The Dissertation Committee for Benjamin Leonard Sperisen
certifies that this is the approved version of the following dissertation:

Essays on Reputation and Repeated Games

Committee:

Thomas Wiseman, Supervisor

Lorenzo Alvisi

V. Bhaskar

Dale Stahl

Caroline Thomas

Essays on Reputation and Repeated Games

by

Benjamin Leonard Sperisen, B.A.; B.S.; M.S. Econ.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2015

Acknowledgments

I am deeply indebted to Thomas Wiseman for his continuous advice, guidance, and encouragement, as I could not have asked for a better advisor. I am very grateful for the help of the other members of my committee: Lorenzo Alvisi, V. Bhaskar, Dale Stahl, and Caroline Thomas. I thank Ricardo Ainslie, Cecilia Ballí, Neal Ghosh, Florian Kuhn, Melinda Petre, Viridiana Rios, Thomas Roderick, Maxwell Stinchcombe, and Ina Taneva for helpful comments and discussions.

I also thank my parents, Franz and Susi Sperisen, who encouraged curiosity and supported my educational pursuits throughout my life, and Chelsea Couch, for her love and support during both the good and (especially) the tough times of writing this dissertation. Finally, I express gratitude to Carl Enomoto, Christopher Erickson, and James Peach, who kindled the interest in economics that led me here.

Essays on Reputation and Repeated Games

Publication No. ____________________

Benjamin Leonard Sperisen, Ph.D.
The University of Texas at Austin, 2015

Supervisor: Thomas Wiseman

This dissertation consists of three essays on reputation and repeated games. Reputation models typically assume players have full memory of past events, yet in many applications this assumption does not hold. In the first chapter, I explore two different relaxations of the assumption that history is perfectly observed in the context of Ely and Välimäki's (2003) mechanic game, where reputation (with full history observation) is clearly bad for all players. First I consider "limited history," where short-run players see only the most recent T periods. For large T, the full history equilibrium behavior always holds due to an "echo" effect (for high discount factors); for small T, the repeated static equilibrium exists. Second I consider "fading history," where short-run players randomly sample past periods with probabilities that "fade" toward zero for older periods. When fading is faster than a fairly lax threshold, the long-run player always acts myopically, a result that holds more generally for reputation games where the long-run player has a strictly dominant stage game action. This finding suggests that reputational incentives may be too weak to affect long-run player behavior in some realistic word-of-mouth environments.

The second chapter develops general theoretical tools to study incomplete information games where players observe only finitely many recent periods. I derive a recursive characterization of the set of equilibrium payoffs, which allows analysis of both stationary and (previously unexplored) non-stationary equilibria. I also introduce "quasi-Markov

perfection," an equilibrium refinement which is a necessary condition of any equilibrium that is "non-fragile" (purifiable), i.e., robust to small, additively separable and independent perturbations of payoffs.

These tools are applied to two examples. The first is a product choice game with 1-period memory of the firm's actions, obtaining a complete characterization of the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on an "honest" Stackelberg commitment type, which shows that non-stationary equilibria expand the equilibrium set. The second is the same game with long memory: in all stationary and purifiable equilibria, the long-run player obtains exactly the Stackelberg payoff so long as the memory is longer than a threshold dependent on the prior. These results show that the presence of the honest type (even for arbitrarily small prior beliefs) qualitatively changes the equilibrium set for any fixed discount factor above a threshold independent of the prior, thereby not requiring extreme patience.

The third chapter studies the question of why drug trafficking organizations inflict violence on each other, and why conflict breaks out under some government crackdowns and not others, in a repeated games context. Violence between Mexican drug cartels soared following the government's anti-cartel offensive starting in 2006, but not under previous crackdowns. I construct a theoretical explanation for these observations and previous empirical research. I develop a duopoly model where the firms have the capacity to make costly attacks on each other. The firms use the threat of violence to incentivize inter-cartel cooperation, and under imperfect monitoring, violence occurs on the equilibrium path of a high payoff equilibrium. When a "corrupt" government uses the threat of law enforcement as a punishment for uncooperative behavior, violence is not needed as frequently to achieve high payoffs. When government cracks down indiscriminately, the firms may return to frequent violence as a way of ensuring cooperation and high payoffs, even if the crackdown makes drug trafficking otherwise less profitable.

Table of Contents

Chapter 1

Reputation under Limited and Fading History

1.1 Introduction

The reputation literature has shown that even very small uncertainty about a player's type can have dramatic effects on equilibrium behavior and payoffs. Building on the seminal work of Kreps and Wilson (1982) and Milgrom and Roberts (1982), Fudenberg and Levine (1989; 1992) show that introducing such uncertainty assures the long-run player of a payoff arbitrarily close to the payoff that would be achievable by a credible commitment to an action of her choice. Ely and Välimäki (2003) (henceforth EV) construct a model where reputation has a similarly dramatic but negative effect on payoffs (of all players). These models typically assume that short-run players see the full history of past signals.

In reality, agents often perceive reputation through only limited excerpts of the past, raising the question: how robust are these results to relaxing that assumption? The focus of this chapter is answering this question with regard to EV's model, considering two different forms of relaxation: "limited history" (modeling a public list of recent reviews) and "fading history" (modeling word-of-mouth). I find that the full history equilibrium behavior is robust to short-run players seeing many (but not all) past periods in both cases. When short-run players see relatively few past periods, behavior differs between the two models: limited history yields the repeated one-shot equilibrium, while fading history yields myopic long-run player behavior but strictly higher ex ante payoffs for the short-run players. The fading history results also apply to the chain store game (a typical example of the games considered by Fudenberg and Levine (1989), where reputation is good for the long-run player).

EV's model, the mechanic game, has the feature that the long-run player is clearly harmed by reputation.[1] The rational "good" long-run player, who offers expert services to short-run players, has payoffs that perfectly coincide with those of short-run players, and she wants to separate herself from a "bad" type that harms short-run players; this temptation to separate harms the short-run player, causing the whole market to fail. EV point out these dynamics could be a concern in a number of asymmetric information settings involving expert sellers, such as auto mechanics, lawyers, management consultants and medical doctors. These markets generally involve consumers who are not perfectly informed about the seller's past. Consider motorists who solicit information about an auto mechanic through word-of-mouth, patients who choose to see a dentist after reading the first few reviews listed on Yelp, or a consultant who provides prospective clients with a list of only her most recent references on her resumé. The robustness of the EV result in settings where consumers have a limited view of the past may shed light on both positive questions, such as when and why experts are hired in real markets with this feature, and normative issues, such as the optimal design and welfare effects of review websites.

In the mechanic game, the "good" mechanic (rational long-run player) and the motorist (short-run player) have coinciding interests in the stage game: the motorist's car has a problem, and both want the problem fixed correctly. Motorists do not know which repairs their cars need (either a cheap tune-up c or an expensive engine replacement e), but the mechanic does. The motorist would like to hire the good mechanic instead of an outside option $\varnothing$, if she does the right repair. However, the introduction of even a tiny probability that the mechanic is a commitment type (the "bad" mechanic, who performs an expensive engine replacement no matter what problem the car has) impedes the ability of the mechanic and motorists to interact when the motorist prefers the outside option to hiring the bad mechanic.

When motorists can see the entire history of hiring decisions and repairs, a history with sufficiently many engine replacements and no tune-ups yields a belief that the mechanic

[1]Short-run players are also harmed, but this aspect is more common in other reputation games, such as the chain store game.

is likely enough to be bad that the outside option is necessarily preferable to hiring. All subsequent motorists avoid the mechanic, "freezing" the "bad belief" and preventing the mechanic from ever being hired again. At a "critical" history, where the mechanic is just one engine replacement away from a frozen bad belief, a sufficiently patient good mechanic is inevitably tempted into performing a tune-up (even when an engine replacement is needed), signaling that she is good. Such signaling behavior is harmful to the motorist receiving the unnecessary tune-up, whose best response is to avoid the mechanic. Before the critical history, the mechanic's anticipation of the critical motorist's decision not to hire must lead to a certain (possibly unnecessary) tune-up even earlier, so this previous motorist also does not hire, and so on by backward induction, leading to a complete unraveling of the market and no hiring on the equilibrium path of all renegotiation-proof Nash equilibria.

However, real mechanics likely do not expect that any particular action (or the signal it generates) is certain to be observed by every subsequent customer. Even if a mechanic knew with certainty that a customer would immediately report her actions to the world in a review on Yelp, she knows that many future potential customers may not see the review, either because they do not check Yelp at all, or because with time, the review is eventually pushed out of sight on the first page. Word-of-mouth seems particularly unlikely to yield fully informed customers due to its decentralized, random nature.

Motivated by this observation, I relax the assumption of full history observation through two different types of limitations on the history seen by short-run players. The first type is "limited history," where each motorist sees a fixed number of periods into the past, but no further. I find that when this "memory" is long enough, exactly the same equilibrium behavior as the full history model is obtained, because the events in the beginning periods "echo" forever through the participation decisions of the motorists. If the mechanic signals she is good early on with a tune-up, all motorists who see that first tune-up hire her, and the next "generation" of motorists, who do not see that tune-up but see all the hiring that followed, infer her type and also hire (even if they only see engine replacements), and so on forever. By contrast, if the mechanic sends bad signals by performing many engine replacements (and no tune-ups) early on, the first generation of motorists eventually stop hiring, the next generation sees this lack of hiring and follows suit, and so on.

Making this memory too short for an individual motorist to learn much about the mechanic allows an equilibrium that avoids the bad reputation result, and the stage game equilibrium is repeatedly played. But in this equilibrium, reputation is also rendered worthless: it does not help the long-run or short-run players because it is too uninformative to have any effect on behavior at all.

The second restriction on history observation I consider is "fading history," in which motorists see the last period with probability $\lambda \in (0, 1)$, the second-to-last period with probability λ^2, and so on for all past periods. This can be thought of as modeling the decentralized randomness of word-of-mouth. Like limited history with high T, fading history yields the bad reputation result for high λ. When λ is small enough, reputational incentives are too weak to cause bad reputation, but reputation still sometimes helps the motorists avoid the bad mechanic; this differs from the low T limited history case, where reputation is always useless to the motorists. This result seems reasonably realistic: the good mechanic is not diverted from serving customers by extremely strong reputation incentives, good and bad mechanics are both sometimes hired, and some of the more discerning customers hire the good mechanic while avoiding the bad mechanic.

In fact, this result for fading history with low λ applies to a more general class of reputation games: when the long-run player has a strictly dominant action in the stage game, it causes the long-run player to behave myopically and always choose that dominant action (as though reputation did not exist), while the payoffs of short-run players are often greater than under the static game because they are sometimes well-informed. The upper bound on λ is not trivially small: for a patient mechanic, it corresponds to a given motorist talking to an average of $\frac{1}{2}$ future potential customers about their experience, high enough to cover scenarios with significant (but not totally ubiquitous) word-of-mouth communication.[2]

[2] Appendix 1.2.2 gives a higher upper bound (for the mechanic game) that allows talking to an average of $\frac{2}{3}$ future motorists, given some reasonable restrictions on equilibria. This upper bound can be even

The result is also robust to correlation between observations, which allows applications to more centralized communication like online forums where public messages "fade" over time (see Remark 1.2.1 for a discussion). This suggests reputation may be too weak to affect long-run player behavior in many real word-of-mouth situations.

Though fading history is intended as a model of word-of-mouth communication, it differs importantly from existing work on word-of-mouth (e.g. Ellison and Fudenberg (1995), Banerjee and Fudenberg (2004)) where players randomly sample some fixed number of past events. Key differences include that under fading history, the "sample size" is random, and that players are more likely to observe the recent past than distant past. The fading of past events is critical for ruling out the never-ending echo that occurs in the limited (but long) history case.

To better understand the role of these restrictions on history monitoring, I also consider them in the context of the chain store game of Selten (1978), variants of which have been widely used to study reputation (for example, Pitchik (1993), Aoyagi (1996) and Wiseman (2008)). The chain store game is a typical example of a Stackelberg-type game, where reputation bounds the long-run player's payoff from below. The (general) result for fading history with low λ applies directly to the chain store game, and the effect on short-run player payoffs is actually more dramatic because they need only observe a single past event to learn the long-run player's type. The other results do not carry over so simply, and in fact a "myopic equilibrium" does not exist at all for the limited history chain store game, even when limited to seeing just one previous period.

Bounded memory repeated games without reputation have been the topic of a number of papers — for example, Sabourian (1998), Mailath and Olszewski (2011a), and Bhaskar, Mailath, and Morris (2013a) — but such games with reputation are relatively unexplored. I use the term "limited history" to distinguish it from the similar "limited records" studied by Liu and Skrzypacz (2014a), the only difference being that limited records allow

higher (allowing talking to up to but not including an average of 1 future motorist) depending on the parameterization of the stage game.

short-run players to see only the long-run player's actions in each observed period, whereas limited history lets them see the full outcome of each period.[3] This difference can produce starkly different results because under limited records, the long-run player can unilaterally "clean" her history. Liu and Skrzypacz find cyclical equilibrium behavior — "riding of reputation bubbles" — that is qualitatively different from both the one-shot and complete history cases. Liu (2011a) also finds cyclical equilibria in an environment where short-run players incur a cost to observe limited records of past long-run player actions. These papers show reputation being continually accumulated, exploited and then replenished. By contrast, this chapter finds *non*-cyclical behavior under limited history, because reputation inevitably gets "stuck" in a particular state, despite only viewing the recent past.

Limited records are more realistic in certain settings; Liu and Skrzypacz point to the example of the Better Business Bureau, which reports complaints on businesses from the last 36 months, but does not report on the business's volume or types of transactions. However, in many cases long-run players need the cooperation of short-run players to establish a desirable reputation. For example, a consultant must be hired by today's client in order to provide a favorable reference tomorrow; she cannot do this unilaterally. In such environments, limited history is a more appropriate assumption. A more detailed discussion of the differences between limited history and limited records is given in Subsection 1.2.1.1.

This chapter also relates to other extensions of the mechanic game. EV show that when the motorist is also a long-run player, an equilibrium exists where the mechanic and motorist are able to interact. Mailath and Samuelson (2006a) consider the possibility of random "captive consumers," who hire no matter the history. Ely, Fudenberg, and Levine (2008) extend bad reputation to a broader class of games, illustrating the difference between bad and good reputation. For example, they allow a larger set of commitment types; bad reputation is robust to the introduction of a sufficiently small probability of a Stackelberg commitment type (who always performs the correct repair), but if the probability is high enough relative to the probability of the bad type, reputation is no longer bad. This

[3] When motorists do not hire, this prevents observation of the long-run player's action.

assumption seems reasonable in markets where consumers have faith that experts are more often non-strategically honest than bad, but is more problematic where consumers are more suspicious.[4] Though my results only consider limited and fading history as applied to the original EV mechanic game, there is no apparent reason why similar results would not apply to such generalizations.

The rest of the chapter is organized as follows. Section 1.2 discusses the main results for the mechanic game, with Subsection 1.2.1 covering limited history and Subsection 1.2.2 covering fading history. Section 1.3 considers applications and implications for the chain store game. Section 1.4 concludes. The Appendix contains omitted proofs.

1.2 The Mechanic Game

In the mechanic game, reputation leads to a lower payoff for both the long-run and short-run players than in the static game. A long-lived car mechanic faces a different short-lived motorist each period. Each motorist's car is in one of two states, each requiring a different repair: either a cheap tune-up c or an expensive engine replacement e. The states are drawn iid each with probability $\frac{1}{2}$. The motorist does not know which repair is needed, but the mechanic does.

In each period, the motorist first chooses to either hire the mechanic or choose an outside option $\varnothing$ with payoff zero. If hired, the mechanic observes the state of the car, either θ_c or θ_e. The motorist benefits if the mechanic performs the correct repair, receiving payoffs according to the following table:

	θ_c	θ_e
c	u	$-w$
e	$-w$	u
$\varnothing$	0	0

[4] Surveys by Gallup (2013) show that such faith varies widely across some of the applications suggested by EV. Medical doctors are among those most trusted, with 69% of respondents in December 2013 rating their "honesty and ethical standards" as "high" or "very high." Auto mechanics are viewed much less favorably (29%), and lawyers (20%) rank even further down the list.

Assume $w > u > 0$. This insures that if the mechanic chooses the repair independent of the state, the motorist will prefer the outside option.

The mechanic can be one of two types: good (g) and bad (b). The mechanic's type is denoted s. The good mechanic has the same stage payoff as the motorist (in the table above), and wants to maximize her expected discounted average payoff, discounted at rate $\delta \in (0, 1)$. The bad mechanic is non-strategic and simply performs engine replacements, regardless of the state.[5] Motorists observe the full history of repair and hiring decisions, but not the previous motorists' states (i.e., it is not known whether the repairs were correct), as public knowledge. Beginning at period 0, the first motorist has prior belief μ^0 that the mechanic is bad, and subsequent motorists update their beliefs about the mechanic's type according to Bayes' rule.

In the one-shot game, the motorist's expected payoff for hiring is simply $\frac{1}{2}\mu(u - w) + (1 - \mu)u$ where μ is the belief that the mechanic is bad (doing the right repair is strictly dominant for the good mechanic). She will hire the mechanic only if this expected payoff is nonnegative, which is clearly false when the belief μ is greater than critical value $p^* \equiv 2u/(u + w)$ since $u < w$.[6]

EV prove that the supremum of the mechanic's Nash equilibrium payoffs must converge to zero for δ close enough to one, so that equilibria where the mechanic is hired must have the mechanic hired only infrequently. They point out that equilibria with such infrequent hiring have an implausible feature: once the mechanic performs a single tune-up, she reveals herself to be good (with certainty) to all future motorists. After a tune-up, it makes sense that all subsequent motorists (knowing the mechanic is good) will want to hire, and

[5] EV also show their result holds for a strategic bad mechanic who receives a discounted sum of period payoffs that do not depend on the motorists' states, receiving u for performing an engine replacement , $-w$ for a tune-up and 0 when not hired.

[6] p^* is defined as the belief such that the motorist is indifferent about hiring when the good mechanic always does the correct repair, i.e.

$$p^* \frac{u - w}{2} + (1 - p^*)u = 0.$$

the mechanic will want to perform correct repairs for them. For this reason, EV use the following renegotiation-proofness assumption to rule out such dubious behavior.

Assumption 1 (Renegotiation-Proofness)**.** *The mechanic is hired at any history on the equilibrium path at which she is known to be good by the motorist.*

EV then find the following dramatic result.

Theorem 1.2.1. *Let $\mu^0 > 0$ be given. When δ is close enough to one, any Nash equilibrium satisfying Assumption 1 has a unique equilibrium outcome where the mechanic is never hired.*

Without going into the proof here, the intuition behind it is that in any equilibrium, if the mechanic performs some number L engine replacements and no tune-ups, the motorists' beliefs must rise above p^* and they do not hire. If a motorist hires at any history, the mechanic must perform an engine replacement with sufficient probability (otherwise the motorist's expected payoff from hiring would be negative), and this means that, with positive probability, the mechanic performs L consecutive engine replacements on the equilibrium path. After performing $L - 1$ engine replacements (and no tune-ups), an engine replacement gives a continuation payoff of 0, compared with a continuation payoff of u for a tune-up (since all future motorists hire). When she is sufficiently patient, the mechanic always performs a tune-up, so she cannot be hired at a "critical" history (i.e., after $L - 1$ engine replacements without tune-ups); backwards induction leads to the result of no hiring on the equilibrium path.

1.2.1 Limited History

I relax the full history monitoring assumption by allowing motorists to view finite T previous periods. First I consider the situation when T is large, obtaining the same behavior as Theorem 1.2.1.

At the heart of this result is making sure that a critical history can "fit" into the memory of motorists. I formalize this with the following notation. An infinite history

$h = (h_0, h_1, ...)$ is an infinite sequence of events, where h_k is the outcome at period k, either a repair (c or e) or a no-hiring decision ($\varnothing$). Often it is useful to look at a history ending just before some period t, which I denote with a superscript: a history h^t (at some period t) is a t-length sequence $(h_0^t, ..., h_{t-1}^t)$ where h_k^t is the event at period $k < t$. Since many short-run players do not see all of the past history, I sometimes refer to this as the "full" or "complete history" (to contrast with the "observable subhistory" defined below). I also sometimes denote the event at period k as η_k. This notation is used when discussing expectations for future events, given a particular history; for example, the probability that the mechanic is not hired at period k given history h^k is written $P(\eta_k = \varnothing | h^k)$, instead of $P(h_k = \varnothing | h^k)$, to avoid confusion about what is already part of the history and what is yet to happen.

For periods $0, ..., T$, motorists observe the full history. Given a history h^t at any $t \in \{0, ..., T\}$ where the mechanic is hired, the expected payoff of hiring must be nonnegative (otherwise the motorist chooses the outside option):

$$\mu^t(h^t)\left(\frac{u-w}{2}\right) + (1-\mu^t(h^t))(\beta^t(h^t)u - (1-\beta^t(h^t))w) \geq 0,$$

where $\mu^t(h^t)$ is the posterior belief of motorist t that the mechanic is bad and $\beta^t(h^t)$ is the probability that the good mechanic performs the correct repair at that history. Solving for β^t gives

$$\beta^t(h^t) \geq \frac{1}{u+w}\left[w + \frac{\mu^t(h^t)}{1-\mu^t(h^t)}\left(\frac{w-u}{2}\right)\right] \geq \frac{w}{u+w}. \tag{1.2.1}$$

Let $\beta_a^t(h^t)$ be the probability that the mechanic does repair a conditional on a being needed, so that $\beta^t(h^t) = \frac{1}{2}\beta_c^t + \frac{1}{2}\beta_e^t \leq \frac{1}{2} + \frac{1}{2}\beta_c^t$, which can be substituted into (1.2.1) to get the lower bound

$$\beta_c^t \geq \frac{w-u}{u+w} \equiv \beta^* \tag{1.2.2}$$

on the probability that a mechanic performs a needed tune-up.

Each period's motorist has a posterior belief, but after period T motorists do not know the beliefs of previous motorists (because they do not see what those motorists saw), which means it is not possible to calculate posteriors by simply updating the previous motorist's posterior, one after another. I use the following notation to denote the limited history that a particular motorist sees.

Definition 1.2.1. Given history h^t at any period t, the *observable subhistory* $\hat{h}^t$ at t is the sequence $(h^t_{t_0}, ..., h^t_{t-1})$, where $t_0 \equiv \max\{0, t-T\}$. The event at some period $t' \in \{t_0, ..., t-1\}$ is denoted $h^t_{t'}$.

Since the posterior after an observable subhistory is not a simple update on the previous motorist's posterior, it will be useful to separate the calculation into steps, period by period across the observable subhistory.

Definition 1.2.2. For any observable subhistory $\hat{h}^t$ at period t, the *partial posterior belief* $\mu^t_{t'}(\hat{h}^t)$ of motorist t is the probability that the mechanic is bad given prior μ^0 and the observed periods $t_0, ..., t'-1$, ignoring periods $t', ..., t-1$. Note that $\mu^t_{t_0}(\hat{h}^t) = \mu^0$ because it ignores all observations.

For $t \leq T$ when motorists still observe the complete history, the observable subhistory $\hat{h}^t$ and full history h^t and are the same, and the partial posterior evolves according to Bayes' rule as it does in the full history case. If a tune-up is observed at $t' < t$, then $\mu^t_{t'+1}(\hat{h}^t) = 0$; if a no-hire event is observed, then $\mu^t_{t'+1}(\hat{h}^t) = \mu^t_{t'}(\hat{h}^t)$; if an engine replacement is observed, then

$$\mu^t_{t'+1}(\hat{h}^t) = \frac{\mu^t_{t'}(\hat{h}^t)}{\mu^t_{t'}(\hat{h}^t) + (1 - \mu^t_{t'}(\hat{h}^t))[\frac{1}{2}\beta^{t'}_e(\hat{h}^{t'}) + \frac{1}{2}(1 - \beta^{t'}_c(\hat{h}^{t'}))]}.$$

Define

$$\Upsilon(\mu) \equiv \frac{\mu}{\mu + (1-\mu)[\frac{1}{2} + \frac{1}{2}(1 - \beta^*)]}$$

so that $\Upsilon(\mu^t_{t'}(\hat{h}^t))$ is a lower bound for $\mu^t_{t'+1}(\hat{h}^t)$ (for $t \leq T$, not $t > T$), and inductively define $\Upsilon^1(\mu) \equiv \Upsilon(\mu)$ and $\Upsilon^{k+1}(\mu) \equiv \Upsilon(\Upsilon^k(\mu))$, so that $\Upsilon^k(\mu)$ is a lower bound for the posterior after observing k engine replacements and no tune-ups at $t \leq T$. Finally, define

$$L(\mu^0) \equiv \min k \text{ such that } \Upsilon^k(\mu^0) > p^* \tag{1.2.3}$$

as an upper bound on the number of engine replacements that can be performed (without any tune-ups) in the first $T+1$ periods before the posterior exceeds p^*.

