework extends to situations where some agents play ad hoc decision rules via DeGroot learning. Suppose that each agent i may be of some type $\eta\in H:=\{\eta_1,\eta_2,...,\eta_K\}$. For example, take the type space to be

$$H=\{\text{Bayesian, Eigenvector, Uniform, Degree}\} \tag{1.17}$$

so each agent may be either a Bayesian agent or a DeGroot agent who constructs simple linear indexes from the past actions taken by neighbors. In particular, suppose that agent i has type $\eta_i=$ Degree. In world $\overline{w}$ and time $t+1$, she observes actions $\mathbf{a}_t^{(i)}$. Based on this, she defines the following index:

$$Degree_{i,t}(\overline{w}):=\sum_{j\in N(i)}\mathbf{a}_{j,t}(\overline{w})\,T_{ij}^{Pref}. \tag{1.18}$$

Therefore, the corresponding action rule is

$$\mathbf{a}_{i,t+1}^{P}(\overline{w})=\begin{cases} 1 \text{ if } Degree_{i,t}(\overline{w})>\frac{1}{2} \\ 0 \text{ if } Degree_{i,t}(\overline{w})<\frac{1}{2} \\ a\in\{0,1\} \text{ if } Degree_{i,t}(\overline{w})=\frac{1}{2}. \end{cases} \tag{1.19}$$

Similarly, we can construct $\mathbf{a}_{i,t+1}^{N}(\overline{w})$, $\mathbf{a}_{i,t+1}^{E}(\overline{w})$ using $T^{Uniform}$ and T^{Eig}, respectively.

As long as *agents' types are common knowledge*, the algorithm described so far can handle heterogeneity in agents' types without changing the nature of the Bayesian updating.

1.8.4 Algorithm to find Action Rules

Step 1:

- Initiate the algorithm with $\mu_{i,1}^{\theta=1}(w,\overline{w})$ and $\mu_1^{\theta=0}(w,\overline{w})$ given by (1.3) and (1.4), action rule $\mathbf{a}_{i,1}(\overline{w})$ as in (1.6) and $p_{i,1}(\overline{w})$ as in (1.5).

Step 2:

- At period $t+1$, start with $\left\{\mu_{i,t}^{\theta=1}(w,\overline{w}), \mu_{i,t}^{\theta=0}(w,\overline{w}), p_{i,t}(\overline{w}), \mathbf{a}_{i,t}(\overline{w})\right\}$. Derive $\mu_{i,t+1}^{\theta=1}(w)$ and $\mu_{i,t+1}^{\theta=0}(w)$ using using equations (1.11) and (1.10).
- Obtain $p_{i,t+1}(\overline{w})$ from (1.12), and then derive the action that each agent takes, $\mathbf{a}_{i,t+1}(\overline{w})$ depending on the agent's type.

Step 3:

- Repeat Step 2 until $t=T$.

1.9 Appendix B - Incomplete Information Bayesian Algorithm

We now describe a modification of the above environment wherein individuals can be either DeGroot or Bayesian. The main difference comes from the initiation of the algorithm with an expanded type space:

$$\mu_0^{\theta=1}(w) = \mu_0^{\theta=1}(s,\eta) = \left(p^{\sum_{i=1}^n s_i}(1-p)^{n-\sum_{i=1}^n s_i}\right)\prod_{i=1}^n \pi_i^{\eta_i}(1-\pi_i)^{1-\eta_i}, \tag{1.20}$$

$$\mu_0^{\theta=0}(w) = \mu_0^{\theta=1}(s,\eta) = \left((1-p)^{\sum_{i=1}^n s_i} p^{n-\sum_{i=1}^n s_i}\right)\prod_{i=1}^n \pi_i^{\eta_i}(1-\pi_i)^{1-\eta_i}. \tag{1.21}$$

Note that the assumption of independent types is immaterial to the description of the algorithm: we could substitute in principle the term $\prod_{i=1}^n \pi_i^{\eta_i}(1-\pi_i)^{1-\eta_i}$ for some function $F(\eta)=\Pr(\eta)$.

The extended type space, or signal, is now a draw of the signal and a draw of the learning process. After each agent sees her "signal" (s_i,η_i), we can calculate the derived measures over worlds as

$$\begin{aligned}
\mu_{i,1}^{\theta=1}(w,\overline{w}) &= \Pr\left((s,\eta),(\overline{s_i},\overline{\eta_i}) \mid (s_i,\eta_i),\theta=1\right) \\
&= \frac{\Pr(w\mid\theta=1)\Pr((s_i,\eta_i)\mid w,\theta=1)}{\sum_{z\in W}\Pr(z)\Pr((s_i,\eta_i)\mid z,\theta=1)} = \begin{cases} 0 \text{ if } s_i\neq\overline{s_i}, \text{ or } \eta_i'\neq\overline{\eta_i} \\ \frac{\mu_0^{\theta=1}(w)}{\sum_{z\in W}\mu_0^{\theta=1}(z)\Pr(s_i,\eta_i|z,\theta=1)} \text{ otherwise}\end{cases}
\end{aligned}$$

and similarly for $\mu_{i,1}^{\theta=0}(w,\overline{w})$. For the other needed objects for the algorithm, we use:

$$\mathbf{a}_{i,1}(\overline{w}) = \begin{cases} 1 \text{ if } \overline{s_i} = 1 \\ 0 \text{ otherwise} \end{cases}$$

$$p_{i,1}(\overline{w}) = \begin{cases} p \text{ if } \overline{s_i} = 1 \\ 1 - p \text{ if } \overline{s_i} = 0. \end{cases}$$

For the rest of the algorithm, the action rule will depend on the type. Let $\mathbf{a}_{i,t}^{B}(\overline{w})$ denote the probability with which agent i at stage t plays $a = 1$ in world $\overline{w}$ if she acts as a Bayesian and $\mathbf{a}_{i,t}^{M}(\overline{w})$ be the analogous if the agent acts as an $M-$weighter. The action profile at time t is

$$\mathbf{a}_{i,t}(\overline{w}) = \begin{cases} \mathbf{a}_{i,t}^{B}(\overline{w}) \text{ if } \overline{\eta_i} = 1 \\ \mathbf{a}_{i,t}^{M}(\overline{w}) \text{ if } \overline{\eta_i} = 0. \end{cases}$$

1.10 Appendix C - Filtering

Here we describe a filter to estimate the probability that an agent is Bayesian. Let

$$F_0(\eta) := \Pr(\eta \mid I_0, \hat{\pi})$$

be the probability of a given agent being Bayesian, where I_0 is the information set of the statistician at $t = 0$. By design of the experiment we know that $\{s, \theta\} = I_0$. Call s^* and θ^* the chosen values by the experimentalist. For example, if type endowments are independent of both θ and s and the location on the network, we have

$$F_0(\eta) = \prod_{i=1}^{n} \pi_i^{\eta_i} (1 - \pi_i)^{1-\eta_i}. \tag{1.22}$$

Now suppose that we have calculated such a probability through time $t - 1$,

$$F_{t-1}(\eta) = \Pr(\eta \mid I_{t-1}, \hat{\pi}).$$

Next, define

$$\mathcal{A}_{i,t-1}^{*}(\eta) := \Pr(a_{i,t} = 1 \mid I_{t-1}, (s^*, \eta), \hat{\pi}) = \begin{cases} 1 - \varepsilon \text{ if } A_{i,t-1}(\eta, s^*) = 1 \\ \varepsilon \text{ if } A_{i,t-1}(\eta, s^*) = 1. \end{cases} \tag{1.23}$$

This is the probability distribution that the statistician has over actions if she knew the true type endowment. (The probability ε is to ensure that as statisticians, we put positive weight on every history. We may later take $\varepsilon \to 0$.) This will not be a problem empirically, since all histories have positive probability empirically.

Observe that $\mathcal{A}_{i,1}^{*}(\eta) = \mathbf{1}(s_i = s_i^*)$ for all $\eta \in H$ for any (reasonable) model that for any agent, if they see only their signal, they choose their own signal. Let $\mathbf{a}_t^*$ be the $n \times 1$ action vector observed by the experimenter at time t. This function allows us to get the conditional probabilities over actions of all agents as:

$$\begin{aligned}\Pr\left(\mathbf{a}_t \mid I_{t-1},(s^*,\eta),\hat{\pi}\right) &= \left(\prod_{j:a_j=1}\Pr\left(a_{j,t}=1 \mid I_{t-1},(s^*,\eta),\hat{\pi}\right)\right)\left(\prod_{j:a_j=0}\left(1-\Pr\left(a_{j,t-1}=1 \mid I_{t-1},(s^*,\eta)\right)\right)\right) \qquad (1.24)\\ &= \left(\prod_{j:a_j=1}\mathcal{A}^*_{i,t-1}(\eta)\right)\left(\prod_{j:a_j=0}\left(1-\mathcal{A}^*_{i,t-1}(\eta)\right)\right).\end{aligned}$$

We have used in (1.24) the fact that agents randomize over actions independently. Then, after observing action vector $a^*_t \in \{0,1\}^n$, the statistician updates her beliefs over types as

$$\begin{aligned}F_t(\eta) &= \Pr(\eta \mid I_{t-1},a^*) = \frac{\Pr(\eta \mid I_{t-1},\hat{\pi})\Pr(\mathbf{a}^*_t \mid I_{t-1},\eta,\hat{\pi})}{\sum_{\tilde{\eta}}\Pr(\tilde{\eta} \mid I_{t-1})\Pr(\mathbf{a}^*_t \mid I_{t-1},\tilde{\eta},\hat{\pi})} \qquad (1.25)\\ &= \frac{F_t(\eta)\left(\prod_{j:a^*_j=1}\mathcal{A}^*_{i,t-1}(\eta)\right)\left(\prod_{j:a^*_j=0}\left(1-\mathcal{A}^*_{i,t-1}(\eta)\right)\right)}{\sum_{\eta\in H}F_t(\tilde{\eta})\left(\prod_{j:a^*_j=1}\mathcal{A}^*_{i,t-1}(\tilde{\eta})\right)\left(\prod_{j:a^*_j=0}\left(1-\mathcal{A}^*_{i,t-1}(\tilde{\eta})\right)\right)}.\end{aligned}$$

Finally, to finish up the algorithm, we need to calculate $\mathcal{A}^*_{i,t}(\eta)$. This comes directly from the algorithm described above. The algorithm then to get the distribution of η conditional on the whole set of information is:

Step 1: Initiate algorithm with $F_0(\eta)$ as in (1.22)and an action function $\mathcal{A}^*$ as described above. Moreover, introduce information about s^* (only thing that we actually need)

Step $\mathbf{t<T}$: Taking $\mathcal{A}^*_{i,t-1}(\eta)$ as given, run learning code as in the previous section and calculate $\mathcal{A}^*_{i,t}(\eta)$ as in (1.23)

Step $\mathbf{t=T}$: Once $\mathcal{A}^*_{i,T-1}(\eta)$ is calculated, calculate likelihood over type endowments as

$$F_T(\eta) = \frac{F_{T-1}(\eta)\left(\prod_{j:a^*_{T_j}=1}\mathcal{A}^*_{i,T-1}(\eta)\right)\left(\prod_{j:a^*_T=0}\left(1-\mathcal{A}^*_{i,T-1}(\eta)\right)\right)}{\sum_{\tilde{\eta}\in H}F_{T-1}(\tilde{\eta})\left(\prod_{j:a^*_{T_j}=1}\mathcal{A}^*_{i,T-1}(\tilde{\eta})\right)\left(\prod_{j:a^*_T=0}\left(1-\mathcal{A}^*_{i,T-1}(\tilde{\eta})\right)\right)} \qquad (1.26)$$

Last Step : Get the probability of being Bayesian of agent i as

$$\Pi_{i,t,v}(\pi) := \Pr(\eta_i = 1) = \sum_{\tilde{\eta}\in H:\tilde{\eta}_i=1} F_T(\tilde{\eta}). \qquad (1.27)$$

1.11 Appendix D - Stuck Nodes and Social Quilt Trees

This section describes some results on how nodes become stuck and develops an example of a social quilt tree.

1.11.1 Stuck Nodes

Given an undirected graph $G = (V, E)$ and a subset of nodes $v \subseteq V$ we define $G_v = (v, E_v)$as the induced subgraph for subset v, where $E_v = \{(ij) \in E : \{i,j\} \subseteq v\}$. Given a subgraph G_v, let $d_i(G_v)$ be the degree of node i in subgraph G_v. Let $a_{it} \in \{0,1\}$ be the action that node $i \in V$ takes at round $t \in \mathbb{N}$, which we will assume follows the uniform DeGroot action model; i.e. $a_{i,t} = \mathbf{1}\left\{\frac{1}{d_i+1}\sum_{j \in N_i} a_{j,t-1} > \frac{1}{2}\right\}$. We allow for any tie-breaking rule when $\frac{1}{d_i+1}\sum_{j \in N_i} a_{j,t-1} = \frac{1}{2}$.

Lemma 5. *Take a subset of individuals $v \subseteq V$ such that there exists $h \in \mathbb{N}$ with*

$$h \leq d_i(G_v) \leq d_i < 2h + 1 \text{ for all } i \in v$$

If agents behave according to the uniform weighting DeGroot action model, and at some $T \in \mathbb{N}$ we have $a_{i,T} = a \in \{0,1\}$ for all $i \in v$, then $a_{i,t} = a$ for all $t \geq T$.

Proof. The proof is by induction: without loss of generality, suppose $a_{i,T} = 1$ for all $i \in v$. Of course, for $t = T$ the result is trivially true. Suppose now that $a_{i,t} = 1$ for all $i \in v$ and $t \geq T$, and we need to show that $a_{i,t+1} = 1$ too. Let $I_{i,t+1} = \frac{1}{d_i+1}\sum_{j \in N_i} a_{j,t}$ be the index of uniform weighting. We then now that $I_{i,t} \geq \frac{1}{2}$ for all nodes in v, and it suffices to show that this implies $I_{i,t+1} \geq \frac{1}{2}$. Observe,

$$\begin{aligned} I_{i,t+1} &= \frac{\sum_{j \in N_i} a_{j,t}}{d_i+1} = \frac{\sum_{j \in v \cap N_i} \underbrace{a_{j,t}}_{=1} + \sum_{j \in N_i - v} a_{j,t}}{d_i+1} \underset{(i)}{\geq} \frac{h+1+\sum_{j \in N_i - v} a_{j,t}}{d_i+1} \\ &\geq \frac{h+1}{d_i+1} \underset{(ii)}{>} \frac{1}{2}. \end{aligned}$$

We have used in (i) that $d_i(G_v) \geq h$ and $a_{j,t} = 1$ for all $j \in v$. Inequality (ii) comes from the fact that

$$\frac{h+1}{d_i+1} > \frac{1}{2} \Longleftrightarrow d_i < 2h+1.$$

Therefore, we have that $I_{i,t+1} > \frac{1}{2}$ for any $i \in v$, implying then that $a_{i,t+1} = 0$, as we wanted to show. □

This lemma says that whenever we find a subset of nodes v such that each node has more connections to nodes in v than it has outside v, then whenever they reach consensus, they would remain there forever. We present an useful corollary of Lemma 5, which we will use when studying the family SQT_r.

Lemma 6. *Take a family of nodes $v \in V$ such that there exists $k \in \mathbb{N}$ such that*

1. G_v is a k-regular graph
2. $d_i < 2k + 1$ for all $i \in v$.

Then, if agents behave according to the uniform weighting DeGroot action model, and at some $T \in \mathbb{N}$ we have $a_{i,T} = a \in \{0, 1\}$ for all $i \in v$, then $a_{i,t} = a$ for all $t \geq T$.

Proof. Simply take $h = k$ and apply Lemma 5. □

See that any triangle in SQT_r is a $2 -$ regularsubgraph, and that each node in it has $d_i = 4 < 2 \cdot 2 + 1 = 5$, so we apply Corollary 6 with $k = 2$. So, whenever a triangle achieves concensus, it remains there forever.

1.11.2 Social Quilt Trees: Preliminaries

We define $S_r = \{i \in V_r : i \text{ gets stuck}\}$ and let $N_r = \#(V_r)$, the number of nodes in SQT_r. Our object of interest is the random variable

$$\mathcal{F}_r(r) = \text{Fraction of nodes in } SQT_r \text{ that gets stuck} \equiv \frac{\#S_r}{N_r}$$

which is a random variable. Our objective is to get an asymptotic bound on $\mathcal{F}_r$. Since we do not yet know whether $\mathcal{F}_r$ has a limit for almost every realization, we define $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$as

$$\underline{\mathcal{F}} = \liminf_{r \to \infty} \mathcal{F}_r \text{ and } \overline{\mathcal{F}} = \limsup_{r \to \infty} \mathcal{F}_r \tag{1.28}$$

which is well defined for all realizations of the sequence $\mathcal{F}(r)$ and so it is a well defined random variable. Namely, we want to get the tightest asymptotic lower and upper bounds for the fraction of stuck nodes. Our objective is to get a number $F \in [0, 1]$ such that $\underline{\mathcal{F}} \geq \underline{F}$ and $\overline{\mathcal{F}} \leq \overline{F}$ almost surely; i.e. $\mathrm{P}\left\{\underline{F} \leq \underline{\mathcal{F}} \leq \overline{\mathcal{F}} \leq \overline{F}\right\} = 1$.

Any SQT_r has an initial parent triangle $P = (i_0, i_1, i_2)$. For any node $i \in V_r$ we define the distance to parent $d(i, P) = \min\{d(i, i_0), d(i, i_1), d(i, i_2)\}$ where $d(i, j)$is the minimum number of links we have to go through to connect node i with node j. Likewise, given a triangle $T = (i_s, i_k, i_j)$, we define the distance between triangle T and the parent triangle P as

$$d(T, P) = \max_{i \in T}\{d(i, P)\}$$

Definition 7. Given a graph $SQT_r = (V_r, E_r)$ and $s \in \{1, ..., r\}$ we define R_s, the level s ring of SQT_r as the subgraph $R_s = (V_r^s, E_r^s)$ of nodes that lie in triangles with distance to parent $d(T, P) = s - 1$.

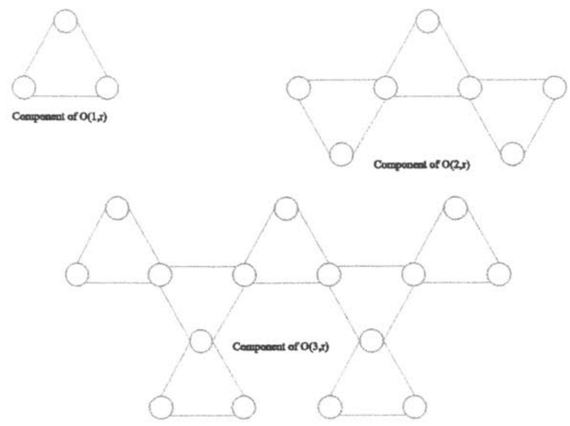


Figure 1-12: Outer Rings.

Intuitively, a ring is just the collections of triangles that lie in the $s-th$ level of the social quilt tree as seen in Figure 1-12. Note that for all s, R_s is a graph that consists of disconnected triangles.

Define

$$OR_r(k) := \bigcup_{s=0}^{s=k} R_{r-s}$$

as the subgraph formed by the outer rings from $r-k$ to r. This subgraph is also disconnected, with a lot of components, which now are no longer triangles, but rather "trees of triangles" as pictured in Figure 1-12.

Let $C \subset OR_r(k)$ be a component (a subgraph as shown in Figure 1-12). The last level of nodes in every component correspond to terminal nodes. The most important property of these components is that **the only connection between each component C and the rest of the graph is the parent node of the component** C, denoted by i_C . This will be the key property of these components, which we will try to explore.

Define

$$\Psi_r(k) = \frac{\#\{OR_r(k) \cap S_r\}}{\#\{OR_r(k)\}}$$

to be the fraction of stuck nodes in $OR_r(k)$

$$\underline{\Psi}(k) := \liminf_{r\to\infty} \frac{\#\{OR_r(k) \cap S_r\}}{\#\{OR_r(k)\}}$$

and

$$\overline{\Psi}(k) := \limsup_{r\to\infty} \frac{\#\{OR_r(k) \cap S_r\}}{\#\{OR_r(k)\}}$$

which is also a well defined random variable. These are the tightest asymptotic lower and upper bounds on the fraction of nodes stuck in the last k rings. That is, a lower bound on the fraction of nodes in $OR_r(k)$ that get stuck.

Lemma 8. *For all $k \in \mathbb{N}$, $\lim_{r \longrightarrow \infty} \frac{\#\{OR_r(k)\}}{N_r} = \frac{2^{k+1}-1}{2^{k+1}}$.*

Proof. Let $L_r =$ number of terminal triangles in ring r. Of course, we have that $L_r = T_{r-1}$. Because of how CSQ_r grows, we have the following recursion for L_r :

$$L_{r+1} = 2L_r \text{ and } L_2 = 3.$$

It can be easily shown that

$$L_r = 3 * 2^{r-2}. \tag{1.29}$$

We also need to calculate $N_r = \#(V_r)$. Again, because of how CSQ_r is generated, we have the following recursion for N_r :

$$N_{r+1} - N_r = 2L_{r+1}$$

and it can be also easily shown that

$$N_r = 3\left(2^r - 1\right) \tag{1.30}$$

Finally, let n_k be the number of nodes in a component $C \subset OR_r(k)$. It is also easy to show that

$$n_k = 2^{k+1} - 1$$

Now, we can state the result. Observe that

$$\begin{aligned}
\frac{\#\{O_r(k)\}}{N_r} &= \frac{\overbrace{n_k}^{\text{nodes per component}} \times \overbrace{L_{r-k+1}}^{\text{number of components}}}{N_r} = \left(2^{k+1} - 1\right) \frac{3 \times 2^{r-k+1-2}}{3\left(2^r - 1\right)} \\
&= \frac{2^{k+1} - 1}{2^{k+1}} \left(\frac{2^r}{2^r - 1}\right) \longrightarrow_{r \longrightarrow \infty} \frac{2^{k+1} - 1}{2^{k+1}}
\end{aligned}$$

as we wanted to show. □

The following proposition is the key to understand how to get bounds on $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$ by getting bounds on $\underline{\Psi}(k)$ and $\underline{\Psi}(k)$

Proposition 9. *Suppose there exist functions $\overline{\psi}, \underline{\psi} : \mathbb{N} \rightarrow [0,1]$ such that for all k we have*

$$\underline{\psi}(k) \leq \underline{\Psi}(k) \leq \overline{\Psi}(k) \leq \overline{\psi}(k) \text{ almost surely.}$$

Then, for all $k \in \mathbb{N}$ almost surely,

$$\underline{\mathcal{F}} \geq \frac{2^{k+1} - 1}{2^{k+1}} \underline{\psi}(k) \tag{1.31}$$

and

$$\overline{\mathcal{F}} \leq 1 - \left(\frac{2^{k+1} - 1}{2^{k+1}}\right) \left[1 - \overline{\psi}(k)\right]. \tag{1.32}$$

Proof. Lets focus only on inequality 1.31, since 1.32 follows the same reasoning. See that

$$\begin{aligned}\mathcal{F}(r) = & \frac{\#\{O_r(k)\}}{N_r}\left(\frac{\#\{O_r(k)\cap S_r\}}{\#\{O_r(k)\}}\right)+\frac{\#\{S_r-O_r(k)\}}{N_r} \\ \geq & \frac{\#\{O_r(k)\}}{N_r}\left(\frac{\#\{O_r(k)\cap S_r\}}{\#\{O_r(k)\}}\right)\end{aligned}$$

so, for all realizations,

$$\begin{aligned}\underline{\mathcal{F}} &= \liminf_{r\to\infty}\mathcal{F}(r)\geq\left(\lim_{r\longrightarrow\infty}\frac{\#\{O_r(k)\}}{N_r}\right)\left(\liminf_{r\to\infty}\left(\frac{\#\{O_r(k)\cap S_r\}}{\#\{O_r(k)\}}\right)\right) \\ &= \left(\lim_{r\longrightarrow\infty}\frac{\#\{O_r(k)\}}{N_r}\right)\underline{\Psi}(k)=\frac{2^{k+1}-1}{2^{k+1}}\underline{\Psi}(k).\end{aligned}$$

This, together with the fact that $\underline{\Psi}(k)\geq\underline{\psi}(k)$ almost surely, finishes the proof. □

Note that this proposition is true for any learning model (Bayesian or DeGroot). The learning model plays a role when calculating the bounds $\underline{\psi}$and $\overline{\psi}$. See that condition 1.31 and 1.32 are bounds on $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$, which do not depend on k: therefore, these are bounds for all k: the higher k, the tighter the bound we get.

1.11.3 Bounding stuck nodes in the Uniform Weighting model

To normalize, we will assume that the true state of nature is $\theta = 1$, which implies that as $r\to\infty$ the fraction of nodes with true signals is $p>\frac{1}{2}$. The idea is pretty simple: take a component $C=(V_C,E_C)\subset OR(k,r)$. As we mentioned before, the only connection between C and the rest of the graph is through the parent node i_C (as seen in Figure 1-12). Let $W_C=\{0,1\}^{n_k}$ be the set of signal endowments for nodes in C. We will try to find a lower bound $\underline{\psi}_k(w)$ for each signal endowment realization in C such that, when signal endowment is w, the fraction of stuck nodes in C is larger that $\underline{\psi}_k(w)$fraction of stuck nodes in C if endowment is $w\geq\underline{\psi}_k(w)$. If we can find such $\underline{\psi}_k(w)$, then we can use a law of large numbers to argue that

$$\underline{\Psi}(k)\geq\underline{\psi}(k)\equiv\mathbb{E}_{w\in W_C}\left\{\underline{\psi}_k(w)\right\}\text{ almost surely}$$

because the realizations of w in each component C is independent of each other. Likewise, if we can find a function $\overline{\psi}_k(w)$ to bound from above the fraction of stuck nodes, and then

$$\overline{\Psi}(k)\geq\overline{\psi}(k)\equiv\mathbb{E}_{w\in W_C}\left\{\overline{\psi}_k(w)\right\}\text{ almost surely}$$

Imagine first that the signal endowment of the upper triangle in C is $(0,0,0)$. Then, using Lemma 5we know that the upper triangle of C will get stuck from period $t=1$ on, and we can get the expected value of stuck nodes in C from there on. See that the fraction of nodes that get stuck in this component is only a function of the realization of $w\in W_C$, which is independent of the realization of the signal endowment of other components on $OR_r(k)$

When the signal endowment of the upper triangle in C is different from $(0,0,0)$, we make use of the other property we knew from C : that the only connection to the rest of the graph is through i_C,the uppermost node in C. Therefore, **a way of getting a lower bound on the number of nodes that get it wrong, is assuming that from round $t=2$ on, node i_C knows the truth, and plays $a_{i_C,t}=1$ for all $t \geq 2$.** Intuitively, we are making the graph to have the biggest effect possible in convincing nodes in C that actually, $\theta=1$, which can only do by making $a_{i_C,t}=1$ for all rounds other than $t \geq 2$. Once we have that, we can simulate the learning model on C, and calculate $\underline{\psi}_k(w)$ and $\overline{\psi}_k(w)$in this way.

There are two ways of calculating $\mathbb{E}_{w \in W_C}\left\{\underline{\psi}_k(w)\right\}$:

1. **Doing it explicitly:** This can be done for $k=2$ and $k=3$, because $\#\{W_C\}=128$. The bound when $k=3$ is the one we present in this paper.

2. **Monte-Carlo:** Of course, as k goes bigger, it is computationally unfeasible to calculate the expected value of $\underline{\psi}_k(w)$ explicitly, since

$$\#\{W_C\}=2^{2^{k+1}-1}=O(\exp(\exp(k)))$$

which grows super-exponentially. However, we can simulate random draws of $w \in W_C$ and get an estimate for $\mathbb{E}_{w \in W_C}\left\{\underline{\psi}_k(w)\right\}$ using law of large numbers.

The above method will also work for different learning models on the SQT_r family.

1.12 Appendix E - Proofs

Proof of Proposition 2. The first part follows from Mossel et al. (2012). The second part follows from (Golub and Jackson 2010), since every node has degree 2 or 4, Corollary 1 applies. Namely, $\max_{1\leq i\leq n}\frac{d_i}{\sum d_i}\to 0$ along our sequence and therefore the social learning process is wise. The third part follows from Lemma 8. □

Proof of Proposition 4. Let $g(n,T)$ be the number of calculations for the original model (with no trembles). It is given by

$$g(n,t)=\sum_{t=1}^{t=T} n(2^{2n}+2^{n+1})=nT(4^n+2^{n+1})=\Theta(nT4^n).$$

Meanwhile, let $f(n,T)$ be the number of calculations that needs to be done for a network of size n and played for T rounds.

$$\begin{aligned}
f(n,T) &= \sum_{t=1}^{T} n2^{nt+1}(1+2^{nt}) = 2n\left(\sum_{t=1}^{T} 2^{nt} + \sum_{t=1}^{T} 4^{nt}\right) \\
&= 2n\left[\frac{2^{n(T+1)}-2^{n}}{2^{n}-1} + \frac{4^{n(T+1)}-4^{n}}{4^{n}-1}\right] = \Theta\left(n4^{nT}\right).
\end{aligned}$$

Thus, the complexity ratio between the model with trembles and the model with no trembles is

$$\frac{f(n,T)}{g(n,T)} = \frac{2n\left[\frac{2^{n(T+1)}-2^{n}}{2^{n}-1} + \frac{4^{n(T+1)}-4^{n}}{4^{n}-1}\right]}{nT(4^{n}+2^{n+1})} = \Theta\left(\frac{1}{T}4^{n(T-1)}\right)$$

which completes the proof. □

Chapter 2

A Note on Payments in Experiments of Infinitely Repeated Games with Discounting

2.1 Introduction

Many lab experiments study behavior in repeated and dynamic games. The goal is to study behavior in a game specified by theory, which we call the target game. Typically, a participant plays a round of a game which then continues to the subsequent round with a given probability. In order to incentivize participant behavior, the experimenter pays the participant as a function of the history of play. Initially, the experiments in the literature paid participants for every round that they took part in. However, in order to compensate for self-insurance across rounds, some experiments moved to paying individuals for one randomly chosen round. An example of the argument offered was that "this payment structure prevents individuals from self-insuring income risk across rounds. The utility maximization problem of the experiment matches that of the theoretical model" ((?)). Examples of studies in the literature that either paid participants for all rounds or random rounds include Murnighan and Roth (1983), (?), (?), (?), (?; ?; ?), (?), (?), (?), Chandrasekhar et al. (2011), among others.[1]

In this paper, we present a general framework to think about payment in experiments. We argue that the payment scheme used may induce a game different from the target game; researchers must be cognizant of the game they induce their participants to play. As such, researchers ought to check whether their mode of payment implements the same subgame perfect equilibria (SPE) outcomes as the target game, and if it does not, they ought to account for this in their analysis.

[1](?) note that paying individuals for a randomly chosen game introduces a different form of discounting and show that the discount factor converges to the right one mandated by the target game. However, they argue that this alone suffices to show that payment for a randomly chosen round approximately implements the target game – a claim which we show to be false.

Our core results are as follows. First, we establish a test based on (?) to see if a payment scheme implements a particular SPE: the set of implementable actions at any given history must be independent of the history. Second, we argue that there is a simple payment scheme, paying individuals for the last round in which they participated, that implements the target game. Third, we discuss the payment scheme which pays individuals for all rounds, show that it does not implement the target game, and even induces individuals to become asymptotically indifferent about their actions. Fourth, we discuss the payment scheme which pays individuals for a round chosen at random and argue that it does not alleviate the problems. We show that it has a set of implementable outcomes that is generally different from the target game; people will typically behave differently along the equilibrium path. Moreover, we characterize the mechanics of the behavior induced by this payment scheme. Paying individuals for a randomly chosen round induces them to discount the future too much, from any period, relative to the target game. Individuals also become asymptotically indifferent between decisions.

To make this more concrete, we provide examples where the equilibrium is changed drastically simply when the researcher changes the payment scheme. Consider repeated Prisoner's Dilemma in which, for a certain level of discounting, under the target game cooperation is sustained in equilibrium. Consequently, paying participants for the last round only generates exactly the same SPE. However, paying for a randomly chosen round rules out cooperation in equilibrium for a certain range of the discounting parameter.

In general, our results demonstrate that distortions are sizeable. For any discount rate and any dynamic game, the "virtual" net present value (NPV) of a constant stream of consumption at the first round, under round at random payment, would be at the most half of the actual NPV. In particular, if the discount rate is 0.95, the "virtual" NPV would be less than 30% the value under the theoretical model. Moreover, paying in all rounds creates distortions when we allow for curvature in the utility for wealth.[2] In an example, we show that for utility arbitrarily close to linear (CES utility with almost zero elasticity of intertemporal substitution) the implementable outcomes are indeed very different and have very different asymptotic properties. Finally, we note that the implementation result can be easily generalized for any dynamic game with discounting.

A windfall of our results is that it allows a researcher to correct her theoretical predictions if she was interested in estimating a structural model form the lab experiment data. By explicitly characterizing the discounting induced by the payment mechanisms as well as its impact on the induced game, we suggest that a researcher interested in structurally estimating parameters simply use the corrected model when performing structural estimation.

The papers closest to ours are (?) and Sherstyuk et al. (2012). (?) study an infinite-horizon risk-sharing game note that the discounting behavior under random payments can be different, but argue as the incentive compatibility constraints in their game converges to those of the target game,

[2] Under risk-neutrality, all round payment would implement the same set of SPE outcomes, though we show in Section 2.6.2 that this result is not robust.

later rounds should be representative of equilibrium behavior in the target game. Sherstyuk et al. (2012) concurently and independently of us study the three payments schemes – all round, random round, and last round – but do so in an experimental study of repeated prisoner's dilemma. Their findings complement ours as they show that last round payment and all round payment generate comparable behavior whereas random round payments generated less cooperation. Our general treatment of the theory nests the discussion of both the risk-sharing games as well as risk-neutral examples such as the prisoner's dilemma as special cases.

The remainder of the paper is organized as follows. In Section 2.2 we present a simple example to provide intuition about our results. In this example we show an example where the round at random payment scheme induces a game which generates potentially different equilibrium paths than what would have transpired under the target game. Specifically, if individuals were payed correctly, cooperation would have been sustained in equilibrium but for a range of the parameter space, employing the wrong payment scheme rules out cooperation. Section 2.3 establishes the general framework and develops a test of whether a payment scheme implements the target game. Section 2.4 demonstrates that the last period payment scheme implements the target game. In Section 2.5 we study the payment schemes from the previous literature, payment for all rounds and payment for a round chosen at random, and characterize the mechanics of the problem. Section 2.6 discusses our results in the context of a simple model of savings behavior. Section 2.7 concludes.

2.2 A Simple Example

We begin with a simple example of infinitely-repeated Prisoner's Dilemma. Let the stage game be described by the following per-period payoffs in utility space, where $a > 1$, $a - b < 2$, as in (?).

	C	D
C	1,1	$-b, a$
D	$a, -b$	0,0

It is easy to see that under grim trigger punishment, cooperation is enforceable if and only if $\beta > 1 - \frac{1}{a}$.[3] To begin, let us suppose that $a = 3$, $b = 3/2$, and $\beta = 0.7$. Cooperation is sustainable if and only if $\beta \geq \frac{2}{3}$.

Assume that a participant is told that she must play the game in period 1, and that the game will continue with probability 0.7 (in every period when the game is still not over); she will continue to play until the game ends. She will be paid her payoff from a (uniformly) randomly selected round.

To show that the equilibria of the theoretical game, and the experimental game are not the same, suppose that cooperating forever, and playing grim trigger in any history different from $h^t = (C, C)^t$, is an equilibrium. This will generate a contradiction.

[3] This follows from $(1 - \beta) + \beta > (1 - \beta) a \Longleftrightarrow \beta > 1 - \frac{1}{a}$

In round 10, after having cooperated for 9 periods, if players have gotten there, the payoff of playing to the proposed equilibrium is 1. This is because the history will have been $(C, C)^t$, which paid 1 for each player. However, the payoff of playing D is larger than 1.

Finish at t	10	11	12
Probability	$1-\beta$	$(1-\beta)\beta$	$(1-\beta)\beta^2$
Payment of defection	1.2	1.09	1

One can check that

$$\sum_{t=0}^{\infty}(1-\beta)\beta^t\left(\frac{9}{10+t}+\frac{3}{10+t}\right)=\sum_{t=0}^{\infty}(0.7)^t\cdot(0.3)\cdot\left(\frac{9}{10+t}+\frac{3}{10+t}\right)=1.0116$$

so cooperation is not an equilibrium.

This implies is that paying participants for a randomly chosen round (RCR) induces a different discounting process, which may generate different equilibrium outcomes. In Section 2.5.1, we provide a complete characterization of the induced discounting.

For our Prisoner's Dilemma example, note that in the theoretical model, cooperation is sustainable if and only if $\beta \geq \underline{\beta}^* = 1 - \frac{1}{a}$, whereas random round payment induces a different threshold for cooperation, which we denote $\underline{\beta}^{rcr}$. It follows that if the researcher selects $\beta \in \left(\underline{\beta}^*, \underline{\beta}^{rcr}\right)$, though cooperation would be enforeable under the theoretical model, the random round payment scheme cannot sustain a fully cooperative equilibrium.

To illustrate this, we repeat the (infinitedly repeated) Prisoner's Dilemma exercise for several values of a and b. We present threshold values of β above which cooperation can be sustained.

	Theoretical Model $\underline{\beta}^*$	Randomly Chosen Round $\underline{\beta}^{rcr}$
$a=2$	0.50	0.57
$a=3$	0.67	0.82
$a=4$	0.75	0.90
$a=5$	0.80	0.92
$a=6$	0.83	0.94

The remainder of the paper extends this intuition into a far more general context. We characterize the biases that emerge from random round payment as well as all round payment. Furthermore, we develop a metric to measure the extent of the bias. However, moving to a general framework requires more machinery, which we develop in the next section.

2.3 Framework

2.3.1 Setup and Notation

The researcher is interested in conducting an experiment to test behavior in a repeated game Γ, specified as

$$\Gamma := \left\{ \left\{ A_i,\ u_i : \mathbb{R}_{\geq 0} \to \mathbb{R},\ r_i : \prod_{i=1}^{n} A_i \to \mathbb{R}_{\geq 0} \right\}_{i=1}^{n},\ \beta \in (0,1) \right\}$$

where A_i is the strategy space for agent i, and $A = \prod_{i=1}^{n} A_i$ the set of action profiles. Utility of agent i over sequences $\mathbf{a} = \{a_t\}_{t=0}^{\infty} \in A^{\infty}$ is given by the time separable utility function,

$$U_i(\mathbf{a}) = (1-\beta) \sum_{t=1}^{\infty} \beta^t u_i \left[r_i(a_t) \right]$$

where β is a common discount rate across agents, $r_i(a_t)$ is the monetary reward of the game and $u_i(c)$ is the utility of wealth in the theoretical model.

Let $\mathcal{H} = \cup_{t=1}^{\infty} A^t$ be the set of all possible histories and define a *pure strategy for agent* i as a function $\sigma_i : \mathcal{H} \to A_i$, which specifies after each possible history of play $h^t = (a_1, a_2, ..., a_{t-1})$, an action $a_{i,t} = \sigma_i(h^t)$. We will also denote $\sigma_i | h^t$ to be the conditional strategy on history h^t. Given a strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$, the *outcome of* σ is a sequence $\mathbf{a}(\sigma) \in A^{\infty}$ of actions prescribed by σ.

We write $\mathbf{SPE}(\Gamma)$ to denote the set of all subgame perfect equilibrium profiles[4]. A sequence $\mathbf{a} = \{a_t\}_{t=0}^{\infty}$ is an *implementable outcome* if there is an SPE profile $\sigma \in \mathbf{SPE}(\Gamma)$ such that $\mathbf{a} = \mathbf{a}(\sigma)$. The set of implementable outcomes is denoted by $\mathbf{O}(\Gamma) \subseteq A^{\infty}$. Given a history h^t, let $\mathbf{O}(\Gamma | h^t)$ be the set of implementable outcomes of the subgame starting from the node at h^t.

2.3.2 Payment Schemes and a Test of Implementation

The researcher wants to test Γ, which has an infinite horizon and in which agents have exponential discounting. As such, she has to design an alternative finite repeated game $\widehat{\Gamma}$ with the same strategy space, to test the predictions of game Γ. It is a well known fact that the infinite horizon and exponential discounting nature of Γ can be replicated by a game that ends in T periods, where T is a geometric random variable with prbability of termination β, whose realization is not known to the players; $\Pr(T = t) = (1-\beta)\beta^t$. We write $T \sim \text{Geom}(\beta)$.

To test the predictions of game Γ, the researcher chooses a *payment scheme* $\mathbf{R} = \{R_i(h^t)\}_i$ that specifies, for each history h^t, the payment that agent i should receive, in terms of a distribution $R_i(h^t)$ of monetary rewards *at the end of the game.*

[4] A profile σ is a SPE if and only if, for all histories $h^t \in \mathcal{H}$ we have $U_i \left[\mathbf{a}(\sigma \mid h^t) \right] \geq U_i \left[\mathbf{a}(\hat{\sigma}_i, \sigma_{-i} \mid h^t) \right]$ for all i and all $\hat{\sigma}_i \in \Sigma_i$

For example, she might choose the following payment scheme:

$$R_i(h^t) = \begin{cases} r_i(a_s) \text{ with prob. } \frac{1}{t} \text{ for all } s \leq t & \text{if} \quad t = T \\ 0 & \quad\quad \text{otherwise.} \end{cases}$$

We will call this the *randomly chosen round* payment scheme (where the choice is uniform at random) and we will denote it $\mathbf{R}_{rcr}$. The payment scheme induces a new repeated game, which we will call $\widehat{\Gamma}(\mathbf{R})$. At any history h^t, the individual evaluates future sequences according to:

$$\widehat{U}_i^{\mathbf{R}}(\mathbf{a} \mid h^t) = \mathbb{E}\left\{u_i\left[R_i\left(h^T \mid h^t\right)\right]\right\}.$$

If the researcher wants to test the predictions of game Γ, the payment scheme used *must not change the predictions of subgame perfect equilibria of the original game* Γ. This is formalized in the next definition.

Definition 10 (Implementation). We say a *payment scheme* $\mathbf{R} = \left\{R_i(h^t)\right\}_i$ *implements* Γ if the set of implementable outcomes from Γ and from $\widehat{\Gamma}(\mathbf{R})$ coincide, i.e. $\mathbf{O}\left(\widehat{\Gamma}(\mathbf{R})\,|h^t\right) = \mathbf{O}\left(\Gamma|h^t\right)$ for every $h^t \in \mathcal{H}$.

We refer to such schemes as *implementing schemes.* Given a payment scheme $\mathbf{R}$, we are interested in studying whether it implements Γ.It is easy to show, following (?), that under some mild conditions, any implementing scheme must be "memoryless" in the sense that the set of implementable outcomes under any history h^t is independent of the history itself. If the the payment scheme has history dependence, we might suspect that it is not an implementing scheme.

