Essays on Dynamic Games and Reputations

by

Di Pei

B.A. Mathematics, Tsinghua University (2011)
M.A. Economics, Toulouse School of Economics (2013)

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

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

© Di Pei, MMXVIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Author . . . .
Department of Economics
June 1, 2018

Certified by . . . .
Daron Acemoglu
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Certified by . . . .
Drew Fudenberg
Paul A. Samuelson Professor of Economics
Thesis Supervisor

Certified by . . . .
Juuso Toikka
Gary Loveman Career Development Associate Professor of Economics
Thesis Supervisor

Accepted by . . . .
Ricardo Caballero
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays on Dynamic Games and Reputations

by

Di Pei

Submitted to the Department of Economics
on June 1, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

This thesis consists of three essays on dynamic games with incomplete information. In Chapter 1, I study reputation effects when individuals have persistent private information that matters for their opponents' payoffs. I examine a repeated game between a patient informed player and a sequence of myopic uninformed players. The informed player privately observes a persistent state, and is either a strategic type who can flexibly choose his actions or is one of the several commitment types that mechanically plays the same action in every period. Unlike the canonical models on reputation effects, the uninformed players' payoffs depend on the state. This interdependence of values introduces new challenges to reputation building, namely, the informed player could face a trade-off between establishing a reputation for commitment and signaling favorable information about the state. My results address the predictions on the informed player's payoff and behavior that apply across all Nash equilibria. When the stage game payoffs satisfy a monotone-supermodularity condition, I show that the informed long-run player can overcome the lack-of-commitment problem and secure a high payoff in every state and in every equilibrium. Under a condition on the distribution over states, he will play the same action in every period and maintain his reputation for commitment in every equilibrium. If the payoff structure is unrestricted and the probability of commitment types is small, then the informed player's return to reputation building can be low and can provide a strict incentive to abandon his reputation.

In Chapter 2, I study the dynamics of an agent's reputation for competence when the labor market's information about his performance is disclosed by an intermediary who cannot commit. I show that this game admits a unique Markov Perfect Equilibrium (MPE). When the agent is patient, his effort is inverse U-shaped, while the rate of information disclosure is decreasing over time. I illustrate the inefficiencies of the unique MPE by comparing it with the equilibrium in the benchmark scenario where the market automatically observes all breakthroughs. I characterize a tractable subclass of non-Markov Equilibria and explain why allowing players to coordinate on payoff-irrelevant events can improve efficiency on top of the unique MPE and the exogenous information benchmark. When the intermediary can commit, her optimal Markov disclosure policy has a deadline, after which no breakthrough will be disclosed. However, deadlines are not incentive compatible in the game without commitment, illustrating a time inconsistency problem faced by the intermediary. My model can be applied to professional service industries, such as law and consulting. My results provide an explanation to the observed wage and promotion patterns in Baker, Gibbs and Holmström (1994).

In Chapter 3, I study repeated games in which a patient long-run player (e.g. a firm) wishes to win the trust of some myopic opponents (e.g. a sequence or a continuum of consumers) but has a strict incentive to betray them. Her benefit from betrayal is persistent over time and is her private information. I examine the extent to which persistent private information can overcome this lack-of-commitment problem. My main result characterizes the set of payoffs a patient long-run player can attain in equilibrium. Interestingly, every type's highest equilibrium payoff only depends on her true benefit from betrayal and the *lowest* possible benefit in the support of her opponents' prior belief. When this lowest possible benefit vanishes, every type can approximately attain her Stackelberg commitment payoff. My finding provides a strategic foundation for the (mixed) Stackelberg commitment types in the reputation models, both in terms of the highest attainable payoff and in terms of the commitment behaviors. Compared to the existing approaches that rely on the existence of *crazy types* that are either irrational or have drastically different preferences, there is common knowledge of rationality in my model, and moreover, players' ordinal preferences over stage game outcomes are common knowledge.

Thesis Supervisor: Daron Acemoglu
Title: Elizabeth and James Killian Professor of Economics

Thesis Supervisor: Drew Fudenberg
Title: Paul A. Samuelson Professor of Economics

Thesis Supervisor: Juuso Toikka
Title: Gary Loveman Career Development Associate Professor of Economics

Acknowledgments

I am deeply indebted to my advisors Drew Fudenberg, Daron Acemoglu and Juuso Toikka for their continuous guidance, encouragement and support in the past five years.

I was fortunate enough to take Drew's game theory course at Harvard in my first year. His lectures stimulated my interests in repeated games and reputations, and more importantly, equipped me with the technical tools to do frontier research in economic theory. Drew became the natural choice for my primary advisor after I decided to write my job market paper on reputation effects. I was constantly impressed by the sharpness of his mind, the clarity of his thinking and the speed with which he learns new things. Drew has always been approachable to graduate students and enjoys a reputation for being generous with his time and belongings. Every time I emailed him a 10-page proof I wrote, he would always get back to me within half an hour with very detailed comments. I also appreciate his continuous investments into the Harvard-MIT theory community, by, for example, running weekly theory reading groups, inviting Harvard students to present in the MIT theory lunch, hosting annual theory parties in his house, etc.

My story with Daron started after finishing my general exams. I emailed him in June and asked for a summer RAship. He happily agreed and we started talking several potential applications of mechanism design to political economy problems, which eventually developed into a coauthored paper on constitutional design. Daron is the one of the broadest economists I have ever met: the conversations in his office could start from a discussion on some deep mathematical theorems and end with mapping the results I just obtained to features in the Bill of Rights or the Peloponnesian War. Every discussion with Daron is an invaluable experience. I learnt a lot in terms of how to ask an economically interesting question, how to abstract a complicated economic problem into a simple model and how to justify the economic significance of the results being derived.

Juuso has always been my hero, both as an economic theorist and as a teacher/advisor. He is a deep thinker, very rigorous in stating theorems and writing proofs, yet at the same time, being accessible and has the talent of explaining deep theoretical ideas in intuitive and non-technical ways. I was lucky enough to be his teaching assistant in contract theory for three consecutive years, from which I learnt a lot on how to give good lectures and clear presentations. We spent hours in his office, going through the modeling assumptions and the proof techniques of nearly every project I have attempted in graduate school, no matter whether it has eventually turned into a paper or not. I have always admired his dedication to scientific research, his colleagues, his students and his family, and I hope I could follow his example after starting my job at Northwestern University.

I would also like to thank Christian Hellwig, Jean Tirole, Robert Gibbons and Alexander Wolitzky, who have all been instrumental in helping me grow as an economist. I would definitely not be in this position had it not been for Christian and Jean, my advisors in Toulouse. They introduced me to research in economics, showed me how to become a good scholar and had given me numerous advice on how to write papers, how to give talks, how to get along with colleagues, etc. Bob is a great mentor and has always been my first choice whenever I am seeking for career advice. He was always open to discussing research ideas even when it became clear that I was not on the path to becoming an organizational economist. In my opinion, Alex is one of the best and most productive theorists in his generation, yet at the same time, he is always so humble and hard-working. Despite he is not officially my advisor, he would always email me constructive comments after my talks and has read through the drafts of almost every paper I wrote in graduate school. During the stressful job market season, he was one of the most encouraging person starting from the preparation stage to the time I got my first job offer.

Aside from the people I just mentioned, I have also been fortunate to receive helpful comments on my research from many others, including Jie Bai, Abhijit Banerjee, Alessandro Bonatti, Vivek

Bhattacharya, Sydnee Caldwell, Gonzalo Cisternas, Mehmet Ekmekci, Glenn Ellison, Ben Golub, Kevin He, Bengt Holmström, Daria Khromenkova, Ernest Liu, Birger Wernerfelt and Muhamet Yildiz. I am also grateful to the help and support from my friends, such as Seth Blumberg, Elaine Chung, Daniel Clark, Jetlir Duraj, Masao Fukui, Chishio Furukawa, Arda Gitmez, Omer Karaduman, Shuo Liu, Yueran Ma, Yusuke Narita, Cara Nikolaus, Lauren Russell, Elizabeth Setren, Kosti Takala, Linh Tô, Carmen Wang, and the list goes on.

Finally, I would like to thank my family for their unconditional love and support over the past five years, especially, of course, my parents and my grandparents. They have all sacrificed a lot to help me succeed, and no words can express my gratitude to their love and care.

Contents

List of Figures

List of Tables

Chapter 1

Reputation Effects under Interdependent Values

This chapter studies reputation effects when individuals have persistent private information that matters for their opponents' payoffs. I examine a repeated game between a patient informed player and a sequence of myopic uninformed players. The informed player privately observes a persistent state, and is either a strategic type who can flexibly choose his actions or is one of the several commitment types that mechanically plays the same action in every period. Unlike the canonical models on reputation effects, the uninformed players' payoffs depend on the state. This *interdependence of values* introduces new challenges to reputation building, namely, the informed player could face a trade-off between establishing a reputation for commitment and signaling favorable information about the state. My results address the predictions on the informed player's payoff and behavior that apply across all Nash equilibria. When the stage game payoffs satisfy a *monotone-supermodularity* condition, I show that the informed long-run player can overcome the lack-of-commitment problem and secure a high payoff in every state and in every equilibrium. Under a condition on the distribution over states, he will play the same action in every period and maintain his reputation for commitment in every equilibrium. If the payoff structure is unrestricted and the probability of commitment types is small, then the informed player's return to reputation building can be low and can provide a strict incentive to abandon his reputation.

I am indebted to Daron Acemoglu, Drew Fudenberg, Juuso Toikka and Alex Wolitzky for guidance and support. I thank Jie Bai, Abhijit Banerjee, Vivek Bhattacharya, Alessandro Bonatti, Gonzalo Cisternas, Dan Clark, Jetlir Duraj, Mehmet Ekmekci, Glenn Ellison, Chishio Furukawa, Bob Gibbons, Ben Golub, Yingni Guo, Kevin He, Christian Hellwig, Daria Khromenkova, Annie Liang, Shuo Liu, Ben Roth, Bruno Strulovici, Jean Tirole, Birger Wernerfelt, Muhamet Yildiz and my seminar participants at MIT, Princeton, Harvard, Northwestern, UC Berkeley, Kellogg School of Management and USC for helpful comments. All errors are mine.

1.1 Introduction

Economists have long recognized that good reputations can lend credibility to people's threats and promises. This intuition has been formalized in a series of works starting with Kreps and Wilson (1982), Milgrom and Roberts (1982), Fudenberg and Levine (1989) and others, who show that having the option to build a reputation dramatically affects a patient individual's gains in long-term relationships. Their reputation results are *robust* as they apply across all equilibria, which enables researchers to make robust predictions in many decentralized markets where there is no mediator helping participants to coordinate on a particular equilibrium.

However, previous works on *robust* reputation effects all restrict attention to private value environments. This excludes situations where reputation builders have persistent private information that directly affects their opponents' payoffs. For example in the markets for food and custom software, merchants can benefit from a reputation for providing good customer service, but they also want to signal their products have high quality. The latter directly affects consumers' willingness to pay and is usually the merchants' private information (Banerjee and Duflo 2000, Bai 2016). In the pharmaceutical, cable TV and passenger airline industries, incumbent firms could benefit from committing to fight potential entrants, but are also better informed about the market demand curve, such as the price elasticities, the effectiveness and spillover of advertising (Ellison and Ellison 2011, Seamans 2013, Sweeting, Roberts and Gedge 2016), etc. As a result, incumbent firms' choices of prices, quantities and the intensity of advertising not only show their resolve to fight entrants but also signal their private information about demand. Understanding how the interactions between reputation building and signalling affect economic agents' reputational incentives is important both for firms in designing business strategies and for policy makers in evaluating the merits of quality-control programs and anti-trust regulations.

Motivated by these applications, this paper addresses the robust predictions in reputation games where a player has persistent private information about his opponents' payoffs. In my model, a patient long-run player (player 1, he, seller, incumbent) interacts with a sequence of short-run players (player 2, she/they, buyers, entrants). Unlike the canonical reputation models, I study *interdependent value* environments in which player 1 privately observes a perfectly persistent state (product quality, market demand) that directly affects player 2's payoff. Player 1 is either one of the *strategic types* who maximizes his discounted payoff and will be referred to by the state he observes, or is *committed* to play a state-contingent stationary strategy. Player 2 updates her belief by observing all the past actions. I show that (1) the robust reputation effects on player 1's payoffs extend to a class of interdependent value games despite the existence of a trade-off between commitment and signalling, (2) reputation can also lead to robust and accurate predictions on player 1's equilibrium behavior.

To illustrate the challenges, consider an example of an incumbent firm (player 1) facing a sequence of potential entrants. Every entrant chooses between staying out (O) and entering the market (E). Her preference between O and E depends not only on the incumbent's business strategy, which is either fight (F) or accommodate (A), but also on the market demand curve (the state θ, can be price elasticity, market size, etc.), which is fixed over time and is either high (H) or low (L). This is modeled as the following entry deterrence game:

$\theta = \text{High}$	Out	Enter
Fight	$2, 0$	$0, \mathbf{-1}$
Accommodate	$3, 0$	$1,\ 2$

$\theta = \text{Low}$	Out	Enter
Fight	$2-\eta, 0$	$-\eta, \mathbf{1}$
Accommodate	$3, 0$	$1, 2$

where $\eta \in \mathbb{R}$ is a parameter. When $\theta = H$ is common knowledge (call it the *private value benchmark*), the incumbent faces a lack-of-commitment problem in the stage game: His payoff from the unique Nash equilibrium (A, E) is 1. This is strictly lower than his payoff by committing to fight, which provides his opponent an incentive to stay out and he will receive his *commitment payoff* equal to 2. Fudenberg and Levine (1989) show that reputation can solve this lack-of-commitment problem by establishing the following *commitment payoff theorem*: if the incumbent is non-strategic and fights in every period with positive probability, then a patient strategic incumbent can secure his commitment payoff in *every* Nash equilibrium of the repeated game. Intuitively, if the strategic incumbent imitates the non-strategic one, then he will eventually convince the entrants that F will be played with high enough probability and the latter will best respond by staying out.

The above logic no longer applies when θ is the incumbent's private information. This is because an entrant's best reply to F depends on θ (it is O when $\theta = H$ and E when $\theta = L$), which is signalled through the incumbent's past actions. In situations where fighting is interpreted as a signal of state L,[1] an entrant will have an incentive to play E despite being convinced that F will be played. As a result, the incumbent's return from always fighting will be low. Furthermore, obtaining robust and accurate predictions on the incumbent's equilibrium behavior faces additional challenges as he is *repeatedly signalling* the state. This could lead to multiple possible behaviors. Even the commitment payoff theorem cannot imply that he will maintain his reputation for fighting in every equilibrium, as a strategy that can secure himself a high payoff is not necessarily his optimal strategy.

In Section 1.3, I examine when the commitment payoff theorem applies to *every* payoff function of the long-run player (i.e. it is *fully robust*) without any restrictions on the game's payoff structure.[2] Theorem 1.1 provides a sufficient and (almost) necessary condition for full robustness, which requires that the prior likelihood ratio between each *bad strategic type* and the commitment type be below a cutoff.[3] According to this result, securing the commitment payoff from a mixed action occurs under more demanding conditions than that from a nearby pure action. This implies that small trembles of pure commitment types can lead to a large decrease in the strategic long-run

[1] This is a serious concern since player 1's action today can affect players' future equilibrium play. Equilibria in which player 2 attaches higher probability to state L after observing F are constructed in Appendix A.7.

[2] Full robustness is an important property of Fudenberg and Levine (1989)'s result, which ensures the validity of the commitment payoff bound against (1) modeling misspecifications of the long-run player's payoff function, (2) short-run players entertaining incorrect beliefs about the long-run player's payoff function. This includes, for example, the incumbent's cost of production and returns from advertising, the seller's cost of exerting high effort, all of which are hard to know from an outsider's perspective.

[3] Formally, a strategic type is *bad* if player 2's best reply to the commitment action under his state is different from her best reply when she is facing the commitment type. My conditions are 'almost necessary' as they leave out a degenerate set of beliefs.

player's guaranteed payoff. Another interesting observation is that playing some actions in the support of the mixed commitment action can *increase* the aforementioned likelihood ratios. Therefore, mixed commitment payoffs cannot be guaranteed by replicating the commitment strategy, making the existing techniques in Fudenberg and Levine (1989, 1992), Gossner (2011) inapplicable. To overcome these difficulties, my proof of the sufficiency part makes use of martingale techniques and the central limit theorem to construct a *non-stationary* strategy such that player 1 can achieve three goals simultaneously: (1) avoiding negative inferences about the state, (2) matching the frequency of his actions to the mixed commitment action, (3) player 2's prediction about his actions is close to the mixed commitment action in all but a bounded number of periods.

Theorem 1.1 has two interpretations. First, starting with the private value reputation game in Fudenberg and Levine (1989), it evaluates the robustness of their main insight under a richer set of perturbations. Namely, player 2 can entertain the possibility that her opponent is another strategic type who has private information about her payoff. My result implies that their fully robust reputation result extends when these interdependent value perturbations are unlikely compared to the commitment types, and vice versa. Second, one can also start with a repeated incomplete information game with interdependent values and perturb it with commitment types. According to this view, every commitment type is arbitrarily unlikely compared to every strategic type. Theorem 1.1 then implies that in some equilibria, player 1's return from reputation building is low, and in fact, he will have a strict incentive to abandon his reputation. Therefore, reputation cannot guarantee that player 1 can overcome the lack-of-commitment problem even when he is arbitrarily patient.

This second interpretation motivates the study of games with more specific payoff structures. In Section 1.4, I focus on stage games with *monotone-supermodular* payoffs (*MSM* for short). This requires that the states and every player's actions be ranked such that (1) player 1's payoff is strictly increasing in player 2's action but is strictly decreasing in his own action (or *monotonicity*), and (2) the action profile and the state are complements in player 1's stage game payoff function, and player 2 has a stronger incentive to play a higher action when the state is higher or when player 1's action is higher (or *supermodularity*). In the entry deterrence example, if we rank the states and actions according to $H \succ L$, $F \succ A$ and $O \succ E$, then MSM translates into $\eta > 0$, which is the case when θ is the price elasticity of demand, the market size, the effectiveness of advertising, etc. MSM is also satisfied in buyer-seller games where providing good service is less costly for the seller when his product quality is high, which fits into the custom software industry and the restaurant industry.

My results establish robust predictions on player 1's equilibrium payoff and behavior when there exists a commitment type that plays the highest action in every period. I consider two cases. When the high states are relatively more likely compared to the low states (the *optimistic prior* case), Theorem 1.2 shows that a patient player 1 can guarantee his commitment payoff from playing the highest action in every state and in every equilibrium. In the example, when state H is more likely than state L, player 1 receives at least 2 in state H and $\max\{2-\eta, 1\}$ in state L. This payoff bound applies even when every commitment type is *arbitrarily unlikely* relative to every strategic type. It is also tight in the sense that no strategic type can guarantee himself a strictly higher equilibrium payoff by establishing a reputation for playing another pure commitment action.[4]

In the complementary scenario (the *pessimistic prior* case), Theorem 1.3 shows that when player 1 is patient and the probability of commitment is small (1) his equilibrium payoff equals to the highest equilibrium payoff in the benchmark game without commitment types (Theorem 1.3 and Proposition 1.2); (2) his on-path behavior is the same across all Nash equilibria.[5] According to this unique behavior, there exists a cutoff state (in the example, state L) such that the strategic player 1 plays the highest action in every period if the state is above this cutoff, plays the lowest action in every period if the state is below this cutoff, and mixes between playing the highest action in every period and playing the lowest action in every period at the cutoff state. That is to say, player 1 will behave consistently and maintain his reputation for commitment in all equilibria.

The intuition behind this behavioral uniqueness result is the following *disciplinary effect*: (1) player 1 can obtain a high continuation payoff by playing the highest action, (2) but it is impossible for him to receive a high continuation payoff after he has failed to do so, as player 2's belief about the state will become even more pessimistic than her prior. The first part is driven by the commitment type and the second is because the low states are more likely. This contrasts with Fudenberg and Levine (1989) and the optimistic prior case where deviating from the commitment action may lead to an optimistic posterior, after which a patient player 1 can still receive a high continuation payoff. As a result, player 1 can have multiple on-path behaviors, and in many sequential equilibria, he may have a strict incentive to behave inconsistently and abandon his reputation.

Conceptually, the above comparison suggests that interdependent values can contribute to the *sustainability of reputation*. This channel is novel compared to those proposed in the existing literature, such as impermanent commitment types (Mailath and Samuelson 2001, Ekmekci, et al. 2012), competition between informed players (Hörner 2002), incomplete information about the informed player's past behavior (Ekmekci 2011) and others.[6]

A challenge to prove Theorems 1.2 and 1.3 comes from the observation that a repeated supermodular game is *not* supermodular. This is because player 1's action today can have persistent effects on future equilibrium play. I apply a result in a companion paper (Liu and Pei 2017) which states that if a *1-shot signalling game* has MSM payoffs, then the sender's equilibrium action must be non-decreasing in the state. In a repeated signalling game with MSM stage game payoffs, this result implies that in equilibria where playing the highest action in every period is optimal for player 1 in a low state (call them *regular equilibria*), then he must be playing the highest action with probability 1 at every on-path history in every higher state. Therefore, in every regular equilibrium, player 2's posterior about the state will never decrease if player 1 has always played the highest action. Nevertheless, there can also exist *irregular equilibria* where playing the highest action in

[4]This conclusion extends to any other mixed commitment action if in the stage game, the long-run player strictly prefers the highest action profile to the lowest action profile in every state.

[5]Theorem 1.3 states the result when there is one commitment action. Theorem A.1 (Appendix A.4.2) allows for multiple commitment actions and shows that when the total probability of commitment is small enough, player 1's payoff and on-path behavior are *almost* the same across all equilibria. If all commitment types are pure, then payoff and on-path behavior are the same across all equilibria.

[6]In contrast to these papers and Cripps, Mailath and Samuelson (2004), I adopt a more robust standard for reputation sustainability by requiring that it be sustained in every equilibrium.

every period is *not* optimal in any low state, and it is possible that at some on-path histories, it will lead to a deterioration of player 2's belief about the state. To deal with this complication, my proof shows that in every irregular equilibrium, if player 1 has never deviated from the highest action, then player 2's belief about the state can never fall below a cutoff.

To summarize, we know that in the optimistic prior case, player 2's posterior cannot become too pessimistic given that player 1 has always played the highest action, no matter whether the equilibrium is regular or irregular. Therefore, if player 1 plays the highest action in every period, he can convince player 2 that the highest action will be played in the future and at the same time, player 2's posterior belief about the state will remain optimistic, which leads to the commitment payoff theorem. However, due to the existence of irregular equilibria, player 1 has multiple equilibrium behaviors. In the pessimistic prior case, the necessary condition for irregular equilibria is violated in the first period. Therefore, irregular equilibria do not exist and every regular equilibrium will lead to the same equilibrium payoff and on-path behavior.

My work contributes to the existing literature from several different angles. From a modeling perspective, it unifies two existing approaches to the study of reputation, differing mainly in the interpretation of the informed player's private information. Pioneered by Fudenberg and Levine (1989), the literature on reputation refinement focuses on private value environments and studies how a reputation for commitment affects a patient informed player's payoff in all equilibria.[7] A separate strand of works on dynamic signalling games, including Bar-Isaac (2003), Lee and Liu (2013), Pei (2015) and Toxvaerd (2017), examines the effects of persistent private information about payoff-relevant variables (such as talent, quality, market demand) on the informed player's behavior. However, these papers have focused on some particular equilibria rather than on the common properties of all equilibria. In contrast, I introduce a framework that incorporates commitment over actions and persistent private information about the uninformed players' payoffs. In games with MSM payoffs, I derive robust predictions on the informed player's payoff and behavior that apply across all Nash equilibria.

In the study of repeated Bayesian games with interdependent values,[8] my reputation results can be interpreted as an equilibrium refinement, just as Fudenberg and Levine (1989) did for the repeated complete information games studied in Fudenberg, Kreps and Maskin (1990). By allowing the informed long-run player to be non-strategic and mechanically playing a state-contingent stationary strategy, Theorems 1.1 and 1.2 show that reputation effects can sharpen the predictions

[7]The commitment payoff theorem has been extended to environments with imperfect monitoring (Fudenberg and Levine 1992, Gossner 2011), frequent interactions (Faingold 2013), long-lived uninformed players (Schmidt 1993a, Cripps, Dekel and Pesendorfer 2005, Atakan and Ekmekci 2012), weaker solution concepts (Watson 1993), etc. Another strand of works characterizes Markov equilibria (in infinite horizon games) or sequential equilibria (in finite horizon games) in private value reputation games with a (pure) stationary commitment type, which includes Kreps and Wilson (1982), Milgrom and Roberts (1982), Barro (1986), Schmidt (1993b), Phelan (2006), Liu (2011), Liu and Skrzypacz (2014), etc. See Mailath and Samuelson (2006) for an overview.

[8]This is currently a challenging area and not much is known except for 0-sum games (Aumann and Maschler 1995, Pęski and Toikka 2017), undiscounted games (Hart 1985), belief-free equilibrium payoff sets in games with two equally patient players (Hörner and Lovo 2009, Hörner et al. 2011). In ongoing work (Pei 2016), I characterize the limiting equilibrium payoff set in a repeated Bayesian game between a patient long-run player and a sequence of short-run players when the stage game has MSM payoffs.

on a patient player's equilibrium payoff. Theorem 1.3 advances this research agenda one step further by showing that reputation effects can also lead to accurate predictions on a patient player's equilibrium behavior, which is a distinctive feature of interdependent value models.

In terms of the applications, my result offers a robust explanation to Bain (1949)'s classical observation that "*...established sellers persistently ... forego high prices ... for fear of thereby attracting new entry to the industry and thus reducing the demands for their outputs and their own profit*". This will only happen in some non-renegotiation proof equilibria under private values, but will happen in every equilibrium when the incumbent has private information about demand and the potential entrants are optimistic about their prospects of entry. Similarly, in the study of firm-consumer relationships, my result provides a robust foundation for Klein and Leffler (1981)'s reputational capital theory, which assumes that consumers will coordinate and punish the firm after observing low effort. This will happen in every equilibrium when consumers are skeptical enough about the product quality, which the firm privately knows. I will elaborate more on these in subsection 1.4.5.

1.2 The Model

Time is discrete, indexed by $t = 0, 1, 2....$ An infinitely-lived long-run player (player 1, he) with discount factor $\delta \in (0, 1)$ interacts with a sequence of short-run players (player 2, she), one in each period. In period t, players simultaneously choose their actions $(a_{1,t}, a_{2,t}) \in A_1 \times A_2$. Both A_1 and A_2 are finite sets with $|A_i| \geq 2$ for $i \in \{1, 2\}$. Players have access to a public randomization device, with $\xi_t \in \Xi$ as the realization in period t.

States, Strategic Types & Commitment Types: Let $\theta \in \Theta$ be the state of the world, which is perfectly persistent and is player 1's private information. I assume that Θ is a finite set. Player 1 is either *strategic*, in which case he can flexibly choose his action in every period, or he is *committed* to play the same action in every period, which can be pure or mixed and can be state contingent.

I abuse notation by using θ to denote the *strategic type* who knows that the state is θ (or *type* θ). As for commitment, every *commitment type* is defined based on the (mixed) action he plays. Formally, let $\Omega^m \subset \Delta(A_1)$ be the set of actions player 1 could possibly commit to, which is assumed to be finite. I use $\alpha_1 \in \Omega^m$ to represent the commitment type that plays α_1 in every period (or commitment type α_1). Let $\phi_{\alpha_1} \in \Delta(\Theta)$ be the distribution of θ *conditional on* player 1 being commitment type α_1, with $\phi \equiv \{\phi_{\alpha_1}\}_{\alpha_1 \in \Omega^m}$.

Let $\Omega \equiv \Theta \cup \Omega^m$ be the set of types, with $\omega \in \Omega$ a typical element. Let $\mu \in \Delta(\Omega)$ be player 2's prior belief, which I assume has full support. The pair (μ, ϕ) induces a joint distribution over θ and player 1's characteristics (committed or strategic), which I call a distributional environment.

Note that the above formulation of commitment accommodates the one in which player 1 commits to play a state-contingent stationary strategy. To see this, let $\gamma : \Theta \to \Delta(A_1)$ be a state-contingent commitment plan, with Γ the finite set of commitment plans. Player 2 has a prior over Θ as well as the chances that player 1 is strategic or is committed to follow each plan in Γ. To convert this to my formulation, let

$$\Omega^m \equiv \{\alpha_1 \in \Delta(A_1) | \text{ there exist } \gamma \in \Gamma \text{ and } \theta \in \Theta \text{ such that } \gamma(\theta) = \alpha_1\},$$

which is the set of actions that are played under (at least) one commitment plan. The probability of every $\alpha_1 \in \Omega^m$ and its correlation with the state ϕ_{α_1} can be computed via player 2's prior. My formulation is more general, as it allows for arbitrary correlations between the state and the probability of being committed.

Histories & Payoffs: All past actions are perfectly monitored. Let $h^t = \{a_{1,s}, a_{2,s}, \xi_s\}_{s=0}^{t-1} \in \mathcal{H}^t$ be the public history in period t with $\mathcal{H} \equiv \cup_{t=0}^{+\infty} \mathcal{H}^t$. Let $\sigma_\omega : \mathcal{H} \to \Delta(A_1)$ be type ω's strategy, with the restriction that $\sigma_{\alpha_1}(h^t) = \alpha_1$ for every $(\alpha_1, h^t) \in \Omega^m \times \mathcal{H}$. Let $\sigma_1 \equiv (\sigma_\omega)_{\omega \in \Omega}$ be player 1's strategy. Let $\sigma_2 : \mathcal{H} \to \Delta(A_2)$ be player 2's strategy. Let $\sigma \equiv (\sigma_1, \sigma_2)$ be a typical strategy profile, and let Σ be the set of strategy profiles.

Player i's stage game payoff in period t is $u_i(\theta, a_{1,t}, a_{2,t})$, with $i \in \{1, 2\}$, which is naturally extended to the domain $\Delta(\Theta) \times \Delta(A_1) \times \Delta(A_2)$. Unlike Fudenberg and Levine (1989), my model has *interdependent values* as player 2's payoff depends on θ, which is player 1's private information. Strategic type θ maximizes $\sum_{t=0}^{\infty} (1-\delta)\delta^t u_1(\theta, a_{1,t}, a_{2,t})$. The player 2 who arrives in period t maximizes his expected stage game payoff.

Let $\text{BR}_2(\alpha_1, \pi) \subset A_2$ be the set of player 2's pure best replies when a_1 and θ are independently distributed with marginal distributions $\alpha_1 \in \Delta(A_1)$ and $\pi \in \Delta(\Theta)$, respectively. For every $(\alpha_1^*, \theta) \in \Omega^m \times \Theta$, let

$$v_\theta(\alpha_1^*) \equiv \min_{a_2^* \in \text{BR}_2(\alpha_1^*, \theta)} u_1(\theta, \alpha_1^*, a_2^*), \tag{1.1}$$ [9]

be type θ's (complete information) *commitment payoff* from playing α_1^*. If α_1^* is pure, then $v_\theta(\alpha_1^*)$ is a *pure commitment payoff*. Otherwise, α_1^* is mixed and $v_\theta(\alpha_1^*)$ is a *mixed commitment payoff*.

Solution Concept & Questions: The solution concept is Bayes Nash equilibrium (or *equilibrium* for short). The existence of equilibrium follows from Fudenberg and Levine (1983), as Θ, A_1 and A_2 are all finite sets and the game is continuous at infinity. Let $\text{NE}(\delta, \mu, \phi) \subset \Sigma$ be the set of equilibria under parameter configuration (δ, μ, ϕ). Let $V_\theta^\sigma(\delta)$ be type θ's discounted average payoff under strategy profile σ and discount factor δ. Let $\underline{V}_\theta(\delta, \mu, \phi) \equiv \inf_{\sigma \in \text{NE}(\delta,\mu,\phi)} V_\theta^\sigma(\delta)$ be type θ's worst equilibrium payoff.

I am interested in two sets of questions. First, can we find good lower bounds for a patient long-run player's guaranteed payoff, i.e. $\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu, \phi)$? In particular, can we extend Fudenberg and Levine (1989)'s insights that reputation can overcome the lack-of-commitment problem (when the reputation builder is patient) to interdependent value environments. Formally, for a given

[9] Abusing notation, I will use θ to denote the Dirac measure on θ. The same rule applies to degenerate distributions on A_1 and A_2.

$(\alpha_1^*, \theta) \in \Omega^m \times \Theta$, is it true that:

$$\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu, \phi) \geq v_\theta(\alpha_1^*)? \tag{1.2}$$

Furthermore, when is the above commitment payoff bound *fully robust*, that is, inequality (1.2) applies to every payoff function of the long-run player? Second, can we obtain robust predictions on player 1's equilibrium behavior? In particular, will he play the commitment strategy and maintain his reputation in every equilibrium?

My first set of questions examines player 1's guaranteed payoff when he can build a reputation. When u_2 does not depend on θ, inequality (1.2) is implied by the results in Fudenberg and Levine (1989, 1992) and player 1 can guarantee the payoff on the RHS by playing α_1^* in every period. In interdependent value environments, however, player 1 may receive a low payoff by playing α_1^* in every period, as convincing player 2 that α_1^* will be played does not determine her best reply. I address the robustness against equilibrium selection and against misspecifications of the long-run player's payoff function. Both are desirable properties of the results in Fudenberg and Levine (1989, 1992) as (1) reputation models are often applied to decentralized markets where there are no mediators helping participants to coordinate on a particular equilibrium, and (2) the modeler and the short-run players may entertain incorrect beliefs about the long-run player's payoff function.

My second set of questions advances the reputation literature one step further by examining the robust predictions on the long-run player's equilibrium behavior. Nevertheless, delivering robust behavioral predictions in this infinitely-repeated signalling game is challenging, as the conventional wisdom suggests that both infinitely-repeated games and signalling games have multiple equilibria with diverging behavioral predictions. Note that the commitment payoff bound does not imply that the long-run player will play his commitment strategy in every equilibrium, as a strategy that can secure him a high payoff is not necessarily his optimal strategy.

1.3 Fully Robust Commitment Payoff Bounds

I characterize the set of distributional environments under which the commitment payoff bound is fully robust. My conditions require that the likelihood ratios between some strategic types and the commitment type be below some cutoffs. My result evaluates the robustness of the commitment payoff bound in private value games against interdependent value perturbations. It also examines the validity of the commitment payoff bound in interdependent value environments without any restrictions on the long-run player's payoff function.

Saturation Set & Strong Saturation Set Throughout this section, I make the generic assumption that for every $(\alpha_1^*, \theta) \in \Omega^m \times \Theta$, $\mathrm{BR}_2(\alpha_1^*, \theta)$ is a singleton.[10] Let a_2^* be the unique

[10] $\mathrm{BR}_2(\alpha_1^*, \theta)$ being a singleton is satisfied under generic $u_2(\theta, a_1, a_2)$. This assumption will be relaxed in the Online Appendix, where I develop generalized fully robust commitment payoff bounds.

element in $\mathrm{BR}_2(\alpha_1^*,\theta)$. For every $(\alpha_1^*,\theta)\in\Omega^m\times\Theta$, let

$$\Theta^b_{(\alpha_1^*,\theta)}\equiv\big\{\tilde{\theta}\in\Theta\big|a_2^*\notin\mathrm{BR}_2(\alpha_1^*,\tilde{\theta})\big\}, \tag{1.3}$$

be the set of *bad states* (with respect to (α_1^*,θ)). Let $k(\alpha_1^*,\theta)\equiv\left|\Theta^b_{(\alpha_1^*,\theta)}\right|$ be its cardinality, with all private value models satisfying $k(\alpha_1^*,\theta)=0$. If $\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}$, then type $\tilde{\theta}$ is a 'bad' strategic type. For every $\tilde{\mu}\in\Delta(\Omega)$ with $\tilde{\mu}(\alpha_1^*)>0$, let $\tilde{\lambda}(\tilde{\theta})\equiv\tilde{\mu}(\tilde{\theta})/\tilde{\mu}(\alpha_1^*)$ be the likelihood ratio between type $\tilde{\theta}$ and commitment type α_1^*. Let $\tilde{\lambda}\equiv\Big(\tilde{\lambda}(\tilde{\theta})\Big)_{\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}}\in\mathbb{R}_+^{k(\alpha_1^*,\theta)}$ be the likelihood ratio vector. The best response set for $(\alpha_1^*,\theta)\in\Omega^m\times\Theta$ is defined as:[11]

$$\overline{\Lambda}(\alpha_1^*,\theta)\equiv\Big\{\tilde{\lambda}\in\mathbb{R}_+^{k(\alpha_1^*,\theta)}\Big|\{a_2^*\}=\arg\max_{a_2\in A_2}\big\{u_2(\phi_{\alpha_1^*},\alpha_1^*,a_2)+\sum_{\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}}\tilde{\lambda}(\tilde{\theta})u_2(\tilde{\theta},\alpha_1^*,a_2)\big\}\Big\}. \tag{1.4}$$

Intuitively, a likelihood ratio vector belongs to the best response set if a_2^* is player 2's strict best reply to α_1^* when she only counts the bad strategic types and the commitment type α_1^* in her calculations, while ignoring all the other strategic types and commitment types.

The *saturation set* for $(\alpha_1^*,\theta)\in\Omega^m\times\Theta$ is:

$$\Lambda(\alpha_1^*,\theta)\equiv\Big\{\tilde{\lambda}\Big|\lambda'\in\overline{\Lambda}(\alpha_1^*,\theta)\text{ for every }\mathbf{0}\ll\lambda'\ll\tilde{\lambda}\Big\}, \tag{1.5}$$

in which '$\ll$' denotes weak dominance in product order on $\mathbb{R}^{k(\alpha_1^*,\theta)}$ and $\mathbf{0}$ is the null vector in $\mathbb{R}^{k(\alpha_1^*,\theta)}$. Intuitively, $\tilde{\lambda}$ belongs to the saturation set if and only if every likelihood ratio vector equal or below $\tilde{\lambda}$ belongs to the best response set $\overline{\Lambda}(\alpha_1^*,\theta)$. By definition, $\Lambda(\alpha_1^*,\theta)\neq\{\varnothing\}$ if and only if $\mathbf{0}\in\Lambda(\alpha_1^*,\theta)$, or equivalently, $\mathrm{BR}_2(\alpha_1^*,\theta)=\mathrm{BR}_2(\alpha_1^*,\phi_{\alpha_1^*})=\{a_2^*\}$.

If $\Lambda(\alpha_1^*,\theta)\neq\{\varnothing\}$, then for every $\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}$, let $\psi(\tilde{\theta})$ be the largest $\psi\in\mathbb{R}_+$ such that:

$$a_2^*\in\arg\max_{a_2\in A_2}\Big\{u_2(\phi_{\alpha_1^*},\alpha_1^*,a_2)+\psi u_2(\tilde{\theta},\alpha_1^*,a_2)\Big\}.$$

By definition, $\psi(\tilde{\theta})$ is the intercept of $\Lambda(\alpha_1^*,\theta)$ on the $\lambda(\tilde{\theta})$-coordinate, which is strictly positive and finite. The *strong saturation set* for $(\alpha_1^*,\theta)\in\Omega^m\times\Theta$ is:

$$\underline{\Lambda}(\alpha_1^*,\theta)\equiv\begin{cases}\Big\{\tilde{\lambda}\Big|\sum_{\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}}\tilde{\lambda}(\tilde{\theta})/\psi(\tilde{\theta})<1\Big\} & \text{if }\Lambda(\alpha_1^*,\theta)\neq\{\varnothing\}\\ \{\varnothing\} & \text{if }\Lambda(\alpha_1^*,\theta)=\{\varnothing\}\end{cases} \tag{1.6}$$

Intuitively, the strong saturation set contains every non-negative vector that lies below the $k(\alpha_1^*,\theta)-1$ dimensional hyperplane that contains all the intersections between $\Lambda(\alpha_1^*,\theta)$ and the coordinates. In general, when $\mathrm{BR}_2(\alpha_1^*,\theta)$ may not be a singleton, $\underline{\Lambda}(\alpha_1^*,\theta)$ is defined as

$$\underline{\Lambda}(\alpha_1^*,\theta)\equiv\mathbb{R}_+^{k(\alpha_1^*,\theta)}\Big\backslash\mathrm{co}\Big(\mathbb{R}_+^{k(\alpha_1^*,\theta)}\Big\backslash\Lambda(\alpha_1^*,\theta)\Big),$$

[11] A relevant argument $\phi_{\alpha_1^*}$ is suppressed in the expression $\overline{\Lambda}(\alpha_1^*,\theta)$ to simplify notation.

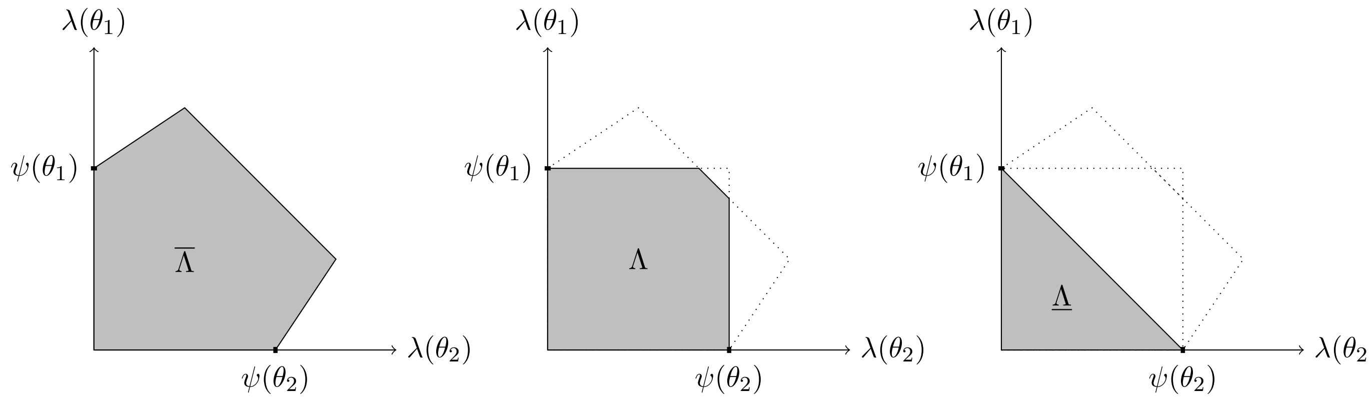

Figure 1-1: The best response set (left), the saturation set (middle) and the strong saturation set (right).

where co(·) denotes the convex hull. I show in the Online Appendix that $\widehat{\lambda} \in \underline{\Lambda}(\alpha_1^*, \theta)$ if and only if there exists $\phi \in (0, +\infty)^{k(\alpha_1^*,\theta)}$ such that:

$$\widehat{\lambda} \in \Big\{\tilde{\lambda}\Big| \sum_{\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}} \tilde{\lambda}(\tilde{\theta})/\phi(\tilde{\theta}) < 1 \text{ and } \tilde{\lambda}(\tilde{\theta}) \geq 0, \ \forall\, \tilde{\theta} \in \Theta^b_{(\alpha_1^*,\theta)}\Big\} \subset \Lambda(\alpha_1^*, \theta).$$

Figure 1-1 depicts the three sets in an example with two bad strategic types. I summarize some geometric properties of these sets for future reference. First, despite $\overline{\Lambda}(\alpha_1^*, \theta)$ can be unbounded, both $\Lambda(\alpha_1^*, \theta)$ and $\underline{\Lambda}(\alpha_1^*, \theta)$ are bounded sets. Furthermore, they are convex polyhedrons with characterizations independent of both player 1's payoff function and the probabilities of commitment types other than α_1^*. Second, as suggested by the notation, $\underline{\Lambda}(\alpha_1^*, \theta) \subset \Lambda(\alpha_1^*, \theta) \subset \overline{\Lambda}(\alpha_1^*, \theta)$. Third, if there is only one bad strategic type, i.e. $k(\alpha_1^*, \theta) = 1$ and $\Lambda(\alpha_1^*, \theta) \neq \{\varnothing\}$, then there exists a scalar $\psi^* \in (0, +\infty)$ such that:

$$\underline{\Lambda}(\alpha_1^*, \theta) = \Lambda(\alpha_1^*, \theta) = \overline{\Lambda}(\alpha_1^*, \theta) = \{\tilde{\lambda} \in \mathbb{R} | 0 \leq \tilde{\lambda} < \psi^*\}. \tag{1.7}$$

When $k(\alpha_1^*, \theta) \geq 2$, however, these three sets can be different, as I show in Figure 1-1.

1.3.1 Statement of Result

My first result characterizes the set of (μ, ϕ) under which the commitment payoff bound is fully robust i.e. it applies to every u_1. Let μ_t be player 2's belief in period t. Let λ and λ_t be the likelihood ratio vectors induced by μ and μ_t, respectively. For a set $X \subset \mathbb{R}^n$, recall that $\text{co}(X)$ is its convex hull and let $\text{cl}(X)$ be its closure.

Theorem 1.1. *For every $(\alpha_1^*, \theta) \in \Omega^m \times \Theta$ with α_1^* being pure,*

1. *If $\lambda \in \Lambda(\alpha_1^*, \theta)$, then $\liminf_{\delta\to 1} \underline{V}_\theta(\delta, \mu, \phi) \geq v_\theta(\alpha_1^*)$ for every u_1.*

2. *If* $\lambda \notin cl\Big(\Lambda(\alpha_1^*,\theta)\Big)$ *and* $BR_2(\alpha_1^*,\phi_{\alpha_1^*})$ *is a singleton, then there exists* u_1 *such that* $\limsup_{\delta\to 1}\underline{V}_\theta(\delta,\mu,\phi) < v_\theta(\alpha_1^*)$.

For every $(\alpha_1^*,\theta)\in\Omega^m\times\Theta$ *with* α_1^* *being mixed,*

3. *If* $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$*, then* $\liminf_{\delta\to 1}\underline{V}_\theta(\delta,\mu,\phi)\geq v_\theta(\alpha_1^*)$ *for every* u_1.

4. *If* $\lambda \notin cl\Big(\underline{\Lambda}(\alpha_1^*,\theta)\Big)$*,* $BR_2(\alpha_1^*,\phi_{\alpha_1^*})$ *is a singleton and* $\alpha_1^*\notin co\Big(\Omega^m\Big\backslash\{\alpha_1^*\}\Big)$*, then there exists* u_1 *such that* $\limsup_{\delta\to 1}\underline{V}_\theta(\delta,\mu,\phi) < v_\theta(\alpha_1^*)$.

According to Theorem 1.1, full robustness requires that the likelihood ratio between every bad strategic type and the relevant commitment type be below some cutoff, while it does not depend on the probabilities of the other strategic types and commitment types. Intuitively, this is because type θ needs to come up with a history-dependent action plan under which the likelihood ratio vector will remain low forever along every dimension. When the commitment payoff bound is fully robust, such action plans should exist regardless of player 2's belief about the other strategic types' strategies. This includes the adverse belief in which all the good strategic types separate from, while all the bad strategic types pool with, the commitment type.

However, unlike the private value benchmark, player 1 cannot guarantee his mixed commitment payoff by replicating the mixed commitment strategy. This is because playing some actions in the support of the mixed commitment strategy can *increase* some likelihood ratios, after which player 2's belief about the persistent state becomes pessimistic and player 1 cannot guarantee a high continuation payoff. Moreover, as $\underline{\Lambda}(\alpha_1^*,\theta)\subset\Lambda(\alpha_1^*,\theta)$, overcoming the lack-of-commitment problem and securing the commitment payoff requires more demanding conditions when the commitment strategy is mixed. This implies that small trembles by a pure commitment type can lead to a large decrease in player 1's guaranteed equilibrium payoff. This highlights another distinction between private and interdependent values, which I formalize in the Online Appendix.

This theorem has two interpretations. First, it evaluates the robustness of reputation effects in private value reputation games against a richer set of perturbations. Starting from Fudenberg and Levine (1989) in which θ is common knowledge and there is a positive chance of a commitment type, one can allow the short-run players to entertain the possibility that their opponent is another strategic type who may have private information about their preferences. My result implies that the fully-robust commitment payoff bound extends when these interdependent value perturbations are relatively less likely compared to the commitment type, and vice versa.

Second, it points out the limitations of reputation effects in repeated incomplete information games with interdependent values and unrestricted payoffs. According to this view, the modeler is perturbing a repeated game with interdependent values with commitment types. Therefore, every commitment type is *arbitrarily unlikely* compared to any strategic type. As a result, my conditions fail whenever $k(\alpha_1^*,\theta)>0$. This motivates the study of games with specific payoff structures in Section 1.4, which allows one to further explore the robust implications of reputation effects in interdependent value environments.

The proof of Theorem 1.1 appears in Appendices A.1 and A.2 as well as the Online Appendix. I make several remarks on the conditions before explaining the proof. First, Theorem 1.1 left out two degenerate sets of beliefs, which are the boundaries of $\Lambda(\alpha_1^*, \theta)$ and $\underline{\Lambda}(\alpha_1^*, \theta)$. In these knife-edge cases, the attainability of the commitment payoff bound depends on the presence of other mixed strategy commitment types and their correlations with the state. Second, the assumption that $\mathrm{BR}_2(\alpha_1^*, \phi_{\alpha_1^*})$ is a singleton in statements 2 and 4 is satisfied under generic parameter values, and is only required for the proof when $\Lambda(\alpha_1^*, \theta) = \{\varnothing\}$, which is used to rule out pathological cases where $a_2^* \in \mathrm{BR}_2(\alpha_1^*, \phi_{\alpha_1^*})$ but $\{a_2^*\} \neq \mathrm{BR}_2(\alpha_1^*, \phi_{\alpha_1^*})$. An example on this issue is presented in Appendix A.2. Third, according to the separating hyperplane theorem, the requirement that $\alpha_1^* \notin \mathrm{co}\big(\Omega^m \backslash \{\alpha_1^*\}\big)$ guarantees the existence of a payoff function $u_1(\theta, \cdot, \cdot)$ under which type θ's commitment payoff from any alternative commitment action in Ω^m is strictly below $v_\theta(\alpha_1^*)$. This convex independence assumption cannot be dispensed, as no restrictions are made on $\mu(\Omega^m \backslash \{\alpha_1^*\})$ and $\{\phi_{\alpha_1}\}_{\alpha_1 \neq \alpha_1^*}$. Therefore, commitment types other than α_1^* are allowed to occur with arbitrarily high probability and can have arbitrary correlations with the state.

1.3.2 Proof Ideas of Statements 1 & 3

I start with the case in which α_1^* is pure and then move on to those in which α_1^* is mixed.

Pure Commitment Payoff: Since α_1^* is pure, $\lambda_t(\tilde{\theta})$ will not increase if player 2 observes a_1^* for every $\tilde{\theta} \in \Theta^b_{(\alpha_1^*, \theta)}$. Therefore, $\lambda_t(\tilde{\theta}) \leq \lambda(\tilde{\theta})$ for every $t \in \mathbb{N}$ if player 1 imitates the commitment type. By definition, if $\lambda_t \in \Lambda(\alpha_1^*, \theta)$ and a_2^* is not a strict best reply (call period t a *bad period*), then the strategic types must be playing actions other than a_1^* in period t with probability bounded from below, after which they will be separated from the commitment type. As in Fudenberg and Levine (1989), the number of bad periods is uniformly bounded from above, which implies that player 1 can secure his commitment payoff as $\delta \to 1$.

Mixed Commitment Payoff when $\boldsymbol{k(\alpha_1^*, \theta) = 1}$: Let $\Theta^b_{(\alpha_1^*, \theta)} \equiv \{\tilde{\theta}\}$. Recall from equation (1.7) in subsection 3.1 that when $\underline{\Lambda}(\alpha_1^*, \theta) \neq \{\varnothing\}$, there exists $\psi^* > 0$ such that $\underline{\Lambda}(\alpha_1^*, \theta) = \{\tilde{\lambda} | 0 \leq \tilde{\lambda} < \psi^*\}$. The main difference from the pure commitment action case is that λ_t can *increase* after player 2 observes some actions in the support of α_1^*. As a result, type θ *cannot* secure his commitment payoff by replicating α_1^* since he may end up playing actions that are more likely to be played by type $\tilde{\theta}$, in which case λ_t will exceed ψ^*.

The key step in my proof shows that for every equilibrium strategy of the short-run players, one can construct a *non-stationary strategy* for the long-run player under which the following three goals are achieved simultaneously: (1) To avoid negative inferences about the state, i.e. $\lambda_t < \psi^*$ for every $t \in \mathbb{N}$. (2) In expectation, the short-run players believe that actions within a small neighborhood of α_1^* will be played for all but a bounded number of periods. (3) Every $a_1 \in A_1$ will be played with occupation measure close to $\alpha_1^*(a_1)$.[12]

[12]There is a remaining step after this to deal with correlations between action and state, with details shown in

To understand why one can make such a construction, note that $\{\lambda_t\}_{t\in\mathbb{N}}$ is a non-negative supermartingale conditional on α_1^*. Since $\lambda_0 < \psi^*$, the probability measure over histories (induced by α_1^*) in which λ_t never exceeds ψ^* is bounded from below by the Doob's Upcrossing Inequality.[13] When δ is close to 1, the Lindeberg-Feller Central Limit Theorem (Chung 1974) ensures that the set of player 1's action paths, in which the discounted time average frequency of every a_1 being close to $\alpha_1^*(a_1)$, occurs with probability close to 1 under the measure induced by α_1^*. Each of the previous steps defines a subset of histories, and the intersection between them occurs with probability bounded from below. Then I derive a uniform upper bound on the expected sum of relative entropy between α_1^* and player 2's predicted action conditional on only observing histories at the intersection. According to Gossner (2011), the unconditional expected sum is bounded from above by a positive number that does not explode as $\delta \to 1$. Given that the intersection between the two sets has probability bounded from below, the Markov Inequality implies that the conditional expected sum is also bounded from above. Therefore, the expected number of periods that player 2's predicted action is far away from α_1^* is bounded from above.

Mixed Commitment Payoff when $\boldsymbol{k(\alpha_1^*, \theta) \geq 2}$: Let $S_t \equiv \sum_{\tilde{\theta}\in\Theta^b_{(\alpha_1^*,\theta)}} \lambda_t(\tilde{\theta})/\psi(\tilde{\theta})$, which is a non-negative supermartingale conditional on α_1^*. The assumption that $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$ implies that $S_0 < 1$. Doob's Upcrossing Inequality provides a lower bound on the probability measure over histories under which S_t is always strictly below 1, i.e. $\lambda_t \in \underline{\Lambda}(\alpha_1^*, \theta)$ for every $t \in \mathbb{N}$. The proof then follows from the $k(\alpha_1^*, \theta) = 1$ case.

To illustrate why $\lambda \in \Lambda(\alpha_1^*, \theta)$ is insufficient when $k(\alpha_1^*, \theta) \geq 2$ and α_1^* is mixed, I present an example in Appendix A.7.8 where $\lambda \in \Lambda(\alpha_1^*, \theta)$ but type θ's equilibrium payoff is bounded below his commitment payoff. The idea is to construct equilibrium strategies for the bad strategic types, under which playing every action in the support of α_1^* will increase the likelihood ratio along some dimensions. As a result, player 2's belief in period 1 is bounded away from $\overline{\Lambda}(\alpha_1^*, \theta)$ regardless of the action played in period 0.

1.3.3 Proof Ideas of Statements 2 & 4

To prove statement 2, let α_1^* be the Dirac measure on $a_1^* \in A_1$. Let player 1's payoff be given by:

$$u_1(\widetilde{\theta}, a_1, a_2) = \mathbf{1}\{\widetilde{\theta} = \theta, a_1 = a_1^*, a_2 = a_2^*\}. \tag{1.8}$$

I construct an equilibrium in which type θ obtains a payoff strictly bounded below 1 even when δ is arbitrarily close to 1. The key idea is to let the bad strategic types pool with the commitment type (with high probability) and the good ones separate from the commitment type. As a result,

Appendix A.1.2 Part II.

[13] In private value reputation games with noisy monitoring, Fudenberg and Levine (1992) use the upcrossing inequality to bound the number of *bad periods* when player 1 imitates the commitment strategy. In contrast, I use the upcrossing inequality to show that player 1 can *cherry-pick* actions in the support of his mixed commitment strategy in order to prevent $\lambda(\tilde{\theta})$ from exceeding $\psi(\tilde{\theta})$ while simultaneously making his opponents believe that actions close to α_1^* will be played in all but a bounded number of periods.

type θ cannot simultaneously build a reputation for commitment while separating away from the bad strategic types.

However, the proof is complicated by the presence of other commitment types that are playing mixed strategies. To understand this issue, consider an example where $\Theta = \{\theta, \tilde{\theta}\}$ with $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$, $\Omega^m = \{a_1^*, \alpha_1\}$ with α_1 non-trivially mixed, attaching positive probability to a_1^* and $\{a_2^*\} = \text{BR}_2(a_1^*, \phi_{a_1^*}) = \text{BR}_2(\alpha_1, \phi_{\alpha_1})$. The naive construction in which type $\tilde{\theta}$ plays a_1^* all the time does not work, as type θ can then obtain a payoff arbitrarily close to 1 by playing $a_1 \in \text{supp}(\alpha_1)\backslash\{a_1^*\}$ in period 0 and a_1^* in every subsequent period.

To circumvent this problem, I construct a sequential equilibrium in which type θ's action is deterministic on the equilibrium path. Type $\tilde{\theta}$ plays a_1^* in every period with probability $p \in (0,1)$ and plays strategy $\sigma(\alpha_1)$ with probability $1-p$, with p being large enough that λ_1 is bounded away from $\overline{\Lambda}(\alpha_1^*, \theta)$ after observing a_1^* in period 0. The strategy $\sigma(\alpha_1)$ is described as follows: At histories that are consistent with type θ's equilibrium strategy, play α_1; at histories that are inconsistent, play a completely mixed action $\hat{\alpha}_1(\alpha_1)$ which attaches strictly higher probability to a_1^* than to any element in $\Omega^m\backslash\{a_1^*\}$.

To verify incentive compatibility, I keep track of the likelihood ratio between the fraction of type $\tilde{\theta}$ who plays $\sigma(\alpha_1)$ and the commitment type α_1. If type θ has never deviated before, then this ratio remains constant. If type θ has deviated before, then this ratio increases every time a_1^* is observed. Therefore, once type θ has deviated from his equilibrium play, he will face a trade-off between obtaining a high stage-game payoff (by playing a_1^*) and reducing the likelihood ratio. This uniformly bounds his continuation payoff after any deviation from above, which is strictly below 1. Type θ's on-path strategy is then constructed such that his payoff is strictly between 1 and his highest post-deviation continuation payoff. This can be achieved, for example, by using a public randomization device that prescribes a_1^* with probability less than 1 in every period.

The proof of statement 4 involves several additional steps, with details shown in the Online Appendix. First, the payoff function in equation (1.8) is replaced by one that is constructed via the separating hyperplane theorem, such that type θ's commitment payoff from every other action in Ω^m is strictly lower than his commitment payoff from playing α_1^*. Second, I show that there exists an integer T (independent of δ) and a T-period strategy for the strategic types other than θ such that the likelihood ratio vector in period T is bounded away from $\overline{\Lambda}(\alpha_1^*, \theta)$ regardless of player 1's behavior in the first T periods. Third, the continuation play after period T modifies the construction in the proof of statement 2. The key step is to construct the bad strategic types' strategies under which type θ's continuation payoff after any deviation is bounded below his commitment payoff from playing α_1^*.

1.4 Games with Monotone-Supermodular Payoffs

Motivated by the discussions in Section 1.3, I study stage games where players' payoffs satisfy a *monotone-supermodularity* (or MSM) condition. I explore the robust predictions on the long-run player's equilibrium payoff and on-path equilibrium behavior. All the results in this section apply even when the commitment types are arbitrarily unlikely compared to any strategic type.

1.4.1 Monotone-Supermodular Payoff Structure

Let Θ, A_1 and A_2 be finite ordered sets. I use '$\succ$', '$\succsim$', '$\prec$' and '$\precsim$' to denote the rankings between pairs of elements. The stage game has MSM payoffs if it satisfies the following pair of assumptions:

Assumption 1.1 (Monotonicity). *$u_1(\theta, a_1, a_2)$ is strictly decreasing in a_1 and is strictly increasing in a_2.*

Assumption 1.2 (Supermodularity). *$u_1(\theta, a_1, a_2)$ has strictly increasing differences (or SID) in (a_1, θ) and increasing differences (or ID) in (a_2, θ). $u_2(\theta, a_1, a_2)$ has SID in (a_1, a_2) and (θ, a_2).*[14]

I focus on games where player 2's decision-making problem is binary, which have been a primary focus of the reputation literature, for example, Kreps and Wilson (1982), Milgrom and Roberts (1982), Mailath and Samuelson (2001), Ekmekci (2011), Liu (2011) and others.[15]

Assumption 1.3. $|A_2| = 2$.

I will discuss how my assumptions fit into the applications to business transactions (or product choice game) and monopolistic competition (or entry deterrence game) in Subsection 1.4.5.

Preliminary Analysis: Let $\overline{a}_i \equiv \max A_i$ and $\underline{a}_i \equiv \min A_i$, with $i \in \{1, 2\}$. For every $\pi \in \Delta(\Theta)$ and $\alpha_1 \in \Delta(A_1)$, let

$$\mathcal{D}(\pi, \alpha_1) \equiv u_2(\pi, \alpha_1, \overline{a}_2) - u_2(\pi, \alpha_1, \underline{a}_2). \tag{1.9}$$

I classify the states into *good*, *bad* and *negative* by partitioning Θ into the following three sets:

$$\Theta_g \equiv \Big\{\theta \Big| \mathcal{D}(\theta, \overline{a}_1) \geq 0 \text{ and } u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2)\Big\},$$

$$\Theta_p \equiv \Big\{\theta \notin \Theta_g \Big| u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2)\Big\} \text{ and } \Theta_n \equiv \Big\{\theta \Big| u_1(\theta, \overline{a}_1, \overline{a}_2) \leq u_1(\theta, \underline{a}_1, \underline{a}_2)\Big\}.$$

Intuitively, Θ_g is the set of *good states* in which $\overline{a}_2$ is player 2's best reply to $\overline{a}_1$ and player 1 strictly prefers the commitment outcome $(\overline{a}_1, \overline{a}_2)$ to his minmax outcome $(\underline{a}_1, \underline{a}_2)$. Θ_p is the set of *bad states* in which player 2 has no incentive to play $\overline{a}_2$ but player 1 strictly prefers $(\overline{a}_1, \overline{a}_2)$ to his minmax outcome. Θ_n is the set of *negative states* in which player 1 prefers his minmax outcome to the commitment outcome. Lemma 1.4.1 shows that every good state is higher than every bad state, and every bad state is higher than every negative state:

Lemma 1.4.1. *If the stage game payoff satisfies Assumption 1.2, then:*

[14] First, given Assumption 1.2, the case in which $u_1(\theta, a_1, a_2)$ is strictly increasing in a_1 and strictly decreasing in a_2 can be analyzed similarly by reversing the orders of the states and each player's actions. Second, I only require u_1 to have ID in (a_2, θ) in order to accommodate the classic *separable case*, in which player 1's return from player 2's action does not depend on the state. Assumption 1.2 can be further relaxed, which can be seen in the conclusion (Assumption 1.4) and the Online Appendix.

[15] The results extend to games with $|A_2| \geq 3$ under extra conditions on u_1. The details can be found in the Online Appendix.

1. *For every* $\theta_g \in \Theta_g$, $\theta_p \in \Theta_p$ *and* $\theta_n \in \Theta_n$, $\theta_g \succ \theta_p$, $\theta_p \succ \theta_n$ *and* $\theta_g \succ \theta_n$.

2. *If* $\Theta_p, \Theta_n \neq \{\varnothing\}$, *then* $\mathcal{D}(\theta_n, \overline{a}_1) < 0$ *for every* $\theta_n \in \Theta_n$.

Proof of Lemma 1.4.1: For statement 1, since $\mathcal{D}(\theta_g, \overline{a}_1) \geq 0$ and $\mathcal{D}(\theta_p, \overline{a}_1) < 0$, SID of u_2 with respect to (θ, a_2) implies that $\theta_g \succ \theta_p$. Since $u_1(\theta_p, \overline{a}_1, \overline{a}_2) > u_1(\theta_p, \underline{a}_1, \underline{a}_2)$ and $u_1(\theta_n, \overline{a}_1, \overline{a}_2) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, we know that $\theta_p \succ \theta_n$ due to the SID of u_1 in (θ, a_1) and ID of u_1 in (θ, a_2). If $\Theta_p \neq \{\varnothing\}$, then statement 1 is proved. If $\Theta_p = \{\varnothing\}$, then since $u_1(\theta_g, \overline{a}_1, \overline{a}_2) > u_1(\theta_g, \underline{a}_1, \underline{a}_2)$ and $u_1(\theta_n, \overline{a}_1, \overline{a}_2) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, we have $\theta_g \succ \theta_n$. For statement 2, if $\Theta_p, \Theta_n \neq \{\varnothing\}$, then $\theta_n \prec \theta_p$. SID of u_2 with respect to (θ, a_2) implies that $\mathcal{D}(\theta_n, \overline{a}_1) < \mathcal{D}(\theta_p, \overline{a}_1) < 0$. □

1.4.2 Statement of Results

My results in this section outline the robust implications on player 1's payoff and behavior when he has the option to build a reputation for playing the highest action. For this purpose, I assume that $\overline{a}_1 \in \Omega^m$ and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, i.e. there exists a commitment type that plays the highest action in every period, and player 2 has a strict incentive to play $\overline{a}_2$ conditional on knowing that she is facing commitment type $\overline{a}_1$.

The qualitative features of equilibria depend on the relative likelihood between the strategic types who know that the state is good (call them *good strategic types*) and the ones who know that the state is bad (call them *bad strategic types*). In particular, player 2's prior belief is *optimistic* if:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0, \tag{1.10}$$

and is *pessimistic* if:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) \leq 0. \tag{1.11}$$

Notice that formulas (1.10) and (1.11) allow type $\overline{a}_1$ to be arbitrarily unlikely compared to every strategic type. In my results, these inequalities can be replaced by $\sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ and $\sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0$, respectively, when the total probability of commitment types is small enough. These expressions will be useful once we compare the reputation game to the benchmark game without commitment types.

Equilibrium Payoff under Optimistic Priors: The main result in the optimistic prior case is the commitment payoff bound for playing the highest action, which is stated as Theorem 1.2:

Theorem 1.2. *If* $\overline{a}_1 \in \Omega^m$, $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$ *and* μ *satisfies (1.10), then for every* $\theta \in \Theta$, *we have:*

$$\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu, \phi) \geq \max\{u_1(\theta, \overline{a}_1, \overline{a}_2), u_1(\theta, \underline{a}_1, \underline{a}_2)\}. \tag{1.12}$$

According to Theorem 1.2, a patient long-run player can overcome the lack-of-commitment problem and guarantee his payoff from $(\overline{a}_1, \overline{a}_2)$ in every state and in every equilibrium. It implies, for example, that a firm can secure high returns by maintaining a reputation for exerting high effort despite his customers' skepticism about product quality; an incumbent who might have unfavorable information about the market demand curve (say, demand elasticities are low) can guarantee high profits by fighting entrants.

The key distinction between condition (1.10) and the distribution conditions in Theorem 1.1 is that the good strategic types can contribute to attaining the commitment payoff. As a result, the commitment type is allowed to be arbitrarily unlikely compared to every bad strategic type. Intuitively, this is driven by the following implication of MSM payoff: when player 1 has an incentive to pool with the commitment type $\overline{a}_1$ in a lower state, he will then play $\overline{a}_1$ with probability 1 at every on-path history in a higher state. According to Lemma 1.4.1, every good state is higher than every bad state. The above property implies that in equilibria where some bad strategic types pool with the commitment type, all the good strategic types will behave like the commitment type on the equilibrium path, which can help to guarantee the commitment payoff.

Nevertheless, there also exist repeated game equilibria in which the commitment strategy is not optimal for any bad strategic type. This undermines the implications of MSM and as a result, the good strategic types will have a strict incentive not to play the commitment strategy in some of those equilibria. Therefore, reputation effects cannot provide accurate predictions on player 1's equilibrium behavior in the optimistic prior case. Moreover, there can exist on-path histories at which every bad strategic type plays $\overline{a}_1$ with strictly higher probability than every good strategic type, so player 2's belief can become more pessimistic after observing $\overline{a}_1$.

In order to establish the commitment payoff bound in those equilibria, I circumvent the aforementioned complications by showing that player 2's posterior cannot become too pessimistic conditional on $\overline{a}_1$ is always being played. This implies that in every period where $\overline{a}_2$ is not player 2's strict best reply, the strategic types must be separating from the commitment type with probability bounded from below. Hence, there can be at most a bounded number of such periods, which validates the commitment payoff bound in those equilibria.

One may also wonder whether player 1 can guarantee a strictly higher payoff by establishing a reputation for playing an alternative commitment action. In the Online Appendix, I adopt a notion of tightness introduced by Cripps, Schmidt and Thomas (1996) and show that when there are bad strategic types, i.e. $\Theta_p \neq \{\varnothing\}$, no type of player 1 can guarantee a strictly higher equilibrium payoff by establishing a reputation for playing another pure commitment action. Furthermore, if $\Theta_p \neq \{\varnothing\}$ and $\Theta_n = \{\varnothing\}$, then player 1 cannot guarantee a strictly higher equilibrium payoff by taking any other (pure or mixed) commitment actions.

Equilibrium Payoff and Behavior under Pessimistic Priors: When μ satisfies condition (1.11), there exists a unique pair of $(\theta_p^*, q(\mu)) \in \Theta_p \times (0,1]$ such that:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + q(\mu)\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \sum_{\theta \succ \theta_p^*} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) = 0. \tag{1.13}$$

Since $\theta_p^* \in \Theta_p$, the definition of Θ_p and Assumption 1.1 imply the existence of $r \in (0,1)$ such that:

$$ru_1(\theta_p^*, \overline{a}_1, \overline{a}_2) + \big(1-r\big)u_1(\theta_p^*, \overline{a}_1, \underline{a}_2) = u_1(\theta_p^*, \underline{a}_1, \underline{a}_2). \tag{1.14}$$

Let

$$v_\theta^* \equiv \begin{cases} u_1(\theta, \underline{a}_1, \underline{a}_2) & \text{if } \theta \precsim \theta_p^* \\ ru_1(\theta, \overline{a}_1, \overline{a}_2) + \big(1-r\big)u_1(\theta, \overline{a}_1, \underline{a}_2) & \text{if } \theta \succ \theta_p^*. \end{cases} \tag{1.15}$$

Let $\mathcal{H}^\sigma(\theta)$ ($\subset \mathcal{H}$) be the set of histories that occurs with strictly positive probability under strategy profile σ conditional on player 1 being strategic type θ. For the sake of exposition, I state the result when $\Omega^m = \{\overline{a}_1\}$, which will be generalized to the case with multiple commitment types in Theorem A.1 (Appendix A.4.2) under the extra requirement that the total probability of commitment types, $\mu(\Omega^m)$, is small enough:

Theorem 1.3. *If $\Omega^m = \{\overline{a}_1\}$ and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, then for every μ satisfying equation (1.11), there exists $\overline{\delta} \in (0,1)$, such that for every $\delta > \overline{\delta}$ and $\sigma \equiv \Big((\sigma_\omega)_{\omega \in \Omega}, \sigma_2\Big) \in NE(\delta, \mu, \phi)$,*

1. *For every $\theta \succ \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\overline{a}_1$ at h^t.*
   For every $\theta \prec \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\underline{a}_1$ at h^t.
   In period 0, type θ_p^* plays $\overline{a}_1$ with probability $q(\mu)$ and $\underline{a}_1$ with probability $1 - q(\mu)$.
   For every $h^t \in \mathcal{H}^\sigma(\theta_p^*)$ with $t \geq 1$:

   - *If h^t contains $\overline{a}_1$, then at h^t, type θ_p^* plays $\overline{a}_1$.*
   - *If h^t contains $\underline{a}_1$, then at h^t, type θ_p^* plays $\underline{a}_1$.*

2. *If $q(\mu) \neq 1$, then $V_\theta^\sigma(\delta) = v_\theta^*$ for every $\theta \in \Theta$.*

According to Theorem 1.3, every strategic type's equilibrium payoff and on-path behavior are (generically) the same across all Nash equilibria when player 1 is patient. Furthermore, his behavior and payoff are independent of the discount factor as long as it lies above some threshold $\overline{\delta}$. On the equilibrium path, every type strictly above θ_p^* plays $\overline{a}_1$ in every period, every type strictly below θ_p^* plays $\underline{a}_1$ in every period, type θ_p^* randomizes (in period 0) between playing $\overline{a}_1$ in every period and playing $\underline{a}_1$ in every period. This suggests that the long-run player will behave consistently over time and maintain his reputation for commitment in every equilibrium. As implied by equation (1.13), player 2 is indifferent between $\overline{a}_2$ and $\underline{a}_2$ starting from period 1 conditional on player 1's always having played $\overline{a}_1$ in the past. When the cutoff type θ_p^* plays a non-trivial mixed strategy, i.e. $q(\mu) \neq 1$,[16] his indifference condition in period 0 pins down every strategic type's equilibrium payoff. In particular, it requires that the occupation measure of $\overline{a}_2$ be r when $\overline{a}_1$ is played in every period and 0 when $\underline{a}_1$ is played in every period, as can be seen from equations (1.14) and (1.15).