The result below establishes a lower bound on T sufficient for the bad reputation result. For $T > L(\mu^0)$, it is straightforward to see that the mechanic cannot perform more than L engine replacements without any tune-ups within the first $T+1$ periods on the equilibrium path. Motorists (before period $T+1$) who arrive after the Lth engine replacement will believe the mechanic is so likely to be bad that the payoff of hiring must be negative, preventing any hiring until at least period $T+1$. Performing more than L engine replacements in the first $T+1$ periods is only possible if the mechanic first performs a tune-up, so if motorist $T+1$ observes an engine replacement in every period $1, ..., T$, the mechanic is known to be good because an unobserved tune-up must have preceded the observable subhistory, i.e. at period 0 (the result of Lemma 1.1.1). In fact, this "echo" effect continues for all future periods; if motorists see the mechanic hired every observed period, these hiring decisions signal that the mechanic must be good, even if only engine replacements are observed.

Having ruled out equilibria where the mechanic performs more than L engine replacements (and no tune-ups) in the first $T+1$ periods, one may wonder if there exist equilibria where the mechanic performs L or fewer engine replacements. Such equilibria mean there are histories on the equilibrium path where the mechanic is not hired for many of the first $T+1$ periods, reducing her continuation payoff from doing an engine replacement. If T is large enough, the mechanic cannot resist the temptation to perform a tune-up that ensures she is hired in all of those periods.

This temptation effectively forces beliefs to be either 0 or greater than p^* by period $T+1$ in any equilibrium where the mechanic is hired with positive probability. Because motorists in periods $T+1$ and beyond know this, they need only look at their observable subhistories to tell whether the mechanic is definitely good (they see a tune-up or hiring in every observed period) or likely-enough-to-be-bad (motorists stop hiring at some point in the first T periods, and never hire again afterwards). Reputation is pinned into one of these two extreme states early on and frozen there by the hiring decisions, giving the large reputational incentive that causes bad reputation.

Theorem 1.2.2. *Let $\mu^0 > 0$ and $L(\mu^0)$ be given. If*

$$T > \left(2 + \frac{w}{u}\right) L - 1, \tag{1.2.4}$$

then for any sequential equilibrium satisfying Assumption 1 there is a unique equilibrium outcome where the mechanic is never hired when δ is close enough to one.

What happens when T is small? A "myopic equilibrium" exists where the good mechanic always performs the correct repair, so that the mechanic always performs an engine replacement with probability $\frac{1}{2}$. Given prior μ, the motorist's posterior after observing an engine replacement is

$$\bar{\Upsilon}(\mu) \equiv \frac{\mu}{\mu + \frac{1}{2}(1-\mu)} = \frac{2\mu}{1+\mu},$$

with $\bar{\Upsilon}^t(\mu)$ defined inductively like $\Upsilon(\cdot)$. Define

$$\bar{L}(\mu^0) \equiv \min t \text{ such that } \bar{\Upsilon}^t(\mu^0) > p^*.$$

Note that $\bar{L}(\mu^0) \leq L(\mu^0)$ because $L(\mu^0)$ is a lower bound that presumes the mechanic performs engine replacements with maximum probability $\frac{1}{2} + \frac{1}{2}(1-\beta^*) > \frac{1}{2}$.

Theorem 1.2.3. *Let $\mu^0 > 0$ and $T < \bar{L}(\mu^0)$ be given. A sequential equilibrium exists where the good mechanic always performs the correct repair and the motorists always hire.*

Proof. It is easy to show that always hiring is a best response for motorists. At any period t, the motorists' posterior is less than or equal to $\bar{\Upsilon}^T(\mu^0) \leq p^*$, so the payoff of hiring is nonnegative. For the mechanic, the continuation payoff of a tune-up is equal to that of an engine replacement (she is always hired, no matter her strategy), so performing the correct repair always yields a greater payoff. □

Theorem 1.2.3 avoids the disaster of Theorem 1.2.2, but it does so by preventing reputation from having *any* effect. Reputation neither tempts the mechanic into a costly tune-up that harms the motorist, nor does it help the motorists sort out a good mechanic from the bad, and the ex ante payoff for every motorist is equal to that of the one-shot game. Reputation is only useful to motorists if it gives them a posterior greater than p^* when the mechanic is bad, so that they can avoid hiring her; the myopic equilibrium avoids the bad reputation outcome precisely by ruling out that possibility.

Can reputation actually be useful in the limited history environment? Though these results do not give a complete answer to that question, they do suggest that the answer is probably "no," at least for equilibria that do not involve highly unrealistic behavior.

They do not tell us what behavior occurs when T is between the lower bound of Theorem 1.2.2 and the upper bound of Theorem 1.2.3, but they do show a tension between helping the motorist avoid the bad mechanic and limiting the reputational benefit of signaling to the mechanic that suggests any "useful reputation equilibria" are at best fragile and likely implausible.

Any equilibrium with useful reputation must delicately resolve this tension. Suppose there is an equilibrium where motorist t has an ex ante payoff higher than the one-shot or myopic equilibria. This equilibrium must avoid the "echo" effect of Theorem 1.2.2 by preventing motorist t's no-hiring decision at high posteriors from triggering further no-hiring events. One way to "dampen" the echo is by making t's decision not to hire only an imperfect signal that his posterior is greater than p^*, which requires that he mixes when the posterior is less than p^*. Yet if he is mixing, he must be have a payoff of zero at the observable subhistories where he mixes,[7] which is less than the myopic payoff. Adding more "noise" in the signal requires mixing (and thus decreasing the payoffs) for more observable subhistories, and it is unclear that the motorist can in fact come out ahead overall. An alternative is to allow motorist t's no-hire choice to be a perfect signal that his posterior is greater than p^*, and dampen the echo by having his successors in periods $t+1, ..., t+T$ mix, but having them mix requires lowering their payoffs instead. This damping also requires very time-specific

[7]This would be because the good mechanic is herself mixing between the right and wrong repair, which requires delicate balancing of her continuation payoffs.

behavior. If motorists' beliefs depend simply on the distribution rather than order of events (i.e. the ratio of engine replacements to no-hire events), then useful reputation equilibria are impossible. All such equilibria, if they exist, involve rather unrealistic coordination between the players (they must know which periods are being used to dampen). For these reasons, the most natural interpretation of the results is that under limited history, reputation is either bad or useless.

1.2.1.1 Limited History vs Limited Records

The results above relate to interesting work by Liu and Skrzypacz (2014a), who study a reputation game under "limited records" with very different behavior, which I call the "reputation bubble game." Their game also features a long-run player ("the firm") facing a sequence of short-run players ("the consumers"), with a stage game which can be interpreted as the consumer choosing how much to trust the firm (how large an order to purchase), followed by the firm choosing how much to honor that trust (the quality of product to deliver). The long-run player is either a strategic type or a commitment type who always honors the consumer's trust (delivering a high quality product). There are a number of differences between their environment and mine,[8] but the most interesting one is that they assume limited records, where short-run players only observe the long-run player's recent actions — not those of past short-run players. By contrast, limited history includes the full outcomes of recent periods, which reveals the short-run player actions (and, in the mechanic game, hides the mechanic's actions when not hired).

Their main result is that all equilibria in their environment feature "reputation bubbles," where the long-run player "cleans" her history to mimic the commitment type, and once the observable history is completely clean (i.e. contains only the commitment type's action), the long-run player exploits the short-run player. As the long-run player

[8] In their stage game, the long-run player has a static incentive to exploit the short-run player's trust (long-run incentives can prevent exploitation); in the mechanic game, there is a static incentive to *honor* the short-run player's trust (it is the long-run incentives that can lead to exploitation). Also, their consumers prefer to interact with the commitment type rather than the strategic type; in the mechanic game, it is the commitment type that they want to avoid.

cleans a “dirty” history, short-run players grant her trust not because they believe she is the commitment type (they know she is not because they see past exploitative behavior), but because the cleaning action itself honors the trust; this is what they call “riding a reputation bubble.” Such behavior is impossible under full history (or records) because the long-run player cannot surprise and exploit a short-run player as the history is impossible to clean. Their results differ from mine qualitatively in the sense that behavior under limited records is different from both complete records and no records; by contrast, the mechanic game has the same equilibrium behavior for both small $T > 0$ and $T = 0$, and the same behavior for large $T < \infty$ as $T = \infty$.

The limited records assumption makes the model much more tractable, and Liu and Skrzypacz point out that the assumption that only the long-run players' actions are observable is realistic in applications like the Better Business Bureau, which does not show how much business a firm gets but does show complaints. In many other settings, however, the long-run player needs short-run players to cooperate in order to send the signals she wants. For example, a consultant or lawyer needs clients to hire her in order to provide references to future clients; she cannot generate an observable “high quality” signal through sheer effort alone. Instead, future prospective clients will observe that she was not hired, and they may interpret that as a bad signal, leading to persistent unemployment. This lack of total control over one's reputation is critical to the bad reputation effect, giving the result that reputation has exactly the same impact when memory is long as when it is complete. To what extent this applies to other limited history reputation games remains an interesting open question (this is discussed further in Subsection 1.3.1).

1.2.2 Fading History

While the limited history model avoids the bad reputation result when records are sufficiently limited, it appears to do so at the expense of reputation being useful. The information structure introduced in this subsection allows reputation to be weakened enough that the mechanic can play myopically, but it is still sometimes informative enough to motorists that they do not hire the mechanic, achieving the "middle ground" that seems to be lacking under limited history. This is achievable in a similar "myopic equilibrium" that does not require such complicated strategies, and although it reduces the motorist payoff conditional on the mechanic being good (because the good mechanic sometimes "looks bad" and is not hired), it is outweighed by the increase in the motorist payoff conditional on the mechanic being bad such that the ex ante motorist payoff is greater than the one-shot payoff.

Under fading history, the motorist observes each previous period with some positive probability. Let $p_t^{t'}$ denote the probability that the motorist in period t' observes the actions in period $t < t'$. By comparison, under full history this probability is always one; under limited history, $p_t^{t'} = 1$ for $t' \in \{t+1, ..., t+T\}$ and $p_t^{t'} = 0$ for $t' > t+T$. Under fading history, this probability is never one nor zero, instead starting relatively high right after the event and exponentially "fading" toward zero: $p_t^{t'} = \lambda^{t'-t}$ for $\lambda \in (0,1)$. It is assumed each observation is independent from the others, but some results are robust to correlation between observations (see Remark 1.2.1).

This can be interpreted as roughly reflecting how word-of-mouth spreads. It is not certain that customers hear about previous experiences, nor is it certain that they do not, but it is more likely that they hear about recent history than the distant past.

I first show the existence of a myopic equilibrium (analogous to Theorem 1.2.3) for fading history when λ is below a threshold. One striking feature is that the upper bound on λ does not depend on μ^0. This is because the proof does not rely on calculating beliefs. The action at some period t affects the payoff at some later period $t' > t$ only if either t' observes t directly or there exists some sequence $(t_1, ..., t_n)$ such that motorist t_1 observes t, motorist t_j observes t_{j-1} for all $j \in \{2, ..., n\}$, and motorist t' observes t_n. The proof bounds the probability of such "observation chains." Of course, the motorists in these chains would also have to change their action in response to their observation to affect the payoff at t' (so the bound is not as high as it could be), but the fact that this technique ignores the actual beliefs allows using essentially the same technique for fading history in other games.

This technique is impossible in the limited history environment because such a chain always connects every period together even for $T = 1$ (period 1 observes period 0, period 2 observes period 1, and so on).

This result is given for a more general set of games — those where the long-run player has a strictly dominant action in the stage game — and is also stronger than Theorem 1.2.3 because it shows that *every* equilibrium has myopic behavior by the long-run player. Since the stage games considered are extensive-form, this strict dominance is needed at decision nodes for the long-run player; I define this notion as "strictly conditionally dominant."[9]

Definition 1.2.3. Let an extensive-form game between players 1 and 2 be given with finite action spaces A_1 and A_2. The payoff for player i of action profile $(a_1, a_2) \in A \equiv A_1 \times A_2$ is denoted $u_i(a_1, a_2)$. Let the set of player 2 actions which lead to a decision node for player 1 be denoted $\tilde{A}_2 \subset A_2$. An action $a_d \in A_1$ for player 1 is *strictly conditionally dominant* if and only if $u_1(a_d, a_2) > u_1(a_1', a_2)$ for all $a_1' \in A_1 \backslash \{a_d\}$ and all $a_2 \in \tilde{A}_2$.

Restricting attention to stage games with strictly conditionally dominant actions gives a positive lower bound for the current period benefit of playing myopically. Even less restrictive assumptions may well be possible, but the assumptions of Theorem 1.2.4 suffice for the games considered in this chapter, and many other participation and simultaneous-move games.[10]

Theorem 1.2.4. *Consider any infinitely repeated reputation game between long-run player 1 and a different short-run player 2 each period, with fading history specified by $\lambda < 1/(2\delta)$. Player 1 is either a rational type θ_0 or one of $N \in \mathbb{N}$ commitment types $\theta_1, ..., \theta_N$, with each short-run player having prior beliefs $\mu^0(\theta)$ on each of the types. Suppose that rational*

[9] The definition used here is equivalent to a unique strategy that is not "conditionally dominated" under Definition 4.2 in Fudenberg and Tirole (1991), who consider iterated conditional dominance as a solution concept for extensive-form games.

[10] Indeed, the argument behind the proof of Theorem 1.2.4 holds even without the presence of commitment types. Of course, the absence of commitment types also means the absence of useful reputation (or any reputation for that matter), which is the motivation for this result.

player 1 has action $a_d \in A_1$ which is strictly conditionally dominant in the extensive-form stage game. Define

$$z \equiv \max_{a \in A} u_1(a) - \min_{a \in A} u_1(a) \qquad \textit{and} \qquad z_d \equiv \min_{(a_1', a_2) \in (A_1 \setminus \{a_d\}) \times \tilde{A}_2} \{u_1(a_d, a_2) - u_1(a_1', a_2)\}.$$

If

$$\lambda < \frac{z_d}{\delta(z + 2z_d)}, \tag{1.2.5}$$

then any sequential equilibrium has rational player 1 playing a_d at every history.

In the case of the mechanic game, $z = z_d = u + w$, so the upper bound (1.2.5) on λ is $1/(3\delta)$; for $\lambda = \frac{1}{3}$ this corresponds to a customer sharing her experience with an average of $\frac{1}{3} + \frac{1}{3^2} + \cdots = \frac{1}{2}$ future customers. Note that this upper bound does not require Assumption 1. For the mechanic game, Appendix 1.2.2 gives an upper bound between $2/(5\delta)$ and $1/(2\delta)$ (depending on the ratio w/u), corresponding to a customer talking to an average between $\frac{2}{3}$ and 1 future customer (for δ close to one), using Assumption 1 and an intuitive restriction on equilibria, suggesting (1.2.5) can generally be improved upon in applications to specific models.

The following example illustrates why the upper bound on λ is sufficient for precluding an "echo" in the mechanic game like that in Theorem 1.2.2.

Example 1.2.1. Let $\lambda = \frac{1}{3}$. If the motorist hires at period 0, the difference in stage payoffs between doing the right repair and the wrong repair is $u + w$. The probability that motorist 1 observes the repair at period 0 is $\frac{1}{3}$. The probability that motorist 2 observes period 1 is $\frac{1}{9}$, and the probability that motorist 2 observes period 1 and period 1 observes period 0 is $\frac{1}{9}$, so the probability of a "chain" of observations between period 0 and period 2 is bounded by $\frac{2}{9}$. The probability that period 3 observes period 0 is $\frac{1}{27}$, the probability that 3 observes 1 and 1 observes 0 is $\frac{1}{9} \cdot \frac{1}{3} = \frac{1}{27}$, and the probability that 3 observes 2 and that period 0 "reaches" period 2 via an observation chain is less than or equal to $\frac{1}{3} \cdot \frac{2}{9} = \frac{2}{27}$, so the probability that such a chain exists between period 0 and period 3 is bounded by $\frac{4}{27}$. This pattern continues so that the probability of a chain from period 0 to period t is bounded from above by $2^{t-1}/3^t$. The maximum difference the period 0 repair can make in the stage payoff at any future period that has a chain of observations back to period 0 is $u + w$,[11] so the discounted sum over these differences is

$$\sum_{t=1}^{\infty} \delta^{t-1} \frac{2^{t-1}}{3^t}(u+w) = \frac{u+w}{3} \cdot \frac{1}{1-\frac{2\delta}{3}} = \frac{u+w}{3-2\delta},$$

which is less than the period 0 benefit of doing the right repair $(u + w)$. Thus, doing the right repair is the only best response.

Remark 1.2.1. Though this section generally assumes observations are independent, Theorem 1.2.4 (and Corollary 1.2.1) are robust to correlation between observations. This is because the proofs rely on Boole's inequality to bound the probability of the current repair being "chained" to any particular future period and then use the expected discounted sum of the effects of these chains at each future period for a bound on the continuation value. So long as the *probabilities* of these observations satisfy the "fading" definition, this correlation does not affect Theorem 1.2.4.

Correlation between observations can be interpreted as certain consumers being more connected to each other than others (a network of friends may be more likely to offer advice to each other than to strangers), and it does not have to be interpreted as decentralized communication. For example, consider messages posted on a centralized medium like an online forum, where it is visible to future customers while on the front page but with probability λ it disappears from view because other unrelated messages have pushed it off the front page.[12] Others may publish replies underneath the post, sharing their own experiences; when the original post is pushed out of view, so are all these replies. In this case, if consumer t's post disappears at the end of period $t' > t$, then all consumers $t + 1, ..., t'$

[11] Intuitively, one would expect only a difference of only u because that is the maximum decrease in the stage payoff going from being hired to not being hired (the difference between the maximum payoff and the minmax payoff). Corollary 1.2.1 shows this intuition holds given some natural restrictions, yielding a higher upper bound for λ.

[12] This, of course, ignores the possibility of searching through old posts, so this online forum model is more appropriate for situations where consumers spend little time researching and casually check a forum to see what others recommend (or warn against).

will see his message, but none of the consumers after period t' do. The ex ante probability that period t is observed by period $t+k$ is still λ^k, which is sufficient for the bounds used by Theorem 1.2.4 to ensure myopic long-run player behavior.

The type of equilibrium described by Theorem 1.2.4 is similar to that of Theorem 1.2.3 in that the mechanic always does the correct repair, but what differs is that the mechanic is not always hired, even when good. Equilibrium behavior of the motorists would be unique except for the possibility that the posterior at some observable subhistory $\hat{h}^t$ makes the motorist indifferent ($\mu^t(\hat{h}^t) = p^*$) and therefore allows mixing at this belief.

Because motorists sometimes receive information that is useful for avoiding the bad mechanic, the ex ante payoff for all but the first few motorists is strictly greater in this equilibrium than under Theorem 1.2.3. Motorist t does not hire if $\mu^t(\hat{h}^t) > p^*$ (where the observable subhistory $\hat{h}^t$ is the (random) set of observations from periods that motorist t sees). Denoting the payoff of motorist t given $\hat{h}^t$ as

$$v^t_{SR}(\hat{h}^t; s) = \begin{cases} \frac{u-w}{2}\alpha^t(\hat{h}^t) & s = b \\ u\alpha^t(\hat{h}^t) & s = g, \end{cases}$$

where $\alpha^t(\hat{h}^t)$ is the probability of hiring given $\hat{h}^t$, the ex ante payoff of the motorist is

$$\begin{aligned} E[v^t_{SR}(\hat{h}^t; s)] &= \sum_{\hat{h}^t \in \hat{\mathscr{H}}^t} P(\hat{h}^t) E[v^t_{SR}(\hat{h}^t; s) | \hat{h}^t] \\ &= \sum_{\hat{h}^t \in \hat{\mathscr{H}}^t} P(\hat{h}^t) \left[P(s = b | \hat{h}^t) \frac{u-w}{2} + P(s = g | \hat{h}^t) u \right] \mathbf{1}\{\mu^t(\hat{h}^t) \leq p^*\} \end{aligned}$$

where $\hat{\mathscr{H}}^t$ is the space of observable subhistories at period t, and the term in brackets is the expected payoff of hiring at $\hat{h}^t$, which is negative when $\mu^t(\hat{h}^t) > p^*$ (when $\mu^t(\hat{h}^t) \leq p^*$, hiring is always a best response because the mechanic does the right repair). If there exists any $\hat{h}^t$ such that $P(\hat{h}^t) > 0$ and $\mu^t(\hat{h}^t) > p^*$, then

$$E[v^t_{SR}(\hat{h}^t)] = \sum_{\hat{h}^t \in \hat{\mathscr{H}}^t} P(\hat{h}^t) \left[\mu^t(\hat{h}^t) \frac{u-w}{2} + (1 - \mu^t(\hat{h}^t)) P(s = g | \hat{h}^t) u \right] \mathbf{1}\{\mu^t(\hat{h}^t) \leq p^*\}$$

$$> \sum_{\hat{h}^t \in \hat{\mathscr{H}}^t} P(\hat{h}^t) \left[\mu^t(\hat{h}^t) \frac{u-w}{2} + (1 - \mu^t(\hat{h}^t)) P(s = g | \hat{h}^t) u \right], \tag{1.2.6}$$

where the right hand side of (1.2.6) is the one-shot ex ante payoff. It is easy to see for every period $t > \bar{L}(\mu^0)$ that such an $\hat{h}^t$ exists (at a minimum, they observe the full history with only engine replacements with positive probability). Thus, the ex ante payoff for every motorist $t > \bar{L}(\mu^0)$ is greater than the one-shot payoff.

When history "fades" too slowly (i.e. λ is high), the bad reputation outcome is recovered. I first present a result that is weaker than Theorem 1.2.2 because it uses a different order of taking limits: instead of holding λ fixed and letting $\delta \to 1$, it holds δ fixed and lets $\lambda \to 1$. A stronger result with the same order of limits as Theorem 1.2.2 is given at the end of the section, using a mild restriction on equilibria.

Theorem 1.2.5. *Let $\mu^0 > 0$ and $\delta > (u+w)/(2u+w)$ be given. Then for λ close enough to one, for any sequential equilibrium satisfying Assumption 1 there is a unique equilibrium outcome where the mechanic is never hired.*

The proof of Theorem 1.2.5 shows that as λ gets arbitrarily close to one, the mechanic who performs a tune-up at a critical history is hired with probability arbitrarily close to one for arbitrarily many periods, while doing an engine replacement yields arbitrarily many periods of not being hired. At some point, the "memory" of the repair will (at least directly) fade away, and this "premium" the mechanic receives for a tune-up will eventually go away (or at least the bounds used cannot rule that out). The proof does not rule out the possibility that this premium for the tune-up is eventually (at periods far in the future) replaced by an even greater premium for the engine replacement. Instead, it simply relies on λ being high enough that any such "reverse premium" is postponed long enough that it is discounted away.

Establishing a lower bound on λ independent of the discount factor clearly requires ruling out such a reverse premium, which seems intuitively implausible because it requires that motorists far into the future are somehow dissuaded from hiring because of a tune-up, rather than an engine replacement, that they never observe directly (if they observed it directly they would hire, of course, because of Assumption 1). By restricting attention to equilibria where tune-ups do not, in expectation, dissuade future motorists from hiring, a lower bound on λ independent of δ is obtained that gives the bad reputation result. A word about notation: I use μ_t (with a subscript instead of superscript t) to denote motorist t's beliefs about the mechanic's type *and* the history (as opposed to μ^t, which is simply the belief on the type).