(i) $\mathbf{SPE}(\Gamma) \neq \varnothing$, (ii) $A = \prod_{i=1}^{i=n} A^i$ is a compact set, and (iii) $u_i(r_i(\cdot))$ are continuous functions of $a \in A$ for all $i = 1, 2, ..., n$.

Proposition 11 (Corollary to (?)). *Under Assumption 2.3.2, if* $\mathbf{R}$ *implements* Γ*, then* $\mathbf{O}\left(\widehat{\Gamma}(\mathbf{R})\,|h^t\right) = \mathbf{O}(\Gamma)$ *for all* $h^t \in \mathcal{H}$.

By making payments $R_{i,t}(h^t)$ depend on t and on past history h^{t-1} we will typically violate this property. We will show that both paying individuals for a randomly chosen round and paying individuals for all rounds (with any amount of curvature of the utility function) will typically violate this property. This is because when we pay for all rounds, the payment scheme explicitly depends on the entire history and when we pay for a randomly chosen round, the incentives faced by the agent will depend explicitly on the round number t. However, in the next section we show the existence of an implementing payment scheme that essentially works for any game.

2.4 Last Round Payment

Define the scheme $\mathbf{R}_{last}$ by

$$R_i(h^t) := \begin{cases} r_i(a_t) \text{ with prob. } 1 & \text{if } t = T \\ 0 & \text{otherwise,} \end{cases}$$

where $T \sim \text{Geom}(\beta)$. Here we pay agents for only the last period of the game. Having arrived at a history h^{t-1}, the payment scheme prescribes that with probability $1-\beta$ the game ends and agents get as final reward $r_i(a_t)$. With probability β the game continues at least for one more period, and whatever was played at time t does not enter into the final payment. As this payment scheme does not depend on the history h^{t-1}, the implementability test implied by Proposition 11 is satisfied.

Proposition 12. *For all $h^t \in \mathcal{H}$ and all $\mathbf{a} \in A^\infty$, we have*

$$\widehat{U}_i^{\mathbf{R}_{last}}\left(\mathbf{a} \mid h^t\right) = U_i\left(\mathbf{a} \mid h^t\right),$$

which implies that $\mathbf{SPE}(\Gamma) = \mathbf{SPE}\left(\widehat{\Gamma}\left(\mathbf{R}_{last}\right)\right)$. *Consequently,* $\mathbf{R}_{last}$ *implements* Γ.

Proof. The proof follows from the fact that, for any history h^t and any sequence $\mathbf{a} \mid h^t$ we have that

$$\widehat{U}_i^{\mathbf{R}_{last}}\left(\mathbf{a} \mid h^t\right) = \mathbb{E}_T\left\{u_i\left[r_i(a_{t+T})\right]\right\} = \sum_{s=0}^{\infty}(1-\beta)\,\beta^t u_i\left[r_i(a_{t+s})\right] = U_i\left(\mathbf{a} \mid h^t\right)$$

because T, being geometric, has no memory. Since in both games Γ and $\widehat{\Gamma}(\mathbf{R}_{last})$ utilities are measured identically for all subgames starting at any history h^t, it follows that $\mathbf{SPE}(\Gamma) = \mathbf{SPE}\left(\widehat{\Gamma}\left(\mathbf{R}_{last}\right)\right)$ and $V(\Gamma) = V\left(\widehat{\Gamma}(\mathbf{R}_{last})\right)$. □

This argument is easily generalizable for *any* multistage game with observable actions, time separable utility and common discount factors.[5]

2.5 Payment Schemes in the Literature

We now turn to payment schemes used in the literature. A number of experiments have paid participants for their total earnings across all rounds in the game. Noting that the curvature of utility of wealth may cause the participant to deviate from the modeled behavior, experiments have turned to paying participants for a single randomly chosen round. First, in section 2.5.1 we review the payment for the randomly chosen round payment scheme all rounds scheme and analyze how it

[5] We note that the utility functions, strategy spaces, and common discount factors can all be history dependent. For instance, this payment scheme is implementing for dynamic games with capital accumulation, savings, etc., which are often of interest.

alters participant behavior. In section 2.5.2 we look at the all round payment scheme and explicitly characterize the curvature effect.

2.5.1 Payment for a Randomly Chosen Round

In this subsection we argue that the round at random payment scheme introduced in Section 2.3.2 also fails the test of Proposition 11 and does not generally implement Γ. Moreover, we characterize the behavior induced by this payment scheme. Individuals discount the future too much in any period and become asymptotically indifferent between their choices. We develop a formal measure of the distortion to be able to quantify these biases.

We begin by defining the functions

$$\eta(\beta, t) := \sum_{k=0}^{\infty} \frac{\beta^k}{t+k} \text{ and } \eta_\beta(\beta, t) := \frac{\partial \eta(\beta,t)}{\partial \beta} \tag{2.1}$$

We catalog their properties in Appendix 2.9. These are modified discount factors and, as we will show in Lemma 22, allow us to write down exactly how randomly chosen round payment can be represented as discounting with η. Specifically, utility of outcome $\mathbf{a}$ from time $t = 0$ onward can be written as

$$\widehat{U}_i^{\mathbf{R}_{rcr}}(\mathbf{a}) = (1-\beta) \sum_{t=0}^{\infty} \eta(\beta, t)\, \beta^t u_i[r_i(a_t)] \tag{2.2}$$

and at any history h^t

$$\begin{aligned} \widehat{U}_i^{\mathbf{R}_{rcr}}(\mathbf{a} \mid h^t) &= (1-\beta)\,\eta(\beta, t) \sum_{k=0}^{t} u_i[r_i(a_k)] \\ &+ \beta \left\{ (1-\beta) \sum_{s=0}^{\infty} \beta^s \eta(\beta, t+1+s)\, u_i[r_i(a_{t+s+1})] \right\}. \end{aligned} \tag{2.3}$$

From (2.2) we see that the implied discount factor is $\beta^t \eta(\beta, t)$ instead of just β^t. This means that individuals in game $\widehat{\Gamma}(\mathbf{R}_{rcr})$ discount future flows of utility too rapidly in the induced game relative to Γ. Moreover, we can see this from the fact that $\lim_{t \to \infty} \eta(\beta, t) = 0$ (see, Lemma 20) that

$$\lim_{t \to \infty} \frac{\eta(\beta, t)\, \beta^t}{\beta^t} = 0.$$

As a practical matter, we should expect agents to behave much more impatiently than in the target model. This comes from the fact that, when choosing a_1 at $t = 1$, the agent should internalize not only the fact that she should receive $(1-\beta)\, u_i[r_i(a_1)]$ utils at $t = 1$, but that this also affects the expected utility at time $t = 2$ by $(1-\beta)\, \beta \frac{1}{2} u_i[r_i(a_1)]$, at time $t = 3$ by $(1-\beta)\, \beta^2 \frac{1}{3} u_i[r_i(a_1)]$, and so on. Ultimately, this increases the weight of time $t = 1$'s decision on lifetime utility as it shows up in every subsequent utility computation. The next result illustrates the shape of the distortions caused by round at random payment.

Proposition 13 (Implementable outcomes in $\widehat{\Gamma}(\mathbf{R}_{rcr})$). *Suppose Assumption 2.3.2 holds. Under* $\mathbf{R}_{rcr}$ *a sequence* $\mathbf{a} = \{a_t\}_{t=0}^{\infty}$ *is an implementable outcome of* $\widehat{\Gamma}(\mathbf{R}_{rcr})$ *if and only if for all* $t \in \mathbb{N}$ *and all* $i \in \{1, 2, ..., n\}$,

$$a_{i,t} \in \underset{\tilde{a}_i \in A_i}{argmax}\,(1-\beta)\, u_i\left[r_i\left(\tilde{a}_i, a_{-i,t}\right)\right] + \beta W_i\left(\tilde{a}_i, a_{-i,t}\right) \tag{2.4}$$

with

$$W_i\left(\tilde{a}_i, a_{-i,t}\right) = \begin{cases} \underline{W}^i(t) & \text{if} \quad \tilde{a}_i \neq a_{i,t} \\ (1-\beta)\sum_{s=0}^{\infty}\left[\frac{\eta(\beta,t+1+s)}{\eta(\beta,t)}\right]\beta^s u_i\left[r_i\left(a_{t+s}\right)\right] & \quad\quad otherwise \end{cases},$$

where $\underline{W}_i(t)$ is the value of agent i's optimal punishment in game $\tilde{\Gamma}_t$ (as defined in Lemma 23),

$$\underline{W}_i(t) = \min\left\{v^i : \exists v^{-i} \in \mathbb{R}^{n-1} \text{ with } \left(v^i, v^{-i}\right) \in V\left(\tilde{\Gamma}_t\right)\right\}.$$

Proposition 13 illustrates the nature of the distortion over the set of implementable outcomes generated by $\mathbf{R}_{rcr}$. Condition (2.4) is nearly identical to the implementability condition of game Γ; the crucial distinction is that future rewards are further discounted by the term $\eta(\beta,t)^{-1} \cdot \eta(\beta, t+s+1)$. As η is decreasing in t, we have that $\eta(\beta,t)^{-1} \cdot \eta(\beta, t+s+1) < 1$. This immediately implies that for all s,

$$\frac{\eta(\beta, t+s+1)}{\eta(\beta,t)}\beta^s < \beta^s.$$

Consequently a participant is more impatient in $\widehat{\Gamma}(\mathbf{R}_{rcr})$ than in Γ. In addition, punishments after deviations, $\underline{W}_i(t)$, are also at least as small as the optimal punishment in game Γ because of the greater discounting which also affects the set of implementable outcomes at time t. That is, $\underline{W}_i(t) < \underline{W}_i$. Moreover, the fact that the implementability condition (2.4) depends on t also may violate the test devised in Proposition 11.

It turns out that in the long run (for t large enough), the implementability condition of sequences in Γ and $\widehat{\Gamma}(\mathbf{R}_{RCR})$ are arbitrarily close. (?) document this phenomenon in a model of rish sharing with limmited commitment. We confirm this in Lemma 21, showing that

$$\lim_{t\to\infty} \frac{\eta(\beta, t+s+1)}{\eta(\beta,t)}\beta^s = \beta^s \text{ and } \lim_{t\to\infty}(1-\beta)\sum_{s=0}^{\infty}\frac{\eta(\beta,t+s+1)}{\eta(\beta,t)}\beta^s = 1.$$

This means as players keep on playing the IC constraints of the actual game and the induced game are not very different. (?) take this as evidence that $\mathbf{R}_{rcr}$ "almost" implements Γ.

We caution that this argument does not follow, for several reasons. First, as shown in the Prisoner's Dilemma example of section 2.2, subgame perfect equilibria may exhibit path dependence. Second, the speed of convergence of the incentive compatibility constraints is rather slow, as discussed below. Third, (?) assess convergence in terms of the implicit discount rate. We argue that the correct measure of convergence ought to use the present value of fixed income streams, which we show to converge at a much slower rate. Fourth, we prove that the participants will exhibit asymptotic indifference: even though the IC constraints of the target and induced games are asymptotically similar, agents simply will not care about what happens in any continuation game if t is high enough. The reason is simple: if a participant has been playing for a long enough time, whatever she does today only negligibly affects the expected value the lottery that she faces. Moreover, as she is discounting, the effect of future payoffs on the expected value of payments is also negligible. We formalize this idea below.

2.5.1.1 Asymptotic Indifference

Definition 14 (Contribution). Let $H : A^{\infty} \to \mathbb{R}$ be some function that has the property that it can be written as

$$H(\mathbf{a}) = \sum_{s=0}^{\infty} F(s, a_s)$$

for some function $F : \mathbb{N} \times A \to \mathbb{R}$. We define the *contribution of* a_s *to* H ,[6]

$$\mathcal{C}(H \mid s)(\mathbf{a}) := F(s, a_s).$$

Likewise, let $I \subset \mathbb{N}$ index set. We define the *contribution of* $\{a_k\}_{k \in I}$ as

$$\mathcal{C}(H \mid I)(\mathbf{a}) := \sum_{k \in I} F(k, a_k).$$

Example. Take $U_t^i(\mathbf{a}) = (1-\beta)\sum_{s=t}^{\infty} \beta^{s-t} u_i[r_i(a_s)]$ as the time t utility for agent i in game Γ at all subgames that start at date t. Then

$$\mathcal{C}\left(U_t^i \mid t\right)(\mathbf{a}) = (1-\beta)\, u_i[r_i(a_t)]$$

and

$$\mathcal{C}\left(U_t^i \mid s > t\right)(\mathbf{a}) = (1-\beta) \sum_{s=t+1}^{\infty} \beta^{s-t} u_i[r_i(a_s)].$$

For the particular case of a stationary path (i.e. $u_i[r_i(a_s)] = \hat{u}_i$ for all s) then we can simplify the above as

$$\mathcal{C}\left(U_t^i \mid s > t\right)(\mathbf{a}) = \beta \hat{u}_i$$

Note that both are time independent and non-negligible.

Proposition 15 (Asymptotic Indifference). *Let* $\bar{u}^i = \max_{a \in A} u_i[r_i(a)]$. *For all histories* h^{∞} *and*

[6]The contribution of a_t to H is simply the Radón-Nikodym derivative of H with respect to the unit measure.

all $a \in A$,

$$\mathcal{C}\left(U_t^i \mid t\right)(\mathbf{a}) \leq \bar{u}^i (1-\beta)\,\eta(\beta,t) = \bar{u}^i \frac{1}{t} + o\left(\frac{1}{t}\right)$$

and for all histories h^∞,

$$\mathcal{C}\left(U_t^i \mid s > t\right)(\mathbf{a}) \leq \bar{u}^i (1-\beta)\,\beta\eta_\beta(\beta,t) = \bar{u}^i \frac{\beta^2}{t(1-\beta)} + o\left(\frac{1}{t}\right)$$

so both expressions converge to 0 as $t \to \infty$

2.5.1.2 Measuring Distortions in Implementability and Payoffs

Equation (2.4) allows us to compare the implementability constraints quite easily. In each of the expressions, the present is evaluated in the same manner, $(1-\beta)\,u_i\left[r_i(a_t)\right]$, and the only differences come from discounting future payoffs, which is time dependent.

We now develop a measure of the distortion. Suppose we consider a constant outcome $a_s = a$ for all $s \geq t$ at time t, which generates a constant stream utility. The theoretical expected present value from any history h^t onwards, which we will denote by W_t, is

$$W_t = (1-\beta)\sum_{s=0}^{\infty} \beta^s u = u.$$

On the other hand, the when we do this computation in the game $\widehat{\Gamma}\left(\boldsymbol{R}_{RCR}\right)$, we have from the incentive compatibility (IC) constraint of Proposition (13),

$$\widehat{W}_t = (1-\beta)\sum_{s=0}^{\infty}\left[\frac{\eta(\beta,t+s+1)}{\eta(\beta,t)}\right]\beta^s u = (1-\beta)\,\frac{\eta_\beta(\beta,t)}{\eta(\beta,t)}u.$$

Then, for any utility level u, we can define the *ratio of present values* ρ_t as

$$\rho_t := \frac{\widehat{W}_t}{W_t} = (1-\beta)\,\frac{\eta_\beta(\beta,t)}{\eta(\beta,t)}. \tag{2.5}$$

We show that $\rho_t \to 1$ and $\rho_t < 1$ for all t, since agents behave as if they discounted the future more than they actually do. With (2.5) we are equipped with an explicit measure of how bad the problem is.[7] This is a measure of the distortion in the implementability condition.

In Figure 2-1 we explore the behavior of this ratio for different values of β, as t grows.

[7] We can calculate the functions η and η_β accurately using the finite integral formulation in 3 and 4 in Lemma 20.

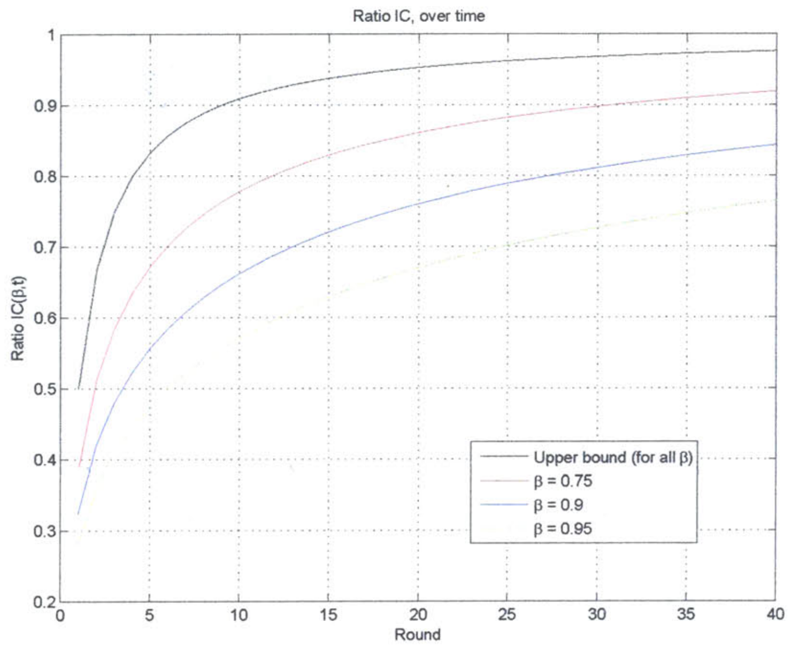


Figure 2-1: Presents the IC Ratio vs Time

We note that for $\beta = 0.9$, at round 1 the ratio is less than 0.3, which implies that agents valuate relative future relative utility streams at 30% the value in the target game, which gives a sizable measure of the distortion of incentives in the implementability condition (2.4). Even by round 10 this distortiuon is on the order of 60%. In addition the figure displays a uniform bound across all discount factors. To demonstrate how the slow convergence relates to the asymptotic indifference, we introduce the relative contribution at time t as

$$\text{Relative contribution at } t := \frac{\mathcal{C}\left(U_t^i \mid s \geq t\right)}{\mathcal{C}\left(U_t^i \mid s < t\right) + \mathcal{C}\left(U_t^i \mid s \geq t\right)}.$$

This captures the share of one's utility which is comprised of today's and subsequent periods' decisions. In the target game notice that the relative contribution is identically 1 at all periods. By studying the behavior of the ratio of present values ρ_t and the relative contribution together against time, we can see that as ρ_t slowly converges to 1 (and the distortion becomes arbitrarily smaller), meanwhile the relative contribution rapidly converges to zero. We show this in Figure (2-2).

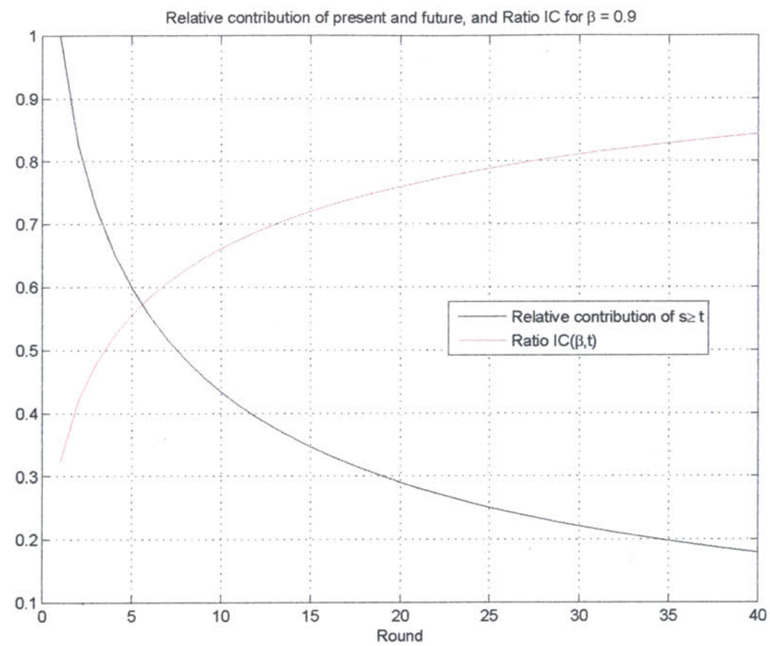


Figure 2-2: Presents the IC Ratio vs Time as well as the Relative Contribution vs Time

Note that by period 10, the relative contribution has dropped to 0.44 and the IC ratio is merely 0.67. This figure suggests that by the time that the valuation of relative future utility streams are close to the target game, agents are "almost indifferent" about the potential continuation histories they could face.

2.5.2 Payment for All Rounds

Paying individuals for all rounds may not implement the model Γ.[8] We establish that the payment schemes may significantly weaken the incentives of the participants as the number of rounds played increases.

The payment scheme $\mathbf{R}_{all}$ is given by

$$R_{i,t}(h^t) = \begin{cases} \sum_{s=0}^{s=t} r_i(a_s) \text{ with prob. } 1 & \text{if} \quad t = T \\ 0 & \quad\quad \text{otherwise} \end{cases}$$

where $T \sim \text{Geom}(\beta)$. Given an outcome $\mathbf{a} = \{a_t\}_{t=0}^{\infty}$ we define $R_{i,t}(\mathbf{a})$ as the accumulated rewards up to time t for agent i; i.e. $R_{it}(\mathbf{a}) = \sum_{s=0}^{t} r_i(a_s)$ and let $R_t(\mathbf{a}) = (R_{1,t}(\mathbf{a}), R_{2,t}(\mathbf{a}), ..., R_{n,t}(\mathbf{a}))$ the vector of all accumulated rewards up to time t. The following proposition gives the obvious characterization of the set of implementable outcomes of $\widehat{\Gamma}(\mathbf{R}_{all})$.

Proposition 16. *Under Assumption 2.3.2, a sequence* $\mathbf{a}$ *is implementable if and only if for all* i *and all* t

[8] This has been noted in (**?**) and (**?**), among others.

$$a_{i,t} \in \underset{\tilde{a}_i \in A_i}{argmax} (1-\beta)\, u_i \left[R_{i,t-1}(\mathbf{a}) + r_i(\tilde{a}_i, a_{-i,t})\right] + \beta W_i(\tilde{a}_i, a_{-i,t}) \tag{2.6}$$

where

$$W_i(\tilde{a}_i, a_{-i,t}) = \begin{cases} \underline{W}_i(R_{i,t}(\mathbf{a})) & \text{if } \tilde{a}_i \neq a_{i,t} \\ (1-\beta)\sum_{s=0}^{\infty} \beta^s u_i(R_{i,t+s}(\mathbf{a})) & \text{otherwise} \end{cases}$$

and the function $\underline{W}_i(R_{i,t})$ is the value of the optimal punishment for agent i as a function of i's accumulated rewards.

Proposition 16 shows why there may be a distortion due to $\mathbf{R}_{all}$. Since agents receive the payment for the experiment only when the experiment ends, the amount earned up to time t generates a stock of earnings not yet consumed. If there is some curvature in the utility function u_i, then the stock of unconsumed earnings may affect incentives of agent i in all subsequent rounds. In particular, if utility over monetary rewards were concave, we should expect to see a diminishing marginal utility of wealth as t increases, which would weaken incentives in the long run. We mention that if utility was linear in earnings, this payment scheme would not cause problems as past earnings would drop from condition (2.6).

We formalize these intuitions in Proposition 17. Let

$$\overline{r}_i = \max_{a \in A} r_i(a) \text{ and } \underline{r}_i = \min_{a \in A} r_i(a)$$

be the best and worst possible stage rewards for agent i, and suppose that $\underline{r}_i \geq 0$.

Proposition 17. *Suppose that $u_i(\cdot)$ is an increasing, concave and differentiable function.*

1. *The range of values for contemporaneous and continuation utilities is decreasing over time. Specifically, for any history h^t and any pair of continuation sequences $\mathbf{a}, \mathbf{a}' \in A^\infty$ we have that*

$$\left| \widehat{U}_i^{\mathbf{R}_{all}}(\mathbf{a} \mid h^t) - \widehat{U}_i^{\mathbf{R}_{all}}(\mathbf{a}' \mid h^t) \right| \leq u_i'\left[(t+1)\underline{r}_i\right](\overline{r}_i - \underline{r}_i).$$

   If, in addition, u_i satisfies the Inada condition $u_i'(\infty) = 0$ and $\underline{r}_i > 0$, then as $t \to \infty$

$$\sup_{\mathbf{a}, \mathbf{a}' \in A^\infty} \left| \widehat{U}_i^{\mathbf{R}_{all}}(\mathbf{a} \mid h^t) - \widehat{U}_i^{\mathbf{R}_{all}}(\mathbf{a}' \mid h^t) \right| \to 0.$$

2. *If u is linear (so agents are risk-neutral) then $\mathbf{R}_{all}$ implements Γ.*

Proposition 17 illustrates the nature of the distortion caused by $\mathbf{R}_{all}$. As time passes the amount by which an agent's utility changes must be decreasing. The payment scheme is only implementing if the participants are modeled as risk-neutral.

This highlights a natural tension that arises in certain situations. If, for instance, one is interested in studying high-stakes infinitely repeated interaction, e.g. risk-sharing in the vein of (?) or (?), there is a tension between having the game have high-stakes payoff to generate realistic behavior and meanwhile not being able to pay for either all rounds or random rounds due to the aforementioned biases induced.

2.6 A Model of Savings

An agent has an initial endowment of assets, $a_0 = 0$. At each t, the agent receives a deterministic endowment of $y_t \geq 0$ units and can save any amount at a constant interest rate $R > 0$. The budget constraint and no-Ponzi conditions are

$$c_t + a_{t+1} = y_t + Ra_t \text{ for all } t \in \mathbb{N} \text{ and } \lim_{t \to \infty} R^{-t} a_t = 0.$$

The agent has preferences given by $U = (1-\beta) \sum_{t=0}^{\infty} \beta^t u(c_t)$ where $u' > 0$ and $u'' < 0$. Arrow's time-zero constraint is

$$\sum_{t=0}^{\infty} R^{1-t} c_t = \sum_{t=0}^{\infty} R^{1-t} y_t \tag{2.7}$$

The usual Euler equation is $u'(c_t) = \beta R u'(c_{t+1})$. In particular $\beta R = 1$ yields $c_t = c_{t+1} = c^*$. The consequence of this, of course, is Friedman's permanent income hypothesis which follows by (2.7), with

$$\sum_{t=1}^{\infty} \beta^{t-1} c^* = \sum_{t=1}^{\infty} \beta^{t-1} y_t \Longleftrightarrow c^* = (1-\beta) \sum_{t=1}^{\infty} \beta^{t-1} y_t.$$

Note that this is independent of the specific preferences we considered.

2.6.1 Round at Random Payment

Suppose that we pay the agent for a randomly chosen round in the above environment. It can be shown, using the recursive method shown in Appendix 2.9.2, that the modified Euler equation is

$$u'(c_t^{rar}) = \underbrace{\frac{\eta(\beta, t+1)}{\eta(\beta, t)}}_{<1} \beta R_t u'(c_{t+1}^{rar}) < \beta R u'(c_{t+1}^{rar})$$

which clearly distorts the natural euler equation from the original model. . Instead of consuming a constant amount, the agent would choose a forever decreasing consumption bundle. To further our intuition, define $\hat{R}_t = \eta(\beta, t)^{-1} \cdot \eta(\beta, t+1) R_t$ as the effective gross interest rate, as the Euler equation is $u'(c_t) = \beta \hat{R}_t u(c_{t+1})$. Observe $\hat{R}_t < R$ for all periods; agents will save less under round at random payment than under the theoretical model. The following proposition shows the extent of this distortion.

Proposition 18. *Let $u(\cdot)$be a strictly concave and differentiable utility function. Then*

$$\frac{u'(c_t^*)}{u'(c_t^{rar})} \propto \eta(\beta, t) \to 0 \text{ as } t \to \infty. \tag{2.8}$$

To illustrate the previous proposition, assume we have a CES utility function: $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. Using the proposition we can show that

$$\frac{c_t^{rar}}{c_t^*} \propto [\eta(\beta, t)]^{\frac{1}{\sigma}} = O\left(t^{-\frac{1}{\sigma}}\right).$$

This implies that consumption under the randomly chosen round payment scheme is infinitely smaller than the theoretical predicted consumption decision when $t \to \infty$. Suppose now that $\beta R = 1$, so that theoretical consumption is constant (i.e. $c_t^* = c^*$for all t) and moreover suppose $\sigma = 1$ (so $u(c) = \log(c)$). Then,

$$c_t = \frac{\eta(\beta, t)}{\eta(\beta, 1)} c_1 \propto \eta(\beta, t).$$

so $c_t \to 0$. Even if the Euler equation does converge to the one of the theoretical model (since $\eta(\beta, t)^{-1} \cdot \eta(\beta, t+1) \to 1$ as $t \to \infty$) the behavior of the solution does not approximate the one in the theoretical model. In particular, the solution at very large t does resemble a constant consumption, but the wrong constant. Instead of $c_t = c^*$, it will become arbitrarily close to zero. Normalizing $c^* = 1$ we compare the consumption sequences under the two models in Figure 2.6.1, under different values of β. In early rounds agents consume more than prescribed by the theoretical model. In particular, for $\beta = 0.9$ consumption at $t = 1$ is 8.26 times bigger than consumption in the first period and at $t = 2$ is 5.6 times bigger.

2.6.2 Payments for All Rounds

We now consider the case where agents are paid for all rounds. In Appendix 2.9.3 we show that the modified Euler equation is now

$$u'(c_t^{all}) = \beta(R_t - 1) u'(c_t^{all}) < \beta R_t u'(c_{t+1}^{all}) \tag{2.9}$$

Again this has the effect of reducing incentives for saving, since the effective gross interest rate for the agent is $\hat{R}_t := R_t - 1$. The extent of the distortion is illustrated by the following proposition.

Proposition 19. *Let $u(\cdot)$ be a strictly concave and differentiable utility function. Then*

$$\frac{u'(c_t^*)}{u'(c_t^{all})} \propto \left(1 + \frac{1}{r}\right)^{-t}. \tag{2.10}$$

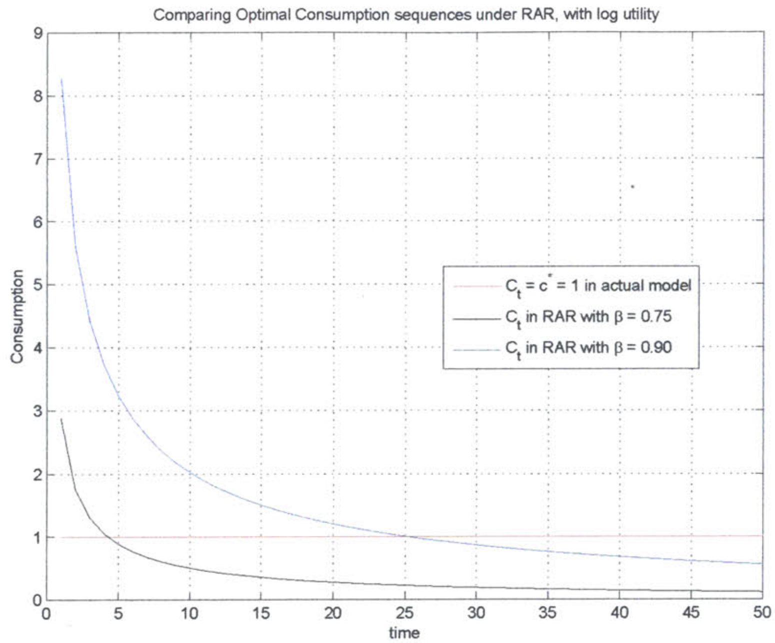


Figure 2-3: Consumption Sequences with log utility for RAR.

To illustrate this proposition, we return to the CES example. Notice

$$\frac{c_t^{all}}{c_t^*} \propto \left(\frac{R-1}{R}\right)^{\frac{t-1}{\sigma}} \to 0 \text{ as } t \to \infty,$$

so no matter how small σ is the ratio between consumptions goes to zero. In particular, consider the case where $1 < R < 1 + \frac{1}{\beta}$. This implies that $\beta R > 1$ and $\beta (R - 1) < 1$. The above equations imply that for any $\sigma > 0$ we have that

$$c_t^{all} \propto [\beta (R-1)]^{\frac{t-1}{\sigma}} \to 0 \text{ since } \beta (R-1) < 1$$

but

$$c_t^* \propto (\beta R)^{\frac{t-1}{\sigma}} \to \infty \text{ since } \beta R > 1.$$

Therefore, we have shown not only that the ratio of consumption goes to zero but also that the behavior of the optimal solution path is extremely different. Furthermore, notice that this is true no matter how much curvature (i.e. $\sigma > 0$) we assume. We highlight this point because we know that when utility is linear (i.e. $\sigma = 0$) payment in all rounds does implement the actual game. However, this result is largely non-generic: allowing for arbitrarily small amounts of curvature (in the CES family) implementable outcomes are starkly different. Moreover, if the researcher is interested in larger-stakes games (see, e.g., Gneezy et al.), it may very well be that curvature enters decisions thereby invalidating all round payment as an alternative to random round payment.

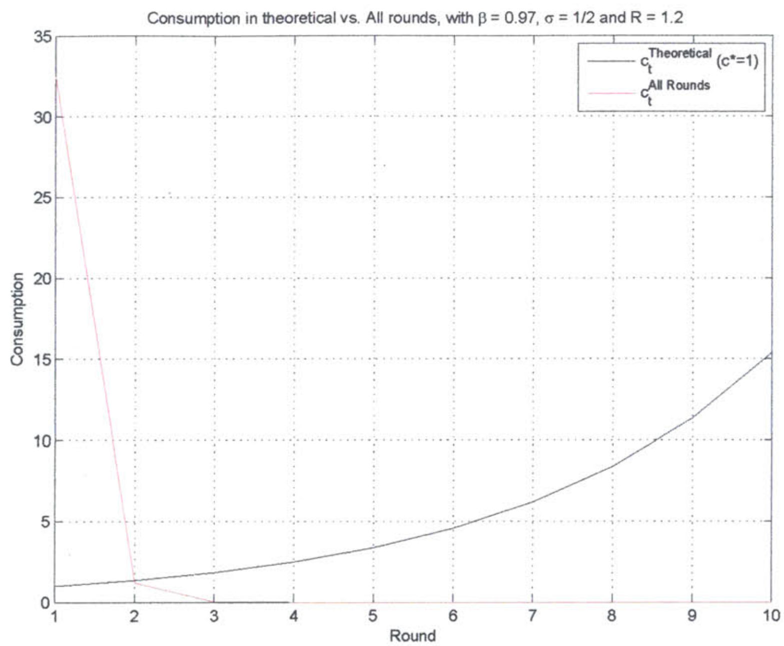


Figure 2-4: Consumption in theoretical model versus all round payment

To illustrate numerically this feature, we set $\beta = 0.97$, $\sigma = \frac{1}{2}$ and $R = 1.2$.[9] Normalizing the net present value of endowments so that $c_1^* = 1$, we present the results in Figure 2.6.2.

For $\beta R = 1$ we have that $c_t^{all} > c_{t+1}^{all}$, so consumption is decreasing even when c_t should be constant, similar to what we observed under the RCR payment scheme. Using log-utility, it is easy to see that

$$\frac{c_{t+1}^{all}}{c_t^{all}} = \beta (R-1) = 1 - \beta \text{ if } R\beta = 1.$$

For β close to one, this is a rapidly decaying function. Figure 2.6.2 presents a numerical illustration. Intuitively, an agent who is very patient (thereby facing a low interest rate) can essentially consume as much as possible in the first period, by borrowing against the flow of future income. Since all future streams of consumption until the end of the game is kept by the agent, the she wants to maximize the amount of consumption early on.

Let $c^* = 1$ and observe that

$$c_1 = \sum_{t=1}^{\infty} \beta^{t-1} y_t = \frac{1}{1-\beta} c^* = \frac{1}{1-\beta} \to \infty \text{ as } \beta \to 1$$

Turning to Figure 2.6.2, when $\beta = 0.9$ the agent consumes $c_1 = 9.1$ while the upper bound of consumption is $\frac{1}{1-0.9} = 10$. This has to be followed by a steep reduction in consumption, as seen

[9]Smaller values of R or σ make results much starker.

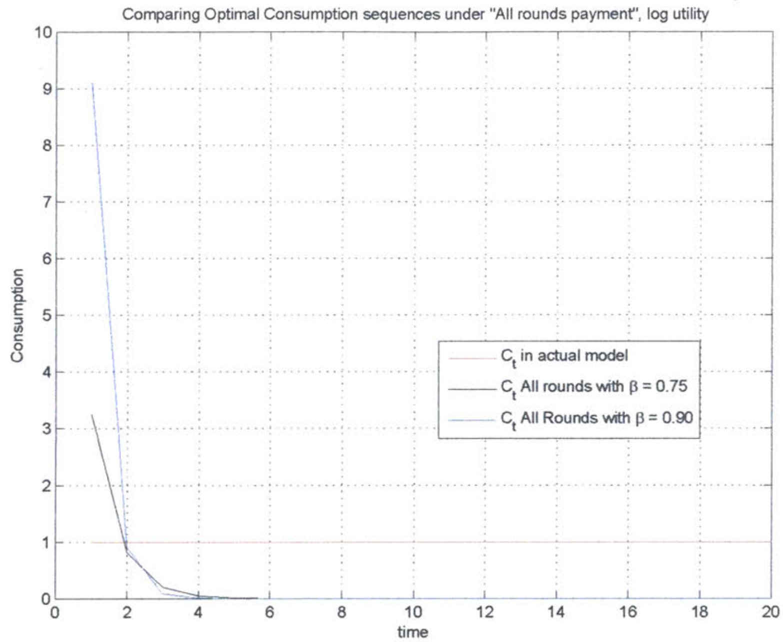


Figure 2-5: Consumption Sequences with log utility for All Round Payment.

in the figure, in order to be able to repay it. When $\beta = 0.99$ the participant gets even closer: consumption in the first period is $c_1 = 99.01$ while $\frac{1}{1-0.99} = 100$.

2.7 Conclusion

We have discussed payment schemes in infinitely repeated games with discounting. When the researcher is interested in implementing a particular game, the payment schemes often used in the literature fail. In particular, we have characterized how the implementability conditions are changed by the payment schemes. We add that since our characterizations are analytically explicit, one can in principle use the modified implementability conditions to then analyze the behavior in data generated from these payment schemes. In addition, we have found a simple payment scheme which implements a much broader class of games by simply paying participants for the last round of play.