Intuitively, the uniqueness of player 1's on-path behavior is driven by the following *disciplinary effect*: he can obtain a high payoff by playing $\overline{a}_1$ in every period thanks to the commitment type,

[16]The condition that $q(\mu) \neq 1$ is satisfied for generic μ. It is also satisfied if we fix the likelihood ratios between the strategic types and focus on cases where the total probability of commitment types is small.

but it is impossible for him to receive a high payoff in the continuation game if he has ever failed to do so. To be more precise, I begin with the useful observation that under a pessimistic prior, always playing $\overline{a}_1$ must be optimal for *some* bad strategic types. Since players' payoffs are MSM, the above statement implies that all the good strategic types will play $\overline{a}_1$ at every on-path history and in every equilibrium. Therefore, player 2's belief about θ deteriorates whenever player 1 fails to play $\overline{a}_1$, after which his continuation payoff will be low. To see why playing the highest action for some time and then switching to a lower action is suboptimal for every strategic type, notice that (1) if his first deviation happened at an optimistic belief, then he could guarantee a strictly higher payoff by playing the highest action in every period thanks to the commitment type; (2) if his first deviation occurred after period 0 when player 2's belief is pessimistic, then he could strictly save the cost of playing $\overline{a}_1$ by playing $\underline{a}_1$ from period 0. The probabilities with which the cutoff type θ_p^* mixes in period 0 can be uniquely pinned down due to the *substitutability* between his return from playing $\overline{a}_1$ and the equilibrium probability with which he plays $\overline{a}_1$. In particular, if type θ_p^* plays $\overline{a}_1$ with higher probability, then it reduces player 2's incentive to play $\overline{a}_2$ after observing $\overline{a}_1$ and hence reduces type θ_p^*'s return from playing $\overline{a}_1$.

In contrast, player 1 exhibits multiple on-path behaviors in Fudenberg and Levine (1989) and behaving inconsistently is strictly optimal in some sequential equilibria. This is because deviating from the commitment action only signals that player 1 is strategic, but cannot preclude him from obtaining a high payoff in the continuation game according to Fudenberg et al. (1990). As a result, he may have an incentive to separate from the commitment type in any given period, depending on which equilibrium players coordinate on.[17] Similarly in the optimistic prior case, deviating from the commitment action can still lead to an optimistic posterior about the state, after which player 1's continuation payoff can still be high, leading to multiple equilibrium behaviors.

For an overview of the extension to multiple commitment types (Theorem A.1 in Appendix A.4.2): If there are only pure strategy commitment types and type $\overline{a}_1$ is the only commitment type under which $\overline{a}_2$ is optimal, then all the conclusions in Theorem 1.3 apply without any further qualifications. If there are only pure commitment types, but there are commitment types other than $\overline{a}_1$ under which player 2 has a strict incentive to play $\overline{a}_2$, then as long as $\mu(\Omega^m)$ is small enough, player 1's equilibrium payoff and behavior are the same across all equilibria. Every strategic type's equilibrium behavior is the same as described in Theorem 1.3 except for the cutoff type θ_p^*, who can play actions other than $\overline{a}_1$ and $\underline{a}_1$ with positive probability. If there are mixed commitment types and $\mu(\Omega^m)$ is small enough, then there exists a cutoff type $\theta_p^* \in \Theta_p$ (which is the same across all equilibria) such that all types above θ_p^* play $\overline{a}_1$ all the time, all types below θ_p^* play $\underline{a}_1$ all the time. Type θ_p^*'s various on-path behaviors in different equilibria will coincide with (ex ante) probability

[17] The behavioral uniqueness conclusion will also fail in repeated incomplete information games where the state only affects player 1's payoff, regardless of how pessimistic the prior belief is. To see this, consider for example that player 1 has persistent private information about his discount factor (Ghosh and Ray 1996) or his cost of taking a higher action (Schmidt 1993b). In these cases, the strategic types who have low discount factors or high costs either have no incentive to pool with the commitment type, in which case the disciplinary effect only works temporarily; or if they play the commitment strategy in equilibrium, then they are equivalent to the commitment type in player 2's best-response problem, in which case they are no longer 'bad'.

of at least $1-\epsilon$, and moreover, he will either play $\overline{a}_1$ all the time or $\underline{a}_1$ all the time with probability of at least $1-\epsilon$, with ϵ vanishing as $\mu(\Omega^m) \to 0$.

I conclude this subsection by adding two caveats. To begin with, my behavior uniqueness result requires that the long-run player be patient. I show by counterexample in Appendix A.7.7 that he can have multiple possible equilibrium behaviors when δ is low. Intuitively, this is because an impatient bad strategic type has no incentive to pay the cost of imitating the commitment type. As a result, the disciplinary effect will disappear. Next, neither Theorem 3 nor its extension (Theorem A.1) can imply the uniqueness of Nash equilibrium or Nash equilibrium outcome. This is because first, Nash equilibrium places no restriction on players' behaviors off the equilibrium path. Second, player 2's behavior on the equilibrium path is not necessarily unique. To see this, assume for example, $\Omega^m = \{\overline{a}_1\}$. Since player 2 is indifferent starting from period 1 conditional on $\overline{a}_1$ always being played, her behavior is only restricted by two sets of constraints. The first is, type θ_p^*'s indifference condition in period 0. The second constraint is, type θ_p^*'s incentives to play $\overline{a}_1$ in period $t \in \mathbb{N}$. The first one only pins down the occupation measure of $\overline{a}_2$ conditional on $\overline{a}_1$ being played in every period, and the second one only requires that $\overline{a}_2$ not be too front-loaded. Under these constraints, there are still multiple ways to allocate the play of $\overline{a}_2$ over time, leading to multiple equilibrium outcomes.

1.4.3 Proof Ideas of Theorems 1.2 and 1.3

The proofs of Theorem 1.2, Theorem 1.3 and Theorem A.1 can be found in Appendices A.3 and A.4, and the counterexamples to my results in which each of my assumptions fails are in Appendices A.7.1, A.7.2 and A.7.3.

To recall the challenges ahead, first, since values are interdependent and the commitment types are allowed to be arbitrarily unlikely compared to every strategic type, Theorem 1.1 suggests that the proofs need to exploit the properties of player 1's payoff function. Therefore, the standard learning-based arguments in Fudenberg and Levine (1989, 1992), Sorin (1999), Gossner (2011) and others cannot be directly applied.

Second, a repeated supermodular game is *not* supermodular, as player 1's action today can affect future equilibrium play. Consequently, the monotone selection result on static supermodular games (see Topkis 1998) is not applicable. Similar issues have been highlighted in complete information extensive form games (Echenique 2004) and 1-shot signalling games (Liu and Pei 2017). For an illustrative example, consider the following 1-shot signalling game where the sender is the row player and the receiver is the column player:

$\theta = H$	l	r
U	$4,8$	$0,0$
D	$\mathbf{2,4}$	$0,0$

$\theta = L$	l	r
U	$-2,-2$	$\mathbf{2,0}$
D	$0,-4$	$5,1$

If we rank the states and players' actions according to $H \succ L$, $U \succ D$ and $l \succ r$, one can verify that both players' payoffs are strict supermodular functions of the triple (θ, a_1, a_2). However, there exists a sequential equilibrium in which the sender plays D in state H and U in state L. The receiver plays l after she observes D and r after she observes U. Therefore, the sender's equilibrium action can be *strictly decreasing* in the state, despite all the complementarities between players' actions and the state.

The game studied in this paper is trickier than 1-shot signalling games, as the sender (or player 1) is *repeatedly signalling* his private information. The presence of intertemporal incentives provides a rationale for many different behaviors and belief-updating processes that cannot be rationalized in 1-shot interactions. For example, even when the stage game has MSM payoffs, there can still exist equilibria in the repeated signalling game where at some on-path histories, player 1 plays $\overline{a}_1$ with higher probability in a lower state compared to a higher state. As a result, player 1's reputation could deteriorate even when he plays the highest action.

Proof Sketch in the Entry Deterrence Game: I illustrate the logic of the proof using the entry deterrence game in the introduction. Recall that players' stage game payoffs are given by:

$\theta = H$	O	E
F	$2, 0$	$0, -1$
A	$3, 0$	$1, 2$

$\theta = L$	O	E
F	$2-\eta, 0$	$-\eta, 1$
A	$3, 0$	$1, 2$

Let $H \succ L$, $F \succ A$ and $O \succ E$. One can check that Assumptions 1.1 and 1.3 are satisfied. I focus on the case where $\eta \in (0,1)$, which satisfies Assumption 1.2 and moreover, $L \in \Theta_p$. I make several simplifying assumptions which are relaxed in Appendices A.3 and A.4. First, player 2 can only observe player 1's past actions, i.e. $h^t = \{a_{1,s}\}_{s=0}^{t-1}$. Second, there is only one commitment type, i.e. $\Omega^m \equiv \{F\}$. Third, let ϕ_F be the Dirac measure on state H.

Two Classes of Equilibria: I classify the set of equilibria into two classes, depending on whether or not playing F in every period is type L's best reply. Formally, let h_F^t be the period t history at which all past actions were F. For any given equilibrium $\sigma \equiv (\{\sigma_\omega\}_{\omega \in \Omega}, \sigma_2)$, σ is called a *regular equilibrium* if playing F at every history in $\{h_F^t\}_{t=0}^{\infty}$ is type L's best reply to σ_2. Otherwise, σ is called an *irregular equilibrium.*

Regular Equilibria: I use a monotone selection result on 1-shot signalling games (Liu and Pei 2017):

- If a 1-shot signalling game has MSM payoffs and the receiver's action choice is binary, then the sender's action is non-decreasing in the state in every Nash equilibrium.

This result implies that in the repeated signalling game studied in this section, if playing the highest action in every period is player 1's best reply in a lower state, then he will play the highest action with probability 1 at every on-path history in a higher state (see Lemma A.3.1 for a formal statement). In the context of the entry deterrence game, if an equilibrium is regular, then playing F in every period is type L's best reply. Since $H \succ L$, the result implies that type H will play F with probability 1 at h_F^t for every $t \in \mathbb{N}$.

Irregular Equilibria: I establish two properties of irregular equilibria. First, at every history h_F^t where player 2's belief attaches higher probability to type H than to type L, either O is her strict best reply, or the strategic types will be separated from the commitment type at h_F^t with significant probability. Next, I show that when δ is large enough, player 2's posterior belief will attach higher probability to type H than to type L at every h_F^t. Let q_t be the ex ante probability that player 1 is type L and he has played F from period 0 to $t-1$, and let p_t be the ex ante probability that player 1 is type H and he has played F from period 0 to $t-1$.

Claim 1. *For every $t \in \mathbb{N}$, if $p_t \geq q_t$ but O is not a strict best reply at h_F^t, then:*

$$(p_t + q_t) - (p_{t+1} + q_{t+1}) \geq \mu(F)/2. \tag{1.16}$$

Proof of Claim 1: Player 2 does not have a strict incentive to play O at h_F^t if and only if: $\mu(F) + p_{t+1} - (p_t - p_{t+1}) - q_{t+1} - 2(q_t - q_{t+1}) \leq 0$, which implies that $\mu(F) + 2p_{t+1} + 2q_{t+1} \leq p_t + 2q_t + q_{t+1} \leq p_t + 3q_t \leq 2p_t + 2q_t$, where the last inequality makes use of the assumption that $p_t \geq q_t$. If we rearrange the terms, the result is inequality (1.16). □

Claim 2. *If δ is large enough, then in every irregular equilibrium, $p_t \geq q_t$ for all $t \geq 0$.*

Claim 2 establishes an important property of irregular equilibria, namely, despite the fact that playing the highest action could lead to negative inferences about the state, player 2's belief about the strategic types cannot become too pessimistic. Intuitively, this is because type L's continuation payoff must be low if he separates from the commitment type in the *last* period with a pessimistic belief, while he can guarantee himself a high payoff by continuing to play F. This contradicts his incentive to separate in that last period.

Proof of Claim 2: Suppose towards a contradiction that $p_t < q_t$ for some $t \in \mathbb{N}$. Given that playing F in every period is not type L's best reply, there exists $T \in \mathbb{N}$ such that type L has a strict incentive to play A at h_F^T.[18] That is to say, $p_s \geq q_s = 0$ for every $s > T$. Let $t^* \in \mathbb{N}$ be the *largest* integer t such that $p_t < q_t$. The definition of t^* implies that (1) player 2's belief at history $(h_F^{t^*}, A)$ attaches probability strictly more than $1/2$ to type L, (2) type L is supposed to play A with strictly positive probability at $h_F^{t^*}$.

Let us examine type L's incentives at $h_F^{t^*}$. If he plays A, then his continuation payoff at $(h_F^{t^*}, A)$ is 1. This is because player 2's belief is a martingale, so there exists an action path played with positive probability by type L such that at every history along this path, player 2 attaches probability strictly more than $1/2$ to state L, which implies that she has a strict incentive to play E, and type L's stage game payoff is at most 1.

If he plays F at $h_F^{t^*}$ and in all subsequent periods, then according to Claim 1, there exists at most $\overline{T} \equiv \lceil 2/\mu(F) \rceil$ periods in which O is not player 2's strict best reply. This is because by definition, $p_s \geq q_s$ for all $s > t^*$. Therefore, type L's guaranteed continuation payoff is close to

[18] This is no longer true when player 2 can condition her actions on her predecessors' actions and the realizations of public randomization devices, in which case it can only imply that type L has a strict incentive to play A at some on-path histories where he has always played F before. These complications will be discussed in Remark II and will be treated formally in Appendix A.3.

$2-\eta$ when δ is large. This is strictly larger than 1. Comparing his continuation payoffs by playing A versus playing F reveals a contradiction. $\square$

Optimistic Prior Belief: When the prior belief is optimistic, i.e. $\mu(F)+\mu(H) > \mu(L)$, I establish the commitment payoff theorem for the two classes of equilibria separately. For regular equilibria, since type H behaves in the same way as the commitment type F, one can directly apply statement 1 of Theorem 1.1 and obtain the commitment payoff bound for playing F. For irregular equilibria, Claims 1 and 2 imply that conditional on playing F in every period, there exist at most $\overline{T}$ periods in which O is not player 2's strict best reply. Therefore, type H can guarantee a payoff close to 2 and type L can guarantee payoff close to $2-\eta$.

Pessimistic Prior Belief: When the prior belief is pessimistic, i.e. $\mu(F)+\mu(H) \leq \mu(L)$, we know that $p_0=\mu(H)<\mu(L)=q_0$. According to Claim 2, there is no irregular equilibria. So every equilibrium is regular, and therefore, type H will play F with probability 1 at every h_F^t.

Next, I pin down the probability with which type L plays F at every h_F^t. I start by introducing a measure of optimism for player 2's belief at h_F^t by letting

$$X_t \equiv \mu(F)\mathcal{D}(H,F)+p_t\mathcal{D}(H,F)+q_t\mathcal{D}(L,F). \tag{1.17}$$

Note that $\{X_t\}_{t=0}^{\infty}$ is a non-decreasing sequence as $\mathcal{D}(H,F)>0$, $\mathcal{D}(L,F)<0$, p_t is constant and q_t is non-increasing. The pessimistic prior assumption translates into $X_0 \leq 0$. The key step is to show that:

Claim 3. *If δ is large enough, then $X_t=0$ for all $t \geq 1$.*[19]

Proof of Claim 3: Suppose towards a contradiction that $X_t<0$ for some $t \geq 1$, then let us examine type L's incentives at h_F^{t-1}. Since $X_t<0$, type L will play F with positive probability at h_F^{t-1}. If he plays F at h_F^{t-1}, then his continuation payoff at h_F^t is 1. If he plays A at h_F^{t-1}, then his continuation payoff at (h_F^{t-1},A) is 1, but he can receive a strictly higher stage game payoff in period $t-1$. This leads to a contradiction.

Suppose towards a contradiction that $X_t>0$ for some $t \geq 1$, then let t^* be the smallest t such that $X_t>0$. Since $X_s \leq 0$ for every $s<t^*$, we know that type L will play A with positive probability at $h_F^{t^*-1}$. In what follows, I examine type L's incentives at $h_F^{t^*-1}$. If he plays A, then his continuation payoff at $(h_F^{t^*-1},A)$ is 1. If he plays F forever, then I will show below that O is player 2's strict best reply at h_F^s for every $s \geq t^*$. Once this is shown, we know that type L's guaranteed continuation payoff at $h_F^{t^*}$ is $2-\eta$, which is strictly greater than 1 and leads to a contradiction.

I complete the proof by showing that O is player 2's strict best reply at h_F^s for every $s \geq t^*$. Suppose towards a contradiction that player 2 does not have a strict incentive to play O at h_F^s for

[19] When there are other commitment types playing mixed strategies, X_t is close to albeit not necessarily equal to 0. Nevertheless, the variation of X_t across different equilibria vanishes as the total probability of commitment types goes to 0. When there are no mixed commitment types under which player 2 has a strict incentive to play $\overline{a}_2$, the sequence $\{X_t\}_{t=0}^{\infty}$ is generically unique.

some $s \geq t^*$, then:

$$\mu(F)\mathcal{D}(H,F) + p_s\mathcal{D}(H,F) + q_{s+1}\mathcal{D}(L,F) + (q_s - q_{s+1})\mathcal{D}(L,A) \leq 0, \tag{1.18}$$

$$\Rightarrow \quad q_s - q_{s+1} \geq \frac{X_s}{\mathcal{D}(L,F) - \mathcal{D}(L,A)} \underbrace{\geq}_{\text{since } X_s \geq X_{t^*}} \underbrace{\frac{X_{t^*}}{\mathcal{D}(L,F) - \mathcal{D}(L,A)}}_{>0} \equiv Y. \tag{1.19}$$

Hence, there exist at most $\lceil q_0/Y \rceil$ such periods, which is a finite number. Let period $\bar{t}$ be the *last* of such periods. Let us examine type L's incentive at $h_F^{\bar{t}}$. On one hand, he plays A with positive probability at this history in equilibrium, which results in a continuation payoff close to 1. On the other hand, his continuation payoff from playing F in every period is $2 - \eta$, which results in a contradiction. □

Remark I: When $L \in \Theta_n$, i.e. $\eta \geq 1$, Claim 1 as well as the conclusion on regular equilibria will remain intact. What needs to be modified is Claim 2: despite the fact that p_t can be less than q_t for some $t \in \mathbb{N}$ in some equilibria (think about for example, when the prior attaches very high probability to state L such that $p_0 < q_0$), type H can still guarantee a payoff close to 2 in every equilibrium.

To see this, in every irregular equilibrium where $p_t < q_t$ for some t, let t^* be the largest of such t and let us examine type L's incentives in period 0. For this to be an equilibrium, he must prefer playing F from period 0 to $t^* - 1$ and then A in period t^*, compared to playing A forever starting from period 0. By adopting the first strategy, his continuation payoff is 1 after period $t^* + 1$, his stage game payoff from period 0 to $t^* - 1$ is no more than 1 if O is played, and is no more than $-\eta$ if E is played. By adopting the second strategy, he can guarantee himself a payoff of at least 1. For the first strategy to be better than the second, the occupation measure with which E is played from period 0 to $t^* - 1$ needs to be arbitrarily close to 0 as $\delta \to 1$. That is to say, if type H plays F in every period, then the discounted average payoff he loses from period 0 to $t^* - 1$ (relative to 2 in each period) vanishes as $\delta \to 1$. According to Claim 1, his guaranteed continuation payoff after period t^* is close to 2. Summing up, his guaranteed payoff in period 0 is at least 2 in the $\delta \to 1$ limit.

Remark II: In Appendices A.3 and A.4, I extend the above idea and provide full proofs to Theorems 1.2 and 1.3, which incur two additional complications. First, there can be arbitrarily many strategic types, and in particular, good, bad and negative types could co-exist. Second, player 2's actions can be conditioned on the past realizations of public randomization devices as well as on her predecessors' actions, which could open up new equilibrium possibilities and therefore, can potentially undermine the robust predictions on payoff and behavior.

In terms of the proof, the main difference occurs in the analysis of irregular equilibria, as there may not exist a *last history* at which the probability of the bad strategic types is greater than the probability of the good strategic types. This is because the predecessor-successor relationship is

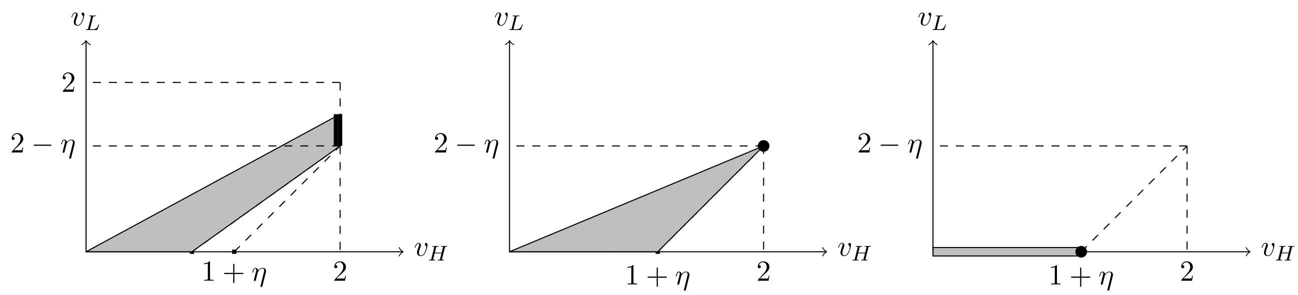

Figure 1-2: Limiting equilibrium payoff set in the interdependent value entry deterrence game without commitment types (in gray) and the selected payoffs under reputation effects (in black) when $\eta \in (0, 1)$. Left panel: $\hat{\mu}(H) > 1/2$. Middle panel: $\hat{\mu}(H) = 1/2$. Right panel: $\hat{\mu}(H) < 1/2$.

incomplete on the set of histories where player 1 has always played $\overline{a}_1$ once $\{a_{2,s}, \xi_s\}_{s \leq t-1}$ is also included in h^t.

My proof overcomes this difficulty by showing that every time a switch from a pessimistic to an optimistic belief happens, the bad strategic types must be separating from the commitment type with ex ante probability bounded from below. This implies that such switches can only happen finitely many times conditional on every positive probability event. On the other hand, the bad strategic types only have incentives to separate at those switching histories when their continuation payoffs from imitating the commitment type are low, which implies that another such switch needs to happen again in the future. This implies that such switches must happen infinitely many times if it happens at least once, leading to a contradiction.

1.4.4 Implications for Equilibrium Refinement

In this subsection, I revisit a classic application of reputation results by studying how they refine equilibrium payoffs and behaviors in repeated incomplete information games with long-run and short-run players. To do this, I study a benchmark repeated Bayesian game without commitment types, that is, player 1 is strategic with probability 1. I compare player 1's equilibrium payoff and behavior in the benchmark game and in the reputation game. This comparison addresses the question of which repeated game equilibria are more plausible when the long-run player can build reputations.

Let Θ be the set of states, which is also the set of player 1's types in the benchmark game. Let $\widehat{\mu} \in \Delta(\Theta)$ be player 2's prior belief. Let $\mathcal{V}(\delta, \widehat{\mu}) \subset \mathbb{R}^{|\Theta|}$ be player 1's equilibrium payoff set with $v \equiv (v_\theta)_{\theta \in \Theta}$ a typical element. I start with the optimistic prior case in which:

$$\sum_{\theta \in \Theta_g \cup \Theta_p} \widehat{\mu}(\theta) \mathcal{D}(\theta, \overline{a}_1) \geq 0. \tag{1.20}$$

To see the relationship between equations (1.20) and (1.10), notice that in the original reputation game where $\mu(\Omega^m)$ is small, (1.10) implies (1.20). Moreover, when we perturb a benchmark game that satisfies (1.20), the reputation game will satisfy (1.10) given that $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$. Let

$$\overline{\delta} \equiv \max_{\alpha_2 \in \Delta(A_2)} \Big\{ \frac{u_1(\overline{\theta}, \underline{a}_1, \alpha_2) - u_1(\overline{\theta}, \overline{a}_1, \alpha_2)}{u_1(\overline{\theta}, \underline{a}_1, \alpha_2) - u_1(\overline{\theta}, \overline{a}_1, \alpha_2) + u_1(\overline{\theta}, \overline{a}_1, \overline{a}_2) - u_1(\overline{\theta}, \underline{a}_1, \underline{a}_2)} \Big\},$$

the result is stated as Proposition 1.1:

Proposition 1.1. *If* $\overline{a}_1$ *is player* 1*'s pure Stackelberg action in state* $\overline{\theta}$*, then for every* $v \in \mathcal{V}(\delta, \widehat{\mu})$*, we have* $v_{\overline{\theta}} \leq u_1(\overline{\theta}, \overline{a}_1, \overline{a}_2)$*. Furthermore, if* $\widehat{\mu}$ *satisfies (1.20) and* $\delta \geq \overline{\delta}$*, then*

$$\sup_{v \in \mathcal{V}(\delta, \widehat{\mu})} v_{\overline{\theta}} \in \Big[(1-\delta) u_1(\overline{\theta}, \overline{a}_1, \underline{a}_2) + \delta u_1(\overline{\theta}, \overline{a}_1, \overline{a}_2), \quad u_1(\overline{\theta}, \overline{a}_1, \overline{a}_2)\Big].$$

The proof can be found in Appendix A.5. To summarize the role of reputation in refining equilibrium payoffs, first, it rules out equilibria with bad payoffs (for example, those with payoff $u_1(\theta, \underline{a}_1, \underline{a}_2)$) and selects equilibria that deliver every strategic type θ a payoff no less than his highest equilibrium payoff in a complete information repeated game where θ is common knowledge (Fudenberg, Kreps and Maskin 1990).

Second, according to Proposition 1.1, reputation effects select the highest equilibrium payoff for type $\overline{\theta}$ in the benchmark incomplete information game. However, types lower than $\overline{\theta}$ can obtain payoff strictly higher than $u_1(\theta, \overline{a}_1, \overline{a}_2)$ in the benchmark game even in the $\delta \to 1$ limit. Figure 2 depicts player 1's limiting equilibrium payoff set in the entry deterrence game, with more details coming in the Online Appendix.

Next, I analyze the pessimistic prior case in which

$$\sum_{\theta \in \Theta_g \cup \Theta_p} \widehat{\mu}(\theta) \mathcal{D}(\theta, \overline{a}_1) < 0. \tag{1.21}$$

The above inequality translates equation (1.11) into the benchmark game without commitment types, given that $\mu(\Omega^m)$ is small enough. Recall the definition of v_θ^* in expression (1.15), I state the result as Proposition 1.2:

Proposition 1.2. *For every* $\widehat{\mu}$ *satisfying (1.21), there exists* $\widehat{\delta} \in (0,1)$ *such that for every* $\delta > \widehat{\delta}$ *and* $\theta \in \Theta$*, we have:*

$$\sup_{v \in \mathcal{V}(\delta, \widehat{\mu})} v_\theta = v_\theta^*. \tag{1.22}$$

The proof is detailed in Appendix A.6. Proposition 1.2 implies that reputation effects select the highest equilibrium payoff for *every* strategic type in the benchmark incomplete information game. Moreover, since the unique equilibrium play for player 1 in Theorem 1.3 constitutes an equilibrium in the benchmark game as well, reputation effects also lead to the selection of a unique on-path behavior for the long-run player.

Unlike Proposition 1.1, Proposition 1.2 does not require $\overline{a}_1$ to be player 1's pure Stackelberg action in state $\overline{\theta}$. This is because under a pessimistic prior belief, playing actions other than $\overline{a}_1$ cannot induce player 2 to play $\overline{a}_2$, while under an optimistic prior belief, such possibilities cannot be ruled out unless we assume that $\overline{a}_1$ is type $\overline{\theta}$'s pure Stackelberg action. Moreover, the equilibrium selection result applies to all strategic types under a pessimistic belief but only applies to the highest type under an optimistic belief, as the pessimistic prior belief condition implies tight upper bounds on every bad strategic type's equilibrium payoff.

1.4.5 Related Applications

I discuss two applications of reputation models as well as how they fit into my MSM assumptions: the product choice game which highlights the lack-of-commitment problem in business transactions (Mailath and Samuelson 2001, Liu 2011, Ekmekci 2011) and the entry deterrence game that studies predatory pricing behaviors in monopolistic competition (Kreps and Wilson 1982, Milgrom and Roberts 1982).

Limit Pricing & Predation with Unknown Price Elasticities: Player 1 is an incumbent choosing between a *low price* (interpreted as limit pricing or predation) and a *normal price*, every player 2 is an entrant choosing between *out* and *enter*. The incumbent has private information about the demand elasticities $\theta \in \mathbb{R}_+$, which measures the increase in his product's demand when he lowers the price. The payoff matrix is given by:

State is θ	Out	Enter
Low Price	$p_L(Q_M+\theta), 0$	$p_L(Q_D+\gamma\theta), \Pi_L(\theta)-f$
Normal Price	$p_N Q_M, 0$	$p_N Q_D, \Pi_N - f$

where p_L and p_N are the low and normal prices, f is the sunk cost of entry, Q_M and Q_D are the incumbent's monopoly and duopoly demands under a normal price, Π_L and Π_N are the entrant's profits when the incumbent's price is low and normal, $\gamma \in (0,1)$ is a parameter measuring the effect of price elasticity on the incumbent's demand in duopoly markets relative to monopoly markets. This parameter is less than 1 as the entrant captures part of the market, which offsets some of the demand increase (of the incumbent's product) from a price cut.

In this example, Assumptions 1.1 and 1.2 require that (1) setting a low price is costly for the incumbent and he strictly prefers the entrant to stay out; (2) the entrant's profit from entering the market is lower when the incumbent sets a low price and when the demand elasticity is higher; (3) it is less costly for the incumbent to set a low price when the demand elasticity is higher. The first and third requirements are natural. The second one is reasonable, since lowering prices leaves the entrant a smaller market share, and this effect is more pronounced when the demand elasticity is higher.

Among other entry deterrence games, my assumptions also apply when the entrant faces uncertainty about the market size or the elasticity of substitution between her product and the incumbent's. It is also valid when the incumbent uses non-pricing strategies to deter entry, such as choosing the intensity of advertising in the pharmaceutical industry where advertising has positive spillovers to the entrant's product (Ellison and Ellison 2011). However, my supermodularity assumption fails in the entry deterrence model of Harrington (1986), in which the incumbent's and the entrant's production costs are positively correlated and the entrant does not know her own production costs before entering the market.

Product Choice Games: Consider an example of a software firm (player 1) and a sequence of clients (player 2). Every client chooses between the custom software (C) and the standardized software (S). In response to his client's request, the firm either exerts high effort (H) which can ensure a timely delivery and reduce the cost overruns, or exerts low effort (L). A client's willingness to pay depends not only on the delivery time and the expected cost overruns, but also on the quality of the software, which can be either good (G) or bad (B), and is the firm's private information. Here, quality is interpreted as the hidden running risks, the software's adaptability to future generations of operation systems, etc. Therefore, compared to delivery time and cost overruns, quality is much harder to observe directly, so it is reasonable to assume that future clients learn about quality mainly through the firm's past behaviors. This is modeled as the following product choice game:

$\theta =$ Good	Custom	Standardized
High Effort	$1, 3$	$-1, 2$
Low Effort	$2, 0$	$0, 1$

$\theta =$ Bad	Custom	Standardized
High Effort	$1 - \eta, 0$	$-1 - \eta, 1$
Low Effort	$2, -2$	$0, 0$

MSM requires that (1) exerting high effort is costly for the firm but it can result in more profit when the client purchases the custom software; (2) clients are more inclined to buy the custom software if it can be delivered on-time and its quality is high; (3) firms that produce higher quality software face lower effort costs. The first and second requirements are natural. The third one is reasonable since both the cost of making timely deliveries and the software's quality are positively correlated with the talent of the firm's employees. Indeed, Banerjee and Duflo (2000) provide empirical evidence in the Indian software industry, showing that firms enhance their reputations for competence via making timely deliveries and reducing cost overruns.

1.5 Concluding Remarks

A central theme of my analysis is that reputation building can be challenging when the uninformed players' learning is confounded. Even though the informed player can convince his opponents about his future actions, he may still fail to teach them how to best reply since their payoffs depend on the state. Similar in spirit is a contemporary work of Deb and Ishii (2017), which revisits the commitment payoff theorem when the uninformed players face uncertainty about the monitoring structure.[20] Their paper is complementary to mine, with the main difference being: the state can be identified through some exogenous public signals in their model (see Assumption 2.3 in their paper), while it can only be learned through the informed player's actions in my model. Under

[20]Related ideas also appear in Wolitzky (2011), who studies reputational bargaining with non-stationary commitment types and shows the failure of the commitment payoff theorem. However, his negative result requires that the uninformed player being long-lived and the commitment types playing non-stationary strategies, none of which are needed for the counterexamples (see Appendix A.7) and negative results (See statements 2 and 4 of Theorem 1.1) in my paper.

their identification assumption, they show that the strategic long-run player cannot guarantee his Stackelberg payoff when there are only stationary commitment types. They also construct a countably-infinite set of non-stationary commitment types, under which the informed player can guarantee his Stackelberg payoff. Their informational assumption fits in environments where informative signals about the state arrive frequently (or, in every period), as for example, when the state is the performance of vehicles, mobile phones, etc.

In contrast, my informational assumption fits into applications where exogenous signals are unlikely to arrive for a long time, or the variations of their realizations are mostly driven by noise orthogonal to the state. For example, when the state is the adaptability of a software to future generations of operating systems, the resilience of an architectural design to earthquakes, the long-run health impact of a certain type of food, the demand elasticity in markets with high sunk costs, the effectiveness of advertising in the NBA Finals, the amount of connection traffic in the hub of a major airline, and the like. My assumption has been adopted in many repeated Bayesian game models with interdependent values, such as Hart (1985), Aumann and Maschler (1995), Hörner and Lovo (2009), Kaya (2009), Hörner et al. (2011), Roddie (2012), Pęski and Toikka (2017).

My work is also related to the papers on bad reputation, for example, Ely and Välimäki (2003), Ely, Fudenberg and Levine (2008). These papers study a class of private value reputation games with imperfect monitoring known as *participation games.* They show that a patient long-run player's equilibrium payoff is low when the *bad commitment types* (i.e. ones that commit to play actions which discourage the short-run players from participating) are relatively more likely compared to the Stackelberg commitment type.

Although both my Theorem 1.1 and their results underscore the possibilities of reputation failure, the economic forces behind them are very different. In their models, reputations are bad due to the tension between the long-run player's forward-looking incentives and the short-run players' participation incentives. In particular, a patient long-run player has a strong incentive to take actions that can generate good signals but harm the participating short-run players. This discourages the short-run players from participating, which prevents the long-run player from signalling and leads to a bad reputation. In contrast, reputation failure occurs in my model as the short-run players' learning is confounded. This is because the long-run player's actions signal the payoff-relevant state. In different equilibria, these signals are interpreted in different ways, which will affect the short-run players' best reply to the commitment action. If the bad strategic types are believed to be pooling with the commitment type with high probability, then the strategic long-run player cannot simultaneously build a reputation for commitment while separating from the bad strategic types.

Extensions: I conclude by discussing several extensions of my results. First, players move sequentially in the stage game rather than simultaneously in some applications, such as a firm that chooses its service standards after consumers decide which product to purchase; an incumbent sets prices before or after observing the entrant's entering decision. My results are robust when the long-run player moves first. When the short-run players move first, my results are valid when every commitment type's strategy is independent of the short-run players' actions. This requirement is not redundant, as the short-run players cannot learn the long-run player's reaction following an unchosen a_2.

Along this line, my analysis can be applied to the following repeated bargaining problem, which models conflict resolution between employers and employees, firms and clients and other contexts. In every period, a long-run player bargains with a short-run player. The short-run player makes a take-it-or-leave-it offer, which is either soft or tough, and the long-run player either accepts the offer or chooses to resolve the dispute via arbitration. The long-run player has persistent private information about both parties' payoffs from arbitration, which can be interpreted as the quality of his supporting evidence. The short-run players observe the long-run player's bargaining postures in the previous periods and update their beliefs about their payoffs from arbitration.[21] In this context, my results provide accurate predictions on the long-run player's payoff and characterize his unique Nash equilibrium behavior when his (ex ante) expected payoff from arbitration is below a cutoff.

In some other applications where the uninformed players move first, the informed player cannot take actions at certain information sets. For example, the firm cannot exert effort when its client refuses to purchase, the incumbent cannot fight if the entrant stays out. My results in Section 1.4 apply to these scenarios as long as the informed long-run player can make an action choice in period t if $a_{2,t} \neq \overline{a}_2$. This condition allows for entry deterrence games but rules out participation games defined in Ely, Fudenberg and Levine (2008).

Second, my results are robust against the presence of *non-stationary* commitment types given that (1) all the non-stationary commitment types are pure, (2) different commitment strategies behave differently on the equilibrium path. When there are non-stationary commitment types playing mixed strategies, the attainability of the commitment payoff bound also depends on the probabilities of these non-stationary commitment types and their correlations with the state. To see this, consider the entry deterrence game. If there is a commitment type who plays F in every period and another one who plays strategy $\widehat{\sigma}_1$, which is defined as:

$$\widehat{\sigma}_1(h^t) \equiv \begin{cases} \frac{1}{2}F + \frac{1}{2}A & \text{if } t = 0 \\ F & \text{otherwise.} \end{cases}$$

Conditional on commitment type $\widehat{\sigma}_1$, state L occurs with certainty. If the probability of type $\widehat{\sigma}_1$ is three times larger than that of type F, then the conclusions in Theorems 1.2 and 1.3 will fail. This is because player 2 has no incentive to play O even conditional on the event that F will be played in every future period and player 1 is committed.

Third, Assumption 1.2 can be replaced by the following weaker condition as players' incentives remain unchanged under affine transformations on player 1's state contingent payoffs.

[21] Lee and Liu (2013) study a similar game without commitment types, but the short-run players observe their realized payoffs in addition to the long-run player's past actions. Their model applies to litigation, where the court's decisions are publicly available. My model applies to arbitration, as arbitration hearings are usually confidential and the final decisions are not publicly accessible.

Assumption 1.4. *There exists $f : \Theta \to (0, +\infty)$ such that $\widetilde{u}_1(\theta, a_1, a_2) \equiv f(\theta)u_1(\theta, a_1, a_2)$ has SID in (a_1, θ) and ID in (a_2, θ). u_2 has SID in (θ, a_2) and (a_1, a_2).*

To see how this generalization expands the applicability of Theorems 1.2 and 1.3, consider for example a repeated prisoner's dilemma game between a patient long-run player (player 1) and a sequence of short-run players (player 2s) in which players are reciprocal altruistic. As in Levine (1998), every player maximizes a weighted average of his monetary payoff and his opponent's monetary payoff, with the weight on his opponent be a strictly increasing function of his belief about his opponent's level of altruism. This can be applied to a number of situations in development economics, for example, a foreign firm, NGO or missionary (player 1) trying to cooperate with different local villagers (player 2s) in different periods. When player 1's level of altruism is his private information, this game violates Assumption 1.2 as his cost from playing a higher action (cooperate) and his benefit from player 2's higher action (cooperate) are both decreasing with his level of altruism. I show in the Online Appendix that the game satisfies Assumption 1.4 under an open set of parameters. I also provide a full characterization of Assumption 1.4 based on the primitives of the model.

Chapter 2

Reputation with Strategic Information Disclosure

This chapter studies the dynamics of an agent's reputation for competence when the labor market's information about his performance is disclosed by an intermediary who cannot commit. I show that this game admits a unique Markov Perfect Equilibrium (MPE). When the agent is patient, his effort is inverse U-shaped, while the rate of information disclosure is decreasing over time. I illustrate the inefficiencies of the unique MPE by comparing it with the equilibrium in the benchmark scenario where the market automatically observes all breakthroughs. I characterize a tractable subclass of non-Markov Equilibria and explain why allowing players to coordinate on payoff-irrelevant events can improve efficiency on top of the unique MPE and the exogenous information benchmark. When the intermediary can commit, her optimal Markov disclosure policy has a deadline, after which no breakthrough will be disclosed. However, deadlines are not incentive compatible in the game without commitment, illustrating a time inconsistency problem faced by the intermediary. My model can be applied to professional service industries, such as law and consulting. My results provide an explanation to the observed wage and promotion patterns in Baker, Gibbs and Holmström (1994).

I am indebted to Daron Acemoglu, Drew Fudenberg and Juuso Toikka for guidance and support. For useful comments, I thank Abhijit Banerjee, Sydnee Caldwell, Gonzalo Cisternas, Mira Frick, Robert Gibbons, Sinem Hidir, Bengt Holmström, Johannes Hörner, Irina Kholodenko, Yucheng Liang, Jean Tirole, Muhamet Yildiz, seminar participants at Harvard, MIT, Stanford SITE, Toulouse, Warwick, and in particular, Alessandro Bonatti, Daria Khromenkova and Alexander Wolitzky. All errors are mine.

2.1 Introduction

Reputation concern is an important driver of incentives in many professions, ranging from consultants, lawyers, judges and fund managers, to scientists, scholars and professional athletes. In particular, the incentive to *establish a name* is more substantial in the early stages of one's career: junior people work hard at entry-level jobs, hoping their professions will recognize their talents. However, the labor market rarely receives information about the performance of these lower-level people. Moreover, this information is mostly revealed by their current employers or direct supervisors.[1] This raises concerns that the latter might have incentives to manipulate the market's expectations by releasing information *strategically*.

Motivated by these phenomena, I analyze a reputation building model with the innovation that the market's information about an agent's performance is released strategically by an informational '*intermediary*', who also has a private interest in the game. Different from other dynamic information disclosure models, the intermediary *cannot commit* to disclosure policies.[2] This generates interesting dynamic interactions between the intermediary's incentive to release information and the agent's incentive to build up his reputation.

Related circumstances abound, especially in professional service industries, where team work among junior people makes it hard for the labor market to infer each individual's contribution. As a result, employers (or supervisors) enjoy substantive discretion in revealing information about their subordinates. Consider the example of a law firm, where junior associates work hard on cases but rarely have the opportunity to present their results in court. Their talents are recognized only after they have been given such chances or being promoted. In European soccer clubs (or in other professional sports), youth team players want to impress first team managers in order to gain opportunities to play in higher level matches (such as the Premier League). But the person who monitors their training and knows their abilities is a youth team coach. Similar stories happen between junior consultants and consulting firms, research assistants and professors, judicial clerks and judges, etc.

I examine a continuous time game between an agent (he), an intermediary (she) and a competitive labor market. The agent's wage is determined by the market's willingness to pay for his service, and the latter depends on his talent, which is his private information. Signals about his performance take the form of *conclusive good news* (or '*breakthroughs*') and arrive according to a Poisson process, with arrival rate increasing in the agent's talent as well as his unobserved effort.[3]

[1] Asymmetric learning between the current employers and the labor market about a worker's ability is a well-known fact documented in the labor economics literature, including the theoretical works of Waldman (1984), Milgrom and Oster (1987) etc. In a recent paper, Kahn (2013) uses National Longitudinal Survey of Youth data and finds empirical evidence supporting asymmetric learning.

[2] Examples where the intermediary can commit include Ekmekci (2011), Ely (2015), Halac et al.(2015), Kremer et al.(2015), Che and Hörner (2015), Hörner and Lambert (2015), etc.

[3] As argued in Bonatti and Hörner (2015), learning via infrequently arrived good news is a distinctive feature of '*creative industries*', in which '*creativity and originality are essential for success*'. This applies to professional service industries, namely, law and consulting, R&D, academia, professional sports, etc. in which a reputation is established at several defining moments of one's career. This includes, for example, bringing in a new client to the firm, making a breakthrough innovation, publishing a paper, scoring a hat-trick in an important match, etc.

Moreover, talent and effort are complements (Dewatripont et al.1999).

The novelty of my model is that the market observes a breakthrough only after the intermediary discloses it. Motivated by the aforementioned applications, I assume that the intermediary benefits from the agent's effort as well as from establishing him in front of the public.[4] Since one breakthrough is sufficient to convince the market about the agent's competence, no interesting strategic interaction takes place once a breakthrough is disclosed, after which both the intermediary and the agent receive a fixed continuation value.

In Section 2.3, I characterize the *unique* Markov Perfect Equilibrium (MPE) of this game. I find that when players are patient, the agent's effort is inverse U-shaped and the disclosure rate is decreasing over time. Intuitively, since the intermediary cannot commit, her incentive to disclose information only depends on the comparison between the future revenue she can milk from the agent and the lump sum payoff from disclosure. Hence, she has more incentive to withhold information when the agent's future effort is higher. On the other hand, since the market learns via infrequently arrived good news, the agent's incentive to exert effort decreases with his current continuation value (which is increasing in his reputation) and increases with the current disclosure rate. As a result, the intermediary has an incentive to suppress information only when the market becomes pessimistic and the rate of disclosure is decreasing over time. In response to this, the agent's effort eventually decreases. My result predicts that conditional on staying at the entry level job, both the agent's real wage as well as his chances of being promoted are inverse U-shaped, which echoes the empirical findings in Baker, Gibbs and Holmström (1994a,b).[5]

Comparing this with the unique equilibrium in the benchmark scenario in which the market observes all breakthroughs automatically (or *exogenous information benchmark*), I find that in both cases, the agent's effort is too low at the beginning, relative to social first best, exhibiting a *procrastination inefficiency*. Interestingly, the agent's continuation value is not monotone in disclosure rate. In particular, when effort cost is low, despite the intermediary will withhold information when the market becomes pessimistic, the agent's continuation value at optimistic beliefs can stil-

[4] The intermediary benefits from establishing her agent is a distinct and indispensable feature of my model, which is relevant in the aforementioned applications. In professional sports, soccer clubs receive transfer fees by selling their players to bigger clubs. In professional service industries, namely law and consulting, an established former employee is more likely to work in-house for a potential client, rather than working for a competing firm in the same industry. Once an alumni becomes an in-house attorney or counsel for a client, this also creates a network benefit for his former law or consulting firm.

For example, a former Boston Consulting Group (BCG) employee, who now runs Red Hat, a big software company, said that "*I wouldn't say I am blindly loyal, but I do use BCG more than any other firm.*" (excerpted from an Economist article in 2014) An SRZ (a large law firm in NYC) alumni, who is now the General Counsel at Sterling Stamos Capital Management, said in an interview that "*so while you need to have great knowledge of the law, you also need great resources. This is where SRZ comes in...*" (excerpted from SRZ alumni website) Network benefits can also take various other forms. For example, an article about professional service firms in *The Economist* suggests that "*... former employees are increasingly treated as assets, not turncoats... such firms are trying to stay in touch with departed workers, hoping to turn them into brand ambassadors, recruiters and salespeople*".

[5] Baker, Gibbs and Holmström (1994a,b) use 20 years of personnel data of management employees from a large US firm and find these two empirical facts. First, the real wage of a worker conditional on no promotion is first increasing and then decreasing over time (Figure IV, page 951). Second, the promotion rate from level 2 (entry level for management employees) to level 3 (intermediate level) is also an inverse U-shaped function of time stayed in level 2 (Table IV, page 902).

l be higher under endogenous information relative to the exogenous information benchmark, i.e. withholding information can encourage procrastination.

Intuitively, this is because low disclosure rate has two effects on the agent's payoff. A direct effect which makes it hard for the agent to establish himself. An indirect effect which *slows down market learning*. When effort cost is low, the intermediary's equilibrium disclosure rate is also low, and the agent can maintain a good reputation for a long time, although no breakthrough has been disclosed. This is because the market attributes the lack of news to the low disclosure rate, instead of the agent's incompetence. As a result, the agent can sustain a high flow payoff, which discourages him to exert effort.

In Section 2.4, I explore the possibility of mitigating the procrastination inefficiency by allowing players' strategies to depend on payoff irrelevant state variables, for example, whether breakthroughs have been concealed in the past or not. In particular, I characterize a tractable subclass of (non-Markov) Perfect Bayesian Equilibrium, which I call Semi-Markov Equilibrium (SME), in which players' strategies are only required to be Markov on the equilibrium path. This solution concept minimally departs from MPE and the equilibrium strategies have intuitive interpretations: every SME is characterized by a *cutoff belief*, such that the intermediary withholds information only when the market's belief falls below this cutoff. This cutoff can take any value within a compact interval, with the unique MPE being the SME with the highest cutoff and the equilibrium under exogenous information being the SME with the lowest cutoff.

Since the time at which the intermediary starts to withhold information matters for the agent's incentive, SME improves efficiency by offering flexibility in choosing this cutoff belief. For example, when effort cost is low, in order to prevent the agent from receiving high flow payoffs before establishing himself, withholding information should only happen at sufficiently pessimistic beliefs.

In Section 2.5, I characterize the optimal Markov policy when the intermediary can commit. When the market's prior belief is not too pessimistic,[6] there exists a '*deadline*' in the optimal policy, after which no information is disclosed and the agent has no incentive to exert effort. This is in sharp contrast to the no commitment case, in which the incentive to exert effort never vanishes before a breakthrough is disclosed. This deadline minimizes the agent's continuation value upon reaching it, which facilitates incentive provision early on in the game. However, due to the intermediary's sequential rationality constraint, deadlines are never implementable in any Perfect Bayesian Equilibrium in absence of commitment.

In Section 2.6, I examine the robustness of my result along several directions, including cases in which the intermediary can disclose past breakthroughs, the market can also learn from public signals in addition to the intermediary's private signals, the intermediary is the agent's direct supervisor instead of his current employer, etc. I also discuss how to enrich my model in order to incorporate more realistic features.

Related Literature: This paper contributes to a burgeoning literature on the dynamic provision of incentives via information disclosure in principal-agent relationships. The majority of these

[6] Notice that the optimal Markov policy depends on the market's prior belief.

papers focus on how to motivate an agent by providing him informational feedback. Some prominent examples include Campbell et al.(2014), Ely (2015), Kremer et al.(2015), Che and Hörner (2015), Halac et al.(2015). Comparing with these papers, I ask a novel question, that is, how to motivate an agent by releasing information to a *third party* (namely, the market).

Two papers that share a similar motivation with mine are Ekmekci (2011) and Hörner and Lambert (2015). The former constructs a rating system that sustains reputation building incentives. The latter studies the Holmström (1999) career concern model and examines the optimal design of public information structure to maximize the agent's effort in stationary equilibria.

The main differences are as follows. First, in terms of the modeling choice, the informational intermediary can commit to dynamic disclosure policies in both of these papers, which assumes away the strategic issues in information disclosure. In contrast, I focus on the complementary case in which the intermediary *cannot commit* and is a strategic long run player. While the commitment benchmark fits better into applications such as online platforms and credit rating agencies, the non-commitment case is more coherent to disclosure problems in firms and organizations, where employers enjoy substantive discretion in releasing information about their employees. Second, in terms of the main result, I focus on the *dynamics* of effort and disclosure rate, instead of examining the sustainability of reputation building incentives or the maximal *stationary* effort level.

My model builds on continuous time reputation models with Poisson good news, such as Faingold and Sannikov (2011).[7] The main difference is that there are two long run players, giving rise to a multiplicity of equilibria. As a result, more restrictive solution concepts (MPE and SME) are required to make sharp predictions. The intertemporal substitutability of effort is also reported in the strategic experimentation literature as well as its applications in venture capital financing (for example, Bergemann and Hege 1998, 2005, Hörner and Samuelson 2013). Common in these papers, if an agent shirks today, he retains the option value of succeeding tomorrow, which causes inefficient delays. These papers offer solutions when formal contracts are available: by making the agent's share of surplus a decreasing function of the elapsed time. My paper proposes a complementary solution when output is not contractible: by decreasing the publicity of the agent's performance after the market's belief falls below an endogenously chosen cutoff.[8]

2.2 The Baseline Model

I introduce a baseline model in this section, which highlights the mechanisms at work and will be the main focus of this paper. I will discuss the robustness of my results as well as how to enrich this model to incorporate more realistic features in Section 2.6.

[7] Other reputation models with Poisson good news include Board and Meyer-ter-Vehn (2013), Halac and Prat (2015), etc. After removing the intermediary, my model can be viewed as a continuous time analogue of Mailath and Samuelson (2001), in which the competent agent exerts effort to distinguish himself from the inept type.

[8] Another difference between these papers and mine is that the agent does not know his type in strategic experimentation models, while in my reputation building model, he knows his type. I will discuss the '*career concern*' case in Section 2.6.

Players & Actions: There is an agent (he, junior worker), an intermediary (she, current employer of the worker) and a competitive labor market (or '*market*'). Time $t \in [0, +\infty)$ is continuous. The agent's type is denoted by $\theta \in \{0, 1\}$, which is constant over time and is either high ($\theta = 1$) or low ($\theta = 0$). At every time t, he chooses an effort level $\tilde{a}_t \in [0, 1 - \phi]$, where $\phi \in (0, 1)$, and produces outputs called '*breakthroughs*', which are generated according to a Poisson process with arrival rate $\mu\theta(\tilde{a}_t + \phi)$, where $\mu > 0$ is a parameter. As we can see, effort and talent are complements (Dewatripont et al.1999).

Whenever a breakthrough arrives, the intermediary decides at that instant between disclosing it publicly and withholding it.[9] Importantly, the intermediary *cannot* commit to dynamic disclosure policies and *cannot* disclose when breakthroughs do not exist.[10] Let $\chi_t \in [0, 1]$ be the probability of disclosing a breakthrough conditional on its arrival at time t. Throughout the paper, I will be focusing on the '*informational intermediary*' role of the employer while abstracting away from the others.

Information Structure & History: The agent's effort and his type are only known by himself, i.e unbeknownst to the market and the intermediary. Let $\pi_0 \in (0, 1)$ be the probability that their prior belief attaches to $\theta = 1$. Let π_t be *the market's* posterior belief at time t, which I will refer to as '*the agent's reputation*'. Whenever a breakthrough arrives, it is automatically observed by the agent and the intermediary. The novelty of my model is that the market can observe a breakthrough *if and only if* the intermediary discloses it. Since the low type can never produce any breakthroughs, the market knows $\theta = 1$ after a disclosure. I say that the agent '*establishes himself*' when π_t reaches 1.

A public history, $h^t \in [0, t]$, consists of a sequence of past disclosure dates $0 \leq t_1 < ... < t_n \leq t$, with $\{\varnothing\}$ the history that no breakthrough has been disclosed. The intermediary's private history, $h^t_m \equiv (h^t, \overline{h}^t_m)$, consists of the public history, as well as $\overline{h}^t_m \subset [0, t]$, which is a sequence of breakthrough arrival times. Since breakthroughs cannot be forged, $h^t \subset \overline{h}^t_m$. The agent's private history, $h^t_a \equiv (h^t_m, a^t, \theta)$, consists of the intermediary's private history, his past effort choices, $a^t \equiv \{a_{t'}\}_{t' \in [0,t]}$, as well as his type. Let h^{t-}, h^{t-}_m and h^{t-}_a be the (public and private) histories up to, but not including time t, and let H^{t-}, H^{t-}_m and H^{t-}_a be the set of histories up to time t. Let $H \equiv \bigcup_{t=0}^{\infty} H^t$, $H_m \equiv \bigcup_{t=0}^{\infty} H^t_m$ and $H_a \equiv \bigcup_{t=0}^{\infty} H^t_a$ be the entire set of histories. I use '$\succ$' to denote the successor relationship between two histories.

Payoffs: Both the agent and the intermediary are risk neutral and discount their future payoffs at rate $r > 0$. For a player receiving flow payoffs $\{U_t\}_{t \geq 0}$, his or her (normalized) *continuation*

[9] As I will explain later, the equilibrium in my baseline model remains robust when we allow for disclosing past breakthroughs (or '*delayed disclosure*'). Moreover, in other equilibria that arise under delayed disclosure, several important qualitative features of the equilibrium in the baseline model remain robust. I will discuss this in Section 2.6.

[10] Scenarios in which the intermediary cannot announce '*forged breakthroughs*' include: when information is verifiable, or when she faces significant penalties (for example, in terms of reputation costs) for disclosing false information, etc. See Milgrom (2008) for more discussions on the applications of disclosure models.

value at t is:

$$r\int_t^\infty e^{-r(s-t)}U_s ds.$$

As in other reputation or career concern models, for example, Holmström (1999), Mailath and Samuelson (2001), I assume that the agent's wage, w_t, is determined by the labor market's expected willingness to pay for his output produced at time t.

The agent's flow payoff is $w_t - c\tilde{a}_t$, where $c \in (0, 1)$ is his marginal cost of effort. Since the low type can never produce any breakthroughs, his effort is always 0. Hereafter, I will focus exclusively on the high type. The market's willingness to pay for each breakthrough is $\frac{1}{\mu}$, which normalizes the high type's marginal product of effort to 1 and implies that it is always socially efficient for him to exert effort. Let $a_t \equiv \tilde{a}_t + \phi \in [\phi, 1]$, we have $w_t = \pi_t a_t$.[11] Hereafter, I abuse terminology and refer a_t as the agent's effort.

Next, I specify the intermediary's payoff. According to the interpretation that she is the agent's current employer in professional service industries, she receives the agent's output and pays his wage when $\pi_t < 1$, i.e. her flow payoff is $\theta a_t - w_t$ before the agent establishes himself.[12] After $\pi_t = 1$, the intermediary's flow payoff is constantly b. In the law and consulting firm application, the agent leaves the entry level job offered by the intermediary after establishing himself, and b is her network benefit for having one of her former employees working in-house for a potential client. In what follows, I will focus on the case in which $b \in (\phi, 1)$.[13] I will discuss other values of b as well as alternative specifications of the intermediary's payoff in Section 2.6.

Strategies: The agent chooses an *effort plan*, $\mathbf{a} \equiv \{a(h_a^t)\}_{h_a^t \in H_a}$. The intermediary chooses a *disclosure plan* $\boldsymbol{\chi} \equiv \{\chi(h_m^t)\}_{h_m^t \in H_m}$, with $\chi(h_m^{t-})$ being the probability of disclosing information at time t conditional on $t \in h_m^t$. Truncating the effort plan and the disclosure plan at t, we get $\{a(h_a^{t'}), \chi(h_m^{t'})\}_{t' \leq t}$. This together with the event $\{\theta = 1\}$ induce a probability measure over $H^t \times H_m^t \times H_a^t$, which I denote by $\mathcal{P}_t^{\mathbf{a},\boldsymbol{\chi}}$. The projection of $\mathcal{P}_t^{\mathbf{a},\boldsymbol{\chi}}$ on H^t induces a probability measure over the public histories, with $\mathbb{E}_t^{\mathbf{a},\boldsymbol{\chi}}[\cdot]$ the expectation taken under this measure.

The market's believed effort plan and disclosure plan are $\hat{\mathbf{a}} \equiv \{\hat{a}(h_a^t)\}_{h_a^t \in H_a}$ and $\hat{\boldsymbol{\chi}} \equiv \{\hat{\chi}(h_m^t)\}_{h_m^t \in H_m}$ respectively. These together with π_0 govern the joint distribution over θ and the public histories. Let $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}$ be the probability measure over $H^t \times H_m^t \times H_a^t$ induced by $\hat{\mathbf{a}}$, $\hat{\boldsymbol{\chi}}$ and π_0. Let $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[h^t] \in \Delta(H_m^t \times H_a^t)$ be the projection of $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}$ on $H_m^t \times H_a^t$ conditional on the public history being h^t, which is the market's '*conditional belief*' over the private histories. Let $\mathbb{E}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[\cdot|h^t]$ be the expectation under $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[h^t]$. Let $\boldsymbol{\pi} : H \to \Delta(\Theta)$ be a '*market belief system*', with $\pi(h^t)$ being the probability the market attaches to $\theta = 1$ after observing public history h^t.

[11] As in other reputation or career concern models, w_t depends on the market's expected a_t, instead of the true a_t. I will distinguish between these two in the '*Strategies*' paragraph, in which I formally introduce the notation for '*believed effort*'.

[12] Her valuation for the agent's output (or breakthrough) is the same as the market's.

[13] When $b \in (\phi, 1)$, conditional on knowing $\theta = 1$, the intermediary has a strict incentive to disclose information if the agent's future effort is constantly ϕ (at its minimum) and has a strict incentive to withhold information if the agent's future effort is constantly 1.

Policies: A '*policy*', $(\mathbf{a}, \boldsymbol{\chi})$, consists of an effort plan and a disclosure plan. I introduce two classes of policies: Markov and Semi-Markov Policies.

Definition 2.1. *A policy is Markov if for every* (h_a^{t-}, h_m^{t-}) *and* $(\hat{h}_a^{t-}, \hat{h}_m^{t-})$*, if* $\pi(h^{t-}) = \pi(\hat{h}^{t-})$*, then*

$$\Big(a(h_a^{t-}), \chi(h_m^{t-})\Big) = \Big(a(\hat{h}_a^{t-}), \chi(\hat{h}_m^{t-})\Big).$$

When analyzing Markov policies, I use $(a(\pi_t), \chi(\pi_t))$ or (a_t, χ_t) instead of $(a(h_a^{t-}), \chi(h_m^{t-}))$ for notation simplicity, where π_t is the market's belief at t when $h^{t-} = \{\varnothing\}$. Furthermore, $\mathbb{E}_t^{\mathbf{a},\boldsymbol{\chi}}[\cdot]$ and $\mathbb{E}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[\cdot|h^{t-} = \{\varnothing\}]$ are induced by Poisson processes with instantaneous arrival rates $\mu a_t \chi_t$ and $\mu \pi_t a_t \chi_t$.

Next, I define Semi-Markov Policies.

Definition 2.2. *A policy is Semi-Markov if for every* (h_a^{t-}, h_m^{t-}) *and* $(\hat{h}_a^{t-}, \hat{h}_m^{t-})$*, if* $\pi(h^{t-}) = \pi(\hat{h}^{t-})$ *and* $\mathcal{P}_t^{\mathbf{a},\boldsymbol{\chi}}(h_a^{t-}, h_m^{t-}) > 0, \mathcal{P}_t^{\mathbf{a},\boldsymbol{\chi}}(\hat{h}_a^{t-}, \hat{h}_m^{t-}) > 0$*, then*

$$\Big(a(h_a^{t-}), \chi(h_m^{t-})\Big) = \Big(a(\hat{h}_a^{t-}), \chi(\hat{h}_m^{t-})\Big).$$

In a nutshell, a Semi-Markov policy only requires that $\mathbf{a}$ and $\boldsymbol{\chi}$ to be Markov *on the equilibrium path*, while allowing for non-Markov plans off-path. By definition, the set of Semi-Markov policies contains the set of Markov policies.

Solution Concepts: A *Perfect Bayesian Equilibrium* (or PBE) consists of an equilibrium policy $\{\mathbf{a}, \boldsymbol{\chi}\}$, the market's conditional belief over private histories, $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[h^t]$ (induced by the believed policy $\{\hat{\mathbf{a}}, \hat{\boldsymbol{\chi}}\}$) and a wage process $\mathbf{w} : H \to \mathbb{R}_+$, such that:

1. $\mathbf{w}$ is consistent with the market's belief, i.e. $w_t \equiv \mathbf{w}(h^{t-}) = \mathbb{E}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[\theta a(h_a^{t-})|h^{t-}]$.
2. $a(h_a^{t-})$ is optimal for the agent for every h_a^{t-} given $\boldsymbol{\chi}$, $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[h^t]$ and $\mathbf{w}$.
3. $\chi(h_m^{t-})$ is optimal for the intermediary for every h_m^{t-} given $\mathbf{a}$, $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[h^t]$ and $\mathbf{w}$.
4. For every $h^t \in H$, $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}[h^t]$ is derived from $\mathcal{P}_t^{\hat{\mathbf{a}},\hat{\boldsymbol{\chi}},\pi_0}$ according to Bayes Rule.[14]
5. The market's belief is correct, i.e. $\{\mathbf{a}, \boldsymbol{\chi}\} = \{\hat{\mathbf{a}}, \hat{\boldsymbol{\chi}}\}$.

Optimality requirements in 2 and 3 imply that at every private history, the agent and the intermediary choose a_t and χ_t respectively to maximize their own expected continuation values, which are given by:

$$\mathbb{E}_t^{\mathbf{a},\boldsymbol{\chi}}\Big[r\int_t^\infty e^{-r(s-t)}\Big(w_s - c(a_s - \phi)\Big)ds\Big|h_a^{t-}\Big]$$

and

$$\mathbb{E}_t^{\mathbf{a},\boldsymbol{\chi}}\Big[r\int_t^\infty e^{-r(s-t)}U_{m,s}ds\Big|h_m^{t-}\Big],$$

[14]Bayes Rule always applies since all public histories occur with strictly positive probability on the equilibrium path.

where $U_{m,s} = \theta a_s - w_s$ if $h^s = \{\varnothing\}$ and $U_{m,s} = b$ otherwise. The intermediary's payoff is evaluated conditional on $\theta = 1$ since she has no decision to make until she knows that $\theta = 1$.[15]

Two PBEs are '*outcome equivalent*' if they induce the same joint distribution over $\{\pi_t, w_t, a_t\}_{t\in[0,+\infty)}$. I introduce two refinements of PBE, which will be the solution concepts examined in this paper:

Definition 2.3. *A Markov Perfect Equilibrium (or MPE) is a PBE in which the equilibrium policy is Markov.*[16] *A Semi-Markov Equilibrium (or SME) is a PBE in which the equilibrium policy is Semi-Markov.*

Let $(\hat{\mathbf{a}}, \hat{\boldsymbol{\chi}}) \equiv \{\hat{a}(\pi_t), \hat{\chi}(\pi_t)\}_{\pi_t\in(0,1)}$ be the believed policy in an MPE, or the believed on-path policy in an SME. I make the technical restriction that both $\hat{a}(\pi_t)$ and $\hat{\chi}(\pi_t)$ are left-continuous functions of π_t. This implies that $a(\pi_t)$ and $\chi(\pi_t)$ are also left-continuous since the market's belief is correct. Let $\mathbf{A}$ and $\mathbf{X}$ be the set of left-continuous Markov effort plans and disclosure plans.

2.3 Markov Perfect Equilibrium

In this section, I characterize the unique MPE of this game. I show that effort is inverse U-shaped and the disclosure rate is decreasing over time given that no breakthrough has been disclosed in the past. I highlight the agent's and the intermediary's incentives through their best response correspondences in Subsection 2.3.1. I characterize the unique equilibrium when the market can automatically observe all breakthroughs (or the '*exogenous information benchmark*') in Subsection 3.3.2. I state the main characterization result (Proposition 2.2) and discuss its implications in Subsection 2.3.3, which is also compared with the exogenous information benchmark in subsection 2.3.5. I sketch the proof of Proposition 2.2 in Subsection 2.3.4.

2.3.1 Preliminaries

Belief Updating: The first step is to specify the evolution of π_t. Similar to Poisson bandit models, for example Keller, Rady and Cripps (2005), if a breakthrough is disclosed, π_t jumps to 1. Otherwise, the evolution of π_t is characterized by the following ordinary differential equation

[15] Formally, since a_t is unobservable to the intermediary, we also need to specify the intermediary's believed effort plan $\overline{a}(h_a^{t-})$, and evaluate her payoff under the probability measure induced by $\{\overline{\mathbf{a}}, \boldsymbol{\chi}\}$ and $\theta = 1$. However, since first, $\mathbf{a} = \overline{\mathbf{a}}$ in equilibrium; and second, the intermediary only moves when her private belief is 1, after which her belief updating process is trivial, so omitting $\overline{\mathbf{a}}$ and evaluating the intermediary's payoff using $\mathbf{a}$ is without loss. I hope this '*inconsistency*' will not cause confusion. However, I need to specify the market's believed effort since it matters for its belief updating process.

[16] Although there has been no agreed upon definition of MPE when actions are unobservable, my definition is in the spirit of Maskin and Tirole (2001) in which players' strategies are only conditioned on the coarsest information partition such that if all other players use measurable strategies, each player's decision-making problem is also measurable. This is partly because the agent knows his own type and the intermediary only acts after knowing the agent's type, which shuts down the channel for private learning.

(ODE):[17]

$$\dot{\pi}_t = -\pi_t(1-\pi_t)Y_t, \tag{2.1}$$

where $Y_t \equiv \mu\chi_t a_t$ is the arrival rate of publicly disclosed breakthroughs conditional on $\theta = 1$.

The Agent's Incentives: Let $V_a(\pi_t)$ be the agent's (equilibrium) continuation value when his reputation is π_t, which can be decomposed into a weighted average of his flow payoff in the time interval $(t, t+dt]$ and his continuation value at $t+dt$:[18]

$$V_a(\pi_t) = r\Big(\pi_t a_t - c(a_t-\phi)\Big)dt + (1-rdt)\Big(\underbrace{Y_t dt}_{\text{prob. of disclosure}} V_a(1) + \underbrace{(1-Y_t dt)}_{\text{prob. of no disclosure}} V_a(\pi_{t+dt})\Big). \tag{2.2}$$

Expanding this equation and ignoring higher order terms, we get:

$$V_a(\pi_t) = \Big(\pi_t a_t - c(a_t-\phi)\Big) + \frac{Y_t}{r}\Big(V_a(1) - V_a(\pi_t)\Big) + \frac{1}{r}\frac{dV_a(\pi_t)}{dt}. \tag{2.3}$$

The law of motion of π_t implies that when $Y_t \neq 0$,

$$\frac{dV_a(\pi_t)}{dt} = -\pi_t(1-\pi_t)Y_t V_a'(\pi_t).$$

Next, suppose the agent deviates and chooses a_t different from the market's believed effort $\hat{a}_t$, his continuation payoff at π_t is:[19]

$$\Big(\pi_t \hat{a}_t - c(a_t-\phi)\Big) + \frac{Y_t}{r}\Big(V_a(1) - V_a(\pi_t)\Big) - \frac{\hat{Y}_t}{r}\pi_t(1-\pi_t)V_a'(\pi_t), \tag{2.4}$$

with $\hat{Y}_t \equiv \mu\chi_t\hat{a}_t$. This is because the agent's effort is unobservable, deviations only affect his cost of effort as well as the arrival rate of breakthroughs. But conditional on no disclosure, π_t is still updated according to the '*believed effort*'. Choosing a_t to maximize (2.4) obtains the agent's best response correspondence:

$$a_t \begin{cases} = \phi & \text{when } \frac{\mu\chi_t}{r}(V_a(1) - V_a(\pi_t)) < c \\ \in [\phi, 1] & \text{when } \frac{\mu\chi_t}{r}(V_a(1) - V_a(\pi_t)) = c \\ = 1 & \text{when } \frac{\mu\chi_t}{r}(V_a(1) - V_a(\pi_t)) > c \end{cases}. \tag{2.5}$$

The term $\frac{\mu\chi_t}{r}(V_a(1) - V_a(\pi_t))$ measures his '*reputational incentive*', which is increasing in the publicity of his performance ($\mu\chi_t$); decreasing in the discount rate (r) and increasing in the difference between his continuation value when $\pi_t = 1$ and his current continuation value. After a breakthrough

[17] Admissibility requires that both $a(\pi_t)$ and $\chi(\pi_t)$ are left continuous, so is $\mu a(\pi_t)\chi(\pi_t)$. In a good news model, no news is bad news, i.e. $\dot{\pi}_t \leq 0$, so according to Klein and Rady (2011), there exists a solution to ODE (2.1) for any given π_0. If there are multiple solutions to this initial value problem, I select the one that is consistent with discrete time approximation, which is shown to be unique.

[18] The following ODE uses the fact that in equilibrium, the market's belief is always correct, i.e.$\hat{a}_t = a_t$.

[19] Expression (2.4) deals with the case in which $\hat{Y}_t \neq 0$. When $\hat{Y}_t = 0$, the agent's continuation payoff following a deviation is $(\pi_t\hat{a}_t - c(a_t-\phi)) + \frac{Y_t}{r}(V_a(1) - V_a(\pi_t))$.

is disclosed, $\pi_t = 1$ and the agent's reputational incentive disappears. So $a_t = w_t = V_a(1) = \phi$.[20] Since χ_t is bounded below 1, he has an incentive to work only if his continuation value is low enough, i.e.

$$V_a(\pi_t) \leq \overline{V}_a \equiv \phi - \frac{cr}{\mu}. \tag{2.6}$$

The Intermediary's Incentives: Let $V_m(\pi_t)$ be the intermediary's continuation value conditional on $\theta = 1$. Similar to $V_a(\pi_t)$, $V_m(\pi_t)$ satisfies the following ODE:

$$V_m(\pi_t) = a_t(1 - \pi_t) + \frac{Y_t}{r}\Big(b - V_m(\pi_t)\Big) + \frac{1}{r}\frac{dV_m(\pi_t)}{dt}. \tag{2.7}$$

Since her deviations cannot be observed by the market either, her best response correspondence is:

$$\chi_t \begin{cases} = 1 & \text{when } V_m(\pi_t) < b \\ \in [0,1] & \text{when } V_m(\pi_t) = b \\ = 0 & \text{when } V_m(\pi_t) > b \end{cases}. \tag{2.8}$$

From (2.8), the intermediary is more reluctant to disclose information when $V_m(\pi_t)$ is high and vice versa. Intuitively, due to the lack of commitment, her incentive to disclose information only depends on the revenue she can milk from the agent by continuing the relationship after she knows that $\theta = 1$.