Criterion 1. *Let a sequential equilibrium be given with strategy σ_t^* and beliefs μ_t for each motorist t. Let σ_g be any best response strategy (not necessarily the equilibrium strategy) for the mechanic, and let $\tilde{\sigma}_g^{h^t}$ be the strategy identical to σ_g except that the mechanic does a tune-up with certainty at history h^t. The equilibrium satisfies Criterion 1 if and only if doing a tune-up at h^t does not decrease the probability of being hired at any future period k given the motorists' strategies and beliefs, i.e.*

$$P(\eta_k \neq \varnothing | h^t, \sigma_g, (\sigma_{t'}^*, \mu_{t'})_{t'}) \leq P(\eta_k \neq \varnothing | h^t, \tilde{\sigma}_g^{h^t}, (\sigma_{t'}^*, \mu_{t'})_{t'}),$$

for all $k > t$, where η_k is the event at period k, for all t, h^t and σ_g.

Criterion 1 is similar in spirit to the D1 Criterion (see, for example, Section 11.2 of Fudenberg and Tirole (1991)), but it is about actions instead of beliefs. The proof of Theorem 1.2.2 shows that any equilibrium satisfying its assumptions must satisfy Criterion 1, as do the myopic equilibria of Theorems 1.2.3 and 1.2.4.

Theorem 1.2.6. *Let $\mu^0 > 0$ and $L(\mu^0)$ be given. There exists λ^* such that for any $\lambda \in (\lambda^*, 1)$, for all sequential equilibria satisfying Assumption 1 and Criterion 1, there is a unique equilibrium outcome where the mechanic is never hired for δ close enough to one.*

1.3 The Chain Store Game

Selten's (1978) chain store game, depicted in Figure 1.3.1, is a typical example of a Stackelberg-type game that has been widely used to study the effects of reputation. In

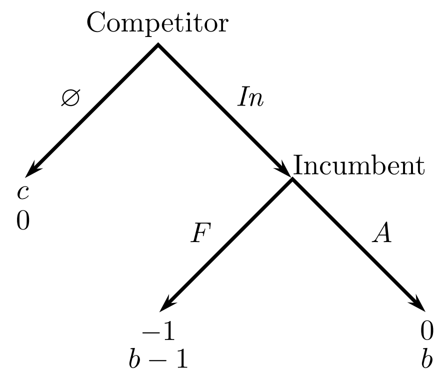

Figure 1.3.1: The chain store stage game, with payoffs at each node for the incumbent on top and for the competitor on bottom.

light of the results of Section 1.2, this section considers the infinitely repeated chain store game where there is probability μ^0 that the incumbent is a "tough" commitment type that plays F every time entry occurs and probability $1-\mu^0$ that the incumbent is a normal or "weak" (rational) type.

The classic result of Fudenberg and Levine (1989) shows that when the full history is observed, there is a lower bound on long-run player payoffs that approaches the Stackelberg payoff (which in this game is c) as δ approaches one. Given prior belief μ that the incumbent is tough, with full history the posterior must increase every time the incumbent fights (so long as she has not acquiesced in the past) to at least

$$\Upsilon(\mu) \equiv \frac{\mu}{\mu + (1-\mu)b}, \tag{1.3.1}$$

because the weak incumbent must play A with at least probability $1-b$ for entry to be a best response by the competitor. Letting $\Upsilon^1(\mu) \equiv \Upsilon(\mu)$ and $\Upsilon^k(\mu) \equiv \Upsilon(\Upsilon^{k-1}(\mu))$ be defined inductively (similarly to Section 1.2), then there can be at most $L(\mu^0)$ periods with entry on the equilibrium path if A is never played, for

$$L(\mu^0) \equiv \min t \text{ such that } \Upsilon^t(\mu^0) > b.$$

Let $\underline{v}_I(\mu^0, \delta)$ be the infimum over the set of the incumbent's payoff in any Nash equilibrium given μ^0, δ. Fudenberg and Levine (1989) establish the following lower bound:

$$\underline{v}_I(\mu^0, \delta) \geq \delta^{L(\mu^0)} c - (1 - \delta^{L(\mu^0)}). \tag{1.3.2}$$

The intuition of their result is that the incumbent can always play a (possibly deviation) strategy of playing F the first $L(\mu^0)$ times there is entry, thereby raising the belief above the critical value p^*, at which the competitor would be indifferent between In and $\varnothing$ if the weak incumbent plays A. This precludes any further entry and gives a payoff that is at least the right hand side of (1.3.2).

The results of Section 1.2 suggest that when history is more "transparent" (meaning short-run players likely see lots of history), equilibrium behavior is similar to that of the full history case, and when it is more "opaque" (short-run players likely see little history) a myopic equilibrium exists where the long-run player plays as though reputation did not exist. I find similar outcomes for the fading history chain store game for both the transparent and opaque cases. By contrast, the limited history chain store game has crucial differences that prevent application of the same techniques used in Subsection 1.2.1. I discuss the nature of these differences and show that a myopic equilibrium for limited history cannot exist, no matter how short the memory is.

1.3.1 Limited History

As in Subsection 1.2.1, suppose that the short-run players observe only the past T periods. It may seem that the techniques employed in the mechanic game could be used for the chain store game to get analogous results, but such a straightforward application is not possible.

The difficulty is that for large T in the mechanic game, participation is only informative to the extent that if motorists $t > T$ see hiring in all observable periods, they know the mechanic is good. If they do not see hiring every period, they know the mechanic is too likely to be bad to be hired (their posterior is greater than p^*). Thus, hiring cannot "subtly" signal the mechanic's type.

The analogue of Assumption 1, that entry occurs whenever competitors know the incumbent is weak, would be very useful here, but it cannot be justifiably assumed in the chain store game, so using the same method for ruling out subtle informativeness of participation is not reasonable.[13] A way of circumventing this problem is only considering strategies involving pure actions for the competitors, so that given a subhistory, a competitor will choose either $\varnothing$ or In with certainty. This makes the equilibrium path for a tough incumbent deterministic, allowing the use of $\Upsilon(\cdot)$ as a lower bound on the partial posterior following an F event and giving a lower bound on incumbent payoffs similar to Fudenberg and Levine (1989)'s (1.3.2), but it is not clear that allowing mixed actions also gives such a bound.

Since the arguments of Subsection 1.2.1 do not carry over directly, the dynamics of the limited history chain store game with reputation remain unclear, and a full analysis is beyond the scope of this chapter. However, there do appear to be qualitative differences, at least for small T.

Because of Theorem 1.2.3, one might expect the existence of an equilibrium where the one-shot equilibrium (A, In) is played every period for small T. Surprisingly, this is not the case, even if $T = 1$. To see why, suppose by contradiction such an equilibrium exists. Consider a deviation by the incumbent of playing F in period 0. Competitor 1 must then believe the incumbent is tough with probability 1 and must play $\varnothing$ as a best response at that history. This means that competitor 2 also believes the incumbent is tough with probability 1 (even if $T = 1$) because she observes $h_1 = \varnothing$ and $P(h_1 = \varnothing | \, Weak) = 0$, so she also does not enter, and so on for every subsequent period. This means that the continuation payoff for playing F in period 0 is c, while the continuation payoff for playing A is 0. For δ close enough to one, $-(1-\delta) + \delta c > 0$, so A is not a best response, a contradiction. Thus, there

[13]The problem the good mechanic and the motorists face is simply the temptation to signal, which is why the mechanic game has a unique equilibrium if uncertainty about the mechanic's type is removed. Assumption 1 says that if the mechanic's type is revealed in the middle of a repeated game that began with uncertainty, play then proceeds as if there had never been such uncertainty because all players are better off, i.e. renegotiation-proofness. There is no such coincidence of interests in the chain store game.

is no result for the limited history chain store game equivalent to Theorem 1.2.3, so the incumbent's payoff is bounded away from 0 (the one-shot payoff as well as minmax payoff) due to reputation, even if history is as limited as possible without eliminating the history observation entirely.

Using their stage game, Liu and Skrzypacz (2014a) point to the existence of an equilibrium under limited history (not limited records) where the long-run player always mimics the commitment type on the equilibrium path with a simple grim trigger threat of Nash reversion. A similarly simple grim trigger equilibrium does not exist for the chain store game, but a more complicated equilibrium without entry on the equilibrium path exists in the $T = 1$ case, constructed in Appendix 1.3.1. This equilibrium requires mixing off the equilibrium path such that the incumbent always be indifferent between fighting and acquiescing. By contrast Liu and Skrzypacz's limited history equilibrium does not require such indifference by the long-run player (nor do the equilibria of Theorems 1.2.2 and 1.2.3). This similarity on the equilibrium path and dissimilarity off of it raise interesting questions about how limited history affects reputation games more generally.

1.3.2 Fading History

This subsection assumes the fading history assumed in Subsection 1.2.2 specified by λ. Unlike limited history, the intuition of fading history largely carries over to the chain store game. Theorem 1.2.4 applies directly to the chain store game, and Theorem 1.2.5 uses arguments that do not rely crucially on the specifics of the mechanic game, so an analogous result in the chain store game can be found using these techniques. (Theorem 1.2.6 does rely crucially on the particulars of the mechanic game and also on Assumption 1, for which there is no justifiable analogue for the chain store game, so I do not attempt a similar result here.)

For low λ, Theorem 1.2.4 shows that in any equilibrium the weak incumbent always acquiesces. In this case, $z = c-1$ and $z_d = 1$, so the upper bound (1.2.5) on λ is $1/[\delta(c+1)]$. As with the mechanic game, reputation in these equilibria increases the ex ante payoffs of

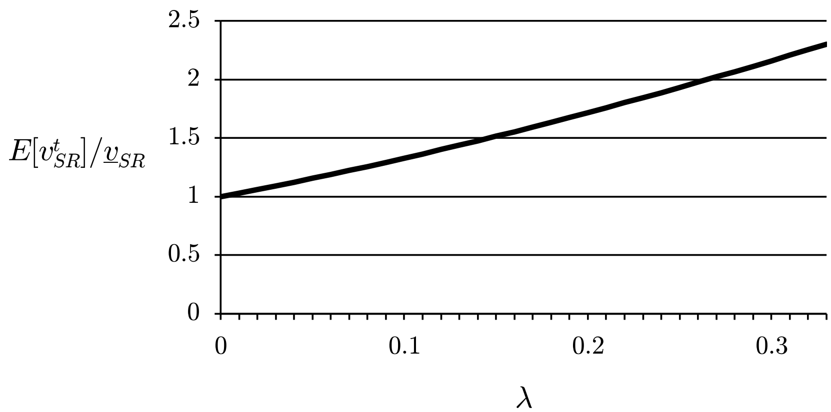

Figure 1.3.2: The ratio of the ex ante payoff $E[v_{SR}^t]$ for competitor $t = 20$ to the one-shot payoff $\underline{v}_{SR}$ in any chain store game equilibrium with $\mu^0 = \frac{1}{5}, b = \frac{1}{4}$ (note that $E[v_{SR}^t] = \underline{v}_{SR}$ at $\lambda = 0$) is plotted for values of $\lambda \in [0, \frac{1}{3}]$, which satisfy (1.2.5) for $c = 2$ and any δ. These payoffs do not significantly change for periods past 20 (because $\psi(\lambda, 20) \approx \psi(\lambda, \infty)$).

the short-run players above that of the one-shot game. The increase here is more dramatic than in the mechanic game because a competitor need only observe one previous period to know if the incumbent is tough (all periods will be either F or $\varnothing$) or weak (all periods are A). Thus, the ex ante payoff of competitor t is

$$E[v_{SR}^t] = \psi(\lambda, t)(\mu^0(b-1) + (1-\mu^0)b) + (1 - \psi(\lambda, t))(1-\mu^0)b,$$

where $\psi(\lambda, t) \equiv \prod_{k=1}^{t}(1-\lambda^k)$ is the probability of competitor t observing no history at all, which is strictly greater than the one-shot payoff for all competitors except at period 0.[14] This is straightforward to calculate and plotted in Figure 1.3.2 for some example parameters.

Finally I find a lower bound on incumbent payoffs similar to (1.3.2) for high λ.

Theorem 1.3.1. *Let $\mu^0 > 0$, $L(\mu^0)$ and any $\varepsilon > 0$ be given. Then there exists $\lambda^* \in (0,1)$ such that for any $\lambda \in (\lambda^*, 1)$, the infimum $\underline{v}_I$ of the set of incumbent payoffs in any sequential equilibrium satisfies*

$$\underline{v}_I(\mu^0, \delta, \lambda) \geq \delta^L c - (1-\delta^L) - \varepsilon.$$

[14] $\psi(\lambda, t)$ converges absolutely as $t \to \infty$ to a value in the set $(0, 1)$ when $\lambda \in (0, 1)$ (Apostol, 1976).

1.4 Conclusion

For the mechanic game with limited history, reputation is bad when short-run players have a long enough memory T. This is because early events that tarnish the mechanic's reputation "echo" for all following periods through the refusal of subsequent motorists to hire, which is observed by the following motorists who consequently also refuse to hire, and so on. When T is small enough, a myopic equilibrium exists, where the good mechanic always plays her stage game dominant strategy (doing the correct repair). This equilibrium avoids the bad reputation result at the expense of making reputation irrelevant — motorists never see enough information to change their hiring decisions. In summary, for limited history, equilibrium behavior is the same for small T as the one-shot game, and the same for large T as the full history game. This differs qualitatively from the cyclical behavior for limited records found by Liu and Skrzypacz (2014a).

Under fading history, when λ is less than a critical value, an equilibrium with myopic behavior by the mechanic exists, but reputation still has an effect — sometimes motorists are informed enough that they do not hire. This increases the short-run players' ex ante payoffs. The result holds generally for reputation games where the long-run player has a strictly dominant action in the stage game: when λ is less than a critical value, the long-run player's equilibrium strategy is always to play the dominant action. This is because fading history bounds the probability of an "observation chain" from the current period t to future period $\hat{t}$, where $t' > t$ observes t, $t'' > t'$ observes t', etc., which bounds the reputational payoffs of any signaling strategy. By contrast, such a chain always exists in limited history, even when $T = 1$, because period 1 always observes 0, period 2 always observes 1, etc. For high λ, the bad reputation result is recovered. The result for fading history with small λ applies directly to the chain store game, leading to a more dramatic increase in short-run player payoffs because they need only observe one past period to learn the long-run player's type.

Equilibria under limited history seem qualitatively different in the chain store game versus the mechanic game; in particular, no myopic chain store equilibrium exists like that of the mechanic game for small T. The folk theorems of Mailath and Olszewski (2011a) and the purifiability result of Bhaskar, Mailath, and Morris (2013a) offer intriguing clues for an investigation. An interesting question is whether limited history equilibria in Stackelberg-type games (like the chain store game) can exhibit the cyclical behavior under limited records found by Liu and Skrzypacz or have the non-cyclical behavior of the mechanic game. More generally, behavior in other limited history reputation games remains largely unknown and is an interesting topic for future research.

Chapter 2

Bounded Memory, Reputation, and Impatience

2.1 Introduction

Consider a market where a seller faces a sequence of different buyers. The buyers choose how much to trust the seller (that is, how large of an order to place), and the seller chooses whether to honor that trust by incurring a cost to provide a high quality good, or instead betray it with low quality; the one-shot outcome is low trust and low quality. Because the seller never faces the same buyer again, any incentive to provide high quality must come from the threat of punishment by future buyers informed about today's action. Hence, markets often maintain records of past performance to counteract the myopic temptation to exploit.

Most reputation models assume that the full history of past behavior is observable, yet in many real world settings this glimpse into the past goes only so far. For example, events on credit histories are deleted after a certain time period in many countries, as are infractions on driving records; workers typically provide only recent references to prospective employers when applying for jobs; and many online markets display only recent reviews of sellers.[1] Even when the full list of reviews is available, online markets often make the most recent ones most prominent on their website (e.g. eBay); it may be safe to assume buyers simply glance at a few of the latest reviews instead of reading all of them. Recently the Court of Justice of the European Union ruled in *Google v. Costeja* that individuals have a "right to be forgotten" and may demand that search engines remove links to certain old

[1] A number of these examples are pointed out by Liu and Skrzypacz (2014b), who also study a bounded memory reputation environment.

information "in light of the time that has elapsed." What happens when agents know that today's behavior will some day be forgotten?

The first half of this chapter introduces general tools to study such environments. The primary contribution is a recursive characterization of the set of equilibrium payoffs for a general class of bounded memory games with incomplete information (multiple player types), presented in Section 2.3. This dynamic programming method allows analysis of both stationary and non-stationary equilibria for the first time. Section 2.4 introduces an equilibrium refinement I call "quasi-Markov perfection" (an extension of the standard notion of Markov perfection from complete information games), which rules out some fragile equilibria that are not "purifiable," meaning they do not survive the addition of small, independent private shocks to payoffs.

The second half of the chapter (Section 2.5) demonstrates these tools in two applications. This first is a product choice game (between a firm and a sequence of consumers) with a Stackelberg ("honest") commitment type and 1-period memory of the firm's actions, where the recursive method yields a complete characterization of the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on the commitment type, showing that allowing non-stationary equilibria expands the set of equilibrium payoffs. The second application looks at the same game with very long memory, where the dynamic programming state space grows very large and so studying non-stationary equilibria is difficult. Fortunately, stationary equilibria have a very simple interpretation in my framework, which is used to show that when memory is sufficiently long, the firm receives exactly the Stackelberg payoff in all purifiable, stationary equilibria, given any fixed discount factor (above a threshold dependent on the stage game payoffs) and a positive prior. Both results show that introducing even very little incomplete information has a big impact on the equilibrium set; a difference from previous results is showing that this is true even when the long-run player is not particularly patient.

The recursive characterization builds on the dynamic programming methods of Abreu, Pearce, and Stacchetti (1990) (hereafter APS) and Doraszelski and Escobar (2012) (hereafter DE). For the complete information case (e.g., the seller's type is known), APS characterize the equilibrium set for full memory, while DE characterize it for bounded memory; in both full and bounded memory, the recursive structure of these games allows the set of equilibrium payoffs to be calculated as the largest fixed point of a "generating (set) operator" that transforms sets of objects containing payoffs;[2] this fixed point is the "largest self-generating set." However, the assumption of complete information is restrictive: even a slight relaxation can dramatically change the equilibria (as shown in the applications).

I extend these techniques to incomplete information ("reputation") under bounded memory. More formally, I characterize the set of weak Perfect Bayesian Equilibrium (wPBE) payoffs of repeated games under imperfect monitoring where a long-run player, who is one of finitely many commitment types, faces a sequence of short-run players who observe only the K most recent periods. This presents two main challenges. The first is that in addition to playing best responses, players must also form beliefs consistent with Bayes' rule. I show how these games also have a recursive structure, allowing the construction of an analogous generating operator that, roughly speaking, transforms sets of objects containing both payoffs and beliefs. This also yields an algorithm for computing the largest fixed point by repeatedly applying the generating operator. The second challenge is that the first K periods (where players still see the full history) are qualitatively different from later periods, and so the largest fixed point of the generating operator (largest self-generating set) does not directly give the equilibrium payoffs. Instead, the full game's equilibrium payoffs are found by solving for certain equilibria of a set of finitely repeated K-period games, with payoffs augmented according to this fixed point.

This framework is necessary for studying non-stationary equilibria, where strategies may depend on the calendar date. Previous papers studying bounded memory reputation assume stationary strategies, which requires hiding the date from short-run players.[3] Though this makes the analysis much simpler, there are a variety of real-world applications

[2]In APS, the objects *are* payoffs. This is discussed in greater detail on page 37 and in Section 2.3.

[3]This is because although the strategies are time-independent, the beliefs and therefore best responses may not be.

where short-run players know the time: creditors know the age of borrowers even when credit histories are bounded, auto insurers know the age of drivers, and buyers can observe the age of a seller's account on eBay. The framework enables us to explore the impact of assuming stationarity (the application to a 1-period memory example shows that assuming stationarity is restrictive for some priors). Nevertheless, for long memory, this dynamic programming method becomes increasingly intractable due to the curse of dimensionality, and so solving for non-stationary equilibria may not be possible in practice. Stationary equilibria have a particularly simple interpretation in this context as "self-generating *points*" rather than "self-generating sets;" that is, instead of the more complicated task of finding the largest set of many points that generates itself, computing stationary equilibria means searching for individual points which generate themselves (I use this method in the long-memory example).

To simplify application of the recursive framework, I introduce the notion of *quasi-Markov perfection*. For complete information dynamic games, attention is often restricted to Markov perfect equilibria, where players do not condition on payoff-irrelevant histories, instead conditioning only on the payoff relevant "Markov state." Quasi-Markov perfection naturally extends this notion to the incomplete information environment.[4] It particularly simplifies games with one commitment type and perfect monitoring of the long-run player's actions. To support the argument that the simplicity of quasi-Markov equilibria does not come at the expense of realism, I show that all non-quasi-Markov equilibria are "fragile," meaning they are not purifiable in the sense of Harsanyi (1973) because there are no nearby equilibria if we add small, independent (across actions and time) private shocks to the payoffs. That is, even very tiny private payoff information destroys all non-quasi-Markov equilibria. This result is an extension of Bhaskar, Mailath, and Morris (2013b) (hereafter BMM), who show that Markov perfection is a necessary condition for purifiability in a general class of complete information, sequential-move games with bounded memory. As BMM

[4] I use the term "quasi-Markov" instead of "Markov" to distinguish it from the use of beliefs as Markov states, as is often done in the literature (see Section 18.4.2 of Mailath and Samuelson (2006b) for an example). Using beliefs as states is too coarse for the results presented here, as discussed in Section 2.4.

argue, purifiability can be motivated by the notion that games are only approximations of reality and so real payoffs are generally at least slightly different from the model.

These tools are then applied to a repeated sequential-move product choice game with a sequence of short-lived consumers, who first choose between either a small or large order, facing a long-lived firm, choosing between providing low and high quality. The firm is either a "normal" strategic type (with a myopic incentive to exploit) or an "honest" Stackelberg commitment type always providing high quality. The K most recent firm's actions are observed but those of the consumers are not.[5]

To put the application results in context, what does the existing literature tell us about this game? The complete information case (only the strategic type) is well understood. Under full memory, cooperation is simple to achieve via grim trigger strategies. In fact, all payoffs between the low one-shot equilibrium payoff and the high Stackelberg payoff are achievable with full memory (using arguments from Fudenberg, Kreps, and Maskin (1990)). However, imposing a bound on memory is disastrous. BMM show that any bound K, no matter how high K is, means all cooperative equilibria are fragile (non-purifiable); the only purifiable equilibrium is the repeated one-shot equilibrium.

The literature has less to say about the incomplete information case (adding the honest type). For full memory, the standard reputation result (due to Fudenberg and Levine (1989; 1992) and improved by Gossner (2011)) is that the firm is guaranteed a payoff close to the high Stackelberg payoff when very patient (the discount factor $\delta \to 1$). The intuition is that by persistently playing "honestly," the firm could eventually convince consumers to expect honest behavior (high quality), thereby guaranteeing a payoff close to the honest (Stackelberg) payoff when sufficiently patient. Such full memory games are difficult to solve beyond such bounds. With K-period memory, it is possible to achieve a similar bound via similar arguments, by making memory long and then making the firm very patient (i.e., "$\lim_{\delta\to 1} \lim_{K\to\infty}$"). For a similar product choice game, Liu and Skrzypacz (2014b) improve

[5] This type of monitoring is called "limited records" by Liu and Skrzypacz (2014b).

on this with a time-independent bound on stationary equilibrium payoffs. However, in both cases the order of limits is crucial: in reverse order ("$\lim_{K\to\infty}\lim_{\delta\to 1}$"), these arguments provide no meaningful bound. I also know of no papers that bound the payoffs for small K.[6]

The first application considered is the 1-period memory case, which is appealing because it is both the most restrictive limit on memory possible and yields a simple state space in which to apply the recursive algorithm, yielding a complete characterization of the minimum and maximum quasi-Markov equilibrium payoffs for all prior beliefs on the honest type and almost all discount factors.