2.8 Appendix A - Proofs

Proof of Proposition 11. Under Assumption 2.3.2, it is a direct corollary of (?) (see also Corollary 2.6.1 of (?)) which says that the set of implementable outcomes is history independent; i.e. $\mathbf{O}\left(\Gamma(h^t)\right) = \mathbf{O}\left(\Gamma\right)$ for all h^t as we can always think of the game as being reset. This, together with implementation, delivers the result. $\square$

Proof of Proposition 17. Observe that

$$\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}\mid h^t\right)-\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}'\mid h^t\right)=(1-\beta)\sum_{s=0}^{\infty}\beta^s\left[u_i\left(R_{i,t}+\sum_{k=1}^{s}r_{t+k}\right)-u_i\left(R_{i,t}+\sum_{k=1}^{s}r'_{t+k}\right)\right]$$

where $r_{i,t+k}=r_i\left(a_{t+k}\right)$ and likewise for $r'_{i,t+k}$. By concavity and differentiability of $u_i(.)$, we know that for all $x,y\in\mathbb{R}$ we have that

$$u'(x)(x-y)\leq u(x)-u(y)\leq u'(y)(x-y).$$

Using $x=R_{i,t}+\sum_{k=1}^{s}r_{t+k}$ and $y=R_{i,t}+\sum_{k=1}^{s}r'_{t+k}$,

$$\begin{aligned}(1-\beta)\sum_{s=0}^{\infty}\beta^s u_i'\left(R_{i,t}+\sum_{k=1}^{s}r_{i,t+k}\right)\left(\sum_{k=1}^{s}\left(r_{i,t+k}-r'_{i,t+k}\right)\right)&\leq\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}\mid h^t\right)-\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}'\mid h^t\right)\\&\leq(1-\beta)\sum_{s=0}^{\infty}\beta^s u_i'\left(R_{i,t}+\sum_{k=1}^{s}r'_{i,t+k}\right)\left(\sum_{k=1}^{s}\left(r_{i,t+k}-r'_{i,t+k}\right)\right)\end{aligned}$$

As u_i is concave, $u_i'\left(R_{i,t}+\sum_{k=1}^{s}r_{i,t+k}\right)\leq u'(R)$ and the same for $r'_{i,t+k}$, so

$$\left|\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}\mid h^t\right)-\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}'\mid h^t\right)\right|\leq u'\left(R_{i,t}\right)(1-\beta)\sum_{s=0}^{\infty}\beta^s\left|r_{i,t+k}-r'_{i,t+k}\right|$$

If $\underline{r}_i>0$, then $R_{i,t}\geq(t+1)\underline{r}_i$, implying that

$$\left|\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}\mid h^t\right)-\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}'\mid h^t\right)\right|<u_i'\left[(t+1)\underline{r}_i\right]\left[\left|(1-\beta)\sum_{s=0}^{\infty}\beta^s\left|\overline{r}_i-\underline{r}_i\right|\right|\right]\leq u_i'\left[(t+1)\underline{r}_i\right]\left(\overline{r}_i-\underline{r}_i\right).$$

as we wanted to show. To show the contemporary utility result, without loss of generality take $r,\widehat{r}\in r_i(A)$, such that $r>\widehat{r}$. Then,

$$\begin{aligned}u_i'\left(R_{i,t}+\overline{r}_i\right)(r-\widehat{r})&\leq u_i'\left(R_{i,t}+r\right)(r-\widehat{r})\leq(1-\beta)u_i\left(R_{i,t}+r\right)-(1-\beta)u_i\left(R_{i,t}+\widehat{r}\right)\\&\leq u_i'\left(R_{i,t}+\widehat{r}\right)(r-\widehat{r})\leq u_i'\left(R_{i,t}+\overline{r}_i\right)\left(\overline{r}_i-\underline{r}_i\right)\end{aligned}$$

which implies that

$$\left|u_i\left(R_{i,t}+r\right)-u_i\left(R_{i,t}+\hat{r}\right)\right|\leq u_i'\left(R_{i,t}+\underline{r}_i\right)\left(\overline{r}_i-\underline{r}_i\right)\leq u_i'\left[(t+1)\underline{r}_i\right]\left(\overline{r}_i-\underline{r}_i\right).$$

as we wanted to show. □

Proof of Proposition 13. (Sketch) The proposition is a direct consequence of Lemma 23: $\{a_t\}$ is an implementable outcome iff is the outcome of some SPE strategy σ, and such strategy has to be such that $\{a_{t+s}\}_{s=0}^{\infty} \in \mathbf{O}(\Gamma_t)$ (i.e. is a SPE outcome of the game starting at sub-tree h^t, which is identical to Γ_t, up to difference in intercepts in the utility functions), which we showed in Lemma 23, together with the fact that it is continuous at infinity, and has a compact $V(\Gamma_t)$. Then, following the same reasoning as in (?) we conclude that we can implement any outcome by using optimal simple punishments, which has payoff $\underline{W}^i(t)$. Continuation values in Γ_t are calculated using the discount function $B(\beta, t+1, s)$ for all $s \geq 1$, finishing the proof. □

Proof of Proposition 15. From Lemma 22, equation 2.14, we have that for any a_t the contribution of t period strategies is $\mathcal{C}(U_t^i \mid t)(a_t) = (1-\beta)\eta(\beta,t)\,u_i[r_i(a_t)]$ and the result follows from there. Also, Property 5 in Lemma 20 gives the approximation result (and convergence to zero). For any history a_{t+s}, using again equation 2.14, we have

$$\begin{aligned}
\mathcal{C}\left(U_t^i \mid \{s \geq t+1\}\right)\left(\{a_{t+s}\}_{s=1}^{\infty}\right) &= \beta(1-\beta)\sum_{s=0}^{\infty}\eta(\beta, t+1+s)\,u_i[r_i(a_{t+1+s})] \\
&\leq \beta\bar{u}_i(1-\beta)\sum_{s=0}^{\infty}\sum_{k=s}^{\infty}\frac{\beta^k}{t+1+k} = \beta\bar{u}_i(1-\beta)\sum_{k=0}^{\infty}\sum_{s=0}^{k}\frac{\beta^k}{t+1+k} \\
&= \beta\bar{u}_i(1-\beta)\sum_{k=0}^{\infty}\frac{(k+1)\beta^k}{t+(k+1)} = \beta\bar{u}_i(1-\beta)\left(\sum_{j=1}^{\infty}\frac{j\beta^{j-1}}{t+j}\right) \\
&= \beta\bar{u}_i(1-\beta)\left(\sum_{j=0}^{\infty}\frac{j\beta^{j-1}}{t+j} - 0\right) = \beta\bar{u}_i(1-\beta)\,\eta_\beta(\beta,t)
\end{aligned}$$

and the approximation result comes from property 6 in Lemma 20. □

Proof of 18. Using the Euler equation for both the theoretical and the RCR sequence of optimal consumption, we get

$$u'\left(\overset{*}{c_t}\right) = \frac{1}{\beta^{t-1}R^{t-1}}u'\left(\overset{*}{c_1}\right)$$

and

$$u'(c_t^{rcr}) = \frac{1}{\beta^{t-1}R^{t-1}}u'(c_1^{rcr})\frac{\eta(\beta,1)}{\eta(\beta,t)}$$

then

$$\frac{u'(c_t^*)}{u'(c_t^{rcr})} = \frac{u'(c_1^*)}{u'(c_1^{rcr})}\frac{\eta(\beta,t)}{\eta(\beta,1)} \propto \eta(\beta,t)$$

□

Proof of 19. From before we knew that

$$u'\left(c_t^*\right) = \frac{1}{\beta^{t-1}R^{t-1}}u'\left(c_1^*\right)$$

and likewise, we show that

$$u'(c_t^{all}) = \frac{1}{\beta^{t-1}(R-1)^{t-1}}u'(c_1^{all})$$

Therefore

$$\frac{u'(c_t^*)}{u'(c_t^{all})} = \left(\frac{R-1}{R}\right)^{t-1}\frac{u'(c_1^*)}{u'(c_1^{all})} \propto \left(1+\frac{1}{r}\right)^{-t}$$

□

2.9 Appendix B - Auxiliary results

2.9.1 Technical Results

Lemma 20. *Let η and ϕbe defined as in 2.1 and let $\eta_\beta\left(\beta,t\right) := \frac{\partial\eta(\beta,t)}{\partial\beta}$. Then*

1. For all $\beta \in (0,1)$ and all $t \geq 1$ we have $\eta\left(\beta,t+1\right) = \frac{1}{\beta}\left(\eta\left(\beta,t\right) - \frac{1}{t}\right)$
2. For $t \geq 1$ we can write $\eta\left(\beta,t\right)$ as $\eta\left(\beta,t\right) = \frac{1}{\beta^t}\ln\left(\frac{1}{1-\beta}\right) - \sum_{i=1}^{t-1}\frac{1}{(t-i)\beta^i}$
3. For all β, t we can write $\eta\left(\beta,t\right) = \frac{1}{\beta}\int_0^\beta\left(\frac{z}{\beta}\right)^t\frac{1}{z(1-z)}dz$
4. For all β, t we can write η_β as $\eta_\beta\left(\beta,t\right) = \sum_{k=0}^{\infty}\frac{k\beta^{k-1}}{t+k} = \int_0^\beta\left(\frac{z}{\beta}\right)^t\frac{1}{(1-z)^2}dz$
5. For givenβ: $\eta\left(\beta,t\right) = \frac{1}{t(1-\beta)} + o\left(\frac{1}{t}\right)$ and therefore $\lim_{t\to\infty}\eta\left(\beta,t\right) = 0$
6. For all β, $\eta_\beta\left(\beta,t\right) = \frac{\beta}{t(1-\beta)^2} + o\left(\frac{1}{t}\right)$, and therefore $\lim_{t\to\infty}\eta_\beta\left(\beta,t\right) = 0$
7. For all β, t, s we have that $\sum_{k=s}^{\infty}\frac{\beta^k}{t+k} = \beta^s\eta\left(\beta,t+s\right)$

Proof. We first prove 2. This proof is by induction. For $t = 1$ we have that

$$\eta\left(\beta,1\right) = \sum_{k=0}^{\infty}\frac{\beta^k}{1+k}$$

it can be shown, using integration and Abel's Theorem, that

$$\eta\left(\beta,1\right) = \frac{1}{\beta}\log\left(\frac{1}{1-\beta}\right)$$

To prove it, we need to prove the following recursion

$$\eta\left(\beta,t+1\right) = \frac{1}{\beta}\left[\eta\left(\beta,t\right) - \frac{1}{t}\right] \tag{2.11}$$

and is easy to see that η as defined in 2.1 satisfies this recursion. To show that this recursion is true, we do some algebra:

$$\eta(\beta, t+1) = \sum_{s=0}^{\infty} \frac{\beta^s}{t+(1+s)} = \sum_{j=1}^{\infty} \frac{\beta^{j-1}}{t+j} = \frac{1}{\beta}\left[\sum_{j=0}^{\infty} \frac{\beta^j}{t+j} - \frac{\beta^0}{t+0}\right] = \frac{1}{\beta}\left[\eta(\beta, t) - \frac{1}{t}\right]$$

For 3, see that

$$\sum_{k=0}^{\infty} \beta^k = \frac{1}{1-\beta} \Longrightarrow \sum_{k=0}^{\infty} \beta^{t+k-1} = \frac{\beta^{t-1}}{1-\beta} \Longrightarrow \sum_{k=0}^{\infty} \frac{\beta^{t+k}}{t+k} = \int_0^{\beta} \frac{z^{t-1}}{(1-z)} dz \Longleftrightarrow$$

$$\eta(\beta, t) = \frac{1}{\beta} \int_0^{\beta} \left(\frac{z}{\beta}\right)^t \frac{1}{1-z} dz$$

which is valid since η is a power series. For 4 we can also use this to differentiate η:

$$\eta_\beta(\beta, t) = \frac{\partial}{\partial \beta}\left(\sum_{k=0}^{\infty} \frac{\beta^k}{t+k}\right) = \sum_{k=0}^{\infty} \frac{k\beta^{k-1}}{t+k}$$

also

$$\sum_{k=0}^{\infty} k\beta^{k-1} = \frac{d\left(\frac{1}{1-\beta}\right)}{d\beta} = \frac{1}{(1-\beta)^2} \Longrightarrow \sum_{k=0}^{\infty} k\beta^{t+k-1} = \frac{\beta^t}{(1-\beta)^2} \Longleftrightarrow$$

$$\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} \frac{k\beta^{k-1}}{t+k} = \int_0^{\beta} \left(\frac{z}{\beta}\right)^t \frac{1}{(1-z)^2} dz$$

For 5, we must show that $t\eta(\beta, t) \to \frac{1}{1-\beta}$ as $t \to \infty$. We can write $t\eta(\beta, t) = \sum_{k=0}^{\infty} \frac{t}{t+k}\beta^k$. Defining the sequence of sequences $f_t(k) := \frac{t}{t+k}\beta^k$ is easy to see that $f_t \nearrow \beta^k$ pointwise. Therefore, we can use the Dominated convergence theorem to show that $\lim_{t\to\infty} t\eta(\beta, t) = \sum_{k=0}^{\infty}\left(\lim_{t\to\infty} \frac{t}{t+k}\beta^k\right) = \frac{1}{1-\beta}$. The convergence to 0 of η is straightforward and omitted.

For 6 we follow the same strategy, and note that $t\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} \frac{tk}{t+k}\beta^k$. We have that $\frac{tk}{t+k}\beta^k \nearrow k\beta^k$ pointwise, which implies that $\lim_{t\to\infty} t\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} k\beta^k = \frac{\beta}{(1-\beta)^2}$

Finally, for 7 see that $\phi(\beta, t, s) = \sum_{k=s}^{\infty} \frac{\beta^k}{t+k} = \sum_{j=0}^{\infty} \frac{\beta^{s+j}}{t+s+j} = \beta^s \eta(\beta, t+s)$ □

Lemma 21. *Define the function $B(\beta, t, s)$, as*

$$B(\beta, t, s) = \beta^s \frac{\eta(\beta, t+1+s)}{\eta(\beta, t)}. \tag{2.12}$$

Then, the following hold:

1. *$B(\beta, t, s) < \beta^s$ for all $t, s \in \mathbb{N}$.*

2. *$B(\beta, t, s)$ is increasing in t and decreasing in s.*

3. $(1-\beta)\sum_{s=0}^{\infty} B(\beta,t,s) = (1-\beta)\frac{\eta_\beta(\beta,t)}{\eta(\beta,t)} \to 1$ *as* $t\to\infty$.

4. $\lim_{s\to\infty} B(\beta,t,s) = 0$ *for all* $t\in\mathbb{N}$.

5. $\lim_{t\to\infty} B(\beta,t,s) = \beta^s$*, so*

$$B(\beta,t,s) \nearrow \beta^s \text{ for all } s, \text{ as } t\to\infty. \tag{2.13}$$

Proof. (1) is obvious, since $\eta(\beta,t)$is decreasing in t.We first show (5). See that we can write

$$B(\beta,t,s) = \frac{\sum_{k=s}^{\infty}\frac{\beta^k}{t+1+k}}{\sum_{k=0}^{\infty}\frac{\beta^k}{t+k}} = \frac{\sum_{k=s}^{\infty}\left(\frac{t}{t+1+k}\right)\beta^k}{\sum_{k=0}^{\infty}\left(\frac{t}{t+k}\right)\beta^k}$$

so

$$\lim_{t\to\infty} B(\beta,t,s) \underset{(i)}{=} \frac{\sum_{k=s}^{\infty}\left(\lim_{t\to\infty}\frac{t}{t+1+k}\beta^k\right)}{\sum_{k=0}^{\infty}\left(\lim_{t\to\infty}\frac{t}{t+k}\beta^k\right)} = \frac{\sum_{k=s}^{\infty}\beta^k}{\sum_{k=0}^{\infty}\beta^k} =$$

$$= \frac{\left(\frac{\beta^s}{1-\beta}\right)}{\left(\frac{1}{1-\beta}\right)} = \beta^s$$

Using in (i) the Uniform Convergence theorem (the summand sequences are monotone decreasing in k). Moreover, is easy to show (with some tedious algebra) that B is decreasing in t and increasing in s (proving (2)). Facts (2) with (5) implies (1). That $B(\beta,t,s)\to 0$ as $s\to\infty$ follows directly from the fact that $\sum_{k=s}^{\infty}\frac{1}{t+k}\beta^k \to 0$ as $s\to\infty$. Finally,

$$\sum_{s=0}^{\infty} B(\beta,t,s) = \sum_{s=0}^{\infty}\frac{\sum_{k=s}^{\infty}\left(\frac{1}{t+1+k}\right)\beta^k}{\eta(\beta,t)} = \frac{1}{\eta(\beta,t)}\sum_{s=0}^{\infty}\sum_{k=s}^{\infty}\left(\frac{1}{t+1+k}\right)\beta^k =$$

$$\frac{1}{\eta(\beta,t)}\sum_{k=0}^{\infty}\sum_{s=0}^{k}\left(\frac{1}{t+1+k}\right)\beta^k = \frac{1}{\eta(\beta,t)}\sum_{k=0}^{\infty}\frac{(k+1)\beta^k}{t+1+k} = \frac{\eta_\beta(\beta,t)}{\eta(\beta,t)}$$

□

Lemma 22. *After history h^t, time t, the utility of agent i of outcome $\mathbf{a}=\{a_t\}_{t=0}^{\infty}$ in $\hat{\Gamma}(\boldsymbol{R}_{RAR})$ can be written as*

$$\hat{U}_i\left(\mathbf{a}\mid h^t\right) = (1-\beta)\,\eta(\beta,t)\sum_{k=0}^{t} u_i\left[r_i(a_k)\right] + \beta\left\{(1-\beta)\sum_{s=0}^{\infty}\beta^s\eta(\beta,t+1+s)\,u_i\left[r_i(a_{t+s+1})\right]\right\} \tag{2.14}$$

Proof. We can decompose the utility as

$$U_t^i = (1-\beta)\frac{1}{t}\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right] + \beta W^i \tag{2.15}$$

, where $W^i$is the discounted present value of future periods payoffs. We can calculate it as

$$W^i = (1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\left(\sum_{k=0}^{t}u_i\left[r_i\left(a_k\right)\right] + \sum_{k=0}^{j}u^i\left[r_i\left(a_{t+k+1}\right)\right]\right) =$$

$$(1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\sum_{k=0}^{t}u_i\left[r_i\left(a_t\right)\right] + (1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\sum_{k=0}^{j}u_i\left[r_i\left(a_{t+k+1}\right)\right] =$$

$$(1-\beta)\,\eta\left(\beta,t+1\right)\sum_{k=0}^{t}u_i\left[r_i\left(a_t\right)\right] + (1-\beta)\sum_{j=1}^{\infty}\sum_{k=0}^{j}\frac{\beta^j}{t+1+j}u_i\left[r_i\left(a_{t+k+1}\right)\right] =$$

$$(1-\beta)\,\eta\left(\beta,t+1\right)\left(\sum_{k=0}^{t}u_i\left[r_i\left(a_t\right)\right]\right) + (1-\beta)\sum_{k=0}^{\infty}u_i\left[r_i\left(a_{t+k+1}\right)\right]\left(\sum_{j=k}^{\infty}\frac{\beta^j}{t+1+j}\right) \iff$$

$$W^i = (1-\beta)\,\eta\left(\beta,t+1\right)\left(\sum_{k=0}^{t}u_i\left[r_i\left(a_t\right)\right]\right) + (1-\beta)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left[r_i\left(a_{t+s+1}\right)\right] \tag{2.16}$$

Therefore, putting together equations 2.15 and 2.16 we get

$$U_t^i = (1-\beta)\left[\frac{1}{t}+\beta\eta\left(\beta,t+1\right)\right]\left(\sum_{k=0}^{t}u_i\left[r_i\left(a_t\right)\right]\right) + \beta\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left[r_i\left(a_{t+s+1}\right)\right]$$

and using the fact that $\eta\left(\beta,t+1\right) = \frac{1}{\beta}\left[\eta\left(\beta,t\right)-\frac{1}{t}\right] \iff \frac{1}{t}+\beta\eta\left(\beta,t+1\right) = \eta\left(\beta,t\right)$, we then get that

$$U_t^i = (1-\beta)\,\eta\left(\beta,t\right)\sum_{k=0}^{t}u_i\left[r_i\left(a_t\right)\right] + \beta\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left[r_i\left(a_{t+s+1}\right)\right]$$

as we wanted to show. □

Lemma 23. *Define the repeated game* $\Gamma_t = \left\{\left\{A,\tilde{U}_t^i\right\}_{i=0}^{n}\right\}$ *which is identical to* Γ *in every aspect, only that sequences* $\mathbf{a} = \{a_s\}_{s=0}^{\infty}$ *are valuated as*

$$\tilde{U}_t^i\left(\mathbf{a}\right) = (1-\beta)\,u_i\left(a_0\right) + (1-\beta)\sum_{s=1}^{\infty}\left[\frac{\eta\left(\beta,t+1+s\right)}{\eta\left(\beta,t\right)}\right]\beta^s u_i\left[r_i\left(a_s\right)\right] \tag{2.17}$$

Then, under assumptions 2.3.2, 2.3.2 and 2.3.2, game Γ_t *is continuous at infinity (so the single deviation principle applies),* $\mathbf{SPE}\left(\Gamma_t\right) \neq \emptyset$*, the value set* $V(\Gamma_t)$ *is compact, and the following rule* *holds*

$$\bar{a} = \{a_s\}_{s=0}^{\infty} \in \mathbf{O}(\Gamma_t) \Longleftrightarrow \exists a^{t-1} : (a^{t-1}\bar{a}) \in \mathbf{O}(\Gamma) \tag{2.18}$$

Proof. Tychonoff's theorem implies that the set $A^{\infty}$is a compact set in the product topology. The functions $(1-\beta)\frac{\eta(\beta,t+s+1)}{\eta(\beta,t)}\beta^s u_i[r_i(.)]$is continuous, which implies that $\tilde{U}_t^i$ is continuous on A^{∞}in the product topology. Using these two facts, we replicate the proof of Proposition 2 in (?) to show that $V(\Gamma_t)$ is compact. Condition 2.18 follows from Proposition 13 and the fact that Γ_t is continuous at infinity. □

2.9.2 Recursive Method for RCR payment

The typical dynamic programming program involves solving

$$V(x_0) = \sup_{\{x_{t+1}\}_{t=0}^{\infty}} (1-\beta)\sum_{t=0}^{\infty}\beta^t F(x_t, x_{t+1})$$

such that

$$\begin{cases} x_{t+1} \in G(x_t) & \forall t \in \mathbb{N} \\ x_0 & \text{given.} \end{cases}$$

The usual Bellman equation is

$$V(x) = \sup_{x' \in G(x)} (1-\beta)F(x,x') + \beta V(x').$$

If we allow for some random variable z and x' to be a function of z,

$$V(x,z) = \sup_{x' \in G(x,z)} (1-\beta)F(x,x',z) + \beta\mathbb{E}_{z'}\{V(x',z') \mid z\}.$$

Observe that, in contrast, when paying for a round at random, the problem is

$$V(x_0) = \sup_{\{x_{t+1}\}_{t=0}^{\infty}} (1-\beta)\sum_{t=0}^{\infty}\beta^t \eta(\beta,t+1)F(x_t,x_{t+1})$$

such that

$$\begin{cases} x_{t+1} \in G(x_t) & \forall t \in \mathbb{N} \\ x_0 & \text{given.} \end{cases}$$

We will define $H(x_t, x_{t+1}, \eta_t) = \eta_t F(x_t, x_{t+1})$. In addition, given β, $\eta(\beta,t)$ is a strictly decreasing function of t. Therefore, let $T(\eta,\beta)$ be the inverse.[10] Using (2.11) and augmenting the state space

[10] $T(\eta,\beta)$ satisfies $\eta(\beta, T(\eta,\beta)) = \eta$.

with η, which has a known law of motion, yields

$$V(x_0, \eta_0) = \sup_{\{x_t, \eta_t\}_{t=0}^{\infty}} (1-\beta) \sum_{t=0}^{\infty} \beta^t H(x_t, x_{t+1}, \eta_t)$$

such that

$$\begin{cases} x_{t+1} \in G(x_t) & \forall t \in \mathbb{N} \\ \eta_{t+1} = \frac{1}{\beta}\left(\eta_t - \frac{1}{T(\eta_t, \beta)}\right) & \forall t \in \mathbb{N} \\ \eta_0 = \frac{1}{\beta} \ln\left(\frac{1}{1-\beta}\right) \text{ and } x_0 & \text{given.} \end{cases}$$

The Bellman equation for this problem is simply

$$V(x, \eta) = \sup_{x' \in G(x)} (1-\beta)\, \eta F(x, x') + \beta V(x', \eta')$$

such that

$$\eta' = \frac{1}{\beta}\left(\eta - \frac{1}{T(\eta, \beta)}\right).$$

2.9.3 Recursive Method for All Round Payment

We want to get the euler equation in 2.9. We can characterize the optimal allocation by means of the following bellman equation:

$$V(a, S, y) = \max_{a'} (1-\beta)\, u(S + y + Ra - a') + \beta V(a', S + y + Ra - a', y')$$

where $S = \sum_{s=0}^{t} c_s$ is the sum of consumptions the agent would be paid if the game ended today. Assuming that the optimum is in the interior, we have

$$(1-\beta)\, u'(S + y + Ra - a') = \beta \left\{ V_a(a', S + y + Ra - a', y') - V_S(a', S + y + Ra - a', y') \right\} \tag{2.19}$$

and by the envelope conditions we have

$$V_a(a, S, y) = R(1-\beta)\, u'(S + y + Ra - a') \tag{2.20}$$

$$V_S(a, S, y) = (1-\beta)\, u'(S + y + Ra - a'). \tag{2.21}$$

Using 2.19 and substituting for c_t and c_{t+1}, we then get that

$$u'(c_t) = \beta (R-1)\, u'(c_{t+1})$$

as we wanted to show.

ff

Chapter 3

Credible reforms: a robust mechanism design approach

3.1 Introduction

In the mind of policy makers, a reputation for credibility is a delicate and hard-won situation. Policy shifts are discussed with great care and concerns regarding how the public will react. By contrast, formal models of reputation employing insights from repeated games typically assume a perfect degree of certainty and coordination. The purpose of this paper is to build on this literature to model reputation in a way that reflects the uncertainty faced by policy makers.

Following the seminal papers of (Kydland and Prescott 1977) and (Barro and Gordon 1983), macroeconomic theory has extensively studied the so-called time inconsistency problem of government policy. In essence, all time inconsistency problems consist of an authority who needs some agents in the economy (e.g., consumers, financial sector) to trust her with a decision that will be taken on their behalf. In the canonical example of monetary policy, a policy maker wants the public to trust her announcements of inflation policy. The fundamental problem resides in the fact that even if the decision maker is allowed to announce what policy she plans to take, the agent's decision to trust decision ultimately depends on their *perception* or beliefs about the (ex-post) incentives the policy maker will face *after* they trust her. This creates a wedge between the ideal policies the authority would want to implement and the ones that she can credibly promise. In the context of our inflation example, since agents know that the government will have ex-post incentives to boost employment by increasing inflation and reducing real wages, this will result in an inefficiently high equilibrium inflation.

The literature has dealt with this problem in two ways. First, some authors have argued that the government should *forfeit some flexibility* through a formal arrangement (e.g., inflation targeting/caps and tax restrictions). Second, others have argued that the government should *modify its incentives*, either by having reputational concerns through repeated interactions ((Stokey 1989;

Stokey 1991) and (Chari and Kehoe 1990; Chari and Kehoe 1993a; Chari and Kehoe 1993b)) or by delegating the decision to an agency with different incentives that will limit its time inconsistency bias. Examples of such delegation include appointing a conservative central banker ((Rogoff 1985)) or making the monetary authorities subject to a formal or informal incentive contract ((Lohmann 1992), (Persson and Tabellini 1993), (Walsh 1995)). If full commitment to contingent policies is not available and flexibility is socially desirable, these "*incentive reforms*" may become a desirable solution. The key difference between both approaches is that policies that are enforced by incentive reforms are very sensitive to the assumption that the public knows exactly what the reformed incentives are. If the public believed that with a high enough probability, the government still has a time inconsistency problem, then the situation would remain unsolved. I will model this uncertainty as the public having incomplete information about the policy makers incentives, as in (Barro 1986; Phelan 2006), but also allowing the public to have uncertainty about the governments expectations for the continuation game. The main goal of our paper will be to investigate if, through repeated interactions, the government can convince the public about its reformed incentives.

Using equilibrium analysis to answer this question typically relies on rather strong common knowledge assumptions as to how agents play, their priors on the government's type, as well as how all parties revise their beliefs. In the particular case of repeated games, predictions of a particular equilibrium may be extremely sensitive to assumptions as to how agents update their expectations about the continuation game *on all potential histories* that might be observed. These are complicated, high dimensional objects, of which the policy maker may have little information about. This may be due to the difficulty in eliciting both the higher order beliefs from the public as well as contingent beliefs on nodes that may never be reached.

The approach I use is conceptually related to the robust mechanism design literature ((Bergemann and Morris 2005; Bergemann and Morris 2009)).The policy is required to implement trust along its path for all feasible agent beliefs within a large class. The class of beliefs that are deemed feasible is crucial to our exercise, since a larger set makes the analysis more robust, a smaller set makes it trivial; the feasible set I consider is discussed below. I study the case where the only constraints on the sets of beliefs are that they are consistent with common certainty of rationality: every agent knows that the other agents are rational, they know everyone knows this, and so on. This is the rationale behind rationalizability, which consists on an iterative deletion of dominated strategies. The present analysis requires beliefs to be such that agents not to question the government's rationality, unless proven otherwise, which is given by the solution concept of *strong* or *extensive-form rationalizability* of (Pearce 1984; Battigalli and Siniscalchi 2002).

I show that this policy exhibits endogenous transitory gains and losses of reputation. Moreover, the policy achieves permanent separation (i.e. public is convinced about the success of the reform from then on) almost surely and it does so with an exponential arrival rate. As the discount factor of the policy maker increases, the expected payoff of this robust policy approximates the full commitment first-best benchmark. This policy will also be the *max-min* strategy for the policy maker regardless of their particular beliefs and hence provides a lower bound both for payoffs as well as the speed of separation of any strategy that is consistent with extensive form rationalizability.

To understand the intuition behind the results, suppose the public hypothesizes that he faces a time inconsistent policy maker and has observed her taking an action that did not maximized her spot utility. To fix ideas, suppose she took an action such that, if she was the time inconsistent type, gave her 10 utils. Meanwhile, she could have reacted in a manner that gave her 25 in spot utility instead. The implied opportunity cost paid by her would only be consistent with her being rational if she expected a net present value of at least 15 utils, and therefore the opportunity cost paid would have been a profitable investment. This further implies that the government beliefs and planned course of action from tomorrow onward must deliver (from the government's point of view) more than 15 utils, which is a constraint that rationality imposes on the goverment's expected future behavior. If however, the maximum feasible net present value attainable by a time inconsistent government was 10 utils, the public should then infer that the only possible time inconsistent type that they are facing is an irrational one. However, if such a history was actually consistent with the policy maker being time consistent (e.g. she had an opportunity cost of 2) then the public should be fully convinced from then on that they are facing the reformed government. Therefore, the implied spot opportunity cost paid by the time inconsistent type will be a measure of reputation that places restrictions on what the public believes the policy maker will do in the future. I emphasize that this will be independent of their particular beliefs and relies only on an assumption of rationality. I also show that this is in fact the *only* robust restriction that strong common certainty of rationality imposes, making the implied opportunity cost paid the *only* relevant reputation measure in the robust policy. Moreover, I show that the optimal robust policy can be solved as a dynamic contracting problem with a *single promise keeping constraint*, analogous to (Thomas and Worrall 1988), (Kocherlakota 1996) and (Alvarez and Jermann 2000) in the context of optimal risk sharing with limited commitment, which makes the analysis of the optimal robust policy quite tractable.

The rest of the paper is organized as follows. Section 3.2 describes two macroeconomic applications of time inconsistency, monetary policy and capital taxation, which are informed by our theoretical results. Section 3.3 provides a brief literature review. Section 3.4 introduces the stage binary action repeated game, and introduces the concepts of weak and strong rationalizability. Section 3.5 defines the concept of robust implementation and studies robust implementation for all weak and strong rationalizable outcomes. In section 3.6 I study the basic properties of the optimal robust strategy and the reputation formation process as well as the limiting behavior as policy makers become more patient. I also study how further restrictions on the set of feasible beliefs can help accelerate the reputation formation process and in particular find a characterization of restrictions that generate monotone robust policies (i.e. policies that exhibit only permanent gains of reputation on its path). In Section 3.7 I study some extensions to our model and discuss avenues for future research. Finally, Section 3.8 concludes.

3.2 Examples

I start with some time-inconsistency examples from the literature and use them to motivate my model and analysis. I focus on two of the most commonly studied questions in the macroeconomic literature: *capital taxation* and *monetary policy.* I will illustrate that even if the policy maker undertakes a reform that solves for her time inconsistent bias, when agents have imperfect knowledge about the government objectives, a time inconsistency problem of government policy arises.

3.2.1 Capital Taxation

I use a modified version of (Phelan 2006) and (Lu 2012), where the time inconsistent type is a benevolent goverment, instead of just opportunistic. Consider an economy with two type of households: workers (w) and capitalists (k). There is a continuum of measure one of identical households, for each type. Capitalist households have an investment possibility and can invest $q \in [0, \overline{q}]$ units in a productive technology with a constant marginal benefit of 1 and a constant marginal cost of I. Workers do not have access to this technology and can only consume their own, fixed endowment of $e > 0$.

There is also a public good that can be produced by a government that has a marginal value of z_k to capitalist households and z_w to workers, where $z = (z_k, z_w)$ is a joint random variable. The government taxes a portion $\tau(z)$ of capital income after the shock is realized, in order to finance the production of $r(z)$ units of public good. Given the expected policy $\{\tau(z), r(z)\}_{z \in Z}$, workers and capitalists households utilities are given by

$$U_w = e + \mathbb{E}_z [r(z) z_w] \tag{3.1}$$

$$U_k = (1 - \tau^e) q - Iq + \mathbb{E}_z [r(z) z_k] \tag{3.2}$$

where $\tau^e = \mathbb{E}[\tau(z)]$. A leading example is the case where the "public good" is simply redistribution from the capitalists to the workers. In this case $z_w > 0$ and $z_k = 0$.

The optimal investment decision for a capitalist is to invest $q_i = \overline{q}$ if $1 - \tau^e < I$, and 0 otherwise, since they do not internalize their marginal effect on the production of public good. As a benchmark, we will first solve for the policy $\{\tau_k(z), r_k(z)\}_{z \in Z}$ that maximizes only the capitalist households expected utility, subject to the government's budget constraint:

$$\max_{q \in [0,\overline{q}], \tau_k(\cdot), r_k(\cdot)} (1 - \mathbb{E}[\tau_k(z)]) q - Iq + \mathbb{E}[r_k(z) z_k] \text{ s.t. } r_k(z) \leq \tau_k(z) q \text{ for all } z \tag{3.3}$$

Given q, the optimal policy involves full expropriation ($\tau_k(z) = 1, r_k(z) = q$) when $z_k \geq 1$ and zero taxes otherwise, which induces an expected tax rate of $\tau^e = \Pr(z_k \geq 1)$. If

$$I < \Pr(z_k \leq 1) \tag{3.4}$$

then households expecting policy $\{\tau_k(z), r_k(z)\}_{z\in Z}$ will choose $q = \overline{q}$. However, this will not be the policy chosen by a benevolent government that also values workers. After the households investment decision, and the state of nature has been realized, the government chooses public good production $\tilde{r}$ and tax rate $\tilde{\tau}$ to solve:

$$\max_{\tilde{r},\tilde{\tau}} \tilde{r}(z_k + \alpha z_w) + (1-\tilde{\tau}) q \text{ s.t. } \tilde{r} \leq \tilde{\tau} q \tag{3.5}$$

where $\alpha \geq 0$ is the relative weight that the government puts on workers welfare.

Defining $z_g := z_k + \alpha z_w$, the marginal value of the public good between capitalists and the government will typically be different, unless $\alpha = 0$. Solving 3.5 gives $\tau_g^e = \Pr(z_k + \alpha z_w > 1)$. I will assume that $I > \Pr(z_k + \alpha z_w \leq 1)$, so capitalist households will optimally decide not to invest ($q = 0$) and no public good production will be feasible. Finally, I assume that the parameters of the model are such that a benevolent government would want to commit to the capitalist's most preferred policy $\{\tau_k(z), r_k(z)\}_{z\in Z}$ if she was given the possibility. [1]

To solve the "time inconsistency" problem, I first explore the possibility of introducing a cost to raise taxes. This means that if taxes are increased, the government has to pay a cost of $c > 0$. The government would then optimally choose taxes $\tau = 0$ and increase them only when needed. She solves

$$\max_{\tilde{r},\tilde{\tau}} \tilde{r}(z_k + \alpha z_w) - 1\{\tilde{\tau} > 0\} c \text{ s.t. } \tilde{r} \leq \tilde{\tau} q.$$

In this case, the expected tax rate is now $\tau^e(c) = \Pr\left(z_k + \alpha z_w > \frac{c}{q}\right)$. By setting $c = \overline{c}$ to solve $1 - \tau^e(c) = I$, the time inconsistent government can now induce households to invest, by credibly distorting its tax policy.

Another way to deal with the problem is to make an institutional reform and delegate the public good provision to a different policy maker, who has incentives aligned with the capitalist households. The new policy maker type now solves

$$\max_{\tilde{r},\tilde{\tau}} \tilde{r}(z_k + \alpha_{new} z_w) + (1-\tilde{\tau}) q \text{ s.t. } \tilde{r} \leq \tilde{\tau} q.$$

By introducing a "pro-capitalist government" with $\alpha_{new} = 0$, the capitalists most desirable policy $\{\tau_k(z), r_k(z)\}_{z\in Z}$ would be credibly implemented without the need of setting a cost to increase capital taxes. Under some parametric assumptions, it will be socially desirable for the benevolent government (without taken into account the commitment cost payed) to delegate policy making to the "pro-capitalist" type that does not need to impose tax increase costs to convince households to invest[2].

However, if households were not convinced that they are indeed facing a reformed, pro-capitalist

[1]This happens if $\Pr(z_k > 1)\mathbb{E}(z_k + \alpha z_w \mid z_k > 1) + \Pr(z_k \leq 1) > 0$.

[2]This happens if $\Pr(z > 1)\mathbb{E}(z_k + \alpha z_w \mid z_k > 1) + \Pr(z_k \leq 1) > \Pr\left(z_k + \alpha z_w > \frac{c}{q}\right)\mathbb{E}\left(z_k + \alpha z_w \mid z_k + \alpha z_w > \frac{c}{q}\right) + \Pr(z_k + \alpha z_w \leq 1)$

government, they will need some assurance (i.e. some restrictions to ex post increase taxes) in order to *trust* that the government will not expropriate their investments too often. In (Phelan 2006), the analog to a pro-capitalist type is a commitment type (as in (Fudenberg and Levine 1989)) that always pick the same tax rate. In (Lu 2012), the government can make announcements, and can be either a committed type (i.e. one that is bound by the announcement) or a purely opportunistic type that may choose to deviate from the promised policy, which is analog to the benevolent type in our setting.

Formally, Let $\pi \in (0,1)$ be the probability that capitalist households assign to the new government to actually be a pro-capitalist type. Then, if there is complete flexibility to increase taxes, the expected tax rate would be

$$\tau^{e}(\pi) = \pi \Pr(z_k > 1) + (1-\pi)\Pr(z_k + \alpha z_w > 1) \tag{3.6}$$

Condition 3.6 implies that for sufficiently low π, we would have $1 - \tau^{e}(\pi) < I$ and capitalists will decide not to invest. Thus, as long as capitalists perceive that the new government might still be time inconsistent (modeled by a low π), it will be necessary to set some cost to raise taxes in order to induce capitalists to invest, even though the government is now a pro-capitalist type.

3.2.2 Monetary Policy

I use the framework in (Obstfeld and Rogoff 1996)[3]. I assume that total output (in logs) y_t depends negatively on the real wage and some supply side shock z_t, according to

$$y_t = \overline{y} - [w_t - p_t(z_t)] - z_t \tag{3.7}$$

where $\overline{y}$ is the flexible price equilibrium level, z_t is a supply shock with $\mathbb{E}(z_t) = 0$ and $p_t(z_t)$ is the nominal price level at time t set by the monetary authority. In equilibrium, nominal wages are set according to $w_t = \mathbb{E}_{t-1}[p_t(z_t)]$, to match expected output to its natural level $\overline{y}$. A benevolent monetary authority observes the shock z_t and decides the inflation level in order to minimize deviations of output with respect to a social optimal output y^* and deviations of inflation from a zero inflation target:

$$\mathcal{L} = \frac{1}{2}(y_t - y_t^*)^2 + \frac{\chi}{2}\pi_t^2. \tag{3.8}$$

I assume that $k := y_t^* - \overline{y} > 0$. This measures the wedge between the output level targeted by authorities and the natural level of output, which are different due to market inefficiencies, even under flexible prices.[4] Defining inflation as $\pi_t(z_t) := p_t(z_t) - p_{t-1}$, and using equation 3.7, together

[3]Section 9.5, pp 634-657.

[4]See (Rogoff 1985) and (Obstfeld and Rogoff 1996) for a discussion of such potential inefficiencies.

with the wage setting rule, the loss function simplifies to

$$\mathcal{L}(\pi, \pi^e, z) = \frac{1}{2}[\pi - \pi^e - z - k]^2 + \frac{\chi}{2}\pi^2, \tag{3.9}$$

where $\pi^e = \mathbb{E}[\pi(z)]$ are the expectations formed by the private sector about inflation (which should be correct under rational expectations). The *full commitment benchmark*, in which the monetary authority can commit, ex-ante, to a state contingent inflation policy $\pi(z)$, to solve

$$\min_{\pi(\cdot), \pi^e} \mathcal{L}(\pi, \pi^e, z) \text{ s.t: } \pi^e = \mathbb{E}[\pi(z)]$$

with solution

$$\pi_c(z) = \frac{z}{1+\chi} \quad \text{and} \quad \pi^e = \mathbb{E}_z[\pi_c(z)] = 0. \tag{3.10}$$

In contrast, when the monetary authority cannot commit to a state contingent policy, conditional on π^e and z, she chooses π to solve:

$$\min_{\pi \in \mathbb{R}} \mathcal{L}(\pi, \pi^e, z) \iff \pi_{nc}(z) = \frac{\pi^e + z + k}{1+\chi} \tag{3.11}$$

By taking expectations on both sides of 3.11 we get $\pi^e = \frac{k}{\chi}$, which I will refer to the *time inconsistency bias.* Equilibrium inflation is then

$$\pi_{nc}(z) = \frac{k}{\chi} + \pi_c(z) \tag{3.12}$$

Output $y(z)$ is identical in both cases, for all shocks. However, $\mathbb{E}\left[\pi_{nc}^2(z)\right] = \mathbb{E}\left[\pi_c^2(z)\right] + \frac{k^2}{\chi^2}$ so the outcome with no commitment is strictly worse than the full commitment benchmark.

How can the monetary authority solve this problem? A first approach is to formally limit the flexibility of monetary policy by restricting the set of inflation levels the monetary authority can choose from. (Athey, Atkeson, and Kehoe 2005) show that this can be optimally done by choosing an *inflation cap* $\overline{\pi}$[5], such that $\pi(z) \leq \overline{\pi}$ for all z. Inflation policy is now

$$\pi(z \mid \overline{\pi}) = \min\left\{\frac{\pi^e(\overline{\pi}) + z + k}{1+\chi}, \overline{\pi}\right\}$$

[5] In their paper, (Athey, Atkeson, and Kehoe 2005) solve for the optimal dynamic mechanism for a time inconsistent policy maker, that has private information about the state of the economy, which is i.i.d across periods, and show that any optimal mechanism exhibits a constant inflation cap in all periods. In a static setting, shocks can be thought as private information for the monetary authority, so an inflation cap would also be a characteristic of the more general mechanism design problem:

$$\max_{\pi(\cdot), \pi^e} \mathbb{E}_z[\mathcal{L}(\pi(z), \pi^e, z)]$$
$$s.t \; : \; \mathcal{L}[\pi(z), \pi^e, z] \geq \mathcal{L}[\pi(z'), \pi^e, z] \text{ for all } z, z' \in Z$$

where $\pi^e(\overline{\pi})$ solves the fixed point equation $\pi^e(\overline{\pi}) = \mathbb{E}_z\left[\min\left\{\frac{\pi^e(\overline{\pi})+z+k}{1+\chi}, \overline{\pi}\right\}\right]$.

An alternative approach, first suggested by (Rogoff 1985), is to introduce *institutional reforms to the monetary authority*, with the purpose of alleviating the time inconsistency bias by inducing changes in their preferences. Imagine first that the government can delegate the monetary policy to a policy maker type $\theta = new$, that wants to minimize a modified loss function

$$\mathcal{L}_{new}(\pi, \pi^e, z) = \frac{1}{2}\left[\pi - \pi^e - z - k^{new}\right]^2 + \frac{\chi^{new}}{2}\pi^2 \tag{3.13}$$

(Rogoff 1985) suggests placing a "conservative central banker", that has $k^{new} = k$ but $\chi^{new} > \chi$, so that it places a greater importance on inflation stabilization than society does. From 3.12 we see that increasing the weight χ makes the effective inflation bias smaller, and hence may alleviate the time inconsistency problem at the expense of a milder reaction to supply shocks, as evidenced by equation 3.11.

By setting $k^{new} = 0$ the optimal policy with no commitment for $\theta = new$ would implement the full commitment solution. This would correspond to having a monetary authority that believes there are no market inefficiencies and wants to stabilize output around its flexible price equilibrium level $\overline{y}$. The same outcome can be implemented if, instead of changing the preference parameters, we add a linear term to the loss function:

$$\mathcal{L}_{new} = \mathcal{L} + \alpha\left[\eta\pi\right]$$

(Walsh 1995) and (Persson and Tabellini 1993) argue that this can be done by offering a contract to the central bank governor. This can be either a formal monetary contract[6] or an informal relational contract under which realized levels of inflation affect the continuation values for the monetary authority (e.g. the governor could be fired if inflation reaches sufficiently high levels, as in (Lohmann 1992)). Here $\alpha > 0$ represents the relative weight of his self-interest payoffs relative to the social welfare. By picking $\eta = \frac{k}{\alpha}$ the full commitment inflation policy would be implemented.

While the institutional reform route may seem desirable, *these institutional reforms may not be perfectly observed by the private sector.* The public might not be convinced that the monetary authority is now more conservative or have a smaller time inconsistency bias than the previous one. Such a problem is likely to be particularly acute because these are changes in preferences, which involve either delegation or perhaps informal relational contracts that are imperfectly observed. If this was the case, restrictions such as inflation caps might still be necessary. For example, take the institutional reform with $k^{new} = 0$, and no inflation caps are set. If the public assigns probability

[6] Suppose the monetary authority minimizes $\hat{\mathcal{L}}_{old} = \mathcal{L}_{old} - \alpha u\left[\phi(\pi)\right]$ where $\phi(\pi)$ is a monetary reward function depending on realized inflation, and $\alpha > 0$ of monetary incentives relative to the monetary authorities "benevolent" incentives. See that by picking $\phi(\pi) = u^{-1}(-\eta\pi)$, a decreasing function of inflation, the contract will induce the linear component in 3.13, which coincides with the optimal contract in this setting.