Patience Level Conditions: As in other dynamic game models, the discount rate r matters for the equilibrium outcomes. Therefore, I introduce the following patience level condition:

Condition 2.1. *Players' patience level is high if $r < \frac{\mu\phi}{c}(1 - c)$ and is low otherwise.*

Patience level is high is equivalent to $\overline{V}_a > c\phi$. This condition is less demanding when the talent premium ϕ is higher, the arrival rate μ is higher or the cost of effort c is lower.

Indifference & Value Invariance Curves: I define two curves, which are of critical importance in analyzing the long run players' dynamic incentives. First, for a given time interval (t_0, t_1), suppose the intermediary is indifferent between disclosing and withholding information for all $t \in (t_0, t_1)$, then $V_m(\pi_t) = b$ and $\frac{dV_m(\pi_t)}{dt} = 0$. Equation (2.7) implies that $a_t(1 - \pi_t) = b$. Let

$$a^*(\pi_t) \equiv \min\left\{1, \frac{b}{1 - \pi_t}\right\} \tag{2.9}$$

be the intermediary's '*indifference curve*', which is increasing in π_t and becomes flat once π_t exceeds $1 - b$.

[20] By replacing $V_a(1)$ with any other constant V^*, it is straightforward to extend my result to the case in which the agent receives continuation value V^* after a breakthrough is disclosed, which is assumed in Bonatti and Hörner (2015).

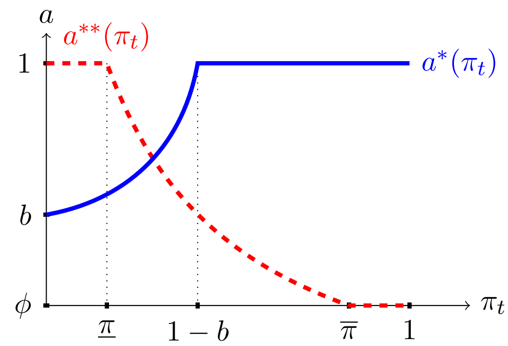

Figure 2-1: Indifference Curve (solid) and Value Invariance Curve (dashed)

Second, when patience level is high i.e. $\overline{V}_a > c\phi$, let $a^{**}(\pi_t)$ be defined as:

$$a^{**}(\pi_t) \equiv \begin{cases} 1 & \text{when } \pi_t \leq \underline{\pi} \\ \frac{\phi(1-c)-\frac{rc}{\mu}}{\pi_t} & \text{when } \pi_t \in (\underline{\pi}, \overline{\pi}] \\ \phi & \text{when } \pi_t > \overline{\pi}, \end{cases} \tag{2.10}$$

in which $\underline{\pi} \equiv \phi(1-c) - \frac{cr}{\mu}$ and $\overline{\pi} \equiv (1-c) - \frac{cr}{\mu\phi}$. The high patience level condition ensures that $0 < \underline{\pi} < \overline{\pi} < 1$. I call $a^{**}(\pi_t)$ the agent's '*value invariance curve*', which is flat at both ends and is strictly decreasing when $\pi_t \in [\underline{\pi}, \overline{\pi}]$. Intuitively, if $\pi_t \in [\underline{\pi}, \overline{\pi}]$, $V_a(\pi_t) = \overline{V}_a$ and $\chi(\pi_t) = 1$, then the agent's continuation value remains unchanged (i.e. $\frac{dV_a(\pi_t)}{dt} = 0$) if $a(\pi_t) = a^{**}(\pi_t)$. I depict the two curves together in Figure 2-1.

2.3.2 Exogenous Information Benchmark

In this subsection, I consider the benchmark scenario in which $\chi_t = 1$ for all t, i.e. the market directly observes every breakthrough the agent has achieved.[21] I call this the '*exogenous information*' benchmark, which will later be compared with the strategic disclosure case. Proposition 2.1 characterizes the unique MPE in this benchmark:

Proposition 2.1 (Exogenous Information)**.** *There exists a unique MPE. When patience level is low, $a(\pi_t) = \phi$ for all $\pi_t \in (0, 1]$. When patience level is high, there exists $\pi^{\ddagger} \in (\underline{\pi}, 1)$ such that:*

$$a(\pi_t) = \begin{cases} 1 & \textit{when } \pi_t \leq \pi^{\ddagger} \\ a^{**}(\pi_t) & \textit{when } \pi_t > \pi^{\ddagger}. \end{cases}$$

The agent's effort path when patience level is high is shown in Figure 2-2. Since this benchmark scenario fits into the definition of '*Poisson good news model*' in Faingold and Sannikov (2011), the unique MPE is also the unique Nash Equilibrium. According to Proposition 2.1, the agent's effort is increasing over time when r is low enough. As in other Poisson good news models, for example,

[21] Several alternative interpretations of this benchmark scenario include: when the agent can directly communicate with the market, or when the intermediary is required to disclose all information.

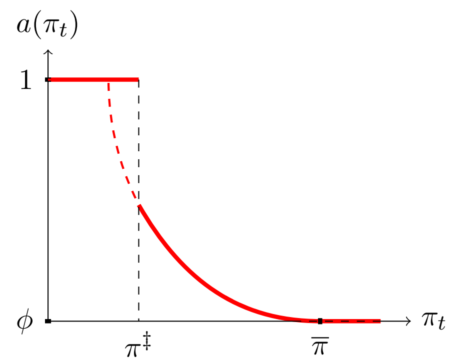

Figure 2-2: Effort Path under Exogenous Information (High Patience)

Board and Meyer-ter-Vehn (2013), the agent works too little when π_t is high (relative to social first best), leading to '*procrastination inefficiencies*'.

This is because under exogenous information, the agent's marginal benefit from exerting effort is $\frac{\mu}{r}(\phi - V_a(\pi_t))$ and $V_a(\pi_t)$ is non-decreasing in π_t. The cutoff belief, $\pi^\ddagger$, is pinned down by the '*promise keeping condition*', i.e. his continuation value at $\pi^\ddagger$ is $\overline{V}_a$. Effort is interior when $\pi_t \in (\pi^\ddagger, \overline{\pi})$ since the agent's incentive to exert effort is decreasing with the market's '*believed effort*'. This uniquely pins down the equilibrium effort level.

2.3.3 Equilibrium Characterization

I characterize the unique MPE when information is strategically disclosed by an intermediary. The main result is stated as Proposition 2.2.

Proposition 2.2. *This game admits a unique MPE.*

- *When patience level is low,* $\chi(\pi_t) = 1$ *and* $a(\pi_t) = \phi$ *for all* $\pi_t \in (0, 1]$.
- *When patience level is high, there exists a cutoff belief* $\pi^\dagger \in (\underline{\pi}, 1)$, *such that:*

$$a(\pi_t) \equiv \begin{cases} a^{**}(\pi_t) & \text{when } \pi_t > \pi^\dagger \\ a^*(\pi_t) & \text{when } \pi_t \leq \pi^\dagger \end{cases}, \quad \chi(\pi_t) \equiv \begin{cases} 1 & \text{when } \pi_t > \min\{1-b, \pi^\dagger\} \\ \frac{cr}{\mu(\phi - V_a(\pi_t))} & \text{when } \pi_t \leq \min\{1-b, \pi^\dagger\} \end{cases}.$$

Moreover, $a^*(\pi^\dagger) > a^{**}(\pi^\dagger)$ *and* $\chi(\pi_t)$ *is strictly increasing in* π_t *when* $\pi_t \leq \min\{1-b, \pi^\dagger\}$.

From now on, I will focus on the high patience case, due to the presence of non-trivial reputation building behavior, which is economically interesting. Effort and disclosure rate when patience level is high are depicted in Figures 2-3 and 2-4 as functions of π_t, depending on whether $\pi^\dagger$ is below (interior case) or above (corner case) $1-b$. Common in both figures, effort is inverse U-shaped while the rate of disclosure decreases over time. Moreover, the equilibrium effort path coincides with the intermediary's indifference curve when the market is pessimistic and coincides with the agent's value invariance curve otherwise.

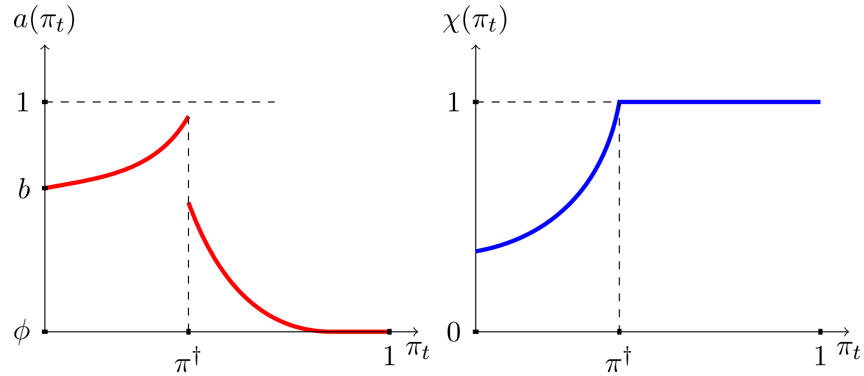

Figure 2-3: Effort and Disclosure Dynamics in the unique MPE: High Patience & Interior Case ($\pi^\dagger \leq 1 - b$)

For some intuition, since the intermediary benefits from extracting the agent's effort, she has more incentive to disclose information when future effort is lower and vice versa. When patience level is low, the agent has no incentive to exert effort and the intermediary cannot milk much revenue from the agent. Because of this, she strictly prefers to disclose information. When patience level is high, the agent has more incentive to work when π_t is lower, as in the exogenous information benchmark. Anticipating this, the intermediary will have an incentive to withhold information when π_t is low.

This gives rise to the following countervailing incentive structure: the agent has more incentive to exert effort when his performance is more visible while the intermediary has more incentive to disclose information when the agent shirks. When the intermediary is required to stochastically withhold information, effort and disclosure rate must be chosen to make the other player indifferent, which results in the described dynamic pattern. The cutoff belief $\pi^\dagger$ is pinned down by the '*promise keeping condition*': the agent's continuation value at $\pi^\dagger$ must be $\overline{V}_a$, the highest continuation value under which he can be motivated to exert effort.

Mapping this back into the application to professional service industries, my result speaks for the anecdotal evidence that firms tend to establish their star employees early on in their careers. But for those who fail to succeed in the beginning, their latter successes will be discounted more and more heavily and are less visible to the outside market. This distinction is driven by the employer's strategic motives: the worker's eagerness to build up his reputation enables his employer to milk more revenue from his hard work, which explains why he is not being promoted for a long time. This lack of transparency in performance will frustrate the worker in the long run and his effort eventually decreases.

The predictions of my model in terms of wages (w_t) and promotion rates ($a_t\chi_t$) match the empirical findings in Baker, Gibbs and Holmström (1994a,b), who study the wage and promotion dynamics empirically using 20 years of personnel data from a large US firm.[22] First, the real wage

[22]The predictions of my model, that effort is inverse U-shaped and the disclosure rate is decreasing are robust when the market can also learn from infrequently arrived public signals, when the intermediary's benefit from disclosure

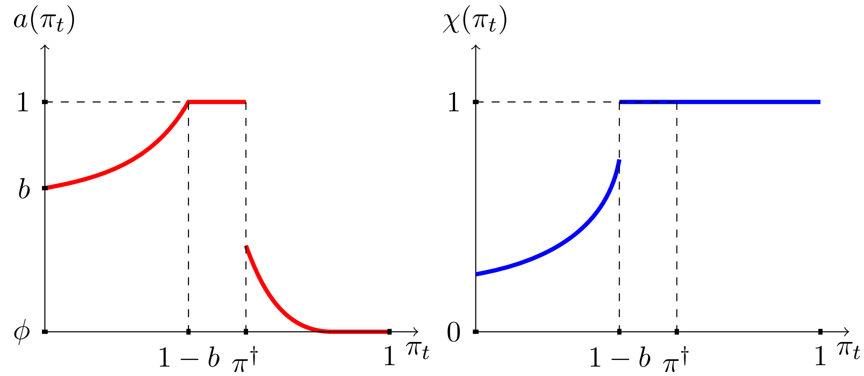

Figure 2-4: Effort and Disclosure Dynamics in the unique MPE: High Patience & Corner Case ($\pi^\dagger > 1-b$)

of a worker conditional on remaining at the entry level job (level 2 in their paper) is first increasing and then decreasing as a function of time (Figure IV on page 951). Second, the promotion rate from level 2 to level 3 is also an inverse U-shaped function of tenure (the number of years at level 2), which is documented in their Table IV (page 902). I provide a novel explanation for these facts based on the dynamic interactions between the agent's incentive to build up his reputation and the intermediary's incentive to release information to the labor market.[23]

Why MPE: The unique MPE is robust to private monitoring, i.e. when the agent cannot perfectly observe the arrival of breakthroughs. Private monitoring is prevalent in applications where performance evaluations are subjective (Fuchs 2007). In my model, MPE is attractive since players' strategies only depend on the public belief and their incentives do not rely on the fine details of their private histories. As a result, their incentive constraints are still satisfied even when breakthroughs are based on the intermediary's subjective assessment.[24]

Furthermore, this unique MPE remains to be an equilibrium when the intermediary can disclose past breakthroughs. This is because her incentive to withhold information is weakly increasing over time, i.e. given she has an incentive to withhold information today, she also has an incentive to withhold it in the future. Nonetheless, the possibility of disclosing past breakthroughs introduces additional equilibria, which will be discussed in Section 2.6.

is low, i.e. $b \in (0, \phi]$, when the intermediary has different time preferences (for example, having a different discount factor or is finitely lived), when the agent does not know θ, etc. These will be discussed in Section 2.6. They are also robust when we consider other PBEs that are non-Markov (Section 2.4).

[23] Alternative explanations in the existing literature include, for example, the inverse U-shaped promotion rate can also be explained by a combination of market learning and human capital accumulation.

[24] The robustness of equilibria under private monitoring has been discussed extensively in the repeated games literature, for example, Mailath and Morris (2002), etc. The gist of this literature is: coordinating current play on past histories becomes more challenging when monitoring is private and histories are not commonly known among players.

Remark: My result that effort is inverse U-shaped is different from, albeit complementary to, both career concern models with Gaussian learning and reputation building models with Poisson good news.

In Holmström (1999), the agent's effort is decreasing over time, which is driven by the Gaussian information structure, i.e. the market's signal is highly sensitive to effort. As time elapses, the precision of market belief increases and the impact of current effort decreases.[25] Contrary to Holmström (1999), I examine cases in which the market receives information *infrequently*. This is relevant to reputation building of junior people, as the market rarely knows about their performance before they become famous.

In Poisson good news models, for example, Board and Meyer-ter-Vehn (2013), the agent's effort is increasing over time as the market's belief deteriorates since $\phi - V_a(\pi_t)$ is decreasing in π_t. In my model, the agent's effort is inverse U-shaped.[26] This is driven by the intermediary's incentive to release information: as the market becomes more pessimistic, the intermediary can capture a larger share of the agent's surplus once she knows that $\theta = 1$. Effort is decreasing over time along the intermediary's indifference curve in order to motivate her to stochastically disclose information.

2.3.4 Proof of Proposition 2.2

I prove the result in seven steps, with proofs of the lemmas relegated to Appendix B.1 and the Online Appendix. To approach the problem, I use the observation that changes in w_t vanishes when π_t goes to 0, that is, the agent's problem becomes '*approximately stationary*' in the limit. I pin down his limiting continuation value when $\pi_t \to 0$ and characterize players' limiting equilibrium behaviors. The agent's continuation value at low enough π_t is characterized by a limiting value problem, which admits a unique solution. This solution is then used to compute $\pi^\dagger$.

Step 1: Market Learning First, I show that the market eventually learns about the agent's type. This result validates my approach of analyzing players' values and behaviors at the limiting belief.

Lemma 2.3.1. *In every MPE, $\pi_t(1-\pi_t)$ converges to 0 in probability.*[27]

[25] In a recent paper, Hörner and Lambert (2015) examine a variant of the Holmström model in which the agent's talent is changing over time, following a mean reverting process. In these '*changing type*' models, the agent's effort level can be constant over time, which is the case in their stationary equilibria.

[26] Inverse U-shaped effort is also reported in Bonatti and Hörner (2015), although the driving forces are very different. In their model, the agent does not know his type, as in the strategic experimentation literature. His effort decreases as his confidence in his own ability falls. In my model, declining effort is driven by the intermediary's incentives to suppress information, which comes from the strategic interaction between the intermediary and the agent, not through the agent's private learning.

[27] In my model, the unique MPE also coincides with the unique public equilibrium, i.e. players' strategies depend only on the public history. Aside from π_t, the market also observes calendar time. In the Online Appendix, I show $\chi(\pi_t) > 0$ under a weaker assumption that players' strategies are public, which implies the existence of a 1-to-1 mapping between t and π_t. As a result, t does not contain any extra information in addition to π_t.

An implication is that $\chi_t \neq 0$ for all t and hence, $Y_t > 0$. Therefore, we can re-write (2.3) as:

$$V_a(\pi_t) = \Big(\pi_t a_t - c(a_t - \phi)\Big) + \frac{Y_t}{r}\Big(V_a(1) - V_a(\pi_t) - \pi_t(1-\pi_t)V_a'(\pi_t)\Big).$$

Step 2: Continuation Value & Behavior at the Limiting Belief Let

$$V_a(0) \equiv \begin{cases} c\phi & \text{when } r < \frac{\mu\phi(1-c)}{c} \\ \frac{\mu\phi^2}{r+\mu\phi} & \text{when } r \geq \frac{\mu\phi(1-c)}{c} \end{cases}, \quad V_m(0) \equiv \begin{cases} b & \text{when } r < \frac{\mu\phi(1-c)}{c} \\ \frac{r\phi+\mu b}{r+\mu} & \text{when } r \geq \frac{\mu\phi(1-c)}{c} \end{cases}.$$

Lemma 2.3.2 shows that the limits of players' continuation values, $\lim_{\pi_t \to 0} V_a(\pi_t)$ and $\lim_{\pi_t \to 0} V_m(\pi_t)$, exist. It also explicitly computes these values as well as players' equilibrium behaviors at the limit:

Lemma 2.3.2. *Players' limiting continuation values exist. Furthermore, in every MPE,*

$$\lim_{\pi_t \to 0} V_a(\pi_t) = V_a(0) \text{ and } \lim_{\pi_t \to 0} V_m(\pi_t) = V_m(0),$$

and there exists $\varepsilon > 0$ such that for all $\pi_t \in (0, \varepsilon)$,

- $\chi(\pi_t) = 1$ *and* $a(\pi_t) = \phi$ *when patience level is low.*
- $a(\pi_t) = a^*(\pi_t)$ *and* $\chi(\pi_t) = \frac{cr}{\mu(\phi - V_a(\pi_t))}$ *when patience level is high.*

In what follows, I will focus on the high patience case. Lemma 2.3.2 implies that when patience level is high and π_t is close to 0, the intermediary must be indifferent between disclosing and withholding information and the agent's effort must be $a^*(\pi_t)$. Hence, the agent's continuation value when π_t is low enough is characterized by a solution to the following limiting value problem:

$$V_a(\pi_t) = \Big(\pi_t a^*(\pi_t) - c(a^*(\pi_t) - \phi)\Big) + \frac{\mu\chi(\pi_t)a^*(\pi_t)}{r}\Big(\phi - V_a(\pi_t) - \pi_t(1-\pi_t)V_a'(\pi_t)\Big) \tag{2.11}$$

with $\lim_{\pi_t \to 0} V_a(\pi_t) = c\phi$. Accordingly, $\chi(\pi_t)$ must be chosen to provide him an incentive to choose an interior effort level, i.e.

$$\frac{\mu\chi(\pi_t)}{r}\Big(\phi - V_a(\pi_t)\Big) = c. \tag{2.12}$$

Plugging (2.12) into (2.11) results in the following ODE:

$$V_a(\pi_t) = \Big(\pi_t a^*(\pi_t) + c\phi\Big) - \frac{ca^*(\pi_t)\pi_t(1-\pi_t)V_a'(\pi_t)}{\phi - V_a(\pi_t)}, \tag{2.13}$$

with $\lim_{\pi_t \to 0} V_a(\pi_t) = c\phi$. Lemma 2.3.3 shows that this problem admits a unique solution:

Lemma 2.3.3. *The above limiting value problem admits a unique solution $V_a^*(\pi_t)$. When patience level is high, there exists $\varepsilon > 0$ such that $V_a(\pi_t) = V_a^*(\pi_t)$ for all $\pi_t \in (0, \varepsilon)$.*[28]

[28]The ODE in (2.13) can be transformed into a Bernoulli Equation, by letting $Z(\pi_t) \equiv \phi - V_a(\pi_t)$. This admits a closed form solution in limiting form. However, this formula is inconvenient both for characterizing the equilibrium as well as for doing comparative statics. Hence, throughout my analysis, I use an indirect approach to establish the uniqueness as well as other properties of MPE.

The proof uses the observation that different values of $V_a(\cdot)$ at any interior belief will lead to diverging values when $\pi_t \to 0$. As a result, the limiting value uniquely pins down the value at every interior belief.

Step 3: Value Invariance Curve The next Lemma exploits the implications of the agent's value invariance curve.

Lemma 2.3.4. *Suppose $\chi_t = 1$ and $V_a(\pi_t) = \overline{V}_a$,*

- *If $\pi_t < \underline{\pi}$, then $V_a'(\pi_t) < 0$ for all $a_t \in [\phi, 1]$. If $\pi_t > \overline{\pi}$, then $V_a'(\pi_t) > 0$ for all $a_t \in [\phi, 1]$.*
- *If $\pi_t \in [\underline{\pi}, \overline{\pi}]$, then*

$$V_a'(\pi_t) \begin{cases} > 0 & \text{when } a_t > a^{**}(\pi_t) \\ = 0 & \text{when } a_t = a^{**}(\pi_t) \\ < 0 & \text{when } a_t < a^{**}(\pi_t). \end{cases}$$

The proof is straightforward from (2.3), which is omitted. Intuitively, fixing $V_a(\pi_t)$, higher believed effort in $(t-dt, t]$ leads to higher wages in $(t-dt, t]$ and therefore, higher $V_a(\pi_{t-dt}) - V_a(\pi_t)$. Since π_t is decreasing over time, this implies that $V_a'(\pi_t)$ is larger.

Step 4: Constructing an MPE I construct an MPE in the high patience case and later show its uniqueness. Let

$$\pi^\dagger \equiv \sup\left\{\pi_t \middle| V_a^*(\pi) < \overline{V}_a \text{ for all } \pi \in (0, \pi_t)\right\}. \tag{2.14}$$

By definition, if $\pi^\dagger < 1$,[29] then $V_a(\pi^\dagger) = \overline{V}_a$ and $\lim_{\pi_t \uparrow \pi^\dagger} V_a'(\pi_t) > 0$. I claim that the strategy profile displayed in Proposition 2.2 and its induced conditional belief constitute an MPE. This also implies that the agent's continuation value when $\pi_t < \pi^\dagger$ is $V_a^*(\pi_t)$, which is the unique solution to the limiting value problem in Lemma 2.3.3.

I check players' incentive constraints. The agent's incentive constraints are satisfied when $\pi_t \le \min\{1-b, \pi^\dagger\}$ since the choice of $\chi(\pi_t)$ makes him indifferent between working and shirking. When $\pi_t \in (\min\{1-b, \pi^\dagger\}, \pi^\dagger]$, the definition of $\pi^\dagger$ implies that $V_a(\pi_t) \le \overline{V}_a$ and the intermediary fully discloses information, implying that he has a strict incentive to exert effort. When $\pi_t > \pi^\dagger$,

- If $a^{**}(\pi_t) > \phi$, then his continuation value remains $\overline{V}_a$ and he is indifferent.
- If $a^{**}(\pi_t) = \phi$, then his continuation value is weakly above $\overline{V}_a$ and he weakly prefers to shirk.

The intermediary's incentive constraint when $\pi_t \le \min\{1-b, \pi^\dagger\}$ is satisfied since her flow payoff is constantly b thereafter. To verify her incentive constraints when $\pi_t > \min\{1-b, \pi^\dagger\}$, I only need to show:

$$a^*(\pi^\dagger) > a^{**}(\pi^\dagger). \tag{2.15}$$

Since $\lim_{\pi_t \uparrow \pi^\dagger} V_a'(\pi_t) > 0$, (2.15) is then implied by Lemma 2.3.4.

[29] I will defer the proof of $\pi^\dagger < 1$ to Lemma 2.3.8, the proof of which does not rely on any of the previous Lemma.

Step 5: Monotonicity of Disclosure Rates I show that $\chi(\pi_t)$ is strictly increasing in π_t for all $\pi_t \le \min\{1-b, \pi^\dagger\}$. It is equivalent to show that $V_a^*(\pi_t)$ is strictly increasing in π_t, or equivalently, the agent's equilibrium continuation value is decreasing over time, conditional on no disclosure.

Lemma 2.3.5. *When $\pi_t \le \pi^\dagger$, $V_a^*(\pi_t)$ is strictly increasing in π_t.*

Step 6: Uniqueness of MPE First, I show that if $\pi_t \le \min\{1-b, \pi^\dagger\}$, then $a(\pi_t) = a^*(\pi_t)$ and $\chi(\pi_t)$ must be chosen to make the agent indifferent. Let

$$\pi^1 \equiv \sup\left\{\pi_t \middle| \pi_t \in (0, \min\{1-b, \pi^\dagger\}],\ a(\pi) = a^*(\pi) \text{ for all } \pi < \pi_t\right\}.$$

I show the following Lemma:

Lemma 2.3.6. *In every MPE, $\pi^1 = \min\{1-b, \pi^\dagger\}$.*

Next, I show that the intermediary always fully discloses information when $\pi_t > \min\{1-b, \pi^\dagger\}$. If $\pi^\dagger > 1-b$ and $\pi_t \in (1-b, \pi^\dagger)$, then $V_a^*(\pi_t) \le \overline{V}_a$ and the intermediary always strictly prefers to disclose, so $a(\pi_t) = \chi(\pi_t) = 1$. When $\pi_t > \pi^\dagger$, Lemma 2.3.7 shows that effort must be $a^{**}(\pi_t)$ in every MPE. Let

$$\pi^2 \equiv \sup\left\{\pi_t \middle| \pi_t \in [\pi^\dagger, 1) \text{ and } a(\pi) = a^{**}(\pi) \text{ for all } \pi \in [\pi^\dagger, \pi_t)\right\}.$$

Lemma 2.3.7. *In every MPE, $\pi^2 = 1$.*

Step 7: Range of $\pi^\dagger$ It is already known from Step 4 that $a^{**}(\pi^\dagger) < a^*(\pi^\dagger)$, which gives a strictly positive lower bound on $\pi^\dagger$, i.e. the lowest belief at which $a^*(\cdot)$ and $a^{**}(\cdot)$ intersect. In this step, I establish an upper bound:

$$\pi^\dagger \le 1 - \frac{cr}{\mu\phi}. \tag{2.16}$$

I show the following claim which implies (2.16): when the agent's reputation is π_t, given that the market's belief is correct, his continuation value is at least $\pi_t\phi$ for any Markov disclosure plan. This minimum is achieved when $\chi(\pi_t) = 0$ and $a(\pi_t) = \phi$, i.e. no information is disclosed and market's belief is π_t forever. This result is also interesting by itself since it characterizes the harshest possible punishment to the agent when the intermediary can commit to disclosure plans.

Formally, for all $\mathbf{a}, \hat{\mathbf{a}} \in \mathbf{A}$, $\boldsymbol{\chi} \in \mathbf{X}$ and $\pi_t \in (0,1)$, let $\Pi^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t, \mathbf{a})$ be the agent's continuation payoff when he adopts effort plan $\mathbf{a}$ under disclosure rule $\boldsymbol{\chi}$, believed effort process $\hat{\mathbf{a}}$ and initial reputation π_t. Let

$$V^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t) \equiv \max_{\mathbf{a}\in\mathbf{A}} \Pi^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t, \mathbf{a})$$

be his continuation value. Consider the following program, which characterizes his lowest continuation value under the restriction that $\hat{\mathbf{a}}$ is optimal for the agent under $\hat{\mathbf{a}}$ and $\boldsymbol{\chi}$:

$$\mathcal{V}(\pi_t) \equiv \min_{(\hat{\mathbf{a}},\boldsymbol{\chi})\in\mathbf{A}\times\mathbf{X}} V^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t), \tag{2.17}$$

subject to:

$$V^{\hat{\mathbf{a}},\chi}(\pi_t) = \Pi^{\hat{\mathbf{a}},\chi}(\pi_t, \hat{\mathbf{a}}).$$

Lemma 2.3.8. *When patience level is high, $\mathcal{V}(\pi_t) = \pi_t \phi$ for all $\pi_t \in (0,1)$,*

As a direct implication, $V_a(\pi_t) > \overline{V}_a$ when $\pi_t > 1 - \frac{cr}{\mu\phi}$, which implies that $\pi^\dagger \leq 1 - \frac{cr}{\mu\phi}$. This result is not trivial since when $\chi(\pi_t) > 0$, higher effort level becomes incentive compatible. When the market anticipates this, π_t also declines faster conditional on no disclosure.[30] As a result, playing the best response to $\chi(\pi_t) = 0$ and $\hat{a}(\pi_t) = \phi$, which is $a(\pi_t) = \phi$, cannot guarantee him payoff $\pi_t \phi$.

2.3.5 Comparisons Between Exogenous and Endogenous Information

I compare the effort path in the unique MPE with the equilibrium effort path under the exogenous information benchmark to assess the impact of strategic information disclosure. First, effort is inverse U-shaped under endogenous information as opposed to monotone increasing under exogenous information. This is because effort needs to be low enough in order to provide the intermediary an incentive to disclose information. Second, the cutoffs at which effort jumps up, $\pi^\dagger$ and $\pi^\ddagger$, are also different.

The next result compares $\pi^\dagger$ with $\pi^\ddagger$. To see why this is interesting, notice that the two effort paths (as well as the paths of disclosure rates) coincide when $\pi_t > \max\{\pi^\dagger, \pi^\ddagger\}$. Moreover, effort is strictly lower under endogenous information when π_t is small enough. The remaining question is: whether strategic disclosure can lead to higher effort at some intermediate beliefs. This is only the case when $\pi^\dagger > \pi^\ddagger$. If so, the presence of a strategic intermediary motivates the agent to front-load effort.

Intuitively, one would expect that front-loading will always happen since χ_t is decreasing over time, that is, the intermediary withholds information at a higher rate at more pessimistic beliefs. This reduces the agent's continuation value at optimistic beliefs. Anticipating this, he will work harder early on since if he shirks today, his performance will be less visible tomorrow.

Unfortunately, the above logic is flawed, since it ignores the impact of equilibrium disclosure rate on the speed of market learning, which affects the dynamics of wages. To see this, when c is very close to 0, χ_t is also very close to 0 even for fairly high π_t, so is $\mu\chi_t a_t$, the rate of market learning. If this is the case, the market will attribute the absence of news to the intermediary's low disclosure rate, instead of the agent's incompetence. As a result, the agent's flow payoff, $a_t\pi_t - c(a_t - \phi)$, will remain high for a long time. In contrast, π_t deteriorates much faster when information is exogenous. So the agent's short run payoff is higher under endogenous information albeit his long run payoff is lower. Since $r > 0$, for c small enough, there exists an interior belief such that the agent's continuation value is higher under endogenous information when π_t exceeds this belief and vice versa. In this case, having a strategic intermediary censoring information can exacerbate the

[30] This effect is also discussed in Cisternas (2015) using a career concern model, in which the agent has more incentive to exert higher effort when the market anticipates higher effort, albeit his equilibrium payoff can be lower.

procrastination problem. This effect is more pronounced when r and μ are high, c and ϕ are low. Proposition 2.3 provides a sufficient condition:

Proposition 2.3. *For every r and ϕ, there exist $\bar{c} > 0$ and $\underline{\mu} > 0$ satisfying*

$$\phi - \bar{c}\phi - \frac{r\bar{c}}{\underline{\mu}} > 0$$

such that for every $c < \bar{c}$ and $\mu > \underline{\mu}$, there exists an open subset $B \subset (\phi, 1)$, such that $\pi^{\ddagger} > \pi^{\dagger}$ when $b \in B$.

To summarize, the unique MPE exhibits two sources of inefficiencies. First, the agent's effort needs to be low enough to encourage the intermediary to disclose information, and the latter is necessary to sustain reputation building incentives. Second, withholding information when $\pi_t < \pi^{\dagger}$ does not necessarily lower the agent's continuation value: when c is low and withholding information happens too early, it is actually encouraging the agent to procrastinate more.

Remark: It is worth clarifying that Proposition 2.3 is about the incentives to front-load effort instead of welfare. Welfare comparisons are hard to obtain due to the inconvenience of expressing the agent's equilibrium continuation value and the intermediary's equilibrium disclosure rate explicitly, which makes computing social surplus not tractable. Even when $\pi^{\dagger} > \pi^{\ddagger}$, despite effort is front-loaded, there is no guarantee that endogenous disclosure can improve social welfare. This is because first, welfare depends on the prior belief π_0 at which to evaluate payoffs; and second, welfare depends not only on the timing of effort, but also on its entire dynamics.

2.4 Semi-Markov Equilibria

In this section, I show that allowing players' strategies to condition on additional payoff irrelevant state variables can mitigate the inefficiencies of the unique MPE. In particular, I focus on a solution concept that minimally departs from MPE, Semi-Markov Equilibrium (SME, see Definition 2.3), which allows the intermediary and the agent to coordinate on payoff irrelevant variables off the equilibrium path.

I characterize the set of SME outcomes, which turns out to be tractable and contains both the unique MPE and the exogenous information equilibrium as special cases. To characterize this set, I construct a subclass of SMEs that covers the entire set of SME outcomes.

Proposition 2.4. *For every $\pi^{\S} \in [0, \min\{1-b, \pi^{\dagger}\}]$, there exist π^* and $\pi^{**} \in [\pi^{\S}, 1)$, such that the following 'three phase strategy profile' forms an SME.*

- ***Phase I:*** *If $\pi_t \leq \pi^{\S}$, then $a(\pi_t) = a^*(\pi_t)$ and $\chi(\pi_t)$ is chosen to make the agent indifferent.*
- ***Phase II:*** *If $\pi_t > \pi^{\S}$ and the intermediary has never concealed a breakthrough in the past,*

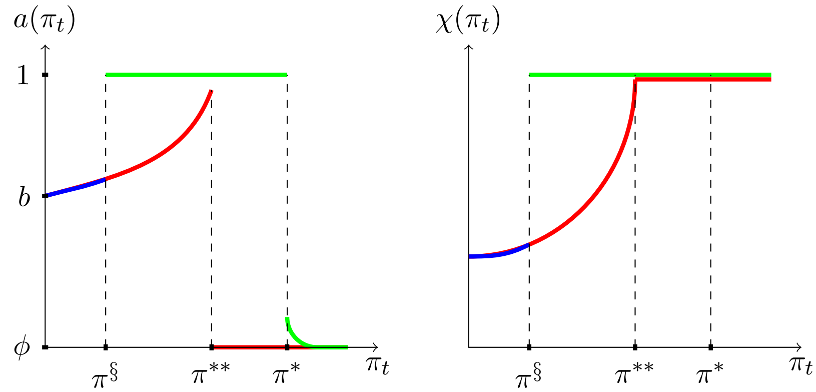

Figure 2-5: Effort and disclosure rate in an SME (in three phase strategy profile) with cutoff $\pi^{\S}$, with Phase I in blue, Phase II in green and Phase III in red.

then $\chi(\pi_t) = 1$ *and*

$$a(\pi_t) \equiv \begin{cases} 1 & \textit{when } \pi_t \in (\pi^{\S}, \pi^*] \\ a^{**}(\pi_t) & \textit{when } \pi_t > \pi^*. \end{cases}$$

- ***Phase III:*** *If* $\pi_t > \pi^{\S}$ *and the intermediary has concealed a breakthrough in the past, then*

$$a(\pi_t) \equiv \begin{cases} a^*(\pi_t) & \textit{when } \pi_t \in (\pi^{\S}, \pi^{**}] \\ \phi & \textit{when } \pi_t > \pi^{**}. \end{cases}$$

$\chi_t = 1$ *when* $\pi_t > \min\{1-b, \pi^{**}\}$ *and* χ_t *is chosen to make the agent indifferent when* $\pi_t \leq \min\{1-b, \pi^{**}\}$.

Moreover, every SME is outcome equivalent to an SME described above.[31]

Figure 2-5 depicts the effort and disclosure dynamics in a generic three phase strategy profile. For some intuition, SME allows players' strategies to condition on another state variable, that is, whether the intermediary has concealed breakthroughs in the past or not. Every SME is characterized by a cutoff belief, $\pi^{\S}$, below which withholding information happens on the equilibrium path. When $\pi_t \leq \pi^{\S}$ (Phase I), the equilibrium play resembles that in the unique MPE in which effort and disclosure rate are chosen to make both players indifferent. When $\pi_t > \pi^{\S}$ and the intermediary has never deviated before (Phase II), the agent exerts high effort and the intermediary fully discloses information. Suppose the intermediary has deviated before (Phase III), players coordinate on the low effort low disclosure rate equilibrium. π^* and π^{**} are chosen such that the agent's on-path continuation value at π^* and off-path continuation value at π^{**} are both $\overline{V}_a$. A formal characterization of the pair (π^*, π^{**}) is presented in Appendix B.3, along with the proof of Proposition 2.4.

Mapping back into the applications to law and consulting industries, SMEs capture the following effect which is absent in the unique MPE: the agent becomes frustrated if his past successes have

[31]The set of three phase strategy profiles characterize all possible on-path behaviors in SMEs. However, there exist other SMEs which differ in terms of their off-path play.

been ignored and in response to that, his future effort will decrease. Suppressing information today leads to lower effort in the future. This is reflected in the equilibrium behavior that effort jumps downward once a breakthrough is concealed.[32]

In what follows, I discuss the properties of SMEs. First, notice that the unique MPE is the SME with the highest $\pi^{\S}$, i.e. $\pi^{\S} = \min\{1-b, \pi^{\dagger}\}$, while the unique equilibrium in the exogenous information benchmark is outcome equivalent to the SME with $\pi^{\S} = 0$. SME allows $\pi^{\S}$ to take any value between 0 and $\min\{1-b, \pi^{\dagger}\}$, which enables players to control the timing at which withholding information starts and implies an efficiency gain when the optimal cutoff does not coincide with the unique MPE cutoff. In particular, when c is low, players can coordinate on a low enough $\pi^{\S}$ so that the market learns rapidly when π_t is high. This circumvents the problem identified in Proposition 2.3 since the agent cannot sustain high flow payoff at any interior belief. Moreover, SME boosts effort to 1 when $\pi_t \in (\pi^{\S}, \min\{1-b, \pi^{\dagger}\}]$. This arrangement is incentive compatible since the intermediary has no decision to make before the first breakthrough arrives, so increasing the agent's effort before that does not upset her incentive constraints. Some other properties of SME are summarized below.

Front-loading Effort: Since the unique MPE and the equilibrium under exogenous information are both special cases of SME, the optimal SME must (weakly) outperform both.

Corollary 2.1. *There exists $\varepsilon > 0$ such that for all $\pi^{\S} \in (0, \varepsilon)$, $\pi^*(\pi^{\S}) > \pi^{\ddagger}$.*

Corollary 2.1 says that SME with $\pi^{\S}$ small enough strictly outperforms the unique equilibrium in the exogenous information benchmark in terms of front-loading effort, i.e. the agent starts working harder at a higher belief.

Simple Belief Updating Rule: SMEs have the attractive property that the market does not need to compute the *probability that a breakthrough has been withheld in the past* in order to formulate its posterior belief about θ. This is because in every SME, the market always correctly anticipates the agent's on-path effort and the intermediary's on-path disclosure rate.[33]

Robustness: Every SME induced by a three phase strategy profile remains robust when the intermediary can disclose past breakthroughs. This is because her incentive to withhold information is weakly increasing over time, both on and off the equilibrium path. However, SMEs (aside from the unique MPE) are not robust to private monitoring, since they rely on the arrival times of breakthroughs being common knowledge between the intermediary and the agent.

[32] To comment more on the jump, effort jumps from 1 to $a^*(\pi_t)$ when $\pi_t \in (\pi^{\S}, \pi^{**}]$; from 1 to ϕ when $\pi_t \in (\pi^{**}, \pi^*]$; from $a^{**}(\pi_t)$ to ϕ when $\pi_t \in (\pi^*, \overline{\pi}]$.

[33] Nevertheless, the set of PBEs in which players' strategies can condition on π_t as well as whether breakthrough has been withheld in the past or not is much larger than the set of SMEs, since in general, future effort and disclosure rate can differ whenever the intermediary has withheld a breakthrough in the past, regardless of whether play has gone off-path or not. However, complicated market belief updating and wage formulas will occur, which are not tractable to analyze.

2.5 Optimal Markov Policy with Commitment

In this section, I examine the optimal Markov policy when the intermediary can commit to dynamic disclosure plans: the intermediary designs a Markov policy, $(\mathbf{a}, \boldsymbol{\chi}) \in \mathbf{A} \times \mathbf{X}$, to maximize her expected payoff evaluated at belief π_0, subject to the market's and the agent's incentive constraints.[34]

Recall the definitions of $V^{\mathbf{a},\boldsymbol{\chi}}(\pi)$ and $\Pi^{\mathbf{a},\boldsymbol{\chi}}(\pi, \mathbf{a})$ in Subsection 2.3.4. The intermediary's maximization problem is:

$$\sup_{(\mathbf{a},\boldsymbol{\chi})\in\mathbf{A}\times\mathbf{X}} \Big\{ \pi_0 \underbrace{\left(r\int_0^\infty e^{-rt-(y_t-y_0)}(a_t - w_t - b)dt + b \right)}_{\text{gain from high type}} - (1-\pi_0) \underbrace{r\int_0^\infty e^{-rt} w_t dt}_{\text{loss from low type}} \Big\}, \tag{2.18}$$

subject to:

$$\underbrace{V^{\mathbf{a},\boldsymbol{\chi}}(\pi_0) = \Pi^{\mathbf{a},\boldsymbol{\chi}}(\pi_0, \mathbf{a})}_{\text{Effort plan } \mathbf{a} \text{ is optimal for the agent}}, \quad \dot{\pi}_t = -\pi_t(1-\pi_t)\mu a(\pi_t)\chi(\pi_t) \text{ and } w(\pi_t) = \pi_t a(\pi_t),$$

with $y_t \equiv \ln \frac{1-\pi_t}{\pi_t}$. Re-write the intermediary's objective function as:

$$r(1-\pi_0)\int_0^\infty e^{-rt-y_t}(a_t - w_t - b)dt - r(1-\pi_0)\int_0^\infty e^{-rt} w_t dt + b\pi_0. \tag{2.19}$$

Plugging in the expression for w_t and ignoring positive affine transformations, the intermediary's problem is to minimize:

$$\int_0^\infty e^{-rt}\frac{\pi_t}{1-\pi_t}dt. \tag{2.20}$$

subject to the agent's and the market's incentive constraints. The optimal policy is characterized below:

Proposition 2.5. *An optimal Markov policy exists. For every $\pi_0 \in (0,1)$, there exists $0 \le \pi' < \pi'' \le \pi_0$ such that the optimal Markov policy with commitment has three phases:*

- ***Shirking Phase:*** *If $\pi_t > \pi''$, then $\chi(\pi_t) = 1$ and $a(\pi_t) = \phi$.*
- ***Working Phase:*** *If $\pi_t \in (\pi', \pi'']$, then $a(\pi_t) = \chi(\pi_t) = 1$.*
- ***Deadline:*** *If $\pi_t \le \pi'$, then $\chi(\pi_t) = 0$ and $a(\pi_t) = \phi$.*

Moreover, $\pi' > 0$ if and only if $\pi_0 > \pi^\ddagger$.

The proof uses standard optimal control techniques, which is relegated to Online Appendix C. As long as π_0 is sufficiently large, the shirking phase is not degenerate ($\pi'' < \pi_0$) and the deadline is not trivial ($\pi' > 0$). The agent's effort is still inverse U-shaped, with no effort when the market's belief is extremely high or low and maximal effort when the market's belief is intermediate.

[34] It is worth emphasizing that there is a loss of generality when focusing on Markov Policies. This is because potentially, effort and disclosure rate can also depend on players' private histories as well as calendar time, on top of the public belief.

Comparing with the unique MPE or the SMEs, the optimal commitment solution has three interesting features. First, the intermediary commits to fully disclose information when π_t is high. Second, the agent's effort is always bang-bang. Third, when the prior π_0 is high enough, she commits to a deadline. This reduces the agent's continuation value and makes $a_t = 1$ incentive compatible at more optimistic beliefs.[35]

These features contrast sharply with the equilibria without commitment. Aside from the exogenous information equilibrium ($\pi^{\S} = 0$), information disclosure rate and effort must both be interior when π_t is low enough in every other SME. Moreover, the third feature (deadline) is not replicable in any *PBE* in absence of commitment, as formally stated in the following Corollary:

Corollary 2.2. *In every PBE, there exists no $h_m^t \in H_m$, such that $\chi(h_m^{t'}) = 0$ for all $h_m^{t'} \succeq h_m^t$.*

Intuitively, this is because the agent knows the intermediary's private history, and he will shirk forever after reaching h_m^t. Anticipating this, the intermediary is tempted to disclose information at h_m^t. Lack of commitment unravels the deadline arrangement, making it harder to motivate the agent when π_t is high.

2.6 Discussions & Extensions

In this section, I enrich and extend the baseline model in several directions and examine the robustness of my results. In Subsection 2.6.1, I discuss the properties of Markov Equilibria when the intermediary can disclose past breakthroughs. In Subsection 2.6.2, I allow the market to learn from a public signal in addition to the intermediary's private signal. In Subsection 2.6.3, I discuss several variants of the model, by investigating different preferences of the intermediary and the agent, alternative informational assumptions, etc.

2.6.1 Disclosing Past Breakthroughs

In this subsection, I allow the intermediary to disclose past breakthroughs. Despite every SME (including the unique MPE) remains robust to disclosing past breakthroughs, enriching the intermediary's strategy space opens up new equilibrium possibilities. This is true even for Markov solution concepts, since the possibility of disclosing past breakthroughs changes the set of payoff relevant state variables.

Formally, when disclosing past breakthroughs is allowed, I adopt the '*multi-stage game*' formulation in Murto and Välimäki (2013). In what follows, I will proceed at an intuitive level and will formally define the game in Online Appendix D.2. Let $x_t \in \{0, 1\}$ be defined as:

$$x_t \equiv \begin{cases} 1 & \text{when } h_m^t \neq \{\varnothing\} \\ 0 & \text{when } h_m^t = \{\varnothing\}, \end{cases}$$

[35] As shown in Lemma 2.3.8, deadline is the harshest possible punishment.

which determines the intermediary's capability of disclosing information. Once $x_t = 1$, whether the breakthrough arrives at t or before is payoff irrelevant. The market's belief about θ as well as x_t are both deterministic functions of t. Hence, t and x_t are the only payoff relevant state variables.

The solution concept is weak Markov Perfect Equilibrium (or wMPE), in which at time t, the agent's effort depends only on t and x_{t-}, but the intermediary's disclosure rate depends not only on t and x_{t-}, but also on $\xi_t \equiv \mathbf{1}\{t \in h_m^t\}$, i.e. whether a signal has arrived at t or not.[36] By definition, every SME constructed in Section 2.4 (therefore, also the unique MPE) is a wMPE. Despite the set of wMPEs is much larger than the set of SMEs and is not tractable to analyze, the next Lemma identifies some common properties of wMPEs:

Lemma 2.6.1. *In every wMPE, there exists $\pi^\P \in [0, 1)$, such that:*

- *The intermediary always fully discloses the first breakthrough when $\pi_t > \pi^\P$.*
- *The agent's effort is $a^*(\pi_t)$ after the first breakthrough is concealed when $\pi_t \leq \pi^\P$.*

There are multiple equilibria even under a given $\pi^\P$ since effort and disclosure rate are not uniquely pinned down when $\pi_t \leq \pi^\P$ but before a breakthrough has been concealed.

Although the possibility of disclosing past breakthroughs generates additional equilibria, in which the rate of learning, the cutoff beliefs and the effort paths are *quantitatively* different from the SMEs in the baseline model, two *qualitative* features remain robust. First, the intermediary only withholds information when π_t is low. Second, if π_t falls below the cutoff and a breakthrough has been concealed, the agent's effort is decreasing over time and coincides with $a^*(\pi_t)$ in *every* wMPE.

2.6.2 Learning from Public Signals

In this subsection, the market can also learn from public signals, arrive according to Poisson rate $\mu_0 \theta a_t$, in addition to the private signals disclosed by the intermediary, which arrive at Poisson rate $\mu_1 \theta a_t$, with parameters $\mu_0 \geq 0, \mu_1 > 0$. The market automatically observes the public signal, but can only observe the private signal after the intermediary discloses it. Let $\mu \equiv \mu_0 + \mu_1$ be the '*net arrival rate*', I introduce the following condition, which measures the public signal arrival rate.

Condition 2.2. *Public signal arrival rate is low if $\mu_0 \leq \frac{cr}{\phi(1-c)}$.*

Otherwise, we say that public signal arrival rate is high. Intuitively, the intermediary has less control over the market's information if public signal arrival rate is high and vice versa. The patience level condition is re-defined as follows:

Condition 2.3. *Patience level is high if $r < \frac{\mu\phi}{c}(1-c)$. Patience level is low if $r \geq \frac{\mu\phi}{c}(1-c)$.*

[36] If we insist on more restrictive solution concepts, for example, by requiring that the agent's effort to depend only on t and x_{t-} and the intermediary's disclosure rate to depend only on t and x_t, only a trivial equilibrium exists, in which the intermediary always discloses information. I will state and show this result in the Online Appendix.

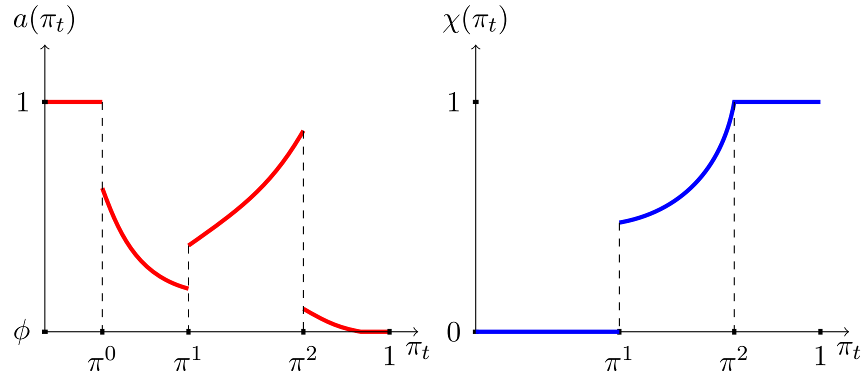

Figure 2-6: Effort and Disclosure Rate when Public Signals Arrive Frequently

Since $\mu > \mu_0$, patience level must be high when public signal arrival rate is high. I start with the low arrival rate case, in which the unique MPE and the set of SME outcomes coincide with those in the baseline model.

Lemma 2.6.2. *When* $\mu_0 \leq \frac{cr}{\phi(1-c)}$,

- *The unique MPE with public signals is outcome equivalent to the unique MPE in the baseline model.*
- *Every SME with public signals is outcome equivalent to an SME in the baseline model.*

Intuitively, fixing the net arrival rate μ and given that μ_0 is small enough, the exact decomposition between public and private signals has no impact on the equilibrium outcome since the net disclosure rate $\mu_0 + \mu_1\chi_t$ is unchanged: we can always find $\chi_t \in [0, 1]$ to match the equilibrium net disclosure rate in the baseline model ($\mu_0 = 0$). In this case, public and private signals are perfect substitutes.

Once μ_0 exceeds $\frac{cr}{\phi(1-c)}$, there exists no χ_t that can match the net disclosure rate when π_t is small. Due to frequent arrival of public signals, the agent has a strict incentive to exert effort when π_t is low even if the intermediary ceases to disclose private signals. As a result, $\chi_t \to 0$ and $a_t \to 1$ when π_t is close to 0. Moreover, public and private signals are no longer perfect substitutes: an increase in μ_0 increases the agent's continuation value at some beliefs, which leads to an increase in the net disclosure rate required to motivate the agent, so the disclosure rate of private signal can increase with μ_0 through this indirect intertemporal effect. The equilibrium effort path and disclosure rate are shown in Figure 2-6.[37]

2.6.3 Other Extensions & Discussions

In this subsection, I discuss several alternative specifications of the intermediary's and the agent's payoffs, the robustness of my result to alternative informational assumptions as well as how to enrich the baseline model to account for more realistic features.

[37]The proofs of these results are available upon request.

The Intermediary's Time Preference: The effort and disclosure dynamics in every SME has nothing to do with the intermediary's discount factor. As a result, it remains robust when the intermediary faces a different discount rate. The unique MPE is robust even when the intermediary is finitely lived. This is because whenever she is supposed to partially disclose information, her flow payoff is b from then on; whenever she is supposed to fully disclose information, her flow payoff is weakly below b at every instant.[38]

The Intermediary's Disclosure Benefit is Low: When $b \in (0, \phi]$,[39] withholding information is the intermediary's dominant strategy when π_t is sufficiently low. The agent ceases to exert effort and market learning stops when π_t falls below the following cutoff: $\pi^* \equiv 1 - \frac{b}{\phi}$, at which the intermediary's continuation payoff is b and the agent's continuation value is $\pi_t \phi$. However, in the unique MPE, effort is still inverse U-shaped and the disclosure rate is still decreasing over time (both reaching 0 when $\pi_t \leq \pi^*$), which are qualitatively similar to the baseline model.

The Intermediary as a Supervisor: In academia and professional sports, the intermediary who has private information about the junior agent's performance is usually his direct supervisor, instead of his current employer. Different from employers, although supervisors gain private benefits from extracting her agent's effort and can obtain a network benefit from establishing her agent in front of the public, they do not pay their agent's wages. As a result, their flow payoff when $\pi_t < 1$ is θa_t, instead of $\theta a_t - w_t$.

When patience level is high, the game still admits a unique MPE, in which the effort and disclosure rate dynamics are shown in Figure 2-7, with the red dashed line being the agent's value invariance curve (same as the baseline model) and the blue dashed line being the intermediary's indifference curve. Similar to the baseline model, the intermediary's disclosure rate is decreasing over time and the agent's effort coincides with the value invariance curve when $\pi_t > \pi^\dagger$ and coincides with the intermediary's indifference curve when $\pi_t \leq \pi^\dagger$. However, since the intermediary only cares about the agent's effort, her indifference curve is flat. As a result, the agent's effort remains unchanged when belief is low.

Convex Effort Cost: When the agent's effort cost is convex instead of linear, the baseline model still admits a unique MPE. Several features of the unique MPE under linear cost remain robust, including inverse U-shaped effort and decreasing disclosure rate. However, there will be no '*jump*' in effort under convex cost, which is the main difference from the linear cost model.[40]

Comments on Informational Assumptions: The qualitative features of the effort and disclosure dynamics identified in my baseline model are robust to other variants of informational

[38] All SMEs in three phase strategy profiles constructed in Section 2.4 remain robust to finitely lived intermediaries if the next intermediary inherits all her predecessors' information.

[39] Cases in which $b \geq 1$ and $b < 0$ are trivial, since the intermediary always has a strict incentive to disclose or to withhold information

[40] However, there will be kinks in the equilibrium effort path.

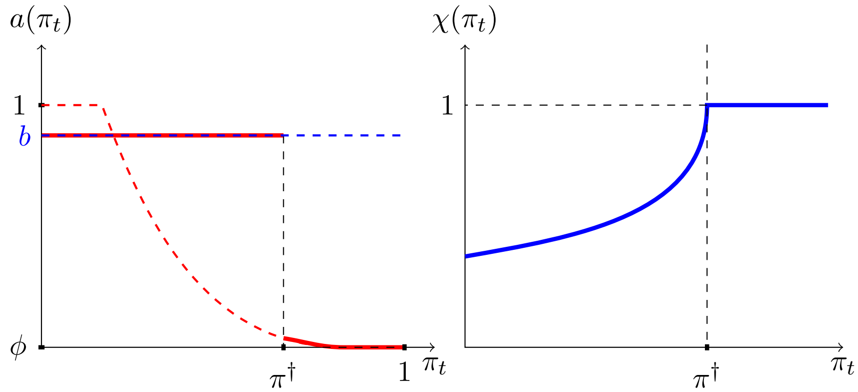

Figure 2-7: Effort and Disclosure Rate when the Intermediary is the Agent's Supervisor

assumptions. For example, even if the intermediary can observe effort or the agent cannot observe the breakthrough, the unique MPE remains robust. When the agent does not know his type, as in career concern models (for example, Holmström 1999), the agent's equilibrium effort can condition not only on the market's belief, but also on his private belief as well as the market's belief about his private belief, all of these become payoff relevant.[41] Despite characterizing the set of MPEs is a formidable task, but nonetheless, the agent's effort will still be inverse U-shaped. To see this, if the agent knows his type, his effort path will be the same as in the '*reputation concern case*'. If he does not know his type, his effort is also inverse U-shaped, as shown in Bonatti and Hörner (2015). This is because the agent will become more pessimistic about his ability as no breakthrough has arrived over time. Due to the complementarity between effort and ability, he has less incentive to exert effort. To summarize, we would still anticipate inverse U-shaped effort and decreasing disclosure rate.

Multiple Job Levels: My model can be enriched to account for the fact that lots of '*promising future stars*' (especially in academia and professional sports) work hard despite having favourable public beliefs. To see this, suppose there are three job levels: $1, 2$ and 3, and three types: high, medium and low. The high type and the medium type can produce breakthroughs in level 1 while only the high type can produce breakthroughs in level 2. After a breakthrough in level k is disclosed, the intermediary in level k receives a lump sum payoff and the agent is promoted to level $k+1$, working for a new intermediary.

In this variation of my model, high type agents have stronger incentives to work hard in level 1 for two reasons.[42] First, he can be further promoted after reaching level 2, thus his continuation value of entering level 2 is higher than the medium type. Second, being promoted to level 2 at an earlier date distinguishes himself from the medium type, leading to a higher market belief and increases the intermediary's information disclosure rate at level 2. As a result, high type agents

[41] The intermediary's private belief is trivial, since she can only disclose information after her private belief is 1.

[42] As in the baseline model, the agent will eventually stop working hard once reaching the final level, i.e. level 3.

work harder in level 1. In terms of predictions, those who have succeeded earlier in level 1 are more likely to be promoted sooner in level 2, which is the well-known '*fast track*' phenomena documented in the personnel economics literature. Moreover, '*fast tracks*' are more salient when information is disclosed by a strategic intermediary, due to the decreasing disclosure rate over time.

2.7 Conclusion

This paper studies the impact of endogenous information disclosure by a strategic intermediary on an agent's incentive to build up his reputation. In the unique MPE, the agent's effort is inverse U-shaped and the information disclosure rate is decreasing over time. Surprisingly, the agent's continuation value can be higher and can have more incentives to procrastinate in the unique MPE comparing with the exogenous information benchmark. This is because withholding information also slows down the rate of market learning, which allows the agent to enjoy high flow payoff for a long time, even without producing any breakthroughs. Relaxing the Markov restriction and allowing players to coordinate on payoff irrelevant events such as whether breakthroughs have been withheld in the past or not, can mitigate this inefficiency. This is because it enables players to flexibly choose the cutoff belief, below which the intermediary starts to withhold information.

Chapter 3

Repeated Interactions without Commitment

This chapter studies repeated games in which a patient long-run player (e.g. a firm) wishes to win the trust of some myopic opponents (e.g. a sequence or a continuum of consumers) but has a strict incentive to betray them. Her benefit from betrayal is persistent over time and is her private information. I examine the extent to which persistent private information can overcome this lack-of-commitment problem. My main result characterizes the set of payoffs a patient long-run player can attain in equilibrium. Interestingly, every type's highest equilibrium payoff only depends on her true benefit from betrayal and the *lowest* possible benefit in the support of her opponents' prior belief. When this lowest possible benefit vanishes, every type can approximately attain her Stackelberg commitment payoff. My finding provides a strategic foundation for the (mixed) Stackelberg commitment types in the reputation models, both in terms of the highest attainable payoff and in terms of the commitment behaviors. Compared to the existing approaches that rely on the existence of *crazy types* that are either irrational or have drastically different preferences, there is common knowledge of rationality in my model, and moreover, players' ordinal preferences over stage game outcomes are common knowledge.

3.1 Introduction

Imagine a politician running for president pledging for massive tax cuts. Once elected, he might be tempted to breach his promise due to the growth in mandatory spending and rising budget deficits. Anticipating such possibilities of future betrayal, should the electorate vote for this candidate in the first place?[1] Alternatively, firms would like to convince consumers about their high quality

I thank Daron Acemoglu, Mehmet Ekmekci, Drew Fudenberg, Jean Tirole, Juuso Toikka and Alex Wolitzky for helpful conversations. All errors are mine.

[1]A classic example is ex-president George H.W. Bush's 1988 acceptance speech at the New Orleans convention "*Read my lips, no new taxes.*" But after becoming president, he agreed to increase several existing taxes in order to reach a compromise with the Democrat-controlled Congress. Breaching this promise has hurt Bush politically during his 1992 campaign, as both Pat Buchanan and Bill Clinton cited his quotation and questioned his trustworthiness.

standards. But after receiving the upfront payments, they are tempted to undercut quality, especially on aspects that are hard to verify. Similar plights happen to central banks when showing their resolve to fight hyperinflation and to entrepreneurs when persuading venture capitalists to fund their projects.

An important feature shared by these examples is a *lack-of-commitment problem* faced by the politicians, firms, central banks, entrepreneurs, etc. This problem is prevalent in many economic activities and has significant impact on social welfare. For example, it can block profitable trade, cause capital mis-allocation and limit the effectiveness of public policies. As a result, exploring ways to lend credibility to promises and threats has become a central question in non-cooperative game theory. One prominent idea dates back to Thomas Schelling (1960), who argues that an individual's problem of commitment can be solved by a so-called *reputational equilibrium*, in which she could *lose her favorable reputation by deviating from expected norms of behavior.* This intuition is formalized in a series of papers starting from Kreps and Wilson (1982), Milgrom and Roberts (1982), Fudenberg and Levine (1989, 1992) in which a player's behavior in the current stage is linked to her future benefits through others' perceptions about her *type*. Importantly, this reputation building player is *crazy* with strictly positive probability, in which case she is either irrational or having non-standard payoff functions.

This chapter revisits Thomas Schelling's classic argument by exploring the extent to which persistent private information can overcome these lack-of-commitment problems in repeated interactions. Compared to the aforementioned models with crazy types, I adopt a complementary approach by requiring all types of the reputation building player to be rational and to have *reasonable* stage game payoff functions. The way I interpret '*reasonableness*' in this paper is that all types of the reputation building player sharing the same ordinal preference over stage-game outcomes.[2] This has the advantage of maintaining the sensible assumptions on the game's payoffs (such as providing high quality is costly for the firm), which are likely to be common knowledge in reality. It can also evaluate whether the insights from the canonical reputation models rely on the presence of those crazy types.

To capture the key economic forces in the aforementioned applications, I study stage games with one-sided lack-of-commitment (or *trust games*), which is played repeatedly over the infinite horizon between a patient long-run player (e.g. a firm) and a sequence of myopic short-lived opponents (e.g. consumers).[3] In every period, the long-run player wishes to win her opponent's trust by promising high effort, but has a strict incentive to exert low effort and betray them once trust is granted. Her

[2]Nevertheless, the exact notion of reasonableness should depend on the application. The notion I adopt in this paper fits well into the leading application of my model, namely, interactions between firms and consumers. In that scenario, it is likely to be common knowledge that providing high quality is costly for the firm, but the fine details of its production cost tends to be the firm's private information.

[3]This assumption on myopia only affects the upper bound result on the long-run player's equilibrium payoff set, but the result on the attainability of Stackelberg commitment payoff (i.e. overcoming the lack-of-commitment problem) applies regardless of the uninformed players' discount factor(s). In the applications, the uninformed players are close to being myopic either because they are sufficiently small (such as citizens interacting with the government, a large firm serving a large pool of clients) or when they interact with the informed player only once (such as a sequence of consumers purchasing a durable good).

cost of high effort is persistent over time and is her private information. Every short-run player perfectly observes the entire sequence of past plays, and is willing to trust the long-run player if and only if he expects effort to be high with probability above some cutoff.

For some useful benchmarks, first, when the long-run player can commit, her payoff is maximized when committing to a mixed action that makes her opponents indifferent between trust and not trust. The resulting payoff is called her *Stackelberg commitment payoff.* Second, if the long-run player cannot commit and her cost of effort is common knowledge, then her highest payoff in the repeated complete information game is strictly bounded below her Stackelberg commitment payoff.[4] This is because her opponents' myopia imposes constraints on the action profiles that can be played in every period, and consequently, some feasible and individually rational payoffs are not attainable in any equilibrium regardless of the long-run player's discount factor. In another word, repeated interactions under complete information *cannot* fully solve the lack-of-commitment problem.

My main result (Theorem 3.1) characterizes the set of equilibrium payoffs the patient long-run player can attain in the repeated incomplete information game. I provide a tractable formula for every type's highest equilibrium payoff, which equals to the product of her (complete information) Stackelberg commitment payoff and an *incomplete information multiplier*. The latter is a decreasing function of the *lowest* possible cost in the support of her opponents' prior belief which is a sufficient statistic for the effect of incomplete information. When this lowest possible cost vanishes to 0, the multiplier converges to 1, or equivalently, every type of the long-run player can approximately attain her Stackelberg commitment payoff.

Theorem 3.1 has the following set of implications. First, it identifies the aspect of the type distribution that matters for a patient long-run player's payoff. According to my characterization, the equilibrium payoff set only depends on the lowest possible cost in the support. Intuitively, this is because the lowest cost type has no good candidate to imitate, so the uninformed players' myopia introduces an upper bound on her equilibrium payoff in the repeated incomplete information game, same as the one in the repeated complete information game where her type is common knowledge. This in turn leads to an upper bound for *every* other type's equilibrium payoff, the exact value of which only depends on the maximal frequency with which low effort can be played to induce the uninformed players' trust and has nothing to do with other details of the distribution. Intuitively, this is because exerting low effort as frequently as possible minimizes the disadvantage of the high-cost types relative to the lowest cost type, leading to higher equilibrium payoffs for the former.

Second, it implies that incomplete information can help the patient long-run player overcome her lack-of-commitment problem and (approximately) attain her Stackelberg commitment payoff. Different from the existing reputation results, I obtain this insight without relying on the presence of crazy types that are mechanically playing *mixed commitment strategies*, which are hard to rationalize via standard payoff functions. My approach addresses the concerns raised by Weinstein and Yildiz (2007, 2013, 2016) that when all the incomplete information perturbations are allowed, one can rationalize almost every outcome by constructing crazy types that have qualitatively dif-

[4]The highest payoff in the complete information repeated game is the long-run player's *pure Stackelberg commitment payoff*, which in generic games, is strictly less than her Stackelberg commitment payoff.

ferent payoff functions and well-calibrated belief hierarchies.[5] Their results call for a more careful selection of perturbations, and ideally, the perturbed types should also have reasonable payoffs and beliefs.[6] In context of trust games, I show that the Stackelberg commitment payoffs are attainable even when these aspects of the stage-game payoff are required to be common knowledge: (1) the long-run player values her opponents' trust; (2) she faces strictly positive cost to exert effort. Nevertheless, her exact cost of effort, which depends on the fine details of her production technology, talent, etc. is still assumed to be her private information.

Third, my result provides a partial strategic foundation for the crazy types that are playing (mixed) Stackelberg strategies, the presence of which is a key component of the canonical reputation models. Theorem 3.1 implies that in terms of attaining the Stackelberg commitment payoff, these types can be replaced by a strategic type that has the same ordinal preference over stage game outcomes but has a low cost to exert high effort.[7]

Along this line of inquiry, I then study the *behavior* of the lowest cost type in equilibria that attain the Stackelberg commitment payoff. This also evaluates which of the many Stackelberg commitment strategies can be rationalized by low-cost types. Somewhat surprisingly, I show that in every such equilibrium, any type that has sufficiently low cost *will not* play a (non-trivial) mixed action at every history. This implies that the *stationary* commitment strategy of playing the Stackelberg action at every history cannot be rationalized by low-cost types. To understand why, suppose towards a contradiction that such equilibria exist, then the highest cost type can exert low effort in every period and attain the same payoff as the lowest cost type. Applying the upper bound result in Fudenberg and Levine (1992) that every type's payoff cannot exceed her Stackelberg commitment payoff, one can obtain a contradiction when the lowest possible cost is sufficiently small relative to the highest possible cost.

The above finding also suggests that equilibrium behaviors of the low-cost types exhibit the following features: (1) along every action path played by the low-cost types, the discounted average frequency of high effort is no less than the probability of high effort in the Stackelberg commitment action; (2) with probability close to 1, the discounted average frequency of high effort is close to the probability of high effort in the Stackelberg commitment action. Intuitively, this is to say that the low-cost types *cherry-pick* their actions in order to avoid exerting low effort with frequency above

[5]Weinstein and Yildiz (2007, 2013, 2016) show in various settings (e.g. static games, infinitely repeated games, finitely repeated games) that every outcome within some large sets (e.g. interim correlated rationalizable outcomes, Bayes Nash Equilibrium payoffs, pure strategy non-stationary commitment behaviors, etc.) is uniquely rationalizable in some *nearby* games according to the product topology if *all* incomplete information perturbations are allowed.

[6]In this spirit, I adopt an alternative interpretation of the well-known critique in Wilson (1987) that:

> *Game theory has a great advantage in explicitly analyzing the consequences of trading rules that presumably are really common knowledge; it is deficient to the extent that it assumes other features to be common knowledge, such as one agent's assessment about another's preferences or information.*

Namely, one should relax the stringent informational assumptions that are poor descriptions of the real world while maintaining the common knowledge assumptions on aspects that are likely to be common knowledge in reality.

[7]This is only a *partial* foundation as there are equilibria in which the long-run player's payoff is strictly bounded below her Stackelberg commitment payoff. In reputation models with crazy types that are mechanically playing Stackelberg actions, the long-run player can approximately attain her Stackelberg commitment payoff in every Nash equilibrium.

some cutoff. Achieving this objective requires their strategies to be non-stationary and to exhibit non-trivial history dependence.

The rest of the chapter is organized as follows. I will setup the baseline model and will analyze several useful benchmarks in section 3.2. I will state the main result and outline its key economic implications in section 3.3. I will sketch a proof of the theorem in section 3.4 and review several related ideas in the existing literature. Section 3.5 concludes and examines some natural extensions of the baseline model.

3.2 The Baseline Model

In this section, I introduce a repeated *trust game* which can capture the lack-of-commitment problem in many socioeconomic interactions. Different from the existing approaches that introduce *crazy types* to show that incomplete information can lend credibility to the long-run player's promises, my model features common knowledge of rationality and moreover, there is no type that has a drastically different stage game payoff in the sense that players' ordinal preferences over stage game outcomes are common knowledge.

The Stage Game: In every period, a firm (player 1, she) interacts with a client (or player 2, he). The client moves first, deciding whether to purchase a product from the firm (i.e. *trusting the firm*, taking action T) or not (i.e. *not trusting the firm*, taking action N). If he takes action N, then both players' stage game payoffs are 0 and play moves on to the next period. If he takes action T, then the firm chooses between high effort (action H) and low effort (action L). If the firm chooses L, then her game payoff is normalized to 1 and the client's payoff is $-c$; if the firm chooses H, then her stage game payoff is $1-\theta$ and the client's payoff is b, where:

- $c > 0$ is the client's loss from the firm's betrayal;
- $b > 0$ is the client's net benefit from the firm's high effort;
- $\theta \in \Theta \equiv \{\theta_1, ...\theta_m\} \subset (0,1)$ is the firm's cost of exerting high effort, or under more general interpretations, the long-run player's temptation to betray or her cost to honor promises. Without loss of generality, I assume that $0 < \theta_1 < \theta_2 < ... < \theta_m < 1$.

The benefit and cost parameters, b and c, are common knowledge. The cost of exerting high effort is the firm's private information.[8] The firm is of *type* θ_i if she knew that $\theta = \theta_i$.

The Repeated Game: Consider an infinitely repeated version of the above stage game. Time is discrete, indexed by $t = 0, 1, 2, ...$ The firm is interacting with an infinite sequence of clients, arriving one in each period and plays the game only once. In period t, players choose their actions

[8]This assumption is relevant in many industries as firms usually have private information about their production technologies, see for example, the seminal work of Baron and Myerson (1982).

according to the pre-specified order. They also have access to a public randomization device, with $\xi_t \in [0,1]$ a typical realization.