The analysis yields several insights. First, this technique obtains the *actual* minimum and maximum payoffs rather than lower and upper bounds because it relies on the convergence of the recursive algorithm instead of the traditional argument bounding the payoff of repeatedly playing the commitment action (which may not be an equilibrium strategy). Second, even a little bit of incomplete information (a small but positive prior belief on the honest type) resurrects non-fragile cooperation, allowing a purifiable equilibrium with the Stackelberg payoff.[7] Third, assuming stationary equilibria is restrictive; that is, allowing non-stationary equilibria expands the set of equilibrium payoffs. For a range of priors, the minimum payoff is not given by stationary equilibria, but rather by equilibria where players and beliefs condition on the time in periodic cycles. When consumers have a sufficiently high prior on the honest type, the maximum payoff (higher than the Stackelberg payoff) is given by a non-stationary equilibrium where strategies have a two-period cycle. In even periods (starting with period 0), the firm exploits "naive" cooperative customers; in odd periods, customers know the firm is not honest, but still cooperate knowing the firm will provide them high quality in order to exploit the next customer in the following even

[6]In their version of the product choice game, Liu and Skrzypacz (2014b) show that equilibrium *behavior* changes, but do not show how the payoffs change for small K. This appears to be because their continuous action spaces allow qualitatively different behavior whose corresponding payoffs are more difficult to calculate.

[7]Although the general result shows only that purifiability implies quasi-Markov perfection (rather than the converse), I show in Appendix 2.3.2 that these minimum and maximum payoffs correspond to purifiable equilibria for almost all priors.

period.

The second application studies the effect of K growing large. For long memory, the state space for the algorithm becomes intractably large, so I restrict attention to stationary equilibria by finding self-generating points instead of self-generating sets. When the memory is long enough (K exceeds some threshold dependent on the prior), the long-run player receives exactly the Stackelberg payoff in any stationary, purifiable equilibrium when the discount factor is above a bound that depends only on the stage game payoffs (not the prior). Imposing purifiability shows that reputation effects are even stronger than the previous literature suggests, since the standard "patience lower bound" on equilibrium payoffs allows the possibility that the complete information game is robust to slightly incomplete information so long as the long-run player is not extremely patient. In this game, less than extreme patience is not enough to allow low payoffs — the memory must also be sufficiently short (or totally unbounded).

This work relates to a variety of papers on repeated games and reputation. As mentioned above, the recursive characterization is closest to DE, who extend the APS tools (for full memory in complete information games) to equilibria where players condition only on summary statistics of the histories — bounded memory is a special case of this.[8] To expand on the previous discussion slightly more formally (detailed discussion is saved for Section 2.3), for full memory APS show that the equilibrium payoffs are given by the largest self-generating set of *payoff vectors* (with a payoff for each long-run player). For bounded memory, DE show the equilibrium set is given by the largest self-generating set of vector-valued *payoff functions* (mapping histories to payoff vectors). The extension to incomplete information shows that the appropriate notion is self-generating sets of objects I call *HBPs*, each containing a *History distribution*, *Belief mapping*, and *Payoff function*. A history distribution is a vector of probability distributions on the space of (bounded) histories

[8]The APS framework has been extended to a variety of other settings in the literature; a few examples are Atkeson (1991) and Phelan and Stacchetti (2001), who study games with dynamic payoff relevant states, and Ely, Hörner, and Olszewski (2005), who characterize belief-free equilibrium payoffs in games with private monitoring.

conditional on each type of long-run player; the belief mapping, which maps beliefs on those types to each history, is only necessary for histories not in the support of the history distribution (i.e. off the hypothetical equilibrium path). The payoff function serves the same purpose as in DE — to keep track of continuation payoffs while breaking dependence on past play beyond the bounds of the memory.

Other work on bounded memory and complete information under perfect monitoring includes Barlo, Carmona, and Sabourian (2009), who prove a folk theorem for 1-period memory with rich action sets, and Mailath and Olszewski (2011b), who prove a folk theorem for bounded memory strategies. It is worth noting that bounded memory with complete information is effectively a restriction on strategies, while bounded memory with incomplete information is a restriction on learning as well. For complete information, all bounded memory equilibria are also full memory equilibria; for incomplete information, this is not true. Under incomplete information, Monte (2013) uses the term "bounded memory" in a different sense, but his result shows how limits on learning can lead to qualitatively different equilibria.[9] Ekmekci (2011) also studies a product choice game under incomplete information, constructing a finite rating system that translates the history into a rating observed by short-run players. Both papers show how restricted learning can result in permanent reputation even under imperfect monitoring, in contrast to the "temporary reputation" result of Cripps, Mailath, and Samuelson (2004).

The product choice game application of Section 2.5 is closely related to that of Liu (2011b) and particularly Liu and Skrzypacz (2014b), who both study stationary equilibria (calendar dates are unobserved) in product choice type games. Liu (2011b) studies behavior (rather than payoffs) in a model where monitoring of past firm actions is endogenous as a costly action available to consumers; monitoring is limited by the increasing cost of obtaining older information (instead of a fixed bound), yielding "random auditing" and reputation cycles. Liu and Skrzypacz (2014b), who have a continuous stage game, assume

[9] Monte models "bounded memory" for a long-run player as a finite set of memory states, where a player's strategy is to choose an action for each state and transition rules between the states.

a fixed bound on monitoring of past firm actions (as in this chapter), showing that all (stationary) equilibria have consumers "riding reputation bubbles" by helping a firm that they know is not honest build reputation to exploit future consumers. Their focus is also on behavior, but also show a time-independent bound on payoffs which has bite for the "$\lim_{\delta\to 1}\lim_{K\to\infty}$" limit discussed above. I focus on characterizing payoffs rather than behavior, but the proofs indicate that behavior in the game studied here differs from that in Liu and Skrzypacz's model, suggesting that a continuous action space may allow substantially different dynamics.[10] Where stationarity is assumed (the long-memory case), I follow Liu (2011b) in assuming that short-run players have the improper uniform prior on the calendar date. Liu and Skrzypacz (2014b) provide results more generally for an arbitrary prior on the date, treating the improper uniform prior as an interesting and particularly tractable special case.

2.2 Model

I consider a two player sequential-move stage game G, with the infinite repetition of G denoted G^∞, starting at period 0. G^∞ is referred to as the *full game*. In keeping with (perhaps here counter-intuitive) convention, player 1 is a long-run player (who moves second) and player 2 is a short-run player who moves first, choosing an action from finite action space A_2. Player 1 observes player 2's action $a_2 \in A_2$ and then player 1 chooses action a_1 from finite action space A_1. A public signal y from finite set Y is generated according to probability distribution $\rho(y|a_2, a_1)$.

The space of action profiles is $A \equiv A_2 \times A_1$ with typical action profile a. For any finite set X, let ΔX be the set of probability distributions over X. Denote a mixed action profile as $\alpha \in \Delta A$.

Player 1 observes the full history of actions and signals (formalized in the next

[10] See Footnote 28 for a more detailed explanation. It is interesting that while reputation bubble behavior is ruled out for stationary equilibria, the 1-period memory non-stationary maximum payoff equilibrium for "naive" consumers (with very high priors) has a similar flavor, with odd-period consumers knowingly helping the non-honest firm exploit even-period consumers.

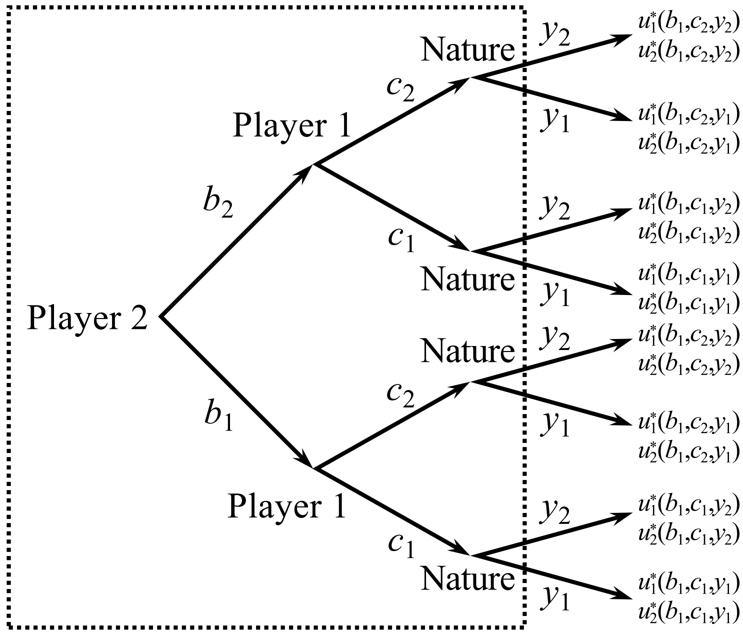

Figure 2.2.1: An example stage game with two player 2 actions $A_2 \equiv \{b_1, b_2\}$, two player 1 actions $A_1 \equiv \{c_1, c_2\}$, and two signals $Y \equiv \{y_1, y_2\}$. The dotted box is used for reference in Figure 2.3.1, enclosing the part of the game where players move rather than nature.

paragraph). Player 2 observes the K most recent public signals, but receives no other information about past play. Player i receives ex-post payoff $u_i^*(a, y)$, with ex-ante payoffs given by

$$u_i(a) = \sum_{y \in Y} u_i^*(a, y)\rho(y|a).$$

The ex ante payoff given mixed action profile $\alpha \in \Delta A$ is $u_i(\alpha)$. Figure 2.2.1 depicts a simple example stage game.

The set of *full histories* at period t is $\mathbf{H}^t \equiv (A \times Y)^t$; let $\mathbf{H} \equiv \bigcup_t \mathbf{H}^t$. The focus will primarily not be on full histories (as discussed below), so I do not use the term "history" to refer to these. The set of *full semipublic histories* at some period t is $\mathcal{H}^t \equiv Y^t$ with typical element $\mathfrak{h}^t$ (superscript t is used to make clear the length of such a history), with the set of all full semipublic histories $\mathcal{H} \equiv \bigcup_{t=0}^{\infty} \mathcal{H}^t$. These are called "semipublic" because each element of the history was public at some point in the past. Denote the concatenation of full semipublic history $\mathfrak{h}^t$ and some signal y as $\mathfrak{h}^t y$.

Both players observe the date, thereby allowing non-stationary behavior.[11] Player 2 observes only the last K periods, which I call the *public history*. For period t, the set of public histories is Y^t for $t < K$ and Y^K for $t \geq K$. Since short-run players are in information sets containing a public history and the date, such pairs are called *date-histories*. Let $H^t \equiv \{t\} \times Y^{\min\{t,K\}}$ be the set of all date-histories at period t, and let $H \equiv \bigcup_{t=0}^{\infty} H^t$ denote the set of all date-histories. This chapter is primarily concerned with the public histories instead of the full histories (due to Lemma 2.2.1), so I often refer to a public history simply as a "history."

To reflect the fact that the elements of the public history h have happened in the past, I index its elements with negative indices: $h \equiv (h_{-K}, h_{-K+1}, ..., h_{-1})$. For periods $t \geq K$, at the end of each period the oldest element h_{-K} is deleted, every subsequent element is "pushed back" one space, and the newly generated signal y from the current period's play is appended. I denote this as $hy \equiv (h_{-K+1}, ..., h_{-1}, y)$ and say y is *pushed on* h.

The long-run player is one of the types in the set $\Theta \equiv \{\theta_0\} \cup \hat{\Theta}$, where θ_0 is the "normal type" with payoffs $u_1^*(a, y)$ given above, and $\hat{\Theta}$ is a finite set of "commitment types," where each $\hat{\theta} \in \hat{\Theta}$ is committed to playing a (possibly mixed) action $\hat{\alpha}_\theta \in \Delta A_1$ every period. Each player 2 has prior belief $\mu^0(\theta) \in [0, 1]$ for each type $\theta \in \Theta$, and updates those beliefs based on the date-history according to Bayes' rule.

A strategy for player 2 at period t is a mapping $\sigma_2^t : H^t \to \Delta A_2$; that is, it depends only on the date and public history, which I call a *public strategy*. For convenience, denote the vector $(\sigma_2^0, \sigma_2^1, ...)$ of all player 2 strategies as $\sigma_2 : H \to \Delta A_2$, so that $\sigma_2(t, h) = \sigma_2^t(h)$. In general, a strategy for player 1 is a mapping

$$\sigma_1 : \Theta \times \left(\bigcup_{t=0}^{\infty} \mathbf{H}^t \right) \times A_2 \to \Delta A_1,$$

[11] I discuss stationary equilibria in an alternative specification in Section 2.3.5, where player 2 does not observe the date and instead has the improper uniform prior on the date where she enters.

but because of the focus on payoffs I can and will restrict attention to only public strategies $\sigma_1 : \Theta \times H \times A_2 \to \Delta A_1$ (see Lemma 2.2.1 below). I denote the value of a strategy profile $\sigma \equiv (\sigma_1, \sigma_2)$ to player 1 as $V(\sigma)$.

Definition 2.2.1. (σ^*, μ^*) is a weak Perfect Bayesian Equilibrium (wPBE) if σ^* are mutual best responses with beliefs μ^*, and μ^* is consistent with Bayes' rule on the equilibrium path. σ^* is *stationary* if strategies at periods $t \geq K$ are independent of the calendar date.

Denote the set of strategy profiles as Σ, the set of wPBE strategy profiles as Σ^*, the set of public strategy profiles as $\hat{\Sigma}$, and the set of wPBE public strategy profiles as $\hat{\Sigma}^*$. Focus on public strategy profiles is not restrictive with respect to payoffs because of the following lemma.

Lemma 2.2.1. *Let any wPBE with strategy profile $\tilde{\sigma} \in \Sigma^*$ be given. There exists public wPBE strategy profile $\bar{\sigma} \in \hat{\Sigma}^*$ such that $V(\bar{\sigma}) = V(\tilde{\sigma})$.*

2.3 Recursive Characterization of Equilibrium Payoffs

The literature on bounded memory reputation thus far has restricted attention to stationary equilibria, where strategies depend on the public history but not the calendar date (starting at period K). In the model here, this assumption greatly simplifies the analysis by making the strategy space finite.[12] By contrast, allowing non-stationary behavior means players may also condition on the calendar date, making the strategy space infinite.

Assuming stationary equilibria generally requires hiding the calendar date from the short-run player, since the equilibrium distribution of play need not be constant through time (even with stationary strategies), so the beliefs and therefore expected payoffs (for the short-run player) also depend on time. Hiding the date ensures that beliefs are also constant through time, so the set of best responses is the same at all periods $K, K+1, ...$

[12] The short-run player chooses an action at each of the $\sum_{t=0}^{K-1} |H^t| = \sum_{t=0}^{K-1} |Y|^t$ initial histories plus the $|Y|^K$ possible histories at periods $t \geq K$, while the long-run player chooses $|A_2|$ times as many actions.

Such models must assume short-run players arriving on and after period K have a common prior on the current date, typically the improper uniform prior, as is done in Sections 2.3.5 and 2.5.2.

The problem of finding equilibrium payoffs in an infinite strategy space is resolved by transforming it into a dynamic programming problem. In full-memory, complete information repeated games — the environment of APS — the strategy space is also infinite (the space of histories is itself infinite). The key to their framework is the strategic equivalence of the full game and the continuation subgame; the strategies starting at any (full) history constitute a perfect public equilibrium (PPE), and so continuation payoffs must themselves be equilibrium payoffs.

The framework introduced here is best understood by drawing analogies between its definitions and those of APS. I start by stating the APS approach with deliberately vague language (with more concrete descriptions in parentheses) to hint at the intuition of my approach.

1. Let a set of hypothetical summary statistics of future equilibrium play (continuation payoffs) be given.

2. Given these hypothetical futures, what sort of current play (action profiles) is possible today?

3. Combine the possible current play and hypothetical futures into a new set of hypothetical summary statistics of equilibrium play (average the payoffs of the current action profile and the hypothetical continuation payoffs to get a new set of payoffs).

APS show that if the output of step 3 yields the hypothetical input in step 1, that hypothetical input describes (non-hypothetical) equilibria; this is because full memory, complete information games have a recursive structure. I show that bounded memory, incomplete information games also have a recursive structure, and so it is possible to use more complicated "hypothetical summary statistics" and "current play" to find actual equilibrium values. To assist in drawing these analogies, I provide brief sketches of APS[13] and DE before discussing the incomplete information framework in Section 2.3.2.

2.3.1 Recap of APS and DE

For simplicity, assume two long-run players 1 and 2 playing a simultaneous-move stage game with finite action spaces A_1, A_2, respectively, and restrict to pure strategies, letting $A \equiv A_1 \times A_2$. Assume perfect monitoring so the appropriate solution concept is subgame-perfect equilibrium (SPE). The stage payoff for player i is $u_i(a)$. Let $\mathscr{E} \subset \mathbb{R}^2$ be the set of pure-action SPE payoffs for each player. For any $v \in \mathscr{E}$, there exists SPE σ^v with payoffs $v = V(\sigma)$, where $V(\sigma)$ is the vector of values of σ to players 1 and 2.

Let $\mathscr{F}^\dagger \subset \mathbb{R}^2$ denote the set of feasible payoffs, and let $\mathscr{W} \subset \mathscr{F}^\dagger$ be some set of feasible but not necessary equilibrium payoffs. If we can construct a continuation payoff function $\gamma : A \to \mathscr{W}$ on the set of hypothetical (i.e. not necessarily credible) continuation payoffs $\mathscr{W}$ so that action profile a is incentive compatible for both players, the action profile a is *enforced* by γ on $\mathscr{W}$.

Definition 2.3.1. Let $\mathscr{W} \subset \mathbb{R}^2$ be given. An action profile $a \in A$ is *enforced by* $\gamma : A \to \mathscr{W}$ if

$$(1-\delta)u_i(a) + \delta\gamma_i(a) \geq (1-\delta)u_i(a_i', a_{-i}) + \delta\gamma_i(a_i', a_{-i})$$

for all i, $a_i' \in A_i$. We say that a is *enforceable on* $\mathscr{W}$ if such a function γ exists.

For a given payoff vector $v \in \mathscr{F}^\dagger$, if there exists an action profile a enforced by some γ on $\mathscr{W}$ such that the discounted average payoffs for a and the continuation payoffs $\gamma(a)$ are equal to v, i.e. for both i

$$v_i = (1-\delta)u_i(a) + \delta\gamma_i(a), \tag{2.3.1}$$

then v is *decomposable on* or "generated by" $\mathscr{W}$. APS construct the set operator $\mathscr{B}$ so that

[13] Here I follow the notation of Section 2.5.1 of Mailath and Samuelson (2006b).

$\mathscr{B}(\mathscr{W})$ is the set of all payoff vectors decomposable on $\mathscr{W}$.

Definition 2.3.2. Let $\mathscr{W} \subset \mathbb{R}^2$ be given. Define

$$\begin{aligned}\mathscr{B}(\mathscr{W}) \equiv \{&v \in \mathbb{R}^2 : \exists a \in A, \exists \gamma : A \to \mathscr{W} \text{ such that} \\ &a \text{ is enforced by } \gamma \text{ and } v_i = (1-\delta)u_i(a) + \delta\gamma_i(a)\}.\end{aligned}$$

Any set $\mathscr{W}$ that generates a superset of itself, i.e. $\mathscr{W} \subset \mathscr{B}(\mathscr{W})$, is a *self-generating set*. Every self-generating set is a subset of the equilibrium payoffs, and the set of equilibrium payoffs is the largest fixed point of $\mathscr{B}$.

Proposition 2.3.1 (Theorems 1 and 2, APS)**.** *The following holds:*

1. ***Self-generation:*** *Let any set $\mathscr{W} \subset \mathbb{R}^2$ be given. If $\mathscr{W} \subset \mathscr{B}(\mathscr{W})$, then $\mathscr{W} \subset \mathscr{E}$.*
2. ***Factorization:*** *$\mathscr{B}(\mathscr{E}) = \mathscr{E}$.*

APS also give an algorithm for computing $\mathscr{E}$, showing that repeatedly applying $\mathscr{B}(\cdot)$ to the set of feasible payoffs converges to the set of equilibrium payoffs. This algorithm is the basis of Judd, Yeltekin, and Conklin (2003), who develop a numerical implementation of Proposition 2.3.2.

Proposition 2.3.2 (Theorem 5, APS)**.** $\bigcap_{m=0}^{\infty} \mathscr{B}^m(\mathscr{F}^\dagger) = \mathscr{E}$.

DE extend the APS tools to equilibria which condition on summary statistics of past play, where bounded memory is a special case. They show that the appropriate notion is self-generating sets of vector-valued (continuation) payoff functions, rather than payoff vectors. Why?

Consider the example model depicted in Figure 2.2.1 satisfying the specification in Section 2.2, but leave out reputation by only allowing the normal type θ_0. I recycle the notation above by letting $\mathscr{E} \subset \mathbb{R}$ be the set of player 1 PPE payoffs. For expositional clarity, this discussion uses the term "decompose" informally for 1-period memory since Section 2.3.4 (which incorporates reputation) formalizes this as a special case.

When dealing with sets of payoffs instead of payoff functions, the APS framework allows the freedom to choose any equilibrium payoff as a continuation payoff for any action profile played today. Consider the "Full Game G^{∞}" part of Figure 2.3.1 (for now ignoring the "$P_{\sigma^*}(\cdots)$" notation and "Variant Game" parts). The dotted boxes indicate the player actions of each stage game, corresponding to the dotted box in Figure 2.2.1. With full memory, continuation play at the history y_1y_2 at period 2 can be different from continuation play at the history y_2y_2, since players may always condition on the outcome of period 0. Hence the continuation payoff $V(y_1y_2)$ can differ from $V(y_2y_2)$. Define functions $\gamma^{y_1} : Y \to \mathbb{R}, \gamma^{y_2} : Y \to \mathbb{R}$ so that $\gamma^y(y') = V(yy')$. Thus, the continuation payoffs following period 1 for history y_1, specified by γ^{y_1}, may be different from those at history y_2, specified by γ^{y_2}; that is, APS allow $\gamma^{y_1} \neq \gamma^{y_2}$. Instead of keeping track of these continuation payoff functions, APS keep track of the individual payoffs $\{\gamma^{y_1}(y_1), \gamma^{y_1}(y_2), \gamma^{y_2}(y_1), \gamma^{y_2}(y_2)\} \subset \mathscr{E}$, constructing (possibly different) continuation payoff functions at each history for each period. Knowing that four payoffs $\mathscr{W}^2 \equiv \{\tilde{v}^1, \tilde{v}^2, \tilde{v}^3, \tilde{v}^4\} \subset \mathscr{E}$ is sufficient to know that any two payoffs $\mathscr{W}^1 \equiv \{v^1, v^2\}$ decomposable on $\mathscr{W}^2$ are also PPE payoffs starting at period 1, which can serve as continuation payoffs for play at period 0, decomposing some v as the payoff for the whole equilibrium.