$\mu \in (0,1)$ to the incentive reform being successful, expected inflation would then be

$$\pi^e = (1-\mu)\,\frac{k}{\chi} > 0$$

Thus, as long as the public perceives there might still be a time inconsistency bias, institutional reforms might not be enough, and inflation caps might be necessary to implement smaller inflation expectations. The literature has studied the case of

3.3 Literature Review

The literature on time inconsistency of government policy is extensive, beginning with the seminal papers by (Kydland and Prescott 1977) and (Barro and Gordon 1983), where the idea of the commitment solution (i.e. choosing policy first) was first introduced. The reputation channel was first explored by (Backus and Driffil 1984) and (Barro 1986), who studied policy games in which a rational (albeit time inconsistent) government living for finitely many periods may find it optimal to imitate a "commitment type". This commitment type is an irrational type that plays a constant strategy at all histories. They show (following the arguments in (Kreps, Milgrom, Roberts, and Wilson 1982; Kreps and Wilson 1982; Milgrom and Roberts 1982)) that for long enough horizon, the unique sequential equilibrium of the game would involve the government imitating the commitment type for the first periods, and then playing mixed strategies, which imply a gradual reputation gain if she keeps imitating.

In an infinite horizon setting, (Fudenberg and Levine 1989) show that a long lived agent facing a sequence short lived agents can create a reputation for playing as the commitment type. By consistently playing the commitment strategy, the long lived agent can eventually convince the short lived agents that she will play as a committed type for the rest of the game. (Celentani and Pesendorfer 1996) generalized this idea to the case of a government playing against a continuum of long-lived small players, whose preferences depend only on aggregate state variables. The atomistic nature of the small players allows them to use (Fudenberg and Levine 1989) results to get bounds on equilibrium payoffs. (Phelan 2006) studies the problem of optimal linear capital taxation, in a model with impermanent types, which can accommodate occasional losses of reputation. Rather than obtaining bounds, he characterizes the optimal Markovian equilibrium of the game, as a function of the posterior the public has about the government's type.

A second strand of the literature on reputation focuses on a complete information benchmark with the goal of characterizing sustainable policies. These are policies that are the outcome of a subgame perfect equilibrium of the policy game, starting with (Stokey 1989; Stokey 1991) and (Chari and Kehoe 1990; Chari and Kehoe 1993a; Chari and Kehoe 1993b). In such environments, governments may have incentives to behave well under the threat of punishment by switching to a bad equilibrium afterwards. See (Sargent and Ljungqvist 2004)[7] for a tractable unified framework to study these issues.[8]

This paper studies reputation formation on both dimensions: in terms of payoff heterogeneity and in terms of equilibrium punishments. The main point of departure is that instead of designing the optimal policy for a time inconsistent policy maker (that wants to behave as if she was time consistent), I focus on the opposite case. I consider the problem of a trustworthy policy maker (with no time inconsistency bias) who nevertheless may be perceived as opportunistic by the agents. Therefore, its goal is essentially to separate itself from the time inconsistent, untrustworthy type, if possible. The most related papers in spirit to mine are (Debortoli and Nunes 2010),(King, Lu, and Pasten 2012) and particularly (Lu 2012). (Debortoli and Nunes 2010) study the optimal policy problem of a benevolent government that has access to a "loose commitment" technology, under which not all announcements can be guaranteed to be fulfilled. (Lu 2012) explores the optimal policy of a committed government that worries she might be perceived as a government that cannot credibly commit to her announced policies. This paper also focuses on characterizing the optimal policy for the time consistent type (the committed type in her setting) instead of just studying the optimal policy of a time inconsistent type imitating a consistent one. (King, Lu, and Pasten 2012) apply these ideas in the context of the standard New Keynesian model, similar to our setup in subsection 3.2.2. Ultimately these papers study equilibrium in an environment where all players involved know that the government is ex-ante either a type that can commit or not (which holds for all subsequent periods). They then study a particular equilibrium refinement that happens to select the best equilibrium for the able-to-commit type. They also show that other equilibrium refinements such as the intuitive criterion (e.g., (Cho and Kreps 1987)) select a different equilibrium. In macroeconomics, the most related paper to mine that studies robustness to specific refinements is (Pavan and Angeletos 2012). They study the robust predictions of any equilibria in a global game setting with incomplete information.

The literature on robust mechanism design is fairly recent, starting with partial robust implementation in (Bergemann and Morris 2005), and robust implementation in (Bergemann and Morris 2009). The latter focuses on finding conditions on environments and social choice functions such that they are implemented under implemented for all possible beliefs, if the only thing that the mechanism designer knows about the agent's beliefs is that they share common knowledge (or certainty) of rationality. When the environment is dynamic, different concepts of rationalizability may

[7] Chapter 16, pp 485-526

[8] This principle is also exploited in the relational contract literature ((Levin 2003),(Baker and Murphy 2002) ,(Bull 1987)) where a principal announces a payment scheme after income is realized (the state-contingent policy) but has no commitment to it other than the one enforced by the threat of retaliation by the agent (not making effort, strike, quit, etc). Similar themes are studied in the literature on risk sharing with limited commitment ((Thomas and Worrall 1988), (Kocherlakota 1996), (Ligon 1998) and (Ligon, Thomas, and Worrall 2000)) where a transfer scheme conditional on the realization of income (the contingent policy) is enforced by threating agents who deviate of excluding them from the social contract.

be used, like normal form Interim Correlated Rationalizability (as in (Weinstein and Yildiz 2012)) and Interim Sequential Rationalizability ((Penta 2011; Penta 2012)), among others. This paper focuses on the stronger assumption of common strong certainty of rationality ((Battigalli and Siniscalchi 1999; Battigalli and Siniscalchi 2002; Battigalli and Siniscalchi 2003)) which is also equivalent to (Pearce 1984) notion of "Extensive-form rationalizability". In a similar vein, the paper most related in spirit to mine is (Wolitzky 2012). He studies reputational bargaining in a continuous time setting in which agents announce bargaining postures that they may become committed to with a given positive probability. He characterizes the minimum payoff consistent with mutual knowledge of rationality between players (i.e., one round of knowledge of rationality), and the bargaining posture that she must announce in order to guarantee herself a payoff of at least this lower bound. A crucial difference to my setting is the commitment technology, which ensures certain expected payoffs to the other party, regardless whether they think they are facing a rational opponent or not. I characterize optimal robust policy in a repeated setting in which one can guarantee themselves the best payoff that is consistent with (strong) common knowledge of rationality.

3.4 The Model

I now introduce the framework and model. Section 3.4.1 describes the stage game and shows the multiplicity of equilibria. I then setup the repeated game in Section 3.4.2 and develop the concept of system of beliefs in Section 3.4.3. Section 3.4.4 introduces weak and strong rationalizability and Section 3.4.5 argues why we must turn to robustness relative to equilibrium refinements.

3.4.1 Stage Game

There are two players: a policy maker d (she) and an agent p (he). The agent represents the public. In the no-commitment benchmark, p is asked to trust a state-contingent decision to d, who after a state of nature $z \in Z$ is realized, has to choose a policy that affects both parties payoffs. For simplicity, I will assume that d has only two options: a "normal" policy that is optimal most of the time and an "emergency" policy that needs to be taken in certain instances. As a mnemonic device, we will write g for the normal policy (pushing a "$g$reen button") and r for the emergency policy (pushing a "$r$ed button"). The extensive form game is described in Figure 3-1.

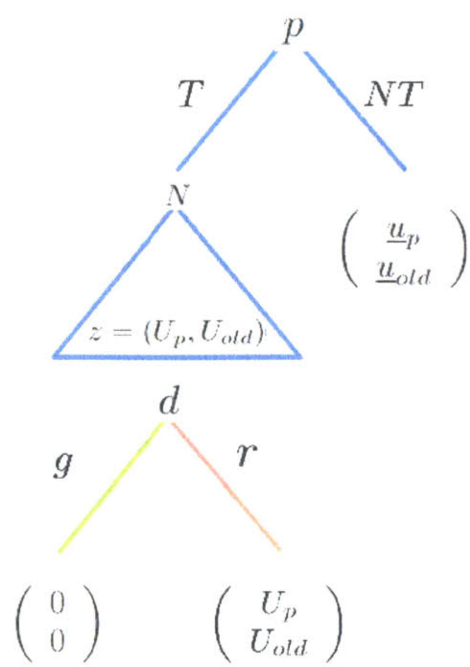


Figure 3-1: Stage game with time inconsistent type

The random shock $z = (U_p, U_{old})$ is the profile of relative utilities of the emergency action r with respect to g, for both the public p and the decision maker d. I assume this random shock to be an absolutely continuous random variable over $Z := [\underline{U}, \overline{U}]^2 \subset \mathbb{R}^2$, with density function $f(z)$. The subscript "old" serves to remind the reader that these are the preferences of the policy maker before a reform is undertaken, which will be described below.

To make this concrete, consider the capital taxation example described in Section 3.2.1. The government decides whether to expropriate capital to finance public good provision (the emergency policy with $r = q$) or to keep taxes at zero (the normal policy with $r = 0$). The government gets to make this decision only when capitalist households decide to invest $q_i = \overline{q}$ which corresponds to the public placing trust in the government as without any investment the government cannot finance public good provision in the first place. As such, T corresponds to "invest $\overline{q}$". The shock z can be written as

$$z = (U_p, U_{old}) := \overline{q}\,(z_k - 1, z_b - 1) \tag{3.14}$$

Returning to the general model, I will additionally assume that

$$\mathbb{E}_z\left[\max(0, U_p)\right] > \underline{u}_p > \int_{U_{old}>0} U_p f(z)\, dz \tag{3.15}$$

This implies that if p was guaranteed his most desirable policy, he would place trust in d.[9] I will

[9] In the context of the capital taxation problem, by setting $\underline{u}_p = -(1 - I)\,\mathbb{E}(z_k \mid z_b \geq 1)(\overline{q} - 1)$, the left hand side inequality of 3.15 corresponds to the solution to 3.3 in the capital taxation problem, together with assumption 3.4. Notice also that $\underline{u}_p < 0$

further assume that

$$\underline{u}_p, \underline{u}_{old} < 0 \tag{3.16}$$

so that both parties would benefit from ensuring that the decision maker plays the normal action g for all shocks (and hence losing all flexibility), if such a policy was enforceable. In the capital taxation problem, this would correspond to a ban on positive taxes while in our inflation example this would correspond to a commitment to zero inflation.[10] Although this would induce the public trust, this would come at the cost of losing all flexibility to optimally react to shocks.

As I discussed in the introduction, I will explore two (potentially complementary) ways to solve this *time inconsistency problem:* by credibly loosing *some* flexibility to react to the economic shock, or by reforming the incentives of the decision maker, in order to alleviate the time inconsistency problem.

First, I introduce a commitment cost technology, under which the decision maker can choose, before the game starts, a cost $c \geq 0$ of taking the emergency action r. This can be interpreted as a *partial commitment* to the normal policy g, that includes an escape clause to break the commitment, forcing the decision maker to suffer a cost of $c \geq 0$ utils (as in (Lohmann 1992)). Although d cannot commit to a complete contingent rule, I assume that the commitment cost chosen is binding. In the capital taxation model, this corresponds to the cost of increasing taxes chosen by the time inconsistent government, while in the inflation setting model this would intuitively translate to the inflation cap $\overline{\pi}$, in that it is a partial commitment chosen by the monetary authority. The modified stage game is illustrated in Figure 3-2.

[10] Although this seems to be to an extreme policy to be seen in practice, hyperinflation stabilization programs usually involve drastic measures, that resemble losing all flexibility to stabilize output. For example, Zimbabwe in 2009 decided to abandon its currency (and hence most of its monetary policy) within the context of a severe hyperinflation (which reached a peak of 79,600,000% per month in November of 2008).

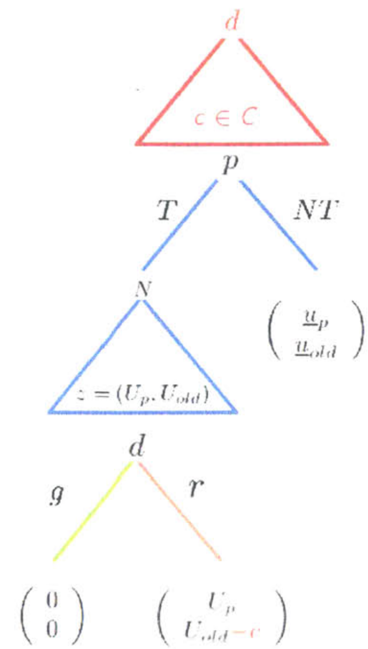


Figure 3-2: Stage game with commitment cost choice

I will allow $c = \infty$, the case where d decides to shut down the emergency action r, as I previously discussed, so that the commitment cost set is $C = [0, \max_{z \in Z} \max(|z_p|, |z_{old}|)] \cup \{\infty\}$. In this game, the decision maker would then choose the commitment cost $c = \overline{c}$ that makes p indifferent between trusting and not trusting:

$$\overline{c} := \min \left\{ c \geq 0 : \int_{U_{old} > c} U_p f(z)\, dz \geq \underline{u}_p \right\} < \infty \tag{3.17}$$

While this technology is available, I will consider the case in which the policy maker makes a reform, by effectively changing the ex-post incentives that the decision maker faces, by either delegating the decision to a different agent (like the conservative central banker of (Rogoff 1985)) or by designing a contract for the decision maker (as in (Walsh 1995) and (Persson and Tabellini 1993)). I model this reform by creating a new policy maker type, $\theta = new$, with ex-post payoffs given by

$$U_{new}(z) = U_p \tag{3.18}$$

i.e. the reformed decision maker has the same ex-post incentives as the public. Condition 3.15 implies that if p knew he was facing this type of decision maker, he would trust her even with $c = 0$, so that no commitment cost would be necessary. This corresponds to the "pro-capitalist" government in the capital taxation model, which removes the time inconsistency of government policy.

The key problem I will study in this paper is the government's lack of credibility. Even though an incentive reform has been carried out, the public may remain unconvinced that the reform has been effective. For instance, investors might still believe that the government is not pro-capitalist enough, and will expropriate them too often, or that the new appointed central banker may not be a conservative type. I model this situation by introducing *payoff uncertainty* from the public side: p believes he is facing a reformed, time consistent decision maker $\theta = new$ with probability $\pi \in (0, 1)$ and otherwise faces the old, time inconsistent type $\theta = old$. This results on an incomplete information game, described in Figure 3-3.

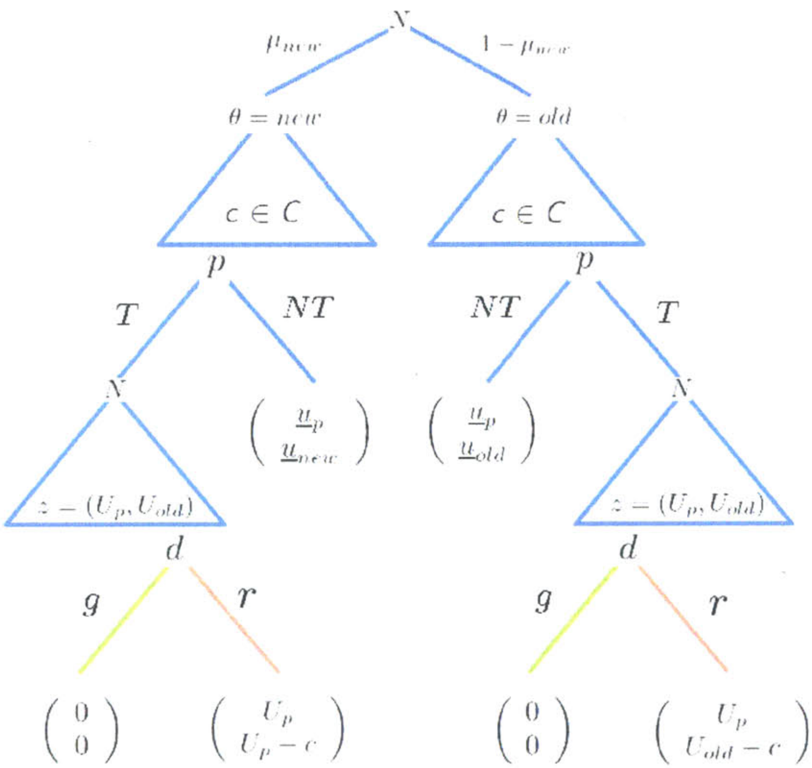


Figure 3-3: Incomplete information game

This model follows an *executive approach* to optimal policy, where it is the decision maker herself who decides the commitment cost and the policy rule. This contrasts with the *legislative approach* studied in (Canzoneri 1985) and (Athey, Atkeson, and Kehoe 2005) who instead solve for the optimal mechanism design problem from the point of view of p. In Section (3.7) I briefly explore this route and find that in our setting, it will be detrimental to welfare, conditional on the government being of type $\theta = new$.

Because the commitment cost choice is taken after the type has been realized, this is effectively a *signaling game*. The choice of commitment cost could in principle help p to infer d's type. As a prelude to what follows it is useful to characterize the set of Perfect Bayesian Equilibria (PBE) of this game.

Let $\underline{c}(\pi)$ be the minimum cost that induces p to trust d in any pooling equilibrium:

$$\underline{c}(\pi) := \min\left\{c \geq 0 : \pi \int_{U_p > c} U_p f(z)\, dz + (1-\pi) \int_{U_{old} > c} U_p f(z)\, dz \geq \underline{u}_p\right\}. \tag{3.19}$$

It is easy to see that $\underline{c}(\pi)$ is decreasing in π and that $\underline{c}(\pi) < \underline{c}(0) = \overline{c}$ according to 3.17. Proposition 24 characterizes the set of all PBE of the stage game.

Proposition 24. *All PBE of the static game are pooling equilibria. For any $\hat{c} \in [\underline{c}(\pi), \overline{c}]$ there exists a PBE in which both types choose $\hat{c}$ as the commitment cost.*

Proof. Any equilibrium must induce p to trust since d can always choose $c > \overline{U}$ (so it will never be optimal to take the emergency action) and get a payoff of $0 > \underline{u}_d$. There cannot be any pooling equilibrium with $c < \underline{c}(\pi)$, since the definition of $\underline{c}(\pi)$ implies it would give p less than his reservation utility $\underline{u}_p$. It cannot happen either if $c > \overline{c}$, since either type would deviate and choose $\overline{c}$ and induce p to trust, regardless of his updated beliefs $\pi_p(\overline{c})$. This follows from

$$\pi_p(\overline{c}) \int_{U_P > \overline{c}} U_p f(z)\, dz + [1 - \pi_p(\overline{c})] \int_{U_{old} > \overline{c}} U_p f(z)\, dz \underset{(1)}{>}$$

$$\pi_p(\overline{c})\, \mathbb{E}[\max(U_p - \overline{c}, 0)] + [1 - \pi_p(\overline{c})]\, \underline{u}_p \underset{(2)}{>} \pi_p(\overline{c})\, \underline{u}_p + [1 - \pi_p(\overline{c})]\, \underline{u}_p = \underline{u}_p \tag{3.20}$$

where (1) follows from definition 3.17 and (2) from the fact that $0 > \underline{u}_p$. I will now show that for any $\hat{c} \in [\underline{c}(\pi), \overline{c}]$ there exist a pooling equilibrium in which both $\theta = new$ and $\theta = old$ find it optimal to choose $c = \hat{c}$. Conjecture the following belief updating rule:

$$\pi_p^{\hat{c}}(c) := \begin{cases} 0 & \text{if } c < \hat{c} \\ \pi & \text{if } c \geq \hat{c}. \end{cases} \tag{3.21}$$

Under a pooling equilibrium, since $\hat{c} \geq \underline{c}(\pi)$, p will trust d. Neither type will deviate from $\hat{c}$ since the optimal deviation that would make p trust would be to choose $\hat{c} = \overline{c}$. The non-existence of other PBE is left to the appendix. □

Suppose now that, in light of the results of Proposition 24, the reformed decision maker $\theta = new$ is considering what commitment cost to choose. If we are thinking about a policy prescription that is supported by some PBE of this game, the multiplicity of equilibria in Proposition 24 requires some equilibrium refinement. If the policy prescription was the commitment cost supported by the *best equilibrium* of the game, Proposition 24 shows that $c = \underline{c}(\pi)$ is selected for either $\theta = new$ or $\theta = old$. We will argue that this will not be a particularly "robust" policy in the sense that expecting p to trust after observing $c = \underline{c}(\pi)$ relies on strong and sensitive assumptions about p 's beliefs, which may be imperfectly known by the decision maker.

First, observe that the best PBE policy is very sensitive to the prior. If $p's$ true prior were

$\tilde{\pi} = \pi - \epsilon$ for some $\epsilon > 0$, then p would not trust after observing $c = \underline{c}(\pi)$. Second, even if π were commonly known, after the commitment cost has been chosen, p updates beliefs to $\pi_p(c) \in (0,1)$. As illustrated by the proof of Propostion 24 (in particular the belief updating rule in 3.21), the indeterminacy of beliefs following zero probability events generates a large set of potential beliefs that can arise in equilibrium. As such, p's behavior will depend on the complete specification of her updates beliefs for *all off-equilibrium costs*, not just the candidate equilibrium one. Therefore, small changes in the updating rule (for example, by changing $\hat{c}$ in 3.21) generates potentially very different behavior for p.[11]

Our main question becomes whether we can choose a policy that is robust to mis-specifications of both the prior π and the updating rule $\pi_p(c)$. It is clear that by choosing $c = \infty$ and effectively removing all flexibility, a rational p would trust d independently of his beliefs. However, we can do better. Inequality 3.20 implies that if $c = \overline{c}$ and p will find it optimal to trust irrespective of the updating rule $\pi_p(c)$ if he still believes he is facing a rational decision maker. I will show that in fact $c = \overline{c}$ is the *only* robust policy, when the only assumption we make about $p's$ beliefs is that they are consistent with strong common certainty of rationality; i.e. p believes he is facing a rational d if her observed past behavior is consistent with common knowledge of rationality.

Since the difference between the reformed and the time inconsistent type is about their ex-post incentives, Proposition 24 gives a negative result: types cannot separate in any equilibrium of the stage game, by their choice of c. Only by having repeated interactions can the reformed decision maker hope to convince p of the success of the reform, trying to signal her type through her reactions to the realized shocks. Throughout the remainder of the paper I will investigate whether robust policies, such as the one I found in the static game, can eventually convince p that $\theta = new$, regardless of his particular belief updating rule.

3.4.2 Repeated game: Setup and basic notation

I extend the stage game to an infinite horizon setting: $\tau \in \{0, 1, ...\}$. I assume that d is infinitely lived and that types are permanent; i.e. at $\tau = 0$ nature chooses $\theta = new$ with probability π_{new}. d has discounted expected utility with discount factor $\beta_\theta \in (0,1)$. For notational ease, I will assume $\beta_{old} = \beta_{new} = \beta$. I will specify when the results are sensitive to this assumption. Shocks are iid across periods: $z_\tau := (U_{p,\tau}, U_{old,\tau}) \sim_{i.i.d} f(z_\tau)$. I assume that there is a sequence of myopic short run players p_τ (or equivalently $\beta_p = 0$) which is a standard assumption in the reputation literature ((Fudenberg and Levine 1989), (Phelan 2006)). This will be without loss of generality for most

[11] More generally, we apply (Aumann and Brandenburger 1995) results to the interim normal form of this game, finding tight sufficient conditions for any particular Bayesian equilibria (not necessarily perfect) to be the expected solution outcome: **(a)** There is common knowledge of rationality, **(b)** the strategies of both $\theta = new$ and $\theta = old$ prescribed by the Bayesian equilibrium are common knowledge, and **(c)** the inference rule $\pi_p(.)$ is also common knowledge

applications to macroeconomic models applications.[12] At every period, d chooses $c_\tau \geq 0$ which is binding only for that period. The policy maker can change its choice freely in every period. I also assume that all past history of actions and shocks (except for $d's$ payoff type) is observed by all players at every node in the game tree. I will further assume that the structure of the game described so far is *common knowledge* for both players and that agents know their own payoff parameters.[13]

A stage τ *outcome* is a 4-tuple $h_\tau = (c_\tau, a_\tau, z_\tau, r_\tau)$ where c_τ is the commitment cost, $a_\tau \in \{0,1\}$ is the trust decision, and $r_\tau \in \{0,1\}$ is the contingent policy, where $r_\tau = 1$ if d chooses the emergency action, and $r_\tau = 0$ otherwise. A *history* up to time τ is defined as

$$h^\tau := (h_0, h_1, ..., h_{\tau-1}).$$

I will refer to a "partial history" as a history plus part of the stage game. For example, p moves at histories (h^τ, c_τ), after the commitment cost is chosen. The set of all partial histories will be denoted as $\mathcal{H}$, and $\mathcal{H}_i \subseteq \mathcal{H}$ is the set of histories in which agent $i \in \{p,d\}$ has to take an action.

A *strategy* for the policy maker is a function $\sigma_d : \mathcal{H}_d \to C \times \{0,1\}^Z$ that specifies, at the start of every period τ, a commitment cost c_τ and the contingent choice provided p trusts. Then, we can always write a strategy as a pair:

$$\sigma_d(h^\tau) = \left(c^{\sigma_d}(h^\tau), r^{\sigma_d}(h^\tau, \cdot) : C \times Z \to \{0,1\}\right) \tag{3.22}$$

where the choice is a commitment cost $c^{\sigma_d}(h^\tau) \in C$ and a *policy rule function* $r^{\sigma_d}(h^\tau, c_\tau, z_\tau)$ of the shock, given commitment cost c_τ The superscript σ_d serves to remind the reader these objects are part of a single strategy σ_d. Likewise, a strategy σ_p for p is a function $\sigma_p : \mathcal{H}_p \to \{0,1\}$ that assigns to every observed history, his trust decision

$$\sigma_p(h^\tau, c_\tau) = a^{\sigma_p}(h^\tau, c_\tau) = \begin{cases} 1 & \text{if } p \text{ trusts} \\ 0 & \text{if } p \text{ does not trust.} \end{cases}$$

Write the set of strategies of each agent as Σ_i for $i \in \{d,p\}$ Also let $\Sigma = \Sigma_d \times \Sigma_p$ be the set of strategy profiles $\sigma = (\sigma_d, \sigma_p)$. If player $i \in \{d,p\}$ plays strategy σ_i, the set of histories that will be consistent with σ_i is denoted $\mathcal{H}(\sigma_i) \subset \mathcal{H}$. For a history $h \in \mathcal{H}$ we say a strategy σ_i *is consistent with* h if $h \in \mathcal{H}(\sigma_i)$. Let $\Sigma_i(h) = \{\sigma_i \in \Sigma_i : h \in \mathcal{H}(\sigma_i)\}$ be the set of strategies consistent with h.

Given a strategy profile $\sigma = (\sigma_p, \sigma_d)$ let $W_\theta(\sigma \mid h)$ be the expected continuation utility for $d's$

[12](Celentani and Pesendorfer 1996) show that this assumption is without loss of generality when p is modeled as representative agent for a continuum of atomistic and anonymous patient agents. In particular, the capital taxation model of section 3.2.1 satisfies these assumptions when capitalist households have a common discount rate $\delta_k \in (0,1)$.

[13]These are the basic assumptions in (Penta 2012).

type $\theta \in \{old, new\}$ given history h

$$W_\theta(\sigma \mid h) := (1-\beta)\,\mathbb{E}\left\{\sum_{s=\tau}^{\infty} \beta^{s-\tau}\left[a_s r_s\left(U_{\theta,s} - c_s\right) + (1-a_s)\,\underline{u}_p\right] \mid h\right\} \tag{3.23}$$

where $c_s = c^{\sigma_d}(h^s)$, $a_s = a^{\sigma_p}(h^s, c_s)$ and $r_s = r^{\sigma_d}(h^s, c_s, z_s)$. Likewise, denote $V(\sigma \mid h)$ for the spot utility for agent p at history h

$$V(\sigma \mid h) := a_\tau \mathbb{E}_z\left(r_\tau U_{p,\tau}\right) + (1-a_\tau)\,\underline{u}_p. \tag{3.24}$$

3.4.3 Systems of Beliefs

Agents form beliefs both about the payoff types of the other player, as well as the strategies that they may be planning to play. In static games, such beliefs are characterized by some distribution $\pi \in \Delta(\Theta_{-i} \times S_{-i})$ where Θ_{-i} is the set of types of the other agent and S_{-i} their strategy set. In our particular game, $\Theta_d = \{new, old\}$ and $\Theta_p = \{p\}$. In dynamic settings however, agents may *revise their beliefs* after observing the history of play. This revision is described by a *conditional probability system*, that respects Bayes rule whenever possible. Formally, let $\mathcal{X}_i$ the Borel $\sigma-$algebra generated by the product topology[14] on $\Theta_{-i} \times \Sigma_{-i}$ and $\mathcal{I}_i = \left\{E \in \mathcal{X}_i : \text{proj}_{\Sigma_{-i}} E = \Sigma_{-i}(h) \text{ for some } h \in \mathcal{H}\right\}$ be the class of infomation sets for i. A *system of beliefs* π_i on $\Theta_{-i} \times \Sigma_{-i}$ is a mapping $\pi_i : \mathcal{I}_i \to \Delta(\Theta_{-i} \times \Sigma_{-i})$ such that:

1. Given an information set $E \in \mathcal{E}_i$, $\pi_i(. \mid E)$ is a probability measure over $\Theta_{-i} \times \Sigma_{-i}$.[15]

2. If $A \subseteq B \subseteq C$ with $B, C \in \mathcal{I}_i$, then $\pi_i(A \mid B)\,\pi_i(B \mid C) = \pi_i(A \mid C)$.

I write $\pi_i(E \mid h) = \pi_i(E \mid \Sigma_i(h))$ for $E \subset \Theta_{-i} \times \Sigma_{-i}$ for the probability assessment of event E conditional on history h. Denote $\Delta^{\mathcal{H}}(\Theta_{-i} \times \Sigma_{-i})$ to be the set of all systems of beliefs. Given $\pi_d \in \Delta^{\mathcal{H}}(\Theta_p \times \Sigma_p) = \Delta^{\mathcal{H}}(\Sigma_p)$ and strategy $\sigma_d \in \Sigma_d$, define $W_\theta^{\pi_d}(\sigma_d \mid h)$ as the expected continuation payoff for type $\theta \in \{old, new\}$ conditional on history h, under beliefs π_d:

$$W_\theta^{\pi_d}(\sigma_d \mid h) := \int W_\theta(\sigma_d, \hat{\sigma}_p \mid h)\, d\pi_d(\hat{\sigma}_p \mid h) \tag{3.25}$$

Analogously, given a system of beliefs π_p the expected utility of strategy σ_p conditional on history h is

$$V^{\pi_p}(\sigma_p \mid h) := \int V(\hat{\sigma}_d, \sigma_p \mid h)\, d\pi_p(\hat{\sigma}_d \mid h) \tag{3.26}$$

For a given system of beliefs π_i we write $\sigma_i \in SBR_{\theta_i}(\pi_i)$ as the set of sequential best responses of type θ_i to beliefs π_i.[16]

[14] A sequence $\{\sigma_{i,n}\}_{n\in\mathbb{N}}$ converges to σ_i in the product topology in Σ_i if and only if $\sigma_{i,n}(h) \to \sigma_i(h)$ for all $h \in \mathcal{H}_i$

[15] We endow $\Theta_{-i} \times \Sigma_{-i}$ with the Borel $\sigma-$algebra with respect to the product topology.

[16] A strategy σ_p is a *sequential best response to* π_p for all $(h^\tau, c_\tau) \in \mathcal{H}_p$ and all other strategies $\hat{\sigma}_p \in \Sigma_p$, we have $V^{\pi_p}(\sigma_p \mid h^\tau, c_\tau) \geq V^{\pi_d}(\hat{\sigma}_p \mid h^\tau, c_\tau)$. Likewise, σ_d is a *sequential best response to belief system* π_d *for type*

3.4.4 Weak and Strong Rationalizability

In this subsection I introduce the notions of weak and strong rationalizability. Our goal is to find strategies that are robust to changes in p's beliefs so that p is induced to trust as long as there is *common certainty of rationality*, which means that all agents are rational, all agents are certain that all agents are rational and so on, ad infinitum. In static settings, beliefs that satisfy these common knowledge assumptions have their support over the set of *rationalizable strategies.* This set is characterized by an iterative deletion process described in (Pearce 1984). The set of strategies is refined by eliminating those which are not a best response to *some* beliefs about the other agents strategies, which are themselves best responses to some other beliefs, and so on.

However, the possibility of reaching zero probability events also creates different ways to extend the concept of rationalizability, which hinge upon on our notion of "certainty or rationality". An agent is *certain* about some event E if she believes that this event happens with probability 1.[17] We say that a history $h \in \mathcal{H}$ *is consistent with event* $E \subseteq \Theta_{-i} \times \Sigma_{-i}$ if there exist a strategy $\sigma_{-i} \in \mathrm{proj}_{\Sigma_{-i}} E$ such that $h \in \mathcal{H}(\sigma_{-i})$. Abusing the notation somewhat I will write $h \in \mathcal{H}(E)$ for such histories.

Definition 25 (*Weak Certainty of event E*). A system of beliefs $\pi_i \in \Delta^{\mathcal{H}}(\Theta_{-i} \times \Sigma_{-i})$ is *weakly certain of event* $E \subset \Theta_{-i} \times \Sigma_{-i}$ if $\pi_i\left(E \mid h^0\right) = 1$

Definition 26 (*Strong Certainty in event E*). A system of beliefs $\pi_i \in \Delta^{\mathcal{H}}(\Theta_{-i} \times \Sigma_{-i})$ is *strongly certain of event* $E \subseteq \Theta_{-i} \times \Sigma_{-i}$ if $\pi_i(E \mid h) = 1$ for all $h \in \mathcal{H}(E)$

To illustrate the difference between both concepts, suppose p has a belief system π_p, that is certain of some event E, and is also certain about a smaller event $F = \{(new, \sigma_{new}), (old, \sigma_{old})\} \subset E$. That is, he is certain about what strategy each type of player d chooses (which is the required assumption in the construction of a Bayesian equilibrium in pure strategies). However, π_p may be an incorrect prediction of d's behavior. Take a history h in which p realizes that the observed history is *not* consistent with the strategies in F but it is nevertheless consistent with event E: i.e. $\{\sigma_{new}, \sigma_{old}\} \cap \Sigma_d(h) = \emptyset$ but $h \in \mathcal{H}(E)$. If π_p is *weakly certain* of event E, then after the unexpected move by d, no restrictions are imposed on the updated beliefs from history h on. In particular, he is not required to remain certain about event E, even if the observed history is consistent with it. On the other hand, if π_p is *strongly certain* about event E, he would realize his beliefs about event F were wrong, but his updated beliefs would remain certain about event E. In a way, the concept

---

$\theta \in \{new, bad\}$ if for all histories $h \in \mathcal{H}_d$ and all strategies $\hat{\sigma}_d \in \Sigma_d$ we have $W_\theta^{\pi_d}(\sigma_d \mid h) \geq W_\theta^{\pi_d}(\hat{\sigma}_d \mid h)$.

[17] When the event E is also true we say that the type *knows* E. This admits the possibility that an agent believes with probability one an event that is indeed false. In static games, because the game ends right after the payoffs are realized, there is no substantive difference between *certainty* and *knowledge*. In dynamic games the situation is more subtle, since an agent's beliefs may be proven wrong (or refuted) by the observed path of play. Because of this feature, the literature has focused on the concept of certainty ((Ben-Porath 1997), (Battigalli and Siniscalchi 1999), (Penta 2011; Penta 2012)) instead of knowledge, for dynamic games.

of strong certainty is similar to an agent that knows that event E is true, and her updated beliefs should respect it as a "working hypothesis" ((Battigalli and Siniscalchi 2002))

These two different notions of certainty will give rise to two different notions of rationalizability. Define the set of sequentially rational outcomes $R_i \subset \Theta_i \times \Sigma_i$ as

$$R_i = \left\{(\theta_i, \sigma_i) : \ \sigma_i \in SBR_{\theta_i}(\pi_i) \text{ for some } \pi_i \in \Delta^{\mathcal{H}}(\Theta_{-i} \times \Sigma_{-i})\right\}. \tag{3.27}$$

The set R_i gives all the strategies and payoff types such that σ_i is the sequential best response to some system of beliefs.

I will now formally follow the iterative procedure of (Battigalli and Siniscalchi 2003). For a given set $E \subset \Theta_{-i} \times \Sigma_{-i}$ write $\mathbf{W}_i(E) \subset \Delta^{\mathcal{H}}(\Theta_{-i} \times \Sigma_{-i})$ to be the set of of beliefs π_i that are weakly certain of E. Analogously, define $\mathbf{S}_i(E) \subset \mathbf{W}_i(E)$ for the set of beliefs that are strongly certain of it. I will denote $WCR_i^k \subset \Theta_i \times \Sigma_i$ and $SCR_i^k \subset WCR_i^k$ as the sets of type-strategy pairs for agent i that are consistent with k rounds of mutual weak (strong) certainty of rationality. For $k = 0$, define

$$WCR_i^0 = SCR_i^0 = R_i.$$

For $k > 1$, define iteratively:

$$WCR_i^k := \left\{(\theta_i, \sigma_i) : \begin{cases} (1) : & (\theta_i, \sigma_i) \in WCR_i^{k-1} \\ (2) : & \exists \pi_i \in \mathbf{W}_i\left(WCR_{-i}^{k-1}\right) : \sigma_i \in SBR_{\theta_i}(\pi_i) \end{cases}\right\} \tag{3.28}$$

$$SCR_i^k := \left\{(\theta_i, \sigma_i) : \begin{cases} (1) : & (\theta_i, \sigma_i) \in SCR_i^{k-1} \\ (2) : & \exists \pi_i \in \mathbf{S}_i\left(SCR_{-i}^{k-1}\right) : \sigma_i \in SBR_{\theta_i}(\pi_i) \end{cases}\right\}. \tag{3.29}$$

We start with beliefs that are weakly (strongly) certain of event $E = R_{-i}$ and then we proceed with an iterative deletion procedure, in which the set agent i is weakly (strongly) certain about is the set $E = WCR_i^{k-1}$ and similarly for strong certainty. Finally, the sets of *weak and strong rationalizable outcomes* is defined as

$$WCR_i^\infty = \bigcap_{k \in \mathbb{N}} WCR_i^k \tag{3.30}$$

$$SCR_i^\infty = \bigcap_{k \in \mathbb{N}} SCR_i^k \tag{3.31}$$

The sets WCR_i^∞,$SCR_i^\infty \subset \Sigma_i$ are the sets of strategies for i that are consistent with him having weak (strong) common certainty of rationality. I will denote $\mathcal{B}_i^{WR}$ and $\mathcal{B}_i^{SR}$ as the sets of weak and strong rationalizable beliefs for p, respectively

$$\mathcal{B}_i^{WR} := \Delta^{\mathcal{H}}\left(WCR_{-i}^\infty\right) \text{ and } \mathcal{B}_i^{SR} := \Delta^{\mathcal{H}}\left(SCR_{-i}^\infty\right). \tag{3.32}$$

I will say that a strategy-belief pair (σ_i, π_i) is θ_i*–strong rationalizable* (or simply p–strong rationalizable for the case $i = p$) whenever $\pi_i \in \mathcal{B}_i^{SR}$ and $\sigma_i \in SBR_{\theta_i}(\pi_i)$. A history h is θ_i–strong rationalizable whenever $h \in \mathcal{H}(\sigma_i)$ for some weak rationalizable pair (σ_i, π_i). I will refer to such pairs as a θ_i*–strong rationalizations* of h. I also define the analogous notions for weak rationalizability.

3.4.5 Discussion

In the above we have characterized the multiplicity of equilibria in the static game and have established the setup of the repeated game including belief systems and notions of rationalizability. The next section will turn to robust implementation. We briefly connect the concepts now, and argue that equilibrium refinements are not robust to a variety of perturbations we might think of.

First, and most importantly for our applications is the one considered in this paper which is robustness to *strategic uncertainty.* Recall that the static game had multiple equilibria. The dynamic game only exacerbates this problem via classical folk theorem like arguments.[18] This suggests that in order for either to form predictions or to make policy recommendations, some equilibrium refinement is needed, as selecting optimal or efficient equilibria. In some contexts this may be reasonable: e.g. if agents could meet and agree upon a desired outcome before the game started and are able to decide both the expected behavior by all agents, the punishments that should be sanctioned to deviators, subject to the constraint that these should be self-enforceable. However, in this environment the public has no reason to agree with the time inconsistent type and, as such, selecting the optimal equilibrium seems suspect. A second limitation of equilibrium refinements is that they are very sensitive to common knowledge assumptions about the payoff structure of the game. If we allow the set of feasible payoff structures to satisfy a richness condition, [19]and we pick a Nash equilibrium of the game and the belief systems that support it, then arbitrarily small perturbations on the beliefs may pick *any* other weak rationalizable outcome as the unique equilibrium of the perturbed game (e.g., (Weinstein and Yildiz 2012; Weinstein and Yildiz 2007; Penta 2012)). Under these assumptions, the only concept that is robust to small perturbations of beliefs is weak rationalizability, and hence only predictions that hold for all weak rationalizable strategy profiles are robust to these perturbations. [20]However, the richness assumption may be too a stringent condition for our robustness exercise, since we are ultimately interested in modeling robustness to strategic uncertainty. Strong rationalizability, being a stronger solution concept may not be robust to all of these perturbations in payoff structures, but we briefly study some in Section 3.7 how to create policies that are robust to richer payoff type spaces.

One of the main implications of strong rationalizability is that agents can be convinced at

[18] See (Mailath and Samuelson 2006) for an exhaustive review on these topics.

[19] Formally, for every strategy σ_i there exist a type $\hat{\theta}_i(\sigma_i) \in \hat{\Theta}_i$ such that σ_i is conditionally dominant for type $\hat{\theta}_i(\sigma_i)$ at every history consistent with it: i.e. $W_{\theta_i}(\sigma_i, \sigma_{-i} \mid h) > W_{\theta_i}(\hat{\sigma}_i, \sigma_{-i} \mid h)$ for all $\hat{\sigma}_i \in \Sigma_i, \sigma_{-i} \in \Sigma_{-i}, h \in \mathcal{H}_i(\sigma_i)$.

[20] (Weinstein and Yildiz 2012) show that when we relax the restriction that all players know their own type at the beginning of the game (and never abandon this belief), then the only robust solution concept is normal form interim correlated rationalizability (ICR), extending their previous result on static games ((Weinstein and Yildiz 2007)).

some histories that certain payoff types are not consistent with the history observed. Suppose that p reaches a history that is not consistent with both strong common certainty of rationality and $\theta = old$, but it is consistent with $\theta = new$. Strong common certainty of rationality implies that at these histories p must be certain that $\theta = new$ for all strong rationalizable continuation histories; it becomes *common knowledge* that $\theta = new$, and the game transforms in practice to a game of complete information.[21] When this happens, we will say that $\theta = new$ has achieved *full or strong separation* from $\theta = old$. This is one of the key ingredients of robust reputation formation: the reformed decision maker can gain reputation by taking actions that $\theta = old$ decision maker would never take, or that at least would be very costly for her.