The firm's cost of high effort, θ, is her private information and is perfectly persistent over time. The client's prior belief over θ is $\pi_0 \in \Delta(\Theta)$, which is assumed to have full support. Both players' actions choices in the past can be perfectly monitored. Let $a_t \in \{N, H, L\}$ be outcome (or *the realized terminal node*) in period t. Let $h^t = \{a_s, \xi_s\}_{s=0}^{t-1} \in \mathcal{H}^t$ be the *public history* in period t with $\mathcal{H} \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}^t$ the set of public histories.

Let $\sigma_2 : \mathcal{H} \rightarrow \Delta\{T, N\}$ be the client's strategy, which maps the set of public histories to his (mixed) actions. Let $\sigma_\theta : \mathcal{H} \rightarrow \Delta\{H, L\}$ be type θ firm's strategy, which specifies her (mixed) actions after receiving her client's trust conditional on the public history. Let $\sigma_1 \equiv (\sigma_\theta)_{\theta \in \Theta}$ be the firm's strategy.

The firm discounts future payoffs by factor $\delta \in (0,1)$. Type θ_i firm maximizes her (expected) discounted average payoff, given by:

$$\mathbb{E}^{(\sigma_\theta, \sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t u_1(\theta_i, a_t)\Big], \tag{3.1}$$

with $\mathbb{E}^{(\sigma_\theta, \sigma_2)}[\cdot]$ the expectation over $\mathcal{H}$ under the probability measure induced by $(\sigma_\theta, \sigma_2)$.

Equilibrium Payoffs: To make the results more convincing, I study the firm's equilibrium payoffs under two solution concepts: Nash equilibrium and sequential equilibrium (Kreps and Wilson 1982), and will later show in Theorem 3.1 that the resulting payoff sets coincide. This implies that my characterization provides a consistent description of a patient firm's equilibrium returns, in the sense that it is not sensitive to the choice of solution concepts.

Formally, let $v \in \mathbb{R}^m$ be a generic payoff vector for the firm, with the ith coordinate being the discounted average payoff of the type θ_i. Let $\underline{V}(\pi_0, \delta)$ and $\overline{V}(\pi_0, \delta) \subset \mathbb{R}^m$ be the firm's sequential equilibrium and Nash equilibrium payoff sets under parameter configuration $(\pi_0, \delta) \in \Delta(\Theta) \times (0,1)$, respectively. Let

$$\underline{V}(\pi_0) \equiv \text{clo}\Big(\liminf_{\delta \rightarrow 1} \underline{V}(\pi_0, \delta)\Big) \tag{3.2}$$

and

$$\overline{V}(\pi_0) \equiv \text{clo}\Big(\limsup_{\delta \rightarrow 1} \overline{V}(\pi_0, \delta)\Big), \tag{3.3}$$

with $\text{clo}(\cdot)$ being the closure of a set. By definition, $\underline{V}(\pi_0, \delta) \subset \overline{V}(\pi_0, \delta)$ for every (π_0, δ), which implies that $\underline{V}(\pi_0) \subset \overline{V}(\pi_0)$ for every π_0. Let $\underline{V}_i(\pi_0)$ and $\overline{V}_i(\pi_0)$ be the projections of $\underline{V}(\pi_0)$ and $\overline{V}(\pi_0)$ on the i-th coordinate. By definition, when the firm is arbitrarily patient, her highest payoff under Nash equilibrium when her effort cost is θ_i is $\max \overline{V}_i(\pi_0)$, her highest payoff under sequential equilibrium is $\max \underline{V}_i(\pi_0)$.

3.2.1 Two Benchmarks

I start from two benchmark scenarios: the commitment benchmark and the complete information benchmark. Both benchmarks will yield interesting comparisons with my main result (Theorem 3.1), which characterizes a patient firm's equilibrium payoffs in a repeated incomplete information game without commitment.

Stackelberg Commitment Payoff: First, let us consider the benchmark scenario in which the firm can commit to an (possibly mixed) action $\alpha_1 \in \Delta\{H, L\}$ before the client chooses between T and N. Every type would optimally commit to play H with probability γ^* and L with complementary probability, with

$$\gamma^* \equiv \frac{c}{b+c}. \tag{3.4}$$

Type θ_i's payoff under her optimal commitment is:

$$v_i^{**} \equiv 1 - \theta_i \gamma^*, \tag{3.5}$$

which is her *Stackelberg commitment payoff* and $\gamma^* H + (1-\gamma^*)L$ is her *Stackelberg commitment action*.[9]

Complete Information Benchmark: Next, let us consider a complete information benchmark in which θ is common knowledge. The unique equilibrium outcome in the stage game N and the firm's payoff is 0. This is because L is the firm's dominant strategy after the client chooses T. This illustrates a lack-of-commitment/time inconsistency problem once compared to the commitment benchmark, which is of first order importance in business transactions (Mailath and Samuelson 2001, Ely and Välimäki 2003, Ekmekci 2011) as well as other economic applications such as fiscal and monetary policies (Barro and Gordon 1983, Barro 1986, Phelan 2006), sovereign debt default (Cole, Dow and English 1995), etc.

Next, suppose the complete information stage game is played repeatedly and future clients can perfectly observe the firm's past action choices. When δ is above some cutoff, outcome (T, H) can be enforced in equilibrium via the following grim-trigger strategy: a client trusts the firm if and only if the latter has never played L before, and at every history on the equilibrium path, the firm exerts high effort. Furthermore, as shown in Fudenberg, Kreps and Maskin (1990), a patient firm's equilibrium payoff set in this complete information repeated game is $[0, 1-\theta]$.

The payoff upper bound in the complete information repeated game, $1-\theta$, is type θ firm's *pure Stackelberg commitment payoff*, i.e. her payoff when she can only commit to play pure actions. In general, this is strictly lower than her Stackelberg commitment payoff, which implies that the firm's patience together with repeated interactions cannot overcome the lack-of-commitment problem

[9] This commitment payoff is also the unique equilibrium payoff in the private value informed principal game *a là* Maskin and Tirole (1990), in which the firm announces a mechanism and the client decides whether to participate (playing T) or not (playing N).

when her opponents cannot be motivated by intertemporal incentives.[10]

3.2.2 Alternative Applications & Interpretations

In this subsection, I provide several alternative interpretations of the baseline model. Readers can skip this subsection and proceed to Section 3.3 for the main result. The common features shared by these examples are (1) a patient long-run player faces a lack-of-commitment problem and (2) she has persistent private information about her temptation to betray her opponents' trust. Other aspects of the game can be different across applications, which include the timing of moves in the stage game, whether player 2 is a sequence of short-lived players or a continuum of long-lived players, whether player 1's action choices are perfectly monitored or not, etc. I will address the validity of my results in these variations in Section 3.5.

Relational Contracts: Player 1 is an agent, for example, a worker, supplier or private contractor. In every period, a principal (player 2, e.g. employer, final good producer) is randomly matched with the agent and decides whether to contract with her or skip the interaction. The principal incurs a fixed cost if he chooses to contract with the agent.[11] The agent can either be reciprocal by providing high quality or shirk and offer low quality. In line with the literature on incomplete contracts, the quality of the agent's service is not verifiable by court but can be observed by all the subsequent principals before they decide whether to contract with her or not. The cost of providing high quality is the agent's persistent private information, as it depends on her talent, production technology, etc.[12]

My main result extends to a variant of the baseline model when the agent's effort choice is a continuum and her past action choices are imperfectly monitored. By choosing effort level $e \in [0, 1]$, the realized quality is high with probability e in that period, and qualities across different periods are independent random variables. The subsequent principals can observe the past realizations of qualities, but not the agent's effort choices. My main result extends when the agent's cost of effort is linear and she has persistent private information about her marginal cost of effort.

Fiscal & Monetary Policies: Consider the following game adapted from Phelan (2006). There is a continuum of investors (player 2s) deciding whether to invest a unit of capital in a developing

[10] As shown in Fudenberg and Maskin (1986), the firm can attain her Stackelberg commitment payoff in repeated complete information games where both players are patient. Nevertheless, in many applications where there are serious lack-of-commitment problems, such as experience good markets, fiscal and monetary policies, sovereign debt default, etc. it is unreasonable to assume that the buyers, citizens, investors or creditors, etc. have strong intertemporal incentives when they are interacting with firms and governments. This is either because each of them only purchases the good once, or because they are anonymous and each one of them has negligible mass and has no impact on aggregate variables.

[11] The literature on organizational economics has provided several interpretations of this fixed cost, which includes an upfront payment the final good producer made to his supplier, a relationship specific investment the principal needs to make, etc. See Gibbons and Roberts (2013) for more details.

[12] Chassang (2010) studies a game with similar incentive structures, besides that the agent's cost of effort is common knowledge but the set of actions that are available in each period (which is i.i.d. across different periods) is the agent's private information. Tirole (1996) uses a similar model with commitment types to study the collective reputations for commercial firms and the corruption of bureaucrats.

country (action T) or not (action N). Player 1 is the government of that country, deciding between expropriate the proceeds (e.g. levying high taxes, action L) or not (action H). The government's benefit from expropriation is her private information. This is because the cost of collecting taxes depends on the state capacity and other factors, which cannot be perfectly observed by the citizens and tend to be persistent over time.

Next, consider a game between a central bank (player 1) and a continuum of households (player 2s). In every period, the central bank chooses the inflation level and at the same time, households form their expectations about inflation. To simplify matters, I assume both the actual and expected inflation are binary. Following Barro and Gordon (1983) and Barro (1986), players' stage game payoffs are given by:

-	Low Expectation	High Expectation
Low Inflation	$0, b$	$-1-d(\theta), -c$
High Inflation	$\theta, -c$	$-1, b$

where $\theta > 0$ is the central bank's private information. To interpret this payoff matrix, households want to match their expectations with the actual inflation. The central bank's payoff decreases with the actual inflation and increases with the amount of surprised inflation (equals to actual inflation minus expected inflation). As argued in Barro and Gordon (1983), the central bank's benefit from surprised inflation originates from the increase in real economic activities, decrease in unemployment rate and increase in governmental revenue. How the central bank trades-off these benefits with the costs of inflation, which is captured by θ, depends on the central banker's ideology and tends to be her persistent private information.

Different from the extensive form stage game in the baseline model, players move simultaneously in this application. I will show in Section 3.5 that every type's highest limiting equilibrium payoff remains the same as in the baseline model. But the limiting equilibrium payoff set expands due to the feasibility of the bad outcome with payoff $-1-d(\theta)$.

3.3 Main Result

In this section, I characterize a patient firm's equilibrium payoff set in this repeated incomplete information game. My theorem provides a clean formula for every type's *highest attainable payoff*. It clarifies the role of incomplete information in repeated games by highlighting the aspects of the type distribution that matter for a patient player's payoffs. My result also provides a (partial) strategic foundation for the Stackelberg commitment types in the reputation literature, both in terms of the patient long-run player's highest attainable payoff and in terms of her behavior in equilibria that approximately attain her Stackelberg commitment payoff.

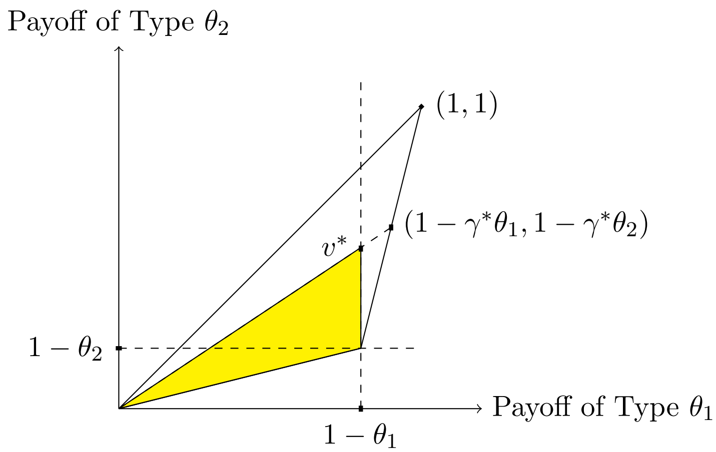

Figure 3-1: Set V^* (in yellow) when $|\Theta| = 2$.

3.3.1 Statement of Result

I start from providing the formula for every type's highest limiting equilibrium payoff. For every $\theta_i \in \Theta$, let

$$v_i^* \equiv \underbrace{(1-\gamma^*\theta_i)}_{\text{Type } \theta_i\text{'s Stackelberg commitment payoff}} \underbrace{\frac{1-\theta_1}{1-\gamma^*\theta_1}}_{\text{incomplete information multiplier}}, \tag{3.6}$$

which is the product of type θ_i's complete information Stackelberg commitment payoff and an *incomplete information multiplier*. Let $v^* \equiv (v_1^*, ..., v_m^*)$. Let V^* be the triangular set with vertices $(0, 0, ..., 0)$, $(1-\theta_1, ..., 1-\theta_m)$ and v^*. Figure 3-1 depicts V^* in a two-type example. My main result claims that V^* is the firm's limiting equilibrium payoff set.[13]

Theorem 3.1. *If π_0 has full support, then $\overline{V}(\pi_0) = \underline{V}(\pi_0) = V^*$.*

The proof of this result is in Appendices C.1 and C.2, with the intuition and ideas behind summarized in Section 3.4. The statement of Theorem 3.1 can be decomposed into a lower bound part and an upper bound part: (1) every payoff vector in the interior of V^* is attainable in sequential equilibria when δ is above some cutoff, or formally, $\underline{V}(\pi_0) \supset V^*$; (2) every payoff vector that is bounded away from V^* is not attainable in Nash equilibria when δ exceeds some cutoff, or formally, $\overline{V}(\pi_0) \subset V^*$. The proof of the first statement constructs equilibria that approximately attains v^* when δ is large enough. In these equilibria, the clients *gradually* learn about the firm's type and are playing myopic best replies. Therefore, v^* is also approximately attainable when the clients are forward looking. Nevertheless, the exact upper bound of the payoff set, namely the payoff upper bound part, depends on the clients' myopia.

To better understand V^*, it is instructive to write every feasible payoff vector $v \in \mathbb{R}^m$ as a convex combination of the payoff vectors from every stage game outcome: N, (T, H) and (T, L).

[13] Despite the theorem is stated in the context of the sequential move stage game with perfect monitoring, it can be generalized to stage games in which players move simultaneously, or the informed long-run player is choosing from a continuum of effort levels and her effort is observed with noise. Both extensions will be addressed in Section 3.5.

The set V^* is characterized by two constraints: (1) The equilibrium payoff of type θ_1 (the lowest cost type) cannot exceed $1-\theta_1$, her pure Stackelberg payoff; (2) The ratio between the convex weight of (T,H) and the convex weight of (T,L) is no less than $\gamma^*/(1-\gamma^*)$. The necessity and sufficiency of both constraints as well as the intuitions behind them will be explained in Section 3.4.

3.3.2 Economic Implications of Theorem 3.1

In this subsection, I outline the economic implications of my result. To draw connections with the reputation literature, I replace firm and clients with *long-run player* and *short-run players*, respectively.

First, the formula for v^* implies that aside from type θ_1 (the lowest cost type), every other type can strictly benefit from her persistent private information in terms of her highest equilibrium payoff. A sufficient statistics for the impact of incomplete information is the multiplier $(1-\theta_1)\big/(1-\gamma^*\theta_1)$, which only depends on the lowest possible cost in the support of her opponents' prior belief. In another word, it is independent of the other realizations of θ and the details of the probability distribution.

Second, the incomplete information multiplier converges to 1 as θ_1 vanishes to 0. This implies that the long-run player can overcome her lack-of-commitment problem and can approximately attain her Stackelberg commitment payoff in every state of the world. This observation is formally stated as Corollary 3.1, which is a straightforward implication of Theorem 3.1:

Corollary 3.1. *For every $\epsilon>0$, there exist $\overline{\delta}\in(0,1)$ and $\overline{\theta}_1>0$ such that when $\delta>\overline{\delta}$, $\theta_1<\overline{\theta}_1$ and $\pi_0(\theta_1)>\epsilon$, there exists a sequential equilibrium in which type θ_i's equilibrium payoff is no less than $v_i^*-\epsilon$ for every $i\in\{1,2,...,m\}$.*

This corollary is reminiscent of a well-known result in the reputation literature, *à la* Fudenberg and Levine (1989, 1992), which claims that a patient long-run player can approximately attain her Stackelberg commitment payoff if with positive probability, she is irrational and mechanically plays her Stackelberg commitment strategy.

Formally, let $\sigma_\theta^*:\mathcal{H}\rightarrow\Delta(A_1)$ be a *commitment strategy* for the long-run player and let $\Sigma_2^*(\sigma_\theta^*)$ be the set of complete information best replies to σ_θ^*.[14] Type θ's commitment payoff from σ_θ^* is:

$$U(\sigma_\theta^*)\equiv\inf_{\sigma_2^*\in\Sigma_2^*(\sigma_\theta^*)}\Big\{\mathbb{E}^{(\sigma_\theta^*,\sigma_2^*)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t u_1(\theta,a_t)\Big]\Big\}. \tag{3.7}$$

I say that σ_1^* is type θ_i's ϵ-*Stackelberg commitment strategy* if $U(\sigma_1^*)\geq 1-\gamma^*\theta-\epsilon$. The canonical example of an ϵ-Stackelberg commitment strategy is the *stationary Stackelberg strategy* in which $\sigma_\theta^*(h^t)[H]=\gamma'$ for every $h^t\in\mathcal{H}$, with $\gamma'\in(\gamma^*,\gamma^*+\epsilon/\theta]$.[15]

[14] The state θ and player 2's belief about θ are irrelevant for player 2s' best reply against σ_θ^* as the game is of private values.

[15] In an incomplete information private value environment, the stationary Stackelberg strategy for every type is to play $\gamma' H+(1-\gamma')L$ at every history, with $\gamma'\in(\gamma^*,\gamma^*+\epsilon/\theta_m]$, where θ_m is the greatest element in Θ.

Fudenberg and Levine (1992) show that if the short-run players' prior attaches strictly positive probability to an *irrational type* who mechanically plays some ϵ-Stackelberg commitment strategy, then in a simultaneous move stage game where the long-run player's actions can be statistically identified, a sufficiently patient long-run player can guarantee herself payoff $1-\gamma^*\theta-2\epsilon$ in every Nash equilibrium. When players move sequentially as in the baseline model, one can show that there exist Nash equilibria in which the patient long-run player obtains payoff at least $1-\gamma^*\theta-2\epsilon$ when the stage game outcomes are perfectly monitored. These results point out the following *reputation effects*, that incomplete information can help a patient long-run player overcome her lack-of-commitment problem and attain her commitment payoff.

Can those irrational types that are playing *mixed commitment strategies* be rationalized?[16] This question is economically important in assessing whether these reputation effects can arise under common knowledge of rationality or they rely on the long-run player's irrationality. Different from the conventional approach of using *crazy types* that have qualitatively different payoff functions, I maintain the realistic assumptions on the long-run player's stage game payoffs by requiring all types to share the *same ordinal preferences* over stage game outcomes.[17] In context of the trust game, it implies that the following aspects are common knowledge, which are reasonable in the applications to business transactions and public policies:

1. The long-run player can (strictly) benefit from the short-run players' trust.

2. Conditional on being trusted, she finds it strictly beneficial to betray them (such as providing low quality, shirking, expropriating returns from investments, setting high inflation rates, etc).

The above requirement makes the conclusion more convincing yet also create new challenges in answering the above question. This is because the long-run player's ϵ-Stackelberg commitment strategies are non-trivially mixed and motivating her to randomize between actions is difficult as she has strict preferences over outcomes.[18]

Corollary 3.1 implies that the irrational types that are playing ϵ-Stackelberg commitment s-trategies can be *partially* rationalized by a strategic type that (1) has a standard stage game payoff function; (2) her cost of honoring her commitment is sufficiently low compared to her benefit from

[16]This distinguishes the research question in my paper with that of Weinstein and Yildiz (2016), in which they seek to rationalize irrational types that are playing pure but non-stationary commitment strategies.

[17]This conventional approach is adopted by Weinstein and Yildiz (2007), who show under a richness assumption (i.e. every action is strictly dominant for some types) that if all payoffs and hierarchies of beliefs are allowed, then every interim correlated rationalizable outcome is uniquely rationalizable in some nearby games according to the product topology. Weinstein and Yildiz (2013,2016) extend this approach to show an unrefineable folk theorem in repeated Bayesian games and to rationalize non-stationary commitment strategies in reputation games. These aforementioned results all rely on the existence of some crazy types that have non-standard stage game payoffs and well-calibrated hierarchies of beliefs.

[18]The standard techniques to construct mixed strategy equilibria in repeated games, such as the belief-free equilibrium approach in Ely, Hörner and Olzewski (2005), Hörner and Lovo (2009) cannot be applied in this context. This is because the uninformed players are myopic, so in every equilibrium that attains payoff close enough to v^*, different types of long-run players are mixing between H and L with different probabilities, i.e. these equilibria are not belief-free.

her opponents' trust. By partially rationalized, I mean that the patient long-run player can attain her Stackelberg commitment payoff in *some equilibria* of the repeated game without irrational types. Nevertheless, the limiting equilibrium payoff sets in the repeated game with and without irrational types may not coincide.

But how will the lowest cost type behave in equilibria that approximately attain the Stackelberg commitment payoff? In particular, which of the ϵ-Stackelberg commitment strategies will she play? This question is motivated by the fact that are many ϵ-Stackelberg commitment strategies. So it is important to know which commitment strategies are more likely to arise when the long-run player is rational. Contrary to the conventional practice of focusing on stationary commitment strategies, I show in Corollary 3.2 that in every equilibrium that approximately attains v^*, the lowest cost type *cannot* play a non-trivial mixed action at every history (where she can make moves). This implies that irrational types that are playing stationary Stackelberg strategies *cannot* be rationalized by a rational type that has arbitrarily low cost to honor her commitment.

Corollary 3.2. *Suppose $\theta_1 < \gamma^* \theta_m$. For every full support $\pi_0 \in \Delta(\Theta)$, there exist $\epsilon > 0$ and $\overline{\delta} \in (0,1)$ such that for every $\delta > \overline{\delta}$, there exists no Nash equilibrium in which (1) player 1's equilibrium payoff is within ϵ of v^*; (2) type θ_1 plays a non-trivially mixed action at every history.*

Proof of Corollary 3.2: Suppose towards a contradiction that such an equilibrium (σ_1, σ_2) exists, then playing L at every history is type θ_1's best reply to σ_2. Since her equilibrium payoff is at least $1 - \theta_1 - \epsilon$, the occupation measure of stage game outcome (T, L) induced by playing L at every history is at least $1 - \theta_1 - \epsilon$. This implies that by playing L in every period, type θ_m (the highest cost type) can obtain payoff at least $1 - \theta_1 - \epsilon$, which is no greater than her equilibrium payoff.

When $\theta_1 < \gamma^* \theta_m$, let $\epsilon \equiv (\gamma^* \theta_m - \theta_1)/3$. The payoff upper bound result in Fudenberg and Levine (1992) implies the existence of $\overline{\delta} \in (0,1)$ such that for every $\delta > \overline{\delta}$, type θ_m's equilibrium payoff cannot exceed $1 - \gamma^* \theta_m + \epsilon$. According to the choice of ϵ, we have:

$$1 - \gamma^* \theta_m + \epsilon < 1 - \theta_1 - \epsilon, \tag{3.8}$$

with the right-hand-side being no greater than type θ_m's equilibrium payoff. This leads to a contradiction. □

As has become clear in the proof, the conclusion in Corollary 3.2 remains valid when $\theta_1 = 0$. This is somewhat surprising as player 1 cannot be mixing at every history despite her cost of exerting high effort being 0. Intuitively, this is because despite her being indifferent between H and L in the stage game, her action choices will affect the (discounted average) frequency with which player 2 trusts her in the future.[19] If the zero-cost type is always indifferent between H and L, then playing L at every history as well as playing H at every history will result in the same frequency with which player 2 plays T. Both of which are arbitrarily close to 1. As a result, all other types will

[19] Conceptually, this distinguishes a zero-cost type with types that are completely indifferent between all stage-game outcomes.

strictly prefer to play L at every history. According to the standard Bayesian learning argument in Fudenberg and Levine (1989,1992), the short-run players will learn that L will be played with very high probability after observing L in the first T periods (with T a bounded number), after which they will play N in all subsequent periods. This contradicts the previous claim that they will play T with frequency close to 1.

Given the conclusion in Corollary 3.2, which ϵ-Stackelberg commitment strategies are played by the lowest-cost type in equilibrium? Let σ_{θ_1} be the lowest-cost type's strategy in an equilibrium that attains payoff close to v^*. For some necessary conditions on σ_{θ_1}, notice that first, type θ_m's equilibrium payoff cannot exceed $1 - \gamma^*\theta_m + \epsilon$ according to the payoff upper bound result in Fudenberg and Levine (1992). This implies that the occupation measure of outcome (T, H) along every action path that occurs with positive probability under $(\sigma_{\theta_1}, \sigma_2)$ is no less than γ^*. Second, if σ_{θ_1} is an ϵ-Stackelberg commitment strategy, then the expected occupation measure of (T, L) must be close to $1 - \gamma^*$. To summarize, when δ is large enough, the distribution over equilibrium action paths induced by $(\sigma_{\theta_1}, \sigma_2)$ satisfies:

- This distribution only attaches strictly positive probability to action paths in which (T, H) has occupation measure no less than γ^*.
- For every $\tau > 0$, the probability measure of action paths in which the occupation measure of (T, H) exceeds $\gamma^* + \tau$ vanishes to 0 as $\delta \to 1$ and $\epsilon \to 0$.

3.4 Proof of Theorem 3.1: Intuition and Ideas

The proof hinges on understanding the sufficiency and necessity of the constraints characterizing the limiting equilibrium payoff set V^*:

1. The equilibrium payoff of type θ_1 cannot exceed $1 - \theta_1$.
2. The ratio between the convex weight of (T, H) and the convex weight of (T, L) is no less than $\gamma^*/(1 - \gamma^*)$.

For illustration purposes, I focus on an example with two types, i.e. $\Theta = \{\theta_1, \theta_2\}$. In subsection 3.4.1, I explain the necessity of these constraints and relate them to the uninformed player's incentives. In subsection 3.4.2, I explain the ideas behind the constructed equilibria that can approximately attain v^*. In subsection 3.4.3, I compare the equilibrium dynamics in my construction to the related ones in the existing literature.

3.4.1 Necessity

Let $\sigma \equiv (\sigma_{\theta_1}, \sigma_{\theta_2}, \sigma_2)$ be a generic Nash equilibrium. To understand the necessity of the first constraint, it is instructive to introduce the *highest action path*. Formally, let $\mathcal{H}(\sigma)$ be the set of on-path histories. For every $h^t \in \mathcal{H}(\sigma)$ such that $\sigma_2(h^t)[T] > 0$, let $\Theta(h^t)$ be the support of player

2's posterior belief at h^t. The highest action path $\overline{\sigma}_1 : \mathcal{H}(\sigma) \to \{H, L\}$ is defined as:

$$\overline{\sigma}_1(h^t) \equiv \begin{cases} H & \text{if } H \in \bigcup_{\theta \in \Theta(h^t)} \operatorname{supp}\big(\sigma_\theta(h^t)\big) \\ L & \text{otherwise .} \end{cases} \tag{3.9}$$

Since player 2 is myopic, he has an incentive to play T at h^t only when $\overline{\sigma}_1(h^t) = H$. By construction, $\overline{\sigma}_1$ is at least one type's best reply to σ_2. Consider two cases separately: (1) Suppose it is type θ_1's best reply, then type θ_1's payoff in every period cannot exceed $1 - \theta_1$, which implies that her discounted average payoff is no more than $1 - \theta_1$. (2) Suppose it is type θ_2's best reply, then type θ_2's payoff in every period cannot exceed $1 - \theta_2$. Since the differences between type θ_1 and type θ_2's payoff is at most $\theta_2 - \theta_1$, type θ_1's discounted average payoff cannot exceed $1 - \theta_1$. The necessity of the first constraint is obtained by summing up the two cases.[20]

To understand the second constraint, I introduce an alternative version of the highest action path based on σ_{θ_2}. Let $\overline{\sigma}_{\theta_2} : \mathcal{H} \to \{H, L\}$ be defined as:

$$\overline{\sigma}_{\theta_2}(h^t) \equiv \begin{cases} H & \text{if } H \in \operatorname{supp}\big(\sigma_{\theta_2}(h^t)\big) \\ L & \text{otherwise,} \end{cases} \tag{3.10}$$

which by construction, is type θ_2's best reply to σ_2.

Suppose type θ_2's payoff from σ is strictly greater than $1-\theta_2$. If player 2 plays according to $\overline{\sigma}_{\theta_2}$, then she can receive stage game payoff strictly greater than $1 - \theta_2$ only at histories where player 2 plays T but σ_{θ_2} prescribes L with probability 1. Player 2's incentive constraints imply that at those histories, σ_{θ_1} prescribes H with sufficiently high probability. Let $\eta(h^t)$ be the probability of type θ_1 at h^t. The above argument implies that if type θ_2 plays L, then she can extract information rent at the expense of revealing information about her type (i.e. $\eta(h^t, L) < \eta(h^t)$); if she plays H, then she will sacrifice her current stage payoff in exchange for a more favorable belief in the future (i.e. $\eta(h^t, H) > \eta(h^t)$).

Now comes the key question: What is the maximal frequency with which (T, L) occurs according to $(\overline{\sigma}_{\theta_2}, \sigma_2)$? Player 2's incentive to play T implies the following upper bound on the relative speed with which η increases after observing H to the speed with which it decreases after observing L:

$$\frac{\eta(h^t, H) - \eta(h^t)}{\eta(h^t) - \eta(h^t, L)} \leq \frac{1 - \gamma^*}{\gamma^*}. \tag{3.11}$$

Inequality (3.11) implies that in order to provide player 2s the incentives to play T in the long-run, the ratio between the frequencies of L and H according to $(\overline{\sigma}_{\theta_2}, \sigma_2)$ cannot exceed $(1 - \gamma^*)/\gamma^*$.

Nevertheless, the first constraint implies that as long as $\theta_1 > 0$, the occupation measure with which player 2 plays T must be bounded away from 1 when play proceeds according to $(\overline{\sigma}_{\theta_2}, \sigma_2)$. For example, in the equilibrium constructed in Appendix C.1, the play according to $(\overline{\sigma}_{\theta_2}, \sigma_2)$ consists of three phases: it starts from a *reputation building phase* in which the outcome is (T, H), followed by a *reputation manipulation phase* in which the outcome alternates between (T, H) and (T, L),

[20] This insight extends whenever there is complementarity between θ and (a_1, a_2) in player 1's payoff function.

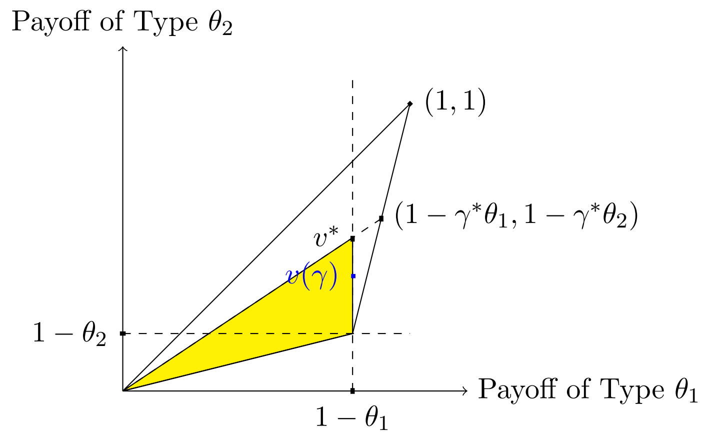

Figure 3-2: V^* in yellow and $v(\gamma)$ in blue for some $\gamma \in (\gamma^*, 1)$.

with the fraction of (T, L) being no more than $1 - \gamma^*$, followed by a *punishment phase* in which the outcome is N. The occupation measure of the first phase goes to 0 as $\delta \to 1$, but the second and third phases have strictly positive occupation measures in the limit.

3.4.2 Sufficiency

In this subsection, I construct a class of sequential equilibria that can approximately attain payoff v^* when δ is large enough. In these equilibria, the long-run player can extract information rent only when her actions are informative about her type and her reputation is *gradually* gained and lost in periods when there is learning.

For every $i \in \{1, 2, ..., m\}$ and $\gamma \in [\gamma^*, 1]$, let

$$v_i(\gamma) \equiv (1 - \gamma\theta_i)\frac{1-\theta_1}{1-\gamma\theta_1} \tag{3.12}$$

and $v(\gamma) \equiv \Big(v_i(\gamma)\Big)_{i=1}^{m}$. An example of $v(\gamma)$ is shown in Figure 3-2. By definition, $v_i(\gamma^*) = v_i^*$ and $v_i(1) = 1 - \theta_i$. Proposition 3.1 claims that every $v(\gamma)$ with $\gamma > \gamma^*$ is attainable in sequential equilibrium when the long-run player is sufficiently patient. Given that other two extreme points of V^* are trivially attainable, namely $(0, 0, ..., 0)$ and $(1 - \theta_1, ..., 1 - \theta_m)$, Proposition 3.1 also implies that every payoff vector in the interior of V^* is attainable when δ is large enough.

Proposition 3.1. *For every $\overline{\eta} \in (0, 1)$ and $\gamma \in (\gamma^*, 1)$, there exists $\overline{\delta} \in (0, 1)$, such that for every $\delta > \overline{\delta}$ and $\pi_0 \in \Delta(\Theta)$ with $\pi_0(\theta_1) \geq \overline{\eta}$, there exists a sequential equilibrium in which player 1's payoff is $v(\gamma)$.*

The rest of this subsection has two parts. In Part I, I provide an overview of the equilibrium construction, and in particular, players' strategies and belief updating process. In Part II, I explain the motivation of the construction and the ideas to overcome the underlying challenges. The details can be found in Appendix C.1.

Overview of Equilibrium Construction: The equilibrium play consists of three phases: a *normal phase* and two *absorbing phases*. I keep track of four state variables, namely, the probability with which player 2's posterior belief attaches to type θ_1, denoted by $\eta(h^t)$, and the remaining occupation measure of each stage game outcome, denoted by $p^N(h^t)$, $p^H(h^t)$ and $p^L(h^t)$, respectively. The initial values of these state variables are given by:

$$\eta(h^0) = \pi_0(\theta_1) \ , \ p^N(h^0) = \frac{\theta_1(1-\gamma)}{1-\gamma\theta_1} \ , \ p^H(h^0) = \frac{(1-\theta_1)\gamma}{1-\gamma\theta_1} \text{ and } p^L(h^0) = \frac{(1-\theta_1)(1-\gamma)}{1-\gamma\theta_1}.$$

One can verify that for every $i \in \{1, 2, ..., m\}$,

$$v_i(\gamma) = p^H(h^0)(1-\theta_i) + p^L(h^0),$$

i.e. the initial value of p^a is the convex weight of outcome a under the target payoff $v(\gamma)$.

Play starts from the *normal phase*, in which player 2 plays T in every period and player 1 is mixing between H and L with every type's mixing probabilities pinned down by the following belief-updating formulas:

$$\eta(h^t, L) - \eta^* = (1 - \lambda\gamma^*)(\eta(h^t) - \eta^*) \tag{3.13}$$

and

$$\eta(h^t, H) - \eta^* = \min\Big\{1 - \eta^*, \big(1 + \lambda(1-\gamma^*)\big)\big(\eta(h^t) - \eta^*\big)\Big\}, \tag{3.14}$$

where η^* is an arbitrary number within $\Big(\gamma^*\eta(h^0), \eta(h^0)\Big)$ and $\lambda > 0$. Intuitively, λ measures the speed of learning and η^* is a lower bound on player 2's normal phase posterior belief. The role of η^* is to satisfy player 2's incentive constraint by the end of the normal phase, with details explained in Part I of Appendix C.1.3. The restriction on λ is specified in (C.5). Intuitively, if the frequency of H is slightly more than $\frac{\gamma^*}{1-\gamma^*}$ times the frequency of L, then player 2's long-run posterior will attach higher probability to type θ_1 compared to his prior given that λ is small enough.

Based on the realized stage game outcome, $p^a(h^t)$ evolves according to:

$$p^a(h^t, a_t) \equiv \begin{cases} p^a(h^t) & \text{if } a_t \neq a \\ p^a(h^t) - (1-\delta)\delta^t & \text{if } a_t = a \end{cases} \tag{3.15}$$

with $a, a_t \in \{N, H, L\}$. Intuitively, player 1 has three separate accounts, each representing the occupation measure of an individual stage game outcome. The state variable $p^a(h^t)$ is then interpreted as the remaining credit in the account for outcome a.

Play transits to the first absorbing phase when $\eta(h^t)$ reaches 1, after which the continuation value of type θ_i equals to $v_1(h^t)\frac{1-\theta_i}{1-\theta_1}$, with

$$v_1(h^t) \equiv \frac{p^H(h^t)(1-\theta_1) + p^L(h^t)}{p^H(h^t) + p^L(h^t) + p^N(h^t)}. \tag{3.16}$$

The resulting payoff vector can be delivered by randomizing between stage game outcomes N and

(T, H).

Play transits to the second absorbing phase when $p^L(h^t)$ is between 0 and $(1-\delta)\delta^t$, or intuitively, $p^L(h^t)$ is close to 0. For illustration purposes, let us focus on the ideal situation in which $p^L(h^t) = (1-\delta)\delta^t$.[21] If player 1 plays L, then play transits to the second absorbing phase and the continuation payoff for type θ_i equals to:

$$v_i(h^t) \equiv \frac{p^H(h^t)(1-\theta_i) + p^L(h^t)}{p^H(h^t) + p^L(h^t) + p^N(h^t)} \tag{3.17}$$

This payoff vector can also be delivered by randomizing between stage game outcomes N and (T, H).

Ideas & Intuitions Behind the Construction: The equilibrium construction is designed to achieve the following objective, namely, making outcome (T, L) incentive compatible for the myopic uninformed players while at the same time, providing incentives for all types of the informed player to mix.

For this purpose, the constructed equilibrium has two defining features (1) the rent extraction outcome (T, L) only occurs at histories where the uninformed players can learn about θ (i.e. in the normal phase); (2) the informed player is facing a trade-off between extracting information rent (by playing L) and building up her reputation (by playing H) throughout the normal phase. As argued before, learning is necessary for rent extraction due to the uninformed players' myopia. Introducing the trade-off between reputation and rent extraction at *all histories* (of the normal phase) makes the informed player indifferent to the *timing of rent extraction*, which holds irrespective of her previous play. This motivates her to play according to the highest action path at every normal phase history, which by construction, backloads the outcome (T, L) by as much as possible.

My construction of the normal phase raises two issues, which motivate the design of the two absorbing phases. The first concern is that the play of H can be too front-loaded, after which there is too much L remaining in player 1's account (relative to H) and the resulting continuation payoff cannot be delivered in an incentive compatible way. The first absorbing phase is designed to address this issue: if player 1 front-loads the play of H, then play will transit to the first absorbing phase, after which type θ_2's continuation payoff is strictly less compared to her payoff from playing L at the transition history. In general, the presence of the first absorbing phase ensures that at every history of the normal phase, player 1's continuation value is always within some proper subsets of V^*, and in particular, the ratio between the remaining occupation measures of L and H cannot exceed some endogenous cutoff. This is stated as Lemma C.1.1. The challenges to prove this result are explained in Appendix C.1.

The second concern is that type θ_2 may have incentives to front-load rent extraction by rarely

[21] Notice that due to integer constraints, $p^L(h^t)$ could be strictly between 0 and $(1-\delta)\delta^t$ in period t. If this is the case, then playing L in period t will result in $p^L(h^{t+1})$ being strictly negative and the continuation payoff vector being outside V^*. I introduce the reshuffling *phase* in Appendix A to deal with this complication. The idea is: the continuation payoff vector can be written as a convex combination of three other payoff vectors in V^*, all of which satisfy the integer constraints.

playing H. This explains the presence of the second absorbing phase. In particular, if she plays L too frequently, then she will reach the second absorbing phase at an earlier date after which player 2 will always play N and no information can be extracted any more in the future.

3.4.3 Comparisons

In this subsection, I compare the equilibrium dynamics in my model to the related ones in the existing literature. This includes the reputation models with behavioral biases (Jehiel and Samuelson 2012), reputation cycles (Phelan 2006, Liu and Skryzpacz 2014) and models of repeated incomplete information games with two patient players (Hart 1985, Aumann and Maschler 1995, Pęski 2014).

Analogical-Based Reasoning Equilibria: The behavioral patterns suggested by type θ_2's highest action path, $\overline{\sigma}_{\theta_2}$ (previously defined in subsection 3.4.1), is reminiscent of the analogical reasoning equilibria in reputation games studied by Jehiel and Samuelson (2012), in which the long-run player switches between her two actions systematically to manipulate her opponents' beliefs about her type.

In their model, there are multiple irrational types of the long-run player that are playing stationary mixed strategies and one rational type. The short-run players' adopt an *analogical based* reasoning process, i.e. they mistakenly believe that the rational long-run player is playing a stationary strategy. Their results imply, in context of the trust game, that the rational long-run player can attain her Stackelberg commitment payoff and her equilibrium behavior will experience a *reputation building* (or *reputation consumption*) phase in which she plays H (or L) for a bounded number of periods, followed by a *reputation manipulation* phase, in which she alternates between H and L so that her opponents are close to being indifferent. Moreover, the short-run players' posterior will fluctuate within a small neighborhood of the cutoff belief, implying that they will never fully learn about the long-run player's type.

In my model, despite type θ_2's behavior following her highest action path exhibits a similar pattern, there are two qualitative differences that highlight the distinctions between rational and analogical-based uninformed players. First, the reputation manipulation phase cannot last forever in my model due to the constraint that type θ_1's equilibrium payoff cannot exceed $1 - \theta_1$. This constraint is driven by the uninformed players' ability to correctly predict the informed player's average action in *every period*, while analogy-based uninformed players can only correctly forecast the informed player's average action *across different periods*. Second, the uninformed players can perfectly learn the state with positive probability in the manipulation phase of my model. In particular, if type θ_2 plays according to $\overline{\sigma}_{\theta_2}$, she can only extract information rent at histories where player 2 can learn $\theta = \theta_1$ perfectly with positive probability. In contrast, the analogy-based agents in Jehiel and Samuelson (2012) can never learn the state perfectly as their posterior beliefs are not responsive enough to the actions they observe.

Reputation Cycles: The equilibrium dynamics of my model are also related to the phenomena called *reputation cycles*, which have been identified in a number of papers when the informed player's

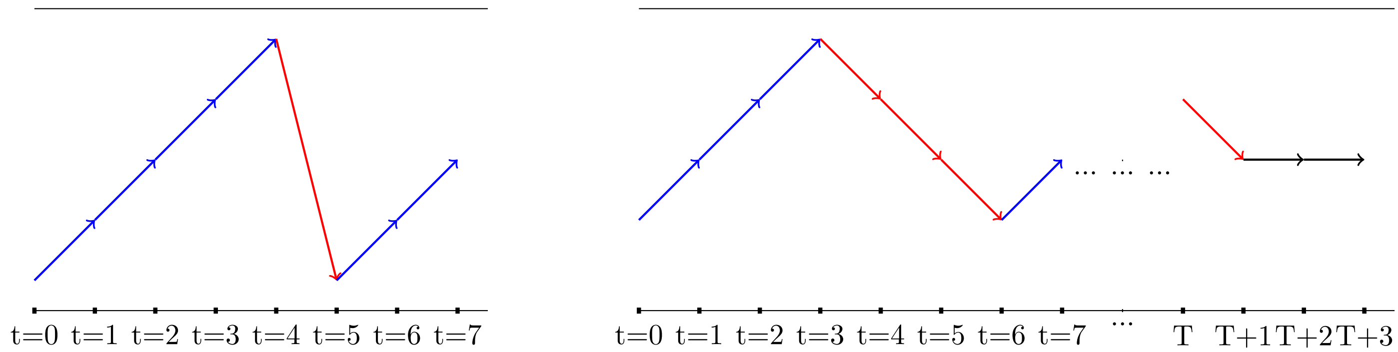

Figure 3-3: The horizontal axis represents the timeline and the vertical axis measures the informed player's reputation, i.e. probability of the commitment type or lowest cost type. Left: Reputation cycles in Phelan (2006). Right: A sample path of the reputation cycle in my model.

type (irrational or rational) is changing over time (Phelan 2006) or when the uninformed players have limited memories (Liu 2011, Liu and Skrzypacz 2014). Nevertheless, there are two important differences in the modeling choices and the resulting dynamics. First, all types of the long-run player are rational and have standard stage game payoff functions in my model, while the papers on reputation cycles require the presence of an irrational type, or types that have qualitatively different stage game payoff functions compared to the normal one.

Second and more importantly, reputations are built and lost *gradually* in my model (captured by a small λ), while in Phelan (2006), Liu (2011), Liu and Skrzypacz (2014), etc. the informed player's reputation will drop to its lower bound whenever she betrays the uninformed players' trust. This gradual reputation building and milking featured in my model is supported empirically from various studies of online markets (Dellarocas 2006). Intuitively, this is because when all types of the long-run player are rational and have strictly positive temptation to renege, even the lowest cost type will shirk with strictly positive probability in the optimal equilibrium, whereas in the canonical reputation models, the irrational types are mechanically playing some stationary strategies which do not exhibit the flexibility of conditioning action choices on the short-run players' belief.

Third, similar to the comparisons to Jehiel and Samuelson (2012), the reputation cycle can last forever in Phelan (2006), Liu (2011) and Liu and Skrzpacz (2014), while it will stop in finite time in my model (after which player 2 will play N forever). To be more precise, the expected occupation measure of the reputation manipulation phase (one that features cycles) will be strictly bounded away from 1, but the expected number of periods of this phase will go to infinity as $\delta \to 1$. These differences are shown in Figure 3-3.

Repeated Incomplete Information Games: A feature of the equilibrium distinguishes my model from the undiscounted or zero-sum repeated incomplete information games (for example, Hart 1985, Aumann and Maschler 1995 Cripps and Thomas 2003), namely, the informed player can only extract information rent in the learning process. In those papers, the patient informed player's equilibrium payoff only depends on the uninformed player's posterior belief after learning stops but the learning phase itself has negligible payoff consequences. In my model, the normal phase's occupation measure is bounded away from 0 even in the $\delta \to 1$ limit and moreover, the informed player's information rent can *only* be delivered in the learning process.

The equilibrium dynamics in my model also contrasts to the ones in discounted non zero-sum repeated games with two patient players (for example, Pęski 2014). In games with two patient players, it is possible for some types of informed players to extract information rent (i.e. obtaining payoff strictly higher than her complete information level) without revealing information about her type, while in my model when the uninformed players are myopic, rent extraction must be accompanied by learning. The uninformed players' lack of intertemporal incentives also introduces a novel constraint on the informed player's equilibrium payoff set, namely, the lowest cost type's equilibrium payoff cannot exceed her pure Stackelberg commitment payoff.

3.5 Extensions & Concluding Remarks

This paper explores the role of incomplete information in repeated interactions when a patient long-run player faces a lack-of-commitment problem. In context of *trust games*, including the product choice game in Mailath and Samuelson (2001), the capital taxation game in Phelan (2006), the monetary policy game in Barro (1986), etc. a patient long-run player can overcome her lack-of-commitment problem and attain her Stackelberg commitment payoff when she has persistent private information about her cost of honoring her commitment.

Compared to the existing reputation models, my model features common knowledge of rationality and does not require any type to have a drastically different stage game payoff function or complicated hierarchies of beliefs and higher order beliefs, i.e. all types share the same belief as well as the same ordinal preference over stage game outcomes. According to this perspective, my result provides a partial strategic foundation for the (mixed strategy) Stackelberg commitment types in the reputation literature and helps to evaluate which of the Stackelberg commitment behaviors are more plausible in strategic environments.

Despite my results are stated when players move sequentially in the stage game and the outcomes in every period can be perfectly monitored, the main insights extend to a number of alternative modeling specifications.

Simultaneous Move Stage Games: Consider the following simultaneous-move stage game:

$\theta = \theta_i$	T	N
H	$1-\theta_i, b$	$-d(\theta_i), 0$
L	$1, -c$	$0, 0$

with $d(\theta_i) \geq 0$ for every $\theta_i \in \Theta$.

Suppose the above simultaneous-move stage game is played repeatedly in discrete time. Players' past action choices are perfectly monitored and the public history $h^t \equiv \{a_{1,s}, a_{2,s}, \xi_s\}_{s=0}^{t-1}$ consists of players' past action choices and the past realizations of the public randomization device. Other features of the game remain the same as in the baseline model. Recall the definition of $v^* \equiv (v_i^*)_{i=1}^m$ in (3.6) and $\underline{V}(\pi_0)$ in (3.2), we have the following result on the attainability of v^* when player 1 is sufficiently patient.

Theorem 3.2. *If π_0 has full support, then $v^* \in \underline{V}(\pi_0)$.*

The proof of Theorem 3.2 follows from that of the payoff lower bound part in Theorem 3.1, which can be found in Appendix C.1. An implication of this theorem is that when the lowest possible cost θ_1 vanishes to 0, then every type can approximately attain her Stackelberg commitment payoff in sequential equilibrium. Furthermore, under a supermodularity condition on player 1's stage game payoff function, i.e. an assumption on the function $d(\cdot)$, one can provide a full characterization on the game's limiting equilibrium payoff set.

Condition 3.1. *u_1 is supermodular if $0 \leq d(\theta_j) - d(\theta_i) \leq \theta_j - \theta_i$ for every $j < i$.*

Intuitively, when we rank the types and players' actions according to $\theta_1 \succ \theta_2 \succ ... \succ \theta_m$, $H \succ L$ and $T \succ N$, Condition 3.1 implies that u_1 is supermodular in θ and (a_1, a_2). In particular, when $d(\theta_j) - d(\theta_i) = 0$ for every i and j, then the stage game payoff function is the same as in the sequential move game. When $d(\theta_j) - d(\theta_i) = \theta_j - \theta_i$ for every i and j, then the stage game payoff is separable, i.e. the cost for player 1 to play H is independent of player 2's action choice. This leads to the following payoff upper bound result

Theorem 3.3. *If u_1 is supermodular, then $\max \overline{V}_i(\pi_0) \leq v_i^*$ for every $i \in \{1, 2, ..., m\}$.*

Theorem 3.3 will be shown together with the payoff upper bound part of Theorem 3.1 in Appendix B. Under the supermodularity condition on player 1's stage game payoff, one can also provide the following characterization of player 1's limiting equilibrium payoff set. To begin with, I say that a payoff vector $v \in \mathbb{R}^m$ is *incentive compatible* if there exists $(\alpha_1, a_2) \in \Delta(A_1) \times A_2$ such that

$$a_2 \in \arg\max_{a_2' \in A_2} u_2(\alpha_1, a_2') \tag{3.18}$$

and $u_1(\theta_i, \alpha_1, a_2) = v_i$ for every $i \in \{1, 2, ..., m\}$. Player 1's limiting equilibrium payoff set V^{**} can be obtained via the following procedure:

1. Take the convex hull of the set of incentive compatible payoff vectors.

2. Truncate this set with two constraints, namely, $v_1 \leq 1 - \theta_1$ and $v_i \geq 0$ for every i.

An example of this set is depicted in Figure 3-4. The proof is similar to that of Theorem 3.1, and in particular, the attainability of the payoff vector marked as the black dot. The details are available upon request.

Stage Games with Continuum of Actions and Noisy Monitoring: The stage game is of sequential-move as in the baseline model, with the difference that if player 2 chooses T, then player 1 chooses between a continuum of effort levels $e \in [0, 1]$ and the output she produces ($y \in \{G, B\}$)

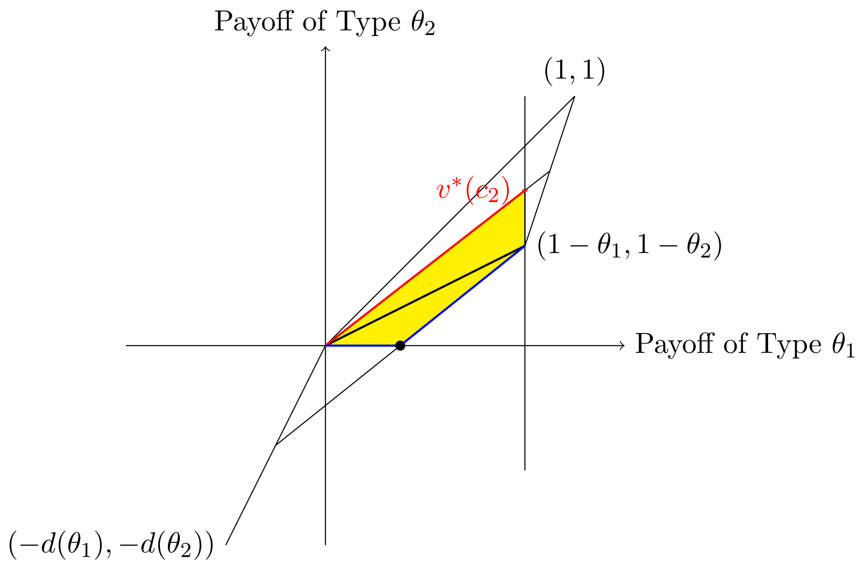

Figure 3-4: In Yellow: A patient player 1's equilibrium payoff set when the stage game is of simultaneous-move and $|\Theta| = 2$.

is good ($y = G$) with probability e. The cost of effort for type θ_i is $\theta_i e$. Her benefit from player 2's trust is normalized to 1, so her stage game payoff under outcome N is 0 and that under outcome (T, e) is $1 - \theta_i e$. Player 2's payoff is 0 if he chooses N. His benefit from good output is b while his loss from bad output is c, with $b, c > 0$. Therefore, player 2 is willing to trust player 1 only when his expectation of effort is no less than γ^*, which has been defined in (3.4).

Next, consider the repeated version of this game in which the public history consists of player 2's actions, the realized outputs and the realizations of public randomization devices in the past, but not the exact effort level. Formally, let $a_{1,t}$, $a_{2,t}$ and y_t be player 1's action, player 2's action and the realization of public signal in period t, respectively. Let $h^t = \{a_{2,s}, y_s, \xi_s\}_{s=0}^{t-1} \in \mathcal{H}^t$ be the public history with $\mathcal{H} \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}^t$ the set of public histories. Player 2's strategy is measurable with respect to the public history and player 1's strategy is measurable with respect to the public history, her type and her past action choices.

In this setting, player 1's Stackelberg commitment payoff, $1 - \gamma^* \theta$, has two interpretations. First, as in the baseline model, it is her equilibrium payoff when she can commit. Second, it is her highest equilibrium payoff in the repeated complete information game when her past action choices are perfectly monitored. In contrast, in the complete information repeated game (i.e. θ is common knowledge) with imperfect monitoring of past actions, her highest equilibrium payoff is $1 - \theta$, which is strictly lower.

The characterization of the patient long-run player's equilibrium payoffs in Theorem 3.1 also applies in this setting. Intuitively, this is because one can substitute player 1's mixed action in the baseline model with a deterministic effort level (equals to the probability of high effort). Given the two interpretations of the Stackelberg commitment payoff, my characterization result implies that persistent private information can overcome the lack-of-commitment problem and/or the imperfect monitoring problem, which enables the patient long-run player to achieve the same payoff as if she can commit or when her actions are perfectly monitored.

Appendix A

Appendix to Chapter 1

A.1 Proof of Theorem 1.1, Statements 1 & 3

I use $a_1^* \in A_1$ to denote the Dirac measure on a_1^*, so when α_1^* is pure, I will replace α_1^* with a_1^*. Recall that $\mathrm{BR}_2(a_1^*, \theta) \equiv \{a_2^*\}$ (or $\mathrm{BR}_2(\alpha_1^*, \theta) \equiv \{a_2^*\}$). Since $\Lambda(a_1^*, \theta) = \{\varnothing\}$ (or $\underline{\Lambda}(\alpha_1^*, \theta) = \{\varnothing\}$) if $\mathrm{BR}_2(a_1^*, \phi_{a_1^*}) \neq \{a_2^*\}$ (or $\mathrm{BR}_2(\alpha_1^*, \phi_{\alpha_1^*}) \neq \{a_2^*\}$), in which case statement 1 (or statement 3) is void. Therefore, it is without loss of generality to assume that $\mathrm{BR}_2(a_1^*, \phi_{a_1^*}) = \{a_2^*\}$ (or $\mathrm{BR}_2(\alpha_1^*, \phi_{\alpha_1^*}) = \{a_2^*\}$).

A.1.1 Proof of Statement 1

When $\Omega^m = \{a_1^*\}$ and $\lambda \in \Lambda(a_1^*, \theta)$, for every $\tilde{\mu}$ with $\tilde{\mu}(\tilde{\theta}) \in [0, \mu(\tilde{\theta})]$ for all $\tilde{\theta} \in \Theta$, we have:

$$\{a_2^*\} = \arg\max_{a_2 \in A_2} \Big\{ \mu(a_1^*) u_2(\phi_{a_1^*}, a_1^*, a_2) + \sum_{\tilde{\theta} \in \Theta} \tilde{\mu}(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2) \Big\}.$$

Let h_*^t be the period t public history such that a_1^* is always played. Let $q_t(\omega)$ be the (ex ante) probability that the history is h_*^t and player 1's type is $\omega \in \Omega$. By definition, $q_t(a_1^*) = \mu(a_1^*)$ for all t. Player 2's maximization problem at h_*^t is:

$$\max_{a_2 \in A_2} \Big\{ \mu(a_1^*) u_2(\phi_{a_1^*}, a_1^*, a_2) + \sum_{\tilde{\theta} \in \Theta} \Big[q_{t+1}(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2) + (q_t(\tilde{\theta}) - q_{t+1}(\tilde{\theta})) u_2(\tilde{\theta}, \alpha_{1,t}(\tilde{\theta}), a_2) \Big] \Big\}$$

where $\alpha_{1,t}(\tilde{\theta}) \in \Delta(A_1 \backslash \{a_1^*\})$ is the distribution of type $\tilde{\theta}$'s action at h_*^t conditional on it is not a_1^*.

Fixing $\mu(a_1^*)$ and given the fact that $\lambda \in \Lambda(a_1^*, \theta)$, there exists $\rho > 0$ such that a_2^* is player 2's strict best reply if

$$\sum_{\tilde{\theta} \in \Theta} q_{t+1}(\tilde{\theta}) > \sum_{\tilde{\theta} \in \Theta} q_t(\tilde{\theta}) - \rho.$$

Let $\overline{T} \equiv \lceil 1/\rho \rceil$, which is independent of δ. There exist at most $\overline{T}$ periods in which a_2^* fails to be a strict best reply conditional on a_1^* has always been played. Therefore, type θ's payoff is bounded from below by:

$$(1-\delta^{\overline{T}})\min_{a\in A}u_1(\theta,a)+\delta^{\overline{T}}v_\theta(a_1^*),$$

which converges to $v_\theta(a_1^*)$ as $\delta\to 1$.

When there are other commitment types, let $\overline{p}\equiv\max_{\alpha_1\in\Omega^m\backslash\{a_1^*\}}\alpha_1(a_1^*)$. which is strictly below 1. There exists $T\in\mathbb{N}$, such that for every $t\geq T$, a_2^* is player 2's strict best reply at h_*^t if: $\sum_{\tilde{\theta}\in\Theta}q_{t+1}(\tilde{\theta})\geq\sum_{\tilde{\theta}\in\Theta}q_t(\tilde{\theta})-\rho/2$. Consider the subgame starting from history h_*^T, we obtain the commitment payoff bound.

A.1.2 Proof of Statement 3

Notation: For every $\alpha_1\in\Omega^m\backslash\{\alpha_1^*\}$, $\theta\in\Theta$ and $\tilde{\mu}\in\Delta(\Omega)$ with $\tilde{\mu}(\alpha_1^*)\neq 0$, let

$$\tilde{\lambda}(\alpha_1)\equiv\tilde{\mu}(\alpha_1)/\tilde{\mu}(\alpha_1^*)\text{ and }\tilde{\lambda}(\theta)\equiv\tilde{\mu}(\theta)/\tilde{\mu}(\alpha_1^*)$$

Abusing notation, let $\tilde{\lambda}\equiv\Big(\big(\tilde{\lambda}(\alpha_1)\big)_{\alpha_1\in\Omega^m\backslash\{\alpha_1^*\}},\big(\tilde{\lambda}(\theta)\big)_{\theta\in\Theta}\Big)$ be the (expanded) likelihood ratio vector. Let $n\equiv|A_1|$ and $m\equiv|\Omega|-1$. For convenience, I write $\Omega\backslash\{\alpha_1^*\}\equiv\{\omega_1,...,\omega_m\}$ and $\tilde{\lambda}\equiv(\tilde{\lambda}_1,...,\tilde{\lambda}_m)$. The proof is divided into two parts.

Part I

Let Σ_2 be the set of player 2's strategies with σ_2 a typical element. Let $\text{NE}_2(\mu,\phi)\subset\Sigma_2$ be the set of player 2's Nash equilibrium strategies for some $\delta\in(0,1)$, i.e.

$$\text{NE}_2(\mu,\phi)\equiv\Big\{\sigma_2\Big|\exists\ \delta\in(0,1)\text{ such that }(\sigma_1,\sigma_2)\in\text{NE}(\delta,\mu,\phi)\Big\}.$$

For every $\sigma_\omega:\mathcal{H}\to\Delta(A_1)$ and player 2's strategy σ_2, let $\mathcal{P}^{(\sigma_\omega,\sigma_2)}$ be the probability measure over $\mathcal{H}$ induced by (σ_ω,σ_2), let $\mathcal{H}^{(\sigma_\omega,\sigma_2)}$ be the set of histories that occur with positive probability under $\mathcal{P}^{(\sigma_\omega,\sigma_2)}$ and let $\mathbb{E}^{(\sigma_\omega,\sigma_2)}$ be its expectation operator. Abusing notation, I use α_1^* to denote the strategy of always playing α_1^*.

For every $\psi\equiv(\psi_1,...\psi_m)\in\mathbb{R}_+^m$ and $\chi\geq 0$, let

$$\underline{\Delta}(\psi,\chi)\equiv\{\tilde{\lambda}|\sum_{i=1}^m\tilde{\lambda}_i/\psi_i=\chi\}$$

Let λ be the likelihood ratio vector induced by player 2's prior belief μ. Let $\lambda(h^t)\equiv(\lambda_1(h^t),...,\lambda_m(h^t))$ be the likelihood ratio vector following history h^t. For every infinite history h^∞, let h_t^∞ be its projection on $a_{1,t}$. Let $\alpha_1(\cdot|h^t)$ be player 2's conditional expectation over player 1's next period action at history h^t. I show the following Proposition:

Proposition A.1. *For every $\chi>0$, $\lambda\in\underline{\Delta}(\psi,\chi)$, $\sigma_2\in NE_2(\mu,\phi)$ and $\epsilon>0$, there exist $\overline{\delta}\in(0,1)$ and $T\in\mathbb{N}$ such that for every $\delta>\overline{\delta}$, there exists $\sigma_\omega:\mathcal{H}\to\Delta(A_1)$ that satisfies:*

1. *$\lambda(h^t)\in\bigcup_{\tilde{\chi}\in[0,\chi+\epsilon)}\underline{\Delta}(\psi,\tilde{\chi})$ for every $h^t\in\mathcal{H}^{(\sigma_\omega,\sigma_2)}$.*

2. For every $h^\infty \in \mathcal{H}^{(\sigma_\omega,\sigma_2)}$ and every $a_1 \in A_1$,

$$\Big|\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{h_t^\infty = a_1\} - \alpha_1^*(a_1)\Big| < \frac{\epsilon}{2(2\chi+\epsilon)} \tag{A.1}$$

3.

$$\mathbb{E}^{(\sigma_\omega,\sigma_2)}\Big[\#\Big\{t\Big|d\big(\alpha_1^*\|\alpha_1(\cdot|h^t)\big) > \epsilon^2/2\Big\}\Big] < T \tag{A.2}$$

Intuitively, Proposition A.1 demonstrates the existence of a strategy for player 1 (for every equilibrium strategy of player 2) such that the following three goals can be achieved simultaneously: first, inducing favorable beliefs about the state; second, the occupation measure of actions is closely matched to α_1^*; and third, the expected number of periods in which player 2's believed action differs significantly from α_1^* is uniformly bounded from above by an integer independent of δ. My proof follows three steps, which is the same as the description in Subsection 1.4.3.

Step 1: Let $A_1^* \equiv \text{supp}(\alpha_1^*)$. Recall that $\mathcal{P}^{(\alpha_1^*,\sigma_2)}$ is the probability measure over $\mathcal{H}$ induced by the commitment type that always plays α_1^*.

Let $\chi(h^t) \equiv \sum_{i=1}^m \lambda_i(h^t)/\psi_i$. Since $\lambda \in \underline{\Lambda}(\psi,\chi)$, we have $\chi(h^0) = \chi$. Using the observation that $\{\lambda_i(h^t), \mathcal{P}^{(\alpha_1^*,\sigma_2)}, \mathcal{F}^t\}_{t\in\mathbb{N}}$ is a non-negative supermartingale for every $i \in \{1,2,...,m\}$, where $\{\mathcal{F}^t\}_{t\in\mathbb{N}}$ is the filtration induced by the public history,[1] we know that $\{\chi_t, \mathcal{P}^{(\alpha_1^*,\sigma_2)}, \mathcal{F}^t\}_{t\in\mathbb{N}}$ is also a non-negative supermartingale. For every $a < b$, let $U(a,b)$ be the number of upcrossings from a to b. According to the Doob's Upcrossing Inequality (Chung 1974),

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}\Big\{U(\chi,\chi+\frac{\epsilon}{2}) \geq 1\Big\} \leq \frac{2\chi}{2\chi+\epsilon}. \tag{A.3}$$

Let $\tilde{\mathcal{H}}^\infty$ be the set of infinite histories that $\chi(h^t)$ is always below $\chi + \frac{\epsilon}{2}$. According to (A.3), it occurs with probability at least $\frac{\epsilon}{2\chi+\epsilon}$.

Step 2: In this step, I show that for large enough δ, there exists a subset of $\mathcal{H}^\infty$, which occurs with probability bounded from below by a positive number, such that the occupation measure over A_1 induced by every history in this set is $\epsilon-$close to α_1^*. For every $a_1 \in A_1^*$, let $\{X_t\}$ be a sequence of i.i.d. random variables such that:

$$X_t = \begin{cases} 1 & \text{when } a_{1,t} = a_1 \\ 0 & \text{otherwise .} \end{cases}$$

Under $\mathcal{P}^{(\alpha_1^*,\sigma_2)}$, $X_t = 1$ with probability $\alpha_1^*(a_1)$, so X_t has mean $\alpha_1^*(a_1)$ and variance $\sigma^2 \equiv \alpha_1^*(a_1)(1-\alpha_1^*(a_1))$. Recall that $n = |A_1|$. I start with the following Lemma:

[1] When α_1^* has full support, $\{\lambda_i(h^t), \mathcal{P}^{(\alpha_1^*,\sigma_2)}, \mathcal{F}^t\}_{t\in\mathbb{N}}$ is a martingale. However, when $A_1^* \neq A_1$ and type ω_i plays action $a_1' \notin A_1^*$ with positive probability, then the expected value of $\lambda_i(h^t)$ can strictly reduce.

Lemma A.1.1. *For any $\varepsilon > 0$, there exists $\overline{\delta} \in (0,1)$, such that for all $\delta \in (\overline{\delta}, 1)$,*

$$\limsup_{\delta \to 1} \mathcal{P}^{(\alpha_1^*, \sigma_2)} \Big(\Big| \sum_{t=0}^{+\infty} (1-\delta)\delta^t X_t - \alpha_1^*(a_1) \Big| \geq \varepsilon \Big) \leq \frac{\varepsilon}{n}. \tag{A.4}$$

Proof of Lemma A.1.1: For every $n \in \mathbb{N}$, let $\hat{X}_n \equiv \delta^n (X_n - \alpha_1^*(a_1))$. Define a triangular sequence of random variables $\{X_{k,n}\}_{0 \leq n \leq k, k, n \in \mathbb{N}}$, such that $X_{k,n} \equiv \xi_k \hat{X}_n$, where

$$\xi_k \equiv \sqrt{\frac{1}{\sigma^2} \frac{1-\delta^2}{1-\delta^{2k}}}.$$

Let $Z_k \equiv \sum_{n=1}^k X_{k,n} = \xi_k \sum_{k=1}^n \hat{X}_n$. By the Lindeberg-Feller Central Limit Theorem (Chung 1974), Z_k converges in law to $N(0,1)$. By construction,

$$\frac{\sum_{n=1}^k \hat{X}_n}{1+\delta+...+\delta^{k-1}} = \sigma \sqrt{\frac{1-\delta^{2k}}{1-\delta^2}} \frac{1-\delta}{1-\delta^k} Z_k,$$

the RHS of this expression converges (in distribution) to a normal distribution with mean 0 and variance

$$\sigma^2 \frac{1-\delta^{2k}}{1-\delta^2} \frac{(1-\delta)^2}{(1-\delta^k)^2}.$$

The variance term converges to $\mathcal{O}\Big((1-\delta)\Big)$ as $k \to \infty$. Using Theorem 7.4.1 in Chung (1974), we have:

$$\sup_{x \in \mathbb{R}} |F_k(x) - \Phi(x)| \leq A_0 \sum_{n=1}^k |X_{k,n}|^3 \sim A_1 (1-\delta)^{\frac{3}{2}},$$

where A_0 and A_1 are constants, F_k is the empirical distribution of Z_k and $\Phi(\cdot)$ is the standard normal distribution. Both the variance and the approximation error goes to 0 as $\delta \to 1$.

Using the properties of normal distribution, we know that for every $\varepsilon > 0$, there exists $\overline{\delta} \in (0,1)$ such that for every $\delta > \overline{\delta}$, there exists $K \in \mathbb{N}$, such that for all $k > K$,

$$\mathcal{P}^{(\alpha_1^*, \sigma_2)} \Big(\Big| \frac{\sum_{i=1}^k \hat{X}_n}{1+\delta+...+\delta^{k-1}} \Big| \geq \varepsilon \Big) < \frac{\varepsilon}{n}.$$

Taking the $k \to +\infty$ limit, one can obtain the conclusion in Lemma A.1.1. □

Step 3: According to Lemma A.1.1, for every $a_1 \in A_1$ and $\epsilon > 0$, there exists $\overline{\delta} \in (0,1)$, such that for all $\delta > \overline{\delta}$, there exists $\mathcal{H}^\infty_{\epsilon, a_1}(\delta) \subset \mathcal{H}^\infty$, such that

$$\mathcal{P}^{(\alpha_1^*, \sigma_2)} (\mathcal{H}^\infty_{\varepsilon, a_1}(\delta)) \geq 1 - \varepsilon / n, \tag{A.5}$$

and for every $h^\infty \in \mathcal{H}^\infty_{\varepsilon,a_1}(\delta)$, the occupation measure of a_1 is ε-close to $\alpha_1^*(a_1)$. Let $\mathcal{H}^\infty_\varepsilon(\delta) \equiv \bigcap_{a_1 \in A_1} \mathcal{H}^\infty_{\varepsilon,a_1}(\delta)$. According to (A.5), we have:

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}^\infty_\varepsilon(\delta)) \geq 1-\varepsilon. \tag{A.6}$$

Take $\varepsilon \equiv \frac{\epsilon}{2(2\chi+\epsilon)}$ and let

$$\widehat{\mathcal{H}}^\infty \equiv \tilde{\mathcal{H}}^\infty \cap \mathcal{H}^\infty_\varepsilon(\delta), \tag{A.7}$$

we have:

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\hat{\mathcal{H}}^\infty) \geq \frac{\epsilon}{2(2\chi+\epsilon)} \tag{A.8}$$

According to Gossner (2011), we have

$$\mathbb{E}^{(\alpha_1^*,\sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^\tau))\Big] \leq -\log\mu(\alpha_1^*). \tag{A.9}$$

The Markov Inequality implies that:

$$\mathbb{E}^{(\alpha_1^*,\sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^\tau))\Big|\hat{\mathcal{H}}^\infty\Big] \leq -\frac{2(2\chi+\epsilon)\log\mu(\alpha_1^*)}{\epsilon}. \tag{A.10}$$

Let $\mathcal{P}^*$ be the probability measure over $\mathcal{H}^\infty$ such that for every $\mathcal{H}_0^\infty \subset \mathcal{H}^\infty$,

$$\mathcal{P}^*(\mathcal{H}_0^\infty) \equiv \frac{\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}_0^\infty \cap \hat{\mathcal{H}}^\infty)}{\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\hat{\mathcal{H}}^\infty)}.$$

Let $\sigma_\omega : \mathcal{H} \to \Delta(A_1)$ be player 1's strategy that induces $\mathcal{P}^*$. The expected number of periods in which $d(\alpha_1^*||\alpha(\cdot|h^t)) > \epsilon^2/2$ is bounded from above by:

$$T \equiv \Big\lceil -\frac{4(2\chi+\epsilon)\log\mu(\alpha_1^*)}{\epsilon^3}\Big\rceil, \tag{A.11}$$

which is an integer independent of δ. The three steps together establish Proposition A.1.