With 1-period memory, the only time players may condition on the outcome of period 0 is period 1. In Figure 2.3.1, the nodes labeled with "B1" are strategically equivalent to each other, as are the nodes labeled with "B2." Thus, 1-period memory imposes the restriction that $V(y_1y_1) = V(y_2y_1)$ and $V(y_1y_2) = V(y_2y_2)$. Thus, we are forced to pick a pair of continuation payoffs at period 1, rather than four, so that the continuation payoff functions above are equal: $\gamma^{y_1} = \gamma^{y_2}$. Put another way, we must choose one continuation payoff function, instead of two. Let $\mathcal{E} \subset \mathbb{R}$ denote the set of 1-period memory PPE payoffs for player 1, and let E denote the set of continuation payoff functions in all PPEs: that is, $\gamma \in E$ if and only if there exists some 1-period memory equilibrium σ' and some period t such that the continuation payoffs $V(\sigma'|t, y)$ at period t satisfy $V(\sigma'|t, y) = \gamma(y)$ for both $y \in \{y_1, y_2\}$. The fact that four payoffs $\mathscr{W}^2 \equiv \{v^1, v^2, v^3, v^4\} \subset \mathcal{E}$ is insufficient to know what payoffs can be decomposed with 1-period memory at period 1. But if we know $\gamma^2 \in E$

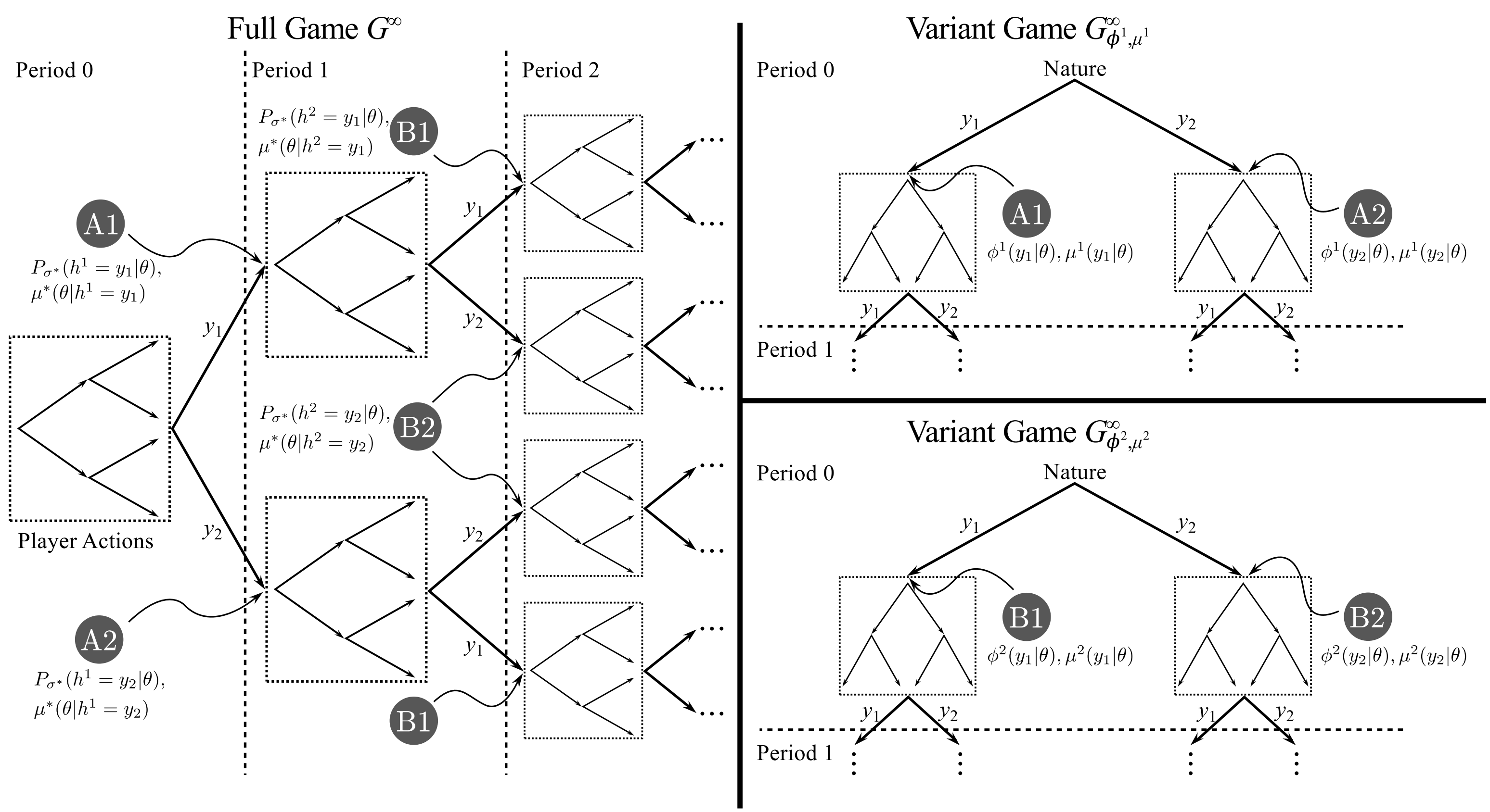

Figure 2.3.1: Depiction of the strategic equivalence between the full game and the variance games for 1-period memory ($K = 1$).

— perhaps so that $\gamma^2(y_1) = v^1, \gamma^2(y_2) = v^2$ — then we can decompose another payoff function γ^1 at period 1, so that γ^1 gives the continuation payoffs for play following period 0. Finally, a single payoff v (rather than another payoff function) for the whole equilibrium can be calculated by finding actions for period 0 that are enforced by γ^1, giving v as the discounted average payoff, similarly to (2.3.1).

2.3.2 Overview of the Framework

The dynamic programming methods of APS, DE, and this chapter are all driven by the recursive aspects of their environments. Summarizing, the key insight used by APS is that the continuation game at any history is strategically equivalent to the full game; the key insight used by DE is more cumbersome to state but similar: with K-period memory, the $|Y|^K$-length vector of continuation games for each public history at period $t \geq K$ is strategically equivalent to the analogous vector at any other period $t' \geq K$.[14]

The central insight used by this chapter's framework is that with bounded memory and incomplete information, continuation play is *almost* strategically equivalent to the full repeated game. In fact, it is equivalent to a slightly modified version of the full game where period 0 is endowed with an exogenous, fictitious initial history of length K, randomly drawn from a distribution dependent on the long-run player type. I show how to rephrase the DE insight in the following useful way, for now assuming complete information (no reputation). Construct a modified version of the full game that I call a *variant game*, where nature randomly picks a fictitious initial history of length K before play begins. Without reputation, this initial history is payoff irrelevant, but players may condition on it. Since nature's choice is payoff irrelevant, the probabilities with which nature picks each initial history are also payoff irrelevant. Suppose that the variant game is defined so that nature picks the initial history according the equilibrium distribution of play at some period $\underline{t} \geq K$ in some equilibrium σ^* of the full game. More formally, let $P_{\sigma^*}((t, h^{\underline{t}}) = (t, h^0))$ denote

[14] These vectors are equivalent to functions mapping public histories to continuation games, hence "self-generating payoff functions."

the probability that the period $\underline{t}$ public history is h^0 under strategy profile σ^*. In the variant game, nature picks initial history h^0 with probability $P_{\sigma^*}((t, h^{\underline{t}}) = (t, h^0))$. Define strategy profile $\tilde{\sigma}$ for the variant game so that for each public history h^k in period k, $\tilde{\sigma}|_{(k,h^k)} = \sigma^*|_{(\underline{t}+k,h^{\underline{t}+k})}$ — that is, $\tilde{\sigma}$ starting at period 0 is identical to σ^* starting at period $\underline{t}$. Then the equilibrium distribution of play in $\tilde{\sigma}$ will be identical to that of σ^*, shifted $\underline{t}$ periods earlier.

The equilibrium distribution of play is payoff irrelevant with complete information, but when the long-run player's type is unknown, it is critical because it determines beliefs. If the fictitious initial history is drawn according to a distribution conditional on the type, the initial history affects beliefs. If these conditional distributions are equal to the conditional distributions of public histories in some period $t \geq K$ in an equilibrium σ^* of the full game, then the beliefs in the variant game at period 0 are the same as the beliefs specified by σ^* at t for histories on the equilibrium path.[15] Furthermore, defining a strategy profile $\tilde{\sigma}$ for the variant game as in the previous paragraph ensures that the conditional equilibrium distributions are identical, shifted $\underline{t}$ periods earlier, so beliefs on the equilibrium path are also identical, shifted $\underline{t}$ periods earlier. Roughly speaking, an equilibrium of the full game starting at period t specifies an equilibrium of a variant game, and an equilibrium of a variant game specifies possible period t continuation play in an equilibrium of the full game whose probability distribution over public histories at period t matches the initial history distribution of the variant game.

2.3.3 Preliminaries and Variant Games

The usefulness of APS stems from the fact that given a set of hypothetical payoffs $\mathscr{W}$, we can prove that these are actually equilibrium payoffs by applying $\mathscr{B}(\cdot)$ and checking that $\mathscr{W} \subset \mathscr{B}(\mathscr{W})$ (i.e. $\mathscr{W}$ is self-generating). Similarly, this section constructs objects that are *hypothetical* properties of an equilibrium — specifically hypothetical probability distributions over public histories, hypothetical beliefs on public histories, and hypothet-

[15] I ignore off-equilibrium beliefs for now, dealing with them later.

ical continuation payoff functions. I begin by introducing these three main hypothetical "primitives," which are subsequently combined.

Definition 2.3.3. A *history distribution (HD)* $\phi : \Theta \to \Delta Y^K$ is a function mapping a type θ to a conditional probability distribution $\phi(\cdot|\theta) \in \Delta Y^K$ over public histories, giving a probability $\phi(h|\theta)$ for observing history h conditional on type $\theta \in \Theta$, such that $\sum_{h \in Y^K} \phi(h|\theta) = 1$. The support of ϕ is $\text{supp}\phi \equiv \{h \in Y^K | \exists \theta \in \Theta, \phi(h, \theta) > 0\}$. The set of all HDs is Φ.

Let some history distribution ϕ be given. In a hypothetical equilibrium whose probability distribution over public histories at period t matches ϕ, for any h on the equilibrium path (i.e. $h \in \text{supp}\phi$) the belief of player 2 that the type is θ upon observing history h is given by Bayes' rule, updating from the prior μ^0. For off-equilibrium histories, beliefs must be defined but are not restricted, so I construct an object to store hypothetical beliefs that are consistent with ϕ on its support. Note that when ϕ has full support ($\text{supp}\phi = Y^K$), this is redundant.

Definition 2.3.4. Let any history distribution $\phi \in \Phi$ be given. A function $\mu : \Theta \times Y^K \to [0, 1]$ is a *belief mapping* (BM) that is *consistent* with ϕ if for each $h \in \text{supp}\phi$, μ satisfies

$$\mu(\theta|h) = \frac{\mu^0(\theta)\phi(h|\theta)}{\sum_{\theta' \in \Theta} \mu^0(\theta)\phi(h|\theta)}, \qquad \forall \theta \in \Theta.$$

Denote the set of all belief mappings consistent with ϕ as M_ϕ.

Note that I do not construct hypothetical beliefs for the full history (unknown to short-run players) because such beliefs are payoff irrelevant: the long-run player (who knows the full history) does not condition on it. Finally, I introduce the third hypothetical "primitive."

Definition 2.3.5. A *payoff function* (PF) is a function $\gamma : Y^K \to \mathbb{R}$ that maps from a public history to a payoff. The set of all PFs is denoted Γ.

Since all three pieces of hypothetical "continuation information" are needed to know what kind of current period behavior can be "enforced," I combine the three primitives into the following composite objects.

Definition 2.3.6. A *history distribution and belief mapping* (HB) is a pair (ϕ, μ) where $\phi \in \Phi$ and $\mu \in M_\phi$. Denote the set of all HBs as $\mathcal{M} \equiv \{(\phi, \mu) : \phi \in \Phi, \mu \in M_\phi\}$.

Definition 2.3.7. A *history distribution, belief mapping and payoff function* object (HBP) is a triplet (ϕ, μ, γ) containing the HB $(\phi, \mu) \in \mathcal{M}$ and a payoff function $\gamma \in \Gamma$. The set of all HBPs is denoted $\mathcal{W}$.

An HBP is the analogue of a payoff vector in APS and the analogue of a payoff function in DE (indeed, with only one long-run player type the HB part becomes irrelevant, effectively simplifying to a payoff function). In APS language, characterizing the "largest self-generating set" of HBPs (defined formally in Section 2.3.4) is the central aim.

For each HB (ϕ, μ), I construct a modified version of the full game where an exogenous, fictitious history is drawn according to ϕ.

Definition 2.3.8. Let any HB $(\phi, \mu) \in \mathcal{M}$ be given. Define the (ϕ, μ)*-variant game* $G^\infty_{\phi,\mu}$ as follows.

1. Starting in period 0, a different short-run player 2 enters each period and plays the stage game G with long-run player 1 (as in G^∞).

2. Just before period 0, nature exogenously sets the public history to some *initial history* $h^0 \in Y^K$ with probability $\phi(h^0|\theta)$ conditional on player 1's type $\theta \in \Theta$. The first short-run player (in period 0) observes h^0, and the period 0 signal y^0 is pushed on h^0, yielding public history $h^0 y^0$ for the period 1 player, and so on (just as in the full game G^∞ at period K and later).

3. For any wPBE (σ^*, μ^*) of the variant game $G^\infty_{\phi,\mu}$, strategies and beliefs are required to be defined at all initial histories, even those chosen with probability 0 by nature, so that $\mu^*(\theta|h^0) = \mu(\theta|h^0)$ for all $h^0 \in Y^K$.

The third condition is a bit unusual because it is nature's choice, not player behavior, that keeps a history $h \notin \text{supp}\phi$ off the equilibrium path, yet I still require strategies and beliefs to be defined there.

Variant games are "real" games in their own right, but they are of interest because continuation play in any equilibrium of the full game is strategically equivalent to the beginning of some variant game. This strategic equivalence is illustrated more precisely in Figure 2.3.1. Focus first on the left half of the figure. Let some wPBE (σ^*, μ^*) be given for the full game G^∞. At the labels A1, A2, B1, and B2, I list the probability $P_{\sigma^*}(h|\theta)$ of reaching the corresponding public history h conditional on type θ under strategy profile σ^*, along with the belief $\mu^*(\theta|h)$ on the type θ upon reaching public history h. Note that both nodes labeled with "B1" have the same beliefs and are strategically equivalent, as are the two nodes labeled "B2."

Turn now to the right half of the figure. Define HB $(\phi^1, \mu^1) \in \mathcal{M}$ so that $\phi^1(h^1|\theta) = P_{\sigma^*}(h^1|\theta)$ and $\mu^1(\theta|h^1) = \mu^*(h^1|\theta)$ for each $h^1 \in Y$ and $\theta \in \Theta$. Similarly define HB $(\phi^2, \mu^2) \in \mathcal{M}$ so that $\phi^2(h^2|\theta) = P_{\sigma^*}(h^2|\theta)$ and $\mu^2(\theta|h^2) = \mu^*(h^2|\theta)$ for each $h^2 \in Y$ and $\theta \in \Theta$. Note that $\sigma^*|_{y_1}$ (i.e. σ^* starting at the A1 node in the full game) defines a PBE for the variant game $G^\infty_{\phi^1,\mu^1}$ starting at the A1 node. Conversely, given any wPBE $(\tilde{\sigma}, \tilde{\mu})$ of the variant game $G^\infty_{\phi^1,\mu^1}$, it can be "plugged into" σ^* starting at period 1 — replacing the strategies and beliefs at periods 1 and later — and the newly merged strategy profile will constitute another wPBE of the full game.[16]

I now construct the primary set of interest $\mathcal{D} \subset \mathcal{W}$, which is shown in the next section to be the "largest self-generating set" in the next section. In APS terms, it is the analogue to the set $\mathscr{E}$ of equilibrium payoffs; in DE terms, it is the analogue to the set E of equilibrium continuation payoff functions. Let $\Sigma_{\phi,\mu}$ denote the set of strategy profiles

[16] This is more precisely stated as follows. Define "merged" wPBE $(\bar{\sigma}, \bar{\mu})$ so that:

1. At period 0, take the strategies from the original full game strategy profile: $\bar{\sigma}(\varnothing) = \sigma^*(\varnothing)$.
2. Period 1 beliefs are taken from the original full game beliefs (which are the same as those specified for the variant game): $\bar{\mu}(h^1|\theta) = \mu^*(h^1|\theta) = \mu^1(h^0|\theta)$.
3. For periods 1, 2, ..., the strategies are taken from the variant game strategy profile, shifted 1 period back: $\bar{\sigma}^t(h) = \tilde{\sigma}^t(h^{t-1})$ for all $t \geq 1, h \in Y$.
4. For periods 2, 3, ..., the beliefs are taken from the variant game beliefs, shifted 1 period back: $\bar{\mu}(\theta|(t,h)) = \tilde{\mu}(\theta|(t-1,h))$ for all $t \geq 2, h \in Y$.

It is straightforward to see that $\bar{\mu}$ must be consistent with $\bar{\sigma}$, and that $\bar{\sigma}_1, \bar{\sigma}_2$ are mutual best responses.

for variant game $G^{\infty}_{\phi,\mu}$, and denote $\Sigma^*_{\phi,\mu}$ as the set of wPBE strategy profiles for $G^{\infty}_{\phi,\mu}$ (*not* the full game). Let $V(\tilde{\sigma}|h^0)$ denote the value of strategy profile $\tilde{\sigma} \in \Sigma_{\phi,\mu}$ to player 1, conditional on the realization of h^0 as the initial history. For each $\tilde{\sigma} \in \Sigma_{\phi,\mu}$, define payoff function $\gamma^{\tilde{\sigma}}_{\phi,\mu} : Y^K \to \mathbb{R}$ so that $\gamma^{\tilde{\sigma}}_{\phi,\mu}(h) = V(\tilde{\sigma}|h)$. Define $\Gamma^*_{\phi,\mu} \equiv \{\gamma^{\tilde{\sigma}}_{\phi,\mu} : \tilde{\sigma} \in \Sigma^*_{\phi,\mu}\}$ as the set of payoff functions for wPBEs of the (ϕ, μ)-variant games; note that without reputation, $\Gamma^*_{\phi,\mu}$ would be the same as E from DE (see Section 2.3.1).

Definition 2.3.9. For each HB $(\phi, \mu) \in \mathcal{M}$, define $\mathcal{D}_{\phi,\mu} \equiv \{(\phi, \mu, \gamma^{\tilde{\sigma}}_{\phi,\mu}) \in \mathcal{W} : \tilde{\sigma} \in \Sigma^*_{\phi,\mu}; \forall y \in Y^K, \gamma^{\tilde{\sigma}}_{\phi,\mu}(y) = V(\sigma|y)\}$ as the set of all HBPs containing HB (ϕ, μ) and the payoff function $\gamma^{\tilde{\sigma}}_{\phi,\mu}$, where for each $h \in Y^K$, the value $\gamma^{\tilde{\sigma}}_{\phi,\mu}(h)$ gives player 1's payoff for a wPBE of the (ϕ, μ)-variant game, conditional on initial history realization h. Define $\mathcal{D} \equiv \bigcup_{(\phi,\mu)\in\mathcal{M}} \mathcal{D}_{\phi,\mu}$ as the set of all such HBPs for all the variant games.

It is worth pausing to clarify the purpose and meaning of the above constructions. An HBP (ϕ, μ, γ) is a hypothetical description of the properties of an equilibrium of the full game G^{∞} at some period $t \geq K$; this is similar to how, in APS, a payoff vector $v \in \mathbb{R}^2$ is a hypothetical description of an equilibrium at some history in a complete information repeated game with full memory. Those hypothetical properties are the probability distribution over the public histories conditional on type, the beliefs at each public history, and the payoffs of the long-run player starting at each of those histories.

Given an HB (ϕ, μ), the (ϕ, μ)-variant game is an *actual* game, but it is useful because of its strategic equivalence to continuation play at some period $t \geq K$ of a *hypothetical* equilibrium of the full game, whose equilibrium distribution of play and beliefs at t "match" the hypothetical description given by (ϕ, μ). The set $\mathcal{D}_{\phi,\mu}$ is the set of all HBPs containing HB (ϕ, μ) and an equilibrium payoff function for the (ϕ, μ)-variant game, meaning a function mapping initial histories to the payoffs conditional on that initial history being realized; this can alternately be stated as $\mathcal{D}_{\phi,\mu} = \{(\phi, \mu)\} \times \Gamma^*_{\phi,\mu}$. Finally, $\mathcal{D}$ is the union over all these HBPs.

Any element $(\phi, \mu, \gamma) \in \mathcal{D}$ tells us that in the (ϕ, μ)-variant game, there exists an equilibrium where player 1 receives payoff $\gamma(h)$ if initial history h is realized. It also tells us that if there exists an equilibrium of the *full* game whose equilibrium distribution and beliefs over the histories at some period t match (ϕ, μ), then there also exists an equilibrium that matches (ϕ, μ) and has payoff $\gamma(h)$ upon arriving at history h in that period.

2.3.4 Self-Generation

This section defines the notions of enforceability, decomposition and self-generation, followed by the main results for Section 2.3. Before proceeding, one more hypothetical primitive must be defined. For full memory, APS use action profiles as a description of "current play," but this is insufficient for my purpose: I cannot pick just any action profile for each full history, since players only condition on the most recent K periods. This is similar to the reason we must use payoff functions rather than payoffs themselves. For complete information, DE use functions mapping from the public history to an action profile. I abuse notation by reusing α, using context to indicate whether it is a mixed action profile versus an "action profile mapping."

Definition 2.3.10. An *action profile mapping* (AM) $\alpha : Y^K \times \Theta \to \Delta A_2 \times (\Delta A_1)^{A_2}$ is a mapping from a public history h and type θ to mixed actions for each player $\alpha(\cdot|h, \theta) \equiv (\alpha_2(\cdot|h), \alpha_1(\cdot|ha_2, \theta))$, where $\alpha(a|h, \theta)$ denotes the probability of pure action profile a being played given h and θ. This is often written as $\alpha(h, \theta)$ for brevity. For each $\hat{\theta} \in \hat{\Theta}$, $\alpha_1(ha_2, \hat{\theta}) = \hat{\alpha}_{\hat{\theta}}$ for all $ha_2 \in Y^K \times A_2$. The set of all AMs is denoted $\mathcal{A}$.

With full memory, the APS notion of "enforceability" captures the requirement that hypothetical current behavior (an action profile) is consistent with hypothetical future behavior (continuation payoffs). With complete information, the only consistency needed is incentive compatibility. Reputation adds the additional issue that current behavior determines future beliefs (and therefore future payoffs), and so requires an additional consistency requirement besides incentive compatibility.

I call this requirement "inducibility." Call the current period "today" and the next "tomorrow." Let some HBP $(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ be given, serving as a description of hypothetical future behavior starting tomorrow, including the hypothetical tomorrow's distributions over histories and beliefs $(\tilde{\phi}, \tilde{\mu})$. Today's hypothetically consistent play must not only be enforced by the payoffs starting tomorrow; it must also yield a distribution of histories that matches what tomorrow's players expect (i.e. $\tilde{\phi}$). However, today's play, described by an action profile mapping α, is not sufficient to give this consistency since players condition on the public history observed today, which is itself random (generated by yesterday's players). Let the HB (ϕ, μ) describe the probability distribution for the history observed *today* (i.e. the probability distribution of yesterday's play) and today's beliefs. Together, (ϕ, μ) and α specify the (unique) probability distribution of the histories observed *tomorrow*. Thus, it will be useful to combine HBs and AMs in a composite object I call an "HBA," analogous to the an action profile in APS.

Definition 2.3.11. A history distribution, belief mapping and action profile mapping object (HBA) (ϕ, μ, α) is a triplet containing an HB (ϕ, μ) combined with an AM α. The set of all HBAs is denoted $\mathcal{X} \equiv \mathcal{M} \times \mathcal{A}$.

Define $\tau(h') \equiv \{h \in Y^K : \forall k \in \{K, ..., 2\}, h_{-k+1} = h'_{-k}\}$ as the set of public histories that can "be followed by h'," i.e. the $K-1$ oldest elements of h' are the same as the $K-1$ most recent elements of h.

Definition 2.3.12. The set of new HBs that are *inducible* by the HBA $x \equiv (\phi, \mu, \alpha)$ is denoted by the correspondence $\Upsilon : \mathcal{X} \rightrightarrows \mathcal{M}$. $\Upsilon(x)$ is the set of HBs $(\tilde{\phi}, \tilde{\mu})$ such that for each $\theta \in \Theta$,

$$\tilde{\phi}(h'y|\theta) = \sum_{h' \in \tau(h)} \sum_{a_2 \in A_2} \sum_{y \in Y} \phi(h|\theta)\alpha_2(a_2|h)\alpha_1(a_1|ha_2, \theta)\rho(y|a_2, a_1) \tag{2.3.2}$$

and $\tilde{\mu} \in M_{\tilde{\phi}}$ (i.e. $\tilde{\mu}$ is pinned down by Bayes' rule on the support of $\tilde{\phi}$).