3.5 Robust Implementation

This section introduces the notion of robust implementation to a given set of restrictions on p's beliefs (subsection 3.5.1) and solves for the robust implementing policies for two important benchmarks: weak and strong rationalizable beliefs. Focusing on strong rationalizable implementation, I characterize the optimal strong rationalizable implementation by solving a recursive dynamic contracting problem with a single promise-keeping constraint. Moreover, for histories where robust separation has not occurred, the relevant reputation measure for d is the *implied spot opportunity cost or sacrifice* for $\theta = old$ of playing $r_{\tau-1}$, so only the immediate previous period matters in terms of building partial reputation. I show that on the outcome path of the optimal robust policy, $\theta = new$ gets both partial gains and (endogenous) losses of reputation until robust separation is achieved. After this, the game essentially becomes one with complete information.

3.5.1 Definition

The decision maker has some information about $p's$ beliefs or may be willing to make some assumptions about them. She considers that $p's$ possible beliefs lie in some subset $\mathcal{B}_p \subset \Delta^{\mathcal{H}}(\Theta_d \times \Sigma_d)$. Write $SBR_p(\mathcal{B}_p) = \bigcup_{\pi_p \in \mathcal{B}_p} SBR_p(\pi_p) \subset \Sigma_p$ as the set of all sequential best responses to beliefs in $\mathcal{B}_p$. We will say that a strategy σ_d is a *robust implementation of trust in* $\mathcal{B}_p$ when it induces p to trust d at all $\tau = 0, 1, 2, ...$, provided d *knows* that **(1)** $p's$ beliefs are in $\mathcal{B}_p$ and **(2)** p is sequentially rational.

Definition 27 (*Robust Implementation*). A strategy $\sigma_d \in \Sigma_d$ robustly implements trust in $\mathcal{B}_p$ if,

[21] If at some continuation history p observes behavior that is inconsistent with $\theta = new$ playing a strongly rationalizable strategy, p abandons the assumption of strong common certainty of rationality, which then allows him to believe that $\theta = old$ after all. When this happens, we apply the "best-rationalization principle" as in (Battigalli and Siniscalchi 2002). It states that whenever p arrives at such a history, she will believe that there are at least $k-$rounds of strong common certainty of rationality, with k the highest integer for which the history is consistent with k rounds of strong rationalizability.

for all histories $(h^{\tau}, c_{\tau}) \in \mathcal{H}_p(\sigma_d)$ we have

$$a^{\sigma_p}(h^{\tau}, c_{\tau}) = 1 \text{ for all } \sigma_p \in SBR_p(\mathcal{B}_p)$$

Under the assumptions on the stage game, for a given belief system π_p , its sequential best response will be generically unique. Therefore, if d knows both that p is rational and that he has beliefs π_p, then she can predict the strategy that p will choose.

3.5.2 Weak Rationalizable Implementation

I begin with the most lax notion of rationalizability at our disposal – weak rationalizability; i.e. $\mathcal{B}_p = \mathcal{B}_p^{WR}$. Here I show that this notion of rationalizability delivers a rather stark and, in some sense, negative result: only by eliminating the emergency action entirely can d robustly implement trust. Since the public cares only about the decision maker's strategy, the multiplicity of commitment costs that are consistent with common certainty of rationality allows for the following weak rationalizable beliefs: believe d is rational only if she takes one of these specified decisions but is actually thought to be irrational if she takes any other. Then it becomes impossible to implement trust in both belief types, unless d gets rid of the emergency action altogether.[22]

Proposition 28. *The unique robust implementing policy in $\mathcal{B}_p = \mathcal{B}_p^{WR}$ involves $c_{\tau} = \infty$ (i.e.prohibiting the emergency action) every period.*

Thus, the result is that weak rationalizability is too weak a concept to be used for our purposes. After an unexpected commitment cost choice, p could believe d to be *irrational* and never trust d again unless r is removed. This sort of reasoning does not take into account a restriction that, say, if p could find some other beliefs under which d would be rational, then this now becomes p's working hypothesis. This is precisely the notion of strong rationalizability, which I explore below.

3.5.3 Strong Rationalizable Implementation

The next three subsections present the main results of the paper in which I characterize the *optimal strong rationalizable implementation.* I will show that an optimal robust implementing strategy corresponds to the sequential best response to some strong rationalizable system of beliefs. In that sense, an optimal robust implementing strategy will be equivalent to finding the most pessimistic beliefs that d could have about p's behavior, that is consistent with common strong certainty of rationality. Most importantly, I will also show that any optimal robust strategy will be in fact the *min-max* strategy for d, delivering the best possible utility that d can guarantee at any continuation history, regardless of her system of beliefs.

[22]This argument extends to any game of *private values* with multiple weak rationalizable outcomes. A Bayesian game is of if utility for each agent depends only on their own payoff parameter, and not about the other agents payoffs. Formally, is of private values if for all mathnorma$\theta \in \Theta \equiv \times_{i=1}^{I} \Theta_i$

To simplify notation, denote $\Sigma_i^{SR} = \text{proj}_{\Sigma_i} SCR_i^{\infty}$ for the set of extensive form rationalizable strategies for agent i. Abusing the notation somewhat I will also write $\Sigma_\theta^{SR} = \text{proj}_{\Sigma_d}\left\{\left(\hat{\theta}, \hat{\sigma}_d\right) \in SCR_d^{\infty} : \hat{\theta} = \theta\right\}$ to represent the set of extensive form rationalizable strategies for type. The goal is to characterize *optimal robust strategies: i.e.* robust strategies that maximize the expected (ex-ante) utility for d, at $\tau = 0$.

In the repeated game setting, in order to guarantee $p's$ trust we need to make the utility of trust to be greater than the outside option value $\underline{u}_p$ for all strong rationalizable beliefs. Since p is myopic and does not care directly about the commitment cost payed, the only relevant object to determine his expected payoff is the way he expects d to react to shocks at time τ. Define then the set of all strong rationalizable policy functions

$$\mathbf{R}(h^\tau, c_\tau) = \left\{r(\cdot) = r^{\sigma_d}(h^\tau, c_\tau, \cdot) \text{ for some } \sigma_d \in \Sigma_d^{SR}(h^\tau, c_\tau)\right\} \tag{3.33}$$

Define also $\mathbf{R}_\theta(h^\tau, c_\tau) \subset \mathbf{R}(h^\tau, c_\tau)$ as those policy functions that are $\theta-$rationalizable. Is easy to show that $a^{\sigma_p}(h^\tau, c_\tau)$ for all strong rationalizable strategies if and only if [23]

$$\int r(z_\tau) U_p(z_\tau) f(z_\tau) dz_\tau \geq \underline{u}_p \text{ for all } r(\cdot) \in \mathbf{R}(h^\tau, c_\tau)$$

which can be rewriten in a single condition as:

$$\underline{V}(h^\tau, c_\tau) := \min_{r(\cdot)\in\mathbf{R}(h^\tau, c_\tau)} \int r(z_\tau) U_p(z_\tau) f(z_\tau) dz_\tau \geq \underline{u}_p \tag{3.34}$$

i.e. the worst rationalizable payoff for p must be higher than the reservation utility. In Appendix 44 we show that $\mathbf{R}(h^\tau, c_\tau)$ and $\mathbf{R}_\theta(h^\tau, c_\tau)$ are compact sets and the objective function in the minimization problem of 3.34 is continuous in the product topology, which makes $\underline{V}(h^\tau, c_\tau)$ a well defined object.

Then, the optimal robust strategy $\sigma_{new}^* = \{c^*(\cdot), r^*(\cdot)\}$ for type $\theta = new$ is the strategy that solves the following programming problem:

$$W_{new}^* = \max_{\{c^*(\cdot), r^*(\cdot)\}} \mathbb{E}\left\{(1-\beta)\sum_{\tau=0}^{\infty} \beta^\tau \left[U_p(z_\tau) - c^*(h^\tau)\right] r^*(h^\tau, c^*(h^\tau), z_\tau)\right\} \tag{3.35}$$

subject to

$$\underline{V}(h^\tau, c^*(h^\tau)) \geq \underline{u}_p \text{ for all } h^\tau \in \mathcal{H}(\sigma_\theta^*) \tag{3.36}$$

and analogously for W_{old}^*. The goal for the rest of the paper is to characterize the solution to 3.35. optimal robust strategy for the reformed payoff type $\theta = new$. Note that restriction 3.36 fully

[23] Because of Fubini's theorem, we can write $\mathbb{E}^{\pi_p}\left[r^{\sigma_d}(h^\tau, z) U_p \mid h^\tau, c_\tau\right] = \mathbb{E}_z\left\{\mathbb{E}_{\hat{\sigma}_d}^{\pi_d}\left[r^{\hat{\sigma}_d}(h^\tau, z_\tau) \mid h^\tau, c_\tau\right] U_p\right\}$ which corresponds to the expected value over a mixed strategy $\hat{\sigma}_d$ with expected policy $\mathbb{E}\left[r^{\hat{\sigma}_d}(h^\tau, c_\tau)\right] = \mathbb{E}_{\hat{\sigma}_d}\left[r^{\hat{\sigma}_d}(h^\tau, z_\tau) \mid h^\tau, c_\tau\right]$. Then, the minimum rationalizable payoff of trusting is the one that assigns probability 1 to the worst rationalizable policy function $r(\cdot)$ from the viewpoint of p, on that history

incorporates the robustness restriction into our programming problem. Theorem 53 shows that Σ_θ^{SR} is a compact set, and so are the subsets $\Sigma_\theta^{SR}(h) \subset \Sigma_\theta^{SR}$ of history consistent strategies, for all θ. This implies that existence of payoff functions $\underline{W}_\theta, \overline{W}_\theta : \mathcal{H}_d \to \mathbb{R}$ such that, for all $h \in \mathcal{H}_d$ and $\theta \in \{new, old\}$

$$\underline{W}_\theta(h) \leq W_\theta^{\pi_d}(\sigma_d \mid h) \leq \overline{W}_\theta(h) \text{ for all } \pi_d \in \mathcal{B}_\theta^{SR}, \sigma_d \in SBR_\theta(\pi_d). \tag{3.37}$$

I will refer to $\underline{W}_\theta(\cdot)$ and $\overline{W}_\theta(\cdot)$ as the *best* and *worst strong rationalizable payoffs* for type θ. I will also write $\underline{\mathbb{W}}_\theta := \underline{W}_\theta(h^0)$ and $\overline{\mathbb{W}}_\theta = \overline{W}_\theta(h^0)$ for the ex-ante worst (and best) rationalizable payoff, from $\tau = 0$ perspective. The first result relates these bounds to robust implementation. Any optimal robust policy is extensive form rationalizable (i.e. it corresponds to the best response of some rationalizable beliefs) and delivers the worst rationalizable payoff $\underline{W}_\theta(h)$ at all histories and types $\theta \in \{new, old\}$.

Lemma 29. *Let σ_θ^* be the optimal robust strategy for type θ. Then $\sigma_\theta^* \in \Sigma_\theta^{SR}$, with rationalizing belief $\underline{\pi}_\theta \in \mathcal{B}_\theta^{SR}$. Moreover, for all histories $h \in \mathcal{H}_d$*

$$W_\theta(\sigma_d^* \mid h) = \underline{W}_\theta(h),$$

i.e., the optimal robust policy delivers the worst strong rationalizable payoff at all histories.

Lemma 29 implies a very important corollary. The optimal robust strategy is the min-max strategies for type $\theta \in \{old, new\}$ (as in (Mailath and Samuelson 2006)) across all beliefs that are consistent with common strong certainty of rationality. The beliefs $\underline{\pi}_\theta$ corresponds to the min-max beliefs for type θ, the most pessimistic beliefs that type θ can have about the strategy that p may be playing. Therefore, at history h^τ, the optimal robust policy gives the *best payoff that type θ can guarantee herself,* regardless of her beliefs, as long as p plays some strong rationalizable strategy. This further implies that the value of program 3.35 at any history satisfies

$$W_\theta^* = \underline{\mathbb{W}}_\theta \tag{3.38}$$

Note here that the worst rationalizable payoff does not coincide with the payoff of the worst Bayesian equilibrium of the extensive form game. Common strong certainty of rationality, strictly speaking, is neither a stronger nor weaker solution concept than Bayesian equilibrium.[24].

[24] Applying (Aumann and Brandenburger 1995) to the interim normal form game, for a particular Bayesian equilibria to be the predicted outcome of the game, we need the common prior assumption (i.e. both players know $\pi = \Pr(\theta = new)$) together with weak common *knowledge* of rationality and beliefs (i.e. weak common certainty, plus the requirement that the beliefs are correct). While the common certainty of rationality is weaker than strong certainty, this characterization implies a much stronger condition. Agents have common knowledge about the *strategies* that each other will play and these beliefs must be correct.

3.5.4 Observed Sacrifice and Strong Rationalizable Policies

The program 3.35 may seem complicated, because of the potentially complex history dependence of the set of strong rationalizable policies $\mathbf{R}_\theta(h^\tau, c_\tau)$. Since $\mathbf{R}(h^\tau, c_\tau) = \mathbf{R}_{new}(h^\tau, c_\tau) \cup \mathbf{R}_{old}(h^\tau, c_\tau)$, characterizing these sets will determine the shape of $\underline{V}(h^\tau, c_\tau)$. I will derive the restrictions that strong rationalizability, together with the observed history, impose on the set of policy functions $r(\cdot)$ that p may expect, and show that we only need to know the previous period implied opportunity cost payed by type θ, to be able to characterize the set $\mathbf{R}_\theta(h^\tau, c_\tau)$. In this sense, the set of strong rationalizable policies $\mathbf{R}_\theta(h^\tau, c_\tau)$ is Markovian, with a state variable that is observable by all agents in the game.

Consider a history $(h^\tau, c_\tau) \in \mathcal{H}_p$ observed by agent p. Suppose first that p hypothesizes that d is of payoff type θ, and that history h^τ is such that $r_{\tau-1} = 0$ and $U_{\theta,\tau-1} - c_{\tau-1} > 0$, so that d played the normal action in the previous period, but she would have preferred to play the emergency action, if she was of type θ. Let $h^\tau(r=1)$ be the continuation history had d chosen $r_{\tau-1} = 1$ instead. Then, a $\theta-$rationalizable pair (σ_d, π_d) is consistent with the observed h^τ if and only if

$$\beta W_\theta^{\pi_d}(\sigma_d \mid h^\tau) \geq (1-\beta)(U_{\theta,\tau-1} - c_{\tau-1}) + \beta W_\theta^{\pi_d}(\sigma_d \mid h^\tau(r=1)). \tag{3.39}$$

To interpret condition 3.43, define first $S_{\theta,\tau-1} := U_{\theta,\tau-1} - c_{\tau-1} > 0$ as the *sacrificed spot utility* for type θ of playing $r_{\tau-1} = 0$ instead of $r_{\tau-1} = 1$. Also, let

$$\mathbf{NPV}_\theta^{\pi_d}(\sigma_d \mid h^\tau) := \frac{\beta}{1-\beta}\left[W_\theta^{\pi_d}(\sigma_d \mid h^\tau) - W_\theta^{\pi_d}(\sigma_d \mid h^\tau(r=1))\right] \tag{3.40}$$

denote the *net present continuation value* under pair (σ_d, π_d) of having played $r_{\tau-1} = 0$. This formulation gives a very intuitive characterization of condition 3.39:

$$S_{\theta,\tau-1} \leq \mathbf{NPV}_\theta^{\pi_d}(\sigma_d \mid h^\tau) \tag{3.41}$$

i.e. it would have been optimal for type θ to "invest" an opportunity cost of utils yesterday (by not following the spot optimal strategy) only if she expected a net present value that would compensate her for the investment. We can further refine condition 3.39 by first showing that

$$W_\theta^{\pi_d}(\sigma_d \mid h^\tau(r=1)) \geq \underline{W}_\theta(\sigma_d \mid h^\tau(r=1)) \geq \underline{\mathbb{W}}_\theta \tag{3.42}$$

Combining 3.42 with 3.39 implies a simple *necessary* condition for $\theta-$rationalizability: if (σ_d, π_d) $\theta-$ rationalizes (h^τ, c_τ), then

$$W_\theta^{\pi_d}(\sigma_d \mid h^\tau) \geq \frac{1-\beta}{\beta} S_{\theta,\tau-1} + \underline{\mathbb{W}}_\theta \tag{3.43}$$

Condition 3.43 also holds for any other history (h^τ, c_τ), where we generalize the definition of sacrificed utility as

$$S_{\theta,\tau-1} := \max_{\tilde{r}\in\{0,1\}} \left(U_{\theta,\tau-1} - c_{\tau-1}\right)\tilde{r} - \left(U_{\theta,\tau-1} - c_{\tau-1}\right) r_{\tau-1} \tag{3.44}$$

3.43 puts restrictions on $\theta-$rationalizing pairs (σ_d, π_d) (and hence over policy functions) based only on the previous period outcome, disregarding the information in the observed history up to $\tau - 1$. A striking feature of strong rationalizability is that in fact, 3.43 is also *sufficient*: whether a policy function $r(\cdot)$ pair is strong rationalizable or not depends only on the observed past sacrificed utility. Proposition 30 states the core result of this paper.

Proposition 30. *Let $(h^\tau, c_\tau) \in \mathcal{H}_p$ be $\theta-$rationalizable. Then $r(\cdot) \in \mathbf{R}_\theta(h^\tau, c_\tau)$ if and only if there exists a measurable function $w : Z \to [\underline{\mathbb{W}}_\theta, \overline{\mathbb{W}}_\theta]$ such that*

$$(1-\beta)\left[U_\theta(z_\tau) - c_\tau\right] r(z_\tau) + \beta w(z_\tau) \geq (1-\beta)\left[U_\theta(z_\tau) - c_\tau\right]\hat{r} + \beta \underline{\mathbb{W}}_\theta \tag{3.45}$$

for all $\hat{r} \in \{0,1\}, z_\tau \in Z,$ *and*

$$\int \left\{(1-\beta)\left[U_\theta(z_\tau) - c_\tau\right] r(z_\tau) + \beta w(z_\tau)\right\} f(z_\tau)\, dz_\tau \geq \frac{1-\beta}{\beta} S_{\theta,\tau-1} + \underline{\mathbb{W}}_\theta \tag{3.46}$$

Condition 3.45 is analogous to the (Abreu, Pearce, and Stacchetti 1990) notion of *enforceability*. A policy $r(\cdot)$ will be "enforceable" at some history only if we can find a continuation payoff function that enforces it on the set of strong rationalizable payoffs $[\underline{\mathbb{W}}_\theta, \overline{\mathbb{W}}_\theta]$. This argument employs the same tools and insights as in (Abreu, Pearce, and Stacchetti 1990). Condition 3.46 is the translation of condition 3.43 into this notation. It resembles a promise keeping constraint in a dynamic contracting problem: the expected value of following a rationalizable strategy σ_d at this history (given by the right hand side of 3.46) must be greater than the value implied by the implied opportunity cost paid in the previous period, which can be thought of as the utility "promised" by some rationalizable pair (σ_d, π_d). Its proof resembles closely the well known "optimal penal codes" argument in (Abreu 1988): any strong rationalizable outcome can be enforced by switching to the worst rationalizable payoff upon observing a deviation from the prescribed path of play. This means that without loss of generality, we can check whether a policy $r(\cdot)$ is $\theta-$rationalizable if it is implementable whenever type θ thinks that if she deviated, she will have to play the optimal robust policy from then on.

Proposition 30 requires the history (h^τ, c_τ) to be $\theta-$rationalizable. In order to be able to use this characterization, we need to determine whether (h^τ, c_τ) is also rationalizable. Because of Lemma 29, we know that all histories reached by the optimal robust policy for $\theta = new$ are $new-$rationalizable. Along its path, the observed history may or may not be $old-$rationalizable as well. Determining whether a history is $old-$rationalizable is equivalent to determining whether we have achieved *robust separation*: i.e. if a history is $new-$rationalizable but is not $old-$rationalizable, then p should be certain he is facing $\theta = new$ in the continuation path of the optimal robust policy. Let

$$S_\theta^{\max} := \frac{\beta}{1-\beta}\left(\overline{\mathbb{W}}_\theta - \underline{\mathbb{W}}_\theta\right) \tag{3.47}$$

be the maximum sacrifice level for type θ, that is consistent with common strong certainty of rationality. Proposition 31 gives necessary and sufficient conditions for robust separation

Proposition 31. *Take a new−rationalizable history* $(h^{\tau}, c_{\tau}) \in \mathcal{H}_p$. *Then, it is also old−rationalizable if and only if* $S_{old,k} \leq S_{old}^{\max}$ *for all* $k \leq \tau - 1$

This proposition characterizes completely the conditions for strong separation from type $\theta = old$, along the path of any strong rationalizable strategy, in particular the optimal robust one. The first result we infer from Proposition 31 is that *robust separation can never be achieved by the commitment cost decision.* (see Lemma 55 in Appendix 3.11), and hence $\theta = new$ can only separate from $\theta = old$ based only on how she reacted to the observed shocks. The second result provides a recursive characterization of robust separation: if separation has not yet occured up to period $\tau - 1$, $\theta = new$ will robustly separate from $\theta = old$ at period τ if and only if $S_{old,\tau-1} > S_{old}^{\max}$. This happens because condition 3.46 cannot be satisfied for any policy function $r(\cdot)$ and hence $\mathbf{R}_{old}(h^{\tau}, c_{\tau}) = \emptyset$. If $S_{old,\tau-1} \leq S_{old}^{\max}$, then at $\tau + 1$ the only relevant information to decide whether $h^{\tau+1}$ is *old*−rationalizable is $S_{old,\tau}$, and hence this property is markovian. Proposition 30 shares this markovian feature: the only relevant information to find the set of strong rationalizable policies $\mathbf{R}_{\theta}(h^{\tau}, c_{\tau})$ is the observed sacrifice $S_{\theta,\tau-1}$.

3.5.5 Characterization of Robust Implementation

In this subsection I will use the characterization of $\mathbf{R}_{\theta}(h^{\tau}, c_{\tau})$ of Proposition 30 to characterize the worst rationalizable payoff of trusting, $\underline{V}(h^{\tau}, c_{\tau})$. Furthermore, to solve for the optimal robust strategy, I will derive a recursive representation of the optimal robust policy, which will allow us to solve Program 3.35 with a standard Bellman equation, using the familiar fixed point techniques of (Stokey, Lucas, and Prescott 1989). Suppose that at a given a θ−rationalizable history (h^{τ}, c_{τ}), p hypothesizes he is facing type $\theta \in \{new, old\}$. Using the characterization of Proposition 30, we can calculate the minimum utility he can expect from trusting as:

$$\underline{V}_{\theta}(S_{\theta,\tau-1}, c_{\tau}) := \min_{r(\cdot),w(\cdot)} \int r(z_{\tau}) U_p(z_{\tau}) f(z_{\tau}) dz_{\tau} \tag{3.48}$$

subject to the incentive compatibility constraint:

$$(1-\beta)[U_{\theta}(z_{\tau}) - c_{\tau}] r(z_{\tau}) + \beta w(z_{\tau}) \geq (1-\beta)[U_{\theta}(z_{\tau}) - c_{\tau}] \hat{r} + \beta \underline{\mathbb{W}}_{\theta} \tag{3.49}$$

$$\text{for all } \hat{r} \in \{0,1\}, z_{\tau} \in Z$$

the "promise keeping" constraint:

$$(1-\beta)\int [U_{\theta}(z_{\tau}) - c_{\tau}] r(z_{\tau}) f(z_{\tau}) dz_{\tau} + \beta \int w(z_{\tau}) f(z_{\tau}) dz_{\tau} \geq \frac{1-\beta}{\beta} S_{\theta,\tau-1} + \underline{\mathbb{W}}_{\theta} \tag{3.50}$$

and a feasibility constraint for continuation payoffs:

$$w(z_\tau) \in \left[\underline{\mathbf{W}}_\theta, \overline{\mathbf{W}}_\theta\right] \text{ for all } z_\tau \in Z \tag{3.51}$$

At a history that is both *new* and *old*−rationalizable, the worst strong rationalizable payoff of trusting is

$$\underline{V}(h^\tau, c_\tau) = \min\left\{V_{old}(S_{old,\tau-1}, c_\tau), V_{new}(S_{new,\tau-1}, c_\tau)\right\}$$

Note that $\underline{V}(h^\tau, c_\tau)$ depends on the observed history only through the sacrifices $(S_{old,\tau-1}, S_{new,\tau-1})$, which makes the robust implementation restriction 3.36 to be markovian. The next propostition completely characterizes $\underline{V}(\cdot)$ for all *new*−rationalizable histories. If incentives between $\theta = old$ and $\theta = new$ satisfy an *increasing conflict* condition, then $\underline{V}(\cdot)$ will be an increasing function of the contemporaneous commitment cost.

Distribution $f(\cdot)$ satisfies the increasing conflict condition if $f(U_p, U_{old})$ is non-decreasing in U_{old} when $U_p < 0$ and non-increasing when $U_p > 0$

Proposition 32. *Take a new−rationalizable history $h^\tau \in \mathcal{H}$.*

1. *If $S_{old,k} \leq S_{old}^{\max}$ for all $k \leq \tau - 1$, then*

$$\underline{V}(h^\tau, c_\tau) \geq \underline{u}_p \iff V_{old}(S_{old,\tau-1}, c_\tau) \geq \underline{u}_p \tag{3.52}$$

2. *If $S_{old,k} > S_{old}^{\max}$ for some $k \leq \tau - 1$, then there is a unique strong rationalizable continuation strategy $\hat{\sigma}$, which corresponds to the repeated spot optimum; i.e.*

$$c^{\hat{\sigma}}(h^\tau) = 0 \textit{ and } r^{\hat{\sigma}}(h^\tau, z_\tau) = \begin{cases} 0 & \textit{if } U_{p,\tau} \leq 0 \\ 1 & \textit{if } U_{p,\tau} > 0 \end{cases} \tag{3.53}$$

and hence, for such stories

$$\underline{V}(h^\tau, c_\tau) = \begin{cases} \mathbb{E}_z\left[\max(0, U_p)\right] & \textit{if } c_\tau = 0 \\ \mathbb{E}_z\left[min(0, U_p)\right] & \textit{if } c_\tau > 0 \end{cases}$$

3. *Under assumption 3.5.5 $V_{old}(\cdot)$ is increasing in c_τ.*

Assumption 3.5.5 states that when p prefers $r = 0$, then states with higher utility of $r = 1$ for $\theta = old$ are more likely. Proposition 32 shows that the implementation restriction can be written as a function of $S_{old,\tau-1}$ only, which makes it the relevant reputation measure. When the implied opportunity cost payed by $\theta = old$ is higher than $S_{old}^{\max}$, the maximum net present value that she could get in the continuation game, the observed history is inconsistent with strong rationalizability (i.e. it is not *old*−rationalizable). At these histories, any system of beliefs must be strongly certain that $\theta = new$ (since there is only two types), and hence robust separation is achieved. This proposition also shows that when this happens, there is an unique strong rationalizable strategy profile, which is to play the repeated spot first best, since there are no confilct of interest between the parties, and both get their most prefered outcomes (see Lemma 56).

When $S_{old,\tau-1} < S_{old}^{\max}$, the "promise keeping" condition 3.50 is tighter for higher values of $S_{old,\tau-1}$, since only continuation strategies with a higher net present value are consistent with the observed history. Therefore, higher sacrifice makes $V_{old}(S_{old,\tau-1}, c_\tau)$ weakly lower, which in turn relaxes the robust implementation constraint 3.36 in the sequential program 3.35. This observation reinforces the notion of sacrifice being the relevant reputation measure for robust implementation program: higher values relax the implementation constraints, which increases the value of the robust policy.

The basic assumptions made about the distribution of z_τ may allow for local non-monotonicities of $V_{old}(S_{old,\tau-1}, c_\tau)$ with respect to the commitment cost c_τ. Under the increasing conflict assumption, higher commitment costs increase the minimum utility of facing $\theta = old$. Defining $\mathbf{c}(s) = \min\{c \in C : V_{old}(s, c) \geq \underline{u}_p\}$, under this assumption we have $V_{old}(S_{old,\tau-1}, c_\tau) \geq \underline{u}_p \iff c_\tau \geq \mathbf{c}(S_{old,\tau-1})$

In Appendix 3.12 I study in detail the solution $(\underline{r}(\cdot), \underline{w}(\cdot))$ to 3.52. In Proposition 61 I show that under assumption there exist a threshold $\hat{S} \in (0, S_{old}^{\max})$ such that if $S_{old,\tau-1} \leq \hat{S}$, the promise keeping constraint does not bind, and hence it is identical to the solution of 3.52 when $S_{old,\tau-1} = 0$, and $V_{old}(S_{old,\tau-1}, c_\tau) = V_{old}(0, c)$. When $S_{old,\tau-1} \in \left(\hat{S}, S_{old}^{\max}\right)$ the promise keeping constraint starts binding, making $V_{old}(S_{old,\tau-1}, c_\tau)$ strictly increasing in this interval. Figure 3-4 illustrates the results.

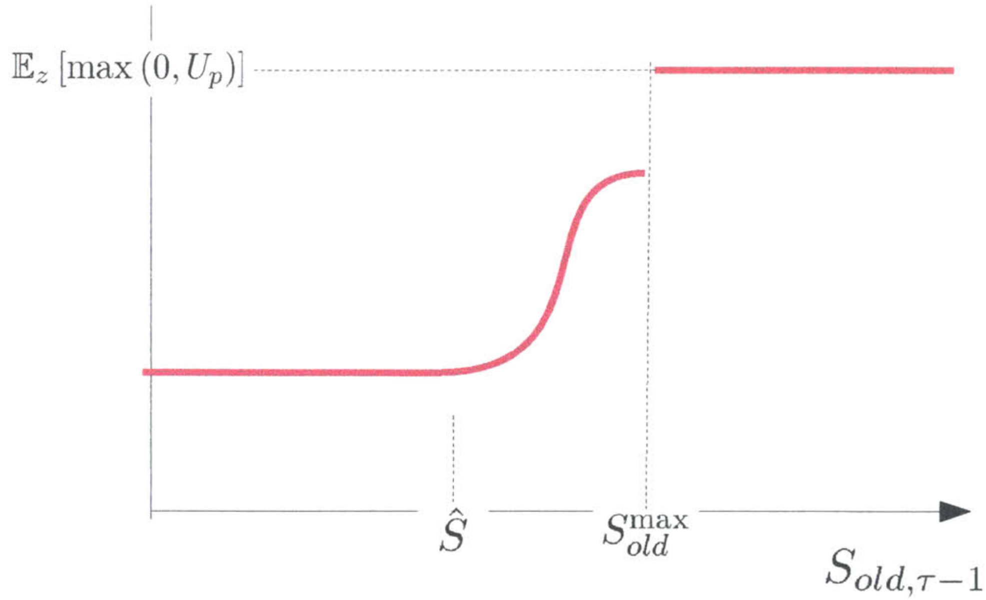


Figure 3-4: Worst Rationalizable Payoff $\underline{V}_{old}(s, c)$

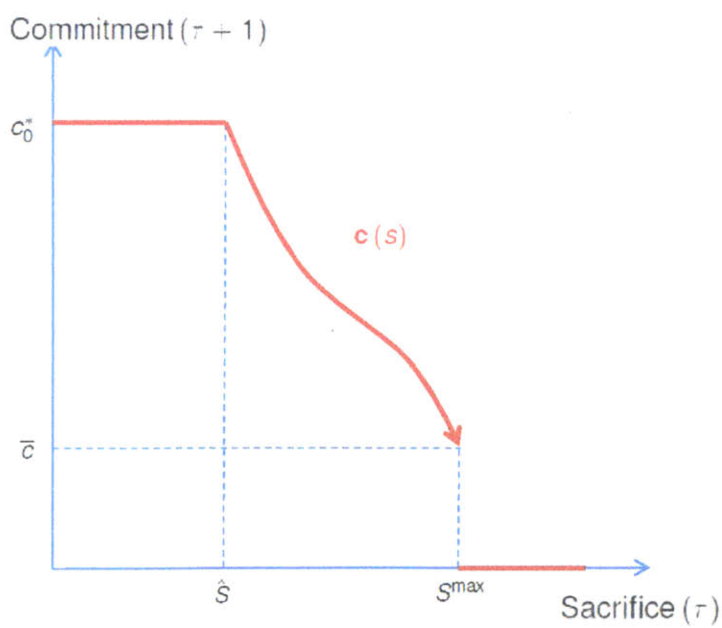


Figure 3-5: Minimum commitment cost function

Intuitively, for small observed sacrifices, p cannot discard the possibility that if $\theta = old$, she is expecting to behave the same as if no sacrifice was observed. Therefore, a robust choice for the commitment cost should prescribe exactly the same solution as in $\tau = 0$: the game basically "resets" and all reputation is lost at these histories. For intermediate sacrifices, p still cannot rule out that $\theta = old$, but can nevertheless impose some restrictions on the set of rationalizable strategies, which are stronger the bigger the sacrifice observed. When sacrifice is bigger than the maximum possible rationalizable net present value gain of any continuation value for $\theta = old$, the decision maker achieves strong separation, and hence she knows p is certain $\theta = new$ for all rationalizable continuation strategies, and therefore play the first best strategy with no commitment costs.

3.5.6 Recursive Representation of Optimal Robust Implementation

Based on the recursive characterization of the implementation restriction, in this section I finally derive a recursive representation of the optimal robust strategy σ^*_{new}. To encode the state of the problem (which depends both on the past sacrifice observed and the rationalizability of the past history) we recursively define the following process: $s_0 = 0$ and for $\tau \geq 1$:

$$s_\tau = \Gamma\left(s_{\tau-1}, c_\tau, z_\tau, r_\tau\right) := \begin{cases} \max_{\hat{r}\in\{0,1\}} \left[U_{old}\left(z_\tau\right) - c_\tau\right]\hat{r} - \left[U_{old}\left(z_\tau\right) - c_\tau\right] r_\tau & \text{if } s_{\tau-1} \leq S^{\max}_{old} \\ s_{\tau-1} & \text{if } s_{\tau-1} > S^{\max}_{old} \end{cases}$$

The state variable $s_{\tau-1}$ gives the sacrifice for $\theta = old$ as long as history h^τ is old-rationalizable. If at some τ the observed history is no longer old-rationalizable, then $s_{\tau+k} = s_\tau > S^{\max}_{old}$, so it also indicates when robust separation occurs. Because of Proposition 32 the robust implementation restriction can be written as a function of $s_{\tau-1}$ alone: $\underline{V}\left(h^\tau, c_\tau\right) \geq \underline{u}_p$ if and only if $\mathcal{V}\left(s_{\tau-1}, c_\tau\right) \geq \underline{u}_p$, where

$$\mathcal{V}\left(s, c\right) := \begin{cases} V_{old}\left(s, c\right) & \text{if } s \leq S^{\max}_{old} \\ \mathbb{E}_z\left[\max\left(0, U_p\right)\right] & \text{if } s > S^{\max}_{old} \text{and } c = 0 \\ \mathbb{E}_z\left[\min\left(0, U_p\right)\right] & \text{if } s > S^{\max}_{old} \text{and } c > 0 \end{cases} \tag{3.54}$$

With these definitions, Proposition 32 allows us to rewrite the optimal robust strategy program 3.35 as:

$$\underline{\mathbb{W}}_{new} = \max_{\{c(\cdot), r(\cdot), s_{\tau-1}(\cdot)\}} (1-\beta)\,\mathbb{E}\left\{\sum_{\tau=0}^{\infty} \beta^\tau \left[U_p\left(z_\tau\right) - c\left(h^\tau\right)\right] r\left(h^\tau, z_\tau\right)\right\} \tag{3.55}$$

$$\text{s.t. :} \quad \begin{cases} \mathcal{V}\left[s_{\tau-1}\left(h^\tau\right), c\left(h^\tau\right)\right] \geq \underline{u}_p & \text{for all } h^\tau \in \mathcal{H}\left(\sigma^*_{new}\right) \\ s_\tau\left(h^{\tau+1}\right) = \Gamma\left[s_{\tau-1}\left(h^\tau\right), c\left(h^\tau\right), z_\tau, r\left(h^\tau, z_\tau\right)\right] & \text{for all } h^{\tau+1} \in \mathcal{H}\left(\sigma^*_{new}\right) \end{cases} \tag{3.56}$$

To get a recursive formulation of $\underline{W}_{new}(h^\tau)$, let $\mathbb{B} = \left\{g : \left[\underline{U}, \overline{U}\right] \to \mathbb{R} \text{ with } g\text{bounded}\right\}$ and define the operator $T : \mathbb{B} \to \mathbb{B}$ as

$$T(g)(s) = \max_{c \in C} \int \left\{ \max_{r(z) \in \{0,1\}} (1-\beta)\left[U_p(z) - c\right] r(z) + \beta g\left[s'(z)\right] \right\} f(z)\, dz \tag{3.57}$$

subject to

$$\mathcal{V}(s,c) \geq \underline{u}_p \tag{3.58}$$

and

$$s'(z) = \Gamma\left[s, c, z, r(z)\right] \text{ for all } z \in Z \tag{3.59}$$

In Lemma 60 I show T is a contraction with modulus β. Since $\mathbb{B}$ is a complete metric space (when endowed with the sup-norm), we can use the contraction mapping theorem to show the existence of a unique function $\mathcal{W}_{new}(\cdot)$ that solves the associated Bellman equation $T(\mathcal{W}_{new})(\cdot) = \mathcal{W}_{new}(\cdot)$; which can expressed as

$$\mathcal{W}_{new}(s) = \max_{c \in C : \mathcal{V}(s,c) \geq \underline{u}_p} \int \left\{ \max_{r \in \{0,1\}} (1-\beta)\left[U_p(z) - c\right] r + \beta \mathcal{W}_{new}\left[s'(z)\right] \right\} f(z)\, dz \tag{3.60}$$

subject to 3.58.

The term inside the integral is the maximization problem that $\theta = new$ faces after having chosen c and after shock z has been realized: she faces a trade-off between short run utility $(1-\beta)\left[U_p(z) - c\right] r$ and reputation gains $\beta \mathcal{W}_{new}\left[s'(z)\right]$, which depend only on the sacrificed that would be observed at the begining of the next period. This is possible since once the commitment cost was chosen, there is no restriction linking ex-post utility in different states. The outer maximization choosing the commitment cost function corresponds to the optimal choice of the commitment cost at the begining of the period. Because of Proposition 30 all past history is completely summarized by the sacrificed observed in the previous period . Notice that s only enters the right hand side problem only through restriction 3.58, which only modifies the set of feasible commitment costs.

Proposition 33. *Let $c(s)$ and $r(s,z)$ be the policy functions associated with the Bellman equation 3.60. Then, for all $h^\tau \in \mathcal{H}(\sigma^*_{new})$*

1. $\underline{W}_{new}(h^\tau) = \mathcal{W}_{new}(s_{\tau-1})$, $c^*(h^\tau) = c(s_{\tau-1})$ *and* $r^*(h^\tau, c_\tau, z_\tau) = r(s_{\tau-1}, z_\tau)$
2. *If* $s_{\tau-1} > S^{\max}_{old}$ *then* $c^*(h^\tau) = 0$ *and* $r^*(s_{\tau-1}, z) = \underset{\hat{r} \in \{0,1\}}{argmax} U_p(z)\, \hat{r}$
3. *If* $s_{\tau-1} \leq S^{\max}_{old}$ *and Assumption 3.5.5 holds,* $c^*(h^\tau) = \mathbf{c}(s_{\tau-1})$

In the remainder of this section, I solve for the optimal robust policy $r^*(z)$ and the law of motion for the sacrifice process $s'(z)$, under the increasing conflict assumption 3.5.5. Figure 3-7 previews the shape of the optimal policy $r^*(z) = r^*(U_p, U_{old})$ over the set of states $Z = [\underline{U}, \overline{U}]^2 \subset \mathbb{R}^2$. Regions where $r^*(U_p, U_{old}) = 1$ (i.e. d takes the emergency action) are depicted in red, and $r^*(U_p, U_{old}) = 0$ in green. In the bottom we include the spot optimum strategy for $\theta = new$ and for agent p. [25] In the right margin, we draw the analogous scale for $\theta = old$.