Part II

Proposition A.1 and $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$ do not imply that player 1 can guarantee himself his commitment payoff. This is because player 2 may not have an incentive to play a_2^* despite $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$ and the average action is close to α_1^*. My proof overcomes this challenge using two observations, which are the two steps of my proof.

1. If $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$, is small in all but at most one entry and player 1's average action is close to α_1^*, then player 2 has a strict incentive to play a_2^* regardless of the correlation. Let Λ^0 be the set of beliefs that has the above feature.

2. If $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$ and player 1's average action is close to α_1^* but player 2 does not have a strict

incentive to play a_2^*, then different types of player 1's actions must be sufficiently different. This implies that there is significant learning about player 1's type after observing his action.

I show that for every $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$, there exists an integer T and a strategy such that if player 1 picks his action according to this strategy in periods with the above feature, then after at most T such periods, player 2's belief about his type will be in Λ^0, which concludes the proof.

Recall that $m \equiv |\Omega^m| + |\Theta| - 1$. Let $\psi \equiv \{\psi_i\}_{i=1}^m \in \mathbb{R}_+^m$ be defined as:

- If $\omega_i \in \Theta^b_{(\alpha_1^*,\theta)}$, then ψ_i equals to the intercept of $\Lambda(\alpha_1^*, \theta)$ on dimension ω_i.
- Otherwise, $\psi_i > 0$ is chosen to be large enough such that

$$\sum_{i=1}^m \lambda_i/\psi_i < 1. \tag{A.12}$$

Such ψ exists given that $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$. Let $\overline{\psi} \equiv \max\{\psi_i | i = 1, 2, ..., m\}$. Recall that Part I has established the existence of a strategy for player 1 under which:

- Player 2's belief always satisfies (A.12), or more precisely, bounded from above by some $\chi < 1$.
- The occupation measure over A_1 at every on-path infinite history is ϵ-close to α_1^*.
- In expectation, there exists at most T periods in which player 2's believed action differs significantly from α_1^*, where T is independent of δ.

Step 1: For every $\xi > 0$, a likelihood ratio vector λ is of '**size ξ**' if there exists $\tilde{\psi} \equiv (\tilde{\psi}_1, ..., \tilde{\psi}_m) \in \mathbb{R}_+^m$ such that: $\tilde{\psi}_i \in (0, \psi_i)$ for all i and

$$\lambda \in \Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^m \tilde{\lambda}_i/\tilde{\psi}_i < 1\Big\} \subset \Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \#\{i | \tilde{\lambda}_i \leq \xi\} \geq m-1\Big\}. \tag{A.13}$$

Intuitively, λ is of size ξ if there exists a downward sloping hyperplane such that all likelihood ratio vectors below this hyperplane have at least $m-1$ entries that are no larger than ξ. Therefore, for every $\xi > \xi' > 0$, if λ is of size ξ', then it is also of size ξ. Proposition A.2 establishes the commitment payoff bound when λ is of size ξ for ξ small enough.

Proposition A.2. *There exists $\xi > 0$, such that if λ is of size ξ, then*

$$\liminf_{\delta \to 1} \underline{V}_\theta(\mu, \delta, \phi) \geq u_1(\theta, \alpha_1^*, a_2^*).$$

In the proof, I show that using the strategy constructed in Proposition A.1, we can ensure that a_2^* is player 2's strict best reply at every h^t where $d(\alpha_1^* || \alpha_1(\cdot|h^t)) < \epsilon^2/2$. This implies Proposition A.2.

Proof of Proposition A.2: Let $\alpha_1(\cdot|h^t,\omega_i)\in\Delta(A_1)$ be the equilibrium action of type ω_i at history h^t. Let

$$B_{i,a_1}(h^t)\equiv\lambda_i(h^t)\Big(\alpha_1^*(a_1)-\alpha_1(a_1|h^t,\omega_i)\Big). \tag{A.14}$$

Recall that

$$\alpha_1(\cdot|h^t)\equiv\frac{\alpha_1^*+\sum_{i=1}^m\lambda_i(h^t)\alpha_1(\cdot|h^t,\omega_i)}{1+\sum_{i=1}^m\lambda_i(h^t)}.$$

is the average action anticipated by player 2. For any $\lambda\in\underline{\Lambda}(\alpha_1^*,\theta)$ and $\epsilon>0$, there exists $\varepsilon>0$ such that at every likelihood ratio vector $\widetilde{\lambda}$ satisfying:

$$\sum_{i=1}^m\widetilde{\lambda}_i/\psi_i<\frac{1}{2}\Big(1+\sum_{i=1}^m\lambda_i/\psi_i\Big), \tag{A.15}$$

a_2^* is player 2's strict best reply to every $\{\alpha_1(\cdot|h^t,\omega_i)\}_{i=1}^m$ satisfying the following two conditions

- $|B_{i,a_1}(h^t)|<\varepsilon$ for all i and a_1.
- $\big\|\alpha_1^*-\alpha_1(\cdot|h^t)\big\|\leq\epsilon$.

This is because when the prior belief satisfies (A.15), a_2^* is player 2's strict best reply when all types of player 1 are playing α_1^*. When ϵ and ε are both small enough, an ϵ-deviation of the average action together with an ε correlation between types and actions cannot overturn this strictness.

According to the Pinsker's Inequality, $\big\|\alpha_1^*-\alpha_1(\cdot|h^t)\big\|\leq\epsilon$ is implied by $d(\alpha_1^*||\alpha_1(\cdot|h^t))\leq\epsilon^2/2$. Pick ϵ and ξ small enough such that:

$$\epsilon<\frac{\varepsilon}{2(1+\overline{\psi})} \tag{A.16}$$

and

$$\xi<\frac{\varepsilon}{(m-1)(1+\varepsilon)}. \tag{A.17}$$

Suppose $\lambda_i(h^t)\leq\xi$ for all $i\geq2$, since $\big\|\alpha_1^*-\alpha_1(\cdot|h^t)\big\|\leq\epsilon$, we have:

$$\frac{\Big\|\lambda_1(\alpha_1^*-\alpha_1(a_1|h^t,\omega_1))+\sum_{i=2}^m\lambda_i\big(\alpha_1^*-\alpha_1(a_1|h^t,\omega_i)\big)\Big\|}{1+\lambda_1+\xi(m-1)}\leq\epsilon.$$

The triangular inequality implies that:

$$\begin{aligned}\Big\|\lambda_1(\alpha_1^*-\alpha_1(a_1|h^t,\omega_1))\Big\| &\leq \sum_{i=2}^m\Big\|\lambda_i(\alpha_1^*-\alpha_1(a_1|h^t,\omega_i))\Big\|+\epsilon\Big(1+\lambda_1+\xi(m-1)\Big)\\ &\leq \xi(m-1)+\epsilon\Big(1+\overline{\psi}+\xi(m-1)\Big)\leq\varepsilon. \end{aligned} \tag{A.18}$$

where the last inequality uses (A.16) and (A.17). Inequality (A.18) implies that $||B_{1,a_1}(h^t)||\leq\varepsilon$, and therefore, when λ is of size ξ, a_2^* is player 2's strict best reply at every history where $d(\alpha_1^*||\alpha_1(\cdot|h^t))\leq\epsilon^2/2$. This further implies that the commitment payoff bound is guaranteed. □

Step 2: In this step, I use Proposition A.2 to show that the mixed commitment payoff is guaranteed for every λ satisfying (A.12). Recall the definition of $B_{i,a_1}(h^t)$ in (A.14). According to Bayes Rule, if $a_1 \in A_1^*$ is observed at h^t, then

$$\lambda_i(h^t) - \lambda_i(h^t, a_1) = \frac{B_{i,a_1}(h^t)}{\alpha_1^*(a_1)} \quad \text{and} \sum_{a_1 \in A_1^*} \alpha_1^*(a_1)\Big(\lambda_i(h^t) - \lambda_i(h^t, a_1)\Big) \geq 0.$$

Let

$$D(h^t, a_1) \equiv \Big(\lambda_i(h^t) - \lambda_i(h^t, a_1)\Big)_{i=1}^m \in \mathbb{R}^m.$$

Suppose $B_{i,a_1}(h^t) \geq \varepsilon$ for some i and $a_1 \in A_1^*$, then $||D(h^t, a_1)|| \geq \varepsilon$ where $|| \cdot ||$ denotes the $\mathcal{L}^2$-norm. Pick $\xi > 0$ small enough to meet the requirement in Proposition A.2. Define two sequences of sets, $\{\Lambda^k\}_{k=0}^{\infty}$ and $\{\widehat{\Lambda}^k\}_{k=1}^{\infty}$, which satisfy $\Lambda^k, \widehat{\Lambda}^k \subset \underline{\Lambda}(\alpha_1^*, \theta)$ for all $k \in \mathbb{N}$, recursively as follows:

- Let Λ^0 be the set of likelihood ratio vectors that are of size ξ,
- For every $k \geq 1$, let $\widehat{\Lambda}^k$ be the set of likelihood ratio vectors in $\underline{\Lambda}(\alpha_1^*, \theta)$ such that if $\lambda(h^t) \in \widehat{\Lambda}^k$, then either $\lambda(h^t) \in \Lambda^{k-1}$ or, For every $\{\alpha_1(\cdot|h^t, \omega_i)\}_{i=1}^m$ such that $||D(h^t, a_1)|| \geq \varepsilon$ for some $a_1 \in A_1^*$, there exists $a_1^* \in A_1^*$ such that $\lambda(h^t, a_1^*) \in \Lambda^{k-1}$.
- Let Λ^k be the set of likelihood ratio vectors in $\underline{\Lambda}(\alpha_1^*, \theta)$ such that for every $\tilde{\lambda} \in \Lambda^k$, there exists $\tilde{\psi} \equiv (\tilde{\psi}_1, ..., \tilde{\psi}_m) \in \mathbb{R}_+^m$ such that: $\tilde{\psi}_i \in (0, \psi_i)$ for all i and

$$\lambda \in \Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^m \tilde{\lambda}_i / \tilde{\psi}_i < 1\Big\} \subset \Big(\bigcup_{j=0}^{k-1} \Lambda^j\Big) \bigcup \widehat{\Lambda}^k. \tag{A.19}$$

  By construction, we know that:

$$\Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^m \tilde{\lambda}_i / \tilde{\psi}_i < 1\Big\} \subset \bigcup_{j=0}^{k} \Lambda^j = \Lambda^k. \tag{A.20}$$

Since $(0, ..., \psi_i - \upsilon, ..., 0) \in \Lambda^0$ for any $i \in \{1, 2, ..., m\}$ and $\upsilon > 0$, so $\text{co}(\Lambda^0) = \underline{\Lambda}(\alpha_1^*, \theta)$. By definition, $\{\Lambda^k\}_{k \in \mathbb{N}}$ is an increasing sequence with $\Lambda^k \subset \underline{\Lambda}(\alpha_1^*, \theta) = \text{co}(\Lambda^k)$ for any $k \in \mathbb{N}$, i.e. it is bounded from above by a compact set. Therefore $\lim_{k \to \infty} \bigcup_{j=0}^k \Lambda^j \equiv \Lambda^\infty$ exists and is a subset of $\text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$. The next Lemma shows that $\text{clo}(\Lambda^\infty)$ coincides with $\text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$.

Lemma A.1.2. *$clo(\Lambda^\infty) = clo\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$*

Proof of Lemma A.1.2: Since $\Lambda^k \subset \underline{\Lambda}(\alpha_1^*, \theta)$ for every $k \in \mathbb{N}$, it is obvious that $\text{clo}(\Lambda^\infty) \subset \text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$. Suppose towards a contradiction that

$$\text{clo}(\Lambda^\infty) \subsetneq \text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big) \tag{A.21}$$

- Let $\widehat{\Lambda} \subset \underline{\Lambda}(\alpha_1^*, \theta)$ be such that if $\lambda(h^t) \in \widehat{\Lambda}$, then either $\lambda(h^t) \in \Lambda^\infty$ or:

- For every $\{\alpha_1(\cdot|h^t,\omega_i)\}_{i=1}^m$ such that $||D(h^t,a_1)|| \geq \varepsilon$ for some $a_1 \in A_1^*$, there exists $a_1^* \in A_1^*$ such that $\lambda(h^t,a_1^*) \in \Lambda^\infty$.

- Let $\breve{\Lambda}$ be the set of likelihood ratio vectors in $\underline{\Lambda}(\alpha_1^*,\theta)$ such that for every $\tilde{\lambda} \in \breve{\Lambda}$, there exists $\tilde{\psi} \equiv (\tilde{\psi}_1,...,\tilde{\psi}_m) \in \mathbb{R}_+^m$ such that: $\tilde{\psi}_i \in (0,\psi_i)$ for all i and

$$\lambda \in \left\{\tilde{\lambda} \in \mathbb{R}_+^m \middle| \sum_{i=1}^m \tilde{\lambda}_i/\tilde{\psi}_i < 1\right\} \subset \left(\Lambda^\infty \bigcup \widehat{\Lambda}\right). \tag{A.22}$$

Since Λ^∞ is defined as the limit of the above operator, so in order for (C.21) to be true, it has to be the case that $\breve{\Lambda} = \Lambda^\infty$, or $\Xi \cap \breve{\Lambda} = \{\varnothing\}$ where

$$\Xi \equiv \text{clo}\Big(\underline{\Lambda}(\alpha_1^*,\theta)\Big)\Big\backslash \text{clo}(\Lambda^\infty). \tag{A.23}$$

One can check that Ξ is convex and has non-empty interior. For every $\varrho > 0$, there exists $x \in \Xi$, $\theta \in (0,\pi/2)$ and a halfspace: $H(\chi) \equiv \left\{\tilde{\lambda} \middle| \sum_{i=1}^m \tilde{\lambda}_i/\chi_i \leq \chi\right\}$ with $\phi > 0$ satisfying:

1. $\sum_{i=1}^m x_i/\psi_i = \chi$.

2. $\partial B(x,r) \bigcap H(\chi) \bigcap \underline{\Lambda}(\alpha_1^*,\theta) \subset \Lambda^\infty$ for every $r \geq \varrho$.

3. For every $r \geq \rho$ and $y \in \partial B(x,r) \cap \underline{\Lambda}(\alpha_1^*,\theta)$, either $y \in \Lambda^\infty$ or $d(y,H(\chi)) > r\sin\theta$, where $d(\cdot,\cdot)$ denotes the Hausdorff distance.

The second and third property used the non-convexity of $\text{clo}(\Lambda^\infty)$. Suppose $\lambda(h^t) = x$ for some h^t and there exists $a_1 \in A_1^*$ such that $||D(h^t,a_1)|| \geq \varepsilon$,

- Either $\lambda(h^t,a_1) \in \Lambda^\infty$, in which case $x \in \breve{\Lambda}$ but $x \in \Xi$, leading to a contradiction.

- Or $\lambda(h^t,a_1) \notin \Lambda^\infty$. Requirement 3 implies that $d(\lambda(h^t,a_1),H(\chi)) > \varepsilon\sin\theta$. On the other hand,

$$\sum_{a_1' \in A_1^*} \alpha_1^*(a_1')\lambda_i(h^t,a_1') \leq \lambda_i(h^t) \tag{A.24}$$

  for every i. Requirement 1 then implies that $\sum_{a_1' \in A_1^*} \alpha_1^*(a_1')\lambda_i(h^t,a_1') \in H(\chi)$, which is to say:

$$\sum_{a_1' \in A_1^*} \alpha_1^*(a_1') \sum_{i=1}^m \lambda_i(h^t,a_1')/\psi_i \leq \chi. \tag{A.25}$$

  According to Requirement 2, $\lambda(h^t,a_1) \notin H(\chi)$, i.e. $\sum_{i=1}^m \lambda_i(h^t,a_1)/\psi_i > \chi + \varepsilon\kappa$ for some constant $\kappa > 0$. Take

$$\rho \equiv \frac{1}{2}\min_{a_1 \in A_1^*}\{\alpha_1^*(a_1)\}\varepsilon\kappa,$$

  (A.24) implies the existence of $a_1^* \in A_1^* \backslash \{a_1\}$ such that $\lambda(h^t,a_1^*) \in H(\chi) \cap B(x,\rho)$. Requirement 2 then implies that $x = \lambda(h^t) \in \breve{\Lambda}$. Since $x \in \Xi$, this leads to a contradiction.

Therefore, (A.21) cannot be true, which validates the conclusion of Lemma A.1.2. $\square$

Lemma A.1.2 implies that for every $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$, there exists an integer $K \in \mathbb{N}$ independent of δ such that $\lambda \in \Lambda^K$. Statement 3 can then be shown by induction on K. According to Proposition A.2, the statement holds when $K = 0$. Suppose the statement applies to every $K \leq K^* - 1$, let us consider the case when $K = K^*$. According to the construction of Λ^{K^*}, there exists a strategy for player 1 such that whenever a_2^* is not player 2's best reply despite $d(\alpha_1^* \| \alpha_1(\cdot|h^t)) < \epsilon^2/2$, then the posterior belief after observing $a_{1,t}$ is in Λ^{K^*-1}, under which the commitment payoff bound is attained by the induction hypothesis.

A.2 Proof of Theorem 1.1, Statement 2

In this Appendix, I prove statement 2. The proof of statement 4 involves some additional technical complication, which is relegated to Online Appendix B. The key intuition behind the distinction of pure and mixed commitment strategies in the construction of low payoff equilibria is summarized in Proposition B.3 and Proposition B.6 in Online Appendix B.

In this section, I replace α_1^* with a_1^*. Let $\overline{\Pi}(a_1^*, \theta)$, $\Pi(a_1^*, \theta)$ and $\underline{\Pi}(a_1^*, \theta)$ be the exteriors of $\overline{\Lambda}(a_1^*, \theta)$, $\Lambda(a_1^*, \theta)$ and $\underline{\Lambda}(a_1^*, \theta)$, respectively. I start with the following Lemma, which clarifies the restriction that $\mathrm{BR}_2(a_1^*, \phi_{a_1^*})$ being a singleton.

Lemma A.2.1. *For every $\lambda \in \Pi(a_1^*, \theta)$, there exist $0 \ll \lambda' \ll \lambda$ and $a_2' \neq a_2^*$ such that $\lambda' \in \overline{\Pi}(a_1^*, \theta)$ and*

$$\sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) \Big(u_2(\tilde{\theta}, a_1^*, a_2') - u_2(\tilde{\theta}, a_1^*, a_2^*) \Big) > 0 \tag{A.26}$$

if either one the following three conditions hold:

1. *$\Lambda(a_1^*, \theta) \neq \{\varnothing\}$.*
2. *$\Lambda(a_1^*, \theta) = \{\varnothing\}$ and $BR_2(a_1^*, \phi_{a_1^*})$ is a singleton.*
3. *$\Lambda(a_1^*, \theta) = \{\varnothing\}$ and $a_2^* \notin BR_2(a_1^*, \phi_{a_1^*})$.*

Proof of Lemma A.2.1: When $\Lambda(a_1^*, \theta) \neq \{\varnothing\}$, by definition of $\Pi(a_1^*, \theta)$, there exists $0 \ll \lambda' \ll \lambda$ and $a_2' \neq a_2^*$ such that:

$$\Big(u_2(\phi_{a_1^*}, a_1^*, a_2') - u_2(\phi_{a_1^*}, a_1^*, a_2^*) \Big) + \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) \Big(u_2(\tilde{\theta}, a_1^*, a_2') - u_2(\tilde{\theta}, a_1^*, a_2^*) \Big) > 0. \tag{A.27}$$

But $\Lambda(a_1^*, \theta) \neq \{\varnothing\}$ implies that $\{a_2^*\} = \mathrm{BR}_2(a_1^*, \phi_{a_1^*})$, so (A.27) implies (A.26).

When $\Lambda(a_1^*, \theta) = \{\varnothing\}$, if $\mathrm{BR}_2(a_1^*, \phi_{a_1^*})$ is a singleton, then $\mathrm{BR}_2(a_1^*, \phi_{a_1^*}) \neq \{a_2^*\}$. Therefore, under condition 2 or 3, $a_2^* \notin \mathrm{BR}_2(a_1^*, \phi_{a_1^*})$, which implies the existence of $\theta' \neq \theta$ and $a_2' \neq a_2^*$ such that $u_2(\theta', a_1^*, a_2') > u_2(\theta', a_1^*, a_2^*)$. By definition, $\theta' \in \Theta^b_{(a_1^*,\theta)}$. Let

$$\lambda'(\tilde{\theta}) \equiv \begin{cases} \lambda(\tilde{\theta}) & \text{if } \tilde{\theta} = \theta' \\ 0 & \text{otherwise} \end{cases}$$

λ' satisfies (A.26) since $\mu(\omega) > 0$ for every $\omega \in \Omega$. □

Remark: Lemma A.2.1 leaves out the case in which $\Lambda(a_1^*, \theta) = \{\varnothing\}$ and $a_2^* \in \mathrm{BR}_2(a_1^*, \phi_{a_1^*})$. In this pathological case, whether player 1 can guarantee his commitment payoff or not depends on the presence of other commitment types. For example, when $\Theta = \{\theta, \theta'\}$, $A_1 = \{a_1^*, a_1'\}$, $A_2 = \{a_2^*, a_2'\}$ and $\Omega^m = \{a_1^*, (1-\epsilon)a_1^* + \epsilon a_1'\}$ with $\phi_{a_1^*}(\theta') = 1$ and $\phi_{(1-\epsilon)a_1^*+\epsilon a_1'}(\theta) = 1$. Suppose $\{a_2^*\} = \mathrm{BR}_2(a_1^*, \theta) = \mathrm{BR}_2(a_1', \theta)$ and $\{a_2^*, a_2'\} = \mathrm{BR}_2(a_1^*, \theta') = \mathrm{BR}_2(a_1', \theta')$. Then type θ can guarantee himself payoff $u_1(\theta, a_1^*, a_2^*)$ by always playing a_1^* even though $\lambda \in \Pi(a_1^*, \theta)$ since a_1' is always player 2's strictly best reply given the presence of commitment type playing $(1-\epsilon)a_1^* + \epsilon a_1'$.

Overview of Two Phase Construction: Let player 1's payoff function be:

$$u_1(\tilde{\theta}, a_1, a_2) \equiv \mathbf{1}\{\tilde{\theta} = \theta, a_1 = a_1^*, a_2 = a_2^*\}. \tag{A.28}$$

By definition, $v_\theta(a_1^*) = 1$. The sequential equilibrium I construct has a '*normal phase*' and an '*abnormal phase*'. Type θ's equilibrium action is pure at every history occurring with positive probability under $(\sigma_\theta, \sigma_2)$. Play starts from the normal phase and remains in it as long as all past actions equal to type θ's equilibrium actions. Otherwise, play switches to the abnormal phase and stays there forever.

Let $A_1 \equiv \{a_1^0, ..., a_1^{n-1}\}$. I show there exists a constant $q \in (0,1)$ (independent of δ) such that:

- After a bounded number of periods (uniform for all δ), type θ obtains expected payoff $1-q$ in every period in the normal phase, i.e. his payoff is approximately $1-q$ when $\delta \to 1$.
- Type θ's continuation payoff is bounded below $1-2q$ in the beginning of the abnormal phase.

Strategies in the Normal Phase: Let $\Theta_{(a_1^*,\theta)} \equiv \Theta \backslash \Theta^b_{(a_1^*,\theta)}$, which are the set of good strategic types.

- **'Mechanical' Strategic Types:** Every strategic type in $\Theta_{(a_1^*,\theta)} \backslash \{\theta\}$ plays $\alpha_1 \in \Omega^m \backslash \{a_1^*\}$ forever, with α_1 being arbitrarily chosen.[2] For every strategic type $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$, he plays α_1 forever with probability $x(\tilde{\theta}) \in [0,1]$ such that conditional on player 2 knowing that
  - player 1 is either a bad strategic type who is *not* always playing α_1; or he is the commitment type that is always playing a_1^*.

[2] If $\Omega^m = \{a_1^*\}$, then all types in $\Theta_{(a_1^*,\theta)} \backslash \{\theta\}$ play some arbitrarily chosen $a_1' \neq a_1^*$.

the likelihood ratio vector induced by her belief equals to λ', with λ' being defined in Lemma A.2.1.

In what follows, I treat the strategic types that are always playing α_1 as the commitment type that is playing α_1. Formally, let

$$\tilde{\Omega}^m \equiv \begin{cases} \{a_1'\} & \text{if } |\Omega^m| = 1 \\ \Omega^m \backslash \{a_1^*\} & \text{otherwise .} \end{cases}$$

Let $l \equiv |\tilde{\Omega}^m|$. By construction, we have $l \geq 1$. Let $\tilde{\phi}_{\alpha_1} \in \Delta(\Theta)$ be the adjusted distribution conditional on player 1 being either commitment type α_1 or strategic type $\tilde{\theta} \in \Theta_{(a_1^*,\theta)} \backslash \{\theta\}$ that always plays α_1.

- **Other Bad Strategic Types:** Conditional on not always playing α_1, type $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$ plays a_1^* forever with probability $p \in [0,1)$, with p chosen such that there exists $a_2' \neq a_2^*$ with

$$u_2(\phi_{a_1^*}, a_1^*, a_2') + \tilde{p} \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2') > u_2(\phi_{a_1^*}, a_1^*, a_2^*) + \tilde{p} \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2^*) \tag{A.29}$$

for all $\tilde{p} \in [p, 1]$. According to the construction of λ', Lemma A.2.1 also implies that

$$\sum_{\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2') > \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2^*). \tag{A.30}$$

For every $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$, type $\tilde{\theta}$ plays a_1^* forever with probability p, he plays $\alpha_1 \in \tilde{\Omega}^m$ in the normal phase with probability $\frac{1-p}{l}$.

Call the bad strategic type(s) who always play $\alpha_1 \in \tilde{\Omega}^m \cup \{a_1^*\}$ in the normal phase *type* $\theta(\alpha_1)$. Let $\mu_t(\theta(\alpha_1))$ be the total probability of such type in period t. By construction, throughout the normal phase, if $\mu_t(\alpha_1) = 0$, then $\mu_t(\theta(\alpha_1)) = 0$; if $\mu_t(\alpha_1) \neq 0$, then $\mu_t(\theta(\alpha_1))/\mu_t(\alpha_1) = \mu_0(\theta(\alpha_1))/\mu_0(\alpha_1)$.

Next, I describe type θ's normal phase strategy:

1. **Preparation Sub-Phase:** This phase lasts from period 0 to $n-1$. Type θ plays a_1^i in period i for all $i \in \{0, 1, ..., n-1\}$. This is to rule out all pure strategy commitment types.

2. **Value Delivery Sub-Phase:** This phase starts from period n. Type θ plays either a_1^* or some $a_1' \neq a_1^*$, depending on the realization of ξ_t. The probability that a_1^* being prescribed is q.

I claim that type θ's expected payoff is close to $1-q$ if he plays type θ's equilibrium strategy when δ is sufficiently close to 1. This is because in the normal phase:

- After period n, player 2 attaches probability 0 to all pure strategy commitment types.

- Starting from period n, whenever player 2 observes player 1 playing his equilibrium action, there exists $\varrho > 1$ such that:

$$\mu_{t+1}(\theta)\Big/\Big(\mu_{t+1}(\alpha_1)+\mu_{t+1}(\theta(\alpha_1))\Big) \geq \varrho\mu_t(\theta)\Big/\Big(\mu_t(\alpha_1)+\mu_t(\theta(\alpha_1))\Big). \tag{A.31}$$

for every $\alpha_1 \in \tilde{\Omega}^m$ satisfying $\mu_t(\alpha_1) \neq 0$.

So there exists $T \in \mathbb{N}$ independent of δ such that in period $t \geq T$, a_2^* is player 2's strict best reply conditional on ξ_t prescribing a_1^* and play remains in the normal phase. Therefore, type θ's expected payoff at every normal phase information set must be within the following interval:

$$\Big[(1-\delta^T)0+\delta^T(1-q),(1-\delta^T)+\delta^T(1-q)\Big].$$

Both the lower and the upper bound of this interval will converge to $1-q$ as $\delta \to 1$.

Strategies in the Abnormal Phase: In the abnormal phase, player 2 has ruled out the possibility that player 1 is type θ. Type $\theta(a_1^*)$'s strategy also remains the same. For every $\alpha_1 \in \tilde{\Omega}^m$, type $\theta(\alpha_1)$ plays:

$$\hat{\alpha}_1(\alpha_1) \equiv (1-\frac{\eta}{2})a_1^* + \frac{\eta}{2}\tilde{\alpha}_1(\alpha_1)$$

where:

$$\tilde{\alpha}_1(\alpha_1)[a_1] \equiv \begin{cases} 0 & \text{when } a_1 = a_1^* \\ \alpha_1(a_1)/(1-\alpha_1(a_1^*)) & \text{otherwise .} \end{cases}$$

I choose $\eta > 0$ such that $\max_{\alpha_1 \in \tilde{\Omega}^m} \alpha_1(a_1^*) < 1-\eta$, and for every $\alpha_1' \in \Delta(A_1)$ satisfying $\alpha_1'(a_1^*) \geq 1-\eta$, we have:

$$\sum_{\tilde{\theta}\in\Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta})u_2(\tilde{\theta},\alpha_1',a_2') > \sum_{\tilde{\theta}\in\Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta})u_2(\tilde{\theta},\alpha_1',a_2^*).$$

Such η exists according to (A.30).

Next, I verify that type θ has no incentive to trigger the abnormal phase. Instead of explicitly constructing his abnormal phase strategy, I compute an upper bound on his payoff in the beginning of the abnormal phase. Let $\beta(\alpha_1) \equiv \mu_t(\theta(\alpha_1))/\mu_t(\alpha_1)$. Since $\max_{\alpha_1 \in \tilde{\Omega}^m} \alpha_1(a_1^*) < 1-\eta$, so if a_1^* is observed in period t,

$$\beta_{t+1}(\alpha_1) \geq \frac{1-\eta/2}{1-\eta}\beta_t(\alpha_1),$$

for every $\alpha_1 \in \tilde{\Omega}^m$. Let $\gamma \equiv 1-\min_{\alpha_1 \in \tilde{\Omega}^m} \alpha_1(a_1^*)$. If $a_1 \neq a_1^*$ is observed in period t, by definition of $\tilde{\alpha}_1(\alpha_1)$,

$$\beta_{t+1}(\alpha_1) \geq \frac{\eta}{2\gamma}\beta_t(\alpha_1).$$

Let $\overline{k} \equiv \left\lceil \log \frac{2\gamma}{\eta} \Big/ \log \frac{1-\eta/2}{1-\eta} \right\rceil$. For every $\alpha_1 \in \tilde{\Omega}^m$, let $\overline{\beta}(\alpha_1)$ be the smallest $\beta \in \mathbb{R}_+$ such that:

$$u_2(\tilde{\phi}_{\alpha_1}, \alpha_1, a_2') + \beta \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, \hat{\alpha}_1(\alpha_1), a_2') \geq u_2(\tilde{\phi}_{\alpha_1}, \alpha_1, a_2^*) + \beta \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, \hat{\alpha}_1(\alpha_1), a_2^*)$$

The choice of η and (A.30) ensures the existence of such $\overline{\beta}(\alpha_1)$. Let $\overline{\beta} \equiv 2\max_{\alpha_1 \in \tilde{\Omega}^m} \overline{\beta}(\alpha_1)$ and $\underline{\beta} \equiv \min_{\alpha_1 \in \tilde{\Omega}^m} \frac{\mu(\theta(\alpha_1))}{\mu(\alpha_1)}$. Let $T_1 \equiv \left\lceil \log \frac{\overline{\beta}}{\underline{\beta}} \Big/ \log \frac{1-\eta/2}{1-\eta} \right\rceil$. In the beginning of the abnormal phase (regardless of when it is triggered), $\beta_t(\alpha_1) \geq \underline{\beta}$ for all $\alpha_1 \in \tilde{\Omega}^m$. After player 2 observing a_1^* for T_1 consecutive periods, a_2^* is being strictly dominated by a_2' until he observes some $a_1' \neq a_1^*$. Every time player 1 plays any $a_1' \neq a_1^*$, he can trigger outcome (a_1^*, a_2^*) for at most $\overline{k}$ consecutive periods before a_2^* is being strictly dominated by a_2' again. Therefore, type θ's payoff in the abnormal phase is at most:

$$(1-\delta^{T_1}) + \delta^{T_1}\Big\{(1-\delta^{\overline{k}-1}) + \delta^{\overline{k}}(1-\delta^{\overline{k}-1}) + \delta^{2\overline{k}}(1-\delta^{\overline{k}-1}) + ...\Big\}$$

The term in the curly bracket converges to $\frac{\overline{k}}{1+\overline{k}}$ as $\delta \to 1$. Let $q \equiv \frac{\overline{k}}{2(\overline{k}+1)+1}$, type θ's payoff in beginning of the abnormal phase cannot exceed $1-2q$.

Remark: My construction in the abnormal phase is reminiscent of Jehiel and Samuelson (2012), in which the short-run players mistakenly believe that the strategic long-run player is using a stationary strategy. In the abnormal phase of my construction, player 2's belief attaches positive probability only to types that are playing stationary strategies. This leads to a similar reputation manipulation procedure: Type θ faces a trade-off between playing a_1^* at the expense of his reputation and playing other actions to build-up his reputation in the abnormal phase. My construction ensures that the speed of reputation building is bounded from above while the speed of reputation deterioration is bounded from below. When player 1's reputation is sufficiently bad, player 2 has a strict incentive to play a_2', which punishes type θ for at least one period.

A.3 Proof of Theorem 1.2

I prove Theorem 1.2 for all games satisfying Assumptions 1.1, 1.2 and 1.3 while allowing $a_{2,t}$ to depend on $h^t \equiv \{a_{1,s}, a_{2,s}, \xi_s\}_{s \leq t-1}$. To avoid cumbersome notation, I focus on the case where ξ_t has a finite number of realizations and there are no other commitment types, i.e. $\Omega^m = \{\overline{a}_1\}$. This is without loss of generality since when player 1 always plays $\overline{a}_1$, the probability of other commitment types becomes negligible relative to type $\overline{a}_1$ after a bounded number of periods, and those periods have negligible payoff consequences as $\delta \to 1$.

A.3.1 Several Useful Constants

I start with defining several useful constants which depend only on μ, u_1 and u_2, while making no reference to σ and δ. Let $M \equiv \max_{\theta,a_1,a_2} |u_1(\theta,a_1,a_2)|$ and

$$K \equiv \max_{\theta\in\Theta} \Big\{u_1(\theta,\overline{a}_1,\overline{a}_2) - u_1(\theta,\overline{a}_1,\underline{a}_2)\Big\} \Big/ \min_{\theta\in\Theta} \Big\{u_1(\theta,\overline{a}_1,\overline{a}_2) - u_1(\theta,\overline{a}_1,\underline{a}_2)\Big\}.$$

Since $\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1) > 0$, expression (1.10) implies the existence of $\kappa \in (0,1)$ such that:

$$\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1) + \sum_{\theta\in\Theta} \mu(\theta)\mathcal{D}(\theta,\overline{a}_1) > 0.$$

For any $\kappa \in (0,1)$, let

$$\rho_0(\kappa) \equiv \frac{(1-\kappa)\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)}{2\max_{(\theta,a_1)} |\mathcal{D}(\theta,a_1)|} > 0 \tag{A.32}$$

and

$$\overline{T}_0(\kappa) \equiv \lceil 1/\rho_0(\kappa) \rceil. \tag{A.33}$$

Let

$$\rho_1(\kappa) \equiv \frac{\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)}{\max_{(\theta,a_1)} |\mathcal{D}(\theta,a_1)|}. \tag{A.34}$$

and

$$\overline{T}_1(\kappa) \equiv \lceil 1/\rho_1(\kappa) \rceil. \tag{A.35}$$

Let $\overline{\delta} \in (0,1)$ be close enough to 1 such that for every $\delta \in [\overline{\delta},1)$ and $\theta_p \in \Theta_p$,

$$(1-\delta^{\overline{T}_0(0)})u_1(\theta_p,\overline{a}_1,\underline{a}_2) + \delta^{\overline{T}_0(0)}u_1(\theta_p,\overline{a}_1,\overline{a}_2) > \frac{1}{2}\Big(u_1(\theta_p,\overline{a}_1,\overline{a}_2) + u_1(\theta_p,\underline{a}_1,\underline{a}_2)\Big). \tag{A.36}$$

A.3.2 Random History & Random Path

Let $h^t \equiv (a^t, r^t)$, with $a^t \equiv (a_{1,s})_{s\leq t-1}$ and $r^t \equiv (a_{2,s},\xi_s)_{s\leq t-1}$. Let $a_*^t \equiv (\overline{a}_1,...,\overline{a}_1)$. I call h^t a *public history*, r^t a *random history* and r^∞ a *random path*.

Let $\mathcal{H}$ and $\mathcal{R}$ be the set of public histories and random histories, respectively, with $\succ$, $\succsim$, $\prec$ and $\precsim$ naturally defined. Recall that a strategy profile σ consists of $(\sigma_\omega)_{\omega\in\Omega}$ with $\sigma_\omega : \mathcal{H} \to \Delta(A_1)$ and $\sigma_2 : \mathcal{H} \to \Delta(A_2)$. Let $\mathcal{P}^\sigma(\omega)$ be the probability measure over public histories induced by (σ_ω,σ_2). Let $\mathcal{P}^\sigma \equiv \sum_{\omega\in\Omega} \mu(\omega)\mathcal{P}^\sigma(\omega)$. Let $V^\sigma(h^t) \equiv (V_\theta^\sigma(h^t))_{\theta\in\Theta} \in \mathbb{R}^{|\Theta|}$ be the continuation payoff vector for strategic types at h^t.

Let $\mathcal{H}^\sigma \subset \mathcal{H}$ be the set of histories h^t such that $\mathcal{P}^\sigma(h^t) > 0$, and let $\mathcal{H}^\sigma(\omega) \subset \mathcal{H}$ be the set of histories h^t such that $\mathcal{P}^\sigma(\omega)(h^t) > 0$. Let

$$\mathcal{R}_*^\sigma \equiv \Big\{r^\infty \Big| (a_*^t, r^t) \in \mathcal{H}^\sigma \text{ for all } t \text{ and } r^t \prec r^\infty\Big\}$$

be the set of random paths consistent with player 1 always playing $\overline{a}_1$. For every $h^t = (a^t, r^t)$, let $\overline{\sigma}_1[h^t] : \mathcal{H} \to A_1$ be a continuation strategy at h^t satisfying $\overline{\sigma}_1[h^t](h^s) = \overline{a}_1$ for all $h^s \succsim h^t$

with $h^s = (a^t, \overline{a}_1, ..., \overline{a}_1, r^s) \in \mathcal{H}^\sigma$. Let $\underline{\sigma}_1[h^t] : \mathcal{H} \to A_1$ be a continuation strategy that satisfies $\underline{\sigma}_1[h^t](h^s) = \underline{a}_1$ for all $h^s \succsim h^t$ with $h^s = (a^t, \underline{a}_1, ..., \underline{a}_1, r^s) \in \mathcal{H}^\sigma$. For every $\theta \in \Theta$, let

$$\overline{\mathcal{R}}^\sigma(\theta) \equiv \Big\{ r^t \Big| \overline{\sigma}_1[a_*^t, r^t] \text{ is type } \theta\text{'s best reply to } \sigma_2 \Big\} \text{ and } \underline{\mathcal{R}}^\sigma(\theta) \equiv \Big\{ r^t \Big| \underline{\sigma}_1[a_*^t, r^t] \text{ is type } \theta\text{'s best reply to } \sigma_2 \Big\}.$$

A.3.3 Beliefs & Best Response Sets

Let $\mu(a^t, r^t) \in \Delta(\Omega)$ be player 2's posterior belief at (a^t, r^t) and specifically, let $\mu^*(r^t) \equiv \mu(a_*^t, r^t)$. Let

$$\mathcal{B}_\kappa \equiv \Big\{ \tilde{\mu} \in \Delta(\Omega) \Big| \kappa \mu(\overline{a}_1) \mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} \tilde{\mu}(\theta) \mathcal{D}(\theta, \overline{a}_1) \geq 0 \Big\}. \tag{A.37}$$

By definition, only $\{\tilde{\mu}(\theta)\}_{\theta \in \Theta}$ matters for whether $\tilde{\mu}$ belongs to $\mathcal{B}_\kappa$ or not. Moreover, $\mathcal{B}_{\kappa'} \subsetneq \mathcal{B}_\kappa$ for every $\kappa, \kappa' \in [0,1]$ with $\kappa' < \kappa$.

Let $q^*(r^t)(\omega)$ be the (ex ante) probability that player 1's type being ω and his past actions being a_*^t conditional on r^t. Let $q^*(r^t) \in \mathbb{R}_+^{|\Omega|}$ be the corresponding vector of probabilities. For any δ and $\sigma \in \text{NE}(\mu, \delta)$, Bayes Rule implies that:

- For any a^t and $r^t, \hat{r}^t \succ r^{t-1}$ satisfying $(a^t, r^t), (a^t, \hat{r}^t) \in \mathcal{H}^\sigma$, we have $\mu(a^t, r^t) = \mu(a^t, \hat{r}^t)$.
- For any $r^t, \hat{r}^t \succ r^{t-1}$ with $(a_*^t, r^t), (a_*^t, \hat{r}^t) \in \mathcal{H}^\sigma$, we have $q^*(r^t) = q^*(\hat{r}^t)$.

This is because player 1's action in period $t-1$ depends on r^t only through r^{t-1}, so is player 2's belief at every on-path history. Since the commitment type always plays $\overline{a}_1$, we have $q^*(r^t)(\overline{a}_1) = \mu_0(\overline{a}_1)$.

For future reference, I introduce two set of random histories based on player 2's posterior beliefs. Let

$$\mathcal{R}_g^\sigma \equiv \Big\{ r^t \Big| (a_*^t, r^t) \in \mathcal{H}^\sigma \text{ and } \mu^*(r^t)\big(\Theta_p \cup \Theta_n\big) = 0 \Big\}, \tag{A.38}$$

and let

$$\hat{\mathcal{R}}_g^\sigma \equiv \Big\{ r^t \Big| \exists r^T \succsim r^t \text{ such that } r^T \in \mathcal{R}_g^\sigma \Big\}. \tag{A.39}$$

A.3.4 A Few Useful Observations

I present four Lemmas, which are useful preliminary results towards the final proof. Recall that $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$ is type θ's strategy. The first one shows the implications of MSM on player 1's equilibrium strategy:

Lemma A.3.1. *Suppose $\sigma \in NE(\delta, \mu, \phi)$, $\theta \succ \tilde{\theta}$ and $h_*^t = (a_*^t, r^t) \in \mathcal{H}^\sigma(\theta) \cap \mathcal{H}^\sigma(\tilde{\theta})$,*

- *if $r^t \in \overline{\mathcal{R}}^\sigma(\tilde{\theta})$, then $\sigma_\theta(a_*^s, r^s)(\overline{a}_1) = 1$ for every $(a_*^s, r^s) \in \mathcal{H}^{(\overline{\sigma}_1(h_*^t), \sigma_2)}(\theta)$ with $r^s \succsim r^t$.*
- *if $r^t \in \underline{\mathcal{R}}^\sigma(\theta)$, then $\sigma_{\tilde{\theta}}(a^s, r^s)(\underline{a}_1) = 1$ for every $(a^s, r^s) \in \mathcal{H}^{(\underline{\sigma}_1(h_*^t), \sigma_2)}(\tilde{\theta})$ with $(a^s, r^s) \succsim (a_*^t, r^t)$.*

Proof of Lemma A.3.1: I only prove first part, since the second part can be shown similarly by switching signs. Without loss of generality, I focus on history h^0. For notation simplicity, let

$\overline{\sigma}_1[h^0] = \overline{\sigma}_1$. For every σ_ω and σ_2, let $P^{(\sigma_\omega,\sigma_2)} : A_1 \times A_2 \to [0,1]$ be defined as:

$$P^{(\sigma_\omega,\sigma_2)}(a_1,a_2) \equiv \sum_{t=0}^{+\infty}(1-\delta)\delta^t p_t^{(\sigma_\omega,\sigma_2)}(a_1,a_2)$$

where $p_t^{(\sigma_\omega,\sigma_2)}(a_1,a_2)$ is the probability of (a_1,a_2) occurring in period t under (σ_ω,σ_2). Let $P_i^{(\sigma_1,\sigma_2)} \in \Delta(A_2)$ be $P^{(\sigma_1,\sigma_2)}$'s marginal distribution on A_i.

Suppose towards a contradiction that $\overline{\sigma}_1$ is type $\tilde{\theta}$'s best reply and there exists σ_θ with $P_1^{(\sigma_\theta,\sigma_2)}(\overline{a}_1) < 1$ such that σ_θ is type θ's best reply, then type $\tilde{\theta}$ and θ's incentive constraints require that:

$$\sum_{a_2\in A_2}\Big(P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2)\Big)u_1(\tilde{\theta},\overline{a}_1,a_2)$$

$$\geq \sum_{a_2\in A_2, a_1\neq\overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\tilde{\theta},a_1,a_2) - u_1(\tilde{\theta},\overline{a}_1,a_2)\Big),$$

and

$$\sum_{a_2\in A_2}\Big(P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2)\Big)u_1(\theta,\overline{a}_1,a_2)$$

$$\leq \sum_{a_2\in A_2, a_1\neq\overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\theta,a_1,a_2) - u_1(\theta,\overline{a}_1,a_2)\Big).$$

Since $P_1^{(\sigma_\theta,\sigma_2)}(\overline{a}_1) < 1$ and u_1 has SID in θ and a_1, we have:

$$\sum_{a_2\in A_2, a_1\neq\overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\tilde{\theta},a_1,a_2) - u_1(\tilde{\theta},\overline{a}_1,a_2)\Big)$$

$$> \sum_{a_2\in A_2, a_1\neq\overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\theta,a_1,a_2) - u_1(\theta,\overline{a}_1,a_2)\Big)$$

which implies that:

$$\sum_{a_2\in A_2}\Big(P_2^{(\sigma_\theta,\sigma_2)}(a_2) - P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)\Big)\Big(u_1(\theta,\overline{a}_1,a_2) - u_1(\tilde{\theta},\overline{a}_1,a_2)\Big) > 0. \tag{A.40}$$

On the other hand, since u_1 is strictly decreasing in a_1, we have:

$$\sum_{a_2\in A_2, a_1\neq\overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\tilde{\theta},a_1,a_2) - u_1(\tilde{\theta},\overline{a}_1,a_2)\Big) > 0$$

Type $\tilde{\theta}$'s incentive constraint implies that:

$$\sum_{a_2\in A_2}\Big(P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2)\Big)u_1(\tilde{\theta},\overline{a}_1,a_2) > 0. \tag{A.41}$$

Since $P_2^{(\sigma_\theta,\sigma_2)}$ and $P_2^{(\overline{\sigma}_1,\sigma_2)}$ are both probability distributions,

$$\sum_{a_2\in A_2}\Big(P_2^{(\sigma_\theta,\sigma_2)}(a_2)-P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)\Big)=0.$$

Since $u_1(\theta,\overline{a}_1,a_2)-u_1(\tilde{\theta},\overline{a}_1,a_2)$ is weakly increasing in a_2, (A.40) implies that $P_2^{(\sigma_\theta,\sigma_2)}(\overline{a}_2)-P_2^{(\overline{\sigma}_1,\sigma_2)}(\overline{a}_2)>0$. Since $u_1(\tilde{\theta},\overline{a}_1,a_2)$ is strictly increasing in a_2, (A.41) implies that $P_2^{(\sigma_\theta,\sigma_2)}(\overline{a}_2)-P_2^{(\overline{\sigma}_1,\sigma_2)}(\overline{a}_2)<0$, leading to a contradiction. □

The next Lemma places a uniform upper bound on the number of '*bad periods*' in which $\overline{a}_2$ is not player 2's best reply despite $\overline{a}_1$ has always been played and $\mu^*(r^t)\in\mathcal{B}_\kappa$.

Lemma A.3.2. *If $\mu^*(r^t)\in\mathcal{B}_\kappa$ and $\overline{a}_2$ is not a strict best reply at (a_*^t,r^t), then:*

$$\sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta)-q^*(r^{t+1})(\theta)\Big)\geq\rho_0(\kappa).\tag{A.42}$$

Proof of Lemma A.3.2: If $\mu^*(r^t)\in\mathcal{B}_\kappa$, then:[3]

$$\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta}q^*(r^t)(\theta)\mathcal{D}(\theta,\overline{a}_1)\geq 0.$$

Suppose $\overline{a}_2$ is not a strict best reply at (a_*^t,r^t), then,

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta}q^*(r^{t+1})(\theta)\mathcal{D}(\theta,\overline{a}_1)+\sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta)-q^*(r^{t+1})(\theta)\Big)\mathcal{D}(\theta,\underline{a}_1)\leq 0,$$

or equivalently,

$$\underbrace{\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta}q^*(r^t)(\theta)\mathcal{D}(\theta,\overline{a}_1)}_{\geq 0}+\underbrace{(1-\kappa)\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)}_{>0}$$

$$+\sum_{\theta\in\Theta}\Big(q^*(r^{t+1})(\theta)-q^*(r^t)(\theta)\Big)\mathcal{D}(\theta,\overline{a}_1)+\sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta)-q^*(r^{t+1})(\theta)\Big)\mathcal{D}(\theta,\underline{a}_1)\leq 0,$$

According to (A.32), we have:

$$\sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta)-q^*(r^{t+1})(\theta)\Big)\geq\frac{(1-\kappa)\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)}{2\max_{\theta,a_1}|\mathcal{D}(\theta,a_1)|}=\rho_0(\kappa).$$

□

Lemma A.3.2 implies that for any $\sigma\in\mathrm{NE}(\delta,\mu,\phi)$ and along any $r^\infty\in\mathcal{R}_*^\sigma$, the number of r^t such that $\mu^*(r^t)\in\mathcal{B}_\kappa$ but $\overline{a}_2$ is not a strict best reply is at most $\overline{T}_0(\kappa)$. The next Lemma obtains an upper bound for player 1's *drop-out* payoff at any *unfavorable belief*.

[3] According to Bayes Rule, $\mu^*(r^t)(\theta)\geq q^*(r^t)(\theta)$ for all $\theta\in\Theta$ and $\frac{\mu^*(r^t)(\theta)}{q^*(r^t)(\theta)}$ is independent of θ as long as $q^*(r^t)(\theta)\neq 0$.

Lemma A.3.3. *For every $\sigma \in NE(\delta,\mu,\phi)$ and $h^t \in \mathcal{H}^\sigma$ with*

$$\mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1) + \sum_\theta \mu(h^t)(\theta)\mathcal{D}(\theta,\overline{a}_1) < 0. \tag{A.43}$$

Let $\underline{\theta} \equiv \min\Big\{supp\big(\mu(h^t)\big)\Big\}$, then:

$$V_{\underline{\theta}}(h^t) = u_1(\underline{\theta},\underline{a}_1,\underline{a}_2).$$

Proof of Lemma A.3.3: Let

$$\Theta^* \equiv \Big\{\tilde{\theta} \in \Theta_p \cup \Theta_n \Big| \mu(h^t)(\tilde{\theta}) > 0\Big\}.$$

According to (A.43), $\Theta^* \neq \{\varnothing\}$. The rest of the proof is done via induction on $|\Theta^*|$. When $|\Theta^*| = 1$, there exists a pure strategy $\sigma^*_{\underline{\theta}} : \mathcal{H} \to A_1$ in the support of $\sigma_{\underline{\theta}}$ such that (A.43) holds for all h^s satisfying $h^s \in \mathcal{H}^{(\sigma^*_{\underline{\theta}},\sigma_2)}$ and $h^s \succsim h^t$. At every such h^s, $\underline{a}_2$ is player 2's strict best reply. When playing $\sigma^*_{\underline{\theta}}$, type $\underline{\theta}$'s stage game payoff is no more than $u_1(\underline{\theta},\underline{a}_1,\underline{a}_2)$ in every period.

Suppose towards a contradiction that the conclusion holds when $|\Theta^*| \leq k-1$ but fails when $|\Theta^*| = k$, then there exists $h^s \in \mathcal{H}^\sigma(\underline{\theta})$ with $h^s \succsim h^t$ such that

- $\mu(h^\tau) \notin \mathcal{B}_\kappa$ for all $h^s \succsim h^\tau \succsim h^t$.
- $V_{\underline{\theta}}(h^s) > u_1(\underline{\theta},\underline{a}_1,\underline{a}_2)$.
- For all a_1 such that $\mu(h^s,a_1) \notin \mathcal{B}_\kappa$, $\sigma_{\underline{\theta}}(h^s)(a_1) = 0$.[4]

According to the martingale property of beliefs, there exists a_1 such that $(h^s,a_1) \in \mathcal{H}^\sigma$ and $\mu(h^s,a_1)$ satisfies (A.43). Since $\mu(h^s,a_1)(\underline{\theta}) = 0$, there exists $\tilde{\theta} \in \Theta^*\backslash\{\underline{\theta}\}$ such that $(h^s,a_1) \in \mathcal{H}^\sigma(\tilde{\theta})$. Our induction hypothesis suggests that:

$$V_{\tilde{\theta}}(h^s) = u_1(\tilde{\theta},\underline{a}_1,\underline{a}_2).$$

The incentive constraints of type $\underline{\theta}$ and type $\tilde{\theta}$ at h^s require the existence of $(\alpha_{1,\tau},\alpha_{2,\tau})_{\tau=0}^\infty$ with $\alpha_{i,\tau} \in \Delta(A_i)$ such that:

$$\mathbb{E}\Big[\sum_{\tau=0}^\infty (1-\delta)\delta^\tau \Big(u_1(\underline{\theta},\alpha_{1,\tau},\alpha_{2,\tau}) - u_1(\underline{\theta},\underline{a}_1,\underline{a}_2)\Big)\Big] > 0 \geq \mathbb{E}\Big[\sum_{\tau=0}^\infty (1-\delta)\delta^\tau \Big(u_1(\tilde{\theta},\alpha_{1,\tau},\alpha_{2,\tau}) - u_1(\tilde{\theta},\underline{a}_1,\underline{a}_2)\Big)\Big],$$

where $\mathbb{E}[\cdot]$ is taken over probability measure $\mathcal{P}^\sigma$. However, the supermodularity condition implies that,

$$u_1(\underline{\theta},\alpha_{1,\tau},\alpha_{2,\tau}) - u_1(\underline{\theta},\underline{a}_1,\underline{a}_2) \leq u_1(\tilde{\theta},\alpha_{1,\tau},\alpha_{2,\tau}) - u_1(\tilde{\theta},\underline{a}_1,\underline{a}_2),$$

leading to a contradiction. □

[4] I omit $(a_{2,s},\xi_s)$ in the expression for histories since they play no role in the posterior belief on Ω at every on-path history.

The next Lemma outlines an important implication of $r^t \notin \hat{\mathcal{R}}_g^\sigma$.

Lemma A.3.4. *If $r^t \notin \hat{\mathcal{R}}_g^\sigma$ and $(a_*^t, r^t) \in \mathcal{H}^\sigma$, then there exists*

$$\theta \in \big(\Theta_p \cup \Theta_n\big) \bigcap supp\big(\mu^*(r^t)\big)$$

such that $r^t \in \overline{R}^\sigma(\theta)$.

Proof of Lemma A.3.4: Suppose towards a contradiction that $r^t \notin \hat{\mathcal{R}}_g^\sigma$ but no such θ exists. Let

$$\theta_1 \equiv \max\Big\{\big(\Theta_p \cup \Theta_n\big) \bigcap \mathrm{supp}\big(\mu^*(r^t)\big)\Big\}.^5$$

Let $(a_*^{t_1}, r^{t_1}) \succsim (a_*^t, r^t)$ be the history at which type θ_1 has a strict incentive not to play $\overline{a}_1$ with $(a_*^{t_1}, r^{t_1}) \in \mathcal{H}^\sigma$. For any $(a_*^{t_1+1}, r^{t_1+1}) \succ (a_*^{t_1}, r^{t_1})$ with $(a_*^{t_1+1}, r^{t_1+1}) \in \mathcal{H}^\sigma$, on one hand, we have $\mu^*(r^{t_1+1})(\theta_1) = 0$. On the other hand, the fact that $r^t \notin \hat{\mathcal{R}}_g^\sigma$ implies that $\mu^*(r^{t_1+1})(\Theta_n \cup \Theta_p) > 0$.

Let

$$\theta_2 \equiv \max\Big\{\big(\Theta_p \cup \Theta_n\big) \bigcap \mathrm{supp}\big(\mu^*(r^{t_1+1})\big)\Big\}.$$

Examine type θ_1 and θ_2's incentive constraints at $(a_*^{t_1}, r^{t_1})$. According to Lemma A.3.1, there exists $r^{t_2} \succ r^{t_1}$ such that type θ_2 has a strict incentive not to play $\overline{a}_1$ at $(a_*^{t_2}, r^{t_2}) \in \mathcal{H}^\sigma$.

Therefore, we can iterate this process and obtain $r^{t_3} \succ r^{t_4}$... Since

$$\Big|\mathrm{supp}\Big(\mu^*(r^{t_{k+1}})\Big)\Big| \leq \Big|\mathrm{supp}\Big(\mu^*(r^{t_k})\Big)\Big| - 1,$$

for any $k \in \mathbb{N}$, there exists $m \leq |\Theta_p \cup \Theta_n|$ such that $(a_*^{t_m}, r^{t_m}) \in \mathcal{H}^\sigma$, $r^{t_m} \succsim r^t$ and $\mu^*(r^{t_m})(\Theta_n \cup \Theta_p) = 0$, which contradicts $r^t \notin \hat{\mathcal{R}}_g^\sigma$. □

A.3.5 Positive Types

In this part, I show the following Proposition:

Proposition A.3. *If $\Theta_n = \{\varnothing\}$ and $\mu \in \mathcal{B}_\kappa$, then for every θ, we have:*

$$V_\theta(a_*^0, r^0) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - 2M(1 - \delta^{\overline{T}_0(\kappa)}).$$

Despite Proposition A.3 is stated in terms of the prior belief, the conclusion applies to all r^t and $\theta \in \Theta_g \cup \Theta_p$ when $\mu^*(r^t) \in \mathcal{B}_\kappa$ and $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta) \backslash \cup_{\theta_n \in \Theta_n} \mathcal{H}^\sigma(\theta_n)$. The proof is decomposed into Lemma A.3.5 and Lemma A.3.6, which together imply Proposition A.3. Let $\sigma \in \mathrm{NE}(\delta, \mu, \phi)$.

Lemma A.3.5. *If $\mu^*(r^t) \in \mathcal{B}_\kappa$ for all $r^t \in \hat{\mathcal{R}}_g^\sigma$, then for any $r^\infty \in \mathcal{R}_*^\sigma$,*

$$\Big|\Big\{t \in \mathbb{N}\Big| r^\infty \succ r^t \text{ and } \overline{a}_2 \text{ is not a strict best reply at } (a_*^t, r^t)\Big\}\Big| \leq \overline{T}_0(\kappa). \quad \text{(A.44)}$$

[5] When $r^t \notin \hat{\mathcal{R}}_g^\sigma$, then the intersection of $\Theta_p \cup \Theta_n$ and $\mathrm{supp}\big(\mu^*(r^t)\big)$ cannot be empty by definition.

PROOF OF LEMMA A.3.5: Pick any $r^{\infty} \in \mathcal{R}^{\sigma}_{*}$, if $r^0 \notin \hat{\mathcal{R}}^{\sigma}_{g}$, then let $t^* = -1$. Otherwise, let

$$t^* \equiv \max\Big\{t \in \mathbb{N} \cup \{+\infty\}\Big| r^t \in \hat{\mathcal{R}}^{\sigma}_{g} \text{ and } r^{\infty} \succ r^t\Big\}.$$

Using the argument in Lemma A.3.2, for any $t \leq t^*$, if $\overline{a}_2$ is not a strict best reply at (a^t_*, r^t), inequality (A.42) holds.

Next, I show that $\mu^*(r^{t^*+1}) \in \mathcal{B}_{\kappa}$. If $t^* = -1$, this is a direct implication of (1.10). If $t^* \geq 0$, then there exists $\hat{r}^{t^*+1} \succ r^{t^*}$ such that $\hat{r}^{t^*+1} \in \hat{\mathcal{R}}^{\sigma}_{g}$. Let $r^{t^*+1} \prec r^{\infty}$, we have $q^*(r^{t^*+1}) = q^*(\hat{r}^{t^*+1})$. Moreover, since $\mu^*(r^t) \in \mathcal{B}_{\kappa}$ for every $r^t \in \hat{\mathcal{R}}^{\sigma}_{g}$, we have $\mu^*(r^{t^*+1}) = \mu^*(\hat{r}^{t^*+1}) \in \mathcal{B}_{\kappa}$.

Since $r^{t^*+1} \notin \hat{\mathcal{R}}^{\sigma}_{g}$, Lemma A.3.4 implies the existence of

$$\theta \in \big(\Theta_p \cup \Theta_n\big) \bigcap \text{supp}\big(\mu^*(r^{t^*+1})\big)$$

such that $r^{t^*+1} \in \overline{R}^{\sigma}(\theta)$. Since $\theta_g \succ \theta$ for all $\theta_g \in \Theta_g$, Lemma A.3.1 implies that for every θ_g and $r^{\infty} \succ r^t \succsim r^{t^*+1}$, we have $\sigma_{\theta_g}(a^t_*, r^t) = 1$, and therefore, $q^*(r^t)(\theta_g) = q^*(r^{t+1})(\theta_g)$. This implies that $\mu^*(r^t) \in \mathcal{B}_{\kappa}$ for every $r^{\infty} \succ r^t \succsim r^{t^*+1}$. If $\overline{a}_2$ is not a strict best reply at (a^t_*, r^t) for any $t > t^*$, inequality (A.42) again applies.

To sum up, for every $t \in \mathbb{N}$, if $\overline{a}_2$ is not a strict best reply at (a^t_*, r^t), then:

$$\sum_{\theta \in \Theta} \Big(q^*(r^t)(\theta) - q^*(r^{t+1})(\theta)\Big) \geq \rho_0(\kappa),$$

from which we obtain (A.44). □

The next result shows that the condition in Lemma A.3.5 holds in every equilibrium when δ is large enough.

Lemma A.3.6. *If $\delta > \overline{\delta}$, then $\mu^*(r^t) \in \mathcal{B}_0$ for every $r^t \in \hat{\mathcal{R}}^{\sigma}_{g}$ with $\mu^*(r^t)(\Theta_n) = 0$.*

PROOF OF LEMMA A.3.6: For any given $\delta > \overline{\delta}$, according to (A.36), there exists $\kappa^* \in (0,1)$ such that:

$$(1-\delta^{\overline{T}_0(\kappa^*)})u_1(\theta_p, \overline{a}_1, \underline{a}_2) + \delta^{\overline{T}_0(\kappa^*)}u_1(\theta_p, \overline{a}_1, \overline{a}_2) > \frac{1}{2}\Big(u_1(\theta_p, \overline{a}_1, \overline{a}_2) + u_1(\theta_p, \underline{a}_1, \underline{a}_2)\Big). \tag{A.45}$$

Suppose towards a contradiction that there exist r^{t_1} and r^{T_1} such that:

- $r^{T_1} \succ r^{t_1}$, $r^{T_1} \in \mathcal{R}^{\sigma}_{g}$ and $\mu^*(r^{t_1}) \notin \mathcal{B}_0$.

Since $\mu^*(r^{T_1}) \in \mathcal{B}_0$, let t^*_1 be the largest $t \in \mathbb{N}$ such that $\mu^*(r^t) \notin \mathcal{B}_0$ for $r^{T_1} \succ r^t \succsim r^{t_1}$. Then there exists $a_1 \neq \overline{a}_1$ and $r^{t^*_1+1} \succ r^{t^*_1}$ such that $\mu\big((a^{t^*_1}_*, a_1), r^{t^*_1+1}\big) \notin \mathcal{B}_0$ and $\big((a^{t^*_1}_*, a_1), r^{t^*_1+1}\big) \in \mathcal{H}^{\sigma}$. This also implies the existence of $\theta_p \in \Theta_p \cap \text{supp}\Big(\mu\big((a^{t^*_1}_*, a_1), r^{t^*_1+1}\big)\Big)$.

According to Lemma A.3.3, type θ_p's continuation payoff at $(a^{t^*_1}_*, r^{t^*_1})$ by playing a_1 is at most

$$(1-\delta)u_1(\theta_p, \underline{a}_1, \overline{a}_2) + \delta u_1(\theta_p, \underline{a}_1, \underline{a}_2). \tag{A.46}$$

His incentive constraint at $(a_*^{t_1^*}, r^{t_1^*})$ requires that his expected payoff from $\overline{\sigma}_1$ is weakly lower than (A.46), i.e. there exists $r^{t_1^*+1} \succ r^{t_1^*}$ satisfying $(a_*^{t_1^*+1}, r^{t_1^*+1}) \in \mathcal{H}^\sigma$ and type θ_p's continuation payoff at $(a_*^{t_1^*+1}, r^{t_1^*+1})$ is no more than:

$$\frac{1}{2}\Big(u_1(\theta_p, \overline{a}_1, \overline{a}_2) + u_1(\theta_p, \underline{a}_1, \underline{a}_2)\Big). \tag{A.47}$$

If $\mu^*(r^t) \in \mathcal{B}_{\kappa^*}$ for every $r^t \in \hat{\mathcal{R}}_g^\sigma \cap \{r^t \succsim r^{t_1^*}\}$, then according to Lemma A.3.5, his continuation payoff at $(a_*^{t_1^*}, r^{t_1^*})$ by playing $\overline{\sigma}_1$ is at least:

$$(1-\delta^{\overline{T}_0(\kappa^*)})u_1(\theta_p, \overline{a}_1, \underline{a}_2) + \delta^{\overline{T}_0(\kappa^*)}u_1(\theta_p, \overline{a}_1, \overline{a}_2),$$

which is strictly larger than (A.47) by the definition of κ^* in (A.45), leading to a contradiction.

Suppose on the other hand, there exists $r^{t_2} \succ r^{t_1^*}$ such that:

- $r^{t_2} \in \hat{\mathcal{R}}_g^\sigma$ while $\mu^*(r^{t_2}) \notin \mathcal{B}_{\kappa^*}$.

There exists $r^{T_2} \succ r^{t_2}$ such that $r^{T_2} \in \mathcal{R}_g^\sigma$ and $r^{T_2} \succ r^{t_2}$. Again, we can find $r^{t_2^*}$ such that t_2^* be the largest $t \in [t_2, T_2]$ such that $\mu^*(r^t) \notin \mathcal{B}_0$ for $r^{T_2} \succ r^t \succsim r^{t_2}$. Then there exists $a_1 \neq \overline{a}_1$ and $r^{t_2^*+1} \succ r^{t_2^*}$ such that $\mu\big((a_*^{t_2^*}, a_1), r^{t_2^*+1}\big) \notin \mathcal{B}_0$ and $\big((a_*^{t_2^*}, a_1), r^{t_2^*+1}\big) \in \mathcal{H}^\sigma$.

Repeating this argument by iterating the above process, for every $k \geq 1$, in order to satisfy player 1's incentive constraint to play $a_1 \neq \overline{a}_1$ at $(a_*^{t_k^*}, r^{t_k^*})$, we can find the triple $(r^{t_{k+1}}, r^{t_{k+1}^*}, r^{T_{k+1}})$, i.e. this process cannot stop after finite rounds of iteration. Since $\mu^*(r^{t_k}) \notin \mathcal{B}_{\kappa^*}$ but $\mu^*(r^{t_k^*+1}) \in \mathcal{B}_0$ as well as $r^{t_{k+1}} \succ r^{t_k^*+1}$, we have:

$$\sum_{\theta\in\Theta} q^*(r^{t_k})(\theta) - q^*(r^{t_{k+1}})(\theta) \geq \sum_{\theta\in\Theta} q^*(r^{t_k})(\theta) - q^*(r^{t_k^*+1})(\theta) \geq \rho_1(\kappa^*) \tag{A.48}$$

for every $k \geq 2$. (A.48) and (A.35) together suggest that this iteration process cannot last for more than $\overline{T}_1(\kappa^*)$ rounds, which is an integer independent of δ, leading to a contradiction. □

The next Lemma is not needed for the proof of Proposition C.1 but will be useful for future reference.

Lemma A.3.7. *For any $\delta \geq \overline{\delta}$ and any $\sigma \in NE(\delta, \mu, \phi)$, for every r^t such that $(a_*^t, r^t) \in \mathcal{H}^\sigma$, $\mu^*(r^t)(\Theta_n) = 0$, $r^t \notin \hat{\mathcal{R}}_g^\sigma$ and*

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0, \tag{A.49}$$

$\overline{a}_2$ is player 2's strict best reply at every $(a_*^s, r^s) \succsim (a_*^t, r^t)$ with $(a_*^s, r^s) \in \mathcal{H}^\sigma$.

Proof of Lemma A.3.7: Since $\mu^*(r^t)(\Theta_n) = 0$ and $r^t \notin \hat{\mathcal{R}}_g^\sigma$, Lemma A.3.4 implies the existence of $\theta_p \in \Theta_p \cap \text{supp}(\mu^*(r^t))$ such that $r^t \in \overline{R}^\sigma(\theta_p)$. According to Lemma A.3.1, $\sigma_\theta(a_*^s, r^s)(\overline{a}_1) = 1$ for every $(a_*^s, r^s) \in \mathcal{H}^\sigma(\theta)$ with $r^s \succsim r^t$. From (A.49), we know that $\overline{a}_2$ is not a strict best reply only if there exists type $\theta_p \in \Theta_p$ who plays $a_1 \neq \overline{a}_1$ with positive probability. In particular, (A.49) implies the existence of $\overline{\kappa} \in (0,1)$ such that:[6]

$$\overline{\kappa}\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0.$$

According to (A.42), we have:

$$\sum_{\theta\in\Theta_p} \Big(q^*(r^s)(\theta) - q^*(r^{s+1})(\theta)\Big) \geq \rho_0(\overline{\kappa})$$

whenever $\overline{a}_2$ is not a strict best reply at $(a_*^s, r^s) \succsim (a_*^t, r^t)$. Therefore, there can be at most $\overline{T}_0(\overline{\kappa})$ such periods. Hence, there exists r^N with $(a_*^N, r^N) \in \mathcal{H}^\sigma$ such that:

- $\overline{a}_2$ is not a strict best reply at (a_*^N, r^N).
- $\overline{a}_2$ is a strict best reply for all $(a_*^s, r^s) \succ (a_*^N, r^N)$ with $(a_*^s, r^s) \in \mathcal{H}^\sigma$.

Then there exists $\theta_p \in \Theta_p$ that plays $a_1 \neq \overline{a}_1$ in equilibrium at (a_*^N, r^N), his continuation payoff by always playing $\overline{a}_1$ is at least $(1-\delta)u_1(\theta_p, \overline{a}_1, \underline{a}_2) + \delta u_1(\theta_p, \overline{a}_1, \overline{a}_2)$ while his equilibrium continuation payoff from playing a_1 is at most $(1-\delta)u_1(\theta_p, \underline{a}_1, \overline{a}_2) + \delta u_1(\theta_p, \underline{a}_1, \underline{a}_2)$ according to Lemma A.3.3. The latter is strictly less than the former when $\delta > \overline{\delta}$, leading to a contradiction. □

A.3.6 Incorporating Negative Types

Next, we extend the proof by allowing for $\Theta_n \neq \{\varnothing\}$. Lemmas A.3.5 and A.3.6 imply the following result in this general environment:

Proposition A.4. *For any $\delta > \overline{\delta}$ and $\sigma \in NE(\delta, \mu, \phi)$, there do not exist $\theta_p \in \Theta_p$, $r^{t+1} \succ r^t$ and $a_1 \neq \overline{a}_1$ that simultaneously satisfy:*

1. $r^{t+1} \in \hat{\mathcal{R}}_g^\sigma$.
2. $\big((a_*^t, a_1), r^{t+1}\big) \in \mathcal{H}^\sigma(\theta_p)$.
3. $V_{\theta_p}\Big(\big((a_*^t, a_1), \hat{r}^{t+1}\big)\Big) = u_1(\theta_p, \underline{a}_1, \underline{a}_2)$ *for all* $\hat{r}^{t+1} \succ r^t$.

PROOF OF PROPOSITION A.4: Suppose towards a contradiction that such $\theta_p \in \Theta_p$, r^t, r^{t+1} and a_1 exist. From requirement 3, we know that $r^t \in \underline{\mathcal{R}}^\sigma(\theta_p)$. According to Lemma 1.4.1, $\theta_n \prec \theta_p$ for all $\theta_n \in \Theta_n$. The second part of Lemma A.3.1 then implies that $\mu^*(\hat{r}^{t+1})(\Theta_n) = 0$ for all $\hat{r}^{t+1} \succ r^t$ with $(a_*^{t+1}, \hat{r}^{t+1}) \in \mathcal{H}^\sigma$.