This completes the definition of hypothetical description of current play, the HBA (analogous to an action profile in APS), and Section 2.3.3 constructed the hypothetical description of future play, the HBP (analogous to a continuation payoff vector in APS). It is now possible to define the incomplete information notion of enforceability, which requires that current and future play be consistent in terms of incentives (as in APS and DE) as well as in terms of beliefs (captured by inducibility).

Definition 2.3.13. For any $W \subset \mathcal{W}$, an HBA $x \equiv (\phi, \mu, \alpha) \in \mathcal{X}$ is *enforceable* on W if there exists HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\mu}, \tilde{\gamma}) \in W$ such that

1. **Inducibility:** the distribution over histories and beliefs must be consistent: $(\tilde{\phi}, \tilde{\mu}) \in \Upsilon(x)$ (see Definition 2.3.12).

2. **Short-run player incentive compatibility:** player 2 has no profitable deviations:

$$\sum_{\theta\in\Theta} u_2(\alpha(ha_2', \theta))\mu(\theta|h) \geq \sum_{\theta\in\Theta} u_2(\alpha_1(ha_2', \theta), a_2')\mu(\theta|h) \tag{2.3.3}$$

   for all $a_2' \in A_2$.

3. **Long-run player incentive compatibility:** player 1 has no profitable deviations: for all $h \in Y^K$, $a_2 \in A_2$, $a_1' \in A_1$,

$$\begin{aligned} \mathsf{V}_{a_2}(x, \tilde{w})(h) &\equiv (1-\delta)u_1(a_2, \alpha_1(ha_2, \theta_0)) \qquad (2.3.4)\\ &\quad +\delta \sum_{a_1\in A}\sum_{y\in Y} \alpha_1(a_1|ha_2, \theta_0)\rho(y|a_1, a_2)\tilde{\gamma}(hy) \\ &\geq (1-\delta)u_1(a_1', a_2) + \delta \sum_{y\in Y} \rho(y|a_2, a_1)\tilde{\gamma}(hy) \qquad (2.3.5) \end{aligned}$$

   where $\mathsf{V}_{a_2}(x, \tilde{w}) \in \Gamma$ is a payoff function.

The HBP $\tilde{w}$ *enforces* the HBA x.

The function $\mathsf{V}_{a_2}(x, \tilde{w})(h)$ defined in (2.3.4) gives the discounted average player 1 payoff upon arriving at public history h, conditional on player 2 action a_2. Define $\mathsf{V}(x, \tilde{w}) \in \Gamma$ to give the actual expected player 1 payoff upon arriving at h by averaging over player 2's strategy: $\mathsf{V}(x, \tilde{w})(h) = \sum_{a_2\in A_2} \alpha_2(a_2|h)\mathsf{V}_{a_2}(x, \tilde{w})(h)$.

I now define corresponding notion of "decomposability." Roughly speaking, an HBP $w \equiv (\phi, \mu, \gamma)$ is decomposed or "generated" by another HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ if there exists an HBA (ϕ, μ, α) enforced by $\tilde{w}$, so that the payoffs of those in α and $\tilde{\gamma}$ "average" out to those in γ.

Definition 2.3.14. An HBP $w \equiv (\phi, \mu, \gamma) \in \mathcal{W}$ is *decomposable* on $W \subset \mathcal{W}$ if there exists action profile mapping $\alpha \in \mathcal{A}$ such that

1. $x \equiv (\phi, \mu, \alpha)$ is enforced by some $\tilde{w} \equiv (\tilde{\phi}, \tilde{\mu}, \tilde{\gamma}) \in W$, and
2. for all $h \in Y^K$, $\gamma(h) = \mathsf{V}(x, \tilde{w})(h)$.

The HBP w is *decomposed* (or "generated") by the pair $(x, \tilde{w})$ (on W).

The following is my version of the APS $\mathscr{B}(\cdot)$ operator (from Definition 2.3.2) and "self-generation."

Definition 2.3.15. For any $W \subset \mathcal{W}$, define

$$\mathcal{B}(W) \equiv \{(\phi, \mu, \gamma) \in W : \gamma = \mathsf{V}(x, \tilde{w}) \text{ for some } x \in \mathcal{X} \text{ enforced by some } \tilde{w} \in W\}.$$

Definition 2.3.16. A set of HBPs $W \subset \mathcal{W}$ is *self-generating* if $W \subset \mathcal{B}(W)$.

It is now possible to state the self-generation and factorization results (analogous to Proposition 2.3.1 from APS), showing that $\mathcal{D}$ is the largest self-generating set. Recall from Section 2.3.3 that the set $\mathcal{D}$ is, roughly speaking, the set of all HBPs (ϕ, μ, γ) such that for the (ϕ, μ)-variant game ($G^{\infty}_{\phi,\mu}$), $\gamma(h)$ gives the player 1 payoffs in some equilibrium of $G^{\infty}_{\phi,\mu}$ when the (fictitious) initial history h is realized. These HBPs are of interest because they describe equilibrium payoffs for the *full game* G^{∞} starting at period K (when the public history gets "full"), conditional on the existence of a hypothetical equilibrium whose distribution over histories at period K matches the distribution specified by ϕ.

To summarize, the following proposition characterizes $\mathcal{D}$, which specifies the full game equilibrium payoffs at period K for all hypothetical equilibrium behavior in the initial periods $0, ..., K-1$. Further below I will show how to calculate what we are ultimately interested in, the set $\mathcal{E}$ of equilibrium payoffs (at period 0), from $\mathcal{D}$.

Proposition 2.3.3. *The following hold:*

1. ***Self-generation:*** *Suppose that a bounded set $W \subset \mathcal{W}$ is self-generating. Then $W \subset \mathcal{D}$.*

2. ***Factorization:*** $\mathcal{D} = \mathcal{B}(\mathcal{D})$.

Before presenting the algorithm, I define its initial starting point, the feasible set of HBPs. Let $\mathscr{F}^\dagger \subset \mathbb{R}$ be the set of feasible payoffs for player 1, and let $\mathcal{F}^\dagger \equiv \{(\phi, \mu, \gamma) \in \mathcal{W} : \forall h \in Y^K, \gamma(h) \in \mathscr{F}^\dagger\}$ denote the set of feasible HBPs (the set of all HBs paired with all payoff functions with feasible values). Repeatedly applying the $\mathcal{B}(\cdot)$ operator to this set converges to $\mathcal{D}$.

Proposition 2.3.4. *For each $m \in \mathbb{N}$, $\mathcal{B}^{m+1}(\mathcal{F}^\dagger) \subset \mathcal{B}^m(\mathcal{F}^\dagger)$ and $\mathcal{D} = \bigcap_{m=0}^{\infty} \mathcal{B}^m(\mathcal{F}^\dagger)$.*

I carry this algorithm out by hand when applying it to the product choice game in Section 2.5.1. Developing a numerical implementation like that of Judd, Yeltekin, and Conklin (2003) is a particularly interesting avenue for future research, given the complexity of these games.

The difficult part of the analysis is behind us, having characterized the possible equilibrium payoffs at period K, and can now turn to the ultimate objective of calculating the set $\mathcal{E}$ of wPBE payoffs of the full game for player 1. Once the possible equilibrium payoffs at period K are known, finding $\mathcal{E}$ boils down to solving the equilibrium payoffs of a full-memory, finitely repeated game with periods $0, ..., K-1$, with payoffs augmented at the end of period $K-1$ by the discounted continuation payoffs in γ for some HBP $(\phi, \mu, \gamma) \in \mathcal{D}$, with the additional requirement that equilibrium behavior match ϕ so that beliefs are consistent.

For each payoff function, construct such a finitely repeated game, calling it an "antegame" since it represents the first K periods of play in the full game.

Definition 2.3.17. Let any payoff function γ be given. The γ-*antegame* is defined as follows. Player 1's type θ is drawn with probability $\mu^0(\theta)$, and the stage game G is repeated K times (with first period 0), with all players observing the full history. Let $h \equiv (h_0, ..., h_{K-1}) \in Y^K$ denote the public history at the end of this game. Each short-run player 2 in period t receives their (ex-post) stage payoff $u_2^*(a^t, y)$ where a^t, y^t are the action profile played and signal generated at period t, while player 1 receives payoff

$$(1-\delta)\sum_{t=0}^{K-1}\delta^t u_1^*(a^t, y^t) + \delta^K \gamma(h).$$

Let Σ_γ^* denote that set of wPBE strategy profiles for the γ-antegame.

As above, let $V(\tilde{\sigma})$ denote the value of $\tilde{\sigma} \in \Sigma_\gamma^*$ to player 1. Let $P_{\tilde{\sigma}}(h|\theta)$ be the probability of "final history" $h \in Y^K$ given $\tilde{\sigma} \in \Sigma_\gamma^*$ and type $\theta \in \Theta$. For each $w \equiv (\phi, \mu, \gamma)$, define

$$\mathcal{E}_{(\phi,\mu,\gamma)} \equiv \{V(\tilde{\sigma}) : \tilde{\sigma} \in \Sigma_\gamma^*; \forall h \in Y^K, P_{\tilde{\sigma}}(h|\theta) = \phi(h|\theta)\} \tag{2.3.6}$$

as the set of payoffs of equilibria of the γ-antegame whose conditional probability distribution of final histories matches ϕ. The following operator summarizes this process.

Definition 2.3.18. Let a set of HBPs $W \in \mathcal{W}$ be given. Define $\mathcal{V}(W) \equiv \{v \in \mathcal{E}_w : w \in W\} \subset \mathbb{R}$.

Again, applying the $\mathcal{V}$ operator to a set of HBPs means solving the relatively simple task of finding the wPBE payoffs of finitely K-repeated games with full memory. Finally, applying the $\mathcal{V}$ operator to $\mathcal{D}$ yields the equilibrium payoffs of the full game.

Proposition 2.3.5. *For any set of HBPs $W \subset \mathcal{D}$, $\mathcal{V}(W) \subset \mathcal{E}$. Furthermore, $\mathcal{V}(\mathcal{D}) = \mathcal{E}$.*

I conclude by summarizing the full procedure of solving for the equilibrium payoffs. First, repeatedly apply the $\mathcal{B}(\cdot)$ operator to the set $\mathcal{F}^\dagger$ of feasible HBPs, yielding the set $\mathcal{D}$ (as in Proposition 2.3.4). Second, apply the $\mathcal{V}(\cdot)$ operator to $\mathcal{D}$, yielding the set $\mathcal{E}$ of equilibrium payoffs (as in Proposition 2.3.5).

2.3.5 Stationary Equilibria

Previous analysis of bounded memory reputation environments (e.g. Liu (2011b) and Liu and Skrzypacz (2014b)) has generally restricted attention to equilibria with stationary strategies, that is, strategies that do not depend on the calendar date (for periods $t \geq K$). If short-run players observe the calendar date, there may not be a wPBE with such strategies since beliefs, and therefore incentives, may depend on the calendar date. Let two public histories $h^t \in H^t, \tilde{h}^{t'} \in H^{t'}$ for $t, t' \geq K$ which are identical except for the period ($t \neq t'$), i.e. for every $k \in \{K, ..., 1\}$, $h^t_{-k} = \tilde{h}^{t'}_{-k}$; denote this equivalence $h^t \backsimeq \tilde{h}^{t'}$. There is no guarantee that short-run player best response in period t at public history h^t is a best response in a different period t' with $\tilde{h}^{t'} \backsimeq h^t$ since beliefs and thus expected payoffs may differ. This is not a problem if the equilibrium distribution of histories at each period is identical for all periods $t \geq K$, but such an assumption may be quite restrictive.

Definition 2.3.19. Let $\sigma \in \Sigma$ be a strategy profile of the full game. σ is *stationary* if it depends only on the public history and is independent of the calendar date; that is, for two histories $h^t, \tilde{h}^{t'}$such that $h^t \backsimeq \tilde{h}^{t'}$, $\sigma_2(h^t) = \sigma_2(\tilde{h}^{t'})$ and $\sigma_1(h^t a_2) = \sigma_1(\tilde{h}^{t'} a_2)$.

Instead, assume short-run players do not observe the calendar date, except for periods $0, ..., K-1$ where the length of the history gives the date away. Thus, the set of public histories in this specification is $H \equiv \bigcup_{t=0}^{K} Y^K$. Instead, they have the improper uniform prior over all periods $K, K+1, ...$, and so their beliefs are based on the limit of the average probability of play:

$$\mu(h|\theta) = \frac{\mu^0(\theta) \lim_{t\to\infty} \left\{ \frac{1}{t-K} \sum_{s=K}^{t} P^t_\sigma(h|\theta) \right\}}{\sum_{\theta' \in \Theta} \mu^0(\theta') \lim_{t\to\infty} \left\{ \frac{1}{t-K} \sum_{s=K}^{t} P^t_\sigma(h|\theta') \right\}} \tag{2.3.7}$$

where $P^t_\sigma(h|\theta)$ is the probability of history h at period t given strategy profile σ and long-run player type θ.

It turns out that combining the "time-average" history distribution ϕ (given by the limit terms in (2.3.7)) with the corresponding belief mapping μ and payoff function γ, with values are given by the continuation payoffs of σ, yields a self-generating singleton HBP (ϕ, μ, γ).

Proposition 2.3.6. *For any stationary wPBE (σ^*, μ^*), there exists an HBP $w \equiv (\phi, \mu, \gamma)$ such that $w \in \mathcal{B}(\{w\})$, i.e. w is decomposed by some HBA $x \equiv (\phi, \mu, \alpha) \in \mathcal{X}$ and itself, where the following are satisfied: $\alpha_2(h) = \sigma_2^*(h)$, $\alpha_1(ha_2, \theta) = \sigma_1^*(ha_2, \theta)$, $\mu^*(\theta|h) = \mu(\theta|h)$ and $V(h) = \gamma(h)$, for all $h \in Y^K$, $a_2 \in A_2$ and $\theta \in \Theta$.*

This gives stationary equilibria in the game with unobserved calendar date a very simple interpretation within the framework outlined in Section 2.3.4. Characterizing the set of non-stationary equilibrium payoffs requires characterizing the largest self-generating *set* of HBPs, while characterizing the set of stationary equilibrium payoffs only requires searching the HBP space for self-generating *points*.

2.4 Purifiability and Quasi-Markov Equilibria

The previous section provides an algorithm (Propositions 2.3.4 and 2.3.5) to calculate the set of wPBE payoffs. Even when the HBP space has few dimensions, carrying out the algorithm remains a daunting task. To get a sense of this, let an HB (ϕ, μ) and an HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ be given. Finding the set of action profile mappings α that make (ϕ, μ, α) enforceable by $\tilde{w}$ essentially amounts to finding the set of Bayes Nash equilibria of a one-shot sequential move game whose outcome probability distributions match $\tilde{\phi}$.[17] Given even a singleton set of HBPs $\tilde{W} \equiv \{\tilde{w}\}$, finding the set of decomposed HBPs $\mathcal{B}(\tilde{W})$

[17] To clarify, this one-shot game consists of the following four steps:

1. Nature chooses randomly chooses a history $h \in Y^K$ with probability $\phi(h|\theta)$, which is not payoff relevant except that it affects beliefs as it is correlated with player 1's type.
2. Player 2 chooses an action $a_2 \in A_2$.
3. Player 1 chooses an action $a_1 \in A_1$.
4. Nature chooses $y \in Y$ according with probability $\rho(y|a_2, a_1)$.

Then player 2 receives payoff $u_2^*(a_2, a_1, y)$ and player 1 receives $(1-\delta)u_1^*(a_2, a_1, y) + \delta\tilde{\gamma}(hy)$. The task is to find equilibria of this game such that history h and signal y occur with probability $\tilde{\phi}(hy|\theta)$.

means solving the above problem for every possible HB (ϕ, μ) (of which there are infinitely many).[18]

This section introduces an equilibrium refinement, which I call "quasi-Markov perfection," which greatly simplifies the set of strategies (i.e. HBAs) that must be considered in cases where the long-run player's actions are perfectly monitored and there is exactly one commitment type, as is the case with the applications in Section 2.5. In complete information dynamic games with a stochastic payoff-relevant state variable, applied work has often restricted attention to Markov perfect equilibria, defined as having strategies that condition only on the current state. Here, the stage game payoffs are static — there is only one Markov state in this sense — but the short-run player's *expected* payoffs are "dynamic" because beliefs change depending on the public history (and possibly time). Quasi-Markov perfection is the natural extension of Markov perfection to this incomplete information environment.

Besides enhancing tractability, quasi-Markov perfection has another virtue: the equilibria that it rules out are all "fragile" because they are not purifiable in the sense of Harsanyi (1973). Non-purifiable equilibria are not robust to arbitrarily small private, independent payoff shocks. I show this by extending the results of BMM, who prove that for complete information dynamic sequential move games with bounded memory, all purifiable equilibria are Markov perfect. As BMM point out, models cannot hope to describe reality perfectly and so at least some private payoff information is always present, so it is argued that this refinement does not come at the expense of realism.

At first glance, it may seem that beliefs are the appropriate extension of Markov states, as they determine the expected short-run player payoffs. However, such an equivalence class is too coarse for our purposes because two histories with the same beliefs today may lead to different beliefs tomorrow even if today's public signal is the same. For example, suppose (as will be done in Section 2.5) that the long-run player has two actions

[18] It means finding the set of Bayes Nash equilibria for every possible probability distribution for nature's move in step 1 of Footnote 17.

$A_1 \equiv \{C, D\}$, with a single commitment type $\hat{\theta}$ who always plays C, and suppose that the long-run player's action (and only her action) is perfectly monitored (formally, $Y \equiv \{C, D\}$ and $\rho(a_1|a_2, a_1) = 1$ for all $a_2 \in A_2$). Let the memory length be $K = 2$, and assume the belief $\mu(\hat{\theta}|h) = 0$ for any history h containing the non-commitment action D. Consider the situation of having history $h \equiv DD$ versus history $h' \equiv DC$ today. Both histories have belief 0, yet if C is played today, the belief tomorrow is $\mu(hC) = \mu(DC) = 0$ in the former case versus $\mu(h'C) = \mu(CC) \geq \mu^0(\hat{\theta}) > 0$ in the latter.

Quasi-Markov perfection allows different behavior at histories with the same beliefs today as long as they lead to different beliefs sometime in the future. A quasi-Markov state is defined as including all histories h, h' which have the same beliefs today's period t ($\mu^t(h) = \mu^t(h')$), will lead to the same beliefs tomorrow following any signal y^{t+1} today ($\mu^{t+1}(hy^{t+1}) = \mu^t(h'y^{t+1})$), and the same beliefs the next day following tomorrow's signal y^{t+2} ($\mu^{t+1}(hy^{t+1}y^{t+2}) = \mu^t(h'y^{t+1}y^{t+2})$), and so on forever.

Definition 2.4.1. Let any public wPBE (σ, μ) be given. For any period t, two date-histories $(t, h), (t, h') \in H^t$ are in the same *quasi-Markov state for player 2*, denoted $(t, h) \sim (t, h')$, if they have the same beliefs at the current period t and for any given continuation history $\boldsymbol{y}^k \in Y^k$ for any $k \geq 1$; that is, $\mu(\theta|(t, h\boldsymbol{y}^k)) = \mu(\theta|(t{+}k, h'\boldsymbol{y}^k))$ for all $\theta \in \Theta$, $k \in \{1, ...\}$, $\boldsymbol{y}^k \in Y^k$. Two player 2 actions $a_2, a_2' \in A_2$ are *incentive-equivalent* if $\rho(y|a_2, a_1) = \rho(y|a_2', a_1)$ for all $a_1 \in A_1$, and

$$u_1(a_2, a_1) - u_1(a_2, a_1') = u_1(a_2', a_1) - u_1(a_2', a_1') \qquad \forall a_1, a_1' \in A_1.$$

Two player 1 date-histories $(t, ha_2), (t, h'a_2')$ are in the same *quasi-Markov state for player 1* if a_2 and a_2' are incentive equivalent ($a_2 \sim a_2'$) and $(t, h), (t, h')$ have the same beliefs for any continuation history $\boldsymbol{y}^k \in Y^k$ for all $k \geq 1$ (but not necessarily at the current history, i.e. $k = 0$).

The equilibrium (σ, μ) is *quasi-Markov perfect* if strategies are the same within a quasi-Markov state for each player: $\sigma_2(t, h) = \sigma_2(t, h')$ and $\sigma_1(t, ha_2) = \sigma_1(t, h'a_2')$ for all $(t, h), (t, h') \in H^t$ and $a_2, a_2' \in A_2$ such that $(t, h) \sim (t, h')$ and $(t, ha_2) \sim (t, h'a_2')$.

To see how quasi-Markov perfection simplifies the strategy space, consider the example mentioned earlier. The histories CD and DD are in the same quasi-Markov state because both give belief 0 today, and no matter what today's signal y^t is (either C or D), the belief tomorrow (and in fact the whole history) will also be the same (formally, $CDy^t = DDy^t$ for both $y^t \in \{C, D\}$). The other two histories DC, CC are each in distinct quasi-Markov states, so there are three total states for $K = 2$.

For $K = 3$, there are four states. The histories DDD, CDD, DCD, CCD are all in the same state because all have belief 0 in the present period t, belief 0 the next period $t + 1$ following signal C, belief 0 in period $t + 2$ following another signal C, and after a third signal C all the histories become CCC (after signal D they have belief 0, of course). The other three states are $\{DDC, CDC\}$, $\{DCC\}$ and $\{CCC\}$. Note that in this example, quasi-Markov perfection means conditioning on when the most recent D event occurred, a point fleshed out in the application in Section 2.5.

I now construct the (ψ, ε)-perturbed game, largely following the construction given by BMM. Let Z_i be a full-dimensional, closed subset of $[0, 1]^{|A_i|}$ for each player i, and let $Z \equiv Z_2 \times Z_1$. Let $\Delta^*(Z)$ be the set of measures which have support Z generated by strictly positive densities. At each full history $\mathbf{h} \in \mathbf{H}$ a payoff shock $z \equiv (z_2, z_1) \in Z$ is drawn according to $\psi \in \Delta^*(Z)$. The payoff shocks are independent across the two players and the histories. The complete history with shocks at period t is $\tilde{\mathbf{h}} \in \tilde{\mathbf{H}}^t \equiv (A \times Y \times Z)^t$; also denote the set of player 1 full histories with shocks at period t as $\tilde{\mathbf{H}}_1^t \equiv (A \times Y \times Z_1)^t \times A_2$, and let $\tilde{\mathbf{H}}_1 \equiv \bigcup_{t=0}^{\infty} \tilde{\mathbf{H}}_1^t$. If player i chooses action a_i, then $\varepsilon z_i^{a_i}$ is added to her stage payoff, where $\varepsilon > 0$ and z^{a_i}. Player i's (ex-post) payoff for action profile a, signal y and shock z_i is

$$\tilde{u}_i(a, y, z_i) = u_i^*(a, y) + \varepsilon z_i^{a_i}.$$

Players privately observe only their own shocks. A strategy for player 2 at period t is $\tilde{\sigma}_2^t : H^t \times Z_2 \to \Delta A_2$. A strategy for player 1 is a mapping $\tilde{\sigma}_1 : \Theta \times \tilde{\mathbf{H}}_1 \to \Delta A_1$.

In any equilibrium of the perturbed game, players have a strict preference for their strategies for almost all shocks. Strategies with this property are called "essentially sequentially strict." The following definition also extends quasi-Markov perfection to the perturbed game for strategies that behave according to Definition 2.4.1 for almost all shocks.

Definition 2.4.2. A wPBE (σ, μ) is an *essentially sequentially strict equilibrium* if for all $(t, h) \in H$ and almost all payoff shocks $z_2 \in Z_2$, the action $\sigma_2(t, h, z_2)$ is pure and the unique maximizer, and similarly for all $a_2 \in A_2$ and almost all $z_1 \in Z_1$, $\sigma_1(t, ha_2, z_1)$ is also pure and the unique maximizer.