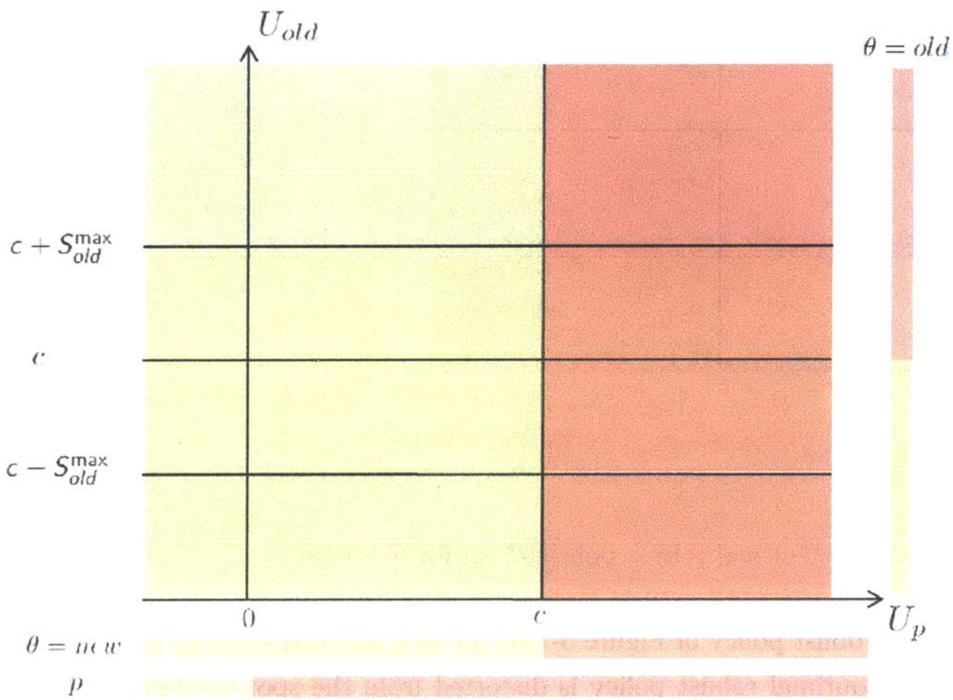


Figure 3-6: Static Best policy $r^{spot}(z)$ for $\theta = new$

[25]The spot optimal policy for p is defined as $r_p^{spot}(z) := 1 \iff U_p \geq 0$. For and for type θ we have $r_\theta^{spot}(z) := 1 \iff U_\theta \geq c$

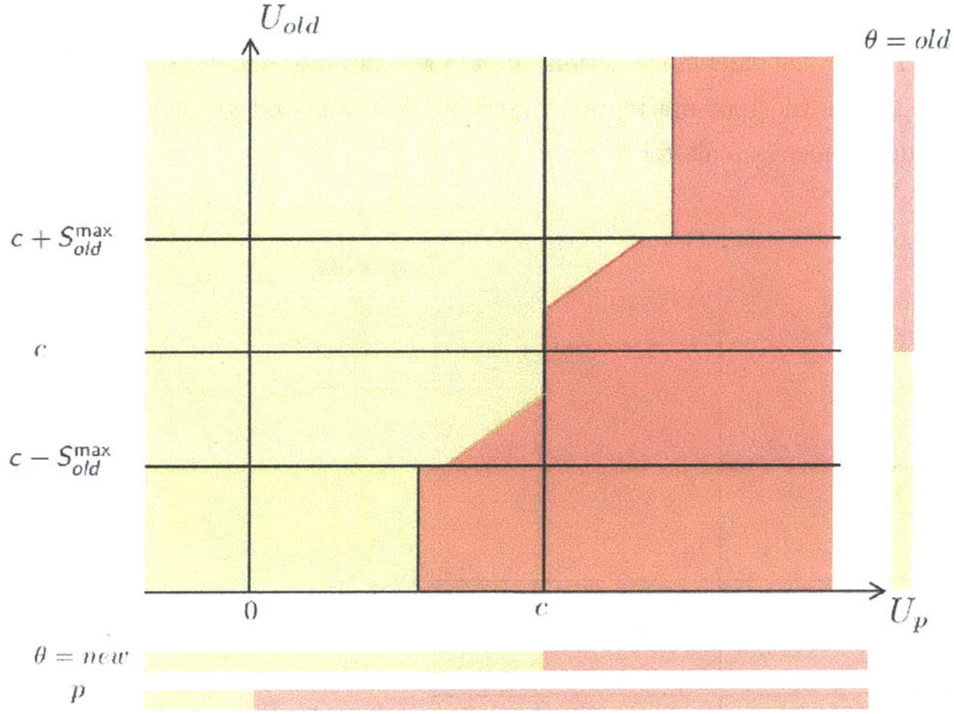


Figure 3-7: Optimal robust policy $r^*(z)$ for $\theta = new$

By comparing the optimal robust policy of Figure 3-7 to the spot optimal strategy for $\theta = new$ in Figure 3-6, we see how the optimal robust policy is distorted from the spot optimum towards actions that are spot inefficient for $\theta = old$. For example, consider the region where $U_{old} > c$, so that $r_{old}^{spot}(z) = 1$. Figure 3-7 shows then that the optimal policy over this region prescribes $r^*(U_p, U_{old}) = 0$ on a strictly larger set of states than $r_{new}^{spot}(z)$. The intuition for this phenomenon is simple: if d plays $r_\tau = 0$, then p will observe an opportunity cost payed by $\theta = old$ of $S_\tau = U_{old} - c > 0$ utils. This will result in a smaller commitment cost at $\tau + 1$ than the one implied by playing $r = 1$ and reseting to the time $\tau = 0$ robust policy from tomorrow on. When the relative reduction in future commitment costs is big enough (i.e. when observed implied sacrifice for $\theta = old$) then the optimal strategy will be to choose $r = 0$. In the rest of this section I will formally characterize both the robust policy $r^*(z)$ and the next period sacrifice $s'(z)$, which governs the reputation formation process.

First, take the region $R_1 = \{z \in Z : U_{old} > c + S_{old}^{\max}\}$, which corresponds to the upper-most horizontal strip of Figure 3-7. For any z in this region, the unique rationalizable policy for $\theta = old$ is to play $r(z) = 1$. This is because if she played $r = 0$, the implied sacrifice $S = U_{old} - c$ would be strictly greater than any potential net present value gain from switching to the best rationalizable payoff (given by $S_{old}^{\max}$). Therefore, if $\theta = new$ chooses $r^*(z) = 0$ she would strongly separate from tomorrow on, achieving the first best payoff $\mathbb{E}_z[\max(0, U_p)]$. However, if she chose $r^*(z) = 0$ then $s'(z) = 0$ and next period the commitment cost gets reset to c_0^*, getting a continuation value of

$\underline{\mathbb{W}}_{new}$. Therefore, $r^*(z) = 0$ over region R_1 if and only if

$$\beta \mathbb{E}_z [\max(0, U_p)] \geq (1-\beta)(U_p - c) + \beta \underline{\mathbb{W}}_{new} \iff$$

$$U_p \leq c + S_{new}^{\max} \tag{3.61}$$

If $U_p \leq c$ then by playing $r^*(z) = 0$ type $\theta = new$ would maximize both her spot and her continuation values, achieving strong separation from $\tau + 1$ on. Even when $U_p > c$, $\theta = new$ could still find it optimal to sacrifice spot gains for the strong separation that would be achieved in the next period. Therefore, when the time inconsistent type has a unique rationalizable strategy, the good type would optimally *invest* in reputation, sacrificing present utility to achieve strong separation in the next period.

Second, take region $R_2 = \left\{z \in Z : U_{old} \in \left(c + \hat{S}, c + S_{old}^{\max}\right)\right\}$. In this region, $\theta = old$ preferred strategy is still $r = 1$, but now $r = 0$ is also old–rationalizable. By playing $r = 0$, $\theta = new$ cannot achieve separation in the next period, but she still can decrease the commitment cost in the next period to $c_{\tau+1} = \mathbf{c}(U_{old} - c)$. Therefore, the same analysis from region R_1 applies here, with the only difference that now the continuation value will be $\mathcal{W}(U_{old} - c) < \mathbb{E}_z[\max(0, U_p)]$. Then, $r^*(z) = 0$ over region R_2 if and only if

$$U_p - c \leq \frac{\beta}{1-\beta}\left[\mathcal{W}(U_{old} - c) - \underline{\mathbb{W}}_{new}\right] := \phi(U_{old} - c)$$

where ϕ is an increasing function of the implied sacrifice $S = U_{old} - c$ of playing $r = 0$ for the time inconsistent type. As before, whenever $U_p < c$ then $r = 0$ will be optimal. When $U_p > c$ her decision will depend on two variables: the spot disutility by not choosing $r = 1$ $(U_p - c)$ and the reputation value gained by choosing $r = 0$, $\phi(U_{old} - c)$. States with very high disutility for $r = 0$ would only prescribe it as an optimal policy for states with high potential sacrifice.

Finally, study region $R_3 = \left\{z \in Z : U_{old} \in \left(c, c + \hat{S}\right)\right\}$. See that regardless of the the action, the sacrifice potential $U_{old} - c$ is too small to make the commitment cost in the next period to be smaller than its maximum possible level, c_0^*. Therefore, regardless of the policy chosen, in the next period reputation will be lost. Therefore the optimal robust policy is just the spot optimal policy: $r^*(z) = 1 \iff U_p > c$.

Is easy to see that when $U_{old} < c$, then the optimal robust policy analysis will be identical, but with the role of each policy reverted, since it will be now when $r = 1$ that sacrifice may be signaled. In Figure 3-8 we illustrate the map $s'(z)$. The dark blue areas correspond to $s'(z) = 0$ (i.e. all reputation is lost in the next period), and white regions are those in which $\theta = new$ achieves full separation from tomorrow on. In the shaded areas, lighter color illustrate higher sacrifice levels, and hence smaller commitment costs in the next period (but not zero, as in full separation).

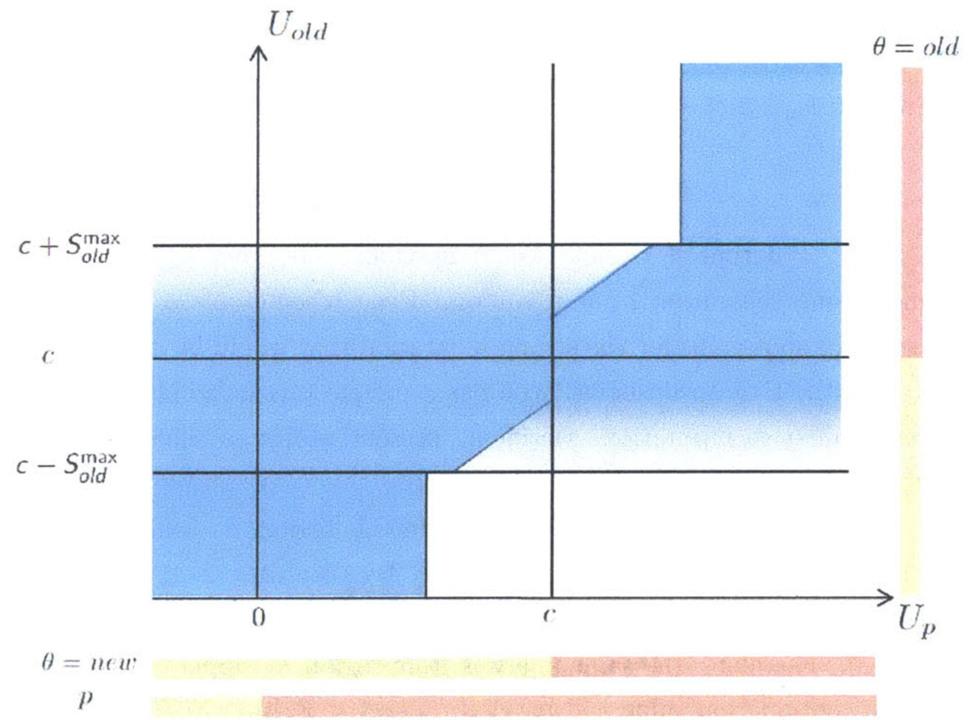


Figure 3-8: Optimal Robust Sacrifice

3.5.7 Discussion

The reason why the robust policy problem ends up being quite tractable is exactly because of the robustness condition: when having to make sure that p trusts in all histories and for all rationalizable beliefs, the *worst types* (in the sense of beliefs) that p might be facing when deciding whether to trust or not, may correspond to very different beliefs about $d's$ behavior. Because of this disconnect , we are able to separate the problems of commitment cost choice and of the optimal robust policy rule $r^*(z)$. When more restrictions are imposed (for example, a belief set $\mathcal{B} = \{\pi_{new}, \pi_{old}\}$, as in any Bayesian equilibrium), this separation will be broken.

In terms of the optimal robust policy, note that there exist regions where p and both types of decision maker would unanimously prefer certain strategy to be played, but because of reputation building motives $\theta = new$ would still want to do exactly the opposite of the unanimous optimal decision. For example, when $U_{old} > c + S_{old}^{\max}$ and $U_p \in (c, c + S_{new}^{\max})$ all agents prefer $r = 1$, but the optimal policy prescribes the normal action $r = 0$.

In the context of the capital taxation model, this would correspond with states where the marginal utility of the public good is sufficiently high for both workers and capitalists, so that both household types would agree that the ex-post optimal strategy would be to expropriate. Through the lens of our model, we can summarize the policy maker's decision by the following argument: "Even as a pro-capitalist government, I am tempted to expropriate capitalists. However, the incentives for a benevolent, time inconsistent government to expropriate would be much higher than mine.

Therefore, by not expropriating, I can show that I am in fact, not the time inconsistent type". Notice also that regardless of the beliefs that d may have, any strong rationalizable strategy of d should also achieve separation at $\tau + 1$, if she decides to play $r = 0$. This then gives a *robust prediction* about $d's$ behavior, as long it is consistent with common strong certainty of rationality.

A perhaps troubling feature of the robust policy is the impermanence of reputation gain: only the sacrifice of the previous period matters, but past sacrifices do not provide relevant information for reputation building. In the next section I find conditions on the set of beliefs $\mathcal{B}$ so that the optimal robust implementation exhibits permanent reputation gains, and hence all past sacrifices give some information about the continuation strategies that the decision maker may be planning to follow.

3.6 Basic Properties of Strong Rationalizable Implementation

In this section I will study some features of the optimal robust policy. I will first show how present potential sacrifices may affect the distribution of future sacrifices, creating "momentum" for reputation formation. I will also show that the observed sacrifice process achieves almost surely the bound $S_{old}^{\max}$. Hence, by playing the robust policy d will eventually convince p that $\theta = new$, with probability one. Moreover, the speed of convergence to the absorbing complete information stage (where p is certain that $\theta = new$) is exponential, which is also the convergence rate of the best equilibrium of the game.

I also study the asymptotic behavior of the robust policy as both the time consistent and the time inconsistent type become more patient, and show that as the discount rate approaches unity, the worst rationalizable payoff $\underline{\mathbb{W}}_{new}$ converges to the first best payoff, and hence the value of the robust policy converges uniformly to the first best payoff (e.g. for all histories). This further implies that the expected value of any strong rationalizable strategy that $\theta = new$ may follow converges to the first best payoff as well, an analog result to (Fudenberg and Levine 1989)

3.6.1 Dynamics of the optimal robust policy

We saw in the previous section that only immediate past behavior builds reputation, and past histories are irrelevant. However, it seems intuitive that there should be some momentum in reputation gaining. The basic idea is that gaining reputation at time τ will lower the commitment cost in the next period. The lowering of commitment cost will allow the reformed type to exploit the difference in ex-post payoffs between both types, which is the source of the difference between her and the time inconsistent decision maker, and therefore making that the commitment cost in $\tau+1$ should also go down even more, in a probabilistic way. However, the degree of generality I have been using so far does not allow for an easy characterization of the stochastic process followed by the commitment cost $\mathbf{c}(S_{\tau-1})$. Therefore, in this section I will show a somewhat weak momentum result, using a plausible assumption on the primitives of the model

In the static version of the game, we have

$$\Pr\left(\underset{\tilde{r}\in\{0,1\}}{\operatorname{argmax}}\,(U_{old}-\overline{c})\,\tilde{r}=0\right)>\Pr\left(\underset{\tilde{r}\in\{0,1\}}{\operatorname{argmax}}\,(U_{old}-\overline{c})\,\tilde{r}=1\right) \tag{3.62}$$

i.e. the optimal static decision rule for $\theta = old$ induces the normal action more often than the emergency action

This assumption further reinforces our interpretation of the green button strategy ($r = 0$) as the status quo: it is the strategy that both the good guy, and a trustworthy bad type would play most often. As we saw in the previous section, the main driver of reputation building is the sacrifice potential $|U_{old,\tau} - c_\tau|$, a exogenous variable for d given the commitment cost chosen. When the sacrifice potential is high is when d may decide to invest in reputation building, and moreover, conditional on observing a sacrifice, higher sacrifice potential imply lower commitment cost in the next period. While I cannot provide a characterization of the commitment cost process, I can show that the expected value of the sacrifice potential goes up when the commitment cost decreases.

Proposition 34. *If Assumption 3.6.1 holds, then*

$$s \geq s' \text{ implies } \mathbb{E}_z\left\{|U_{old} - \mathrm{c}(s)|\right\} \geq \mathbb{E}_z\left\{\left|U_{old} - \mathrm{c}\left(s'\right)\right|\right\} \tag{3.63}$$

so that after higher observed sacrifices, we expect higher potential sacrifices. If $s > s' > \hat{S}$, then the inequality in 3.63 is strict.

To show our second result on the dynamics of the optimal robust policy, I will need this very important lemma.

Lemma 35. *For all old− rationalizable histories $h^\tau \in \mathcal{H}(\sigma_d^*)$, we have that:*

$$\Pr\left(U_p > c^*\left(h^\tau\right) - S_{new}^{\max}, U_{old} < c^*\left(h^\tau\right) - S_{old}^{\max}\right) > 0 \tag{3.64}$$

and

$$\Pr\left(S_\tau \geq S_{old}^{\max}\right) > \underline{q} > 0 \text{ for all } h^{\tau+1} \in \mathcal{H}\left(\sigma_d^*\right) \tag{3.65}$$

where

$$\underline{q} := \Pr\left(U_p > c_0^* - S_{new}^{\max}, U_{old} < \overline{c} - S_{new}^{\max}\right) + \Pr\left(U_p < \overline{c} + S_{new}^{\max}, U_{old} > c_0^* + S_{old}^{\max}\right) \tag{3.66}$$

Proof. See Appendix 3.11 □

Lemma 35 is important on it's own, and states first that all the regions considered in the optimal robust policy have positive probability, and hence separation will surely occur. Moreover, I get a uniform non-zero lower bound on the probability of separating at any history, that can be easily calculated. With it, I can show its speed of convergence to strong separation

Proposition 36. *For all $\tau \in \mathbb{N}$*

$$\Pr\left(\textit{Separating before } \tau \textit{ periods}\right) > 1 - \left(1 - \underline{q}\right)^{\tau} \tag{3.67}$$

Proof. In every trial history (*new* and *old*- rationalizable) there is at least probability $\underline{q}$ of separating. Since shocks are i.i.d this implies that

$$\Pr\left(S_{old,k} < S_{old}^{\max} \text{ for all } k \leq \tau\right) < \left(1 - \underline{q}\right)^{\tau}$$

and hence $\Pr\left(\text{Separating before } \tau \text{ periods}\right) = 1 - \Pr\left(S_{old,k} < S_{old}^{\max} \text{ for all } k \leq \tau\right) = 1 - \left(1 - \underline{q}\right)^{\tau}$ □

This proposition states one of the most important results: the probability of reaching separation from the time inconsistent type is exponentially decreasing in τ. A perhaps even more important corollary is that in fact, for any belief restriction $\mathcal{B}_p$ that is consistent with strong common certainty of rationality, (i.e. $\mathcal{B}_p \subseteq \mathcal{B}_p^s$) we will also achieve separation in exponential time, the probability of separation can only be higher for any smaller belief sets. In the next section we will explore some "reasonable" restrictions we could impose, and see how the solution would be improved.

The second important corollary is that eventually $S_{old,\tau} > S_{old}^{\max}$ almost surely (and states there after separation), so that d will surely separate eventually from the time inconsistent type.

3.6.2 First Best Approximation by patient players

In this subsection the assumption $\beta_{old} = \beta_{new} = \beta$ will be significant, since we will be increasing both discount rates. I will show that as both types become more patient, the payoff of the robust policy for $\theta = new$ converges to the payoff of the stage game after separation. Now, the probability of separation for history h^{τ} will be denoted as $q\left(h^{\tau}, \beta\right)$. I first show that these probabilities are uniformly bound away from zero for all $\delta \in (0,1)$ and all rationalizable histories

Lemma 37. *Let $q\left(h^{\tau}, \beta\right)$ be the ex-ante probability of separation under the optimal robust strategy σ_d^*. Then, there exist $\hat{q} > 0$ such that $q\left(h^{\tau}, \beta\right) > \hat{q}$ for all $h^{\tau} \in \mathcal{H}\left(\sigma_d^*\right), \beta \in (0,1)$*

Proof. See Appendix 3.11 □

The previous lemma shows the existence of a number $\hat{q} > 0$ such that no matter the discount rate β , the probability of reaching separation in any *new* and *old*$-$ rationalizable histories is greater than $\hat{q}$. Since shocks are i.i.d, even if history may exhibit time dependence, we can bound the expected time of separation by a geometric random variable with success probability $\hat{q}$. Since once we reach separation, the unique rationalizable outcome is the First Best (i.e. no commitment, spot optimum policy for $\theta = new$) and the speed of convergence is exponential for this random variable, then for a sufficiently patient decision maker d, the expected payoff of the robust policy will be very close to the first best (i.e. the expected time for separation is very small in utility terms). This is what I show in the following proposition

Proposition 38. *Let* $\mathbb{E}_z\{\max(0,U_p)\}$ *be the first best payoff, corresponding to the case where p is certain that* $\theta = new$ *, and let* $\underline{\mathbb{W}}_{new}(\beta)$ *be the ex-ante expected payoff for the optimal robust policy. Then*

$$\underline{\mathbb{W}}_{new}(\beta) \to \mathbb{E}_z\{\max(0,U_p)\} \text{ as } \beta \to 1 \tag{3.68}$$

Proof. For the robust policy, we always can bound it as

$$\underline{\mathbb{W}}_{new}(\beta) \geq \mathbb{E}_\tau\{\beta^\tau \mathbb{E}_z[\max(0,U_p)]\}$$

where $\tau \sim Geom(\hat{q})$. This is true since $\mathbb{E}_z\{r(h^\tau,z)(U_p(z)-c(h^\tau))\} \geq 0$ by Lemma 56. Since we always have that the contemporaneous utility greater than zero and separation is achieved with a probability greater than $\hat{q}$ in any period, we have that this is a lower bound for the robust policy payoff. See also that

$$\mathbb{E}_\tau(\beta^\tau) = \sum_{\tau=1}^{\infty}\beta^\tau(1-\hat{q})^{\tau-1}\hat{q} = \frac{\hat{q}}{1-\hat{q}}\frac{\beta(1-\hat{q})}{1-\beta(1-\hat{q})} = \frac{\beta\hat{q}}{1-\beta(1-\hat{q})}$$

Therefore

$$\mathbb{E}[\max(0,U_p)] - \underline{\mathbb{W}}_{new}(\beta) \leq \mathbb{E}_z[\max(0,U_p)](1-\mathbb{E}_\tau(\beta^\tau)) =$$

$$= \mathbb{E}_z[\max(0,U_p)]\left(1-\frac{\beta\hat{q}}{1-\beta(1-\hat{q})}\right) = \mathbb{E}_z[\max(0,U_q)]\left(\frac{1-\beta}{1-\beta(1-\hat{q})}\right) \to 0$$

as $\beta \to 1$. □

If d is patient enough, because of discounting and the exponential speed of convergence to separation, the payoff of the robust policy will be very close to the first best payoff. Therefore, if events of distrust are sufficiently bad (as in our infinite cost interpretation of $p's$ distrust), the risk of using a weaker solution concept may be substantial, if we are not quite sure about the restrictions implied by it, while the potential increase in payoffs would be almost irrelevant if d is patient enough.

3.6.3 Restrictions on Beliefs

In certain situations, the policy maker may have more information about the public p's beliefs. I describe how this may be incorporated into the problem. Gains of reputation are not permanent, so a natural question to ask is: what restrictions on beliefs make reputation gains permanent? That is, when is $c_T \geq c_{T+k}$ for every k?

Formally, say a strategy $\sigma_d \in \Sigma_d$ exhibits *permanent reputation gains* if and only if $c^{\sigma_d}(h^{\tau+1}) \leq c^{\sigma_d}(h^\tau)$ for all histories $h^\tau, h^{\tau+1} \in \mathcal{H}(\sigma_d)$. We already know that the optimal robust strategy does not satisfy this property. The goal then is to find what type of restrictions on beliefs should we impose to get a robust implementing strategy that exhibits permanent reputation gains. Say a belief system π_d is $\hat{\sigma}_d$ - *nondecreasing* if and only if, for all histories $h^\tau, h^{\tau+1} \in \mathcal{H}(\hat{\sigma}_d)$,

$$\mathbf{NPV}_{old}^{\pi_d}(\hat{\sigma}_d \mid h^{\tau+1}) \geq \mathbf{NPV}_{old}^{\pi_d}(\hat{\sigma}_d \mid h^\tau).$$

This means that under belief π, $\theta = old$ cannot get less than what she expected in the previous period, by playing strategy $\hat{\sigma}_d$. Denote also $\mathcal{ND}(\hat{\sigma}_d) \subset \Sigma_d$ be the set of best responses to $\hat{\sigma}_d$–nondecreasing beliefs.

Proposition 39. *Take a belief restriction set $\mathcal{B}_p \subseteq \mathcal{B}_p^s$ and σ_d^* the optimal robust policy in $\mathcal{B}_p$. Then*

$$\sigma_d^* \text{ exhibits permanent reputation gains} \Longleftrightarrow \mathcal{B}_p \subseteq \mathbf{S}_p\left[\{\theta = old\} \times \mathcal{ND}(\sigma_d^*)\right].$$

That is, p is strongly certain that $\theta = old$ has σ_d^* – nondecreasing beliefs.

Proof. I show necessity. Take a old–rationalizable pair (σ_d, π_d) and a history $h^\tau \in \mathcal{H}(\sigma_d^*)$ such that $c_k = c_0^*$ for all $k < \tau - 1$ and $c_\tau < c_0^*$. This is a history where there has been only one gain in reputation so far, and which has been realized only in the last period. The fact that the commitment cost decreased has, as a necessary condition, that the observed sacrifice should be higher than certain threshold level $\hat{S}$; i.e.

$$\mathbf{NPV}_{old}^{\pi_d}(\sigma_d \mid h^\tau) \geq S_\tau \tag{3.69}$$

Also, because $h^\tau \in \mathcal{H}(\sigma_d^*)$ condition (3.69) also holds for σ_d^*.Then, the only way the commitment cost could go up in some other history $h^{\tau+s} \in \mathcal{H}(\sigma_d^*)$, is that $\mathbf{NPV}_{old}^{\pi_d}(\sigma_d \mid h^{\tau+s}) < S_\tau$. But since π_d is σ_d^*–nondecreasing, we have

$$\mathbf{NPV}_{old}^{\pi_d}(\sigma_d^* \mid h^{\tau+s}) \geq \mathbf{NPV}_{old}^{\pi_d}(\sigma_d^* \mid h^\tau) \geq S_\tau > \mathbf{NPV}_{old}^{\pi_d}(\sigma_d \mid h^{\tau+s})$$

implying that σ_d is dominated by σ_d^* at $h^{\tau+s}$. Then, the fact that the net present value is always increasing will imply that the resulting commitment cost will be always non-increasing. □

Intuitively, to get a strategy with permanent reputation gains, the assumption we need to make on $p's$ beliefs about the time inconsistent type are the following: if $\theta = old$ and we have observed that $S_{\tau-1} = S$, then the fact that she was willing to sacrifice S utils will "stick", and p will always think that $\theta = old$ will not settle for any smaller net present value. This feature of beliefs are actually pretty common in dynamic adverse selection and signaling problems.

Note that the important restriction is about $p's$ higher order beliefs: they are not about what p thinks d will do, but rather what p believes d believes about the continuation game. While working directly over system of beliefs can always be implemented, assumptions about higher order beliefs are not very transparent in this framework. In Appendix (3.9) I explore a different approach, by modeling restrictions on beliefs as type spaces, which allow the modeling of assumptions on higher order beliefs more tractable.

3.7 Extensions And Further Research

I now address several extensions of the model and strategies for future research. First, a natural alternative is to take a *legislative approach*, as in (Athey, Atkeson, and Kehoe 2005) and (Persson and Tabellini 1993). The policy maker may have delegated the commitment choice to the public. The idea here is that if one delegates the commitment cost to the public then certainly one will have robust implementation. The relevant source of uncertainty in the problem is that the public mistrusts the government. The intuition comes from contract theory: we should give control rights precisely to the party who has the first-order inability to trust. However, this will come at a cost in terms of efficiency. Specifically, the public would always put a higher commitment cost, to make the optimal policy for $\theta = old$ not drive him to indiference between trusting or not. As such, the public would increase commitment costs relative to the levels chosen by the new regime government.Therefore, it is easy to show that, if the government has the same robustness concerns, then the executive approach is superior for her in terms of welfare, given their information.

Second, we may consider robustness to not just a single time inconsistent "old type" but a multitude of time inconsistent types. Is straightforward to see that Proposition (30) would still be true for any type space Θ_d and hence the characterization of $\underline{V}(h^\tau, c_\tau)$ would now be:

$$\underline{V}(h^\tau, c_\tau) = \min_{\theta \in \Theta_d} \mathcal{V}_\theta (S_{\theta,\tau-1}, c_\tau) \tag{3.70}$$

where the function $\mathcal{V}_\theta(c, s)$ is the minimum problem in (3.52) for a given payoff type θ. Therefore, this will be equivalent to our dynamic contracting characterization of the problem above but with multiple types. In the case of a finite type set $\Theta_d = \{\theta_1, \theta_2, ..., \theta_k\}$ where we now have the vector of observed sacrifices $S_{\tau-1} = (S_{\theta_1,\tau-1}, S_{\theta_1,\tau-1}, ..., S_{\theta_k,\tau-1})$ as the state variables for the implied promise keeping constraints. The solution would exhibit separation from certain types across time, and if the other types satisfy the same assumptions made about $\theta = old$, then it will also eventually convince p about her being the time consistent type.

A third extension is to an environment in which d has an imperfect signal about $p's$ perceived incentives of the time inconsistent type. If signals are bounded and its support may be affected by some signal that d observes, then robust policy would be qualitatively identical.

Finally, looking forward, I would like to extend our analysis to situations in which there are a continuum of strategies and policies available. This will allow researchers to apply this robust modeling approach to various macroeconomic applications of interest, as the inflation setting model of subsection (3.2.2). Is easy to see how Proposition (30) would remain valid on more general models, so that the Markovian nature of reputation formation would be a very general characteristic of this type of robustness.

3.8 Conclusions

I have studied the problem of a government with low credibility. A government faces ex-post time inconsistent incentives due to lack of commitment, such as an incentive to tax capital or an incentive to allow for undesirably high levels of inflation. The government undergoes a reform in order to remove these incentives; however, the reform is successful only if the public actually believes that the government has truly reformed its ways. As such, the crux of the problem relies on the government building reputation in the eyes of the public.

After arguing that the typical approach to this problem relies on equilibrium concepts, which are highly sensitive to small perturbations about the public's beliefs, I turned to studying the problem through the lens of optimal robust policy that will implement the public's trust over any rationalizable belief that any party can hold. Focusing on robustness to all extensive-form rationalizable beliefs, I characterize the solution as well as the speed of reputation acquisition.

This is a particularly desirable property from the point of view of macroeconomic mechanism design. Equilibrium type solution concepts rely on every party knowing every higher order belief of every other party involved in the interaction. This is an extremely high dimensional object and in all likelihood it may be very difficult to believe that such an assumption really holds in settings in which one agent is trying to convince the other agent that he is not adversarial. Furthermore, equilibrium concepts rely on high dimensional belief functions off the path of play – that is, nodes or histories that may never be reached. This sort of sensitivity is problematic when advising a policy maker as small deviations in how a party truly conjectures some off the path of play belief may severely affect the policy maker's ability to obtain trust. This sort of analysis, studying optimal robust policy, can be a very powerful tool within macroeconomic policy making.

3.9 Appendix A - Type Spaces

As mentioned before, the decision maker d acting as a "policy maker", may have some information about people's beliefs about what strategy may be played by d , as well as as beliefs d may hold, which may involve assumptions about higher order beliefs (what agent i believes about j, what i believes about j's beliefs about her beliefs, and so on. As reviewed in (Bergemann and Morris 2005) the literature on epistemic game theory distinguishes between two approaches: an explicit and an implicit approach. In the *explicit approach*, beliefs are modeled as subjective probability measures over (1) the other players strategy, and (2) probability measures over the beliefs of the other agents (which are themselves probability measures over the player's own strategy), etc. In the *implicit approach*, beliefs are formed over other player strategies and "types", where a type is directly mapped to a belief over strategies and types played by the other agent. We will argue that for most applications, the implicit approach will be a more tractable modeling assumption.

Formally, the explicit approach consists on modeling people's beliefs as a hierarchy of beliefs over the types and strategies that d could play, the beliefs that d may have about p's beliefs, and so on, ad infinitum. This means that a *belief hierarchy* is a sequence of measures $(\pi_0, \pi_1, \pi_2, ...)$ where π_0 is a CPS over $\{old, new\} \times \Sigma_d$ (the payoff types of d and with the strategies that she can choose from), π_1 is a measure over the space of such probability systems (i.e. $\pi_1 \in \Delta^{\mathcal{H}}\left(\Delta^{\mathcal{H}}\left(\{old, new\} \times \Sigma_d\right)\right)$), $\pi_2 \in \Delta^{\mathcal{H}}\left(\Delta^{\mathcal{H}}\left(\Delta^{\mathcal{H}}\left(\{old, new\} \times \Sigma_d\right)\right)\right)$ and so on. The set of all possible coherent hierarchies of beliefs[26] is denoted by H_p^* , and can be shown to have desirable topological properties[27] Therefore, information about people's beliefs can be then represented as restrictions over the set of all coherent hierarchies; i.e. we can represent our information on beliefs as a subset $I \subset H_p^*$

While this approach has the advantage of being explicit about higher order beliefs, it is cumbersome to work with. Note that we have to consider beliefs over exponentially larger spaces. As an alternative, Harsanyi proposed an implicit approach ((Harsanyi 1967/68)). He suggested that one could bundle all relevant information about beliefs and payoff parameters into different "(epistemic) types" of agents, in the same way that we think about payoff types. As such, we can model the system as a Bayesian game, with a larger type space. We formalize this idea for our context:

Definition 40 (Type Space). A type space $\mathcal{T}$ is a 5-tuple $\left(T_p, T_d, \hat{\theta}(\cdot), \hat{\pi}_p(\cdot), \hat{\pi}_d(\cdot)\right)$ where T_i are sets of types for each agent, $\hat{\theta} : T_d \to \{old, new\}$ is a function that assigns to each type $t_d \in T_d$ to a payoff type $\hat{\theta}(t_d)$, and $\hat{\pi}_i : T_i :\to \Delta^{\mathcal{H}}(T_{-i} \times \Sigma_{-i})$ assigns a CPS $\hat{\pi}_i(t_i)$ over strategy and type pairs (t_{-i}, σ_{-i})

I intersect the approaches of (Bergemann and Morris 2005) and (Battigalli and Siniscalchi 1999), a type encodes it's payoff type (since d knows its type) together with his beliefs about the other agent. The set of "states of the world" Ω_{-i} that agent i form beliefs over is then the set of pairs $\omega_{-i} = (t_{-i}, \sigma_{-i})$ of strategies and types of the other agent. Unlike (Bergemann and Morris 2005) however, agents hold beliefs over other player's types and strategies as well, since in a dynamic environment these are not perfectly observed.[28]This method of representing restrictions on beliefs has the advantage of being more compact (since only first order beliefs have to be specified) and also being a natural generalization of Bayesian games. For example, suppose the information we have about p is that p thinks that $\theta = old$ could have two possible beliefs about the future play of the game: optimistic (expecting her best equilibrium to be played) or pessimistic (expecting her worst equilibrium). If this was the case, we could model this situation by simply augmenting the type space by creating two copies of the old type: an "optimistic old type" t_O and a "pessimistic old type" t_W with the same payoff parameter ($\hat{\theta}(t_O) = \hat{\theta}(t_W) = old$) but with different beliefs $\hat{\pi}_d(t_O) \neq \hat{\pi}_d(t_W)$. The type sets assigned to d would then be $T_d = \{new, t_O, t_P\}$.

In compact static games, (Mertens and Zamir 1985) and (Heifetz 1993) showed that these two approaches are equivalent: for any subset of possible hierarchies of beliefs $I \subset H^*$ there exist a type

[26]See Definition 48 in Appendix 3.10

[27]See Appendix 3.10 for a formal treatment of the topological properties of the set of all coherent hierarchies.

[28]This is the *interactive epistemic characterization* of (Ben-Porath 1997) and (Battigalli and Siniscalchi 1999). This definition also corresponds to (Penta 2012) notion of "conjectures"

space $\mathcal{T}$ that generates the exact same belief hierarchies[29] and vice versa.

In particular, if no restrictions are imposed on hierarchies (i.e. $I = H^*$) there exist a *universal type space* $\mathcal{T}^*$ which is capable of generating all possible hierarchies of beliefs. In another paper, ((Xandri 2012)) I extend (Battigalli and Siniscalchi 1999) to (non-metrizable)[30]topological spaces, to show that this is also true in a relevant class of extensive form games. In Appendix 3.10 I provide an application to our particular setting, and also a formal description of how we can make the mapping between these two approaches.

Because of this equivalence result, I will use the implicit approach throughout this paper, when modeling d's information and assumptions about $p's$ beliefs. I will further consider only *compact type spaces,* where the sets T_i are compact and Hausdorff topological spaces (with some topology) and the belief functions $\pi_i(t_i)$ are continuous in the weak convergence sense: if a sequence $t_{i,n} \to t_i$ then $\pi_i(t_{i,n})(. \mid h)$ converges in distribution to $\pi_i(t_i)(. \mid h)$ for all $h \in \mathcal{H}$. For most applications this will not be restrictive, since any type space that is "closed" is homeomorph to a subset of the universal type space $\mathcal{T}^{*}$[31] which is itself compact (see Theorem 50 in Appendix 3.10), making $\mathcal{T}$ itself compact.

For an epistemic type $t_d \in T_d$ and a strategy $\sigma_d \in \Sigma_d$ define the expected continuation value for type t_d as

$$W^{t_d}(\sigma_d \mid h) = W^{\hat{\pi}}_{\hat{\theta}}(\sigma_d \mid h) \text{ with } \hat{\theta} = \hat{\theta}(t_d), \hat{\pi} = \hat{\pi}_d(t_d) \tag{3.71}$$

Likewise, given a type $t_p \in T_p$ and a strategy $\sigma_p \in \Sigma_p$ define p's expected value as

$$V^{t_p}(\sigma_p \mid h) = V^{\hat{\pi}}(\sigma_p \mid h) \text{ where } \hat{\pi} = \hat{\pi}(t_p)$$

Also, we write $SBR(t_i) = SBR_{\hat{\theta}(t_i)}[\hat{\pi}(t_i)]$. An agent is then *sequentially rational* if the strategy she chooses is a sequential best response to her beliefs: i.e. $\sigma_i \in SBR(t_i)$ The interactive epistemic representation of types allows us to easily write this assumption; as the subset of *sequentially rational states* $R_i \subset T_i \times \Sigma_i$ defined as:

$$R_i := \{(t_i, \sigma_i) \in T_i \times \Sigma_i : \sigma_i \in SBR_i(t_i)\} \tag{3.72}$$

for $i \in \{d, p\}$. We write $\Sigma_i^*(\mathcal{T}) \subseteq \Sigma_i$ as the set of all sequentially rational strategies.

Definition 41 (Robust Implementation). Given a type space $\mathcal{T} = \left(T_d, T_p, \hat{\theta}, \hat{\pi}_p, \hat{\pi}_d\right)$ we say that a strategy σ_d is a *robust implementation of trust* if and only if for all histories $(h^\tau, c_\tau) \in \mathcal{H}_p(\sigma_d)$,

[29] Formally, there is always a *belief morphism* between both types of spaces, as studied in Appendix 3.10

[30] This is a relevant extension, since a very large class of relevant games in economics cannot be modeled with metrizable type spaces. For example, any infinitely repeated dynamic game with a continuous strategy space (such as Cournot duopoly, or most macro applications) are not metrizable.

[31] This is the main lesson from the Universal Type Space Theorem of (Mertens and Zamir 1985),(Heifetz 1993) and (Brandenburger and Dekel 1993). See Theorem 50 in Appendix 3.10

all $t_p \in T_p$ and all $\sigma_p \in SBR(t_p)$ we have $a^{\sigma_p}(h^\tau, c_\tau) = 1$

Besides the information we have about beliefs (modeled by a type space $\mathcal{T}$) we might also know (or be willing to assume) some common certainty restrictions on agents beliefs. Following the construction of subsection 3.4.4 we can extend the definitions of weak and strong rationalizable sets to type spaces, where the sets $WCR_i^k(\mathcal{T})$ and $SCR_i^k(\mathcal{T}) \subset T_i \times \Sigma_i$ correspond to all the weak and strong rationalizable pairs. In Proposition 53 we adapt the result of (Battigalli 2003) to show that for compact type spaces, these sets are compact, which also implies that the set of weak and strong rationalizable sets $WCR_i^\infty(\mathcal{T})$ and $SCR_i^\infty(\mathcal{T})$ are non-empty, compact subsets of T_i. Together with the upper hemicontinuity of the sequential best response correspondence, this implies that the sets of weakly and strongly rationalizable strategies

$$\Sigma_i^w := SBR_i\{\mathbf{W}[WCR_i^\infty(\mathcal{T})]\}, \Sigma_i^s := SBR_i\{\mathbf{S}[SCR_i^\infty(\mathcal{T})]\} \tag{3.73}$$

are compact with respect to the product topology. With this formulation, we can work with common certainty assumptions (of rationality or other assumptions about beliefs) and still retain the type space representation we have been considering. When besides the restrictions on beliefs modeled by the type space $\mathcal{T}$, d is also willing to make assumptions about common certainty of rationality, this can be thought of as refining the type space by getting a subspace $\hat{\mathcal{T}} \subset \mathcal{T}$, formed only by types that have survived the iterative deletion procedure just described. That is $t_i \in \hat{T}_i = WCR_i^\infty(\mathcal{T})$ if we use common weak certainty, and analogously with $SCR_i^\infty(\mathcal{T})$ for strong certainty.

With some abuse of notation, we will denote the type space resulting of this refinement as $\hat{\mathcal{T}} = \mathbf{WCR}^\infty(\mathcal{T})$ and $\hat{\mathcal{T}} = \mathbf{SCR}^\infty(\mathcal{T})$ for the case of strong certainty. When a type space $\mathcal{T}$ can be written as it's own weak rationalizable refinement (i.e. $\mathcal{T} = \mathbf{WCR}^\infty(\mathcal{T})$) we will say that $\mathcal{T}$ is consistent with weak common certainty of rationality, and analogously for strong certainty

Armed with this concepts, we can give a definition of robust implementation that relates to (Bergemann and Morris 2009).

Definition 42 (Weak Robust Implementation). A strategy σ_d is a weak robust implementation of trust if if it implements it for all type spaces $\mathcal{T}$ such that $\mathcal{T} = \mathbf{WCR}^\infty(\mathcal{T})$

In Appendix 3.10 I show how this is actually equivalent to doing robustness with the belief space $\mathcal{B}_p^w$, and likewise for $\mathcal{B}_p^s$

3.10 Appendix B - Universal Type Space and Strong Rationalizable Strategies

In this section I adapt the results in (Xandri 2012) on the characterization of the Universal Type Space theorem and the study of the topological properties of the sets of weak and strong rationalizable strategies to my setup, for any compact type space we might consider. It generalizes (Battigalli and Siniscalchi.1999; Battigalli and Siniscalchi 2003; Battigalli 2003) to general topological spaces, which is necessary because their results do not apply to my paper. Their results require either finite strategies or finite periods, because their are obtained by extending (Brandenburger and Dekel 1993) to dynamic settings, which works with complete metrizable strategy spaces. However, infinitely repeated games typically involve using the weak convergence topology on the set of strategies, which is non-metrizable if, for example, agents have a continuum of actions in each period. This section is organized as follows: in subsection 3.10.1 I introduce and give some results on the topology of strategy spaces. In subsection 3.10.2 I introduce the notion of hierarchies of beliefs, and study their topological properties. In subsection 3.10.3 I provide a version of the Universal Type Space Theorem (as in (Mertens and Zamir 1985; Battigalli and Siniscalchi 1999)) and finally, in subsection 3.10.4 we apply the results we found in the previous sections to characterize the compactness of the set of weak and strong rationalizable strategies, a crucial result for the model studied in this paper.

3.10.1 Topological Properties of Strategy Spaces

We have that the set from which p chooses is clearly Hausdorff, regular and compact $S_p := \{0,1\}$.