If $\mu^*(r^{t+1}) \in \mathcal{B}_\kappa$, then requirement 2 and Proposition A.3 together lead to a contradiction when examining type θ_p's incentive at (a_*^t, r^t) to play a_1 as opposed to $\overline{a}_1$. If $\mu^*(r^{t+1}) \notin \mathcal{B}_\kappa$, since $\delta > \overline{\delta}$ and $r^{t+1} \in \hat{\mathcal{R}}_g^\sigma$, we obtain a contradiction from Lemma A.3.6. □

[6]The reasons why we cannot directly apply Lemma A.3.2 are, first, stronger conclusion is required for Lemma A.3.7, and second, $\overline{\kappa}$ can be arbitrarily close to 1, while κ is uniformly bounded below 1 for any given μ.

The rest of the proof is decomposed into several steps by considering any $\sigma \in \mathrm{NE}(\mu, \delta)$ when δ is large enough. First,[7]

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0 \tag{A.50}$$

for all $t \geq 1$ and r^t satisfying $(a_*^t, r^t) \in \mathcal{H}^\sigma$. This is because otherwise, according to Lemma A.3.3, there exists $\theta \in \mathrm{supp}(\mu^*(r^t))$ such that $V_\theta(a_*^t, r^t) = u_1(\theta, \underline{a}_1, \underline{a}_2)$. But then, at (a_*^{t-1}, r^{t-1}) with $r^{t-1} \prec r^t$, he could obtain strictly higher payoff by playing $\underline{a}_1$ instead of $\overline{a}_1$, leading to a contradiction.

Next comes the following Lemma:

Lemma A.3.8. *$V_\theta(a_*^t, r^t) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - 2M(K+1)(1-\delta)$ for every θ and $r^t \notin \hat{\mathcal{R}}_g^\sigma$ satisfying:*

- $(a_*^t, r^t) \in \mathcal{H}^\sigma$.
- *Either $t = 0$ or $t \geq 1$ but there exists $\hat{r}^t$ such that $r^t, \hat{r}^t \succ r^{t-1}$, $(a_*^t, \hat{r}^t) \in \mathcal{H}^\sigma$ and $\hat{r}^t \in \hat{\mathcal{R}}_g^\sigma$.*

Proof of Lemma A.3.8: If $\mu^*(r^t) \in \mathcal{B}_\kappa$ and $r^t \notin \hat{\mathcal{R}}_g^\sigma$, then Lemmas A.3.1 and A.3.4 suggest that $\mu^*(r^s) \in \mathcal{B}_\kappa$ for all $r^s \succsim r^t$ and the conclusion is straightforward from Lemma A.3.2.

Therefore, for the rest of the proof, I assume that $\mu^*(r^t) \notin \mathcal{B}_\kappa$. I consider two cases. First, when $\mu^*(r^t)(\Theta_n) > 0$, then according to (A.48),[8]

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0.$$

Since $r^t \notin \hat{\mathcal{R}}_g^\sigma$, according to Lemma A.3.4, there exists $\theta \in \Theta_p \cup \Theta_n$ with $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta)$ such that $r^t \in \overline{\mathcal{R}}^\sigma(\theta)$. According to Lemma A.3.1, for all $\theta_g \in \Theta_g$ with $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta_g)$ and every $(a_*^s, r^s) \in \mathcal{H}^\sigma(\theta)$ with $r^s \succsim r^t$, we have $\sigma_{\theta_g}(a_*^s, r^s)(\overline{a}_1) = 1$.

This implies that for every $h^s = (a^s, r^s) \succ (a_*^t, r^t)$ with $a^s \neq a_*^s$ and $h^s \in \mathcal{H}^\sigma$, we have $\mu(h^s)(\Theta_g) = 0$, so for every θ,

$$V_\theta(h^s) = u_1(\theta, \underline{a}_1, \underline{a}_2). \tag{A.51}$$

Let $\tau : \mathcal{R}_*^\sigma \to \mathbb{N} \cup \{+\infty\}$ be such that for $r^\tau \prec r^{\tau+1} \prec r^\infty$, we have:

- $\mu^*(r^\tau)(\Theta_n) > 0$ while $\mu^*(r^{\tau+1})(\Theta_n) = 0$.

Let

$$\overline{\theta}_n \equiv \max\Big\{\mathrm{supp}(\mu^*(r^t)) \bigcap \Theta_n\Big\}.$$

The second part of Lemma A.3.1 and (A.51) together imply that $\mu^*(r^\tau)(\overline{\theta}_n) > 0$. Let us examine type $\overline{\theta}_n$'s incentive at (a_*^t, r^t) to play his equilibrium strategy as opposed to play $\underline{a}_1$ all the time.

[7] Inequality (A.50) trivially applies to r^0 due to (1.10).

[8] To see this, consider three cases. If $\Theta_p = \{\varnothing\}$, then this inequality is obvious. If $\Theta_p \neq \{\varnothing\}$, then $\mathcal{D}(\theta_n, \overline{a}_1) \leq 0$ for all $\theta_n \in \Theta_n$ according to Lemma 1.4.1. When $\mathcal{D}(\theta_n, \overline{a}_1) < 0$ for all θ_n, then the inequality follows from (A.50). When $\mathcal{D}(\theta_n, \overline{a}_1) = 0$ for some $\theta_n \in \Theta_n$, then $\mathcal{D}(\theta_p, \overline{a}_1) = 0$ for all $\theta_p \in \Theta_p$. The inequality then follows from $\mathcal{D}(\theta_g, \overline{a}_1) > 0$ for all $\theta_g \in \Theta_g$ as well as $\hat{\theta} \in \Theta_g$.

This requires that:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}u_1(\overline{\theta}_n,\overline{a}_1,\alpha_{2,s})+(\delta^{\tau-t}-\delta^{\tau+1-t})u_1(\overline{\theta}_n,a_{1,\tau},\alpha_{2,\tau})+\delta^{\tau+1-t}u_1(\overline{\theta}_n,\underline{a}_1,\underline{a}_2)\Big]\geq u_1(\overline{\theta}_n,\underline{a}_1,\underline{a}_2).$$

where $\mathbb{E}[\cdot]$ is taken over $\mathcal{P}^\sigma$ and $\alpha_{2,s}\in\Delta(A_2)$ is player 2's action in period s.

Using the fact that $u_1(\overline{\theta}_n,\underline{a}_1,\underline{a}_2)\geq u_1(\overline{\theta}_n,\overline{a}_1,\overline{a}_2)$, the above inequality implies that:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}\Big(u_1(\overline{\theta}_n,\overline{a}_1,\alpha_{2,s})-u_1(\overline{\theta}_n,\overline{a}_1,\overline{a}_2)\Big)+(\delta^{\tau-t}-\delta^{\tau+1-t})\Big(u_1(\overline{\theta}_n,\underline{a}_1,\alpha_{2,\tau})-u_1(\overline{\theta}_n,\underline{a}_1,\underline{a}_2)\Big)\Big]\leq 0.$$

According to the definitions of K and M, we know that for all θ,

$$\mathbb{E}\Big[\sum_{s=t}^{\tau}(1-\delta)\delta^{s-t}\Big(u_1(\theta_n,\overline{a}_1,\alpha_{2,s})-u_1(\theta_n,\overline{a}_1,\overline{a}_2)\Big)\Big]\leq 2M(K+1)(1-\delta). \tag{A.52}$$

For every $r^\infty\in\mathcal{R}^\sigma_*$, since $r^t\notin\hat{\mathcal{R}}^\sigma_g$, we have:

$$\begin{aligned}\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta}q^*(r^{\tau(r^\infty)+1})(\theta)\mathcal{D}(\theta,\overline{a}_1) &\geq \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta_p\cup\Theta_g}q^*(r^t)(\theta)\mathcal{D}(\theta,\overline{a}_1)\\ &> \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta}q^*(r^t)(\theta)\mathcal{D}(\theta,\overline{a}_1)\geq 0\end{aligned}$$

According to Lemma A.3.7, we know that $V_\theta(a_*^{\tau(r^\infty)+1},r^{\tau(r^\infty)+1})=u_1(\theta,\overline{a}_1,\overline{a}_2)$ for all $\theta\in\Theta_g\cup\Theta_p$ and $r^\infty\in\mathcal{R}^\sigma_*$. This together with (A.52) gives the conclusion.

Second, when $\mu^*(r^t)(\Theta_n)=0$. If $t=0$, the conclusion directly follows from Proposition C.1. If $t\geq 1$ and there exists $\hat{r}^t$ such that $r^t,\hat{r}^t\succ r^{t-1}$, $(a^t_*,\hat{r}^t)\in\mathcal{H}^\sigma$ and $\hat{r}^t\in\hat{\mathcal{R}}^\sigma_g$. Then, since

$$\mu^*(r^t)=\mu^*(\hat{r}^t),$$

we have $\mu^*(\hat{r}^t)(\Theta_n)=0$. Since $\hat{r}^t\in\hat{\mathcal{R}}^\sigma_g$, according to Lemma A.3.6, $\mu^*(\hat{r}^t)=\mu^*(r^t)\in\mathcal{B}_\kappa$. The conclusion then follows from Lemma A.3.7. □

The next Lemma puts an upper bound on type $\theta_n\in\Theta_n$'s continuation payoff at (a^t_*,r^t) with $r^t\notin\hat{\mathcal{R}}^\sigma_g$.

Lemma A.3.9. *For every $\theta_n\in\Theta_n$ such that $\overline{a}_2\notin BR_2(\overline{a}_1,\theta_n)$ and $r^t\notin\hat{\mathcal{R}}^\sigma_g$ with $(a^t_*,r^t)\in\mathcal{H}^\sigma_{\theta_n}$ and $\mu^*(r^t)\notin\mathcal{B}_\kappa$, we have:*

$$V_{\theta_n}(a^t_*,r^t)\leq u_1(\theta_n,\underline{a}_1,\underline{a}_2)+2(1-\delta)M. \tag{A.53}$$

The proof is contained in the proof for the first case in Lemma A.3.8. Let

$$A(\delta)\equiv 2M(K+1)(1-\delta),\quad B(\delta)\equiv 2M(1-\delta^{\overline{T}_0(\kappa)})$$

and

$$C(\delta) \equiv 2MK|\Theta_n|(1-\delta).$$

Notice that when $\delta \to 1$, all three functions converge to 0. The next Lemma puts a uniform upper bound on player 1's payoff when $r^t \in \hat{\mathcal{R}}_g^\sigma$.

Lemma A.3.10. *When $\delta > \overline{\delta}$ and $\sigma \in NE(\delta, \mu, \phi)$, for every $r^t \in \hat{\mathcal{R}}_g^\sigma$,*

$$V_\theta(a_*^t, r^t) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - \Big(A(\delta) + B(\delta)\Big) - 2\overline{T}_1(\kappa)\Big(A(\delta) + B(\delta) + C(\delta)\Big).^{9} \tag{A.54}$$

for all θ such that $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta)$.

Proof of Lemma A.3.10: The non-trivial part of the proof deals with situations where $\mu^*(r^t) \notin \mathcal{B}_\kappa$. Since $r^t \in \hat{\mathcal{R}}_g^\sigma$, Lemma A.3.6 implies that $\mu^*(r^t)(\Theta_n) \neq 0$. Without loss of generality, assume $\Theta_n \subset \text{supp}\Big(\mu^*(r^t)\Big)$. Let me introduce $|\Theta_n| + 1$ integer valued random variables on the space $\mathcal{R}_*^\sigma$.

- $\tau : \mathcal{R}_*^\sigma \to \mathbb{N} \cup \{+\infty\}$ be the first period s along random path r^∞ such that either one of the following two conditions is met.

  1. $\mu^*(r^{s+1}) \in \mathcal{B}_{\kappa/2}$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$.
  2. $r^s \notin \hat{\mathcal{R}}_g^\sigma$.

  In the first case, there exists $a_1 \neq \overline{a}_1$ and $r^{\tau+1} \succ r^\tau$ such that

  - $\Big((a_*^\tau, a_1), r^{\tau+1}\Big) \in \mathcal{H}^\sigma(\tilde{\theta})$ for some $\tilde{\theta} \in \Theta_p \cup \Theta_n$.
  - $\mu\Big((a_*^\tau, a_1), r^{\tau+1}\Big) \notin \mathcal{B}_0$.

  Lemma A.3.3 implies the existence of $\theta \in \Theta_p \cup \Theta_n$ with $\Big((a_*^\tau, a_1), r^{\tau+1}\Big) \in \mathcal{H}^\sigma(\theta)$ such that

  $$V_\theta\Big((a_*^\tau, a_1), r^{\tau+1}\Big) = u_1(\theta, \underline{a}_1, \underline{a}_2).$$

  Suppose towards a contradiction that $\theta \in \Theta_p$, then Lemma A.3.1 implies that $\mu^*(r^{\tau+1})(\Theta_n) = 0$. Since $\mu^*(r^{\tau+1}) \in \mathcal{B}_{\kappa/2}$, Proposition A.3 implies that type θ's continuation payoff by always playing $\overline{a}_1$ is at least

  $$(1 - \delta^{\overline{T}_0(\kappa/2)}) u_1(\theta, \overline{a}_1, \underline{a}_2) + \delta^{\overline{T}_0(\kappa/2)} u_1(\theta, \overline{a}_1, \overline{a}_2),$$

  which is strictly larger than his payoff from playing a_1, which is at most $2M(1-\delta) + u_1(\theta, \underline{a}_1, \underline{a}_2)$, leading to a contradiction.

  Hence, there exists $\theta_n \in \Theta_n$ such that $V_{\theta_n}\Big((a_*^\tau, a_1), r^{\tau+1}\Big) = u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, which implies that $V_{\theta_n}\Big(a_*^\tau, r^\tau\Big) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M$.

  In the second case, Lemma A.3.9 implies that $V_{\theta_n}\Big(a_*^\tau, r^\tau\Big) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M$ for all $\theta_n \in \Theta_n$ with $r^\tau \in \mathcal{H}^\sigma(\theta_n)$.

[9] One can further tighten this bound. However, (A.54) is sufficient for the purpose of proving Theorem 1.2.

- For every $\theta_n \in \Theta_n$, let $\tau_{\theta_n} : \mathcal{R}_*^\sigma \to \mathbb{N} \cup \{+\infty\}$ be the first period s along random path r^∞ such that either one of the following three conditions is met.

  1. $\mu^*(r^{s+1}) \in \mathcal{B}_{\kappa/2}$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$.
  2. $r^s \notin \hat{\mathcal{R}}_g^\sigma$.
  3. $\mu^*(r^{s+1})(\theta_n) = 0$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$, .

By definition, $\tau \geq \tau_{\theta_n}$, so $\tau \geq \max_{\theta_n \in \Theta_n} \{\tau_{\theta_n}\}$. Next, I show that

$$\tau = \max_{\theta_n \in \Theta_n} \{\tau_{\theta_n}\}. \tag{A.55}$$

Suppose on the contrary that $\tau > \max_{\theta_n \in \Theta_n} \{\tau_{\theta_n}\}$ for some $r^\infty \in \mathcal{R}_*^\sigma$. Then there exists $(a_*^s, r^s) \succsim (a_*^t, r^t)$ such that $r^s \in \hat{\mathcal{R}}_g^\sigma$, $\mu^*(r^s) \notin \mathcal{B}_\kappa$ and $\mu^*(r^s)(\Theta_n) = 0$, which contradicts Lemma A.3.6 when $\delta > \overline{\delta}$.

Next, I show by induction over $|\Theta_n|$ that

$$\mathbb{E}\Big[\sum_{s=t}^{\tau} (1-\delta)\delta^{\tau-t}\Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \hat{\alpha}_{2,s})\Big)\Big] \leq 2MK|\Theta_n|(1-\delta), \tag{A.56}$$

for all $\theta \in \Theta$ and

$$V_{\tilde{\theta}_n}(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M, \tag{A.57}$$

for

$$\tilde{\theta} \equiv \min\Big\{\Theta_n \cap \operatorname{supp}\Big(\mu^*(r^{\tau_{\theta_n}+1})\Big)\Big\}$$

with $\theta_n, \tilde{\theta}_n \in \Theta_n$, where $\mathbb{E}[\cdot]$ is taken over $\mathcal{P}^\sigma$ and $\hat{\alpha}_{2,s} \in \Delta(A_2)$ is player 2's (mixed) action at (a_*^s, r^s).

When $|\Theta_n| = 1$, let θ_n be its unique element. Consider player 1's pure strategy of playing $\overline{a}_1$ until r^τ and then play $\underline{a}_1$ forever. This is one of type θ_n's best responses according to (A.55), which results in payoff at most:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1} (1-\delta)\delta^{s-t} u_1(\theta_n, \overline{a}_1, \hat{\alpha}_{2,s}) + \delta^{\tau-t}\Big(u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M\Big)\Big].$$

The above expression cannot be smaller than $u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, which is the payoff he can guarantee by always playing $\underline{a}_1$. Since $u_1(\theta_n, \underline{a}_1, \underline{a}_2) \geq u_1(\theta_n, \overline{a}_1, \overline{a}_2)$, and from the definition of K, we get for all θ,

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1} (1-\delta)\delta^{s-t}\Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \hat{\alpha}_{2,s})\Big)\Big] \leq 2MK(1-\delta).$$

We can then obtain (A.57) for free since $\tau = \tau_{\theta_n}$ and type θ_n's continuation value at (a_*^τ, r^τ) is at most $u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M$ by Lemma A.3.3.

Suppose the conclusion holds for all $|\Theta_n| \leq k-1$, consider when $|\Theta_n| = k$ and let $\theta_n \equiv \min \Theta_n$. If $(a_*^\tau, r^\tau) \notin \mathcal{H}^\sigma(\theta_n)$, then there exists $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \prec (a_*^\tau, r^\tau)$ with $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \in \mathcal{H}^\sigma(\theta_n)$ at which type θ_n plays $\overline{a}_1$ with probability 0. I put an upper bound on type θ_n's continuation payoff at $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$ by examining type $\tilde{\theta}_n \in \Theta_n \backslash \{\theta_n\}$'s incentive to play $\overline{a}_1$ at $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$, where

$$\tilde{\theta} \equiv \min \Big\{ \Theta_n \cap \operatorname{supp}\Big(\mu^*(r^{\tau_{\theta_n}+1})\Big) \Big\}$$

This requires that:

$$\mathbb{E}\Big[\sum_{s=0}^{\infty}(1-\delta)\delta^s u_1(\tilde{\theta}_n, \alpha_{1,s}, \alpha_{2,s})\Big] \leq \underbrace{u_1(\tilde{\theta}_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M}_{\text{by induction hypothesis}}.$$

where $\{(\alpha_{1,s}, \alpha_{2,s})\}_{s\in\mathbb{N}}$ is the equilibrium continuation play following $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$. By definition, $\tilde{\theta}_n \succ \theta_n$, so the supermodularity condition implies that:

$$u_1(\theta_n, \underline{a}_1, \underline{a}_2) - u_1(\tilde{\theta}_n, \underline{a}_1, \underline{a}_2) \geq u_1(\theta_n, \alpha_{1,s}, \alpha_{2,s}) - u_1(\tilde{\theta}_n, \alpha_{1,s}, \alpha_{2,s}).$$

Therefore, we have:

$$\begin{aligned}
V_{\theta_n}(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) &= \mathbb{E}\Big[\sum_{s=0}^{\infty}(1-\delta)\delta^s u_1(\theta_n, \alpha_{1,s}, \alpha_{2,s})\Big] \\
&\leq \mathbb{E}\Big[\sum_{s=0}^{\infty}(1-\delta)\delta^s \Big(u_1(\tilde{\theta}_n, \alpha_{1,s}, \alpha_{2,s}) + u_1(\theta_n, \underline{a}_1, \underline{a}_2) - u_1(\tilde{\theta}_n, \underline{a}_1, \underline{a}_2)\Big)\Big] \\
&\leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M.
\end{aligned}$$

Back to type θ_n's incentive constraint. Since it is optimal for him to play $\overline{a}_1$ until $r^{\tau_{\theta_n}}$ and then play $\underline{a}_1$ forever, doing so must give him a higher payoff than playing $\underline{a}_1$ forever starting from r^t, which gives:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau_{\theta_n}-1}(1-\delta)\delta^{s-t} u_1(\theta_n, \overline{a}_1, \hat{\alpha}_{2,s}) + \delta^{\tau_{\theta_n}}\Big(u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M\Big)\Big] \geq u_1(\theta_n, \underline{a}_1, \underline{a}_2).$$

This implies that:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau_{\theta_n}-1}(1-\delta)\delta^{s-t}\Big(u_1(\theta_n, \overline{a}_1, \overline{a}_2) - u_1(\theta_n, \overline{a}_1, \hat{\alpha}_{2,s})\Big)\Big] \leq 2M(1-\delta),$$

which also implies that for every $\theta \in \Theta$,

$$\mathbb{E}\Big[\sum_{s=t}^{\tau_{\theta_n}-1}(1-\delta)\delta^{s-t}\Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \hat{\alpha}_{2,s})\Big)\Big] \leq 2MK(1-\delta). \tag{A.58}$$

When $\tau > \tau_{\theta_n}$, the induction hypothesis implies that:

$$\mathbb{E}\Big[\sum_{s=\tau_{\theta_n}}^{\tau_\theta - 1}(1-\delta)\delta^{s-\tau_{\theta_n}}\Big(u_1(\theta,\overline{a}_1,\overline{a}_2) - u_1(\theta,\overline{a}_1,\alpha_{2,s})\Big)\Big] \leq 2MK(k-1)(1-\delta). \tag{A.59}$$

According to (A.58) and (A.59).

$$\mathbb{E}\Big[\sum_{s=t}^{\tau}(1-\delta)\delta^{\tau-t}\Big(u_1(\theta,\overline{a}_1,\overline{a}_2) - u_1(\theta,\overline{a}_1,\hat{\alpha}_{2,s})\Big)\Big] \leq 2MKk(1-\delta),$$

which shows (A.56) when $|\Theta_n| = k$. (A.57) can be obtained directly from the induction hypothesis.

Next, I examine player 1's continuation payoff at on-path histories following $(a_*^{\tau+1}, r^{\tau+1}) \in \mathcal{H}^\sigma$. I consider three cases:

1. If $r^{\tau+1} \notin \hat{\mathcal{R}}_g^\sigma$, by Lemma A.3.8, then for every θ,

$$V_\theta(a_*^{\tau+1}, r^{\tau+1}) \geq u_1(\theta,\overline{a}_1,\overline{a}_2) - A(\delta).$$

2. If $r^{\tau+1} \in \hat{\mathcal{R}}_g^\sigma$ and $\mu^*(r^s) \in \mathcal{B}_\kappa$ for all r^s satisfying $r^s \succsim r^{\tau+1}$ and $r^s \in \hat{\mathcal{R}}_g^\sigma$, then for every θ,

$$V_\theta(a_*^{\tau+1}, r^{\tau+1}) \geq u_1(\theta,\overline{a}_1,\overline{a}_2) - B(\delta).$$

3. If there exists r^s such that $\mu^*(r^s) \notin \mathcal{B}_\kappa$ with $r^s \succsim r^{\tau+1}$ and $r^s \in \hat{\mathcal{R}}_g^\sigma$, then repeat the procedure in the beginning of this proof by defining random variables

   - $\tau' : \mathcal{R}_*^\sigma \to \{n \in \mathbb{N} \cup \{+\infty\} | n \geq s\}$
   - $\tau'_{\theta_n} : \mathcal{R}_*^\sigma \to \{n \in \mathbb{N} \cup \{+\infty\} | n \geq s\}$

   similarly as we have defined τ and τ_{θ_n}, and then examine continuation payoffs at $r^{\tau'+1}$...
   Since $\mu^*(r^{\tau+1}) \in \mathcal{B}_{\kappa/2}$ but $\mu^*(r^s) \notin \mathcal{B}_\kappa$, then

$$\sum_{\theta\in\Theta}\Big(q^*(r^{\tau+1})(\theta) - q^*(r^s)(\theta)\Big) \geq \frac{\rho_1(\kappa)}{2}. \tag{A.60}$$

   Therefore, such iterations can last for at most $2\overline{T}_1(\kappa)$ rounds.

Next, I establish the payoff lower bound in case 3. I introduce a new concept called '*trees*'. Let

$$\mathcal{R}_b^\sigma \equiv \Big\{r^t \Big| \mu^*(r^t) \notin \mathcal{B}_\kappa \text{ and } r^t \in \hat{\mathcal{R}}_g^\sigma\Big\}$$

Define set $\mathcal{R}^\sigma(k) \subset \mathcal{R}$ for all $k \in \mathbb{N}$ recursively as follows. Let

$$\mathcal{R}^\sigma(1) \equiv \Big\{r^t \Big| r^t \in \mathcal{R}_b^\sigma \text{ and there exists no } r^s \prec r^t \text{ such that } r^s \in \mathcal{R}_b^\sigma\Big\}.$$

For every $r^t \in \mathcal{R}^\sigma(1)$, let $\tau[r^t] : \mathcal{R}^\sigma_* \to \mathbb{N} \cup \{+\infty\}$ as the first period $s > t$ (starting from r^t) such that either one of the following two conditions is met:

1. $\mu^*(r^{s+1}) \in \mathcal{B}_{\kappa/2}$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$.
2. $r^s \notin \hat{\mathcal{R}}^\sigma_g$.

I call

$$\mathcal{T}(r^t) \equiv \left\{ r^s \middle| r^{\tau[r^{t_1}]} \succsim r^s \succsim r^t \right\}$$

a '*tree*' with root r^t. For any $k \geq 2$, let

$$\mathcal{R}^\sigma(k) \equiv \Big\{ r^t \Big| r^t \in \mathcal{R}^\sigma_b,\ r^t \succ r^{\tau[r^s]} \text{ for some } r^s \in \mathcal{R}^\sigma(k-1)$$

$$\text{and there exists no } r^s \prec r^t \text{ that satisfy these two conditions} \Big\}.$$

Let T be the largest integer such that $\mathcal{R}^\sigma(T) \neq \{\varnothing\}$. According to (A.60), we know that $T \leq 2\overline{T}_1(\kappa)$. Similarly, we can define trees with roots in $\mathcal{R}(k)$ for every $k \leq T$.

In what follows, I show that for every θ and every $r^t \in \mathcal{R}^\sigma(k)$,

$$V_\theta(a_*^t, r^t) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - (T + 1 - k)\Big(A(\delta) + B(\delta) + C(\delta)\Big). \tag{A.61}$$

The proof is done by inducting on k from backwards. When $k = T$, player 1's continuation value at $(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$ is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta) - B(\delta)$ according to Lemma A.3.2 and Lemma A.3.8. His continuation value at r^t is at least:

$$u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta) - B(\delta) - C(\delta).$$

Suppose the conclusion holds for all $k \geq n + 1$, then when $k = n$, type θ's continuation payoff at (a_*^t, r^t) is at least:

$$\mathbb{E}\Big[(1 - \delta^{\tau[r^t]-t})u_1(\theta, \overline{a}_1, \overline{a}_2) + \delta^{\tau[r^t]-t} V_\theta(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})\Big] - C(\delta)$$

Pick any $(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$, consider the set of random paths r^∞ that it is consistent with, let this set be $\mathcal{R}^\infty(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$. Partition it into two subsets:

- $\mathcal{R}^\infty_+(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$ consists of r^∞ such that for all $s \geq \tau[r^t] + 1$ and $r^s \prec r^\infty$, we have $r^s \notin \mathcal{R}^\sigma_b$.
- $\mathcal{R}^\infty_-(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$ consists of r^∞ such that there exists $s \geq \tau[r^t] + 1$ and $r^s \prec r^\infty$ at which $r^s \in \mathcal{R}^\sigma(n+1)$.

Conditional on $r^\infty \in \mathcal{R}^\infty_+(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$, we have:

$$V_\theta(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1}) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta) - B(\delta).$$

Conditional on $r^\infty \in \mathcal{R}^\infty_{-}(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$, type θ's continuation payoff is no less than

$$V_\theta(a_*^s, r^s) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - (T-n)\Big(A(\delta) + B(\delta) + C(\delta)\Big)$$

after reaching $r^s \in \mathcal{R}^\sigma(n)$ according to the induction hypothesis. Moreover, since can he lose at most $A(\delta) + B(\delta)$ before reaching r^s (according to Lemmas A.3.2 and A.3.8), we have:

$$V_\theta(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1}) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - (T+1-n)\Big(A(\delta) + B(\delta) + C(\delta)\Big).$$

which obtains (A.61). (A.54) is implied by (A.61) since player 1's loss is bounded above by $A(\delta) + B(\delta)$ from r^0 to any $r^t \in \mathcal{R}^\sigma(0)$. □

Theorem 1.2 is then implied by the conclusions of Lemma A.3.8, A.3.9 and A.3.10.

A.4 Proof of Theorem 1.3 & Extensions

In this Appendix, I show Theorem 1.3. I also generalize the result by allowing for multiple commitment types. Recall the definitions of $\mathcal{H}^\sigma$ and $\mathcal{H}^\sigma(\omega)$ in the previous section.

A.4.1 Proof of Theorem 1.3

Step 1: Let

$$X(h^t) \equiv \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} q(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1). \tag{A.62}$$

According to (1.11), $X(h^0) < 0$. Moreover, at every $h^t \in \mathcal{H}^\sigma$ with $X(h^t) < 0$, player 2 has a strict incentive to play $\underline{a}_2$. Applying Lemma A.3.3, there exists $\theta_p \in \Theta_p$ with $h^t \in \mathcal{H}(\theta_p)$ such that type θ_p's continuation value at h^t is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, which further implies that always playing $\underline{a}_1$ is his best reply. According to Lemma A.3.1 and using the fact that $X(h^0) < 0$, every $\theta_n \in \Theta_n$ plays $\underline{a}_1$ for sure at every $h^t \in \mathcal{H}(\theta_n)$.

Step 2: Let us examine the equilibrium behaviors of the types in $\Theta_p \cup \Theta_g$. I claim that for every $h^1 = (\overline{a}_1, r^1) \in \mathcal{H}^\sigma$,

$$\sum_{\theta \in \Theta_g \cup \Theta_p} q(h^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0. \tag{A.63}$$

Suppose towards a contradiction that $\sum_{\theta \in \Theta_g \cup \Theta_p} q(h^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$, then $X(h^1) \geq \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)$. According to Proposition **??**, there exists $K \in \mathbb{R}_+$ independent of δ such that type θ's continuation payoff is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)K$ at every $h^1_* \in \mathcal{H}^\sigma$. When δ is large enough, this contradicts the conclusion in the previous step that there exists $\theta_p \in \Theta_p$ such that type θ_p's continuation value at h^0 is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, as he can profitably deviate by playing $\overline{a}_1$ in period 0.

Step 3: According to (A.63), we have $\mu^*(r^1) \notin \mathcal{B}_0$. So Lemma A.3.6 implies that $r^1 \notin \hat{\mathcal{R}}^\sigma_g$. According to Lemma A.3.1, type θ_g plays $\overline{a}_1$ at every $h^t \in \mathcal{H}(\theta_g)$ with $t \geq 1$ for every $\theta_g \in \Theta_g$.

Next, I show that $r^0 \notin \hat{\mathcal{R}}_g^\sigma$. Suppose towards a contradiction that $r^0 \in \hat{\mathcal{R}}_g^\sigma$, then there exists $h^T = (a_*^T, r^T) \in \mathcal{H}^\sigma$ such that $\mu(h^T)(\Theta_p \cup \Theta_n) = 0$. If $T \geq 2$, it contradicts our previous conclusion that $r^1 \notin \hat{\mathcal{R}}_g^\sigma$. If $T = 1$, then it contradicts (A.63). Therefore, we have $r^0 \notin \hat{\mathcal{R}}_g^\sigma$ and we have shown that type θ_g plays $\overline{a}_1$ at every $h^t \in \mathcal{H}(\theta_g)$ with $t \geq 0$ for every $\theta_g \in \Theta_g$.

Step 4: In the last step, I pin down the strategies of type θ_p by showing that $X(h^t) = 0$ for every $h^t = (a_*^t, r^t) \in \mathcal{H}^\sigma$ with $t \geq 1$. First, I show that $X(h^1) = 0$. The argument at other histories follows similarly.

Suppose first that $X(h^1) > 0$, then according to Lemma A.3.7, type θ_p's continuation payoff at (a_*^{t+1}, r^{t+1}) is $u_1(\theta_p, \overline{a}_1, \overline{a}_2)$ by always playing $\overline{a}_1$, while his continuation payoff at (a_*^t, a_1, r^{t+1}) is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, leading to a contradiction. Suppose next that $X(h^1) < 0$, similar to the previous paragraph, there exists type $\theta_p \in \Theta_p$ with $h^1 \in \mathcal{H}(\theta_p)$ such that his incentive constraint is violated. Similarly, one can show that $X(h^t) = 0$ for every $t \geq 1$, $h^t = (a_*^t, r^t) \in \mathcal{H}^\sigma$. Hence, we have established the uniqueness of player 1's equilibrium play.

A.4.2 Generalizations to Multiple Commitment Types

Next, I generalize Theorem 1.3 by accommodating multiple commitment types. For every $\theta \in \Theta$, let $\lambda(\theta)$ be the likelihood ratio between strategic type θ and the lowest strategic type $\underline{\theta} \equiv \min \Theta$ and let $\lambda \equiv \{\lambda(\theta)\}_{\theta \in \Theta}$ be the likelihood ratio vector between strategic types. I use this likelihood ratio vector to characterize the sufficient conditions for behavioral uniqueness as the result under multiple commitment type requires that the total probability of commitment types being *small enough*. The upper bound of this probability depends on the distribution of strategic types. Let

$$\Omega^g \equiv \{\alpha_1 \in \Omega^m \backslash \{\overline{a}_1\} | \mathcal{D}(\alpha_1, \phi_{\alpha_1}) > 0\}$$

which are the set of commitment types under which player 2 has a strict incentive to play $\overline{a}_2$.

For every $t \geq 1$ and $h^t \in \mathcal{H}^t$, let $h_1^t \equiv \{a_{1,0}, ..., a_{1,t-1}\}$ be the projection of h^t on $\times_{s=0}^{t-1} A_{1,s}$. Let $\mathcal{H}_1^t$ be the set of h_1^t. Let $\overline{\mathcal{H}}_1^t \equiv \{(\overline{a}_1, ...\overline{a}_1), (\underline{a}_1, ...\underline{a}_1)\}$. For every probability measure $\mathcal{P}$ over $\mathcal{H}$, let $\mathcal{P}_{1,t}$ be its projection on $\mathcal{H}_1^t$. Recall that $\mathcal{P}^\sigma(\theta)$ is the probability measure over $\mathcal{H}$ under strategy profile σ conditional on player 1 being strategic type θ. For every $\gamma \geq 0$ and two Nash equilibria σ and σ', strategic type θ's on-path behavior is γ-close between these equilibria if for every $t \geq 1$,

$$D_B\Big(\mathcal{P}_{1,t}^\sigma(\theta), \mathcal{P}_{1,t}^{\sigma'}(\theta)\Big) \leq \gamma,$$

where $D_B(p,q)$ denotes the Bhattacharyya distance between distributions p and q.[10] If $\gamma = 0$, then

[10] One can replace the Bhattacharyya distance with the Rényi divergence or Kullback-Leibler divergence in the following way: strategic type θ's on-path behavior is γ-close between σ and σ' if there exists a probability measure $\mathcal{P}$ on $\mathcal{H}$ such that for every $t \geq 1$,

$$\max\Big\{D\Big(\mathcal{P}_{1,t}\Big|\Big|\mathcal{P}_{1,t}^\sigma(\theta)\Big), D\Big(\mathcal{P}_{1,t}\Big|\Big|\mathcal{P}_{1,t}^\sigma(\theta')\Big)\Big\} \leq \gamma,$$

where $D(\cdot||\cdot)$ is either the Rényi divergence of order greater than 1 or the Kullback-Leibler divergence.

type θ_p^*'s on-path behavior is the same between these equilibria. Intuitively, the above distance measures the difference between the *ex ante* distributions over player 1's action paths across different equilibria. The generalization of Theorem 1.3 to multiple commitment types is stated below as Theorem A.1:

Theorem A.1. *Suppose $\overline{a}_1 \in \Omega^m$ and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, then for every $\lambda \in [0, +\infty)^{|\Theta|}$ satisfying:*

$$\sum_{\theta \in \Theta_p \cup \Theta_g} \lambda(\theta) \mathcal{D}(\theta, \overline{a}_1) < 0, \tag{A.64}$$

1. *There exist $\overline{\epsilon} > 0$ and a function $\gamma : (0, \overline{\epsilon}) \to \mathbb{R}_+$ satisfying $\lim_{\epsilon \downarrow 0} \gamma(\epsilon) = 0$, such that for every $\mu \in \Delta(\Omega)$ with $\{\mu(\theta)/\mu(\underline{\theta})\}_{\theta \in \Theta} = \lambda$ and $\mu(\Omega^m) < \overline{\epsilon}$, there exist $\overline{\delta} \in (0, 1)$ and $\theta_p^* \in \Theta_p$ such that for every $\sigma \in NE(\delta, \mu, \phi)$ with $\delta > \overline{\delta}$:*
   - *For every $\theta \succ \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\overline{a}_1$ at h^t.*
   - *For every $\theta \prec \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\underline{a}_1$ at h^t.*
   - *As for type θ_p^*,*
     - ◇ $\mathcal{P}_{1,t}^\sigma(\theta_p^*)(\overline{\mathcal{H}_1^t}) > 1 - \gamma(\overline{\epsilon})$ *for every $t \geq 1$.*
     - ◇ *For every $\sigma' \in NE(\delta, \mu, \phi)$, type θ_p^*'s on-path behavior is $\gamma(\epsilon)$-close between σ and σ'.*
     - ◇ *Furthermore, if there exists no mixed commitment type under which $\overline{a}_2$ is player 2's strict best reply, then type θ_p^*'s on-path behavior is the same across all equilibria under generic parameters. Type θ_p^* plays the same action in every period with (ex ante) probability 1.*
   - *Strategic type θ's equilibrium payoff is v_θ^* for every $\theta \in \Theta$.*

2. *If all commitment types are playing pure strategies, then there exists $\overline{\epsilon} > 0$, such that for every $\mu \in \Delta(\Omega)$ with $\{\mu(\theta)/\mu(\underline{\theta})\}_{\theta \in \Theta} = \lambda$ and $\mu(\Omega^m) < \overline{\epsilon}$, there exist $\overline{\delta} \in (0, 1)$ and $\theta_p^* \in \Theta_p$ such that for every $\sigma \in NE(\delta, \mu, \phi)$ with $\delta > \overline{\delta}$:*
   - *For every $\theta \succ \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\overline{a}_1$ at h^t.*
   - *For every $\theta \prec \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\underline{a}_1$ at h^t.*
   - *There exists $\alpha_1 \in \Delta(\Omega^g \cup \{\overline{a}_1, \underline{a}_1\})$ such that for every $h^t \in \mathcal{H}^\sigma(\theta_p^*)$,*
     - ◇ *Type θ_p^* plays α_1 at h^0.*
     - ◇ *If $t \geq 1$ and there exists $a_1 \in \Omega^g \cup \{\overline{a}_1, \underline{a}_1\}$ such that $a_1 \in h^t$, then type θ_p^* plays a_1 at h^t.*
   - *Strategic type θ's equilibrium payoff is v_θ^* for every $\theta \in \Theta$.*

Let me comment on the conditions in Theorem A.1. First, (A.64) is implied by (1.11) given that $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$. Second, when there are other commitment types under which player 2 has an incentive to play $\overline{a}_2$, then payoff and behavior uniqueness require the total probability of commitment types to be small enough. Intuitively, this is because the presence of multiple good commitment types gives the strategic player 1 many good reputations to choose from, which can lead to multiple behaviors and payoffs. An example is shown in Appendix A.7.5.

Next, I provide a sufficient condition on $\bar{\epsilon}$. Let

$$Y(h^t) \equiv \mu(h^t)(\bar{a}_1)\mathcal{D}(\phi_{\bar{a}_1}, \bar{a}_1) + \sum_{\alpha_1 \in \Omega^g} \mu(h^t)(\alpha_1)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} \mu(h^t)(\theta)\mathcal{D}(\theta, \bar{a}_1), \tag{A.65}$$

which is an upper bound on player 2's incentive to play $\bar{a}_2$ at h^t. I require $\bar{\epsilon}$ to be small enough such that

$$\bar{\epsilon} \max_{\widetilde{\theta} \in \Theta} \{\mathcal{D}(\widetilde{\theta}, \bar{a}_1)\} + (1-\bar{\epsilon}) \sum_{\theta \in \Theta_p \cup \Theta_g} \frac{\lambda(\theta)}{\sum_{\tilde{\theta} \in \Theta} \lambda(\tilde{\theta})} \mathcal{D}(\theta, \bar{a}_1) < 0. \tag{A.66}$$

Such $\bar{\epsilon}$ exists since $\sum_{\theta \in \Theta_p \cup \Theta_g} \lambda(\theta)\mathcal{D}(\theta, \bar{a}_1) < 0$. Inequality (A.66) implies that $Y(h^0) < 0$, which is also equivalent to (1.11) when $\Omega^g = \{\varnothing\}$.

Third, when there are mixed strategy commitment types, the probabilities with which type θ_p^* mixes may not be the same across all equilibria. Intuitively, this is because of two reasons. First, suppose player 2 has no incentive to play $\bar{a}_2$ under any mixed commitment type, then given that all strategic types either always plays $\bar{a}_1$ or always plays $\underline{a}_1$, player 2's incentive to play $\bar{a}_2$ is increasing over time as more $\bar{a}_1$ has been observed. As a result, there will be $T(\delta)$ periods in which player 2 has a strict incentive to play $\bar{a}_2$, followed by at most one period in which she is indifferent between $\bar{a}_2$ and $\underline{a}_2$, followed by periods in which she has a strict incentive to play $\bar{a}_2$, with $T(\delta)$ and the probabilities with which she mix between $\bar{a}_2$ and $\underline{a}_2$ in period $T(\delta)$ pinned down by type θ_p^*'s indifference condition in period 0. Under degenerate parameter values in which there exists an integer T such that type θ_p^* is just indifferent between always playing $\underline{a}_1$ and always playing $\bar{a}_1$ when $\underline{a}_2$ will be played in the first T periods, his mixing probability between always playing $\bar{a}_1$ and always playing $\underline{a}_1$ is not unique. Nevertheless, when the ex ante probability of Ω^m is smaller than ϵ, his probability of mixing cannot vary by more than $\gamma(\epsilon)$ even in this degenerate case, with $\gamma(\cdot)$ diminishes as $\epsilon \downarrow 0$. Second, when there are good mixed strategy commitment types, the probability with which type θ_p^* behaves inconsistently and builds a reputation for being a good mixed strategy commitment type cannot be uniquely pinned down by his equilibrium payoff. Nevertheless, the differences between these probabilities across different equilibria will vanish as the total probability of commitment types vanishes. Intuitively, this is because if type θ_p^* imitates the mixed commitment type with significant probability, then player 2 will have a strict incentive to play $\underline{a}_2$. This implies that as the probability of commitment type vanishes, the probability with which type θ_p^* builds a mixed reputation also vanishes.

A.4.3 Proof of Theorem A.1

Unique Equilibrium Behavior of Strategic Types when $\theta \in \Theta_n \cup \Theta_g$: This part of the proof is similar to the proof of Theorem 1.3, by replacing $X(h^t)$ with $Y(h^t)$. First, I show that every type $\theta_n \in \Theta_n$ will play $\underline{a}_1$ at every $h^t \in \mathcal{H}^\sigma(\theta_n)$ in every equilibrium σ. This is similar to

Step 1 in the proof of Theorem 1.3. Since $Y(h^0) < 0$ and at every $h^t \in \mathcal{H}^\sigma$ with $Y(h^t) < 0$, player 2 has a strict incentive to play $\underline{a}_2$. Applying Lemma A.3.3, there exists $\theta_p \in \Theta_p$ with $h^t \in \mathcal{H}(\theta_p)$ such that type θ_p's continuation value at h^t is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, and hence always playing $\underline{a}_1$ is his best reply. Type θ_n's on-path behavior is pinned down by Lemma A.3.1.

Next, I establish (A.63). Suppose towards a contradiction that $\sum_{\theta \in \Theta_g \cup \Theta_p} q(h^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$, then $Y(h^1) \geq \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)$. According to Theorem 1.2, there exists $K \in \mathbb{R}_+$ independent of δ such that type θ's continuation payoff is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)K$ at every $h^1_* \in \mathcal{H}^\sigma$. When δ is large enough, this contradicts the conclusion in the previous step that there exists $\theta_p \in \Theta_p$ such that type θ_p's continuation value at h^0 is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, as he can profitably deviate by playing $\overline{a}_1$ in period 0. According to (A.63), we have $\mu^*(r^1) \notin \mathcal{B}_0$. Following the same procedure, one can show that $r^1 \notin \hat{\mathcal{R}}^\sigma_g$. and $r^0 \notin \hat{\mathcal{R}}^\sigma_g$. This implies that for every equilibrium σ, type θ_g plays $\overline{a}_1$ at every $h^t \in \mathcal{H}^\sigma(\theta_g)$ for every $\theta_g \in \Theta_g$.

Consistency of Equilibrium Behavior and Generic Uniqueness of Equilibrium Payoff when $\theta \in \Theta_p$: Let Ω^{gm} be the set of mixed strategy commitment types under which player 2 has a strict incentive to play $\overline{a}_2$. I show that when $\Omega^{gm} = \{\varnothing\}$, type θ_p has to behave consistently over time for every $\theta_p \in \Theta_p$. For every $t \geq 1$, let

$$Z(h^t) \equiv \mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\alpha_1 \in \widehat{\Omega}^b} q(h^t)(\alpha_1)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \tag{A.67}$$

where $\widehat{\Omega}^b \equiv \{\alpha_1 \in \Omega^m \backslash \{\overline{a}_1\} | \mathcal{D}(\alpha_1, \phi_{\alpha_1}) < 0\}$. If $\Omega^{gm} = \{\varnothing\}$, then $\mu(h^t)(\Omega^g) = 0$ for every $h^t = (a^t_*, r^t) \in \mathcal{H}^\sigma$ with $t \geq 1$. Therefore, player 2 has a strict incentive to play $\underline{a}_2$ if $Z(h^t) < 0$. Moreover, according to the conclusion in the previous step that type $\theta_g \in \Theta_g$ plays $\overline{a}_1$ for sure at every $h^t = (a^t_*, r^t)$, we know that for every $h^t \succ h^{t-1}$, we have $Z(h^t) \geq Z(h^{t-1})$.

Subcase 1: No Mixed Commitment Types Consider the case where there exists no $\alpha_1 \in \widehat{\Omega}^b$ such that $\alpha_1 \notin A_1$, i.e. there are no mixed strategy commitment types that affect player 2's best reply. By definition, $Z(h^t) = X(h^t)$ for every $t \geq 1$. As shown in Theorem 1.3, we know that $Z(h^t) = 0$ for every $h^t = (a^t_*, r^t) \in \mathcal{H}^\sigma$ and $t \geq 1$. When $\Omega^g \neq \{\varnothing\}$, let $\Omega^g \equiv \{a^1_1, ..., a^{n-1}_1\}$ with $a^1_1 \prec a^2_1 \prec ... \prec a^{n-1}_1 \prec a^n_1 \equiv \overline{a}_1$. There exists $q : \Theta_p \to \Delta(\Omega^g \cup \{\underline{a}_1\})$ such that:

- **Monotonicity:** For every $\theta_p \succ \theta'_p$ and $a^i_1 \in \Omega^g \cup \{\underline{a}_1\}$. First, if $q(\theta_p)(a^i_1) > 0$, then $q(\theta'_p)(a^j_1) = 0$ for every $a^j_1 \succ a^i_1$. Second, if $q(\theta'_p)(a^i_1) > 0$, then $q(\theta_p)(a^j_1) = 0$ for every $a^j_1 \prec a^i_1$.

- **Indifference:** For every $a^i_1 \in \Omega^g \backslash \{\underline{a}_1\}$, we have:

$$\mu(a^i_1)\mathcal{D}(\phi_{a^i_1}, a^i_1) + \sum_{\theta_p \in \Theta_p} \mu(\theta_p)q(\theta_p)(a^i_1)\mathcal{D}(\theta_p, a^i_1) = 0. \tag{A.68}$$

These two conditions uniquely pin down function $q(\cdot)$, and therefore, the behavior of every type in Θ_p. In player 1's unique equilibrium behavior, every strategic type always replicates his action in period 0.

Subcase 2: Presence of Mixed Commitment Types Consider the case where there are mixed strategy commitment types. Player 1's action path $a^t = (a_{1,0}, ..., a_{1,t-1})$ (with $t \geq 1$) is '*inconsistent*' if there exists no $a_1 \in \Omega^g \cup \{\overline{a}_1\} \cup \{\underline{a}_1\}$ such that $a_{1,0} = ... = a_{1,t-1} = a_1$. Otherwise, it is consistent. A history is '*consistent*' or (inconsistent) if the action path it contains is consistent (or inconsistent). Since $\Omega^m = \{\varnothing\}$ and the types in Θ_g are always playing $\overline{a}_1$, so type θ's continuation value at every on-path inconsistent history must be $u_1(\theta, \underline{a}_1, \underline{a}_2)$ for every $\theta \in \Theta$.

I show that in every equilibrium, type θ_p's behavior must be consistent for every $\theta_p \in \Theta_p$. Let

$$W(h^t) \equiv \mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1). \tag{A.69}$$

For every consistent history h^t where a_1 is the consistent action, we know that $W(h^t) \leq Z(h^t)$ since

$$\sum_{\alpha_1 \in \widehat{\Omega}^b} q(h^t)(\alpha_1)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) \leq 0.$$

As shown in the proof of Theorem 1.3, we know that $W(h^t) \geq 0$. Moreover, similar argument shows that:

1. If there exists $\alpha_1 \in \widehat{\Omega}^b$ such that $\alpha_1(a_1) > 0$, then $W(h^t) > 0$.

2. If there exists no such α_1, then $W(h^t) = 0$.

The consistency of type θ_p's behavior at the 2nd class of consistent histories directly follows from the argument in Theorem 1.3. In what follows, I focus on the 1st class of consistent histories.

For every consistent history h^t with $W(h^t) > 0$ and $\mu(h^t)(\Theta_p) \neq 0$, let $\underline{\theta}_p$ be the lowest type in the support of $\mu(h^t)$. According to Lemma A.3.3, his expected payoff at any subsequent inconsistent history is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, i.e. playing $\underline{a}_1$ all the time is his best reply. According to Lemma A.3.1, if there exists $\theta_p \in \Theta_p$ playing inconsistently at h^t, then type $\underline{\theta}_p$ must be playing inconsistently at h^t with probability 1.

Suppose type $\underline{\theta}_p$ plays inconsistently with positive probability at h^t with $Z(h^t) \leq 0$, then his continuation value at h^t is $u_1(\underline{\theta}_p, \underline{a}_1, \underline{a}_2)$. He strictly prefers to deviate and play $\underline{a}_1$ forever at $h^{t-1} \prec h^t$ unless there exists $\hat{h}^T \succ h^{t-1}$ such that $Z(\hat{h}^T) \geq 0$ and type $\underline{\theta}_p$ strictly prefers to play consistently from h^{t-1} to $\hat{h}^T$. This implies that every θ_p plays consistently with probability 1 from h^{t-1} to $\hat{h}^T$, i.e. for every $h^t \succ h^{t-1}$ in which type θ_p plays inconsistently with positive probability and $h^T \succ h^t$, we have $Z(h^T) > Z(\hat{h}^T) \geq 0$. This implies that at h^t, type $\underline{\theta}_p$'s continuation payoff by playing consistently until $Z \geq 0$ is strictly higher than behaving inconsistently, leading to a contradiction.

Suppose type $\underline{\theta}_p$ plays inconsistently with positive probability at h^t with $Z(h^t) > 0$, then according to Lemma A.3.7, his continuation value by playing consistently is at least $u_1(\underline{\theta}_p, a_1, \overline{a}_2)$, which is no less than $u_1(\underline{\theta}_p, \overline{a}_1, \overline{a}_2)$, while his continuation value by playing inconsistently is at most

$(1-\delta)u_1(\underline{\theta}_p, \underline{a}_1, \overline{a}_2) + \delta u_1(\underline{\theta}_p, \underline{a}_1, \underline{a}_2)$, which is strictly less when δ is large enough, leading to a contradiction.

Consider generic μ such that there exist $\theta_p^* \in \Theta_p$ and $q \in (0,1)$ such that:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + q\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \sum_{\theta \succ \theta_p^*} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) = 0; \tag{A.70}$$

as well as generic $\delta \in (0,1)$ such that for every $a_1 \in \Omega^g \cup \{\overline{a}_1\}$, there exists no integer $T \in \mathbb{N}$ such that

$$(1-\delta^T)u_1(\theta_p^*, a_1, \underline{a}_2) + \delta^T u_1(\theta_p^*, a_1, \overline{a}_2) = u_1(\theta_p^*, \underline{a}_1, \underline{a}_2). \tag{A.71}$$

Hence, when $\mu(\Omega^m)$ is small enough such that:

$$\sum_{\theta \succ \theta_p^*} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) + \sum_{\alpha_1 \in \Omega^b} \mu(\Omega^m)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) > 0 \tag{A.72}$$

and

$$(1-q)\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \mu(\Omega^m) \max_{\alpha_1 \in \Omega^m} \mathcal{D}(\phi_{\alpha_1}, \alpha_1) < 0, \tag{A.73}$$

one can uniquely pin down the probability with which type θ_p^* plays $\overline{a}_1$ all the time. To see this, there exists a unique integer T such that:

$$(1-\delta^T)u_1(\theta_p^*, \overline{a}_1, \underline{a}_2) + \delta^T u_1(\theta_p^*, \overline{a}_1, \overline{a}_2) > u_1(\theta_p^*, \underline{a}_1, \underline{a}_2) > (1-\delta^{T+1})u_1(\theta_p^*, \overline{a}_1, \underline{a}_2) + \delta^{T+1} u_1(\theta_p^*, \overline{a}_1, \overline{a}_2).$$

The probability with which type θ_p^* plays $\overline{a}_1$ all the time, denoted by $q^*(\overline{a}_1) \in (0,1)$, is chosen such that:

$$q^*(\overline{a}_1)\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \sum_{\theta \succ \theta_p^*} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1)$$
$$+ \sum_{\alpha_1 \in \Omega^m} \mu(\alpha_1) \underbrace{\alpha_1(\overline{a}_1)^T}_{\text{prob that type } \alpha_1 \text{ plays } \overline{a}_1 \text{ for } T \text{ consecutive periods}} \mathcal{D}(\phi_{\alpha_1}, \alpha_1) = 0. \tag{A.74}$$

Similarly, the probability with which type θ_p^* plays $a_1 \in \Omega^g$ all the time, denoted by $q^*(a_1)$, is pinned down via:

$$q^*(a_1)\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \sum_{\alpha_1 \in \Omega^m} \mu(\alpha_1)\alpha_1(\overline{a}_1)^{T(a_1)}\mathcal{D}(\phi_{\alpha_1}, \alpha_1) = 0.$$

where $T(a_1)$ is the unique integer satisfying:

$$\begin{aligned}(1-\delta^{T(a_1)})u_1(\theta_p^*, a_1, \underline{a}_2) + \delta^{T(a_1)}u_1(\theta_p^*, a_1, \overline{a}_2) \quad &> \quad u_1(\theta_p^*, \underline{a}_1, \underline{a}_2) \\ &> \quad (1-\delta^{T(a_1)+1})u_1(\theta_p^*, a_1, \underline{a}_2) + \delta^{T(a_1)+1}u_1(\theta_p^*, a_1, \overline{a}_2).\end{aligned}$$

The argument above also pins down every type's equilibrium payoff: type $\theta \precsim \theta_p^*$ receives payoff $u_1(\theta, \underline{a}_1, \underline{a}_2)$. Every strategic type above θ_p^*'s equilibrium payoff is pinned down by the occupation measure with which $\overline{a}_2$ is played conditional on player 1 always plays $\overline{a}_1$, which itself is pinned down by type θ_p^*'s indifference condition.

γ-closeness of on-path behavior: Last, I claim that even when $\Omega^{gm} \neq \{\varnothing\}$, (1) All strategic types besides type θ_p^* will either play $\overline{a}_1$ in every period or $\underline{a}_1$ in every period, (2) strategic type θ_p^* will either play $\overline{a}_1$ in every period or $\underline{a}_1$ in every period with probability at least $1-\gamma(\overline{\epsilon})$; (3) his on-path behavior across different equilibria are $\gamma(\overline{\epsilon})$-close, with $\lim_{\overline{\epsilon}\downarrow 0} \gamma(\overline{\epsilon}) = 0$.

Consider the expressions of $Y(h^t)$ in (A.65) and $Z(h^t)$ in (A.67) which provide upper and lower bounds, respectively, on player 2's propensity to play $\overline{a}_2$ at h^t. When $\mu(\Omega^m) < \overline{\epsilon}$, previous arguments imply the existence of $\overline{\gamma}(\overline{\epsilon})$ with $\lim_{\overline{\epsilon}\downarrow 0} \overline{\gamma}(\overline{\epsilon}) = 0$, such that for every equilibrium,

$$Y(h^t), Z(h^t) \in [-\overline{\gamma}(\overline{\epsilon}), \overline{\gamma}(\overline{\epsilon})]$$

for every $h^t \in \mathcal{H}^\sigma$ such that $\overline{a}_1$ has always been played. When $\overline{\epsilon}$ is sufficiently small, this implies the existence of $\theta_p^* \in \Theta_p$ such that type θ_p^* mixes between playing $\overline{a}_1$ in every period and playing $\underline{a}_1$ in every period. This together with Lemma A.3.1 pins down every other strategic type's equilibrium behavior aside from type θ_p^*. Moreover, it also implies that the ex ante probability with which type θ_p^* plays $\overline{a}_1$ in every period or plays $\underline{a}_1$ in every period cannot differ by $2\overline{\gamma}(\overline{\epsilon})/\mu(\theta_p^*)$ across different equilibria. Furthermore, when $\mu(\Omega^m)$ is small enough, player 2 will have a strict incentive to play $\underline{a}_2$ in period 0 as well as in period t if $\underline{a}_1$ has always been played in the past. This and type θ_p^*'s indifference condition pins down every type's equilibrium payoff.

To show that the probability of type θ_p^* behaving inconsistently vanishes with $\mu(\Omega^m)$, notice that first, there exists $s^* \in \mathbb{R}_+$ such that for every $s > s^*$, $\theta_p \in \Theta_p$ and $\alpha_1 \in \Omega^m$,

$$s\mathcal{D}(\theta_p, \overline{a}_1) + \mathcal{D}(\phi_{\alpha_1}, \alpha_1) < 0. \tag{A.75}$$

Therefore, the probability with which every type $\theta_p \in \Theta_p$ playing time inconsistently must be below

$$s^*\overline{\epsilon}\Big\{\min_{\theta_p \in \Theta_p} (1-\overline{\epsilon})\frac{\lambda(\theta_p)}{\sum_{\theta\in\Theta}\lambda(\theta)}\Big\}^{-1}. \tag{A.76}$$

Expression (A.76) provides an upper bound for $\gamma(\overline{\epsilon})$, which vanishes as $\overline{\epsilon} \downarrow 0$. When $\mu(\Omega^m)$ is sufficiently small, Lemma A.3.1 implies the existence of a cutoff type θ_p^* such that all types strictly above θ_p^* always plays $\overline{a}_1$ and all types strictly below θ_p^* always plays $\underline{a}_1$, and type θ_p^* plays consistently with probability at least $1-\gamma(\overline{\epsilon})$, concluding the proof.

A.5 Proof of Proposition 1.1

When $\hat{\mu} \in \Delta(\Theta)$ satisfies (1.20), the equilibrium in which type $\overline{\theta}$ obtains payoff close to $v_{\overline{\theta}}(\overline{a}_1)$ is easy to construct, i.e. every type in $\Theta_g \cup \Theta_p$ always plays $\overline{a}_1$ and every type in Θ_n always plays $\underline{a}_1$. Starting from period 1, player 2 plays $\overline{a}_1$ if and only if player 1 has always been playing $\overline{a}_1$, and plays $\underline{a}_1$ otherwise. In what follows, I show that he cannot obtain payoff greater than $v_{\overline{\theta}}(\overline{a}_1)$ in any equilibrium.

Let $\sigma = (\sigma_1, \sigma_2)$ be an equilibrium under $(\hat{\mu}, \delta)$. Recall the definitions of $\mathcal{H}$ and $\mathcal{H}^\sigma$. Since I will only be referring to on-path histories in this proof, I will replace $\mathcal{H}^\sigma$ with $\mathcal{H}$ from then on. I start with recursively defining the set of '*high histories*'. Let $\overline{\mathcal{H}}^0 \equiv \{h^0\}$ and

$$\overline{a}_1(h^0) \equiv \max\Big\{\bigcup_{\theta\in\Theta} \text{supp}\Big(\sigma_\theta(h^0)\Big)\Big\}.$$

Let

$$\overline{\mathcal{H}}^1 \equiv \{h^1 | \text{ there exists } h^0 \in \overline{\mathcal{H}}^0 \text{ such that } h^1 \succ h^0 \text{ and } \overline{a}_1(h^0) \in h^1\}.$$

For every $t \in \mathbb{N}$ and $h^t \in \overline{\mathcal{H}}^t$, let $\Theta(h^t) \subset \Theta$ be the set of types that occur with positive probability at h^t. Let

$$\overline{a}_1(h^t) \equiv \max\Big\{\bigcup_{\theta\in\Theta(h^t)} \text{supp}\Big(\sigma_\theta(h^t)\Big)\Big\} \tag{A.77}$$

and

$$\overline{\mathcal{H}}^{t+1} \equiv \{h^{t+1} | \text{ there exists } h^t \in \overline{\mathcal{H}}^t \text{ such that } h^{t+1} \succ h^t \text{ and } \overline{a}_1(h^t) \in h^{t+1}\}. \tag{A.78}$$

Let $\overline{\mathcal{H}} \equiv \bigcup_{t=0}^\infty \overline{\mathcal{H}}^t$. For every $\theta \in \Theta$, let $\overline{\mathcal{H}}(\theta)$ be a subset of $\overline{\mathcal{H}}$ such that $h^t \in \overline{\mathcal{H}}(\theta)$ if and only if:

1. For every $h^s \succsim h^t$ with $h^s \in \overline{\mathcal{H}}$, we have $\theta \in \Theta(h^s)$.
2. If $h^{t-1} \prec h^t$, then for every $\tilde{\theta} \in \Theta(h^{t-1})$, there exists $h^s \in \overline{\mathcal{H}}$ with $h^s \succ h^{t-1}$ such that $\tilde{\theta} \notin \Theta(h^s)$.

Let $\overline{\mathcal{H}}(\Theta) \equiv \bigcup_{\theta\in\Theta} \overline{\mathcal{H}}(\theta)$, which has the following properties:

1. $\overline{\mathcal{H}}(\Theta) \subset \overline{\mathcal{H}}$.
2. For every $h^t, h^s \in \overline{\mathcal{H}}(\Theta)$, neither $h^t \succ h^s$ nor $h^t \prec h^s$.

In what follows, I show the following Lemma:

Lemma A.5.1. *For every $h^t \in \overline{\mathcal{H}}$, if $\theta = \max\Theta(h^t)$, then type θ's continuation payoff at h^t is no more than* $\max\{u_1(\theta, \overline{a}_1, \overline{a}_2), u_1(\theta, \underline{a}_1, \underline{a}_2)\}$.

Lemma A.5.1 implies the conclusion in Proposition 1.1 as $h^0 \in \overline{\mathcal{H}}$ and $\overline{\theta} = \max\Theta(h^0)$. A useful conclusion to show Lemma A.5.1 is the following observation:

Lemma A.5.2. *For every $h^t \in \overline{\mathcal{H}}$, if $\theta, \widetilde{\theta} \in \Theta(h^t)$ with $\widetilde{\theta} \prec \theta$, then the difference between type θ's continuation payoff and type $\widetilde{\theta}$'s continuation payoff at h^t is no more than $u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$.*

Proof of Lemma A.5.2: Since u_1 has SID in θ and (a_1, a_2), so for every $\theta \succ \widetilde{\theta}$,

$$(\overline{a}_1, \overline{a}_2) \in \arg\max_{(a_1,a_2)} \Big\{u_1(\theta, a_1, a_2) - u_1(\widetilde{\theta}, a_1, a_2)\Big\} \tag{A.79}$$

which yields the upper bound on the difference between type θ and type $\widetilde{\theta}$'s continuation payoffs. $\square$

For every $h^t \in \overline{\mathcal{H}}(\widetilde{\theta})$, at the subgame starting from h^t, type $\widetilde{\theta}$'s stage game payoff is no more than $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$ in every period and his continuation payoff at h^t cannot exceed $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$. This is because $\overline{a}_1$ is type $\overline{\theta}$'s Stackelberg action, so whenever player 1 plays an action $a_1 \prec \overline{a}_1$, $\underline{a}_2$ is player 2's strict best reply. Lemma A.5.2 then implies that for every $\theta \in \Theta(h^t)$ with $\theta \succ \widetilde{\theta}$, type θ's continuation payoff at h^t cannot exceed $u_1(\theta, \overline{a}_1, \overline{a}_2)$.

In what follows, I prove Lemma A.5.1 by induction on $|\Theta(h^t)|$. When $|\Theta(h^t)| = 1$, i.e. there is only one type (call it type θ) that can reach h^t according to σ, then Lemma A.5.2 implies that type θ's payoff cannot exceed $u_1(\theta, \overline{a}_1, \overline{a}_2)$.

Suppose the conclusion in Lemma A.5.1 holds for every $|\Theta(h^t)| \leq n$, consider the case when $|\Theta(h^t)| = n + 1$. Let $\theta \equiv \max \Theta(h^t)$. Let me introduce set $\overline{\mathcal{H}}^B(h^t)$, which is a subset of $\overline{\mathcal{H}}$. For every $h^s \succsim h^t$ with $h^s \in \overline{\mathcal{H}}$, $h^s \in \overline{\mathcal{H}}^B(h^t)$ if and only if:

- $h^s \in \overline{\mathcal{H}}(\theta)$,
- but $h^{s+1} \notin \overline{\mathcal{H}}(\theta)$ for any $h^{s+1} \succ h^s$ with $h^{s+1} \in \overline{\mathcal{H}}$.

In another word, type θ has a strict incentive not to play $\overline{a}_1(h^s)$ at h^s. A useful property is:

- For every $h^\infty \in \overline{\mathcal{H}}$ with $h^\infty \succ h^t$, either there exists $h^s \in \overline{\mathcal{H}}^B(h^t)$ such that $h^s \prec h^\infty$, or there exists $h^s \in \overline{\mathcal{H}}(\theta)$ such that $h^s \prec h^\infty$.

which means that play will eventually reach either a history in $\overline{\mathcal{H}}^B(h^t)$ or $\overline{\mathcal{H}}(\theta)$ if type θ keeps playing $\overline{a}_1(h^\tau)$ before that for every $t \leq \tau \leq s$.

In what follows, I examine type θ's continuation value at each kind of history.

1. For every $h^s \in \overline{\mathcal{H}}^B(h^t)$, at h^{s+1} with $h^{s+1} \succ h^s$ and $h^{s+1} \in \overline{\mathcal{H}}$, by definition,

$$|\Theta(h^{s+1})| \leq n.$$

   Let $\widetilde{\theta} \equiv \max \Theta(h^{s+1})$. By induction hypothesis, type $\widetilde{\theta}$'s continuation payoff at h^{s+1} is at most $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$. This applies to every such h^{s+1}.

   Type $\widetilde{\theta}$'s continuation value at h^s also cannot exceed $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$ since he is playing $\overline{a}_1(h^s)$ with positive probability at h^s, and his stage game payoff from doing so is at most $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$. Furthermore, his continuation value afterwards cannot exceed $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$.

   Lemma A.5.2 then implies that type θ's continuation value at h^s is at most $u_1(\theta, \overline{a}_1, \overline{a}_2)$.

2. For every $h^s \in \overline{\mathcal{H}}(\theta)$, always playing $\overline{a}_1(h^\tau)$ for all $h^\tau \succsim h^s$ and $h^\tau \in \overline{\mathcal{H}}$ is a best reply for type θ. His stage game payoff from this strategy cannot exceed $u_1(\theta, \overline{a}_1, \overline{a}_2)$, which implies that his continuation value at h^s also cannot exceed $u_1(\theta, \overline{a}_1, \overline{a}_2)$.

Starting from h^t consider the strategy in which player 1 plays $\overline{a}_1(h^\tau)$ at every $h^\tau \succ h^t$ and $h^\tau \in \overline{\mathcal{H}}$ until play reaches $h^s \in \overline{\mathcal{H}}^B(h^t)$ or $h^s \in \overline{\mathcal{H}}(\theta)$. By construction, this is type θ's best reply. Under this strategy, type θ's stage game payoff cannot exceed $u_1(\theta, \overline{a}_1, \overline{a}_2)$ before reaches h^s. Moreover, his continuation payoff after reaching h^s is also bounded above by $u_1(\theta, \overline{a}_1, \overline{a}_2)$, which proves Lemma A.5.1 when $|\Theta(h^t)| = n+1$.

A.6 Proof of Proposition 1.2

Throughout the proof, I normalize $u_1(\theta, \underline{a}_1, \underline{a}_2) = 0$ for every θ. Let $x_\theta(a_1) \equiv -u_1(\theta, a_1, \underline{a}_2)$ and $y_\theta(a_1) \equiv u_1(\theta, a_1, \overline{a}_2)$. Assumptions 1.1 and 1.2 imply that:

- $x_\theta(a_1) \geq 0$, with "=" holds only when $a_1 = \underline{a}_1$.
- $y_\theta(a_1) > 0$ for every $\theta \in \Theta$ and $a_1 \in A_1$.
- $x_\theta(a_1)$ and $-y_\theta(a_1)$ are both strictly increasing in a_1.
- For every $\theta < \widetilde{\theta}$, $x_\theta(a_1) - x_{\widetilde{\theta}}(a_1)$ and $y_{\widetilde{\theta}}(a_1) - y_\theta(a_1)$ are both strictly increasing in a_1.

I start with defining '*pessimistic belief path*' for every $\sigma \in \text{NE}(\delta, \hat{\mu})$. For every $a_1^\infty \equiv (a_{1,0}, a_{1,1}..., a_{1,t}...)$, we say $a_1^\infty \in \mathcal{A}^\sigma(\theta^*)$ if and only if for every $t \in \mathbb{N}$, there exists $r^t \in \mathcal{R}^t$ such that $(a_{1,0}, ..., a_{1,t-1}, r^t) \in \mathcal{H}^\sigma$ and $\sum_{\theta \succsim \theta^*} \mu_t(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0$, where μ_t is player 2's belief after observing $(a_{1,0}, ..., a_{1,t-1})$. For any $\theta \in \Theta$, if $(a_1^\infty, r^\infty) \in \mathcal{H}^\sigma(\theta)$ for some $a_1^\infty \in \mathcal{A}^\sigma(\theta)$, then $V_\theta^\sigma(\delta) = 0$ and always playing $\underline{a}_1$ is type θ's best reply.

For every $\hat{\mu}$ satisfying (1.21), there exist a unique $\theta_p^* \in \Theta_p$ and a unique $q(\hat{\mu}) \in [0,1)$ such that:

$$q(\hat{\mu})\hat{\mu}(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \sum_{\theta \succ \theta_p^*} \hat{\mu}(\theta)\mathcal{D}(\theta, \overline{a}_1) = 0. \tag{A.80}$$

When $\hat{\mu}$ satisfies (1.21), then for every $\sigma \in \text{NE}(\delta, \hat{\mu})$ and $\theta \precsim \theta_p^*$, since $\sum_{\widetilde{\theta} \succsim \theta} \hat{\mu}(\widetilde{\theta})\mathcal{D}(\widetilde{\theta}, \overline{a}_1) < 0$, using the martingale property of beliefs, we know that there exists $a_1^\infty \in \mathcal{A}^\sigma(\theta)$ such that $(a_1^\infty, r^\infty) \in \mathcal{H}^\sigma(\theta)$ for some r^∞. This pins down the unique equilibrium payoff for all $\theta \precsim \theta_p^*$.