An equilibrium (σ, μ) of the perturbed game is *quasi-Markov perfect* if for almost all $z_2 \in Z_2$ and almost all $z_1 \in Z_1$, $\sigma_2(t, h, z_2) = \sigma_2(t, h', z_2)$ and $\sigma_1(t, ha_2, z_1) = \sigma_1(t, h'a_2', z_1)$ for all $(t, h), (t, h') \in H$, $a_2, a_2' \in A_2$ such that $h \sim h'$ and $a_2 \sim a_2'$.

The following result (my version of Proposition 1 from BMM) shows that all equilibria of the perturbed game are essentially sequentially strict, and every essentially sequentially strict equilibrium is quasi-Markov (see Lemma 2.2.2). The intuition is that the continuity of ψ ensures that a player being indifferent occurs with probability zero.

Proposition 2.4.1. *Every wPBE of the perturbed game is quasi-Markov perfect.*

The main purifiability result can now be stated. The following condition is what BMM call "weak purifiability," which is weaker than the purifiability notion of Harsanyi (1973).[19] A sequence of current shock strategies $(\tilde{\sigma}^k)_k$ converges in outcomes to a strategy σ in the unperturbed (full) game if

$$\lim_{k\to\infty} \int \tilde{\sigma}_2^k(a_2|h, z_2)\, d\psi^k(z_2) = \sigma_2(a_2|h) \quad \text{and} \quad \lim_{k\to\infty} \int \tilde{\sigma}_1^k(a_1|ha_2', z_1)\, d\psi^k(z_1) = \sigma_1(a_1|ha_2') \tag{2.4.1}$$

for each public history h and $a_2' \in A_2$.

[19]In the language of BMM, an equilibrium is "weakly purifiable" if there exists a sequence of perturbed games converging to the unperturbed game such that a sequence of corresponding wPBEs converge to the equilibrium. "Harsanyi purifiability" (as BMM call it) requires that for *every* sequence of perturbed games, there exists a sequence of corresponding wPBEs converging to the equilibrium. Harsanyi purifiability implies weak purifiability; see Definitions 6 and 7 of BMM.

Definition 2.4.3. A wPBE (σ^*, μ^*) of the full game G^∞ is *purifiable* if there exists a sequence $(\psi^k, \varepsilon^k)_k \to 0$ with $\psi^k \in \Delta^*(Z)$ and $\varepsilon^k \to 0$ such that there is a sequence of strategy profiles $(\tilde{\sigma}^k)_k$ converging in outcomes to σ^*, with $\tilde{\sigma}^k$ a wPBE of the (ψ^k, ε^k)-perturbed game.

The following is the incomplete information version of Proposition 2 from BMM, showing that quasi-Markov perfection is implied by purifiability.

Proposition 2.4.2. *Every purifiable wPBE is quasi-Markov perfect.*

One may wonder to what extent the converse holds: are quasi-Markov equilibria purifiable? Though I do not attempt a general answer, I show that in the application of Section 2.5, the minimum and maximum quasi-Markov equilibrium payoffs are given by purifiable equilibria for almost all priors.[20]

2.5 Applications

With the theoretical machinery of Sections 2.3 and 2.4 in hand, I apply it to the product choice game depicted in Figure 2.5.1. I assume short-run players have perfect monitoring of the K most recent long-run player actions but no monitoring of past short-run player actions. Liu and Skrzypacz (2014b) call this property "limited records." The long-run player is either normal type θ_0 with payoffs in Figure 2.5.1 or a commitment type $\hat{\theta}$ who always plays C. Short-run players have prior belief $\mu^0(\hat{\theta})$ that player 1 is the commitment type, which are abbreviated as simply μ^0 in this section. For concreteness, I refer to player 1 as the "firm" and player 2 as the "consumer."

Based on the reasoning of Section 2.4, I restrict attention to quasi-Markov perfect equilibria, a restriction that omits only non-purifiable equilibria. For any off-equilibrium

[20]Under complete information, BMM show a partial converse of their proposition for complete information: for a class of games with generic payoffs, all stationary Markov equilibria are purifiable, relying on a result from Doraszelski and Escobar (2010). I conjecture this result carries over to this incomplete information environment, but even if it does, it leaves open the question for non-stationary equilibria and non-generic payoffs, both of which must be considered in this application.

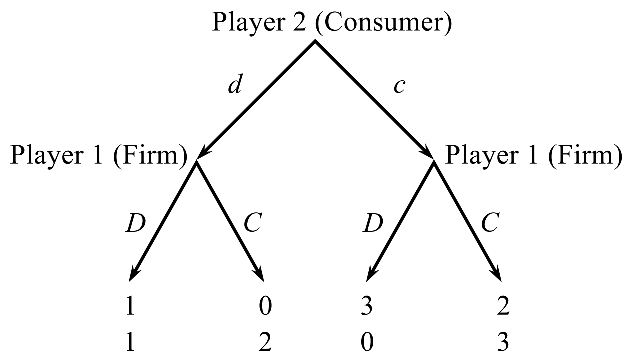

Figure 2.5.1: The product choice game, with player 1 (firm) payoffs on top and player 2 (consumer) payoffs on bottom.

date-history (t, h) where D is played at least once, assume that player 2 has belief 0 on the commitment type; Bayes' rule obviously implies this holds for all histories on the equilibrium path.

Assumption 2. *For any public history h containing D at any period, player 2 believes that player 1 is the normal type θ_0 with probability one.*

This assumption means that quasi-Markov perfection implies that players condition only on the period of the most recent D in the public history (and the calendar date). For any public history $h \in Y^K$, let

$$\iota(h) \equiv K + \min\{k : h_{1-k} = D\} \tag{2.5.1}$$

be the number of Cs since the most recent D, called the "index of h." For each $k \in \{0, ..., K-1\}$, define $I_k \equiv \{h \in Y^K : \iota(h) = k\}$ as the set of histories with index k, and define singleton set $I_K \equiv \{C^K\}$, where C^K is the K-length history containing only "C". For convenience, I use notation of the form $a_1^{k_1}\tilde{a}_1^{k_2} \cdots \bar{a}_1^{k_n}$ to denote the history containing k_1 instances of (player 1) action a_1, followed by k_2 instances of $\tilde{a}_1$, and so on, followed by $\bar{a}_1$ for the k_n most recent periods; for example, $DC^{K-3}D^2$ is the history consisting of one D, followed by $K-3$ periods of C, followed by two periods of D.

In their environment, Liu and Skrzypacz (2014b) show that strategies in all stationary PBEs, including (possibly) non-purifiable ones, depend on only on the index of a history, i.e. the time since the last non-commitment action. Proposition 2.4.2 shows that requiring purifiability also allows such a simplification for all finite stage games with limited records and a single long-run player commitment type, as well as to non-stationary equilibria. For each k, all the histories in I_k are in the same quasi-Markov state by Definition 2.4.1. Furthermore, the fact that the cost of effort is constant across the short-run player's actions (formally, $u_1(c, C) - u_1(c, D) = u_1(d, C) - u_1(d, D) = 1$) implies that both short-run player actions are incentive equivalent; hence, the long-run player does not condition on the short-run player's action. I relax this assumption in my analysis of stationary, long-memory equilibria in Section 2.5.2. I formally add quasi-Markov perfection to Definition 2.3.13 in the following definition.

Definition 2.5.1. Let an HBA $x \equiv (\phi, \mu, \alpha)$ enforced by some HBP $w \equiv (\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ be given. x is *quasi-Markov enforced* by w if $\alpha_2(h) = \alpha_2(h')$ and $\alpha_1(hc) = \alpha_1(hd)$ for any two histories $h, h' \in I_k$ for any k.

The analogous definitions of *quasi-Markov decompose* and *quasi-Markov self-generating*, as well as the analogues of Propositions 2.3.3 and 2.3.4, are given by inserting "quasi-Markov" into the appropriate places; the respective proofs are straightforward modifications and so are omitted.

This partitioning greatly simplifies the analysis of the HBP space. Consider some HBP $(\phi, \mu, \gamma) \in \mathcal{W}$. Since play at periods $K, K+1, ...$ conditions only on the index of the history, the collapsed payoff function space Γ is $(K+1)$-dimensional, one dimension for each index $0, ..., K$. (Indeed, Γ is isomorphic to $\mathbb{R}^{K+1}$.) Similarly partition the history distribution space Φ by index; for a given history distribution ϕ, abuse notation by denoting $\phi(I_k|\theta) \equiv \sum_{h \in I_k} \phi(h|\theta)$. Note that $\phi(I_K|\hat{\theta}) \equiv \phi(C^K|\hat{\theta}) = 1$, so only the history distributions for the normal type θ_0 are non-trivial; abbreviate $\phi(I_k) \equiv \phi(I_k|\theta_0)$. All of this means that Φ is now K dimensional (the requirement that $\sum_{k=0}^{K} \phi(I_k|\theta_0) = 1$ removes the $(K+1)$-th degree of freedom). Since beliefs are pinned down at every history, the belief mapping μ is redundant. Thus, I omit the belief mapping ("(ϕ, γ)" instead of "(ϕ, μ, γ)") except when it is notationally convenient, in which case I abbreviate $\mu(I_K) \equiv \mu(\hat{\theta}|I_K)$ (since beliefs at all other histories are simply zero). Summarizing, purifiability (and hence quasi-Markov perfection) allows us to collapse the $2|Y|^K - 1$ dimensions of HBP space ($|Y|^K - 1$ dimensions of history distributions and $|Y|^K$ dimensions of payoff function space, ignoring any non-redundant belief mappings) to $2K+1$ dimensions (K for history distributions and $K+1$ for payoff functions).

2.5.1 Product Choice Game with 1-Period Records

Consider the case where $K = 1$, which is sufficiently tractable to analytically characterize the exact minimum and maximum quasi-Markov equilibrium payoffs for all priors $\mu^0 \in [0,1]$ and almost all discount factors $\delta \in [0,1)$. There are just two possible histories at any period $t \geq 1$: $Y^K = \{C, D\}$.

As discussed above, the HBP space $\mathcal{W} \equiv \mathcal{M} \times \Gamma$ has three dimensions: one for history distributions and belief mappings ($\mathcal{M}$), and two for payoff functions (Γ). The space Φ of history distributions ϕ (and HB space $\mathcal{M}$, since beliefs are redundant) is isomorphic to the 1-simplex, so I further abbreviate $\phi \equiv \phi(C|\theta_0)$ (since $\phi(C|\theta_0) = 1 - \phi(D|\theta_0)$). For convenience, I use the real numbers in $[0,1]$ interchangeably with history distributions $\phi \in \Phi$.

Since there are only two player 2 actions $A_2 \equiv \{c, d\}$, I abbreviate $\sigma_2(t,h) \equiv \sigma_2(c|t,h) = 1 - \sigma_2(d|t,h)$. I do the same when discussing action profile mappings $\alpha \in \mathcal{A}$: $\alpha_2(h) \equiv \alpha_2(c|h)$. Similar abbreviation is possible for player 1, but it is possible to go further because purifiability requires that player 1 condition only on the calendar date, since either history C or D is immediately erased by the current period's signal (and so leads to the same beliefs for any continuation history) and player 2's actions c and d are incentive-equivalent. In other words, for a given period t, all player 1 date-histories (t, ha_2) are in the same quasi-Markov state. Hence, abbreviate $\sigma_1(t) \equiv \sigma_1(C|t, ha_2)$ for all $h \in Y \equiv \{C, D\}$ and $a_2 \in A_2$, and do the same for an action profile mapping: $\alpha_1 \equiv \alpha_1(C|ha_2)$. As with history distributions, I use real numbers in $[0,1]$ interchangeably with these mixed actions.

The results are presented first, followed by discussion of the algorithm.

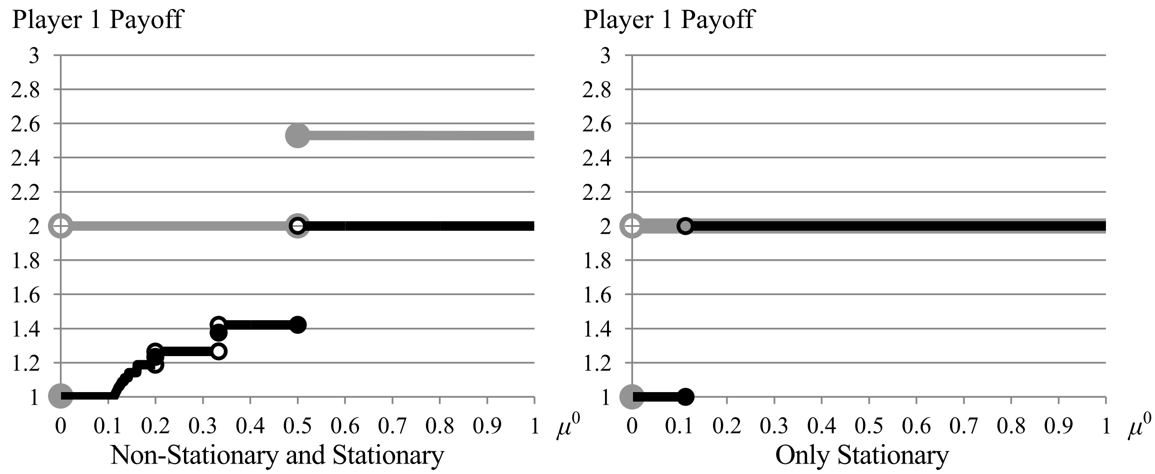

Figure 2.5.2: On the left, the minimum (black) and maximum (gray) quasi-Markov equilibrium long-run player payoffs are plotted for $\delta = 0.9$. On the right, the minimum and maximum payoffs for only stationary equilibria are plotted.

2.5.1.1 Results

The following proposition gives the exact minimum and maximum quasi-Markov equilibrium payoffs for player 1 for all priors μ^0 and all discount factors $\delta \in [0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$. The results are more easily understood with the plot in Figure 2.5.2.

Proposition 2.5.1. *Let* $\mathcal{E}(\delta, \mu^0)$ *be the set of quasi-Markov equilibrium player 1 payoffs of the 1-period records product choice game, and let* $\underline{e}(\delta, \mu^0) \equiv \min \mathcal{E}(\delta, \mu^0), \bar{e}(\delta, \mu^0) \equiv \max \mathcal{E}(\delta, \mu^0)$ *be the minimum and maximum quasi-Markov equilibrium payoffs, respectively. For* $\delta < \frac{1}{2}$*,*

$$\underline{e}(\delta, \mu^0) = \begin{cases} 1 & \mu^0 \le \frac{1}{2} \\ 3(1-\delta) + \delta & \mu^0 > \frac{1}{2} \end{cases} \qquad \bar{e}(\delta, \mu^0) = \begin{cases} 1 & \mu^0 < \frac{1}{2} \\ 3(1-\delta) + \delta & \mu^0 \ge \frac{1}{2}. \end{cases} \tag{2.5.2}$$

For all $\delta > \frac{1}{2}$*,*

$$\underline{e}(\delta, \mu^0) = \begin{cases} 1 & 0 \le \mu^0 \le \frac{1}{9} \\ \lambda^*(\mu^0) & \frac{1}{9} < \mu^0 \le \frac{1}{2} \\ 2 & \frac{1}{2} < \mu^0 \le 1 \end{cases} \qquad \bar{e}(\delta, \mu^0) = \begin{cases} 1 & \mu^0 = 0 \\ 2 & 0 < \mu^0 < \frac{1}{2} \\ \frac{1-\delta}{1-\delta^2}(3 + 2\delta) & \frac{1}{2} \le \mu^0 \le 1, \end{cases} \tag{2.5.3}$$

where $\lambda^*(\mu^0)$ *is defined as follows. Let* $L(\mu^0) = \min_k q^k(0,\mu^0)$ *such that* $q^k(0,\mu^0) \geq \frac{1}{2}$*, where* $q : \mathbb{R} \times [0,1] \to \mathbb{R}$ *is defined by*[21]

$$q(\phi,\mu^0) = \frac{\frac{\mu^0}{1-\mu^0}}{1-2\phi}. \tag{2.5.4}$$

Define

$$\lambda_1^k \equiv \frac{1}{1-\delta^{k+1}}[(1-\delta^k) + 3\delta^k(1-\delta) - \delta^{k-1}(1-\delta)] \tag{2.5.5}$$

for $k \in \{L-1, L\}$*. Finally,*

$$\lambda^*(\mu^0) = \begin{cases} (1-\delta) + \delta\lambda_1^{L(\mu^0)-1} & q^{L(\mu^0)}(0,\mu^0) = \frac{1}{2} \\ \lambda_1^{L(\mu^0)} & \textit{otherwise.} \end{cases}$$

A natural concern about this characterization is that Proposition 2.4.2 proves that quasi-Markov perfection is only a necessary, not necessarily sufficient, condition of purifiability – are the equilibria giving these payoffs actually purifiable? I assuage this concern in Appendix 2.3.2, proving that there is a purifiable equilibrium with the minimum and maximum payoffs given in Proposition 2.5.1 for almost all priors.[22]

In the complete information case ($\mu^0 = 0$), all histories are in the same quasi-Markov state. The BMM result applies here and shows that the only purifiable equilibrium outcome is the repeated static Nash equilibrium with payoff 1.

Going from zero to a slightly positive prior, there is no discontinuity for the minimum equilibrium payoff, but there is for the maximum, which immediately jumps to the Stackelberg payoff of 2. For $0 < \mu^0 \leq \frac{1}{9}$, the minimum payoff is given by stationary equilibria where the long-run player plays C with some probability $\alpha_1 \in [0, \frac{1}{2})$.[23] The lower α_1 is, the more the short-run player prefers d conditional on the normal type θ_0; however, a

[21] Recall that real numbers (in $[0,1]$) and history distributions (in Φ) are used interchangeably. I use $q(\phi,\mu^0)$ only when $\phi \in [0,1]$ and $q(\phi,\mu^0) \in [0,1]$, except to show contradictions in the proof where useful.

[22] The Lebesgue measure zero set of priors ignored in Appendix 2.3.2 correspond to the priors with discontinuities in Figure 2.5.2.

[23] Recall that quasi-Markov perfection means player 1 cannot condition on anything except the calendar date.

smaller α_1 also implies a higher belief at history C the next period. The minimum payoff stationary equilibrium has the long-run player mixing such that these countervailing effects balance and the short-run player is indifferent upon observing history C. Since the short-run player is indifferent at history C (where she has positive belief on the commitment type), she strictly prefers d at history D (when the belief is zero and the long-run player still plays the same mixed action). Since the long-run player is always indifferent, the long-run player's payoff can be calculated with the payoff of always playing D (a best response), which yields flow payoff 1 every period. To keep the long-run player indifferent, the short-run player plays $\alpha_2(D) = 0$ and mixes with positive probability $\alpha_2(C)$ at history C.

The maximum payoff stationary equilibrium is qualitatively similar, with the long-run player mixing with probability $\alpha_1 = \frac{1}{2}$ so that the short-run player is indifferent at history D, while strictly preferring c at history C. At history D, the short-run player mixes so that the long-run player is indifferent. The best response strategy of always playing C yields a flow payoff of 2 every period.

For prior $\mu^0 > \frac{1}{9}$, the payoff 1 stationary equilibria are no longer sustainable. If player 1 mixes today (period t) with probability $\alpha_1(t)$ such that player 2 is indifferent at history C with belief $\mu(t, C)$, the prior is high enough that the belief tomorrow $\mu(t+1, C) > \mu(t, C)$ must be higher. Keeping tomorrow's player 2 indifferent at history C requires player 1 mix at a lower probability $\alpha_1(t+1) < \alpha_1(t)$, so the equilibrium cannot be stationary. Instead, the minimum quasi-Markov equilibrium payoff for $\frac{1}{9} < \mu^0 \leq \frac{1}{2}$ is given by a non-stationary equilibrium where player 1 mixes in periodic cycles, starting with some high probability and gradually playing C more rarely. At the end of the cycle, the belief exceeds $\frac{1}{2}$ and player 2 strictly prefers c, at which point player 1 plays C with high probability, yielding a low belief next period and restarting the cycle. As the prior increases, the belief cycle must become shorter, reaching a high belief more quickly and thus yielding a higher equilibrium payoff. The discontinuities in the graph are a result of the steps in the cycle being eliminated by a higher prior. By contrast, the high payoff stationary equilibrium survives because the short-run player is indifferent at history D, where the belief (zero) is unaffected by the prior.

At $\mu^0 > \frac{1}{2}$, the minimum payoff is the same (in terms of strategies) stationary equilibrium that gave the maximum payoff for $\mu^0 < \frac{1}{2}$. There also exists a non-stationary equilibrium where the short-run player is exploited every other period upon observing history C. The long-run player always plays D in even periods and C in odd periods. The odd-period short-run players clearly have c as a strict best response. The even-period short-run players also have c as a strict best response, because the prior is so high that c gives a higher payoff despite the long-run player playing D with certainty and the fact that the history is uninformative ($\mu(C) = \mu^0$). This equilibrium gives the maximum payoff (approaching 2.5 for δ close to one) for these high priors.

2.5.1.2 Algorithm

To carry out the algorithm of Proposition 2.3.4, the starting point is the set of feasible HBPs $\mathcal{F}^\dagger$, which is the subset of $\mathcal{W}$ where the payoff functions have values in the set of feasible payoffs $\mathscr{F}^\dagger \equiv [0,3]$. This set is drawn in Figure 2.5.3 as a "cube."[24] The algorithm involves applying the $\mathcal{B}(\cdot)$ operator to all the points in this 3-dimensional set, and then each subsequent set $\mathcal{B}(\mathcal{F}^\dagger), \mathcal{B}^2(\mathcal{F}^\dagger), \ldots$.

Fortunately, it is possible to ignore almost all of the points in $\mathcal{F}^\dagger$ because they generate the empty set. In fact, only the gray points in Figure 2.5.3 generate non-empty sets. Note the "vertical" (coming out of the page) plane $\mathcal{I} \equiv \{(\phi, \mu, \gamma) : \gamma(C) = \gamma(D) + \frac{1-\delta}{\delta}\}$, which I call the *indifference plane*, consisting of all points where the long-run player is indifferent between C and D.

To the "northwest" of the indifference plane, player 1 strictly prefers D. This includes the 45° plane (a line in payoff function space) outlined by the diagonal dashed lines in Figure 2.5.3, where there is no intertemporal incentive because the continuation payoffs for both

[24]Though it is drawn as a cube to make the diagrams easier to read, the ϕ dimension is not comparable in any direct way to the payoff function dimensions, so there is no sense in which the cube has the same "length" in this direction as in the other two.

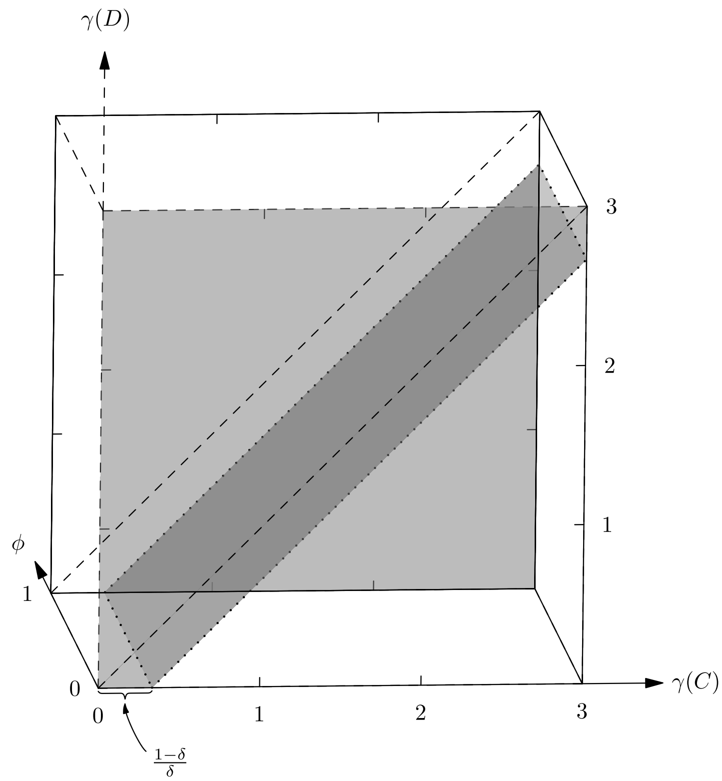

Figure 2.5.3: The "cube" is the set $\mathcal{F}^\dagger$ of feasible HBPs, with the history distribution dimension pointing out of the page. The gray set is the set of "useful" points $\bar{\mathcal{F}}$, which generate non-empty sets.

actions are equal. Consider some HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ in this northwest set. Since she strictly prefers D under $\tilde{w}$, any enforceable HBA (ϕ, α) clearly must have player 1 always playing D: $\alpha_1 = 0$. Since player 1's action is the same as the public signal, inducibility requires that $\tilde{\phi} = \alpha_1 = 0$. Hence, if $\tilde{w}$ is not on the "floor" of the northwest (i.e. $\tilde{\phi} > 0$), there does not exist an HBA enforced by $\tilde{w}$, and so there are no HBPs decomposed by $\tilde{w}$: $\mathcal{B}(\{\tilde{w}\}) = \varnothing$. This is why I only keep track HBPs in the northwest above the floor as the algorithm is carried out.