We will now show that the set from which d chooses is also Hausdorff, compact and regular since it is the product of two compact, regular and Hausdorff spaces:

$$S_d := C \times G \tag{3.74}$$

where $G = \mathcal{M}(Z, \{0,1\})$, the set of measurable functions $g : Z \to \{0,1\}$. $C \subset \mathbb{R}$ is compact by assumption, (and Hausdorff and regular because $\mathbb{R}$ is). We will show that G is also a Hausdorff, compact and regular space with the product topology:i.e. point-wise convergence

$$g_n(.) \to g(.) \iff g_n(z) \to g(z) \text{ for all } z \in Z \tag{3.75}$$

The compactness follows from 3 reasons:

1. $G \subset \{0,1\}^Z = \prod_{z \in Z} \{0,1\} = \hat{G}$ which is a compact space with the product topology described 3.75, because of Tychonoff's Theorem. It is also Hausdorff and regular (Theorem 31.2 in (Munkres 2000)).
2. G is a closed subset of $\hat{G}$, because of the Dominated convergence theorem (Theorem 2.24 in (Folland 1999))
3. G is therefore compact (since it is a closed subset of a Hausdorff compact space, Theorem 26.2 in (Munkres 2000))

The fact that G is regular and Hausdorff is simply because it is a subspace of $\hat{G}$, which is a regular and Hausdorff space itself (Theorem 31.2 in (Munkres 2000)). Because C and G and both Hausdorff, compact and regular, S_d is also Hausdorff, compact and regular. Strategies for d are functions

$$\sigma_d : \mathcal{H}_d \to C \times G$$

which can be written as

$$\Sigma_d \equiv (C \times G)^{\mathcal{H}_d} = S_d^{\mathcal{H}_d}$$

which by Tychonoff's Theorem and Theorem 31.2 in (Munkres 2000), is also Hausdorff, compact and regular, with the product topology; i.e.

$$\sigma_d^{(n)} \to \sigma_d \text{ if and only if } \begin{cases} c_n\left(h^\tau\right) \to c\left(h^\tau\right) & \text{for all histories } h^\tau \in \mathcal{H}_d \\ g_n\left(h^\tau, z\right) \to g\left(h^\tau, z_\tau\right) & \text{for all } \left(h^\tau, c_\tau, 1, z_\tau\right) \in \mathcal{H}_d \end{cases} \tag{3.76}$$

and

$$\sigma_p^{(n)} \to \sigma_p \text{ if and only if } a_n\left(h^\tau, c_\tau\right) \to a\left(h^\tau, c_\tau\right) \tag{3.77}$$

Note that, because of the boundedness of both c and $g(\cdot)$ we can apply the Dominated convergence theorem (Theorem 2.24 in (Folland 1999)) to show that the function $V(\cdot \mid h)$ defined in 3.24 is a continuous function of $\sigma \in \Sigma$, and using Theorem Theorem 2.25 in (Folland 1999), we can also show that the continuation value function $W_\theta(\cdot \mid h)$ defined in 3.23 is also continuous function of $\sigma \in \Sigma$ with the product topology. We summarize the results of this subsection in the following Lemma.

Lemma 43 (**Topology of Σ_i**). *The strategy spaces Σ_i for $i \in \{p, d\}$ are Hausdorff, compact and regular topological spaces, with the topology of point-wise convergence (as in 3.76 and 3.77). Moreover, for all histories $h \in \mathcal{H}_i$, the conditional expected utility functions $V(\sigma \mid h)$ and $W_\theta(\sigma \mid h)$ as defined in 3.24 and 3.23 are continuous*

3.10.2 Hierarchies of Beliefs

Given a topological space (X, τ), define $\Delta(X)$ as the set of all Borel probability measures on X. If X is a compact, Hausdorff space, then $\Delta(X)$ is also a Hausdorff and compact topological space (Theorem 3 in (Heifetz 1993)) with the *weak-* topology.* This is the topology of the convergence in distribution: a sequence $\{\lambda_n\}_{n \in \mathbb{N}}$ converges in distribution to λ (written as $\lambda_n \rightsquigarrow \lambda$) if and only if

$$\int f(x)\, d\lambda_n(x) \to \int f(x)\, d\lambda(x) \text{ for all } f \in \mathcal{M}(X, \mathbb{R}) \tag{3.78}$$

where $\mathcal{M}(X, \mathbb{R})$ is the set of all measurable functions with respect to the Borel $\sigma-$algebra. Moreover, all Borel probability measures on X are also **regular** (Theorem 5, (Heifetz 1993)). Therefore, using Tychonoff's theorem, the set $[\Delta(X)]^{\mathcal{H}}$ is also a compact, Hausdorff space with the product topology (having point-wise the weak topology). The set of conditional probability systems on X, which we write $\Delta^{\mathcal{H}}(X)$ is a closed subset of $[\Delta(X)]^{\mathcal{H}}$ (Lemma 1 in (Battigalli and Siniscalchi 1999)), and therefore inherits compactness and Hausdorff property. We will say a conditional probability system π is **regular** if and only if $\pi(\cdot \mid h)$ is a regular measure over X for all $h \in \mathcal{H}$. These results are summarized in the following Lemma

Lemma 44 (Topology of $\Delta^{\mathcal{H}}(X)$). *Given a Hausdorff and compact topological space X and a family $\mathcal{H}$ of histories, the space $\Delta^{\mathcal{H}}(X)$ of conditional probability systems on X is also a Hausdorff, compact space with the product topology of convergence in distribution: i.e. given a sequence $\{\pi_n\} \in \Delta^{\mathcal{H}}(X)$ and $\pi \in \Delta^{\mathcal{H}}(X)$, we say*

$$\pi_n \to \pi \text{ in } \Delta^{\mathcal{H}}(X) \iff \pi_n(\cdot \mid h) \rightsquigarrow \pi(\cdot \mid h) \text{ for all } h \in \mathcal{H} \tag{3.79}$$

Moreover, every $\pi \in \Delta^{\mathcal{H}}(X)$ is regular.

A useful corollary of Lemma 44 will be needed for characterizing the best reply correspondence. Given a type space $\mathcal{T} = \left(T_p, T_d, \hat{\theta}(\cdot), \hat{\pi}_p(\cdot), \hat{\pi}_d(\cdot)\right)$, define now the functions $\mathcal{V}(\cdot \mid h) : \Sigma_p \times \Delta^{\mathcal{H}}(T_d \times \Sigma_d) \to \mathbb{R}$ and $\mathcal{W}_\theta(\cdot \mid p) : \Sigma_d \times \Delta^{\mathcal{H}}(T_p \times \Sigma_p)$ as

$$\mathcal{V}_p(\sigma_p, \pi_p \mid h) := V^{\pi_p}(\sigma_p \mid h) \tag{3.80}$$

and

$$\mathcal{W}_\theta(\sigma_d, \pi_d \mid h) := W_\theta^{\pi_d}(\sigma_d \mid h) \tag{3.81}$$

Corollary 45 (Continuity of Expected Utility over types). *If T_i are compact topological spaces for $i = 1, 2$, then the functions $\mathcal{V}(\cdot \mid h)$ and $\mathcal{W}_\theta(\cdot \mid h)$ are continuous functions (in the weak topology).*

Proof. Since both functions are linear functionals in the space $\Delta(T_{-i} \times \Sigma_{-i})$, for continuity I only need to show boundedness of both functions. This follows from directly from Lemma 44 and the Dominated convergence theorem (which makes the convergence the weak convergence). The continuity of $\mathcal{V}(. \mid h)$ with respect to σ_d has already been established in Lemma 43 □

I will now replicate here the inductive construction of the set of hierarchies of beliefs, as in (Battigalli and Siniscalchi 1999): Define first

$$X_p^0 := \Theta_d \times \Sigma_d \text{ and } X_d^0 = \Sigma_p \tag{3.82}$$

$$X_i^1 := X_i^0 \times \Delta^{\mathcal{H}_i}\left(X_j^0\right) \text{ for } i \in \{d, p\}$$

and in general

$$X_i^k := X_i^{k-1} \times \Delta^{\mathcal{H}_i}\left(X_j^{k-1}\right) \tag{3.83}$$

Proposition 46. *X_i^k is a Hausdorff and compact topological space for all $k = 0, 1, 2, \ldots.$ and $i \in \{d, p\}$. Moreover, $x \in X_i^k \iff x = (x_{k-1}, \pi_1, \pi_2, \ldots, \pi_{k-1})$ where $x_{k-1} \in X_i^{k-1}$ and π_s is a regular CPS on X_j^s for all $s = 1, 2, \ldots, k-1$.*

Proof. By induction, I will show that X_i^k is compact, Hausdorff, and it consists of regular measures on its previous "level". Clearly is true for $k = 0$, since from 3.82 and Lemma 43, we know that X_i^0 is a Hausdorff and compact topological space. Now, assuming $\left\{X_i^{k-1}\right\}_{i\in\{p,d\}}$ is Hausdorff and compact, I need to show that $\{X_i^k\}_{i\in\{p,d\}}$ is also Hausdorff and compact. Using Lemma 44 we then know that $\Delta^{\mathcal{H}}\left(X_j^{k-1}\right)$ is compact and Hausdorff, and consists of regular measures. This together with definition 3.83 gives the desired result. The second result follows from (Battigalli and Siniscalchi 1999) which show that we can write X_i^k simply as

$$X_i^k = \Sigma_j \times \prod_{s=0}^{k-1} \Delta^{\mathcal{H}_i}\left(X_j^s\right) \tag{3.84}$$

□

Define the set of hierarchies of beliefs for agent $i \in \{p, d\}$ to be the set $H_i = \lim_{k\to\infty} X_i^k$, which can be written (according to 3.84) as

$$H_i = \prod_{k=1}^{\infty} \Delta^{\mathcal{H}_i}\left(X_j^k\right) \tag{3.85}$$

So, an element $\mathbf{h} = (\pi_0, \pi_1, ...) \in H_i$ consists on a CPS π_0 on Σ_j (the strategies of the other agent), a CPS π_1 on $\Delta^{\mathcal{H}_j}(\Sigma_i)$ (the CPS's of j about $i's$ strategies), a CPS π_2 on $\Delta^{\mathcal{H}_j}\left(\Delta^{\mathcal{H}_i}\left(\Delta^{\mathcal{H}_j}(\Sigma_d)\right)\right)$, and so on. Clearly the space H_i is compact and Hausdorff, because of Proposition 46 and Tychonoff's theorem. We summarize these results below

Proposition 47 (**Topology of H_i**). *The set of hierarchies of beliefs H_i for $i \in \{p, d\}$ as defined in 3.85 are Hausdorff and compact topological spaces, with the point-wise convergence in each level:*

$$\mathbf{h}_n = \left(\pi_n^k\right)_{k\in\mathbb{N}} \to \mathbf{h} = \left(\pi^k\right)_{k\in\mathbb{N}} \iff \pi_n^k(\cdot \mid h) \rightsquigarrow \pi^k(\cdot \mid h) \text{ for all } k \in \mathbb{N}, h \in \mathcal{H}_i \tag{3.86}$$

Moreover, for all hierarchies $\mathbf{h} = \left(\pi^k\right)_{k\in\mathbb{N}}$ and all $k \in \mathbb{N}$, we have that $\pi^k \in \Delta^{\mathcal{H}_i}\left(X_j^k\right)$ is a regular CPS

3.10.3 Construction of the Universal Type Space

Not all hierarchies of beliefs will be "rational", in the sense that upper level beliefs (say, $k-$order beliefs) may not be consistent with lower level beliefs. We say that a hierarchy $h \in H_i$ is *coherent* when different levels of beliefs are consistent with each other. The formal definition is given by (Battigalli and Siniscalchi 1999) (Definition 1) :

Definition 48 (**Coherency**). A hierarchy of beliefs $\mathbf{h} \in H_i$ is *coherent* if and only if

$$\operatorname{mrg}_{X_j^{k-1}} \pi^{k+1}(\cdot \mid h) = \pi^k(\cdot \mid h) \text{ for all } h \in \mathcal{H}_i, k \in \mathbb{N} \tag{3.87}$$

where $\mathrm{mrg}_{X^{k-1}}\pi^{k+1}$ is the marginal of measure π^{k+1} on the projection X^{k-1}.

Definition 48 is also identical to the definition of projective sequence of regular Borel probability measures, as in (Heifetz 1993) (Definition 7), since by 47 we know that all measures (and their projections) involved in the hierarchies are regular probability measures. We write H_i^* to mean the set of coherent hierarchies of beliefs for agent i. Is easy to see that coherency is a closed restriction of the space H_i, which readily implies that H_i^* is itself a Hausdorff and compact subspace of H_i.

Define

$$T_p^* = H_p^*, T_d^* = \Theta_d \times H_d^* \tag{3.88}$$

So, the universal type sets are simply the sets of all coherent hierarchies of beliefs for each agent. See also that the identity mappings make sense, in that an element in H_i^* is precisely a coherent CPS on the elements of H_j^*. In order to understand the maximality property of the type space we want to construct, in the sense that all other type spaces are in some way embedded into it, I need to define the concept of *type-morphisms*. Given two type spaces $\mathcal{T} = \left(T_p, T_d, \hat{\theta}(\cdot), \hat{\pi}_p(\cdot), \hat{\pi}_d(\cdot)\right)$ and $\mathcal{T}' = \left(T_p', T_d', \hat{\theta}'(\cdot), \hat{\pi}_p'(\cdot), \hat{\pi}_d'(\cdot)\right)$ and a function $\varphi_i : \Sigma_j \times T_j \to \Sigma_j \times T_j'$ with $i \in \{p, d\}$ and $j \neq i$, define $\tilde{\varphi}_i : \Delta^{\mathcal{H}_i}(\Sigma_j \times T_j) \to \Delta^{\mathcal{H}_i}\left(\Sigma_j \times T_j'\right)$ as the function associating to each CPS π_i on $\Sigma_j \times T_j$ the induced CPS $\tilde{\varphi}_{ii}(\mu_i)$ over $\Sigma_j \times T_j'$, as defined in (Battigalli and Siniscalchi 1999) (Subsection 3.1). Formally, given $\pi_i \in \Delta^{\mathcal{H}_i}(\Sigma_j \times T_j)$ we have

$$\tilde{\varphi}_i(\mu_i)(A \mid h) = \pi_i\left(\varphi_i^{-1}(A) \mid h\right) \text{ for all measurable } A \subset \Sigma_j \times T_j', h \in \mathcal{H}_i \tag{3.89}$$

i.e. it gives events in $\Sigma_j \times T_j'$ the probability according to μ_i in the pre-image of that event in $\Sigma_j \times T_j$.

Definition 49 (**Type-morphisms** ((Battigalli and Siniscalchi 1999))). Given two type spaces $\mathcal{T} = \left(T_p, T_d, \hat{\theta}(\cdot), \hat{\pi}_p(\cdot), \hat{\pi}_d(\cdot)\right)$ and $\mathcal{T}' = \left(T_p', T_d', \hat{\theta}'(\cdot), \hat{\pi}_p'(\cdot), \hat{\pi}_d'(\cdot)\right)$ we say that a pair of functions $\varphi = (\varphi_p, \varphi_d)$ where $\varphi_i : T_i \to T_i'$ is a **type morphism** from $\mathcal{T}$ to $\mathcal{T}'$ if and only if the functions φ_i are continuous, and satisfy

$$\hat{\pi}_i'[\varphi_i(t_i)] = \tilde{\varphi}_i[\hat{\pi}_i(t_i)] \text{ for all } t_i \in T_i, i \in \{p, d\} \tag{3.90}$$

and

$$\hat{\theta}'[\varphi_d(t_d)] = \hat{\theta}(t_d) \text{ for all } t_d \in T_d \tag{3.91}$$

When φ is a homeomorphism we say that $\mathcal{T}$ and $\mathcal{T}'$ are *type-isomorphic*.

Conditions 3.90 and 3.91 state that the beliefs and utility parameters (respectively) of all types in $\mathcal{T}$ can be mapped (in a continuous way) into beliefs and parameters of $\mathcal{T}'$. The intuitive idea of this definition is that $\mathcal{T}$ is "smaller" than $\mathcal{T}'$, since every type in T_i can be mapped to a subset of types in T_i' (i.e. the image of φ) that have essentially the same epistemic properties: same beliefs and same utility parameters.

The following Theorem is a simple consequence of Theorem 8 in (Heifetz 1993) and Proposition 3 in (Battigalli and Siniscalchi 1999), adapted to the modified topological assumptions of this model.

Theorem 50 (Universal Type Space Theorem). *The sets T_p^* and T_d^* defined in 3.88 satisfy:*

$$T_p^* \text{ is homeomorphic to } \Delta^{\mathcal{H}_d}\left(T_d^* \times \Sigma_d\right) \tag{3.92}$$

and

$$T_d^* \text{ is homeomorphic to } \Theta_d \times \Delta^{\mathcal{H}_d}\left(T_p^* \times \Sigma_p\right) \tag{3.93}$$

with homeomorphisms $Q_p : T_p^* \to \Delta^{\mathcal{H}_P}\left(T_d^* \times \Sigma_d\right)$ and $Q_d : T_d^* \to \Theta_d \times \Delta^{\mathcal{H}_d}\left(T_p^* \times \Sigma_p\right)$. The type space $\mathcal{T}^* = \left(T_p^*, T_d^*, \hat{\theta}^*(\cdot), \hat{\pi}_p^*(\cdot), \hat{\pi}_d^*(\cdot)\right)$ with $\left(\hat{\theta}^*(\cdot), \hat{\pi}_d^*(\cdot)\right) = Q_d(\cdot)$ and $\hat{\pi}_p^*(\cdot) = Q_p(\cdot)$ is called the ***Universal Type Space****, and has the following property: for any other type space $\mathcal{T}$ there exists a (unique) type morphism φ between $\mathcal{T}$ and $\mathcal{T}^*$*

Proof. Proposition 47 tells us that all measures in a hierarchy are regular measures. This, together with Theorems 8 and 9 in (Heifetz 1993) proves conditions 3.92 and 3.93, by applying the Generalized consistency theorem to each individual history $h \in \mathcal{H}$ and constructing the homeomorphism by defining it history by history. The universality condition is an almost direct application of Proposition 3 in (Battigalli and Siniscalchi 1999) since we can easily replicate the proof step by step with our topological assumptions. □

3.10.4 Topology of Rationalizable sets

In this section I show that the set of rationalizable strategies for any compact type space is in fact, a compact subset of the set of strategies, which I characterized in subsection 3.10.1. Moreover, the set of strongly rationalizable strategies will be in fact, a subset of the weak rationalizable strategy set, implying that strong rationalizability is a closed, stronger solution concept. This will be useful when using the structure theorems in (Weinstein and Yildiz 2012). The main tool I will be using to prove stated in (Ausubel and Deneckere 1993)

Theorem 51 (Berge's Theorem of the Maximum ((Berge 1963))**).** *Let X and Y be topological spaces, with Y regular, a continuous function $f : X \times Y \to \mathbb{R}$ and a continuous, non-empty and compact valued correspondence $\Gamma : X \rightrightarrows Y$. Then the function*

$$M(x) := \max_{y \in \Gamma(x)} f(x, y)$$

is well defined and continuous, and moreover, the correspondence

$$g(x) := \underset{y \in \Gamma(x)}{\arg\max} f(x, y)$$

is non-empty, compact valued and upper hemi-continuous.

The most important consequence of the theorem of the maximum is the continuity and upper hemicontinuity of the value and best response functions, respectively

Proposition 52 (Continuity of Sequential Best Responses). *For any type space $\mathcal{T} = \left(T_p, T_d, \hat{\theta}(\cdot), \hat{\pi}_p(\cdot), \hat{\pi}_d(\cdot\right.$ and any $t_i \in T_i$ define the sequential best response correspondence $SBR_i : T_i \to \Sigma_i$ that gives the sequential best responses for type i. Then, if $\hat{\pi}_i(\cdot)$ for $i \in \{p, d\}$ are continuous functions, $SBR_i(t_i)$ is a non-empty, compact valued and upper-hemi continuous correspondence.*

Proof. I will only show the continuity of SBR_d, since SBR_p follows a similar (and easier) argument. Corollary 45 tells us that $\mathcal{W}_\theta(\sigma_d, \pi_d \mid h)$ is a continuous function of the CPS π_d, and Proposition 43 tells us that it is also a continuous function of σ_d (taking π_d as given), which makes $\mathcal{W}_\theta(\cdot \mid h)$ a continuous function over $\Sigma_d \times \Delta^{\mathcal{H}_d}(T_p \times \Sigma_p)$ (with the product topology). Proposition 43 also implies that the set Σ_d is regular, Hausdorff and compact. The domain of the program is Σ_d, which is a constant correspondence, hence continuous, non-empty and compact valued (since Σ_d is compact). Therefore, we can apply the theorem of the maximum 51 to show that the correspondence

$$\phi(\theta, \pi_d \mid h) := \arg\max_{\sigma_d \in \Sigma_d} \mathcal{W}_\theta(\sigma_d, \pi_d \mid h)$$

is a continuous, non-empty and compact valued u.h.c correspondence of (θ, π_d) for all $h \in \mathcal{H}$, (continuity on θ comes for free with it's finiteness) which therefore implies that the correspondence $\hat{\phi}(\theta, \pi) = (\phi(\theta, \pi_d \mid h))_{h \in \mathcal{H}}$ is also a continuous, non-empty and compact valued u.h.c correspondence. The desired result then follows from the continuity of $\hat{\pi}_d$, since

$$SBR_d(t_d) = \hat{\phi}\left[\hat{\theta}(t_d), \hat{\mu}_d(t_d)\right]$$

a composition of an u.h.c correspondence with a continuous function, which is also a u.h.c correspondence, as we wanted to show. Compactness also follows from continuity (Weierstrass). $\square$

Now I present the main result of this section

Theorem 53 (Topological Properties of Weak and Strong rationalizability). *Take a compact type space $\mathcal{T} = \left(T_p, T_d, \hat{\theta}(\cdot), \hat{\pi}_p(\cdot), \hat{\pi}_d(\cdot)\right)$ and recall the definitions of $WCR_i^k(\mathcal{T}) \subset \Sigma_i$ and $SCR_i^k(\mathcal{T}) \subset \Sigma_i$ as the set of weak and strong rationalizable strategies for type space $\mathcal{T}$. Then:*

1. *The sets $WCR_i^k(\mathcal{T})$ and $SCR_i^k(\mathcal{T})$ are non-empty, compact, Hausdorff and regular spaces, and satisfy $WCR_i^k(\mathcal{T}) \subseteq SCR_i^k(\mathcal{T})$ for all $k \in \mathbb{N}, i \in \{p, d\}$*

2. *The rationalizable sets $WCR_i^\infty(\mathcal{T}) \subseteq SCR_i^\infty(\mathcal{T})$ are also non-empty, compact, Hausdorff and regular spaces for $i \in \{p, d\}$*

3. *The set of all weak rationalizable strategies $WCR_i^\infty \subset \Sigma_i$ and strong rationalizable strategies $SCR_i^\infty \subset \Sigma_i$ are non-empty, compact, Hausdorff and regular spaces, and satisfy $WCR_i^\infty \subset SCR_i^\infty$ for $i \in \{p, d\}$.*

Proof. (1) follows directly from Propositions 3.5 and 3.6 in (Battigalli 2003), since the strategy space Σ_i is compact (Proposition 3.10.1) and the best response correspondences are u.h.c, non-empty compact valued (Proposition 52). The restrictions on the rationalizable sets can also be mapped as restrictions on the type space, as shown by (Battigalli and Siniscalchi 2003). The fact that $WCR_i^{\infty}(\mathcal{T})$ and $SCR_i^{\infty}(\mathcal{T})$ are non-empty follows from the compactness and non-emptiness proved in follows from (2) and the generalization of Cantor's Theorem, which states that the intersection of a decreasing sequence of non-empty compact sets is non-empty (Theorem 26.9 in (Munkres 2000)). Since $WCR_i^k(\mathcal{T})$ and $SCR_i^k(\mathcal{T})$ are compact, they are also closed sets, which make $WCR_i^{\infty}(\mathcal{T})$ and $SR_i^{\infty}(\mathcal{T})$ closed. Because Σ_i is a Hausdorff space, this also implies that $WCR_i^{\infty}(\mathcal{T})$ and $SCR_i^{\infty}(\mathcal{T})$ are also compact spaces (Theorem 26.2 (Munkres 2000)). Regularity follows from regularity of Σ_i, and therefore we have shown (2). For (3) we use the universal type space theorem 50 to be able to write

$$WCR_i^* \equiv \bigcup_{\mathcal{T}:\mathcal{T} \text{ is a type space}} WCR_i^{\infty}(\mathcal{T}) = WCR_i^{\infty}(\mathcal{T}^*) \tag{3.94}$$

and

$$SCR_i^* \equiv \bigcup_{\mathcal{T}:\mathcal{T} \text{ is a type space}} SCR_i^{\infty}(\mathcal{T}) = SCR_i^{\infty}(\mathcal{T}^*) \tag{3.95}$$

and we use again this theorem to recall that the type space $\mathcal{T}^*$ consists of compact type spaces T_i^* with continuous belief functions $\hat{\pi}_i^*$. Therefore we can apply the result in (3) for the particular case of $\mathcal{T} = \mathcal{T}^*$. □

3.11 Appendix C - Proofs and Supplementary Results

I will need some extra notation for the proofs in this section. Given an appended history $h^s = (h^{\tau}, h^k)$, I write $h^s \sim h^{\tau} = h^k$ for the tail of the history. Also, whenever we can decompose h^s in this manner, I will say that h^{τ} *precedes* h^s and write $h^{\tau} \prec h^s$.

Proof of Lemma 29. The first part is a consequence of Lemma (56) in Appendix 3.11. For the second result, take a robust and strong rationalizable strategy σ_d and suppose there exist a history h and a strong rationalizable pair $(\hat{\sigma}_d, \hat{\pi}_d)$ that deliver an expected payoff that is less than the payoff of the robust policy:

$$W_{\theta}^{\hat{\pi}_d}(\hat{\sigma}_d \mid h) < W_{\theta}(\sigma_d \mid h).$$

However, if $\hat{\pi}_d$ has common strong certainty of rationality, then she is also certain that p plays strong rationalizable strategies (Proposition 3.10 in (Battigalli and Bonanno 1999)), and hence she should be also certain that by following the robust strategy σ_d from history h on she will get a higher expected payoff. Since this is true for any rationalizable belief, $\hat{\sigma}_d$ cannot be the sequential best response for beliefs $\hat{\pi}_d$ (since it is conditionally dominated by σ_d at h), reaching a contradiction □

Lemma 54. *Take a history h^τ and $\theta-$rationalization (σ_d, π_d). Also, let $\nu = (\hat{\sigma}_d, \hat{\pi}_d)$ be another $\theta-$rationalizable pair that satisfies:*

$$W_\theta^{\hat{\pi}_d}\left(\hat{\sigma}_d \mid h^0\right) \geq \frac{1-\beta}{\beta} S_{\theta,\tau-1} + \underline{\mathbb{W}}_\theta \tag{3.96}$$

Then, there exists a pair $(\sigma_d^\nu, \pi_d^\nu)$ that also $\theta-$rationalizes h^τ and is such that

$$\sigma_d^\nu(h^s) = \hat{\sigma}_d(h^s \sim h^\tau), \pi_d^\nu(\cdot \mid h^s) = \hat{\pi}_d(\cdot \mid h^s \sim h^\tau) \tag{3.97}$$

for all histories $h^s \succ h^\tau$

Proof. Define the pair $(\sigma_d^\nu, \pi_d^\nu)$ for any history $\tilde{h}^s$ as

$$\sigma_d^\nu\left(\tilde{h}^s\right) := \begin{cases} \sigma_d\left(\tilde{h}^s\right) & \text{if } s < \tau \text{ or } \tilde{h}^s = h^\tau \\ \sigma_\theta^*\left(\tilde{h}^s \sim \tilde{h}^\tau\right) & \text{if } s \geq \tau \text{ and } h^\tau \not\prec \tilde{h}^s \\ \hat{\sigma}_d\left(\tilde{h}^s \sim h^\tau\right) & \text{if } s \geq \tau \text{ and } h^\tau \prec \tilde{h}^s \end{cases} \tag{3.98}$$

and for any measurable set $A \subset \Sigma_p$

$$\pi_d^k\left(A \mid \tilde{h}^s\right) := \begin{cases} \pi_d\left(A \mid \tilde{h}^s\right) & \text{if } s < \tau \\ \underline{\pi}_\theta\left(A \mid \tilde{h}^s \sim \tilde{h}^\tau\right) & \text{if } s \geq \tau \text{ and } h^\tau \not\prec \tilde{h}^s \\ \hat{\pi}_d\left(A \mid \tilde{h}^s \sim h^\tau\right) & \text{if } s \geq \tau \text{ and } h^\tau \preceq \tilde{h}^s \end{cases} \tag{3.99}$$

so the pair $(\sigma_d^\nu, \pi_d^\nu)$ coincides with (σ_d, π_d) for any histories of length less than $\tau - 1$, and strategies also do it up to time τ. If at history $\left(h^{\tau-1}, c_{\tau-1}, a_{\tau-1}, z_{\tau-1}\right)$ d deviates from $r = r^{\sigma_d}\left(h^{\tau-1}, z_{\tau-1}\right)$ going to h'^τ, then type θ believes that she will switch to the optimal strong rationalizable strategy from then on, to which the best response is σ_θ^* and the expected payoff is

$$W_\theta^{\pi_d^\nu}\left(\sigma_d^\nu \mid h'^\tau\right) = W_\theta^{\underline{\pi}_\theta}\left(\sigma_d^* \mid h^0\right) = \underline{\mathbb{W}}_\theta$$

which is a rationalizable continuation pair. Same is true for the continuations at all histories after h^τ, and so the pair $(\sigma_d^\nu, \pi_d^\nu)$ is rationalizable. Then, to finish our proof, we need to show that it is consistent with h^τ only at $r_{\tau-1}$. Consider first the case where $r_{\tau-1} = 0$ and $S_{\theta,\tau-1} = U_{\theta,\tau-1} - c_{\tau-1} > 0$. Then, the optimal choice under (σ_d^k, π_d^k) is

$$\beta W_\theta^{\pi_d^\nu}\left(\sigma_d^\nu \mid h^\tau\right) \geq (1-\beta)\left(U_{\theta,\tau-1} - c_\tau\right) + \beta \underline{\mathbb{W}}_\theta \iff$$

$$W_\theta^{\hat{\pi}_d}\left(\hat{\sigma}_d \mid h^0\right) \geq \frac{1-\beta}{\beta} S_{\theta,\tau-1} + \underline{\mathbb{W}}_\theta$$

which is the assumption made in 3.96. The other cases are shown in a similar fashion. □

Proof of Proposition 30. Given the functions $(r(\cdot), w(\cdot))$ that satisfy conditions 3.45 and 3.46, I need to construct a $\theta-$rationalizable pair (σ_d, π_d) such that $r^{\sigma_d}(h^\tau, z_\tau) = r(z_\tau)$ for all $z \in Z$. Because the set or rationalizable payoffs is convex, we know that for any $w \in [\underline{\mathbb{W}}_\theta, \overline{\mathbb{W}}_\theta]$ there exist some rationalizable pair (σ_w, π_w) such that

$$W_\theta^{\pi_w}(\sigma_w \mid h^0) = w$$

then, for all $z \in Z$ we can find a rationalizable pair $(\hat{\sigma}_z, \hat{\pi}_z)$ such that

$$W_\theta^{\hat{\pi}_z}(\hat{\sigma}_z \mid h^0) = w(z) \tag{3.100}$$

which are rationalizable continuations from time 0 perspective. Moreover, see that that $r(z)$ solves the IC constraint 3.45 for this continuations, which means that it would be the best response at $\tau = 0$ if θ expected the continuation values $w(z)$ starting from $\tau = 1$. Formally, let $h^1(z) = (c_0, a_0, z_0 = z, r_0 = r(z))$ and define the strategy σ_0 as

$$\sigma_0(h^\tau) = \begin{cases} (c_\tau, r(\cdot)) & \text{if } h = h^0 \\ \sigma_z(h^s \sim h^1(z)) & \text{if } h^1(z) \prec h^s \\ \sigma_\theta^*(h^s \sim h^1) & \text{otherwise} \end{cases}$$

i.e. upon deviations in the first period, goes to the optimal robust strategy, and by following the proposed policy $r(z)$ it continues prescribing strategy σ_z after that history, which gives an expected payoff of $w(z)$. This then implies that the policy function is $\theta-$rationalizable at h^0, and that it's expected payoff is

$$W_\theta^{\pi_0}(\sigma_0 \mid h^0) = \mathbb{E}_z[(1-\beta) r(z)(U_\theta - c_\tau) + \beta w(z)] \geq \frac{1-\beta}{\beta} S_{\theta,\tau-1} + \underline{\mathbb{W}}_\theta$$

But then we can use Lemma 54 for the pair $(\hat{\sigma}_d, \hat{\pi}_d) = (\sigma_0, \pi_0)$, finishing the proof. □

To show Proposition 31 we will need the following Lemma

Lemma 55 (No strong separation by commitment costs). *Take a history h^τ that is strong rationalizable for both types, and a commitment cost $\hat{c}$ such that $(h^\tau, \hat{c})$ is new−rationalizable. Then, $(h^\tau, \hat{c})$ is old−rationalizable as well.*

Proof. Suppose not. Then, at history $(h^\tau, \hat{c})$ type $\theta = new$ would achieve robust separation. I will now construct a system of beliefs $\pi \in \mathcal{B}_d^s$, for any continuation history h :

$$\pi(A \mid h)=\begin{cases}1 & \text{if } h \succ (h^\tau,\hat{c}) \text{ and } \sigma_p^{FB} \in A\\ 1 & \text{if } h \not\succ (h^\tau,\hat{c}) \text{and } \underline{\sigma}_p \in A\\ 0 & \text{otherwise}\end{cases} \tag{3.101}$$

where σ_p^{FB} is the first best strategy for p if he faces $\theta = new$, and $\underline{\sigma}_p(h) = 0$ for all histories (i.e. not trust for all continuation histories). See that because of robust separation, for any continuation history h that is $new-$rationalizable, this will be a rationalizable strategy if p puts measure 1 on $\theta = new$. If a continuation history h is not $new-$rationalizable, then because we assumed it is not $old-$rationalizable either, then strong rationalizability puts no restrictions on beliefs after such histories, and hence $\underline{\sigma}_p$ is a strong rationalizable continuation strategy at these histories. Define $\hat{\sigma}_d$ as

$$\hat{\sigma}_d(h)=\begin{cases}\left(\hat{c}, r_{old}^{spot}(\cdot \mid \hat{c})\right) & \text{if } h = h^\tau\\ \left(0, r_{old}^{spot}(\cdot \mid c=0)\right) & \text{if } h \succ (h^\tau,\hat{c})\\ (\infty, r_g(\cdot)) & \text{if } h \not\succ (h^\tau,\hat{c})\end{cases} \tag{3.102}$$

where $r_\theta^{spot}(z \mid c) = \underset{r\in(0,1)}{\operatorname{argmax}}(U_\theta - c)\, r$ and $r_g(z) = 0$ for all $z \in Z$. Is easy to see that $\hat{\sigma} \in SBR_{old}(\pi)$ since if $c \neq \hat{c}$ then utility will be $\underline{u}_{old}$, and

$$\underline{u}_{old} < 0 < (1-\beta)\,\mathbb{E}\{\max(0, U_{old} - \hat{c})\} + \beta\mathbb{E}\{\max(0, U_{old})\} = W_{old}^{\pi}(\hat{\sigma}_d \mid h^\tau)$$

and clearly it is the best response for the continuation histories. But then choosing $c = \hat{c}$ is a strong rationalizable strategy for $\theta = old$, a contradiction. □

Proof of Proposition 31. We will do it by induction: suppose $k = 0$. Since $h^0 = \emptyset$ is rationalizable for both types, Lemma 55 implies that if c_0 is $new-$rationalizable, history (h^0, c_0) is $old-$rationalizable as well. For $k > 1$, suppose that history (h^{k-1}, c_{k-1}) has been both new and $old-$rationalizable, and we know that (h^k, c_k) is also $new-$rationalizable. Because of Lemma 55 history (h^k, c_k) can be $old-$rationalizable as well if and only if $h^k = (h^{k-1}, c_{k-1}, a_{k-1}, z_{k-1}, r_{k-1})$ is also $old-$rationalizable. Since by the induction step we assumed (h^{k-1}, c_{k-1}) is $old-$rationalizable, we need to rationalize only the choice of r_{k-1} after shock z_{k-1}. But here we can apply directly Proposition 30, getting that h^k is $old-$rationalizable if and only if $S_{old,k-1} = \max_{\tilde{r}\in\{0,1\}}(U_{old,k-1} - c_{k-1})\,\tilde{r} - (U_{old,k-1} - c_{k-1})\, r_{k-1} \leq S_{old}^{\max}$. This concludes the proof. □

To prove Proposition 32, we will need two lemmas first:

Lemma 56. *For any strong rationalizable strategy $\sigma_d \in \Sigma_{new}^{SR}$, and any new−rationalizable history, we have*

$$\mathbb{E}_{z_\tau}\left\{r^{\sigma_d}\left(h^\tau, z_\tau\right)\left[U_p\left(z_\tau\right)-c^{\sigma_d}\left(h^\tau\right)\right]\right\} \geq 0 \tag{3.103}$$

Proof. The proof will follow from 2 steps:

Step 1: Let $\underline{\mathbb{W}}_{new} \geq S_{new}^{\max}$. This is equivalent to showing

$$\underline{\mathbb{W}}_{new} \geq \frac{\beta}{1-\beta}\left\{\mathbb{E}_{z_\tau}\left[\max\left(0, U_p\left(z_\tau\right)\right)\right]-\underline{\mathbb{W}}_{new}\right\} \iff$$

$$\underline{\mathbb{W}}_{new} \geq \beta \mathbb{E}_{z_\tau}\left[\max\left(0, U_p\left(z_\tau\right)\right)\right]$$

Suppose $\underline{\mathbb{W}}_{new} < \beta \mathbb{E}_{z_\tau}\left[\max\left(0, U_p\left(z_\tau\right)\right)\right]$. Then the following strategy would be strongly rationalizable: prohibit $r = 1$ at h^τ and in $\tau + 1$ d separates completely. See that since type $\theta = old$ never prohibits r in any rationalizable strategy, then strong certainty of rationality would imply that $\theta = new$ from then on. Therefore, this strategy would then be a robust one, and therefore $\underline{\mathbb{W}}_{new} \geq \beta \mathbb{E}_{z_\tau}\left[\max\left(0, U_p\left(z_\tau\right)\right)\right]$ from the fact that $\underline{\mathbb{W}}_{new}$ is the maximum utility over robust strategies, and thus reaching a contradiction.

Step 2: $\mathbb{E}_{z_\tau}\left[r^{\sigma_d}\left(h^\tau, z_\tau\right)\left(U_p\left(z_\tau\right)-c^{\sigma_d}\left(h^\tau\right)\right)\right] \geq 0$ for all $\sigma_d \in \Sigma_{new}^{SR}$ and all rationalizable histories h^τ.

For any rationalizable strategy σ_d we have

$$(1-\beta)\,\mathbb{E}_{z_\tau}\left[r^{\sigma_d}\left(h^\tau, z_\tau\right)\left(U_p\left(z_\tau\right)-c^{\sigma_d}\left(h^\tau\right)\right)\right]+\beta \mathbb{E}_{z_\tau}\left[\max\left(0, U_p\left(z_\tau\right)\right)\right] \geq W_{new}^{\sigma_d}\left(h^\tau\right) \geq \underline{\mathbb{W}}_{new}$$

This also implies then that

$$(1-\beta)\,\mathbb{E}_{z_\tau}\left[r^{\sigma_d}\left(h^\tau, z_\tau\right)\left(U_p\left(z_\tau\right)-c^{\sigma_d}\left(h^\tau\right)\right)\right] \geq \beta\underline{\mathbb{W}}_{new}-\beta \mathbb{E}_{z}\left[\max\left(0, U_p\left(z_\tau\right)\right)\right]+(1-\beta)\,\underline{\mathbb{W}}_{new} \iff$$

$$\mathbb{E}_{z_\tau}\left[r^{\sigma_d}\left(h^\tau, z_\tau\right)\left(U_p\left(z_\tau\right)-c^{\sigma_d}\left(h^\tau\right)\right)\right] \geq \underline{\mathbb{W}}_{new}-S_{new}^{\max} \geq 0$$

using Step 1 in the last inequality. □

Lemma 57. *There exist a non-zero measure set $\tilde{S} \subset \left[0, S_{old}^{\max}\right]$ such that $\frac{\partial \mathcal{W}_n}{\partial s}(s) > 0$ for all $s \in \tilde{S}$ and hence $\mathcal{V}\left(s, c^*(s)\right) = \underline{u}_p$ for all $s \in \tilde{S}$*

Proof. Notice first that for all s we have $\mathbb{E}\left[r(z \mid s)\right] = \Pr\left[r(z \mid s) = 1\right] > 0$. This is because if it wasn't, then utility of this policy at s would give utility 0, whereas we could have chosen $c = \mathbf{c}(s)$ and get positive utility, together with positive probability of playing $r = 1$. Suppose not, so that $\frac{\partial \mathcal{W}_n}{\partial s} = 0$ for all s for which the derivative exists (which are almost everywhere). Pick a s such that the constraint is not binding: i.e. $\underline{\mathcal{V}}\left(s, c^*(s)\right) > \underline{u}_p$ (which must necesarily exist given the characterization of the minimum cost function $\mathbf{c}(s)$. Take the optimal policy at that state, which is $r(z) = r^*(z)$, $s(z) = s^*(z)$ and $c = c^*$. We will construct a local feasible deviation: keep the same policy function $r(z)$ and only reduce the commitment cost to $\tilde{c} = c - \epsilon$, which implies that the next period sacrifice would now be

$$s(z,\epsilon) = \max_{\tilde{r}\in[0,1]} (U_{old}(z) - c + \epsilon)\,\tilde{r} - (U_{old}(z) - c + \epsilon)\,r(z)$$

The utility of the right hand side maximized problem was $\square$

$$(1-\beta)\,\mathbb{E}\left[(U_{old}(z) - c)\,r(z)\right] + \beta\mathbb{E}\mathcal{W}\left[s(z)\right]$$

and with the deviation is

$$(1-\beta)\,\mathbb{E}\left[(U_{old} - c + \epsilon)\,r(z)\right] + \beta\mathbb{E}\mathcal{W}\left[s(z,\epsilon)\right]$$

we will show that it is a stricly increasing deviation:

$$(1-\beta)\,\mathbb{E}\left[(U_{old}(z) - c)\,r(z)\right] + \beta\mathbb{E}\mathcal{W}\left[s(z)\right] < (1-\beta)\,\mathbb{E}\left[(U_{old}(z) - c + \epsilon)\,r(z)\right] + \beta\mathbb{E}\mathcal{W}\left[s(z,\epsilon)\right] \iff$$

$$(1-\beta)\Pr\left[r(z) = 1\right]\epsilon + \beta\mathbb{E}\left\{\mathcal{W}\left[s(z,\epsilon)\right] - \mathcal{W}\left[s(z)\right]\right\} > 0 \tag{3.104}$$

Because $\mathcal{W}$ is differentiable almost everywhere, then for almost all $z \in Z$ we can make the differential approximation around $\epsilon = 0$:

$$\mathcal{W}\left[s(z,\epsilon)\right] - \mathcal{W}\left[s(z)\right] \approx \frac{\partial\mathcal{W}}{\partial s}\left[s(z)\right]\left[\frac{\partial s(z,\epsilon)}{\partial\epsilon}\,|_{\epsilon=0}\right]\epsilon$$

and using the envelope theorem

$$\frac{\partial s(z,\epsilon)}{\partial\epsilon} = r_{old}^{spot}(z \mid c - \epsilon) - r(z)$$

so that evaluating it at $\epsilon = 0$ we simplify this condition as

$$\mathcal{W}\left[s(z,\epsilon)\right] - \mathcal{W}\left[s(z)\right] \approx \frac{\partial\mathcal{W}}{\partial s}\left[s(z)\right]\left[r_{old}^{spot}(z) - r(z)\right]\epsilon$$

then for small enough ϵ condition 3.104 is satisfied if and only if

$$(1-\beta)\int r(z)\,f(z)\,dz + \beta\int\left[r_{old}^{spot}(z) - r(z)\right]\frac{\partial\mathcal{W}}{\partial s}\left[s(z)\right]f(z)\,dz > 0 \tag{3.105}$$

The assumption $\Pr(r(z) = 1) > 0$ implies that condition 3.105 will necesarily hold if we can show

$$\int\left[r_{old}^{spot}(z) - r(z)\right]\frac{\partial\mathcal{W}}{\partial s}\left[s(z)\right]f(z)\,dz > 0$$

Because the only potential mass-point for the implied distribution for $s'(z)$ is at $s = 0$ (when there is no sacrifice, sacrifice is zero, and this can happen if $r_{\theta}^{spot} \neq r$ has positive probability) and we already know that $\mathcal{W}$ is locally constant in the interval $[0,\hat{s}]$ we also have that $\frac{\partial \mathcal{W}}{\partial s}(0) = 0$. Therefore,

$$\frac{\partial \mathcal{W}}{\partial s}[s(z)] = 0 \text{ a.e in } z \in Z$$

which given the absolute continuity of Z delivers the desired result.