In what follows, I establish the upper bound in Proposition 1.2. For every $\theta \succ \theta_p^*$, in every action path $a_1^\infty = (a_{1,0}, a_{1,1}, ...)$ which type θ follows with strictly positive probability under $\sigma \in \text{NE}(\delta, \hat{\mu})$, it must be that:

$$V_\theta^\sigma(\delta) = \sum_{a_1, a_2} \mathcal{P}^{a_1^\infty}(a_1, a_2) u_1(\theta, a_1, a_2)$$

and

$$0 = V_{\theta_p^*}^\sigma(\delta) \geq \sum_{a_1, a_2} \mathcal{P}^{a_1^\infty}(a_1, a_2) u_1(\theta_p^*, a_1, a_2)$$

where

$$\mathcal{P}^{a_1^\infty}(a_1, a_2) \equiv \sum_{t=0}^{\infty} (1-\delta)\delta^t p^{(\sigma_1, \sigma_2)}(a_1, a_2)$$

with σ_1 playing a_1^∞ on the equilibrium path. Therefore, $V_\theta^\sigma(\delta)$ must be weakly below the value of the following linear program:

$$\max_{\{\beta(a_1),\gamma(a_1)\}_{a_1\in A_1}} \Big\{\sum_{a_1\in A_1} \beta(a_1)y_\theta(a_1)-\gamma(a_1)x_\theta(a_1)\Big\}, \tag{A.81}$$

subject to

$$\sum_{a_1\in A_1} \gamma(a_1)+\beta(a_1)=1,$$

$$\gamma(a_1),\beta(a_1)\geq 0 \text{ for every } a_1\in A_1,$$

and

$$\sum_{a_1\in A_1} \beta(a_1)y_{\theta_p^*}(a_1)-\gamma(a_1)x_{\theta_p^*}(a_1)\leq 0. \tag{A.82}$$

Due to the linearity of this program, it is without loss of generality to focus on solutions where there exist a_1^* and a_1^{**} such that

$$\beta(a_1)>0 \text{ iff } a_1=a_1^*, \quad \gamma(a_1)>0 \text{ iff } a_1=a_1^{**}.$$

According to (A.82), we have:

$$\beta(a_1^*)y_{\theta_p^*}(a_1^*)\leq (1-\beta(a_1^*))x_{\theta_p^*}(a_1^{**}). \tag{A.83}$$

Plugging (A.83) into (A.81), the value of that expression cannot exceed:

$$\max_{a_1^*,a_1^{**}\in A_1} \Big\{\frac{y_\theta(a_1^*)x_{\theta_p^*}(a_1^{**})-x_\theta(a_1^{**})y_{\theta_p^*}(a_1^*)}{x_{\theta_p^*}(a_1^{**})+y_{\theta_p^*}(a_1^*)}\Big\}. \tag{A.84}$$

Expression (A.84) is maximized when $a_1^*=a_1^{**}=\overline{a}_1$, which gives an upper bound for $V_\theta^\sigma(\delta)$:

$$V_\theta^\sigma(\delta)\leq ru_1(\theta,\overline{a}_1,\overline{a}_2)+(1-r)u_1(\theta,\overline{a}_1,\underline{a}_2), \tag{A.85}$$

with $r\in(0,1)$ satisfying: $ru_1(\theta_p^*,\overline{a}_1,\overline{a}_2)+(1-r)u_1(\theta_p^*,\overline{a}_1,\underline{a}_2)=u_1(\theta_p^*,\underline{a}_1,\underline{a}_2)$. The upper bound in (A.85) is asymptotically achieved when $\delta\to 1$ in an equilibrium where:

- Type θ always plays $\underline{a}_1$ if $\theta\prec\theta_p^*$, always plays $\overline{a}_1$ if $\theta\succ\theta_p^*$.
- Type θ_p^* randomizes between always playing $\overline{a}_1$ and always playing $\underline{a}_1$ with prob $q(\hat{\mu})$ and $1-q(\hat{\mu})$.

A.7 Counterexamples

I present several counterexamples missing from the main text. For convenience, there is only one commitment type in every example besides the one in Appendix A.7.5. The intuition behind the examples still apply when there are multiple pure strategy commitment types. Abusing notation, I use θ to denote the Dirac measure on θ and a_i to denote the Dirac measure on a_i with $i \in \{1, 2\}$.

A.7.1 Failure of Reputation Effects When Supermodularity is Violated

Example 1: To begin with, I construct equilibrium in the entry deterrence game of Harrington (1986) in which the supermodularity condition on u_1 is violated.[11] Let the stage game payoff be:

θ_1	O	I
F	$1, 0$	$-1, -1$
A	$2, 0$	$0, 1$

θ_0	O	I
F	$5/2, 0$	$1/2, 1/2$
A	$3, 0$	$1, 3/2$

One can verify that the monotonicity condition is satisfied. To see this game's payoff fails the supermodularity assumption, let us rank the state and players' actions via $\theta_1 \succ \theta_0$, $F \succ A$ and $O \succ I$. Player 1's cost of fighting is 1 in state θ_1 and is $1/2$ in state θ_0. Intuitively, when the incumbent's and the entrant's costs are positively correlated, the incumbent's loss from fighting (by lowering prices) is higher when his cost is high, and the entrant's profit from entry decreases with the cost and increases with the incumbent's price.

When $\Omega = \{F, \theta_1, \theta_0\}$ and $\mu(F) \leq 2\mu(\theta_0)$, I construct an equilibrium with low payoffs in each of the following three cases, depending on the signs of:

$$X \equiv \frac{\mu(\theta_0)}{2} + \Big(\frac{\mu(F)\phi_F(\theta_0)}{2} - \mu(F)\phi_F(\theta_1)\Big) - \mu(\theta_1) \tag{A.86}$$

and

$$Y \equiv \frac{\mu(F)\phi_F(\theta_0)}{2} - \mu(F)\phi_F(\theta_1). \tag{A.87}$$

1. If $X \leq 0$, then type θ_0 always plays F, type θ_1 mixes between always playing F and always playing A, with the probability of playing F being $1 + X/\mu(\theta_1)$. Player 2 plays I for sure in period 0. Starting from period 1, she plays I for sure if A has been observed before and plays $\frac{1}{2\delta}O + (1 - \frac{1}{2\delta})I$ otherwise. Despite the probability of type θ_1 is large relative to that of type θ_0, type θ_1's equilibrium payoff is 0 and type θ_0's equilibrium payoff is $3/2$, both are lower than their commitment payoffs from playing F.

2. If $X > 0$ and $Y \leq 0$, then type θ_1 always plays A, type θ_0 mixes between always playing F and always playing A, with the probability of playing F being $-Y/\mu(\theta_0)$. Player 2 plays I for sure in period 0. Starting from period 1, she plays I for sure if A has been observed before and plays $\frac{1}{4\delta}O + (1 - \frac{1}{4\delta})I$ otherwise. Type θ_1's equilibrium payoff is 0 and type θ_0's equilibrium payoff is 1.

3. If $X > 0$ and $Y > 0$, then both types always play A. Player 2 plays I no matter what. Type θ_1's equilibrium payoff is 0 and type θ_0's equilibrium payoff is 1.

[11] The case in which u_2 has decreasing differences between a_2 and θ is similar once we reverse the order on the states.

Example 2: Next, I construct equilibrium in the entry deterrence game when the supermodularity condition on u_2 is violated. I focus on the case in which u_2 has *decreasing* differences between a_2 and a_1.[12] Consider the following $2 \times 2 \times 2$ game with payoffs given by:

$\theta = \theta_1$	h	l
H	$1, -1$	$-1, 0$
L	$2, 1$	$0, 0$

$\theta = \theta_0$	h	l
H	$1 - \eta, -2$	$-1 - \eta, 0$
L	$2, -1$	$0, 0$

with $\eta \in (0, 1)$. The states and players' actions are ranked according to $H \succ L$, $h \succ l$ and $\theta_1 \succ \theta_0$. Let $\Omega = \{H, \theta_1, \theta_0\}$. Theorem 1.2 trivially applies as the commitment outcome (H, l) gives every type his lowest feasible payoff. In what follows, I show the failure of Theorem 1.3, i.e. player 1 has multiple equilibrium behaviors. First, there exists an equilibrium in which (L, h) is always played or (L, l) is always played, depending on the prior belief. Second, consider the following equilibrium:

- In period 0, both strategic types play L.
- From period 1 to $T(\delta) \in \mathbb{N}$, type θ_0 plays L and type θ_1 plays H. Player 2 plays h in period $t(\geq 2)$ if and only if $t \geq T(\delta) + 1$ and player 1's past play coincides with type θ_1's equilibrium strategy. The integer $T(\delta)$ is chosen such that:

$$(1 - \delta^{T(\delta)})(-1) + 2\delta^{T(\delta)} > 0 > (1 - \delta^{T(\delta)})(-1 - \eta) + 2\delta^{T(\delta)},$$

  which exists when δ is close enough to 1.

A.7.2 Failure of Reputation Effects When Monotonicity is Violated

I show that the monotonicity condition is indispensable for my reputation result. For this purpose, I consider two counterexamples in which Assumption 1 is violated in different ways.

Example 1: Consider the following $2 \times 2 \times 2$ game:

$\theta = \theta_1$	h	l
H	$3/2, 2$	$0, 0$
L	$1, 1$	$0, 0$

$\theta = \theta_0$	h	l
H	$-1, -1/2$	$1, 0$
L	$0, -1$	$5/2, 1/4$

One can verify that this game satisfies the supermodularity assumption once we rank the states and actions according to $\theta_1 \succ \theta_0$, $H \succ L$ and $h \succ l$.[13] However, the monotonicity assumption is violated as player 1's ordinal preferences over a_1 and a_2 depend on the state.

Suppose $\Omega = \{H, \theta_1, \theta_0\}$ with $4\mu(H) < \mu(\theta_0)$. Consider the following equilibrium in which player 2 plays a '*tit-for-tat*' like strategy. Type θ_1 plays L all the time and type θ_0 plays H all the

[12] The case in which u_2 has decreasing differences between a_2 and θ is similar to the previous example. One only needs to reverse the order between the states.

[13] In fact, the game's payoffs even satisfy a stronger notion of complementarity, that is, both u_1 and u_2 are strictly supermodular functions of the triple (θ, a_1, a_2). The definition of supermodular function can be found in Topkis (1998).

time. Starting from period 1, player 2 plays h in period $t \geq 1$ if L was played in period $t-1$ and vice versa. Both types' equilibrium payoffs are close to 1, which are strictly lower than their pure Stackelberg commitment payoffs, which are $3/2$ and $5/2$ respectively.

To verify that this is an equilibrium when δ is high enough, notice that first, player 2's incentive constraints are always satisfied. As for player 1, if $\theta = \theta_1$, deviating for one period gives him a stage game payoff at most $3/2$ and in the next period his payoff is at most 0. If $\delta > 1/2$, then he has no incentive to deviate. if $\theta = \theta_0$, deviating for one period gives him a stage game payoff at most $5/2$ and in the future, he will keep receiving payoff at most 0 until he plays H for one period. He has no incentive to deviate if and only if for every $t \in \mathbb{N}$,

$$(1-\delta)\frac{5}{2} - (\delta^t - \delta^{t+1}) \leq 1 - \delta^{t+1}. \tag{A.88}$$

which is equivalent to:

$$\frac{5}{2} \leq 1 + \delta + ... + \delta^{t-1} + 2\delta^t.$$

The above inequality is satisfied for every integer $t \geq 1$ when $\delta > 0.9$. This is because when $t \geq 2$, the right hand side is at least $1 + 0.9 + 0.9^2$, which is greater than $5/2$. When $t = 1$, the right hand side equals to 2.8, which is greater than $5/2$.

To see that player 1's equilibrium behavior is not unique, consider another equilibrium in which type θ_1 always plays H, type θ_0 always plays L and for every $t \in \mathbb{N}$, player 2 plays h in period t if H is played in period $t-1$, and plays l in period t if L is played in period $t-1$. This implies that the conclusion in Theorem 1.3 will fail in absence of the monotonicity assumption.

Example 2: Low payoff equilibria can be constructed when player 1's ordinal preference over each player's actions does not depend on the state, but the directions of monotonicity violate Assumption 1. For example, consider the following game:

$\theta = \theta_1$	h	l
H	$2, 2$	$0, 0$
L	$1, 1$	$-1/2, 0$

$\theta = \theta_0$	h	l
H	$1/4, -1/2$	$1/8, 0$
L	$0, -1$	$-1/16, 1/4$

Both players' payoffs are supermodular functions of (θ, a_1, a_2). Player 1's ordinal preferences over a_1 and a_2 are state independent but his payoff is strictly increasing in both a_1 and a_2, which is different from what Assumption 1 suggests. Rank the states and actions according to $\theta_1 \succ \theta_0$, $H \succ L$ and $h \succ l$.

Suppose $\Omega = \{H, \theta_1, \theta_0\}$ with $4\mu(H) < \mu(\theta_0)$. The following strategy profile is an equilibrium. Type θ_1 plays L all the time and type θ_0 plays H all the time. Starting from period 1, player 2 plays h in period $t \geq 1$ if L was played in period $t-1$ and vice versa. Type θ_1 and type θ_0's equilibrium payoffs are close to 1 and $1/8$, respectively as $\delta \to 1$. Their pure Stackelberg commitment payoffs are 2 and $1/4$, respectively, which are strictly higher. The verification of players' incentive constraints is the same as the previous example.

Moreover, contrary to what Theorem 1.3 has suggested, player 1's equilibrium behavior is not unique even when player 2's prior belief is pessimistic, i.e.

$$2\mu(\theta_1) + \mu(H)\Big(2\phi_H(\theta_1) - \frac{1}{2}\phi_H(\theta_0)\Big) - \frac{1}{2}\mu(\theta_0) < 0. \tag{A.89}$$

This is because aside from the equilibrium constructed above, there also exists an equilibrium in which type θ_1 always plays H, type θ_0 mixes between always playing H and always playing L with probabilities such that player 2 becomes indifferent between h and l starting from period 1 conditional on H has always been played. In equilibrium, player 2 plays h in period $t \geq 1$ as long as player 1 has always played H before, and switches to l permanently otherwise.

A.7.3 Failure of Reputation Effects When $|A_2| \geq 3$

I present an example in which the reputation results in Theorems 2 and 3 fail when the stage game has MSM payoffs but player 2 has three or more actions. This motivates the additional conditions on the payoff structure in Online Appendix D. Consider the following $2 \times 2 \times 3$ game with payoffs:

$\theta = \theta_1$	l	m	r
H	$0, 0$	$5/2, 2$	$6, 3$
L	$\epsilon, 0$	$5/2 + \epsilon, -1$	$6 + \epsilon, -2$

$\theta = \theta_0$	l	m	r
H	$0, 0$	$2, -1$	$3, -2$
L	$2\epsilon, 0$	$2 + 2\epsilon, -2$	$3 + 2\epsilon, -3$

where $\epsilon > 0$ is small enough. Let the rankings on actions and states be $H \succ L$, $r \succ m \succ l$ and $\theta_1 \succ \theta_0$. One can check that the stage game payoffs are MSM.

Suppose $\Omega = \{\theta_1, \theta_0, H\}$ with $\mu(\theta_0) = 2\eta$, $\mu(H) = \eta$ and $\phi_H = \theta_1$, with $\eta \in (0, 1/3)$. Type θ_1's commitment payoff from playing H is 6. However, consider the following equilibrium:

- Type θ_0 plays H all the time. Type θ_1 plays L from period 0 to $T(\delta)$ and plays H afterwards, with $1 - \delta^{T(\delta)} \in (1/2 - \epsilon, 1/2 + \epsilon)$. Such $T(\delta) \in \mathbb{N}$ exists when $\delta > 1 - 2\epsilon$.

- Player 2 plays m starting from period 1 if player 1 has always played H in the past. She plays r from period 1 to $T(\delta)$ and plays r afterwards if player 1's past actions are consistent with type θ_1's equilibrium strategy. She plays l at every off-path history.

Type θ_1's equilibrium payoff is approximately $3 + \epsilon/2$ as $\delta \to 1$, which is strictly less than his commitment payoff. To see that player 1 has multiple equilibrium behaviors under a pessimistic prior belief, i.e. $\eta \in [1/4, 1/3)$, there exists another equilibrium in which all types of player 1 plays H at every on-path history. Player 2 plays m if all past actions were H and plays l otherwise.

A.7.4 Time Inconsistent Equilibrium Plays in Private Value Reputation Game

I construct an equilibrium in the private value product choice game in which despite there is a commitment type that always exerts high effort, the strategic long-run player abandons his reputation early on in the relationship and L is played with significant probability. The game's payoff matrix is given by:

$-$	C	S
H	$1,3$	$-1,2$
L	$2,0$	$0,1$

Suppose there is a commitment type that always plays H (which is unlikely compared to the strategic type) and consider the following equilibrium when $\delta > 1/2$:

- The strategic type plays L for sure in period 0. He plays $\frac{1}{2}H + \frac{1}{2}L$ starting from period 1.
- Player 2 plays S for sure in period 0. If H is observed in period 0, then she plays C for sure as long as H has always been played. She plays S for sure in all subsequent periods if L has been played before.

  If L is observed in period 0, C is played for sure in period 1. Starting from period 2, player 2 plays C for sure in period t if H was played in period $t-1$, and $(1-\frac{1}{2\delta})C + \frac{1}{2\delta}S$ in period t if L was played in period $t-1$.

It is straightforward to verify players' incentive constraints. Intuitively, starting from period 1, every time player 1 shirks, he will be punished tomorrow as player 2 will play C with less probability. The probabilities with which he mixes between H and L are calibrated to provide player 2 the incentive to mix between C and S. Despite the strategic type obtains equilibrium payoff δ, which is close to his pure Stackelberg commitment payoff given that he is sufficiently patient. However, the strategic long-run player's equilibrium play is very different from the commitment type's. Perhaps more surprisingly, (i) imitating the commitment type is a strictly dominated strategy, which yields payoff $-(1-\delta)+\delta$, strictly less than his equilibrium payoff; (ii) evaluating the occupation measure of every action ex ante, L is played with significant probability. On average, L is played with occupation measure strictly more than 1/2, which converges to 1/2 as $\delta \to 1$.

A.7.5 Behavioral Uniqueness: Commitment Type Occurs with Low Probability

The following example illustrates why $\mu(\Omega^m)$ being small is not redundant for Theorem 1.3' when there exists $\alpha_1 \in \Omega^m \backslash \{\overline{a}_1\}$ such that $\{\overline{a}_2\} = \text{BR}_2(\alpha_1, \phi_{\alpha_1})$. Consider the following $3 \times 2 \times 2$ stage game:

$\theta = \theta_1$	C	S
H	$1,2$	$-2,0$
M	$2,1$	$-1,0$
L	$3,-1$	$0,0$

$\theta = \theta_0$	C	S
H	$1/2,-1$	$-5/2,0$
M	$3/2,-2$	$-3/2,0$
L	$3,-3$	$0,0$

Let $\Omega \equiv \{H, M, \theta_1, \theta_0\}$ with $\mu(H) = \mu(\theta_1) = 1/20$, $\mu(\theta_0) = 3/10$ and $\mu(M) = 3/5$. Let $\phi_H = \phi_M$ be the Dirac measure on θ_1. One can check that $M \in \Omega^g$ and μ satisfies (1.11). However, for every $\delta > 5/6$, one can construct the following class of equilibria indexed by $T \in \{1, 2, ...\}$:

- **Equilibrium $\boldsymbol{\sigma^T}$:** Type θ_0 plays M forever. Type θ_1 plays M from period 0 to period T, and plays H starting from period $T+1$. Player 2 plays C for sure from period 0 to $T+1$ if player 1's past actions were either all H or all M. For period $t \geq T+2$, player 2 plays C for sure if player 1's past actions were all H or all M from 0 to T and all H afterwards, he plays $\frac{3\delta-1}{3\delta}C + \frac{1}{3\delta}S$ if player 1's past actions were all M. Player 2 plays S for sure at any other history.

One can verify players' incentive constraints. In particular in period $T+1$ conditional on player 1 has always played M in the past, type θ_1 is indifferent between playing H and M while type θ_0 strictly prefers to play M. This class of equilibria can be constructed for an open set of beliefs.[14] As we can see, player 1's equilibrium behaviors are drastically different once we vary the index T, ranging from playing M all the time to playing H almost all the time. Moreover, the good strategic type, namely type θ_1, have an incentive to play actions other than H for a long period of time, contrary to what Theorems 1.3 and A.1 suggest.

A.7.6 Irregular Equilibria in Games with MSM Payoffs

I construct an equilibrium in the repeated product choice game with MSM payoffs in which at some on-path histories, player 1's reputation deteriorates after playing the highest action.[15] Recall that players' stage game payoffs are given by:

$\theta = \theta_1$	h	l
H	$1, 3$	$-1, 2$
L	$2, 0$	$0, 1$

$\theta = \theta_0$	h	l
H	$1-\eta, 0$	$-1-\eta, 1$
L	$2, -2$	$0, 0$

with $\eta \in (0,1)$. Let $\Omega \equiv \{H, \theta_1, \theta_0\}$ with $\mu(H) = 0.06$, $\mu(\theta_0) = 0.04$, $\mu(\theta_1) = 0.9$ and ϕ_H is the Dirac measure on θ_1. Consider the following strategy profile:

- In period 0, type θ_1 plays H with probability $2/45$ and type θ_0 plays H with probability $1/4$. Player 2 plays l.
- In period 1, if the history is (L, l), then use the public randomization device. With probability $(1-\delta)/\delta$, players play (L, l) forever, with complementary probability, players play (H, h) forever. If (H, h) is prescribed and player 1 ever deviates to L, then player 2 plays l at every subsequent history.
- In period 1, if the history is (H, l), then both strategic types play L and player 2 plays h. This is incentive compatible due to the presence of the commitment type.
- In period 2, if the history is (H, l, H, h), then play (H, h) forever on the equilibrium path. If player 2 ever observes player 1 plays L, then she plays l in all subsequent periods.
- In period 2, if the history is (H, l, L, h), then use the public randomization device:

[14] Notice that under a generic prior belief, type θ_1 needs to randomize between always playing H and always playing M in period $T+1$. This can be achieved since he is indifferent by construction.

[15] One can also verify that the constructed strategy profile is also part of a sequential equilibrium under its induced belief system.

◇ With probability $(1-\delta)/\delta$, play (L, l) forever on the equilibrium path.

◇ With probability $1 - \frac{1-\delta}{\delta^2} - \frac{1-\delta}{\delta}$, play (H, h) forever on the equilibrium path. If player 2 ever observes player 1 plays L, then she plays l in all subsequent periods.

◇ With probability $(1-\delta)/\delta^2$, type θ_0 plays L for sure and type θ_1 plays L with probability 1/4, and player 2 plays h.
Following history (H, l, L, h, H, h), play (H, h) forever on the equilibrium path. If player 2 ever observes player 1 plays L, then she plays l in all subsequent periods.
Following history (H, l, L, h, L, h), use the public randomization device again. With probability $(1-\delta)/\delta$, play (L, l) forever. With complementary probability, play (H, h) forever on the equilibrium path. If player 2 ever observes player 1 plays L, then she plays l in all subsequent periods.

In period 0, player 2's belief about θ deteriorates after observing H. This is true no matter whether we only count the strategic types (as strategic type θ_0 plays H with strictly higher probability) or also count the commitment type (probability of θ_1 decreases from 24/25 to 10/11).

A.7.7 Multiple Equilibrium Behaviors when Player 1 is Impatient

I present an example in which the game's payoff satisfies Assumptions 1.1 to 1.3, player 2's prior belief is pessimistic but player 1 has multiple equilibrium behaviors when δ is not high enough. Consider the following product choice game:

$\theta = \theta_1$	C	S
H	$1, 3$	$-1, 2$
L	$2, 0$	$0, 1$

$\theta = \theta_0$	C	S
H	$1-\eta, 0$	$-1-\eta, 1$
L	$2, -2$	$0, 0$

with $\eta \in (0, 1)$, $\Omega^m \equiv \{H\}$ and ϕ_H be the Dirac measure on θ_1. Player 2's prior satisfies:

$$\mu(\theta_0) = 0.7, \mu(\theta_1) + \mu(H) = 0.3 \text{ with } \mu(H) \in (0, 0.1).$$

I construct a class of Nash equilibria when $\delta \in (\frac{1}{2}, \frac{1+\eta}{2})$, in which player 1's on-path equilibrium behaviors are different across these equilibria.[16]

- Type θ_0 always plays L.
- Type θ_1 plays H in every period besides period $t \in \{1, 2, ...\}$, in which he plays L.
- Player 2 plays S in period 0 and period t. In period $s \neq 0, t$, she plays S if player 1 has played L before in any period besides t; she plays C if player 1 has played H in every period or has only played L in period t.

Intuitively, since player 1's discount factor is low, strategic type θ_0 has no incentive to pool with the commitment type. Therefore, after playing H for one period, player 2's belief becomes optimistic which leads to multiple equilibrium behaviors.

[16] One can verify that these Nash equilibrium outcomes can also be implemented in sequential equilibrium.

A.7.8 Why $\lambda \in \Lambda(\alpha_1^*, \theta)$ is not sufficient when α_1^* is mixed?

I use a counterexample to show that $\lambda \in \Lambda(\alpha_1^*, \theta)$ is no longer sufficient to guarantee the commitment payoff bound when α_1^* is mixed. Players' payoffs are given by:

θ_1	l	m	r	θ_2	l	m	r	θ_3	l	m	r
H	$1,3$	$0,0$	$0,0$	H	$0,1/2$	$0,3/2$	$0,0$	H	$0,1/2$	$0,0$	$0,3/2$
L	$2,-1$	$0,0$	$0,0$	L	$0,1/2$	$0,3/2$	$0,0$	L	$0,1/2$	$0,0$	$0,3/2$
D	$3,-1$	$1/2,0$	$1/2,0$	D	$0,0$	$0,0$	$0,0$	D	$0,0$	$0,0$	$0,0$

Suppose $\Omega^m = \{\alpha_1^*\}$ with $\alpha_1^* \equiv \frac{1}{2}H + \frac{1}{2}L$ and $\phi_{\alpha_1^*}$ is the Dirac measure on θ_1, one can apply the definitions and obtain that $v_{\theta_1}(\alpha_1^*) = 3/2$ and $\Theta^b_{(\alpha_1^*,\theta_1)} = \{\theta_2, \theta_3\}$. If $\mu(\alpha_1^*) = 2\mu(\theta_2) = 2\mu(\theta_3) \equiv \rho$ for some $\rho \in (0, 1/2)$, then $\lambda = (1/2, 1/2) \in \Lambda(\alpha_1^*, \theta_1)$. In the following equilibrium, type θ_1's payoff is $1/2$ even when $\delta \to 1$.

- Type θ_1 always plays D. In period 0, type θ_2 plays H and type θ_3 plays L. Starting from period 1, both types play $\frac{1}{2}H + \frac{1}{2}L$. Player 2 plays m in period 0. If she observes H or D in period 0, then she plays m forever. If she observes L in period 0, then she plays r forever.

In the above equilibrium, either $\mu_t(\theta_2)/\mu_t(\alpha_1^*)$ or $\mu_t(\theta_3)/\mu_t(\alpha_1^*)$ will increase in period 0, regardless of player 1's action in that period. As a result, player 2's posterior belief in period 1 is outside $\overline{\Lambda}(\alpha_1^*, \theta_1)$ for sure. This provides him a rationale for not playing l and gives type θ_1 an incentive to play D forever, making player 2's belief self-fulfilling. This situation only arises when α_1^* is mixed and $k(\alpha_1^*, \theta) \geq 2$.

Appendix B

Appendix to Chapter 2

B.1 Proof of Proposition 2.2

In this Appendix, I show the eight lemmas in Subsection 2.3.4 that lead to Proposition 2.2.

B.1.1 Proof of Lemma 2.3.1

Suppose towards a contradiction, that there exists $\pi^* \in (0,1)$ such that $\lim_{t\to\infty} \pi_t = \pi^*$, where π_t is the market's belief conditional on $h^t = \{\varnothing\}$. Rewrite (2.1) as:

$$\frac{d\pi_t}{\pi_t(1-\pi_t)} = -Y_t dt.$$

Integrate both sides from 0 to ∞, we have:

$$\int_0^\infty Y_t dt = \ln \frac{1-\pi^*}{\pi^*} - \ln \frac{1-\pi_0}{\pi_0}, \tag{B.1}$$

Since $\pi^*, \pi_0 \in (0,1)$, so $\int_0^\infty Y_t dt$ is finite. This implies that $\lim_{t\to\infty} Y_t = 0$. Since $Y_t = \mu\chi_t a_t$, $\mu > 0$ and $a_t \geq \phi$, we have $\chi_t \to 0$. Hence, for every $\varepsilon > 0$, there exists $T \in \mathbb{R}_+$ such that $\chi_t < \varepsilon$ for all $t > T$. Pick ε such that:

$$\frac{\mu\varepsilon}{r}\phi < c. \tag{B.2}$$

Since $V_a(\pi_t) \geq 0$, so (2.5) implies that $a_t = \phi$ for all $t > T$, implying that $V_m(\pi_t) < b$. This suggests that $\chi_t = 1$ for all $t > T$, which leads to a contradiction.

B.1.2 Proof of Lemma 2.3.2

Since $\chi(\pi_t) > 0$ for all t, let $\hat{V}_a(t) \equiv V_a(\pi_t)$. Replace $V_a(\pi_t)$ and $V_a'(\pi_t)$ with $\hat{V}_a(t)$ and $\hat{V}_a'(t)$ in (2.3), we have:

$$\hat{V}_a(t) = \pi_t a_t + c\phi + a_t\Big(\frac{\mu\chi_t}{r}(\phi - \hat{V}_a(t)) - c\Big) + \frac{1}{r}\hat{V}_a'(t).$$

The unique bounded solution is:

$$\hat{V}_a(t) = r\int_t^\infty e^{-r(s-t)}\Big[\pi_s a_s + c\phi + a_s\Big(\frac{\mu\chi_s}{r}\big(\phi - \hat{V}_a(s)\big) - c\Big)\Big]ds = c\phi + X_t + W_t, \tag{B.3}$$

where

$$X_t \equiv r\int_t^\infty e^{-r(s-t)}Z_s ds, \quad Z_t \equiv a_t\Big(\frac{\mu\chi_t}{r}\big(\phi - \hat{V}_a(t)\big) - c\Big),$$

and

$$W_t \equiv r\int_t^\infty e^{-r(s-t)}\pi_s a_s ds.$$

This implies that for every $\varepsilon > 0$, there exists $T \in \mathbb{R}^+$ such that for every $t > T$:

$$\pi_t < \varepsilon, \; W_t < \varepsilon,$$

The expression for $\hat{V}_a(t)$ suggests that:

$$X_t < \hat{V}_a(t) - c\phi < X_t + \varepsilon. \tag{B.4}$$

Moreover, according to the agent's incentive constraint, Z_t has the following property:

$$Z_t \begin{cases} < 0 & \text{when } \frac{\mu\chi_t}{r}(\phi - V_a(\pi_t)) < c \\ = 0 & \text{when } \frac{\mu\chi_t}{r}(\phi - V_a(\pi_t)) = c \\ > 0 & \text{when } \frac{\mu\chi_t}{r}(\phi - V_a(\pi_t)) > c \end{cases}$$

The following proof focuses on $t \in \mathbb{R}^+$ large enough such that $\pi_t < 1-b$, i.e. $a^*(\pi_t) \in (\phi, 1)$. Recall the patience level is high if and only if:

$$r < \frac{\mu\phi(1-c)}{c} \text{ or } \overline{V}_a > c\phi.$$

The following Lemma claims that the agent's continuation value converges in the limit.

Lemma B.1.1. *There exists $V_a(0) \in \mathbb{R}_+$ such that:*[1]

$$\lim_{t\to\infty}\hat{V}_a(t) = \lim_{\pi_t\to 0}V_a(\pi_t) = \lim_{\pi_t\to 0}V_a(\pi_t).$$

The proof of this Lemma is in Section B of the Online Appendix. Let $V_a(0) \equiv \lim_{t\to\infty}\hat{V}_a(t)$. In what follows, I will show that $V_a(0) = c\phi$ and then characterize players' behaviors in the limit.

Limiting Continuation Value in High Patience Case: I show that $V_a(0) = c\phi$ by ruling out all other possibilities.

[1]The fact that $\lim_{t\to\infty}\hat{V}_a(t) = \lim_{\pi_t\to 0}V_a(\pi_t)$ is a direct implication of Lemma 2.3.1.

Part I: Suppose $V_a(0) > \overline{V}_a$. Since $V_a(0) > \overline{V}_a > c\phi$, there exists $T \in \mathbb{R}^+$ such that $\hat{V}_a(t) > \overline{V}_a$ and $X_t > 0$ for all $t > T$. But

$$\frac{\mu\chi_t}{r}(\phi - V_a(\pi_t)) < c$$

for all $\chi_t \in [0,1]$ when $V_a(\pi_t) > \overline{V}_a$, and hence, $Z_t < 0$ for all $t > T$. This leads to a contradiction.

Part II: Suppose $V_a(0) < c\phi$. Then there exists $T \in \mathbb{R}^+$ such that $\hat{V}_a(t) < \overline{V}_a$ and $X_t < 0$ for all $t > T$. Hence, there exists $t > T$ such that $Z_t < 0$, i.e.

$$a(\pi_t) = \phi, \quad \frac{\mu\chi_t}{r}\Big(\phi - \hat{V}_a(t)\Big) < c.$$

Since $\hat{V}_a(t) < \overline{V}_a$, so $\chi_t \in (0,1)$, implying that $V_m(\pi_t) \geq b$. Admissibility requires the existence of $\varepsilon_0 > 0$ such that

$$a(\pi_{t+\varepsilon_1}) \in \Big[\phi, a^*(\pi_{t+\varepsilon_1})\Big) \text{ for every } \varepsilon_1 \in (0, \varepsilon_0).$$

Equation (2.7) implies that $V_m(\pi_{t+\varepsilon_1}) > b$, thus $\chi_{t+\varepsilon_1} = 0$, contradicting the conclusion of Lemma 2.3.1.

Part III: Suppose $V_a(0) \in (c\phi, \overline{V}_a)$. Then there exists $T \in \mathbb{R}^+$ such that $\hat{V}_a(t) < \overline{V}_a$ and $X_t > 0$ for all $t > T$. This implies the existence of $t > T$ such that:

$$a(\pi_t) = 1, \quad \frac{\mu\chi_t}{r}\Big(\phi - \hat{V}_a(t)\Big) > c.$$

The intermediary's incentive constraint requires that $\chi(\pi_t) > 0$ and $V_m(\pi_t) \leq b$. Admissibility (left-continuity with respect to belief or equivalently, right-continuity with respect to time) requires the existence of $\varepsilon_0 > 0$ such that

$$a(\pi_{t+\varepsilon_1}) \in \Big(a^*(\pi_{t+\varepsilon_1}), 1\Big) \text{ for every } \varepsilon_1 \in (0, \varepsilon_0).$$

Equation (2.7) implies that $V_m(\pi_{t+\varepsilon_1}) < b$ for all $\varepsilon_1 \in (0, \varepsilon_0)$, which further implies that $\chi_{t+\varepsilon_1} = 1$ for all $\varepsilon_1 \in (0, \varepsilon_0)$. But because $\hat{V}_a(t) < \overline{V}_a$ for all $t > T$, so $a_t = 1$ when $\chi_t = 1$. But then since $a^*(\pi_t) < 1$ for all $t > T$, we have $V_m(\pi_t) > b$, contradicting the previous conclusion that $V_m(\pi_t) \leq b$.

Part IV: Suppose $V_a(0) = \overline{V}_a$. Then there exists $T \in \mathbb{R}^+$ such that $X_t > 0$ for all $t > T$. If $V_a(\pi_t) < \overline{V}_a$ for all $\pi_t \in (0, \pi_T)$, we can obtain the same contradiction as in Part III.

Suppose there exists $t > T$ such that $\hat{V}_a(t) \geq \overline{V}_a$. Since $X_t > 0$ and $Z_t \leq 0$ for all t such that $\hat{V}_a(t) \geq \overline{V}_a$, so there exists $t' > t$ with $V_a(\pi_{t'}) < \overline{V}_a$. Since $V_a(\cdot)$ is a continuous function, there exists $t'' \in [t, t')$ such that

$$\hat{V}_a(t'') = \overline{V}_a, \quad \hat{V}_a'(t'') < 0.$$

Since $\hat{V}_a(t'') = \overline{V}_a$ suggests that $Z_{t''} \leq 0$, we have:

$$\overline{V}_a = c\phi + \underbrace{Z_{t''} + \frac{1}{r}\hat{V}_a'(t'')}_{<0} < c\phi,$$

which is a contradiction.

Behavior at the Limiting Belief: Next, I characterize players' equilibrium behaviors when π_t is close to 0. Since $V_a(0) = c\phi < \overline{V}_a$, there exists $\nu > 0$ such that $V_a(\pi_t) < \overline{V}_a$ for all $\pi_t < \nu$. The entire discussion will be focusing on $\pi_t < \nu$.

I start with showing that $a(\pi_t) < a^*(\pi_t)$ cannot happen when π_t is small enough. Suppose towards a contradiction, that there exists $\pi_t < \nu$ such that $a(\pi_t) < a^*(\pi_t)$, then since $a(\cdot)$ is left continuous, there exists $\varepsilon > 0$ such that $a(\pi') < a^*(\pi')$ for all $\pi' \in (\pi_t - \varepsilon, \pi_t]$. Since $\chi \neq 0$ for all t, and $V_a(\pi') < \overline{V}_a$, it has to be the case that $\chi(\pi') \in (0, 1)$, implying that $V_m(\pi') = b$ for all $\pi' \in (\pi_t - \varepsilon, \pi_t]$, and hence $V_m'(\pi') = 0$ for all $\pi' \in (\pi_t - \varepsilon, \pi_t)$. From (2.7), we have $a(\pi') = a^*(\pi')$ for all $\pi' \in (\pi_t - \varepsilon, \pi_t)$. But this implies that $a(\pi_t) = a^*(\pi_t)$, which is a contradiction.

Then I show that $a(\pi_t) > a^*(\pi_t)$ also cannot happen when $\pi_t \to 0$. The previous step implies the existence of $\nu > 0$ such that $a(\pi_t) \geq a^*(\pi_t)$ for all $\pi_t < \nu$, implying that $V_m(\pi_t) \geq b$. Since $\chi(\pi_t) > 0$, so $V_m(\pi_t) = b$, which is achieved only when $a(\pi_t) = a^*(\pi_t)$ for all $\pi_t \leq \nu$. This completes the proof.

B.1.3 Proof of Lemma 2.3.3

The existence and uniqueness of solution to the initial value problem is established in Lemma B.1.1 and the existence of solution to the limiting value problem is established in a Lemma in the Online Appendix. To show uniqueness in the limiting value problem, let $Z(\pi_t) \equiv \phi - V_a(\pi_t)$, re-write (2.13) as

$$-Z(\pi_t)^2 + Z(\pi_t)\Big(\phi - c\phi - \pi_t a^*(\pi_t)\Big) = ca^*(\pi_t)\pi_t(1-\pi_t)Z'(\pi_t). \tag{B.5}$$

Suppose towards a contradiction that there exists two solutions, $Z_1(\pi_t) \neq Z_2(\pi_t)$. Since

$$\lim_{\pi_t \to 0} Z_1(\pi_t) = \lim_{\pi_t \to 0} Z_2(\pi_t) = \phi - c\phi,$$

there exists π^* small enough such that:

$$Z_1(\pi), Z_2(\pi) \in [\phi - c\phi - \varepsilon, \phi - c\phi + \varepsilon], \text{ for all } \pi < \pi^*$$

and

$$Z_1(\pi^*) = Z_1 \neq Z_2 = Z_2(\pi^*).$$

where $\varepsilon \in (0, \frac{\phi - c\phi}{2})$. Without loss of generality, let $Z_1 > Z_2$. I show that $Z_1(\pi_t) - Z_2(\pi_t) \geq Z_1 - Z_2 > 0$ for all $\pi_t \in (0, \pi^*)$. To see this, differentiating the LHS by $Z(\pi_t)$, we have:

$$-2Z(\pi_t) + \phi - c\phi - \pi_t a(\pi_t) \leq -2(\phi - c\phi - \varepsilon) + \phi - c\phi - \pi_t a(\pi_t) < 0.$$

So, if $Z_1(\pi_t) > Z_2(\pi_t)$ then $Z_1'(\pi_t) < Z_2'(\pi_t)$. So $Z_1(\pi_t) - Z_2(\pi_t)$ is increasing when π_t decreases, i.e.

$$\lim_{\pi_t \to 0} Z_1(\pi_t) - \lim_{\pi_t \to 0} Z_2(\pi_t) \geq Z_1 - Z_2 > 0,$$

contradicting $\lim_{\pi_t \to 0} Z_1(\pi_t) = \lim_{\pi_t \to 0} Z_2(\pi_t) = \phi - c\phi$.

B.1.4 Proof of Lemma 2.3.5

Let $V_a^*(\cdot)$ be the unique solution to (2.13). Re-write the ODE as:

$$\pi_t(1 - \pi_t)V_a^{*'}(\pi_t) = \frac{1}{ca^*(\pi_t)}\big(\phi - V_a^*(\pi_t)\big)\big(c\phi + \pi_t a^*(\pi_t) - V_a^*(\pi_t)\big), \tag{B.6}$$

This implies that $V_a^{*'}(\pi_t) > 0$ if and only if

$$V_a^*(\pi_t) < \pi_t a^*(\pi_t) + c\phi. \tag{B.7}$$

I show that $V_a^*(\pi_t)$ is strictly increasing in π_t for $\pi_t \in (0, \pi^\dagger)$. Suppose towards a contradiction that there exists $\pi < \pi^\dagger$ such that:

$$V_a^*(\pi) \geq \pi a^*(\pi) + c\phi,$$

then $V_a^{*'}(\pi) \leq 0$. Since $\pi_t a^*(\pi_t) + c\phi$ is a strictly increasing function of π_t and $V_a^*(\pi_t)$ is decreasing in π_t for all $\pi_t \in (0, \pi)$, so

$$V_a^*(\pi_t) - \Big(\pi_t a^*(\pi_t) + c\phi\Big)$$

strictly increases as π_t decreases when $\pi_t \leq \pi$. This implies that $\lim_{\pi_t \downarrow 0} V_a^*(\pi_t) > c\phi$, which contradicts the limiting value condition.

B.1.5 Proof of Lemma 2.3.6

According to Lemma 2.3.2, $\pi^1 > 0$. Suppose towards a contradiction that $\pi^1 \in (0, \min\{1 - b, \pi^\dagger\})$. Since $\pi^1 < \pi^\dagger$, Lemma 2.3.5 implies that $V_a(\pi^1) < \overline{V}_a$. So there exists $\varepsilon > 0$ such that $V_a(\pi^1 + \varepsilon_0) < \overline{V}_a$ for all $\varepsilon_0 \in (0, \varepsilon)$. Throughout the proof, I will be focusing on $\pi_t \in [\pi^1, \pi^1 + \varepsilon)$.

By definition of π^1, for every $\varepsilon > 0$, there exists $\pi_t \in (\pi^1, \pi^1 + \varepsilon)$ such that $a(\pi_t) \neq a^*(\pi_t)$. Since $a(\pi_t)$ is left-continuous, we only need to consider the following three cases. Common in all cases, there exists an open interval $\Pi \equiv (\pi_t - \varepsilon_0, \pi_t) \subset (\pi^1, \pi^1 + \varepsilon)$, such that:

- **Case 1:** $a(\pi) \in (\phi, 1)$ but $a(\pi) \neq a^*(\pi_t)$ for all $\pi \in \Pi$. By the choice of ε, the agent's continuation value is strictly below $\overline{V}_a$ for all $\pi \in \Pi$. So $\chi(\pi) \in (0, 1)$ for all $\pi \in \Pi$. However, the intermediary's continuation value function requires that $a(\pi) = a^*(\pi)$ for all $\pi \in \Pi$, which is a contradiction.

- **Case 2:** $a(\pi) = 1$ for all $\pi \in \Pi$. Left-continuity of $a(\cdot)$ implies that $a(\pi_t) = 1$. Since $\chi(\pi) > 0$ for all $\pi \in \Pi \cup \{\pi_t\}$, we have $V_m(\pi) \le b$.

  The ODE in (2.7) implies that if $V_m(\pi_t) \le b$, then $V_m(\pi) < b$ for all $\pi \in \Pi$, and hence, $\chi(\pi) = 1$ by the intermediary's incentive constraint. Since the agent's continuation value is below $\overline{V}_a$ for all $\pi_t \le \pi^1 + \varepsilon$, so $a(\pi) = 1$ whenever $\chi(\pi) = 1$. Let

  $$\pi^2 \equiv \inf\left\{\pi \middle| \pi \in [\pi^1, \pi_t], \quad V_m(\pi') < b \text{ for all } \pi' > \pi\right\},$$

  then $\pi^2 = \pi^1$, implying that $V_m(\pi^1) < b$, contradicting the fact that $a(\pi^1) = a^*(\pi^1)$ and $\chi(\pi^1) \in (0, 1)$.

- **Case 3:** $a(\pi) = \phi$ for all $\pi \in \Pi$. Left-continuity of $a(\cdot)$ implies that $a(\pi_t) = \phi$. Then, $\chi(\pi) < 1$ for all $\pi \in \Pi \cup \{\pi_t\}$, which requires that $V_m(\pi) \ge b$.

  The ODE in (2.7) implies that if $V_m(\pi_t) \ge b$, then $V_m(\pi) > b$ for all $\pi \in \Pi$, so $\chi(\pi) = 0$ by the intermediary's incentive constraint. This contradicts Lemma 2.3.1.

B.1.6 Proof of Lemma 2.3.7

I start with the following Lemma.

Lemma B.1.2. *If $\pi^2 \ge \overline{\pi}$, then $V_a(\pi_t) > \overline{V}_a$ for all $\pi_t > \pi^2$.*

Proof of Lemma B.1.2: Recall the agent's continuation value satisfies:

$$V_a(\pi_t) = \pi_t a_t - (a_t - \phi)c + \frac{\mu\chi_t a_t}{r}\Big(\phi - V_a(\pi_t) - \pi_t(1-\pi_t)V_a'(\pi_t)\Big). \tag{B.8}$$

By definition, if $\pi^2 \ge \overline{\pi}$, then $\chi(\pi_t) = 1$ and $V_a(\pi_t) \ge \overline{V}_a$ for all $\pi_t \in (\pi^\dagger, \pi^2]$. Let

$$\pi^* \equiv \inf\left\{\pi_t \middle| \pi_t \in (\pi^2, 1) \text{ and } V_a(\pi_t) \le \overline{V}_a\right\},$$

Suppose towards a contradiction that $\pi^* < 1$, then $V_a'(\pi^*) < 0$. The intermediary's incentives imply that $\chi(\pi^*) = 1$. Plugging this back to (B.8) and by the fact that $\pi^* \ge \pi^2 \ge \overline{\pi}$, we have $V_a'(\pi^*) \ge 0$, leading to a contradiction. □

Lemma B.1.3. *If $\pi^2 < \overline{\pi}$, then there exists $\pi_t \in (\pi^2, \overline{\pi})$ such that $V_a(\pi_t) \ne \overline{V}_a$.*

Proof of Lemma B.1.3: Suppose towards a contradiction that $V_a(\pi_t) = \overline{V}_a$ for all $\pi_t \in (\pi^2, \overline{\pi}) \equiv \Pi_0$. For any $\pi_t \in \Pi_0$ such that $a(\pi_t) \ne a^{**}(\pi_t)$, since $V_a(\pi_t) = \overline{V}_a$,

- If $\chi(\pi_t) = 1$, then $V_a'(\pi_t) \ne 0$, leading to a contradiction.

- If $\chi(\pi_t) < 1$, then $a(\pi_t) = \phi$, plugging this into (B.8) and using the assumption that $V'_a(\pi_t) = 0$,

$$\overline{V}_a = \pi_t\phi + c\phi\chi_t < (1 - c - \frac{cr}{\mu\phi})\phi + c\phi = \phi - \frac{cr}{\mu} = \overline{V}_a,$$

  which leads to a contradiction.

This finishes the proof. □

Back to Lemma 2.3.7, suppose towards a contradiction that $\pi^2 \in [\pi^\dagger, 1)$. If $\pi^2 > \overline{\pi}$, then according to Lemma B.1.2, $V_a(\pi_t) > \overline{V}_a$ for all $\pi_t > \pi^2$. In this case, only $a(\pi_t) = \phi = a^{**}(\pi_t)$ is incentive compatible, which leads to a contradiction.

When $\pi^2 < \overline{\pi}$. There exists $\pi > \pi^2$, such that $a(\pi) \neq a^{**}(\pi)$. Lemma B.1.3 implies that there exists $\pi_t \in (\pi^2, \overline{\pi})$ such that $V_a(\pi_t) \neq \overline{V}_a$. Two subcases are examined separately.

1. If $V_a(\pi_t) > \overline{V}_a$, then by the continuity of $V_a(\cdot)$ and Lemma 2.3.4, there exists $\pi'_t \in (\pi^2, \pi_t)$ such that

$$a(\pi'_t) > \phi, \quad V_a(\pi') > \overline{V}_a.$$

   This leads to a contradiction since when $V_a(\pi') > \overline{V}_a$, the only incentive compatible effort level is $a(\pi') = \phi$, regardless of $\chi(\pi')$.

2. If $V_a(\pi_t) < \overline{V}_a$, then there exists $\pi' \in (\pi^2, \pi_t)$ such that

$$a(\pi') < a^{**}(\pi'), \quad V_a(\pi') < \overline{V}_a.$$

   Since $a(\cdot)$ is left-continuous and $V_a(\cdot)$ is continuous, there exists $\varepsilon_0 \in (0, \pi_t - \pi^2)$ such that

$$a(\pi'') < a^{**}(\pi''), \quad V_a(\pi'') < \overline{V}_a$$

   for all $\pi'' \in (\pi' - \varepsilon_0, \pi') \equiv \Pi$. This also implies that $\chi(\pi'') < 1$ for all $\pi'' \in \Pi$. The intermediary's incentive constraint requires that $V_m(\pi' - \frac{\varepsilon_0}{2}) \geq b$.

   From the ODE in (2.7), $V_m(\pi'') > b$ for all $\pi'' \in (\pi' - \varepsilon_0, \pi' - \frac{\varepsilon_0}{2})$. This implies that $\chi(\pi'') = 0$, contradicting the conclusion of Lemma 2.3.1.

B.1.7 Proof of Lemma 2.3.8

I start from the following Lemma:

Lemma B.1.4. *If $\pi_0 < \phi$, then for any $t \in \mathbb{R}_+$, $\boldsymbol{\chi} \in \mathbf{X}$, $\mathbf{a} \in \mathbf{A}$ and $\hat{\mathbf{a}} \in \mathbf{A}$:*

$$\mathbb{E}^{\hat{\mathbf{a}}, \boldsymbol{\chi}, \pi_0}[\pi_{t+dt} | \theta = 1, \mathbf{a}] > \pi_t. \tag{B.9}$$

Proof of Lemma B.1.4: Conditional on $\theta = 1$, the probability breakthrough at $[t, t + dt]$ is at least $\mu\chi_t\phi dt$, after which $\pi_t = 1$. With complementary probability, π_t degrades to

$$\pi_t - \mu\chi_t\hat{a}_t\pi_t(1 - \pi_t)dt.$$

The expected belief at $t + dt$ exceeds π_t if:

$$\mu\chi_t\phi dt + (1 - \mu\chi_t\phi dt)(\pi_t - \mu\chi_t\hat{a}_t\pi_t(1 - \pi_t)dt) > \pi_t \tag{B.10}$$

Ignoring higher order terms, we get:

$$\phi > \hat{a}_t\pi_t, \tag{B.11}$$

(B.11) implies that $\pi_0 < \phi$ is sufficient for (B.9). $\square$

Back to the proof of Lemma 2.3.8, for any given $\boldsymbol{\chi}$ and a *correct* market belief $\hat{\mathbf{a}}$, the agent's continuation value is:

$$V_a^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t) = \pi_t\hat{a}(\pi_t) - c\Big(a(\pi_t) - \phi\Big) + \frac{\mu\chi(\pi_t)a(\pi_t)}{r}\Big(\phi - V_a^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t)\Big) - \frac{\mu\chi(\pi_t)\hat{a}(\pi_t)}{r}\pi_t(1 - \pi)V_a^{\hat{\mathbf{a}},\boldsymbol{\chi}\prime}(\pi_t). \tag{B.12}$$

Let $\Gamma^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t) \equiv V_a^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t) - \pi_t\phi$, we have:

$$\begin{aligned}\Gamma^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t) = \pi_t\Big(\hat{a}(\pi_t) - \phi\Big) - c\Big(a(\pi_t) - \phi\Big) + \frac{\mu\chi(\pi_t)a(\pi_t)}{r}\Big(\phi(1 - \pi_t) - \Gamma^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi_t)\Big) \\ -\frac{\mu\chi(\pi_t)\hat{a}(\pi_t)}{r}\pi_t(1 - \pi_t)\Gamma^{\hat{\mathbf{a}},\boldsymbol{\chi}\prime}(\pi_t) - \frac{\mu\phi\chi(\pi_t)\hat{a}(\pi_t)}{r}\pi_t(1 - \pi_t).\end{aligned} \tag{B.13}$$

In what follows, I write $\Gamma(\cdot)$ instead of $\Gamma^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\cdot)$ for notation simplicity. Let

$$\pi^* \equiv \sup\Big\{\pi_t \leq 1\Big| V_a^{\hat{\mathbf{a}},\boldsymbol{\chi}}(\pi) \leq \phi\pi \text{ for all } \pi \leq \pi_t\Big\}.$$

Lemma B.1.4 suggests that when $\pi_0 < \phi$, the agent can guarantee himself payoff $\pi_0\phi$ by playing $a_t = \phi$ forever, so $\pi^* \geq \phi > 0$.

I show that $\pi^* = 1$. Suppose towards a contradiction that $\pi^* < 1$. Since $\Gamma(\pi_t)$ is continuous, we have:

$$\Gamma(\pi^*) = 0, \quad \Gamma'(\pi^*) < 0.$$

From (B.13), we have:

$$0 > \pi^*\Big(\hat{a}(\pi^*) - \phi\Big) - c\Big(a(\pi^*) - \phi\Big) + \frac{\mu\chi(\pi^*)a(\pi^*)}{r}\phi(1 - \pi^*) - \frac{\mu\phi\chi(\pi^*)\hat{a}(\pi^*)}{r}\pi^*(1 - \pi^*). \tag{B.14}$$

I consider three cases, depending on the market's believed effort $\hat{a}$ as well as the strength of reputation concerns, $\frac{\mu\chi(\pi^*)}{r}\phi(1 - \pi^*) - c$. I will obtain a contradiction in each case based on (B.14).

Case 1: If $\hat{a}(\pi^*) = \phi$, then $a(\pi^*) = \phi$ must be a best reply since the market's belief is correct, we have:

$$\begin{aligned}0 \quad &> \quad \frac{\mu\chi(\pi^*)\phi}{r}\phi(1 - \pi^*) - \frac{\mu\phi\chi(\pi^*)}{r}\phi\pi^*(1 - \pi^*) \\ &= \quad \frac{\mu\chi(\pi^*)\phi^2}{r}(1 - \pi^*)^2 \geq 0,\end{aligned}$$

which is a contradiction.

Case 2: If $\hat{a}(\pi^*) \in [\phi, 1]$ and $\frac{\mu\chi(\pi^*)}{r}\phi(1-\pi^*) = c$, then the agent is indifferent between all effort levels, so

$$\begin{aligned} 0 &> \pi^*\Big(\hat{a}(\pi^*) - \phi\Big) + c\phi - \underbrace{\frac{\mu\phi\chi(\pi^*)\hat{a}(\pi^*)}{r}\pi^*(1-\pi^*)}_{=c\pi^*\hat{a}(\pi^*)}. \\ &= (1-c)\pi^*\hat{a}(\pi^*) + \phi(c-\pi^*) \\ &\geq (1-c)\pi^*\phi + \phi(c-\pi^*) = \phi c(1-\pi^*) > 0. \end{aligned}$$

which is a contradiction.

Case 3: If $\hat{a}(\pi^*) = 1$ and $\frac{\mu\chi(\pi^*)}{r}\phi(1-\pi^*) > c$, then $a(\pi^*) = 1$ is the agent's best response. Then

$$\begin{aligned} 0 &> \pi^*\Big(1-\phi\Big) - c\Big(1-\phi\Big) + \frac{\mu\chi(\pi^*)}{r}\phi(1-\pi^*) - \frac{\mu\phi\chi(\pi^*)}{r}\pi^*(1-\pi^*). \\ &= (\pi^*-c)(1-\phi) + \frac{\mu\phi\chi(\pi^*)}{r}(1-\pi^*)^2 \\ &> (\pi^*-c)(1-\phi) + c(1-\pi^*) = \pi^*(1-c) + \phi(c-\pi^*). \end{aligned}$$

If $\pi^* \leq c$, the last expression is strictly positive, which leads to a contradiction. If $\pi^* > c$ and $\pi^* \geq \phi$,

$$\pi^*(1-c) + \phi(c-\pi^*) \geq \pi^*(1-c) + \pi^*(c-\pi^*) = \pi^*(1-\pi^*) > 0,$$

which leads to a contradiction. If $\pi^* > c$ but $\pi^* < \phi$,

$$\pi^*(1-c) + \phi(c-\pi^*) \geq \pi^*(1-c) + (c-\pi^*) = c(1-\pi^*) > 0,$$

which also leads to a contradiction.

B.2 Exogenous Information Benchmark

In this Appendix, I show Proposition 2.1 and Proposition 2.3.

B.2.1 Proof of Proposition 2.1

I start with computing the agent's continuation value when $\pi_t \to 0$ if $\chi_t = 1$ for all t. Applying integral expression (A.3) and plugging in $\chi_t = 1$, we have:

$$\hat{V}_a(t) - c\phi = X_t + W_t.$$

For any $\varepsilon > 0$, there exists $T \in \mathbb{R}_+$ such that $W_t \in (0, \varepsilon)$ for all $t > T$.

Suppose towards a contradiction that $\hat{V}_a(\infty) \leq c\phi$, then $X_t \to 0$ and there exists $\varepsilon > 0$, such that:

$$\frac{\mu}{r}(\phi - \hat{V}_a(t)) > \frac{\mu}{r}(\phi - c\phi) > c + \varepsilon,$$

for t large enough. Hence, $Z_t \geq \varepsilon$, i.e. $X_t \geq \varepsilon$ for t large enough. But then,

$$\lim_{t\to\infty} \hat{V}_a(t) = \lim_{t\to\infty} W_t + \lim_{t\to\infty} X_t + c\phi \geq c\phi + \varepsilon > c\phi,$$

leading to a contradiction.

So $\hat{V}_a(\infty) > c\phi$, implying that $\lim_{t\to\infty} X_t > 0$, i.e. there exists $T \in \mathbb{R}_+$ such that $Z_t > 0$ for all $t > T$. Thus, $\hat{V}_a(\infty)$ satisfies:

$$\hat{V}_a(\infty) = c\phi + \lim_{t\to\infty} X_t = c\phi + r\int_0^\infty \Big(\frac{\mu}{r}(\phi - \hat{V}_a(\infty)) - c\Big)dt, \tag{B.15}$$

which gives:

$$\hat{V}_a(\infty) = V_a(0) = \frac{\mu\phi}{\mu + r} - \frac{rc(1-\phi)}{\mu + r}. \tag{B.16}$$

Next, I construct an MPE of the exogenous information game and apply the result of Faingold and Sannikov (2011) to establish uniqueness.[2] Consider the following limiting value problem:

$$V_a(\pi_t) = \pi_t - c(1-\phi) + \frac{\mu}{r}\Big(\phi - V_a(\pi_t) - \pi_t(1-\pi_t)V_a'(\pi_t)\Big),$$

with (B.16) the limiting value condition. I will show that this problem admits a unique solution in Section A of the Online Appendix. Let $V_a^{**}(\pi_t)$ be the solution. Let $\pi^\ddagger$ be defined by:

$$\pi^\ddagger \equiv \inf\Big\{\pi_t > 0 \Big| V_a^{**}(\pi_t) = \overline{V}_a\Big\}.$$

From Lemma 2.3.4, $\pi^\ddagger > \underline{\pi}$. From Lemma 2.3.8, $\pi^\ddagger < 1 - \frac{cr}{\mu\phi}$. Let

$$a(\pi_t) = \begin{cases} 1 & \text{when } \pi_t \geq \pi^\ddagger \\ a^{**}(\pi_t) & \text{when } \pi_t < \pi^\ddagger. \end{cases} \tag{B.17}$$

I claim that this strategy and its induced market belief system induce an MPE. First, $V_a^{**}(\pi_t)$ is the agent's continuation value when $\pi_t \leq \pi^\ddagger$. By definition of $\pi^\ddagger$, $V_a^{**}(\pi_t) \leq \overline{V}_a$ for all $\pi_t \leq \pi^\ddagger$, implying that $a(\pi_t) = 1$ is incentive compatible for the agent. Second, since $V_a(\pi^\ddagger) = \overline{V}_a$, so

- If $\pi^\ddagger < \overline{\pi}$, then since $a(\pi_t) = a^{**}(\pi_t)$, we have $V_a(\pi_t) = \overline{V}_a$ for all $\pi_t \in [\pi^\ddagger, \overline{\pi}]$ (Lemma 2.3.4), implying that every effort level is optimal. If $\pi_t > \overline{\pi}$, then $V_a(\pi_t) > \overline{V}_a$, implying that $a(\pi_t) = \phi$ is optimal for the agent.

- If $\pi^\ddagger \geq \overline{\pi}$, then $V_a(\pi_t) > \overline{V}_a$ for all $\pi_t > \pi^\ddagger$, i.e. $a(\pi_t) = \phi$ is optimal for the agent.

[2]The exogenous information benchmark fits into the definition of Poisson good news model in Faingold and Sannikov (2011), subsection 9.1 page 823, and the uniqueness result is reported in Theorem 11 of their paper.

B.2.2 Proof of Proposition 2.3

Let $V_1(\cdot)$ be the unique solution to the following limiting value problem:

$$V_a(\pi) = \pi a^*(\pi) + c\phi - \frac{c}{\phi - V_a(\pi)}\pi(1-\pi)V_a'(\pi),$$

with $\lim_{\pi\to 0} V_a(\pi) = c\phi$. Let $V_2(\cdot)$ be the unique solution to the following limiting value problem:

$$V_a(\pi) = \pi - c(1-\phi) + \frac{\mu}{r}\Big(\phi - V_a(\pi) - \pi(1-\pi)V_a'(\pi)\Big),$$

with $\lim_{\pi\to 0} V_a(\pi) = \frac{r}{\mu+r}\overline{V}_a + \frac{\mu}{\mu+r}c\phi$.

Recall that $\pi^\dagger$ and $\pi^\ddagger$ are defined by $V_1(\pi^\dagger) = \overline{V}_a$ and $V_2(\pi^\ddagger) = \overline{V}_a$ respectively. The following Lemma puts upper bounds on $V_1(\cdot)$ and $V_2(\cdot)$.

Lemma B.2.1. *For all $\pi_t \in (0, \pi^\dagger]$,*

$$V_1(\pi_t) \le \pi_t a^*(\pi_t) + c\phi.$$

For all $\pi_t \in (0, \pi^\ddagger]$,

$$V_2(\pi_t) \le \frac{r}{\mu+r}\pi_t + \frac{r}{\mu+r}\overline{V}_a + \frac{\mu}{\mu+r}c\phi.$$

This Lemma will be shown in the Online Appendix.[3] Let

$$\pi^* \equiv \frac{2\mu + r}{2\mu + 2r}\phi, \tag{B.18}$$

which is less than $\frac{1}{2}$ when ϕ is small enough ($\phi < \frac{1}{2}$). Since $V_1(\pi^*) < c\phi + \pi^* a(\pi^*) < c\phi + \pi^*$, we have $V_1(\pi^*) < \overline{V}_a$. I will be focusing on the case in which $b = 1 - \pi^*$. The claim can be extended to an open neighborhood of b by a continuity argument. The key step is the following Lemma:

Lemma B.2.2. *When $\phi < \frac{1}{2}$, for every $\underline{r} > 0$, there exists $\overline{c} \in (0,1)$ and $\underline{\mu} > 0$ satisfying*

$$\phi - \underline{c}\phi - \frac{\underline{r}\overline{c}}{\underline{\mu}} > 0$$

such that for every $c < \overline{c}$, $\mu > \underline{\mu}$ and $r \in \left(\underline{r}, \frac{\underline{\mu}\phi(1-\overline{c})}{\overline{c}}\right)$, we have:

$$V_2(\pi^*) < V_1(\pi^*). \tag{B.19}$$

Proof of Lemma B.2.2: The proof is divided into two parts. In Part I, I show that $V_1(\pi^*)$ is arbitrarily close to π^* when $c \to 0$ and r is sufficiently large. In Part II, I show that when μ and r are both large, $V_2(\pi^*)$ is bounded away from π^* even when $c \to 0$.

[3] The first inequality is a Corollary of Lemma A.3 in the Online Appendix. The second inequality is a Corollary of Lemma A.4 in the Online Appendix.

Part I: Lemma B.2.1 implies that $V_1(\pi^*) \leq \pi^* + c\phi$, so

$$\overline{V}_a - V_1(\pi^*) \geq \phi - c\phi - \frac{cr}{\mu} - \pi^* = \frac{r\phi}{2(\mu+r)} - (c\phi + \frac{cr}{\mu}).$$

Since

$$\frac{\mu\chi(\pi^*)}{r}\Big(\phi - V_1(\pi^*)\Big) = c,$$

we have:

$$\chi(\pi^*) \leq \frac{cr}{\mu}\frac{1}{\phi\frac{r}{2\mu+2r} - c(\phi + \frac{r}{\mu})}.$$

Notice that the agent's flow payoff is bounded above by 1 and below by -1. For any $r > 0$ and $\varepsilon_0 > 0$, there exists $T > 0$ such that

$$r\int_0^T e^{-rt}dt > 1 - \varepsilon_0.$$

Let π_t be defined via the following ODE

$$\dot{\pi}_t = -\mu\pi_t(1-\pi_t)\chi(\pi_t), \tag{B.20}$$

with initial value $\pi_0 = \pi^*$. Then for every $\varepsilon_1 > 0$, there exists c small enough and μ large enough such that

$$\pi^* - \pi_T < \varepsilon_1.$$

Also, by definition of π^*, which is less than ϕ, the agent's flow payoff at time t is at least π_t. Now, we can compute a lower bound on $V_1(\pi^*)$:

$$V_1(\pi^*) \geq r\int_0^T e^{-rt}\min\Big\{\pi^* - \varepsilon_1, \phi\Big\}dt + r\int_T^\infty e^{-rt}(-1)dt \geq (\pi^* - \varepsilon_1)(1-\varepsilon_0) - \varepsilon_0,$$

which converges to π^* as both ε_0 and ε_1 go to 0.

Part II: I will show that $V_2(\pi^*)$ is bounded below by π^*, with difference bounded away from 0 even when $c \to 0$. This can imply that $V_2(\pi^*) < V_1(\pi^*)$. I establish an upper bound on t^*, defined via $\pi_{t^*} = \frac{\phi}{3}$, where π_t is the solution to initial value problem (B.20).

$$t^* \leq \frac{\frac{2\mu+r}{2\mu+2r}\phi - \frac{\phi}{3}}{\frac{\phi}{3}(1-\frac{\phi}{3})\mu} = \frac{1}{2(1-\frac{\phi}{3})}\frac{4\mu+r}{\mu(\mu+r)}.$$

Intuitively, t^* is the time it takes for market to belief to go from π^* to $\frac{\phi}{3}$. Let $C_1 \equiv \frac{1}{2(1-\frac{\phi}{3})}$. We can then obtain an upper bound on $V_2(\pi^*)$:

$$\begin{aligned} V_2(\pi^*) &\leq r\int_0^{C_1\frac{4\mu+r}{\mu(\mu+r)}} e^{-rt}dt + V_2(\frac{\phi}{3})\Big(1 - r\int_0^{C_1\frac{4\mu+r}{\mu(\mu+r)}} e^{-rt}dt\Big) \\ &\leq \Big(1 - e^{-rC_1\frac{4\mu+r}{\mu(\mu+r)}}\Big) + e^{-rC_1\frac{4\mu+r}{\mu(\mu+r)}}\Big(\underbrace{\frac{\phi}{3}\frac{r}{\mu+r} + \frac{\mu\phi}{\mu+r}}_{=\pi^*-\frac{\phi}{6}\frac{r}{\mu+r}} - \frac{rc(1-\phi)}{\mu+r}\Big). \end{aligned}$$

For every $\varepsilon > 0$, there exists $\underline{\mu} > 0$ such that for all $\mu > \underline{\mu}$,

$$\Big(1 - e^{-rC_1\frac{4\mu+r}{\mu(\mu+r)}}\Big) + (\frac{\phi}{2} + c\phi)e^{-rC_1\frac{4\mu+r}{\mu(\mu+r)}} \leq \pi^* - \frac{\phi}{6}\frac{r}{r+\mu} + \varepsilon.$$

□

Lemma B.2.2 has found a belief, π^*, such that $V_2(\pi^*) < V_1(\pi^*) < \overline{V}_a$. What remains to be shown is that $V_1(\pi_t) > V_2(\pi_t)$ for all $\pi_t \in [\pi^*, \min\{\pi^\dagger, \pi^\ddagger\}]$. Since $1 - b = \pi^*$, so when $\pi_t > \pi^*$, $(a_t, \chi_t) = (1, 1)$ in both scenarios, implying that $V_1(\cdot)$ reaches $\overline{V}_a$ at a smaller belief than $V_2(\cdot)$.[4]

B.3 Semi-Markov Equilibrium

B.3.1 Characterizing π^* and π^{**}

In this subsection, I characterize the two belief thresholds, π^* and π^{**}, as functions of the cutoff $\pi^\S$, which we write as $\pi^*(\pi^\S)$ and $\pi^{**}(\pi^\S)$, respectively. I characterize the agent's continuation value in each phase, as well as the thresholds π^* and π^{**}, as functions of $\pi^\S$. Let $V_a^*(\cdot)$ be the unique solution to limiting value problem (2.13), which is the agent's continuation value in Phase I.

Recall that $\pi^\dagger$ is defined in (2.14) by $V_a^*(\pi^\dagger) = \overline{V}_a$. For any $\pi^\S \in [0, \min\{1-b, \pi^\dagger\}]$, Lemma 2.3.5 implies that $V_a^*(\pi^\S) < \overline{V}_a$. Let $V_a^1(\cdot|\pi^\S)$ be the unique solution to the following initial value problem:

$$V_a(\pi_t) = \pi_t - c(1-\phi) + \frac{\mu}{r}\Big(\phi - V_a(\pi_t) - \pi_t(1-\pi_t)V_a'(\pi_t)\Big), \tag{B.21}$$

with $V_a(\pi^\S) = V_a^*(\pi^\S)$. By definition, $V_a^1(\pi_t|\pi^\S)$ is the agent's continuation value in Phase II when $\pi_t \leq \pi^*(\pi^\S)$, with $\pi^*(\pi^\S)$ defined as:

$$\pi^*(\pi^\S) \equiv \sup\Big\{\pi_t \Big| \pi_t \geq \pi^\S,\ V_a^1(\pi|\pi^\S) < \overline{V}_a \text{ for all } \pi < \pi_t\Big\}. \tag{B.22}$$

If this set is empty, then $\pi^*(\pi^\S) = \pi^\S$. By definition,

$$\pi^*(0) = \pi^\ddagger \text{ and } \pi^*\Big(\min\{1-b, \pi^\dagger\}\Big) = \pi^\dagger.$$

[4]This perverse incentive also exists when the intermediary is the agent's direct supervisor, i.e. her flow payoff is θa_t instead of $\theta a_t - w_t$. The proof of this extension is available upon request.

Let $a^1(\pi_t|\pi^{\S})$ be the agent's on-path effort given that punishment phase starts at $\pi^{\S}$. Let $Y^1(\pi_t|\pi^{\S}) \equiv \mu a^1(\pi_t|\pi^{\S})$ be market's on-path learning rate. Let $\hat{V}_a^0(\cdot|\pi^{\S})$ be the (unique) solution to the following initial value problem:

$$V_a(\pi_t) = \pi_t a^1(\pi_t|\pi^{\S}) + c\phi - \frac{Y^1(\pi_t|\pi^{\S})}{r}\pi_t(1-\pi_t)V_a'(\pi_t). \tag{B.23}$$

with $V_a(\pi^{\S}) = V_a^*(\pi^{\S})$. Define $\hat{\pi}^{**}(\pi^{\S})$ as:

$$\hat{\pi}^{**}(\pi^{\S}) \equiv \sup\left\{\pi_t \middle| \pi_t \geq \pi^{\S},\ \hat{V}_a^0(\pi|\pi^{\S}) < \overline{V}_a \text{ for all } \pi < \pi_t\right\}. \tag{B.24}$$

If this set is empty, then $\hat{\pi}^{**}(\pi^{\S}) = \pi^{\S}$. Suppose $\hat{\pi}^{**}(\pi^{\S}) \leq 1-b$, then

$$\pi^{**}(\pi^{\S}) \equiv \hat{\pi}^{**}(\pi^{\S}). \tag{B.25}$$

Otherwise, record the value $\hat{V}_a^0(1-b|\pi^{\S})$. Let $\tilde{V}_a^0(\cdot|\pi^{\S})$ be the (unique) solution to the following initial value problem:

$$V_a(\pi_t) = \pi_t a^1(\pi_t|\pi^{\S}) + c\phi + \frac{\mu}{r}\Big(\phi - V_a(\pi_t)\Big) - \frac{Y^1(\pi_t|\pi^{\S})}{r}\pi_t(1-\pi_t)V_a'(\pi_t). \tag{B.26}$$

with $V_a(1-b) = \hat{V}_a^0(1-b|\pi^{\S})$. Define $\tilde{\pi}^{**}(\pi^{\S})$ as:

$$\tilde{\pi}^{**}(\pi^{\S}) \equiv \sup\left\{\pi_t \middle| \pi_t \geq 1-b,\ \tilde{V}_a^0(\pi|\pi^{\S}) < \overline{V}_a \text{ for all } \pi < \pi_t\right\}. \tag{B.27}$$

Then we have:

$$\pi^{**}(\pi^{\S}) \equiv \tilde{\pi}^{**}(\pi^{\S}). \tag{B.28}$$

B.3.2 Proof of Proposition 2.4

First, the proof of Lemma 2.3.1 directly carries over to the Semi-Markov case which implies that $\pi_t(1-\pi_t) \to 0$ in probability. I use $\{a(\pi_t), \chi(\pi_t)\}$ to represent players' on-path strategies. The next Lemma is the counterpart of Lemma 2.3.2 which characterizes the agent's continuation value in a generic SME when $\pi_t \to 0$.

Lemma B.3.1. *When players' patience level is high, either one of the following two statements is true:*[5]

- $\lim_{\pi_t \to 0} V_a(\pi_t) = c\phi$ *and there exists* $\nu > 0$ *such that for all* $\pi_t \leq \nu$, $a(\pi_t) = a^*(\pi_t)$ *and* $\chi(\pi_t) < 1$.
- $\lim_{\pi_t \to 0} V_a(\pi_t) = \frac{\mu\phi - rc(1-\phi)}{\mu + r}$ *and there exists* $\nu > 0$ *such that for all* $\pi_t \leq \nu$, $a(\pi_t) = 1$ *and* $\chi(\pi_t) = 1$.

[5] Recall that $c\phi$ is the agent's limiting continuation value in the unique MPE and $\frac{\mu\phi - rc(1-\phi)}{\mu+r}$ is his limiting continuation value under exogenous information.

Comparing with Lemma 2.3.2, SME admits another possibility (statement 2). This is because $(a_t, \chi_t) = (1, 1)$ can be sustained by low effort low disclosure rate off the equilibrium path.

Proof of Lemma B.3.1: Let $V_a(\pi_t), V_m(\pi_t), a(\pi_t)$ and $\chi(\pi_t)$ be the agent's and the intermediary's continuation value, effort and disclosure rate on the equilibrium path. Let $\hat{V}_a(t) \equiv V_a(\pi_t)$, $\hat{V}_m(t) = V_m(\pi_t)$. Define X_t, Z_t and W_t as in the proof of Lemma 2.3.2, we have:

$$\hat{V}_a(t) \quad = \quad c\phi + X_t + W_t. \tag{B.29}$$

Since the only relevant off-path is the intermediary withholding information when $\chi(\pi_t) = 1$, then the proofs in Part I, II and IV of Lemma 2.3.2 directly go through. I modify the argument in Part III.

Modified Part III: Suppose $V_a(0) \in (c\phi, \overline{V}_a)$.[6] Then there exists $T \in \mathbb{R}_+$ such that $\hat{V}_a(t) < \overline{V}_a$ and $X_t > 0$ for all $t > T$. So, there exists $t^* > T$ such that:

$$a_{t^*} = 1 \text{ and } \frac{\mu \chi_{t^*}}{r}\left(\phi - \hat{V}_a(t^*)\right) > c.$$

This also suggests that $\chi_t > 0$, which implies $V_m(\pi_t) \leq b$. Two cases are considered

- Suppose $\chi_t \neq 1$, then the Semi-Markov restriction implies that continuation play should not depend on whether a breakthrough has arrived or not, conditional on no disclosure. Admissibility requires the existence of $\varepsilon_0 > 0$ such that

  $$a(\pi_{t+\varepsilon_1}) \in (a^*(\pi_{t+\varepsilon_1}), 1) \text{ for all } \varepsilon_1 \in (0, \varepsilon_0).$$

  The ODE in (2.7) implies that $V_m(\pi_{t+\varepsilon_1}) < b$ for all $\varepsilon_1 \in (0, \varepsilon_0)$, which further implies that $\chi_{t+\varepsilon_1} = 1$ for all $\varepsilon_1 \in (0, \varepsilon_0)$. Since $\hat{V}_a(t) < \overline{V}_a$ for all $t > T$, so $a_t = 1$ and $\chi_t = 1$. But then $1 = V_m(\pi_t) > b$, contradicting $V_m(\pi_t) \leq b$. This contradiction applies as long as for every $T \in \mathbb{R}^+$, there exists $t > T$ such that

  $$Z_t > 0 \text{ and } \chi_t < 1.$$

- Next I consider the case in which $\chi_t = 1$ for all $t > T$ satisfying $Z_t > 0$.

  - If there exists $T \in \mathbb{R}^+$, such that $a_t = 1$ for all $t \geq T$, then we have:

    $$V_a(0) = \frac{\mu\phi - rc(1-\phi)}{\mu + r}.$$

  - If for every $T \in \mathbb{R}^+$, there exists $t > T$ such that $a(\pi_t) < 1$. Since $\hat{V}_a(t) < \overline{V}_a$, it is required that $\chi_t < 1$ and $V_m(\pi_t) = b$. Since $X_t > 0$, for every such t, there exists $t' > t$

[6] This case also treats the situation in which $V_a(0) = \overline{V}_a$ but $\hat{V}_a(t) < \overline{V}_a$ for all t large enough.

such that $a(\pi_{t'}) = 1$. Admissibility requires that for every t, there exists $\varepsilon > 0$ such that

$$a(\pi_{t+\varepsilon_0}) < 1 \text{ for all } \varepsilon_0 \in (0, \varepsilon).$$

This requires that $\chi(\pi_{t+\varepsilon_0}) \in (0,1)$, i.e. $V_m(\pi_{t+\varepsilon_0}) = b$. The ODE in (2.7) then requires that $a(\pi_{t+\varepsilon_0}) = a^*(\pi_{t+\varepsilon_0})$.

This suggests that $a(\pi_t) \geq a^*(\pi_t)$ for all $t \geq T$. Also, for every t such that $a(\pi_t) = a^*(\pi_t)$, there exists $t' > t$ with $a(\pi_{t'}) = 1 > a^*(\pi_{t'})$. But then $V_m(\pi_t) > b$, implying that $\chi_t = 0$, contradicting the conclusion that $\chi_t \neq 0$.

□

Next, I characterize players' *on-path* behavior when π_t is bounded away from 0. The next two Lemmas examine cases when $V_a(\pi_t) < \overline{V}_a$.

Lemma B.3.2. *For any π_t such that $V_a(\pi_t) < \overline{V}_a$, $a(\pi_t) \in \{a^*(\pi_t), 1\}$.*

Proof of Lemma B.3.2: Suppose towards a contradiction that there exists such π_t. Then let

$$\pi^* \equiv \inf \left\{ \pi \middle| V_a(\pi) < \overline{V}_a, \quad a(\pi_t) \notin \{a^*(\pi_t), 1\} \right\}. \tag{B.30}$$

By left-continuity of $a(\cdot)$, there exists $\varepsilon > 0$ such that $a(\pi^* + \varepsilon_0) \notin \{a^*(\pi^* + \varepsilon_0), 1\}$ for all $\varepsilon_0 \in (0, \varepsilon)$. Since $V_a(\pi_t) < \overline{V}_a$, so $\chi(\pi_t) < 1$ for all $\pi_t \in (\pi^*, \pi^* + \varepsilon)$. Since $\chi(\pi_t) \neq 0$, so $\chi(\pi_t) \in (0,1)$, implying that $V_m(\pi_t) = b$ for all $\pi_t \in (\pi^*, \pi^* + \varepsilon)$. But the intermediary's continuation value satisfies:

$$V_m(\pi_t) = a(\pi_t)(1 - \pi_t) + \frac{\mu\chi(\pi_t)a(\pi_t)}{r} \Big\{ b - V_m(\pi_t) - \pi_t(1 - \pi_t)V_m'(\pi_t) \Big\}.$$

Then $V_m(\pi_t) = b$ and $V_m'(\pi_t) = 0$ imply that $a(\pi_t) = a^*(\pi_t)$, leading to a contradiction. □

Lemma B.3.3. *If there exists π^* such that $a(\pi^*) = \chi(\pi^*) = 1$, then for any $\pi_t > \pi^*$ with $V_a(\pi_t) < \overline{V}_a$, we have $a(\pi_t) = \chi(\pi_t) = 1$.*

Proof of Lemma B.3.3: Suppose towards a contradiction that there exists such π_t, then $a(\pi_t) < 1$ while $V_a(\pi_t) < \overline{V}_a$ implies that $\chi(\pi_t) < 1$. Let

$$\pi^{**} \equiv \inf \left\{ \pi \middle| \pi > \pi^*, \quad V_a(\pi) < \overline{V}_a, \quad \chi(\pi) < 1, \quad a(\pi_t) < 1 \right\}. \tag{B.31}$$

By the left-continuity of $a(\cdot)$ and $\chi(\cdot)$, there exists $\varepsilon > 0$ such that $a(\pi^{**} + \varepsilon_0) < 1$ and $\chi(\pi^{**} + \varepsilon_0) < 1$ for all $\varepsilon_0 \in (0, \varepsilon)$. Then since withholding information happens on path before π_t reaches π^{**}, so according to the Semi-Markov restriction as well as Lemma B.3.2, $V_m(\pi^{**}) > b$. But then, $\chi(\pi^{**}) = 0$, which is a contradiction. □

Putting together Lemma B.3.1, Lemma B.3.2 and Lemma B.3.3, for all π_t such that $V_a(\pi_t) < \overline{V}_a$, there exists a cut-off belief $\pi^{\S} \in [0, \min\{1-b, \pi^{\dagger}\}$, such that

$$a(\pi_t) = \begin{cases} a^*(\pi_t) & \text{when } \pi_t \leq \pi^{\S} \\ 1 & \text{when } \pi_t > \pi^{\S} \end{cases}$$

$$\chi(\pi_t) = \begin{cases} \dfrac{cr}{\mu(\phi - V_a(\pi_t))} & \text{when } \pi_t \leq \pi^{\S} \\ 1 & \text{when } \pi_t > \pi^{\S} \end{cases}$$

The limiting value problem in (2.13) and the initial value problem in (B.21) pin down the agent's on-path continuation value. Once it reaches $\overline{V}_a$ at $\pi^*(\pi^{\S})$, then the proofs for uniqueness of on path effort and disclosure rate are exactly the same as in the proof of Lemma 2.3.7.

The last step verifies the players' incentives in Phase III. The agent's incentive to choose $a(\pi_t) = a^*(\pi_t)$ when $\pi_t \in (\pi^{\S}, \pi^{**}]$ and the intermediary's incentive when $\pi_t \leq \min\{1-b, \pi^{**}\}$ is straightforward. To verify the agent's incentive to choose $a(\pi_t) = \phi$ when $\pi_t > \pi^{**}$, we compute his value function:

$$V_a(\pi_t) = \pi_t a(\pi_t) + c\phi + \hat{a}(\pi_t)\Big(\frac{\mu}{r}(\phi - V_a(\pi_t)) - c\Big) - \frac{\mu a(\pi_t)}{r}\pi_t(1-\pi_t)V_a'(\pi_t), \tag{B.32}$$

where $a(\pi_t)$ and $\hat{a}(\pi_t)$ are on-path and off-path effort, respectively. When $\pi_t = \pi^{**}$, we have $V_a(\pi_t) = \overline{V}_a$, which gives:

$$\overline{V}_a = \pi_t a(\pi_t) + c\phi - \frac{\mu a(\pi_t)}{r}\pi_t(1-\pi_t)V_a'(\pi_t),$$

Since $a(\pi_t) \geq a^{**}(\pi_t)$ for all π_t, this implies that $V_a'(\pi_t) > 0$ and the agent has an incentive to shirk at belief higher than π^{**}.

Appendix C

Appendix to Chapter 3

C.1 Proof of Theorem 3.1: Equilibrium Construction

In this Appendix, I show that $V^* \subset \underline{V}(\pi_0)$ by constructing Sequential Equilibria that can attain any payoff vector in the interior of V^* when δ is sufficiently high. For every $\gamma \in [0,1]$, let

$$v_i(\gamma) \equiv (1-\gamma\theta_i)\frac{1-\theta_1}{1-\gamma\theta_1}, \tag{C.1}$$

and let $v(\gamma) \equiv \Big(v_i(\gamma)\Big)_{1\leq i\leq m}$. Let $V(\gamma)$ be the triangular set with vertices $(0,0,...,0)$, $(1-\theta_1,...,1-\theta_m)$ and $v(\gamma)$. By definition, for every $1 \leq i \leq m$, $v_i(\gamma^*) = v_i^*$, $v_i(1) = 1-\theta_i$ and $V^* = V(\gamma^*)$. To establish $V^* \subset \underline{V}(\pi_0)$, it is sufficient to show that for every $\gamma > \gamma^*$, every payoff vector in $V(\gamma)$ is attainable in sequential equilibrium when player 1 is arbitrarily patient. As $(0,0,...,0)$ and $(1-\theta_1,...,1-\theta_m)$ are trivially attainable when δ is high enough and players have access to a public randomization device,[1] it is sufficient to establish the attainability of $v(\gamma)$.

Proposition C.1. *For every $\overline{\eta} \in (0,1)$ and $\gamma \in (\gamma^*,1)$, there exists $\overline{\delta} \in (0,1)$, such that for every $\delta > \overline{\delta}$ and $\pi_0 \in \Delta(\Theta)$ with $\pi_0(\theta_1) \geq \overline{\eta}$, there exists a sequential equilibrium in which player 1's payoff is $v(\gamma)$.*

The proof of Proposition C.1 is decomposed into three subsections. In Subsection C.1.1, I define several variables that are key to my construction. In Subsection C.1.2, I describe players' strategies and belief systems. In Subsection C.1.3, I verify players' incentive constraints and the consistency of their beliefs.

C.1.1 Defining the Variables

In this subsection, I define several variables that are critical for my construction. I will also specify how large $\overline{\delta}$ needs to be for every pair of (π_0, γ).

[1] Payoff vector $(0,0,...,0)$ is attainable by repeating the stage game equilibrium. Payoff vector $(1-\theta_1,...,1-\theta_m)$ is attainable by grim-trigger strategy, namely, player 2 plays N forever after observing L. For every $\gamma' \in [\gamma,1]$, $v(\gamma')$ can be written as a convex combination of $v(1)$ and $v(\gamma)$.

Fixing $\gamma \in (\gamma^*, 1)$, there exists a rational number $\widehat{n}/\widehat{k} \in (\gamma^*, \gamma)$ with $\widehat{n}, \widehat{k} \in \mathbb{N}$. Moreover, there exists an integer $j \in \mathbb{N}$ such that

$$\frac{\widehat{n}}{\widehat{k}} = \frac{\widehat{n}j}{\widehat{k}j} < \frac{\widehat{n}j}{\widehat{k}j - 1} < \gamma.$$

Let $n \equiv \widehat{n}j$ and $k \equiv \widehat{k}j$. Let

$$\widetilde{\gamma} \equiv \frac{1}{2}\Big(\frac{n}{k} + \frac{n}{k-1}\Big), \tag{C.2}$$

and

$$\widehat{\gamma} \equiv \frac{1}{2}\Big(\frac{n}{k} + \gamma^*\Big). \tag{C.3}$$

Let $\overline{\delta}_1 \in (0, 1)$ to be large enough such that for every $\delta > \overline{\delta}_1$,

$$\frac{\delta + \delta^2 + ... + \delta^n}{\delta + \delta^2 + ... + \delta^k} < \widetilde{\gamma} < \frac{\delta^{k-n-1}(\delta + \delta^2 + ... + \delta^n)}{\delta + \delta^2 + ... + \delta^{k-1}}. \tag{C.4}$$

By construction, $\gamma^* < \widehat{\gamma} < \frac{n}{k} < \widetilde{\gamma} < \frac{n}{k-1} < \gamma$. Let $\eta(h^0) \equiv \pi_0(\theta_1)$, which is the probability of type θ_1 according to player 2s' prior belief. Let η^* be an arbitrary real number satisfying:

$$\eta^* \in \Big(\gamma^*\eta(h^0), \eta(h^0)\Big).$$

Let $\lambda > 0$ be small enough such that:

$$\Big(1 + \lambda(1 - \gamma^*)\Big)^{\widehat{\gamma}}\Big(1 - \lambda\gamma^*\Big)^{1-\widehat{\gamma}} > 1. \tag{C.5}$$

Given $\gamma^* < \widehat{\gamma}$, the existence of such λ is implied by the Taylor's Expansion Theorem. Let $X \in \mathbb{N}$ be a large enough integer such that

$$\Big(1 + \lambda(1 - \gamma^*)\Big)^{X-1} > \frac{1 - \eta^*}{\eta(h^0) - \eta^*}. \tag{C.6}$$

Let

$$Y \equiv \frac{1}{2}\underbrace{\Big(\gamma - (1 - \gamma)\frac{\widetilde{\gamma}}{1 - \widetilde{\gamma}}\Big)}_{>0}\frac{1 - \theta_1}{1 - \gamma\theta_1}, \tag{C.7}$$

which is strictly positive. Let $\overline{\delta}_2 \in (0, 1)$ be large enough such that for every $\delta > \overline{\delta}_2$,

$$Y > \max\Big\{1 - \delta^X, \frac{1 - \delta}{1 - \gamma}\Big\} \text{ and } \frac{\delta - \theta_1}{1 - \theta_1} > \frac{1 - \delta}{1 - \gamma}. \tag{C.8}$$

The existence of such $\overline{\delta}_2$ is implied by $\widetilde{\gamma} < \gamma$.

Let $\overline{\delta} \equiv \max\{\overline{\delta}_1, \overline{\delta}_2\}$, which will be referred to as the *cutoff discount factor*. Let v^L, v^H and $v^N \in \mathbb{R}^m$ be player 1's payoff vectors from terminal outcomes L, H and N, respectively. The target payoff vector $v(\gamma)$ can be written as the following convex combination of v^L, v^H and v^N:

$$v(\gamma) = \underbrace{\frac{\theta_1(1-\gamma)}{1-\gamma\theta_1}}_{\equiv p^N} v^N + \underbrace{\frac{(1-\theta_1)\gamma}{1-\gamma\theta_1}}_{\equiv p^H} v^H + \underbrace{\frac{(1-\theta_1)(1-\gamma)}{1-\gamma\theta_1}}_{\equiv p^L} v^L, \tag{C.9}$$

with p^N, p^H and p^L being the convex weights of outcomes N, H and L, respectively.

Importantly, for every $\overline{\delta}$ that meets the above requirements under $\eta(h^0)$, it also meets all the requirements under every $\eta'(h^0) \geq \eta(h^0)$. This is because the required X decreases with $\eta(h^0)$, so an increase in $\eta(h^0)$ only slackens inequality (C.8) while having no impact on other requirements.

C.1.2 Three Phase Equilibrium

In this subsection, I describe players' strategies and player 2s' belief system. Players' sequential rationality constraints and the consistency of their beliefs are verified in the next step. Every type other than type θ_1 follows the same strategy, which is called *high cost types*, while type θ_1 is called the *low cost type*. Let $\eta(h^t)$ be the probability player 2s' posterior belief at h^t attaches to type θ_1. Recall the definition of η^*, which I will refer to as the *belief lower bound*. Let

$$\Delta(h^t) \equiv \eta(h^t) - \eta^*, \tag{C.10}$$

which is the gap between player 2s' posterior belief and the belief lower bound.

State Variables: The equilibrium keeps track of the following set of state variables: $\Delta(h^t)$ as well as $p^a(h^t)$ for $a \in \{N, H, L\}$ such that

$$p^a(h^0) = p^a \text{ and } p^a(h^t) \equiv \begin{cases} p^a(h^{t-1}) & \text{if } h^t \neq (h^{t-1}, a) \\ p^a(h^{t-1}) - (1-\delta)\delta^t & \text{if } h^t = (h^{t-1}, a). \end{cases} \tag{C.11}$$

Intuitively, $p^a(h^t)$ is the remaining occupation measure of outcome a at history h^t, while $p^a(h^0) - p^a(h^t)$ is the occupation measure of a from period 0 to $t-1$. Player 1's continuation value at h^t is

$$v(h^t) \equiv \delta^{-t} \sum_{a \in \{N,H,L\}} p^a(h^t) v^a. \tag{C.12}$$

Equilibrium Phases: The constructed equilibrium consists of three phases: a *normal phase*, an *absorbing phase* and a *reshuffling phase*.

Play starts from the *normal phase*, in which player 2 always plays T. Every type of player 1's mixed strategy at every history can be uniquely pinned down by player 2's belief updating process:

$$\Delta(h^t, L) = (1-\lambda\gamma^*)\Delta(h^t) \quad \text{and} \quad \Delta(h^t, H) = \min\Big\{1-\eta^*, \Big(1+\lambda(1-\gamma^*)\Big)\Delta(h^t)\Big\}. \tag{C.13}$$

Since $\eta(h^0) > \eta^*$, we know that $\Delta(h^t) > 0$ for every h^t in the normal phase.

Play transits to the *absorbing phase* permanently when $\Delta(h^t)$ reaches $1-\eta^*$ for the first time. Recall that $v(h^t) \in \mathbb{R}^m$ is player 1's continuation value at h^t. Let $v_i(h^t)$ be the projection of $v(h^t)$ on the i-th dimension. After reaching the absorbing phase, player 2s' learning stops and the continuation outcome is either (T,H) in all subsequent periods or N in all subsequent periods, depending on the realization of a public randomization device, with the probability of (T,H) being $v_1(h^t)/(1-\theta_1)$.

Play transits to the *reshuffling phase* at h^t if $\Delta(h^t) < 1-\eta^*$ and $p^L(h^t) \in [0,(1-\delta)\delta^t)$.

1. If $p^L(h^t)=0$, then the continuation play starting from h^t randomizes between N and (T,H), depending on the realization of the public randomization device, with the probability of (T,H) being $\frac{v_1(h^t)}{1-\theta_1}$.

2. If $p^L(h^t) \in (0,(1-\delta)\delta^t)$, then the continuation payoff vector can be written as a convex combination of v^H, v^N and

$$(1-\delta)v^L + \widetilde{Q}v^H + (\delta-\widetilde{Q})v^N, \tag{C.14}$$

for some

$$\widetilde{Q} \in \Big[\min\{Y, \frac{\delta-\theta_1}{1-\theta_1}\}, \frac{\delta-\theta_1}{1-\theta_1}\Big]$$

and Y being defined in (C.7). I will show in the next subsection that $\widetilde{Q}$ indeed belongs to this range for every history reaching the reshuffling phase.

If player 1's realized continuation value at h^t takes the form in (C.14), then player 2 plays T at h^t, type θ_1 player 1 plays H for sure while other types mix between H and L with the same probabilities (could be degenerate) such that:

$$\Delta(h^t,L) = -\eta^* \text{ and } \Delta(h^t,H) = \begin{cases} \Delta(h^0) & \text{if } \Delta(h^t) \leq \Delta(h^0) \\ \Delta(h^t) & \text{if } \Delta(h^t) > \Delta(h^0). \end{cases} \tag{C.15}$$

If player 2 observes L at h^t, then he attaches probability 0 to type θ_1 and player 1's continuation value is

$$\delta^{-1}\widetilde{Q}v^H + \delta^{-1}(\delta-\widetilde{Q})v^N, \tag{C.16}$$

which can be delivered by randomizing between outcomes (T,H) and N, with probabilities $\delta^{-1}\widetilde{Q}$ and $1-\delta^{-1}\widetilde{Q}$, respectively.

If player 2 observes H at h^t, then he attaches probability $\Delta(h^t,H)+\eta^*$ to type θ_1 and player 1's continuation value is:

$$\frac{1-\delta}{\delta}v^L + \frac{\widetilde{Q}-(1-\delta)}{\delta}v^H + \frac{\delta-\widetilde{Q}}{\delta}v^N, \tag{C.17}$$

which can be written as a convex combination of v^N and

$$v\Big(1-\frac{1-\delta}{\widetilde{Q}}\Big). \tag{C.18}$$

According to (C.8) and the range of $\widetilde{Q}$,

$$\gamma < 1-\frac{1-\delta}{\widetilde{Q}} < 1, \tag{C.19}$$

which implies that (C.18) can further be written as a convex combination of v^H and $v(\gamma)$.

If the continuation value is v^H or v^N, then the on-path outcome is (T,H) in every subsequent period or is N in every subsequent period. If the continuation value is $v(\gamma)$, then play switches back to the normal phase with belief $\max\{\Delta(h^0),\Delta(h^t)\}$, which is no less than $\Delta(h^0)$.

C.1.3 Verifying Constraints

In this subsection, I verify that the strategy profile and the belief system indeed constitute a Sequential Equilibrium by verifying players' sequential rationality constraints and the consistency of beliefs. This consists of two parts. In Part I, I verify player 2's incentive constraints. In Part II, I verify the range of $\widetilde{Q}$ in Subsection C.1.2. In particular, at every history of the normal phase or reshuffling phase, the ratio between the occupation measure of H and the occupation measure of L must exceed some cutoff.

Part I: Player 2's incentive constraints consist of two parts: the normal phase and the reshuffling phase. If play remains in the normal phase at h^t, then (C.13) implies that the unconditional probability with which H being played is at least γ^*, implying that player 2 has an incentive to play T. If play reaches the reshuffling phase at h^t and at this history, player 1 is playing a non-trivial mixed action, then according to (C.15) and the requirement that $\eta^* > \gamma^*\eta(h^0)$, the unconditional probability with which H is played is at least γ^*. This verifies player 2's incentives to play T.

Part II: In this part, I establish bounds on player 1's continuation value at every history in the normal phase or in the beginning of the reshuffling phase. In particular, I establish a lower bound on the ratio between the convex weight of H and the convex weight of L at such histories, or equivalently, a lower bound on the depleted occupation measure of H and the depleted occupation measure of L. Recall the definitions of n and k in Subsection C.1.1. The conclusion is summarized in the following Lemma:

Lemma C.1.1. *If $\delta > \overline{\delta}$ and $T \geq k+X$, then for every $h^T=(a_0,...,a_{T-1})$, if play remains in the* *normal phase for every $h^t \preceq h^T$, then*

$$\underbrace{(1-\delta)\sum_{t=0}^{T-1}\delta^t \mathbf{1}\{a_t = H\}}_{\textit{depleted occupation measure of } H} - \underbrace{(1-\delta^X)}_{\textit{weight of initial } X \textit{ periods}} \leq \underbrace{(1-\delta)\sum_{t=0}^{T-1}\delta^t \mathbf{1}\{a_t = L\}}_{\textit{depleted occupation measure of } L} \cdot \underbrace{\frac{\widetilde{\gamma}}{1-\widetilde{\gamma}}}_{\textit{multiplier}}. \quad \text{(C.20)}$$

Lemma C.1.1 implies that when play first reaches the reshuffling phase, the remaining occupation measure of H is at least $\widetilde{Q}$. This implies that player 1's continuation value after reshuffling also attaches sufficiently high convex weight on v^H compared to the convex weight of v^L. Using the self-generation argument in Abreu, Pearce and Stacchetti (1990), Chari and Kehoe (1990), one can conclude that payoff vector $v(\gamma)$ is attainable in sequential equilibrium when $\delta > \overline{\delta}$.

The key difficulty to prove this Lemma is that different periods have different weights due to discounting. To see this, for every fixed $\delta > \overline{\delta}$, Lemma C.1.1 requires a uniform bound for every $T \in \mathbb{N}$, including those that are arbitrarily large. Therefore, one could potentially increase the depleted occupation measure of H and decrease that of L by front-loading the play of H, making inequality (C.20) harder to satisfy. Nevertheless, the condition that play never reaching the absorbing phase before h^T constrains on how front-loaded the play of H can be.

Proof of Lemma C.1.1: For every $t \in \mathbb{N}$, let $N_{L,t}$ and $N_{H,t}$ be the number of periods in which L and H are played from period 0 to $t-1$, respectively. The proof is done by induction on $N_{L,t}$.

When $N_{L,t} \leq k-n$, then the conclusion holds as $N_{H,t} \geq n+X$. According to (C.5) and (C.6), we know that $\Delta(h^T)$ will reach $1-\eta^*$ before period T (or equivalently, play reaches the absorbing phase).

Suppose the conclusion holds for when $N_{L,t} \leq N$ with $N \geq k-n$, and suppose towards a contradiction that there exists h^T with $T \geq k+X$ and $N_{L,T} = N+1$, such that play remains in the normal phase for every $h^t \preceq h^T$ but

$$(1-\delta)\sum_{t=0}^{T-1}\delta^t \mathbf{1}\{a_t = H\} - (1-\delta^X) > (1-\delta)\sum_{t=0}^{T-1}\delta^t \mathbf{1}\{a_t = L\} \cdot \frac{\widetilde{\gamma}}{1-\widetilde{\gamma}}, \quad \text{(C.21)}$$

I will obtain a contradiction in three steps.

Step 1: I show that for every $s < T$,

$$(1-\delta)\sum_{t=s}^{T-1}\delta^t \mathbf{1}\{a_t = H\} \geq (1-\delta)\sum_{t=s}^{T-1}\delta^t \mathbf{1}\{a_t = L\}\frac{\widetilde{\gamma}}{1-\widetilde{\gamma}}. \quad \text{(C.22)}$$

Suppose towards a contradiction that the opposite of (C.22) holds, then (C.22) and (C.21) together imply that:

$$(1-\delta)\sum_{t=0}^{s-1}\delta^t \mathbf{1}\{a_t = H\} - (1-\delta^X) > (1-\delta)\sum_{t=0}^{s-1}\delta^t \mathbf{1}\{a_t = L\}\frac{\widetilde{\gamma}}{1-\widetilde{\gamma}} \quad \text{(C.23)}$$

and

$$(1-\delta)\sum_{t=s}^{T-1}\delta^t\mathbf{1}\{a_t=L\}>0. \tag{C.24}$$

According to (C.24), $N_{L,s} < N_{L,T}$. Since $N_{L,T} = N+1$, we have $N_{L,s} \leq N$. Applying the induction hypothesis and (C.23), we know that play reaches the absorbing phase before h^s, leading to a contradiction.

Step 2: I show that for every k consecutive periods

$$\{a_r, ..., a_{r+k-1}\} \subset h^T,$$

the number of H in this sequence is at least $n+1$. According to (C.22) shown in the previous step and (C.4), H occurs at least $n+1$ times in the last k periods, i.e. $\{a_{T-k+1}, ..., a_T\}$.

Suppose towards a contradiction that there exists k consecutive periods in which H occurs no more than n times, then the conclusion above that H occurs at least $n+1$ times in the last k periods implies that there exists k consecutive periods $\{a_r, ..., a_{r+k-1}\}$ in which H occurs exactly n times and L occurs exactly $k-n$ times. According to (C.4), we have

$$(1-\delta)\sum_{t=r}^{r+k-1}\delta^t\mathbf{1}\{a_t=H\}<(1-\delta)\sum_{t=r}^{r+k-1}\delta^t\mathbf{1}\{a_t=L\}\frac{\widetilde{\gamma}}{1-\widetilde{\gamma}}, \tag{C.25}$$

but according to (C.5) and the definition of $\widehat{\gamma}$ in (C.3), we also know that

$$\Delta(h^{r+k}) > \Delta(h^{r+1}). \tag{C.26}$$

Next, let us consider the following new sequence with length $T-k$:

$$\widetilde{h}^{T-k} \equiv \{\widetilde{a}_0, \widetilde{a}_1, ..., \widetilde{a}_{T-k-1}\} \equiv \{a_0, a_1, ..., a_{r-1}, a_{r+k}, ..., a_{T-1}\}$$

which is obtained by removing $\{a_r, ..., a_{r+k-1}\}$ from the original sequence and front-loading the subsequent play $\{a_{r+k}, ..., a_{T-1}\}$. The number of L in this new sequence is at most $N+1-(n-k)$, which is no more than N. According to the conclusion in Step 1:

$$(1-\delta)\sum_{t=r+k}^{T-1}\delta^t\mathbf{1}\{a_t=H\}>(1-\delta)\sum_{t=r+k}^{T-1}\delta^t\mathbf{1}\{a_t=L\}\frac{\widetilde{\gamma}}{1-\widetilde{\gamma}}. \tag{C.27}$$

This together with (C.25) and (C.21) imply that

$$(1-\delta)\sum_{t=0}^{T-k-1}\delta^t\mathbf{1}\{\widetilde{a}_t=H\}-(1-\delta^X)>(1-\delta)\sum_{t=0}^{T-k-1}\delta^t\mathbf{1}\{\widetilde{a}_t=L\}\frac{\widetilde{\gamma}}{1-\widetilde{\gamma}}.$$

According to the induction hypothesis, play will reach the absorbing phase before period $T-k$ if player 1 plays according to $\{\widetilde{a}_0, \widetilde{a}_1, ..., \widetilde{a}_{T-k-1}\}$.

1. Suppose $\widetilde{h}^{T-k}$ reaches the absorbing phase before period r, then play will also reach the absorbing phase before period r according to the original sequence.

2. Suppose $\widetilde{h}^{T-k}$ reaches the absorbing phase in period s, with $s > t$, then according to (C.26), we have $\Delta(\widetilde{h}^s) \leq \Delta(h^{s+k})$, implying that play will reach the absorbing phase in period $s+k$ according to the original sequence.

This contradicts the hypothesis that play has never reached the absorbing phase before period T if play proceeds according to h^T.

Step 3: For every history $h^T \equiv \{a_0, a_1, ..., a_{T-1}\} \in \{H, L\}^T$ and $t \in \{1, ..., T-1\}$, define the operator $\Omega_t : \{H, L\}^T \to \{H, L\}^T$ as:

$$\Omega_t(h^T) = (a_0, ..., a_{t-2}, a_t, a_{t-1}, a_{t+1}, ..., a_{T-1}), \tag{C.28}$$

in another word, swapping the order between a_{t-1} and a_t. Recall the belief updating formula in (C.13) and let

$$\mathcal{H}^{T,*} \equiv \Big\{h^T \Big| \Delta(h^t) < 1-\eta^* \text{ for all } h^t \prec h^T\Big\}. \tag{C.29}$$

If $h^T \in \mathcal{H}^{T,*}$, then $\Omega_t(h^T) \in \mathcal{H}^{T,*}$ unless:

- $a_{t-1} = L$, $a_t = H$.
- *and*, $\Big(1+\lambda(1-\gamma^*)\Big)\Delta(h^{t-1}) \geq 1-\eta^*$.

Next, I show that the above situation cannot occur besides in the last k periods. Suppose towards a contradiction that there exists $t \leq T-k$ such that $h^T \in \mathcal{H}^{T,*}$ but $\Omega_t(h^T) \notin \mathcal{H}^{T,*}$. Then according to the conclusion in step 2, H occurs at least $n+1$ times in $\{a_t, ..., a_{t+k-1}\}$. Now, consider the sequence $\{a_{t-1}, ..., a_{t+k-1}\}$, in which H occurs at least $n+1$ times and L occurs at most $k-n$ times. This implies that:

$$\begin{aligned}
\Delta(h^{t+k}) &\geq \Delta(h^{t-1})\Big(1+\lambda(1-\gamma^*)\Big)^{n+1}\Big(1-\lambda\gamma^*\Big)^{k-n} \\
&= \Delta(h^{t-1})\underbrace{\Big(1+\lambda(1-\gamma^*)\Big)^{n}\Big(1-\lambda\gamma^*\Big)^{k-n}}_{\geq 1}\Big(1+\lambda(1-\gamma^*)\Big) \\
&\geq \Delta(h^{t-1})\Big(1+\lambda(1-\gamma^*)\Big) \\
&\geq 1-\eta^*,
\end{aligned} \tag{C.30}$$

where 2nd inequality follows from $n/k > \widehat{\gamma}$ and (C.5), and the 3rd inequality follows from the hypothesis that $\Omega_t(h^T) \notin \mathcal{H}^{T,*}$. Inequality (C.30) implies that play reaches the high phase before period $t+k \leq T$, contradicting the hypothesis that $h^T \in \mathcal{H}^{T,*}$.

To conclude, for every $t \leq T-k$, if $h^T \in \mathcal{H}^{T,*}$, then $\Omega_t(h^T) \in \mathcal{H}^{T,*}$. For every $t > T-k$, if $h^T \in \mathcal{H}^{T,*}$, then $\Omega_t(h^T) \in \mathcal{H}^{T,*}$ unless $a_{t-1} = L$ and $a_t = H$. Intuitively, this is to say that one can freely swap the order of play before period $T-k$ and can also freely front-load the play the L after period $T-k$.

Let $l \in \mathbb{N}$ be the number of L in the sequence $\{a_{T-k}, ..., a_{T-1}\}$. If $h^T \in \mathcal{H}^{T,*}$, then the following sequence that can be obtained by applying $\{\Omega_t | t > T-k\}$ (only to $a_{t-1} = H$ and $a_t = L$) also belongs to $\mathcal{H}^{T,*}$:

$$\{a_0, ..., a_{T-k-1}, \underbrace{L, L, ..., L}_{\#L=l}, \underbrace{H, H, ...H}_{\#H=k-l}\}. \tag{C.31}$$

Applying $\{\Omega_t | t \leq T-k\}$ to (C.31), we know that the following sequence that can be obtained through these operations also belong to $\mathcal{H}^{T,*}$:

$$\{... \underbrace{L, L, ..., L}_{\#L=k-n}, \underbrace{H, H, ...H}_{\#H=k-l}\}. \tag{C.32}$$

The sequence in (C.32) is feasible due to the induction hypothesis that $N_{L,T} \geq n-k$. The conclusion that (C.32) belongs to $\mathcal{H}^{T,*}$ contradicts the conclusion in step 2, as there exists a sequence with length k in which H occurs at most n times. This contradiction implies that (C.21) cannot be true for any history h^T s.t. $N_{L,T} = N+1$ and play remains in the normal phase for every $h^t \preceq h^T$. This proves Lemma C.1.1. □

C.2 Proof of Theorem 3.1: Payoff Upper Bound

In this Appendix, I show that $\overline{V}(\pi_0) \subset V^*$. In Subsection C.2.1, I establish a payoff upper bound for the lowest cost type that uniformly applies across all discount factors. In Subsection C.2.2, I establish a payoff upper bound for other types that applies in the $\delta \to 1$ limit. To accommodate applications where players move simultaneously, I prove the result under the following stage game:

$\theta = \theta_i$	T	N
H	$1-\theta_i, b$	$-d(\theta_i), 0$
L	$1, -c$	$0, 0$

I assume that players' payoffs are monotone-supermodular (Liu and Pei 2017). In the context of this game, once we rank the states and players' actions according to $\theta_1 \succ \theta_2 \succ ... \succ \theta_m$, $H \succ L$ and $T \succ N$, monotone-supermodularity implies that $d(\theta_i) \geq 0$ for every $\theta_i \in \Theta$ and $|\theta_i - \theta_j| \geq |d(\theta_i) - d(\theta_j)|$ for every $i < j$.

Having players move simultaneously in the stage game and letting future short-run players observing their predecessors' actions introduce new challenges. As discussed in Pei (2017), the predecessor-successor relationship is incomplete on the set of histories where player 1 has always played H. Nevertheless, the adaptation of my proof to sequential move stage games in the baseline model is straightforward.

C.2.1 Payoff Upper Bound for the Lowest Cost Type

I start with recursively defining the set of *high histories*. Let $\overline{\mathcal{H}}^0 \equiv \{h^0\}$ and

$$\overline{a}_1(h^0) \equiv \max\Big\{\bigcup_{\theta\in\Theta} \text{supp}\Big(\sigma_\theta(h^0)\Big)\Big\}.$$

Let

$$\overline{\mathcal{H}}^1 \equiv \{h^1 | \exists h^0 \in \overline{\mathcal{H}}^0 \text{ s.t. } h^1 \succ h^0 \text{ and } \overline{a}_1(h^0) \in h^1\}.$$

For every $t \in \mathbb{N}$ and $h^t \in \overline{\mathcal{H}}^t$, let $\Theta(h^t) \subset \Theta$ be the set of types that occur with positive probability at h^t. Let

$$\overline{a}_1(h^t) \equiv \max\Big\{\bigcup_{\theta\in\Theta(h^t)} \text{supp}\Big(\sigma_\theta(h^t)\Big)\Big\} \tag{C.33}$$

and

$$\overline{\mathcal{H}}^{t+1} \equiv \{h^{t+1} | \exists h^t \in \overline{\mathcal{H}}^t \text{ s.t. } h^{t+1} \succ h^t \text{ and } \overline{a}_1(h^t) \in h^{t+1}\}. \tag{C.34}$$

Let $\overline{\mathcal{H}} \equiv \bigcup_{t=0}^{\infty} \overline{\mathcal{H}}^t$ be the set of high histories. The main result in this subsection is the following Proposition, which shows that at every history, the lowest cost type in the support of player 2s' posterior belief cannot receive a continuation payoff higher than her pure Stackelberg commitment payoff.

Proposition C.2. *For every $h^t \in \overline{\mathcal{H}}$, if $\theta_i = \min \Theta(h^t)$, then type θ_i's continuation payoff at h^t is no more than $1-\theta_i$ in any Nash Equilibrium.*

Since $h^0 \in \overline{\mathcal{H}}$ and $\theta_1 = \min \Theta(h^0)$, a corollary of Proposition C.2 is that type θ_1's payoff cannot exceed $1-\theta_1$ in any Nash Equilibrium.

Proof of Proposition C.2: For every $\theta \in \Theta$, let $\overline{\mathcal{H}}(\theta)$ be a subset of $\overline{\mathcal{H}}$ (could be empty) such that $h^t \in \overline{\mathcal{H}}(\theta)$ if and only if both of the following conditions are satisfied:

1. For every $h^s \succeq h^t$ with $h^s \in \overline{\mathcal{H}}$, we have $\theta \in \Theta(h^s)$.

2. If $h^{t-1} \prec h^t$, then for every $\tilde{\theta} \in \Theta(h^{t-1})$, there exists $h^s \in \overline{\mathcal{H}}$ with $h^s \succ h^{t-1}$ such that $\tilde{\theta} \notin \Theta(h^s)$.

Let $\overline{\mathcal{H}}(\Theta) \equiv \bigcup_{\theta\in\Theta} \overline{\mathcal{H}}(\theta)$. By definition, $\overline{\mathcal{H}}(\Theta)$ possesses the following two properties:

1. $\overline{\mathcal{H}}(\Theta) \subset \overline{\mathcal{H}}$.

2. For every $h^t, h^s \in \overline{\mathcal{H}}(\Theta)$, neither $h^t \succ h^s$ nor $h^t \prec h^s$.

For every $h^t \in \overline{\mathcal{H}}(\theta_i)$, at the subgame starting from h^t, type θ_i's stage game payoff is no more than $1-\theta_i$ in every period if she plays $\overline{a}_1(h^s)$ for every $h^s \succsim h^t$ and $h^s \in \overline{\mathcal{H}}$. Since $h^t \in \overline{\mathcal{H}}(\theta_i)$ implies that doing so is optimal for type θ_i, her continuation payoff at h^t cannot exceed $1-\theta_i$. When the stage game payoff is supermodular, for every $j < i$, the payoff difference between type θ_j and type

θ_i in any period is at most $|\theta_i - \theta_j|$. This implies that for every $\theta_j \in \Theta(h^t)$ with $\theta_j < \theta_i$, type θ_j's continuation payoff at h^t cannot exceed $1 - \theta_j$.

In what follows, I show Proposition C.2 by induction on $\#\Theta(h^t)$. When $\#\Theta(h^t) = 1$, i.e. there is only one type (call it type θ_i) that can reach h^t. The above argument implies that type θ_i's payoff cannot exceed $1 - \theta_i$.

Suppose the conclusion in Proposition C.2 holds for every $\#\Theta(h^t) \leq n$, consider the case when $\#\Theta(h^t) = n + 1$. Let $\theta_i \equiv \min \Theta(h^t)$. Next, I introduce the definition of set $\overline{\mathcal{H}}^B(h^t)$: For every $h^s \succeq h^t$ with $h^s \in \overline{\mathcal{H}}$, $h^s \in \overline{\mathcal{H}}^B(h^t)$ if and only if:

- $h^s \in \overline{\mathcal{H}}(\theta_i)$, but $h^{s+1} \notin \overline{\mathcal{H}}(\theta_i)$ for any $h^{s+1} \succ h^s$ with $h^{s+1} \in \overline{\mathcal{H}}$.

In another word, type θ_i has a strict incentive not to play $\overline{a}_1(h^s)$ at h^s. A useful property of $\overline{\mathcal{H}}^B(h^t)$ is:

- For every $h^\infty \in \overline{\mathcal{H}}$ with $h^\infty \succ h^t$, either there exists $h^s \in \overline{\mathcal{H}}^B(h^t)$ such that $h^s \prec h^\infty$, or there exists $h^s \in \overline{\mathcal{H}}(\theta_i)$ such that $h^s \prec h^\infty$.

which means that play will eventually reach either a history in $\overline{\mathcal{H}}^B(h^t) \bigcup \overline{\mathcal{H}}(\theta_i)$ if type θ plays $\overline{a}_1(h^\tau)$ before that for every $t \leq \tau \leq s$. In what follows, I examine type θ_i's continuation value.

1. For every $h^s \in \overline{\mathcal{H}}^B(h^t)$, at every h^{s+1} satisfying $h^{s+1} \succ h^s$ and $h^{s+1} \in \overline{\mathcal{H}}$, we have:

$$\#\Theta(h^{s+1}) \leq n.$$

   Let $\theta_j \equiv \min \Theta(h^{s+1})$. According to the induction hypothesis, type θ_j's continuation payoff at h^{s+1} is at most $1 - \theta_j$. Since this applies to every such h^{s+1}, type θ_j's continuation value at h^s also cannot exceed $1 - \theta_j$ since she is playing $\overline{a}_1(h^s)$ with positive probability at h^s, and her stage game payoff from doing so is at most $1 - \theta_j$. Therefore, type θ_i's continuation value at h^s is at most $1 - \theta_i$.

2. For every $h^s \in \overline{\mathcal{H}}(\theta_i)$, playing $\overline{a}_1(h^\tau)$ for all $h^\tau \succeq h^s$ and $h^\tau \in \overline{\mathcal{H}}$ is a best reply for type θ_i. Her stage game payoff from this strategy cannot exceed $1 - \theta_i$, which implies that her continuation value at h^s also cannot exceed $1 - \theta_i$.

Starting from h^t, consider the strategy in which player 1 plays $\overline{a}_1(h^\tau)$ at every $h^\tau \succ h^t$ and $h^\tau \in \overline{\mathcal{H}}$ until play reaches $h^s \in \overline{\mathcal{H}}^B(h^t)$ or $h^s \in \overline{\mathcal{H}}(\theta_i)$. By construction, this is type θ_i's best reply. Under this strategy, type θ_i's stage game payoff cannot exceed $1 - \theta_i$ before reaches h^s. Moreover, her continuation payoff after reaching h^s is also bounded above by $1 - \theta_i$, which establishes the conclusion in Proposition C.2 when $\#\Theta(h^t) = n + 1$. □

C.2.2 Payoff Upper Bound for Other Types

I establish an upper bound on the limiting equilibrium payoffs for types other than the lowest cost type. For this purpose, I introduce an auxiliary maximization program. For every $i > j$, let

$$W_i(v_j) \equiv \max_{\{q^k,\alpha_1^k,\alpha_2^k\}_{k\in\mathbb{N}}} \sum_{k=1}^{\infty} q^k u_1(\theta_i, \alpha_1^k, \alpha_2^k), \tag{C.35}$$

subject to $\sum_{k=1}^{\infty} q^k = 1$, and $q^k \in [0,1]$,

$$\alpha_2^k \in \arg\max_{\alpha_2' \in \Delta(A_2)} u_2(\alpha_1^k, \alpha_2') \quad \text{and} \quad \sum_{k=1}^{\infty} q^k u_1(\theta_j, \alpha_1^k, \alpha_2^k) \leq v_j, \tag{C.36}$$

for evert $k \in \mathbb{N}$.

Intuitively, program (C.35) maximizes the high cost type θ_i's expected payoff by choosing a convex combination of action profiles (with weights given by q^k) subject to player 2's (myopic) incentive constraints and type θ_j's expected payoff being no more than v_j. For every $v_j \in [-d(\theta_j), 1-\gamma^*\theta_j]$, one can solve the above linear program and obtain:

$$W_i(v_j) \leq \frac{1-\gamma^* c_i}{1-\gamma^* c_j} v_j, \tag{C.37}$$

with equality holds when $v_j \in [0, 1-\gamma^*\theta_j]$.

For every strategy profile σ, let $v_i^\sigma(h^t)$ be the (discounted average) continuation payoff of type θ_i at history h^t and let $u_i^\sigma(h^t)$ be her stage game payoff. Let P^σ be the probability measure over $\mathcal{H}$ induced by σ. Let

$$D_{i,j}^\sigma(h^t) \equiv v_i^\sigma(h^t) - W_i(v_j^\sigma(h^t)). \tag{C.38}$$

Let $\text{NE}(\delta, h^t)$ be the set of equilibria in the continuation game starting from h^t when the discount factor is δ. The main result of this subsection is stated in the following Proposition:

Proposition C.3. *For every h^t, if $\theta_i, \theta_j \in \Theta(h^t)$ with $\theta_j = \min \Theta(h^t)$, then*

$$\limsup_{\delta \to 1} \sup_{\sigma \in NE(\delta, h^t)} D_{i,j}^\sigma(h^t) \leq 0. \tag{C.39}$$

Proof of Proposition C.3: Without loss of generality, I focus on continuation games that start with h^0. The proof is done by induction on $\#\Theta(h^t)$. I start with cases in which there are two types, with $\theta_i > \theta_j$, before moving on to cases in which there are more than two types.

Two Types: For every given equilibrium σ and history h^t with $P^\sigma(h^t) > 0$, let $\sigma(h^t, \theta_i) \in \Delta(A_1)$ be the action played by type θ_i at h^t and let $\sigma(h^t) \in \Delta(A_1)$ be player 1's weighted average action according to player 2s' posterior belief at h^t. Consider the following strategy $\hat{\sigma}_{\theta_i}$ for type θ_i: For every h^t with $P^\sigma(h^t) > 0$,

1. $\hat{\sigma}_{\theta_i}(h^t) = \sigma(h^t)$ if both types play every action in the support of $\sigma(h^t)$ with positive probability.

2. $\hat{\sigma}_{\theta_i}(h^t) = \sigma(h^t, \theta_i) \backslash \sigma(h^t, \theta_j)$ if it is non-empty.

3. $\hat{\sigma}_{\theta_i}(h^t) = \sigma(h^t, \theta_i)$ otherwise.

By construction, $\hat{\sigma}_i$ is a best response for type θ_i against the equilibrium strategy of player 2. In what follows, I consider the three cases one by one. In the 1st case, one can decompose $D^{\sigma}_{i,j}(h^t)$ according to:

$$D^{\sigma}_{i,j}(h^t) \leq (1-\delta)\Big(u_i^{\sigma}(h^t) - W_i(u_j^{\sigma}(h^t))\Big) + \delta q(h^t)\Big(v_i^{\sigma}(h^t, H) - W_i(v_j^{\sigma}(h^t, H))\Big) + \delta(1-q(h^t))\Big(v_i^{\sigma}(h^t, L) - W_i(v_j^{\sigma}(h^t, L))\Big) \tag{C.40}$$

where $q(h^t)$ is the probability that H is played according to $\sigma(h^t)$. Since $\sigma(h^t)$ is a best reply for both type θ_i and type θ_j, the term $u_i^{\sigma}(h^t) - W_i(u_j^{\sigma}(h^t))$ is non-positive according to the linear program (C.35) that defines W_i.

In the 2nd case, by playing $\sigma(h^t, \theta_i) \backslash \sigma(h^t, \theta_j)$, type θ_i will be separated from type θ_j in period $t+1$, and afterwards, Proposition C.2 implies that $D^{\sigma}_{i,j}(h^t, a_{1,t+1}) \leq 0$. Therefore,

$$D^{\sigma}_{i,j}(h^t) \leq \Big(1 + d(\theta_i)\Big)(1-\delta). \tag{C.41}$$

In the 3rd case, given that type θ_i plays $\sigma(h^t, \theta_i)$ and type θ_j plays $\sigma(h^t, \theta_j)$, we know that if type θ_i only plays H, then:

$$D^{\sigma}_{i,j}(h^t) \leq (1-\delta)\Big(u_i^{\sigma}(h^t) - W_i(u_j^{\sigma}(h^t))\Big) + \delta\Big(v_i^{\sigma}(h^t, H) - W_i(v_j^{\sigma}(h^t, H))\Big). \tag{C.42}$$

If type θ_i only plays L, then:

$$D^{\sigma}_{i,j}(h^t) \leq (1-\delta)\Big(u_i^{\sigma}(h^t) - W_i(u_j^{\sigma}(h^t))\Big) + \delta\Big(v_i^{\sigma}(h^t, L) - W_i(v_j^{\sigma}(h^t, L))\Big). \tag{C.43}$$

In the first subcase, $u_i^{\sigma}(h^t) - W_i(u_j^{\sigma}(h^t)) \leq 0$, and in the second subcase, $u_i^{\sigma}(h^t) - W_i(u_j^{\sigma}(h^t)) \leq 0$ unless T is played with positive probability.

In what follows, I provide a uniform upper bound (interdependent of δ) on the expected number of periods where type θ_i only plays L but player 2 has an incentive to play T. Let P^* be the probability measure over $\mathcal{H}$ induced by $\hat{\sigma}_{\theta_i}$ and σ_2. I establish a uniform upper bound on:

$$Q \equiv \mathbb{E}^{P^*}\Big[\mathbf{1}\Big\{h^t \in \mathcal{H}, \sigma_2(h^t)[T] > 0 \text{ and } \hat{\sigma}_i(h^t) = L\Big\}\Big]. \tag{C.44}$$

This bound is interesting since it can be applied to the previous part and obtain:

$$D^{\sigma}_{i,j}(h^t) \leq 2(1-\delta)\Big(1 + d(\theta_i)\Big)Q \tag{C.45}$$

for every $\sigma \in \text{NE}(h^t, \delta)$, which implies (C.39). For every $h^t \in \mathcal{H}$ with $P^*(h^t) > 0$, let $\eta(h^t)$ be the probability of type θ_j and let

$$\lambda(h^t) \equiv (1 - \eta(h^t))/\eta(h^t).$$

Under the probability measure induced by $(\hat{\sigma}_{\theta_i}, \sigma_2)$, for every h^t with $P^*(h^t) > 0$, we have:

$$\mathbb{E}[\lambda(h^{t+1})|h^t] \geq \lambda(h^t).$$

If $\sigma_2(h^t)[T] > 0$ and only type θ_j plays H at h^t, then according to player 2's incentive constraint, type θ_j plays H with probability at least $\gamma^*(1+\lambda(h^t))$. Therefore, if h^t occurs with probability $P^*(h^t)$, we have:

$$\mathbb{E}[\lambda(h^{t+1})] \geq \frac{\mathbb{E}[\lambda(h^t)]}{1-\gamma^* P^*(h^t)}$$

Since such periods can only occur when $\mathbb{E}[\lambda(h^t)] \leq 1/\gamma^* - 1$ for every $t \in \mathbb{N}$, we have:

$$Q \leq -\frac{1}{\gamma^*} \log \lambda(h^0). \tag{C.46}$$

Three or More Types: I use the result in the previous part as an induction hypothesis to establish the conclusion in Proposition C.3 when there are three or more types. For $n \geq 2$, suppose inequality (C.39) holds at every $\tilde{h}^t$ with $\#\Theta(\tilde{h}^t) \leq n$, then consider a history h^t with $\#\Theta(h^t) = n+1$. Let $\theta_j \equiv \min \Theta(h^t)$ and $\theta_i \in \Theta(h^t)$ with $i > j$. Let h^s be a generic history that succeeds h^t. By definition, we have $\Theta(h^s) \subset \Theta(h^t)$. The induction hypothesis implies that we only need to consider histories in which $\Theta(h^s) = \Theta(h^t)$, which we will be focusing on for the rest of the proof. Let $\sigma(h^s) \in \Delta(A_1)$ be the average action of player 1 at h^s according to σ.

If every type in $\Theta(h^s)$ plays every action on the support of $\sigma(h^s)$ with positive probability, then:

$$\begin{aligned} D^{\sigma}_{i,j}(h^s) &= (1-\delta)\Big(u^{\sigma}_i(h^s) - W_i(u^{\sigma}_j(h^s))\Big) + \delta q(h^s)\Big(v^{\sigma}_i(h^s,H) - W_i(v^{\sigma}_j(h^s,H))\Big) \\ &\quad + \delta(1-q(h^s))\Big(v^{\sigma}_i(h^s,L) - W_i(v^{\sigma}_j(h^s,L))\Big), \end{aligned} \tag{C.47}$$

where $q(h^s)$ is the probability of H at h^s according to σ. Using the same argument as before, we know that $u^{\sigma}_i(h^s) - W_i(u^{\sigma}_j(h^s)) \leq 0$.

If not every type in $\Theta(h^s)$ plays every action on the support of $\sigma(h^s)$ with positive probability, then consider two subcases. First, suppose there exists $a_1 \in A_1$ that is played with positive probability by type θ_i but there exists $\theta_k \in \Theta(h^s)$ such that type θ_k plays a_1 with probability 0, then by playing a_1 at h^s, $\#\Theta(h^s, a_1) \leq n$. The conclusion is established via the induction hypothesis. Second, suppose such action does not exist, then there exists an action on the support of $\sigma(h^s)$ in which $\sigma(h^s, \theta_i)$ attaches 0 probability. In this case, apply the inequality (C.46), we can obtain an upper bound on $D^{\sigma}_{i,j}(h^t)$, which diminishes as $\delta \to 1$.

□

Bibliography

[1] Abreu, Dilip, David Pearce and Ennio Stacchetti (1990) “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica*, 58(5), 1041-1063.

[2] Atakan, Alp and Mehmet Ekmekci (2012) “Reputation in Long-Run Relationships,” *Review of Economic Studies*, 79(2), 751-780.

[3] Aumann, Robert and Michael Maschler (1995) *Repeated Games with Incomplete Information*, MIT Press.

[4] Bai, Jie (2016) “Melons as Lemons: Asymmetric Information, Consumer Learning and Seller Reputation,” Working Paper, Harvard Kennedy School.

[5] Bain, Joe (1949) “A Note on Pricing in Monopoly and Oligopoly,” *American Economic Review*, 39(2), 448-464.

[6] Baker, George, Michael Gibbs and Bengt Holmström (1994a) “The Internal Economics of the Firm: Evidence from Personnel Data,” *Quarterly Journal of Economics*, 109(4), 881-919.

[7] Baker, George, Michael Gibbs and Bengt Holmström (1994b) “The Wage Policy of a Firm,” *Quarterly Journal of Economics*, 109(4), 920-955.

[8] Banerjee, Abhijit and Esther Duflo (2000) “Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry,” *Quarterly Journal of Economics*, 115(3), 989-1017.

[9] Bar-Isaac, Heski (2003) “Reputation and Survival: Learning in a Dynamic Signalling Model,” *Review of Economic Studies*, 70(2), 231-251.

[10] Baron, David and Roger Myerson (1982) “Regulating a Monopolist with Unknown Costs,” *Econometrica*, 50(4), 911-930.

[11] Barro, Robert (1986) “Reputation in a Model of Monetary Policy with Incomplete Information,” *Journal of Monetary Economics*, 17, 3-20.

[12] Barro, Robert and David Gordon (1983) “Rules, Discretion and Reputation in a Model of Monetary Policy,” *Journal of Monetary Economics*, 12, 101-122.

[13] Bergemann, Dirk and Ulrich Hege (1998) "Venture Capital Financing, Moral Hazard, and Learning," *Journal of Banking & Finance*, 22, 703-735.

[14] Bergemann, Dirk and Ulrich Hege (2005) "The Financing of Innovation: Learning and Stopping." *RAND Journal of Economics*, 36(4), 719-752.

[15] Board, Simon and Moritz Meyer-ter-Vehn (2013) "Reputation for Quality," *Econometrica*, 81(6), 2381-2462.

[16] Bonatti, Alessandro and Johannes Hörner (2017) "Career Concerns with Exponential Learning," *Theoretical Economics*, 12(1), 425-475.

[17] Campbell, Arthur, Florian Ederer and Johannes Spinnewijn (2014) "Delay and Deadlines: Freeriding and Information Revelation in Partnerships," *American Economic Journal: Microeconomics*, 6(2), 163-204.

[18] Chassang, Sylvain (2010) "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts," *American Economic Review*, 100(1), 448-465.

[19] Che, Yeon-Koo and Johannes Hörner (2018) "Optimal Design for Social Learning," *Quarterly Journal of Economics*, forthcoming.

[20] Chung, Kai-Lai (1974) *A Course in Probability Theory*, Third Edition, Elsevier.

[21] Cisternas, Gonzalo (2015) "Two-Sided Learning and Moral Hazard," *Review of Economics Studies*, forthcoming.

[22] Cole, Harold, James Dow and Willam English (1995) "Default, Settlement and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt," *International Economic Review*, 36(2), 365-385.

[23] Cripps, Martin, Eddie Dekel and Wolfgang Pesendorfer (2005) "Reputation with Equal Discounting in Repeated Games with Strictly Conflicting Interests," *Journal of Economic Theory*, 121, 259-272.

[24] Cripps, Martin, George Mailath and Larry Samuelson (2004) "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72(2), 407-432.

[25] Cripps, Martin, Klaus Schmidt and Jonathan Thomas (1996) "Reputation in Perturbed Repeated Games," *Journal of Economic Theory*, 69(2), 387-410.

[26] Deb, Joyee and Yuhta Ishii (2017) "Reputation Building under Uncertain Monitoring," Working Paper, Yale School of Management.

[27] Dellarocas, Chrysanthos (2006) "Reputation Mechanisms," *Handbook on Information Systems and Economics*, T. Hendershott edited, Elsevier Publishing, 629-660.

[28] Dewatripont, Mathias, Ian Jewitt and Jean Tirole (1999) "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies," *Review of Economic Studies*, 66(1), 199-217.

[29] Echenique, Federico (2004) "Extensive-Form Games and Strategic Complementarities," *Games and Economic Behavior*, 46, 348-364.

[30] Ekmekci, Mehmet (2011) "Sustainable Reputations with Rating Systems," *Journal of Economic Theory*, 146(2), 479-503.

[31] Ekmekci, Mehmet, Olivier Gossner and Andrea Wilson (2012) "Impermanent Types and Permanent Reputations," *Journal of Economic Theory*, 147(1), 162-178.

[32] Ellison, Glenn and Sara Ellison (2011) "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbets Prior to Patent Expiration," *American Economic Journal-Microeconomics*, 3(1), 1-36.

[33] Ely, Jeffrey (2017) "Beeps," *American Economic Review*, 107(1), 31-53.

[34] Ely, Jeffrey, Drew Fudenberg and David Levine (2008) "When is Reputation Bad?" *Games and Economic Behavior*, 63, 498-526.

[35] Ely, Jeffrey, Johannes Hörner and Wojciech Olszewski (2005) "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73(2), 377-415.

[36] Ely, Jeffrey and Juuso Välimäki (2003) "Bad Reputation," *Quarterly Journal of Economics*, 118(3), 785-814.

[37] Faingold, Eduardo (2013) "Reputation and the Flow of Information in Repeated Games," Working Paper, Yale University.

[38] Faingold, Eduardo and Yuliy Sannikov (2011) "Reputation in Continuous-Time Games," *Econometrica*, 79(3), 773-876.

[39] Fuchs, William (2007) "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97(4), 1432-1448.

[40] Fudenberg, Drew, David Kreps and David Levine (1988) "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44(2), 354-380.

[41] Fudenberg, Drew, David Kreps and Eric Maskin (1990) "Repeated Games with Long-Run and Short-Run Players," *Review of Economic Studies*, 57, 555-573.

[42] Fudenberg, Drew and David Levine (1983) "Subgame Perfect Equilibria of Finite and Infinite Horizon Games," *Journal of Economic Theory*, 31, 251-268.

[43] Fudenberg, Drew and David Levine (1989) "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759-778.

[44] Fudenberg, Drew and David Levine (1992) "Maintaining a Reputation when Strategies are Imperfectly Observed," *Review of Economic Studies*, 59(3), 561-579.

[45] Fudenberg, Drew and Eric Maskin (1986) "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533-554.

[46] Ghosh, Parikshit and Debraj Ray (1996) "Cooperation and Community Interaction Without Information Flows," *Review of Economic Studies*, 63(3), 491-519.

[47] Gibbons, Robert and John Roberts (2013) "Handbook of Organizational Economics", Princeton University Press.

[48] Gossner, Olivier (2011) "Simple Bounds on the Value of a Reputation," *Econometrica*, 79(5), 1627-1641.

[49] Halac, Marina, Navin Kartik and Qingmin Liu (2017) "Contests for Experimentation," *Journal of Political Economy*, 125(5), 1523-1569.

[50] Halac, Marina and Andrea Prat (2016) "Managerial Attention and Worker Engagement," *American Economic Review*, 106(10), 3104-3132.

[51] Harrington, Joseph (1986) "Limit Pricing when the Potential Entrant is Uncertain of Its Cost Function," *Econometrica*, 54(2), 429-437.

[52] Hart, Sirgiu (1985) "Nonzero-Sum Two-Person Repeated Games with Incomplete Information," *Mathematics of Operations Research*, 10(1), 117-153.

[53] Holmström, Bengt (1999) "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169-182.

[54] Hörner, Johannes (2002) "Reputation and Competition," *American Economic Review*, 92(3), 644-663.

[55] Hörner, Johannes and Nicholas Lambert (2015) "Motivational Ratings," Working Paper, Yale University and Stanford University.

[56] Horner, Johannes and Stefano Lovo (2009) "Belief-Free Equilibria in Games With Incomplete Information," *Econometrica*, 77(2), 453-487.

[57] Hörner, Johannes, Stefano Lovo and Tritan Tomala (2011) "Belief-free Equilibria in Games with Incomplete Information: Characterization and Existence," *Journal of Economic Theory*, 146(5), 1770-1795.

[58] Hörner, Johannes and Larry Samuelson (2013) "Incentives for Experimenting Agents," *RAND Journal of Economics*, 44(4), 632-663.

[59] Jehiel, Philippe and Larry Samuelson (2012) "Reputation with Analogical Reasoning," *Quarterly Journal of Economics*, 127(4), 1927-1969.

[60] Kalai, Ehud and Ehud Lehrer (1993) "Rational Learning Leads to Nash Equilibrium," *Econometrica*, 61(5), 1019-1045.

[61] Kaya, Ayça (2009) "Repeated Singalling Games," *Games and Economic Behavior*, 66, 841-854.

[62] Kahn, Lisa, B. (2013) "Asymmetric Information between Employers," *American Economic Journal: Applied Economics*, 5(4), 165-205.

[63] Keller, Godfrey, Sven Rady and Martin Cripps (2005) "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73(1), 39-68.

[64] Klein, Benjamin and Keith Leffler (1981) "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89(4), 615-641.

[65] Klein, Nicolas and Sven Rady (2011) "Negatively Correlated Bandits," *Review of Economic Studies*, 78(2), 693-732.

[66] Kremer, Ilan, Yishay Mansour and Motty Perry (2015) "Implementing the *Wisdom of Crowd*", *Journal of Political Economy*, 122(5), 998-1012.

[67] Kreps, David and Robert Wilson (1982) "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

[68] Lee, Jihong and Qingmin Liu (2013) "Gambling Reputation: Repeated Bargaining with Outside Options," *Econometrica*, 81(4), 1601-1672.

[69] Levine, David (1998) "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1, 593-622.

[70] Liu, Qingmin (2011) "Information Acquisition and Reputation Dynamics," *Review of Economic Studies*, 78(4), 1400-1425.

[71] Liu, Qingmin and Andrzej Skrzypacz (2014) "Limited Records and Reputation Bubbles," *Journal of Economic Theory* 151, 2-29.

[72] Liu, Shuo and Harry Pei (2017) "Monotone Equilibria in Signalling Games," Working Paper, MIT and University of Zurich.

[73] Mailath, George and Stephen Morris (2002) "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102, 189-228.

[74] Mailath, George and Larry Samuelson (2001) "Who Wants a Good Reputation?" *Review of Economic Studies*, 68(2), 415-441.

[75] Maskin, Eric and Jean Tirole (1990) "The Principal-Agent Relationship with an Informed Principal: The Case of Private Values," *Econometrica*, 58(2), 379-409.

[76] Maskin, Eric and Jean Tirole (2001) "Markov Perfect Equilibrium I: Observable Actions," *Journal of Economic Theory*, 100, 191-219.

[77] Milgrom, Paul (2008) "What the Seller Won't Tell You: Persuasion and Disclosure in Markets," *Journal of Economic Perspectives*, 22(2), 115-131.

[78] Milgrom, Paul and Sharon Oster (1987) "Job Discrimination, Market Forces and the Invisibility Hypothesis," *Quarterly Journal of Economics*, 102(3), 453-476.

[79] Milgrom, Paul and John Roberts (1982) "Predation, Reputation and Entry Deterence," *Journal of Economic Theory*, 27, 280-312

[80] Murto, Pauli and Juuso Välimäki (2013) "Delay and Information Aggregation in Stopping Games with Private Information," *Journal of Economic Theory*, 148, 2404-2435.

[81] Myerson, Roger (1991) "Game Theory, Analysis of Conflict," Harvard University Press.

[82] Phelan, Christopher (2006) "Public Trust and Government Betrayal," *Journal of Economic Theory*, 130(1), 27-43.

[83] Pei, Harry Di (2017) "Equilibrium Payoffs in Repeated Games with Interdependent Values," Working Paper, MIT.

[84] Pęski, Marcin (2014) "Repeated Games with Incomplete Information and Discounting," *Theoretical Economics*, 9, 651-694.

[85] Pęski, Marcin and Juuso Toikka (2017) "Value of Persistent Information," *Econometrica*, forthcoming.

[86] Roddie, Charles (2012) "Signaling and Reputation in Repeated Games: Finite Games," Working Paper, University of Cambridge.

[87] Schelling, Thomas (1960) "The Strategy of Conflict," Harvard University Press.

[88] Schmidt, Klaus (1993a) "Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests," *Econometrica*, 61(2), 325-351.

[89] Schmidt, Klaus (1993b) "Commitment through Incomplete Information in a Simple Repeated Bargaining Game," *Journal of Economic Theory*, 60(1), 114-139.

[90] Seamans, Robert (2013) "Threat of Entry, Asymmetric Information and Pricing," *Strategic Management Journal*, 34, 426-444.

[91] Sorin, Sylvain (1999) "Merging, Reputation, and Repeated Games with Incomplete Information," *Games and Economic Behavior*, 29, 274-308.

[92] Sweeting, Andrew, James Roberts and Chris Gedge (2016) "A Model of Dynamic Limit Pricing with an Application to the Airline Industry," Working Paper, University of Maryland and Duke University.

[93] Tirole, Jean (1996) "A Theory of Collective Reputations (with applications to the persistence of corruption and to firm quality)," *Review of Economic Studies*, 63, 1-22.

[94] Topkis, Donald (1998) *Supermodularity and Complementarity*, Princeton University Press.

[95] Toxvaerd, Flavio (2017) "Dynamic Limit Pricing," *RAND Journal of Economics*, 48(1), 281-306.

[96] Waldman, Michael (1984) "Job Assignments, Signalling, and Efficiency," *RAND Journal of Economics*, 15(2), 255-267.

[97] Watson, Joel (1993) "A Reputation Refinement without Equilibrium," *Econometrica*, 61(1), 199-205.

[98] Weinstein, Jonathan and Muhamet Yildiz (2007) "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75(2), 365-400.

[99] Weinstein, Jonathan and Muhamet Yildiz (2013) "Robust Predictions in Infinite-Horizon Games–An Unrefinable Folk Theorem," *Review of Economic Studies*, 80(1), 365-394.

[100] Weinstein, Jonathan and Muhamet Yildiz (2016) "Reputation without Commitment in Finitely Repeated Games," *Theoretical Economics*, 11(1), 157-185.

[101] Wilson, Robert (1987) "Game-Theoretic Analyses of Trading Processes," in Advances in Economic Theory, Fifth World Congress, editted by Truman Bewley, Cambridge University Press, Chapter 2, 3370.

[102] Wolitzky, Alexander (2011) "Indeterminacy of Reputation Effects in Repeated Games with Contracts," *Games and Economic Behavior*, 73, 595-607.