Lemma 58. *If $\underline{\mathcal{V}}(s, c^*(s)) > \underline{u}_p$ for some s, then it also holds for all $s' \in (s, S_o^{\max})$*

Proof. It follows by inspection of the first order conditions of the lagrangian problem, since s only enters the conditions through this constraint, which implies that if it is non-binding at s it is also non-binding at $s' > s$, since increasing the sacrifice only relaxes this constraint, which was not binding in the optimum. □

Corollary. *There exist $\underline{s} > \hat{s}$ such that for all $s \leq \underline{s}$ we have $c^*(s) = \mathbf{c}(s)$ and for $s \geq \hat{s}$ we have $c^*(s) = \mathbf{c}(\hat{s})$*

Lemma 59. *Under the increasing misalignment assumption 3.5.5, given $\epsilon, \delta > 0$, the functions:*

$$G(a, b \mid \epsilon, \delta) := \int_{a-\epsilon}^{a} \left[\int_{b-\delta}^{b} u_p f(u_p, u_o)\, du_p\right] du_o$$

and

$$H(a, b \mid \epsilon, \delta) = \int_{a-\epsilon}^{a+\epsilon} \left[\int_{b-\delta}^{b+\delta} u_p f(u_p, u_o)\, du_p\right] du_o$$

satisfies $\frac{\partial G}{\partial a} \cdot \frac{\partial H}{\partial a} \leq 0$. If $u_p \frac{\partial f}{\partial u_p} \geq 0$ for all z, then we also have $\frac{\partial G}{\partial b} > 0$

Proof. Using Leibnitz rule:

$$\frac{\partial G}{\partial a} = \int_{b-\delta}^{b} u_p f(u_p, a)\, du_p - \int_{b-\delta}^{b} u_p f(u_p, a-\epsilon)\, du_p = \int_{b-\delta}^{b} u_p \left[f(u_p, a) - f(u_p, a-\epsilon)\right] du_p$$

which is negative given our assumption. Moreover,

$$\frac{\partial H}{\partial a} = \int_{b-\delta}^{b+\delta} u_p \left[f(u_p, a+\epsilon) - f(u_p, a-\epsilon)\right] du_p < 0$$

If $u_p \frac{\partial f}{\partial u_p} \geq 0$ for all z, then

$$\frac{\partial G}{\partial b} = \frac{\partial}{\partial b}\left\{\int_{b-\delta}^{b}\left[\int_{a-\epsilon}^{a} u_p f(u_p, u_o)\, du_o\right] du_p\right\} = \int_{a-\epsilon}^{a} b f(b, u_o)\, du_o - \int_{a-\epsilon}^{a} (b-\delta) f(b-\delta, u_o)\, du_o =$$

$$b \int_{a-\epsilon}^{a} \left[f(b, u_o) - f(b-\delta, u_o)\right] du_o + \delta \int_{a-\epsilon}^{a} f(b-\delta, u_o)\, du_o > 0$$

as we wanted to show. □

Proof of Proposition 32. For $S_{old,k-1} > S_{old}^{\max}$ for some $k \leq \tau - 1$, Proposition 31 implies that p should have strong certainty that $\theta = new$. Lemma 56 also implies that,

$$\mathbb{E}_{z_\tau}\left\{r^{\sigma_d}\left(h^\tau, z_\tau\right) U_p\left(z_\tau\right)\right\} \geq \mathbb{E}_z\left\{r^{\sigma_d}\left(h^\tau, z_\tau\right)\left[U_p\left(z_\tau\right) - c^{\sigma_d}\left(h^\tau\right)\right]\right\} \geq 0 > \underline{u}_p$$

Therefore, in any strong rationalizable history where p is strongly certain that $\theta = new$, p strictly prefers to trust. Since the repeated first best is a strong rationalizable continuation strategy (since it maximizes both d and p's utilities), and p will trust regardless of what rationalizable commitment cost is chosen, $\theta = new$ will optimally choose $c_\tau = 0$ and play her first best afterwards, regardless of her beliefs, as long as they are also consistent with common strong certainty of rationality.

When $S_{old,k-1} \leq S_{old}^{\max}$ for all $k \leq \tau - 1$, Lemma 56 also implies that $\underline{\mathcal{V}}_{new}\left(S_{new,\tau-1}, c_\tau\right) \geq 0 > \underline{u}_p$. Therefore, the implementation restriction

$$\underline{V}\left(h^\tau, c_\tau\right) = \min\left\{\underline{\mathcal{V}}_{old}\left(S_{old,\tau-1}, c_\tau\right), \underline{\mathcal{V}}_{new}\left(S_{new,\tau-1}, c_\tau\right)\right\} \geq \underline{u}_p$$

is satisfied if and only if $\underline{\mathcal{V}}_{old}\left(S_{old,\tau-1}, c_\tau\right) \geq \underline{u}_p$, proving the desired result.

To prove the monotonicity of $\underline{\mathcal{V}}_{old}\left(S_{old,\tau-1}, c_\tau\right)$ with respect to c_τ we use the characterization of the solution to program 3.48 in Proposition 61. When $S_{old,\tau-1} \leq \hat{s}$

$$\underline{\mathcal{V}}_{old}\left(S_{old,\tau-1}, c_\tau\right) = \int_{U_o > c_\tau + S_{old,\tau-1}} U_p\left(z_\tau\right) f\left(z_\tau\right) dz + =$$

$$+ \int_{U_o \in \left(c_\tau - S_{old,\tau-1}, c_\tau + S_{old,\tau-1}\right)} \min\left[0, U_p\left(z_\tau\right)\right] f\left(z\right) dz = G\left(c + \overline{U}, \underline{U} \mid \overline{U} - s, \overline{U}\right) + H\left(c, \frac{\underline{U} + \overline{U}}{2} \mid \frac{\overline{U} - \underline{U}}{2}, s\right)$$

using the definitions in Lemma 59, and hence it is decreasing in c, as we wanted to show. ☐

Lemma 60. *T as defined in 3.57 is a contraction mapping with modulus β*

Proof. I use Blackwell's conditions to show the result (see Theorem 3.3 in (Stokey, Lucas, and Prescott 1989)). We only need to check monotonicity and discount. See that if $g \leq h$ then $T(g)(s) \leq T(h)(s)$ for all s, since the integrand is an increasing operator. Moreover, $T(g + a)(s) = T(g)(s) + \beta a$ for all s, and hence T is a contraction mapping of module β, as we wanted to show. ☐

Proof of Proposition 34 . Define $P(c) = \mathbb{E}_z\left[\left|U_{old} - c\right|\right]$. It can be expressed as

$$P(c) = \int_{z \in Z} \left|U_{old} - c\right| f(z)\, dz = \int_{\underline{U}}^{c} (c - u) f_o(u)\, du + \int_{c}^{\overline{U}} (u - c) f_o(u)\, du$$

where $f_o(u) := \int_{\underline{U}}^{\overline{U}} f(U_p, u)\, dU_p$ denotes the partial of U_{old}. Using Leibniz rule

$$P'(c) := \frac{\partial P(c)}{\partial c} = \int_{\underline{U}}^{c} f_o(u)\, du - \int_{c}^{\overline{U}} f_o(u)\, du = \Pr(U_{old} < c) - \Pr(U_{old} > c)$$

so $\frac{\partial P(c)}{\partial c} > 0 \iff \Pr(U_{old} < c) \geq \Pr(U_{old} > c)$ or equivalently $\Pr(U_{old} < c) \leq \frac{1}{2}$. Then, is easy to see that if condition 3.6.1 holds, then for all $c \geq \overline{c}$ we get $P'(c) > 0$ and hence P is increasing in c. Because $c(\cdot) \in [\overline{c}, c_0^*]$ for all $s \in [0, S_{old}^{\max}]$ and is weakly decreasing in s, the result holds. □

Proof of Lemma 35 . I present the proof for the case with $s = 0$, which corresponds to the greatest commitment cost $c_0^* \geq \mathrm{c}(s)$ for all s. For smaller commitment costs the proof will be analogous. It follows from various steps:

Step 1: $\max|U_{old} - c_0^*| > S_{old}^{\max}$.

If this was not the case, then for all z, $c_0^* - S_{old}^{\max} \leq U_{old} \leq c_0^* + S_{old}^{\max}$. If this was the case, using Proposition 61 we have that

$$\underline{V}(h^\tau, c_0^*) = \int_{z \in Z} \min(0, U_p)\, f(z)\, dz \leq \int_{z: U_{old} > 0} U_p dF(z) < \underline{u}_P$$

which violates the definition of c_0^*

Step 2: $\min(U_{old}) = \underline{U} < c_0^* - S_{old}^{\max} < c_0^* < \overline{U} = \max(\overline{U})$

The right hand side inequality follows from the fact that if $\overline{U} \leq c_0^* - S_{old}^{\max}$ then

$$\underline{V}(h^\tau, c_0^*) = 0 > \underline{u}_p$$

which will never hold for c_0^* (since $\theta = old$ can drive them to indifference by decreasing the commitment cost enough). From step 1, we either must have that $c_0^* - S_{old}^{\max} > \underline{U}$ or $\overline{U} > c_0^* + S_{old}^{\max}$ (or both). Suppose that the result is not true, so that$\underline{U} \geq c_0^* - S_{old}^{\max}$. Suppose first that $c_0^* - S_{old}^{\max} < \overline{c}$. Then

$$\underline{V}(h^\tau, c_0^*) = \int_{U_{old} > c_0^* - S_{old}^{\max}} \min(0, U_p)\, f(z)\, dz \tag{3.106}$$

$$= \int_{U_{old} \in \left(c_0^* - S_{old}^{\max}, \overline{c}\right)} \min(0, U_p)\, f(z)\, dz + \int_{U_p > \overline{c}} \min(0, U_p)\, f(z)\, dz \leq$$

$$\int_{U_p > \overline{c}} \min(0, U_p)\, f(z)\, dz < \int_{U_p > \overline{c}} U_p f(z)\, dz = \underline{u}_p$$

violating the definition of c_0^*. If $\overline{c} \leq c_0^* - S_{old}^{\max}$ then

$$\underline{V}(h^\tau, c_0^*) = \int_{U_p > c_0^* - S_{old}^{\max}} \min(0, U_p)\, f(z)\, dz < \int_{U_p > c_0^* - S_{old}^{\max}} U_p f(z)\, dz <$$

$$< \int_{U_p > \overline{c}} U_p f(z)\, dz = \underline{u}_P$$

from the definition of $\overline{c}$ (since it's the minimum cost that achieves $\underline{u}_P$ in the spot game). Therefore, we have shown that if $\underline{U} \leq c_0^* - S_{old}^{\max}$ then we have $\underline{V}(h^\tau, c_0^*) < \underline{u}_p$, violating the definition of c_0^*. Finally, to show $c_0^* > \overline{U}$, suppose that $\overline{U} \leq c_0^*$. Then any strategy consistent with this choice would give the $\theta = old$ an utility of 0, while we know we will make the reservation utility to be binding (i.e. choose the commitment cost a little smaller so that the contrarian behavior is enough to reach the reservation utility).

Step 3 : $\Pr(U_p > c_0^* - S_{new}^{\max}, U_{old} < c_0^* - S_{old}^{\max}) > 0$

Follows from the fact that $\overline{U} > c_0^* > c_0^* - S_{new}^{\max}$, Step 2 and the full support assumption.

Step 4: $\overline{U} > c_0^* + S_{old}^{\max}$

Suppose that this is not the case: then

$$\underline{V}(h^\tau, c_0^*) = \int_{U_{old} \in \left(c_0^* - S_{old}^{\max}, c_0^* + S_{old}^{\max}\right)} \min(0, U_p)\, f(z)\, dz$$

but see that this is identical to expression 3.106. Therefore, replicating the same proof as in Step 2, we conclude the result.

Step 5: $\Pr(U_p < c_0^* + S_{new}^{\max}, U_{old} > c_0^* + S_{old}^{\max}) > 0$

Since $\underline{U} < 0$ we clearly have that $\underline{U} < c_0^* + S_{new}^{\max}$. This, together with the Step 5 and the full support assumption proves the result. □

Proof of Lemma 37 . I first show that for any $old-$rationalizable history h^τ we have $\inf_{\beta \in (0,1)} q(h^\tau, \beta) > 0$. I present the proof for when $c^*(h^\tau) = c_0^*$. Suppose not: then there exists an increasing sequence $\beta_n \in (0,1)$ such that $q(h^\tau, \beta_n) > 0 \forall n \in \mathbb{N}$ and $q(h^\tau, \beta_n) \searrow 0$. For all δ define the expected utility for the people $\underline{v}(\beta_n) := \underline{V}(h^\tau, c_0^*) = \underline{u}_p$. For all n we have:

$$\underline{v}(\beta_n) < \int_{U_{old} \in \left(c_0^*(\beta_n) - S_{old}^{\max}(\beta_n), c_0^*(\beta_n) + S_{old}^{\max}(\beta_n)\right)} \min(0, U_p)\, f(z)\, dz + q(h^\tau, \beta_n) \left[\max_{U_P \in [\underline{U}, \overline{U}]} (0, U_p)\right]$$

where the first term is the utility in the middle region, and the second term is the natural bound on all regions (particularly in separation regions). Taking limits as $n \to \infty:$

$$\underline{u}_p = \lim_{n \to \infty} \underline{v}(\beta_n) \leq \mathbb{E}[\min(0, U_p)] < \underline{u}_P$$

reaching a contradiction. □

3.12 Characterization of $\mathcal{V}(s,c)$

In this section I solve and analyze the solution to the programming problem in subsection (3.52)

Proposition 61 (Rationalizable Contrarian Strategy). *Consider the programming problem 3.52. Then*

1. *We can rewrite it as*

$$\mathcal{V}(s,c) = \max_{r(.),n(.)} \mathbb{E}_z\left[U_p r(z)\right] \tag{3.107}$$

$$s.t: \begin{cases} \mathbb{E}_z\left[(U_{old}-c)\, r(z) + n(z)\right] \geq \frac{1}{\beta}s + \underline{\mathbb{W}}_{old} & (PK \text{ for sacrifice}) \\ r(z)\left[U_{old}-c+n(z)\right] \geq 0 \text{ for all } z \in Z & (IC \text{ for } r=1) \\ \left[1-r(z)\right]\left[n(z)-U_{old}+c\right] \geq 0 \text{ for all } z \in Z & (IC \text{ for } r=0) \\ n(z) \in [0, S_{old}^{\max}] \text{ for all } z \in Z & (Feasibility) \end{cases} \tag{3.108}$$

2. *There exist* $\hat{S} \in (0, S_{old}^{\max})$such that if for $s < \hat{S}$ *then the solution policy* $\underline{r}(z)$ *is*

$$\underline{r}(z) = \begin{cases} 1 & \text{if } U_{old}-c > S_{old}^{\max} \\ 1 & \text{if } U_{old}-c \in (-S_{old}^{\max}, S_{old}^{\max}) \text{ and } U_{old} < 0 \\ 0 & \text{if } U_{old}-c \in (-S_{old}^{\max}, S_{old}^{\max}) \text{and } U_{old} > 0 \\ 0 & \text{if } U_{old}-c < -S_{old}^{\max} \end{cases} \tag{3.109}$$

3. *If* $s \in \left[\hat{S}, S_{old}^{\max}\right]$, *there exist a positive constant* $\alpha(s) \in (0,1)$ *such that*

$$\hat{r}(z) = \begin{cases} 1 & \text{if } U_{old}-c > S_{old}^{\max} \\ 1 & \text{if } U_{old}-c \in (-S_{old}^{\max}, S_{old}^{\max}) \text{ and } U_p < \gamma(s)(U_{old}-c) \\ 0 & \text{if } U_{old}-c \in (-S_{old}^{\max}, S_{old}^{\max}) \text{and } U_p > \gamma(s)(U_{old}-c) \\ 0 & \text{if } U_{old}-c < S_{old}^{\max} \end{cases} \tag{3.110}$$

4. *For all* $s \in (0, S^{\max})$ *we have* $\mathbf{c}(s) \in (\overline{c}, S^{\max})$

Proof. Define $n(z) = \frac{\beta}{1-\beta}\left[w(z) - \underline{\mathbb{W}}_{old}\right]$. If $r(z) = 1$ then we can rewrite the enforceability constraint in 3.45 as $(1-\beta)(U_{old}-c) + \beta w(z) \geq \beta \underline{\mathbb{W}}_{old} \iff U_{old} - c + n(z) \geq 0$. Likewise, if $r(z) = 0$ the IC constraint is $\beta w(z) \geq (1-\beta)(U_{old}-c) + \beta \mathbb{W}_{old} \iff n(z) - U_{old} + c \geq 0$. Finally, rewrite (PK) as

$$\mathbb{E}_z\left[(1-\beta)(U_{old}-c)\, r(z) + \beta\left(w(z) - \underline{\mathbb{W}}_{old}\right)\right] \geq \left(\frac{1-\beta}{\beta}\right)s + (1-\beta)\,\underline{\mathbb{W}}_{old} \iff$$

$$\mathbb{E}_z\left[(U_{old}-c)\, r(z) + n(z)\right] \geq \frac{1}{\beta}s + \underline{\mathbb{W}}_{old}$$

See that for any $z : |U_{old} - c| < S_{old}^{\max}$then any $r \in \{0,1\}$ is implementable. However, if $U_{old} > c + S_{old}^{\max}$ then only $r = 1$ is implementable, and if $c - U_{old} < -S_{old}^{\max}$ then only $r = 0$ is implementable. Then, without the promise keeping constraint (PK) the solution to 3.107 is simple:

$$\underline{r}(z) := \begin{cases} 1 & \text{if } U_{old} > c + S_{old}^{\max} \\ 1 & \text{if } |U_{old} - c| < S_{old}^{\max}, U_p < 0 \\ 0 & \text{otherwise} \end{cases}$$

i.e. whenever both policies $r \in \{0,1\}$ are rationalizable, $\theta = old$ picks the worst policy for p. We will refer to this policy as the *rationalizable contrarian policy.* It will be also the solution when $s = 0$ when the policy $\underline{r}$ satisfies (PK) with strict inequality. Define $\underline{n}(z)$ as the implementing continuation for $\underline{r}(z)$ that maximizes $\mathbb{E}\{((U_{old} - c)\underline{r}(z) + \underline{n}(z))\}$. Then, it will be also the solution of 3.107 if and only if

$$s \leq \beta \mathbb{E}_z [(U_{old} - c)\underline{r}(z) + \underline{n}(z)] - \underline{\mathbb{W}}_{old} \equiv \hat{s}$$

showing (2). For (3), ignoring for now the IC constraints, use the Lagrangian method ((Luenberger 1997))

$$\mathcal{L} = \int U_p r(z) f(z) dz - \gamma \left\{ \int [(U_{old} - c) r(z) + n(z)] - \frac{1}{\beta} s - \underline{\mathbb{W}}_{old} \right\}$$

where $\gamma \geq 0$ is the Lagrange multipliers of the problem.

$$\frac{\partial \mathcal{L}}{\partial r(z)} = U_p - \gamma (U_{old} - c)$$

then, if $r(z) = 1$ is implementable, the optimum will be $r(z) = 1 \iff U_p \leq \gamma (U_{old} - c)$. If we want to implement $r = 1$ we then set $n(z) = \min\{0, c - U_p\}$. Then, given γ we solve for $r(z \mid \gamma)$ and $n(z \mid \gamma)$, and we solve for γ using the promise keeping constraint

$$\int [r(z \mid \gamma)(U_{old} - c) + n(z \mid \gamma)] f(z) = \frac{1}{\beta} s + \underline{\mathbb{W}}_{old}$$

which determines γ as a function of s, showing (3). $\square$

Results are better explained using Figures (3-9) , (3-10) and (3-1) below.

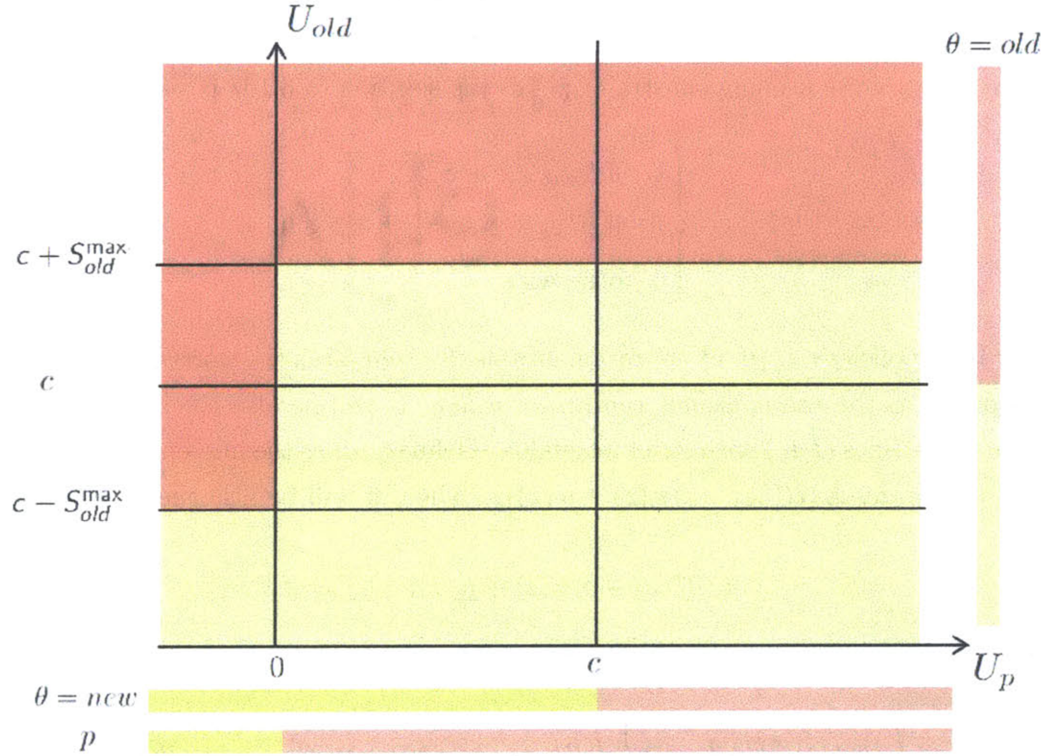


Figure 3-9: Rationalizable Evil Agent strategy, with $s \leq \hat{s}$

See that for U_{old} above $c + S_{old}^{\max}$ and below $c - S_{old}^{\max}$ the unique rationalizable actions for b are $\hat{r}(z) = 1$ (red) and $\hat{r}(z) = 0$ (green) respectively, as we have seen before. When $U_{old} \in (c - S_{old}^{\max}, c + S_{old}^{\max})$, both $r = 0$ and $r = 1$ are rationalizable for any z in this region, by appropriately choosing the expected continuation payoffs. Therefore, the worst strategy that p could expect would be one of a *contrarian*: whenever p wants the green button to be played ($U_p < 0$), then the old type would play the opposite action. We can draw an obvious parallelism to the "evil agent" in the robustness literature of (Hansen and Sargent 2011), with the restriction that instead of a pure evil agent, the *rationalizable evil agent*, that is only contrarian at states in which the utility of doing her most desirable action is not too high.

See that being a rationalizable contrarian is costly for $\theta = old$, since there are regions in which both $p's$ and $\theta = old$ most desired action coincide, as we see in the next figure (regions stressed in darker colors)

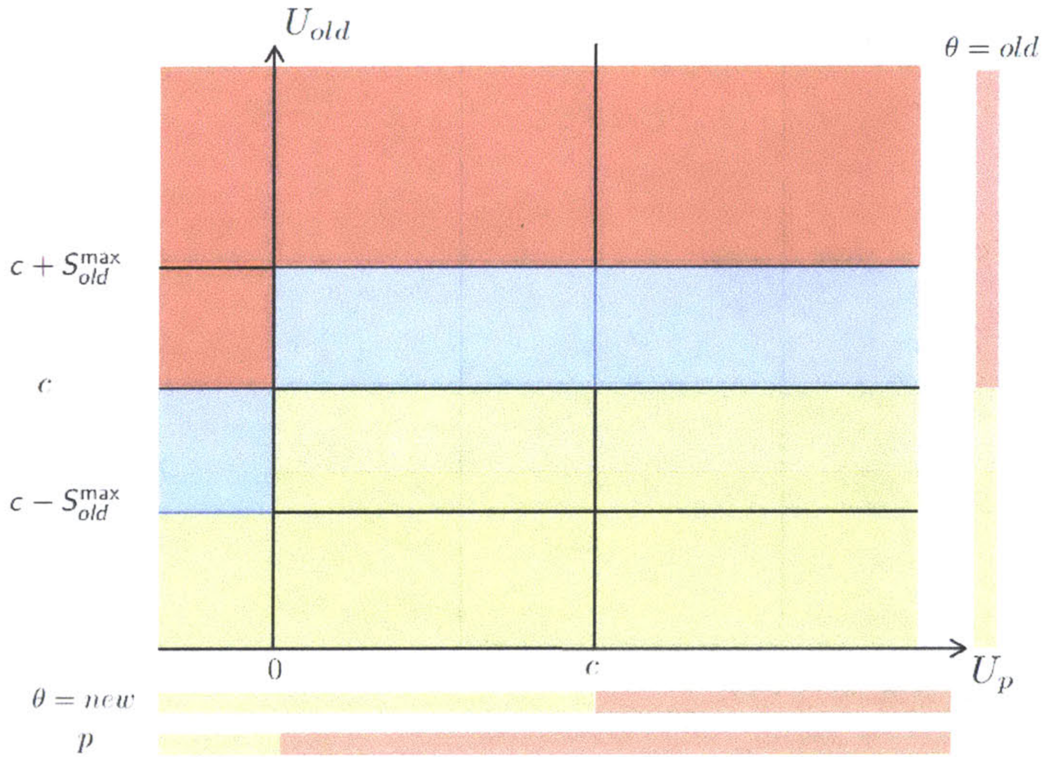


Figure 3-10: Self-contrarian regions for rationalizable evil agent

Then, when sacrifice is high enough, the disutility generated by the dark regions would not be consistent with the observed behavior. Therefore, to satisfy the "promise keeping" constraint, we must allow the "rationalizable evil agent" not to be fully contrarian, and play her desired action in some states, as we see in the figure below.

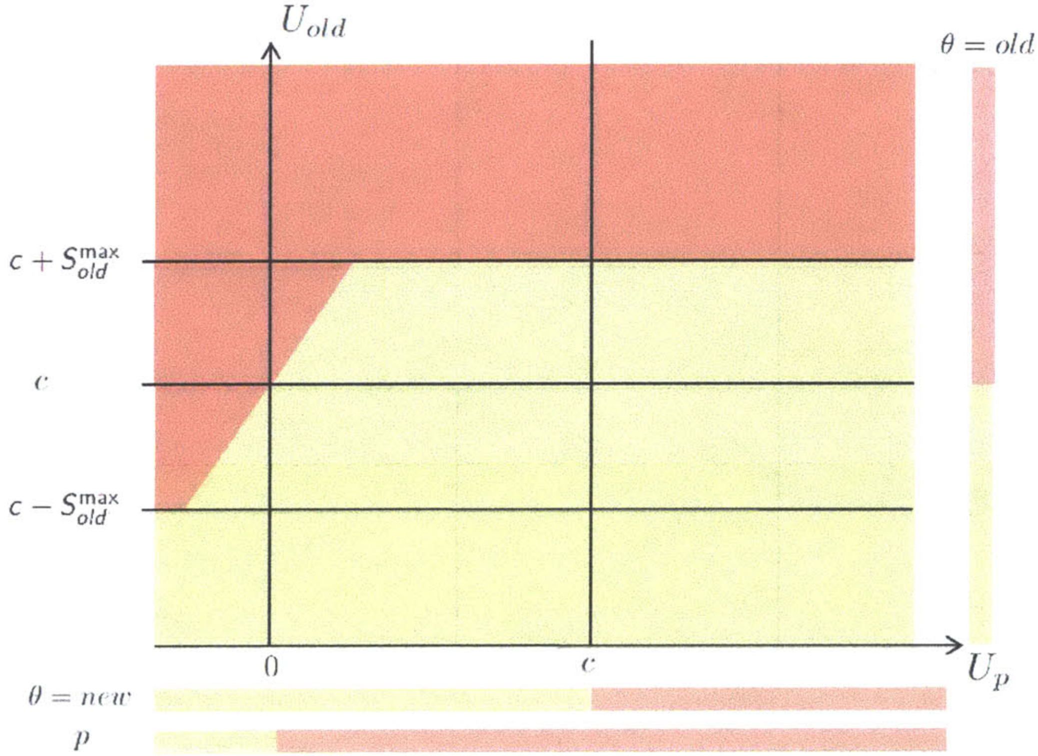


Table 3-1: Rationalizable evil agent strategy, with $s > \hat{s}$

Finally, is easy to see that as the promise keeping constraint becomes more and more binding, the worst type's policy $\hat{r}(.)$ resembles more and more the spot optimum policy $r_{old}^{spot}(.)$. Then, $\mathrm{c}(s) > \overline{c}$ and it approximates it as the promise keeping constraint becomes more binding.

Bibliography

ABREU, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, pp. 383–396.

ABREU, D., D. PEARCE, AND E. STACCHETTI (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, pp. 1041–1063.

ACEMOGLU, D., M. A. DAHLEH, I. LOBEL, AND A. OZDAGLAR (2010): "Bayesian Learning in Social Networks," .

ALVAREZ, F. AND U. J. JERMANN (2000): "Efficiency, Equilibrium, and Asset Pricing with Risk of Default," *Econometrica*, 68, 775–797.

ANDERSON, L. AND C. HOLT (1997): "Information cascades in the laboratory," *The American Economic Review*, 87, 847–862.

ATHEY, S., A. ATKESON, AND P. J. KEHOE (2005): "The Optimal Degree of Discretion in Monetary Policy," *Econometrica*, 73, pp. 1431–1475.

AUMANN, R. AND A. BRANDENBURGER (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, pp. 1161–1180.

AUSUBEL, L. M. AND R. J. DENECKERE (1993): "A generalized theorem of the maximum," *Economic Theory*, 3, 99–107.

BACKUS, D. AND J. DRIFFIL (1984): "Rational Expectations and Policy Credibility Following a Change in Regime," *Review of Economic Studies*, 52, 211–221.

BAKER, G. AND MURPHY (2002): "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 117, 39–84.

BANERJEE, A. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, 797–817.

BANERJEE, A., A. G. CHANDRASEKHAR, E. DUFLO, AND M. JACKSON (2010): "Social networks and microfinance," MIT working paper.

BARRO, R. J. (1986): "Reputation in a model of monetary policy with incomplete information," *Journal of Monetary Economics*, 17, 3 – 20.

Barro, R. J. and D. B. Gordon (1983): "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, 12, 101 – 121.

Battigalli, P. (2003): "Rationalizability in infinite, dynamic games with incomplete information," *Research in Economics*, 57, 1 – 38.

Battigalli, P. and G. Bonanno (1999): "Recent results on belief, knowledge and the epistemic foundations of game theory," *Research in Economics*, 53, 149 – 225.

Battigalli, P. and M. Siniscalchi (1999): "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games," *Journal of Economic Theory*, 88, 188 – 230.

——— (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356 – 391.

——— (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3.

Ben-Porath, E. (1997): "Rationality, Nash Equilibrium and Backwards Induction in Perfect-Information Games," *The Review of Economic Studies*, 64, pp. 23–46.

Berge, C. (1963): *Espaces topologiques (Translated by Patterson, E.M.): Topological Spaces*, Oliver and Boyd.

Bergemann, D. and S. Morris (2005): "Robust Mechanism Design," *Econometrica*, 73, 1771–1813.

——— (2009): "Robust Implementation in Direct Mechanisms," *The Review of Economic Studies*, 76, pp. 1175–1204.

Berger, R. (1981): "A necessary and sufficient condition for reaching a consensus using DeGroot's method," *Journal of the American Statistical Association*, 76, 415–418.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 100.

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

Bull, C. (1987): "The Existence of Self-Enforcing Implicit Contracts," *The Quarterly Journal of Economics*, 102, 147–159.

Canzoneri, M. B. (1985): "Monetary Policy Games and the Role of Private Information," *The American Economic Review*, 75, 1056–1070.

Celen, B. and S. Kariv (2005): "An experimental test of observational learning under imperfect information," *Economic Theory*, 26, 677–699.

Celentani, M. and W. Pesendorfer (1996): "Reputation in Dynamic Games," *Journal of Economic Theory*, 70, 109 – 132.

CHANDRASEKHAR, A. AND R. LEWIS (2010): "Econometrics of sampled networks," MIT working paper.

CHARI, V. V. AND P. J. KEHOE (1990): "Sustainable Plans," *Journal of Political Economy*, 98, pp. 783–802.

——— (1993a): "Sustainable Plans and Debt," *Journal of Economic Theory*, 61, 230–261.

——— (1993b): "Sustainable Plans and Mutual Default," *The Review of Economic Studies*, 60, 175–195.

CHATTERJEE, S. AND E. SENETA (1977): "Towards consensus: Some convergence theorems on repeated averaging," *Journal of Applied Probability*, 89–97.

CHO, I.-K. AND D. M. KREPS (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

CHOI, S., D. GALE, AND S. KARIV (2005): "Behavioral aspects of learning in social networks: An experimental study," *Advances in Applied Microeconomics: A Research Annual*, 13, 25–61.

——— (2009): "Social Learning in Networks: A Quantal Response Equilibrium Analysis of Experimental Data," *WP, May.*

CONLEY, T. AND C. UDRY (2010): "Learning about a new technology: Pineapple in Ghana," *The American Economic Review*, 100, 35–69.

CRUZ, C. (2012): "Social Networks and the Targeting of Illegal Electoral Strategies," *Mimeo.*

DEBORTOLI, D. AND R. NUNES (2010): "Fiscal policy under loose commitment," *Journal of Economic Theory*, 145, 1005 – 1032.

DEGROOT, M. (1974): "Reaching a consensus," *Journal of the American Statistical Association*, 69, 118–121.

DEMARZO, P., D. VAYANOS, AND J. ZWIEBEL (2003): "Persuasion Bias, Social Influence, and Unidimensional Opinions*," *Quarterly journal of economics*, 118, 909–968.

FOLLAND, G. B. (1999): *Real Analysis : modern techniques and their applications - 2nd Edition*, John Wiley & Sons, Inc.

FRIEDKIN, N. AND E. JOHNSEN (1997): "Social positions in influence networks," *Social Networks*, 19, 209–222.

FUDENBERG, D. AND D. K. LEVINE (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, pp. 759–778.

GALE, D. AND S. KARIV (2003): "Bayesian learning in social networks," *Games and Economic Behavior*, 45, 329–346.

Golub, B. and M. Jackson (2009): "How homophily affects learning and diffusion in networks," *Working Papers.*

——— (2010): "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2, 112–149.

Haile, P., A. Hortaçsu, and G. Kosenok (2008): "On the empirical content of quantal response equilibrium," *The American Economic Review*, 98, 180–200.

Hansen, L. and T. Sargent (2011): *Robustness*, Princeton university press.

Harsanyi, J. (1967/68): "Games with Incomplete Information Played by Bayesian Agents," *Management Science*, 14, 159–182, 320–334, 486–502.

Heifetz, A. (1993): "The bayesian formulation of incomplete information - The non-compact case," *International Journal of Game Theory*, 21, 329–338.

Hung, A. and C. Plott (2001): "Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions," *American Economic Review*, 91, 1508–1520.

Jackson, M. (2008a): *Social and economic networks*, Princeton University Press.

Jackson, M., T. Barraquer, and X. Tan (2010): "Social Capital and Social Quilts: Network Patterns of Favor Exchange," .

Jackson, M. O. (2008b): *Social and Economic Networks*, Princeton University Press.

King, R. G., Y. K. Lu, and E. S. Pasten (2012): "Policy design with private sector skepticism in the textbook New Keynesian model," *Working P.*

Kocherlakota, N. R. (1996): "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63, 595–609.

Krackhardt, D. (1987): "Cognitive social structures," *Social Networks*, 9, 109–134.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory*, 27, 245 – 252.

Kreps, D. M. and R. Wilson (1982): "Reputation and imperfect information," *Journal of Economic Theory*, 27, 253 – 279.

Kubler, D. and G. Weizsacker (2004): "Limited depth of reasoning and failure of cascade formation in the laboratory," *Review of Economic Studies*, 71, 425–441.

Kydland, F. E. and E. C. Prescott (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85, pp. 473–492.

Levin (2003): "Relational Incentives Contracts," *The American Economic Review*, 93, 835–857.

Ligon, E. (1998): "Risk Sharing and Information in Village Economics," *Review of Economic Studies*, 65, 847–64.

Ligon, E., J. P. Thomas, and T. Worrall (2000): "Mutual Insurance, Individual Savings, and Limited Commitment," *Review of Economic Dynamics*, 3, 216 – 246.

Lohmann, S. (1992): "Optimal Commitment in Monetary Policy: Credibility versus Flexibility," *The American Economic Review*, 82, pp. 273–286.

Lu, Y. K. (2012): "Optimal Policy with Credibility Concerns," *HK UST Working Paper.*

Luenberger, D. (1997): *Optimization by vector space methods*, Wiley-Interscience.

Mailath, G. J. and L. Samuelson (2006): *Repeated Games and Reputations: Long Run Relationships*, Oxford.

Mertens, J. F. and S. Zamir (1985): "Formulation of Bayesian analysis for games with incomplete information," *International Journal of Game Theory*, 14, 1–29.

Milgrom, P. and J. Roberts (1982): "Predation, reputation, and entry deterrence," *Journal of Economic Theory*, 27, 280 – 312.

Möbius, M., T. Phan, and A. Szeidl (2011): "Treasure Hunt," *Mimeo.*

Mossel, E., A. Sly, and O. Tamuz (Forthcoming): "Asymptotic learning on Bayesian social networks," *Probability Theory and Related Fields.*

Mossel, E. and O. Tamuz (2010): "Effcient Bayesian Learning in Social Networks with Gaussian Estimators," *arXiv:1002.0747.*

Mueller-Frank, M. and C. Neri (2012): "Consensus and Disagreement in Social Networks: An Experimental and Theoretical Study," *Mimeo.*

Munkres, J. R. (2000): *Topology - Second Edition*, Prentice Hall Inc.

Obstfeld, M. and K. Rogoff (1996): *Foundations of International Macroeconomics*, MIT Press.

Pavan, A. and G.-M. Angeletos (2012): "Selection-free predictions in global games with endogenous information and multiple equilibria," *Theoretical Economics, Forthcoming.*

Pearce, D. G. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, pp. 1029–1050.

Penta, A. (2011): "Robust Dynamic Mechanism Design," *Working Paper.*

——— (2012): "Higher Order Uncertainty and Information: Static and Dynamic Games," *Econometrica*, 80, 631–660.

Persson, T. and G. Tabellini (1993): "Designing institutions for monetary stability," *Carnegie-Rochester Conference Series on Public Policy*, 39, 53 – 84.

Phelan, C. (2006): "Public trust and government betrayal," *Journal of Economic Theory*, 130, 27 – 43.

ROGOFF, K. (1985): "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, 100, pp. 1169–1189.

SARGENT, T. J. AND L. LJUNGQVIST (2004): *Recursive Macroeconomic Theory*, MIT Press.

SMITH, L. AND P. SORENSEN (2000): "Pathological outcomes of observational learning," *Econometrica*, 68, 371–398.

STOKEY, N. L. (1989): "Reputation and Time Consistency," *The American Economic Review*, 79, pp. 134–139.

——— (1991): "Credible public policy," *Journal of Economic Dynamics and Control*, 15, 627 – 656.

STOKEY, N. L., R. E. LUCAS, AND E. C. PRESCOTT (1989): *Recursive methods in economics dynamics*, Presidents and Fellows of Harvard College.

THOMAS, J. AND T. WORRALL (1988): "Self-Enforcing Wage Contracts," *Review of Economic Studies (1988)*, 55(4), 541–554.

WALSH, C. (1995): "Optimal Contracts for Central Bankers," *The American Economic Review*, 85, 150–167.

WEINSTEIN, J. AND M. YILDIZ (2007): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75, 365–400.

——— (2012): "Robust Predictions in infinite-horizon games - An unrefinable folk theorem," *Review of Economic Studies, forthcoming.*

WOLITZKY, A. (2012): "Reputational Bargaining With Minimal Knowledge of Rationality," *Econometrica*, 80, 2047–2087.

XANDRI, J. P. (2012): "A Note on Interactive Epistemology in Dynamic Games with Discounting," *MIT Working Paper.*