Exactly the opposite occurs in the "southeast" of the indifference plane, where the long-run player strictly prefers C. Letting $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ in the southeast be given, any enforceable HBA (ϕ, α) must have player 1 always playing C: $\alpha_1 = 1$. Inducibility requires $\tilde{\phi} = \alpha_1 = 1$, and so unless $\tilde{w}$ is on the "ceiling" of the southeast, there is no HBA enforced by $\tilde{w}$ and $\mathcal{B}(\{\tilde{w}\}) = \varnothing$. Thus, I only keep track of the ceiling in the southeast.

The indifference plane itself is special because every point $(\tilde{\phi}, \tilde{\gamma})$ in it generates non-empty sets. Inducibility requires $\tilde{\phi} = \alpha_1$ for an enforceable HBA (ϕ, α), but since player 1 is indifferent, I can always choose α_1 to satisfy this without violating incentive compatibility.

To summarize, the set of "useless points" $\mathcal{F}_{\varnothing} \equiv \{w \in \mathcal{F}^{\dagger} : \mathcal{B}(\{w\}) = \varnothing\}$ — consisting of everything above the floor of the northwest and below the ceiling of the southeast — can be safely ignored. Denote the remainder of the feasible space $\bar{\mathcal{F}} \equiv \mathcal{F}^{\dagger} \backslash \mathcal{F}_{\varnothing}$, the gray space in Figure 2.5.3, called the "useful points."

I construct analogues of the tools of Section 2.3 that deal only with the useful points. Define the set $\bar{\mathcal{D}} \equiv \mathcal{D} \cap \bar{\mathcal{F}}$, it is easy to show that $\mathcal{V}(\bar{\mathcal{D}}) = \mathcal{V}(\mathcal{D}) = \mathcal{E}$ because any equilibrium $\tilde{\sigma}$ of the γ-antegame for some useless HBP $(\phi, \mu, \gamma) \in \mathcal{F}_{\varnothing}$ would clearly fail the requirement that $P_{\tilde{\sigma}}(h|\theta) = \phi(h|\theta)$ in (2.3.6), so $\mathcal{V}(\{(\phi, \mu, \gamma)\}) = \varnothing$. Define the analogous operator $\bar{\mathcal{B}}(W) \equiv \mathcal{B}(W) \cap \bar{\mathcal{F}}$ which ignores the useless points generated by W. It is easy to show that $\bigcap_m \bar{\mathcal{B}}^m(\bar{\mathcal{F}}) = \bar{\mathcal{D}}$, which simplifies executing Proposition 2.3.4's algorithm.

What can the useful points actually generate? Figure 2.5.4 shows the set of generated points for a number of example HBPs $\tilde{w}^1, ..., \tilde{w}^{10}$ all located on the indifference plane (and therefore useful). In the following discussion, I call these the "enforcing points" $\tilde{w}^j \equiv$

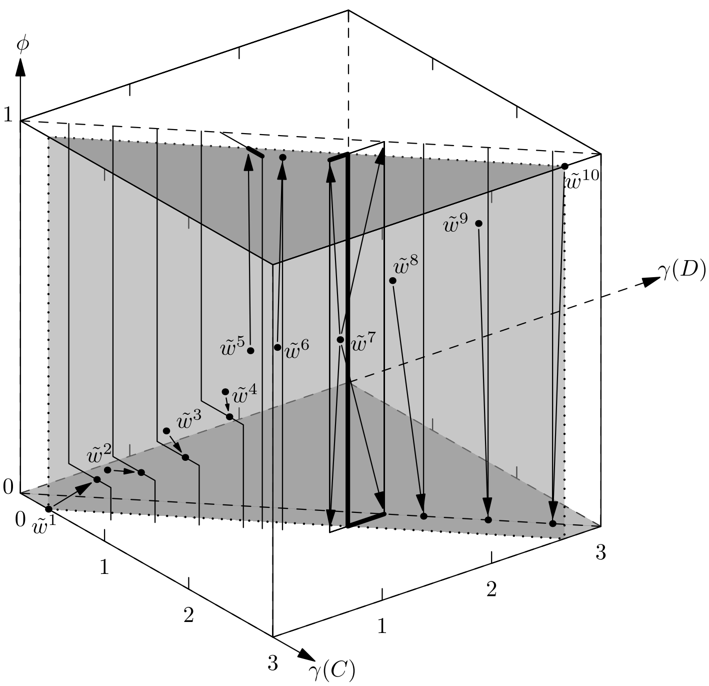

Figure 2.5.4: Ten example points $\tilde{w}^1, ..., \tilde{w}^{10}$ on the indifference plane and the sets of points $\mathcal{B}(\{\tilde{w}^1\}), ..., \mathcal{B}(\{\tilde{w}^{10}\})$ they each generate. The useful points $\bar{\mathcal{B}}(\{\tilde{w}^1\}), ..., \mathcal{B}(\{\tilde{w}^{10}\})$ generated by each respectively is bolded. All points at the same elevation generate qualitatively similar sets, so the examples are purposely selected at distinct elevations.

$(\tilde{\phi}^j, \tilde{\mu}^j, \tilde{\gamma}^j)$, referring to $\tilde{\phi}^j$ as the "enforcing history distribution," and so on. For an HBA $(\phi^j, \mu^j, \alpha^j)$ enforced by $\tilde{w}^j$, I refer to ϕ^j as the "current history distribution," and so on.

First, consider $\tilde{w}^1 \equiv (\tilde{\phi}, \tilde{\gamma})$ on the floor.[25] I start by finding all HBAs $x \equiv (\phi^1, \mu^1, \alpha^1)$ enforced by $\tilde{w}$. Inducibility requires that player 1 always play D (i.e., $\alpha_1^1 = 0$), so if player 2 knows she faces the normal type, she will strictly prefer d. This is the case at the history D, so $\alpha_2^1(D) = 0$. She also strictly prefers d at history C so long as the belief $\mu^1(C)$ is low enough, which is true so long as the probability ϕ^1 of C for the normal type is high enough. The proof shows this threshold is $\frac{\mu^0}{1-\mu^0}$. Below the threshold ($\phi^1 < \frac{\mu^0}{1-\mu^0}$), the belief on the commitment type is so high that she strictly prefers c. When $\phi^1 = \frac{\mu^0}{1-\mu^0}$, player 2 is indifferent and may choose any action $\alpha_2^1(C) \in [0, 1]$.

Having found the HBAs x enforced by $\tilde{w}$, I can find the set of HBPs $w \equiv (\phi^1, \mu^1, \gamma^1)$ decomposed by x and $\tilde{w}$. Since $\alpha_2(D) = 0$, it is known that $\gamma^1(D) = (1-\delta)u_1(d, D) + \delta\tilde{\gamma}^1(D) = (1-\delta) + \delta\tilde{\gamma}^1(D)$. For $\phi > \frac{\mu^0}{1-\mu^0}$, I have $\gamma^1(C) = \gamma^1(D)$ because player 2 strictly prefers d at history C, which means w is on the 45° plane in the northwest (above the floor). For $\phi^1 < \frac{\mu^0}{1-\mu^0}$, I have $\gamma^1(C) = (1-\delta)u_1(d, C) + \delta\tilde{\gamma}^1(D) = 3(1-\delta) + \delta\tilde{\gamma}^1(D)$, so w is in the southeast (below the ceiling) because $\gamma^1(C) > \gamma^1(D) + \frac{1-\delta}{\delta}$. For $\phi^1 = \frac{\mu^0}{1-\mu^0}$, $\gamma^1(C)$ can be anywhere in between — including the indifference plane. This is crucial because the point generated on the indifference plane is the only useful point of $\mathcal{B}(\{\tilde{w}^1\})$, so it is only necessary to keep track of this one for the next iteration of the algorithm.

The HBPs $\tilde{w}^2, \tilde{w}^3, \tilde{w}^4$ have positive but still low enforcing history distributions: $0 < \tilde{\phi}^2 < \tilde{\phi}^3 < \tilde{\phi}^4$. Let $(\phi^j, \mu^j, \alpha^j)$ be an HBA enforced by one of these enforcing points $\tilde{w}^j$ such that the (current) player 2 is indifferent. The set of generated points for each $\tilde{w}^j$ is qualitatively similar but "shifted up" because player 1 is playing C with increasing (current) probability α^j, so belief μ^j at which player 2 is indifferent is lower and hence the current history distribution ϕ^j is higher. Given enforcing distribution $\tilde{\phi}$, the proof shows that this

[25] Figure 2.5.4 shows it on the indifference plane instead of the northwest, but the set of generated points is the same at points directly to the west, i.e. points with the same $\tilde{\gamma}^1(D)$ coordinate. The reason is that only the continuation payoff for D is all that matters since $\tilde{\phi} = 0$ implies player 1 always plays D.

shifting up is described by the function $\phi^j = q(\tilde{\phi}^j) = \frac{\mu^0/(1-\mu^0)}{1-2\tilde{\phi}^j}$ (this is described in greater detail in the discussion of Figure 2.5.5).

Consider $\tilde{w}^5$ in Figure 2.5.4. At a high enough enforcing distribution $\tilde{\phi}^5 = q^{-1}(1)$, player 2 is indifferent at current distribution $\phi = 1$, which is the lowest possible posterior $\mu(C) = \mu^0$. Note that because the "lap" of the "wall-sit" figure intersects with the ceiling of the southeast, there is a line segment of multiple useful points in the decomposed set, bolded in the figure. For all enforcing distributions strictly between $\tilde{\phi}^5$ and $\frac{1}{2}$, the set of generated points is simply a vertical line in the southeast, for example $\tilde{w}^6$ and $\mathcal{B}(\{\tilde{w}^6\})$. Since the generated points all lie in the southeast, the only useful one is on the ceiling.

The $\tilde{w}^7 = (\tilde{\phi}^7, \tilde{\gamma}^7)$ enforcing HBP has $\tilde{\phi}^7 = \frac{1}{2}$, which means that player 2 is indifferent at history D when she knows she is facing the normal type and strictly prefers c when she has positive belief on the commitment type (at history C). Thus, $\alpha_2^7(C) = 1$ so $\gamma^7(C) = 3(1-\delta) + \delta\tilde{\gamma}^7(D)$. On the other hand, I can choose any action $\alpha_2^7(D) \in [0,1]$ and so can decompose any $\gamma^7(D) \in [(1-\delta) + \delta\tilde{\gamma}^7(D), \gamma^7(D)]$. This yields the rectangle depicted in Figure 2.5.4, whose intersection with $\bar{\mathcal{F}}$ is bolded. For enforcing distributions strictly greater than $\frac{1}{2}$ (e.g. $\tilde{w}^8, \tilde{w}^9, \tilde{w}^{10}$), player 2 strictly prefers c at both histories, so for example $\gamma^8(C) = \gamma^8(D) = 3(1-\delta) + \delta\tilde{\gamma}^8(D)$, generating a vertical line on the 45°plane in the northwest.[26]

I summarize the set of useful points generated by different HBPs with Figure 2.5.5. Roughly speaking, Figure 2.5.5 "collapses" the 3-dimensional Figure 2.5.4 into two dimensions to convey the mapping from the elevation of an enforcing point to that of the useful generated points. Given an enforcing HBP $(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$, the correspondence $\mathcal{R}(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma}) \equiv \{\phi : (\phi, \mu, \gamma) \in \bar{\mathcal{B}}(\{(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})\})\}$ gives the set of history distributions ϕ for which a useful point generated. In other words, the horizontal axis is the "elevation" of the enforcing point in Figure 2.5.4, and the vertical axis is the elevation of the generated useful point. The correspondence is labeled by line style to indicate the payoffs that are decomposed at each

[26]For $\tilde{w}^{10}$, the analogue of Footnote 25 holds. Though $\tilde{w}^{10}$ is on the indifference plane, any points directly south (with the same $\tilde{\gamma}^{10}(C)$ coordinate) generate exactly the same set.

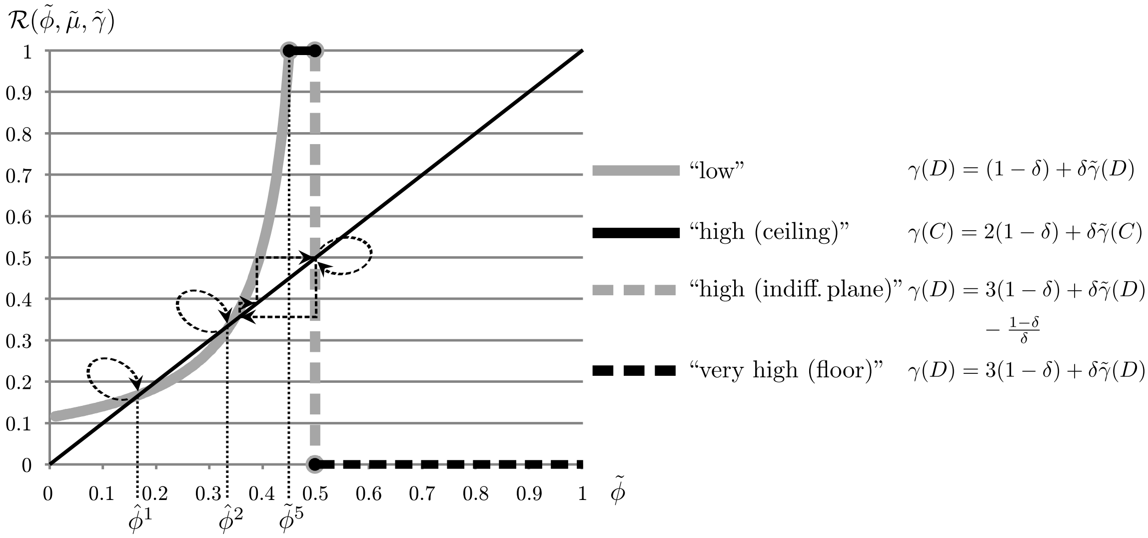

Figure 2.5.5: Given an enforcing HBP $(\tilde{\phi}, \tilde{\gamma})$ with history distribution $\tilde{\phi}$ (horizontal axis), I plot the set of history distributions ϕ (vertical axis) for which a point (ϕ, γ) is generated by $\tilde{w}$. The black line is the 45° line. The different line styles describe the generated payoff function γ.

elevation. Note that $\mathcal{R}$ crosses the 45° line at three points: $\hat{\phi}^1$, $\hat{\phi}^2$ and $\frac{1}{2}$. These indicate the three stationary equilibria, since stationary equilibria correspond to self-generating HBPs (recall Proposition 2.3.6). There are two “low” payoff and one “high” payoff stationary equilibria.

Figure 2.5.6 shows the effect of the prior increasing to $\frac{1}{9}$ and just above. Note that when the prior is $\frac{1}{9}$, the two stationary low payoff equilibria seen in Figure 2.5.5 have “merged” into one. When the prior increases above $\frac{1}{9}$, the stationary equilibrium disappears. Instead, the lowest payoff equilibrium is non-stationary, given by a “cycle” of self-generating HBPs depicted by the arrows. Note that generation “goes backwards in time,” so the belief at history C follows a 5 period cycle, starting from a low value, gradually increasing until the belief is 1, and then starting over again at the low value.

As the prior increases, the number of low payoff steps in the cycle necessarily decreases as the solid gray curve rises, as depicted in 2.5.7. When $\mu^0 = \frac{1}{2}$, the solid gray curve has been reduced to a single point at the upper right corner, because $q(0) = 1$. This allows

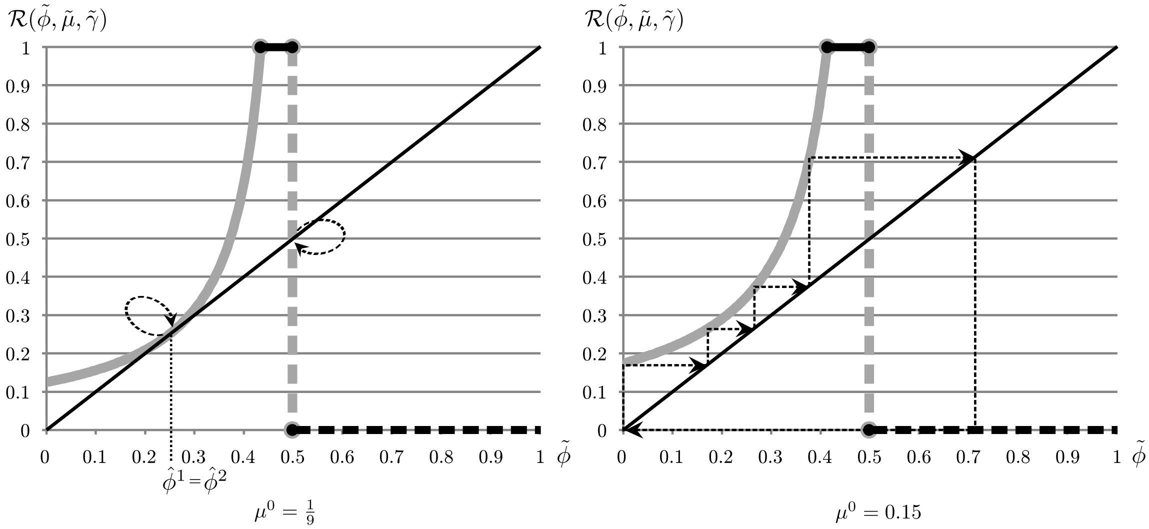

Figure 2.5.6: Left, $\mathcal{R}(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ is plotted for $\mu^0 = \frac{1}{9}$, where there is just a single low payoff stationary equilibrium. Right, $\mathcal{R}(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ is plotted for $\mu^0 = 0.15$, where the minimum equilibrium payoff is given by a non-stationary equilibrium traced out.

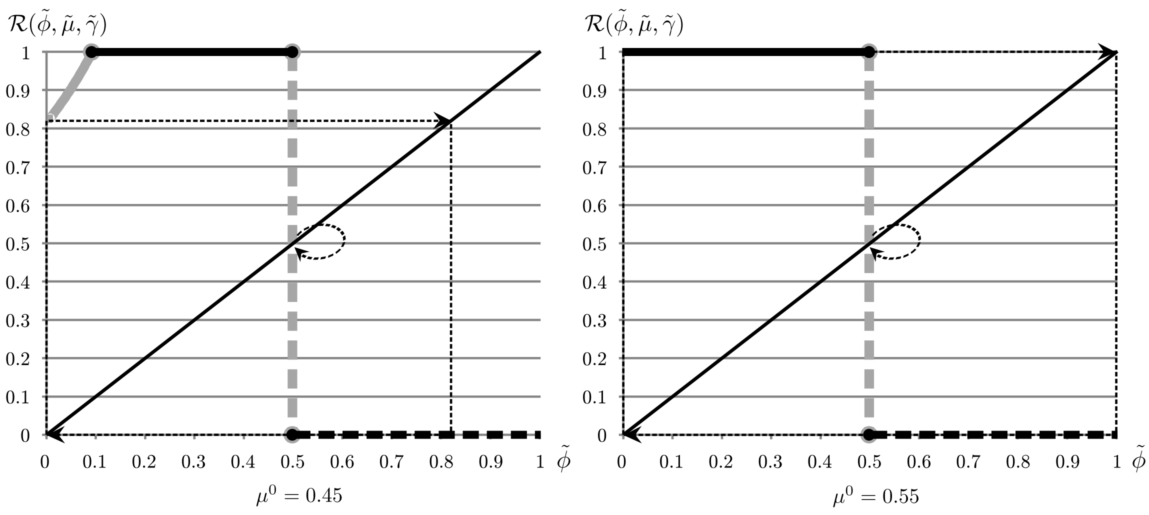

Figure 2.5.7: Left, the graph of $\mathcal{R}$ for $\mu^0 = 0.45$, with the minimum payoff non-stationary equilibrium traced out. Right, the graph of $\mathcal{R}$ for $\mu^0 = 0.55$, with the maximum payoff non-stationary equilibrium traced out. In both cases, the stationary equilibrium with payoff 2 is depicted, which corresponds to the maximum equilibrium payoff on the left and the minimum on the right.

just a two period cycle, alternating between the low payoff and the very high payoff. For $\mu^0 > \frac{1}{2}$, the solid gray curve disappears completely. This allows a higher equilibrium payoff than 2: a cycle alternating between a high payoff of 2 (playing C after player 2 plays c) and very high payoff of 3 (playing D after player 2 plays c). In this case, the lowest payoff equilibrium is the "high" payoff stationary one.

2.5.1.3 Non-Quasi-Markov Equilibria

How restrictive is the assumption of quasi-Markov perfection, and what are equilibria failing this refinement like? Though I do not attempt a general answer because (as discussed at the beginning of Section 2.4) solving for the full set of equilibrium payoffs is very complicated, I construct a class of non-quasi-Markov equilibria that expand the set of equilibrium payoffs for low priors. For all $\mu^0 \in [0, \frac{1}{3}]$, there exist stationary equilibria that give every payoff in the interval $[1, 2]$. In the complete information case ($\mu^0 = 0$), there are such equilibria giving player 1 the "high" payoff of 2 where BMM show the only purifiable payoff is 1. For priors $\mu^0 \in (\frac{1}{9}, \frac{1}{3}]$, there are equilibria with payoff 1 that survive at higher priors.

I now construct these equilibria. For brevity, define $\phi^* \equiv \frac{\mu^0}{1-\mu^0}$. Define player 1's strategy as $\sigma_1(\varnothing a_2) = \frac{1}{2}(1 - \phi^*)$, $\sigma_1(Da_2) = \frac{1}{2}$ and

$$\sigma_1(Ca_2) = \frac{1 - 3\phi^*}{2(1 - \phi^*)}$$

for both $a_2 \in \{c, d\}$. It can be checked that this strategy makes player 2 indifferent at all histories $\varnothing, C, D$ (at all periods). To keep player 1 indifferent, it must be that $\sigma_2^t(C) = \sigma_2^t(D) + \frac{1}{2\delta}$, but otherwise can choose any $\sigma_2^t(C) = \left[\frac{1}{2\delta}, 1\right]$. I am free to choose any $\sigma_2^0(\varnothing)$ without affecting player 1's incentives. The equilibrium payoff can be calculated by evaluating the strategy of playing D every period (a best response):

$$(1 - \delta)\Big[\sigma_2(\varnothing)u_1(c, C) + (1 - \sigma_2(\varnothing))u_1(d, C)$$

$$
\begin{aligned}
& + \sum_{t=1}^{\infty} \delta^t [\sigma_2(C) u_1(c, C) + (1 - \sigma_2(C)) u_1(d, C)] \Bigg] \\
= & \ (1-\delta) \left[2\sigma_2(\varnothing) + \sum_{t=1}^{\infty} 2\delta^t \sigma_2(C) \right] \\
= & \ 2(1-\delta)\sigma_2(\varnothing) + 2\delta\sigma_2(C).
\end{aligned}
$$

I can then choose any $\sigma_2(\varnothing) \in [0,1], \sigma_2(C) \in \left[\frac{1}{2\delta}, 1\right]$ to get an equilibrium with a payoff between 1 and 2.

These equilibria feature relatively complicated and arbitrary mixing by both players in order to keep each other indifferent. At $\mu^0 = 0$, BMM show that the intuition that such behavior is unrealistic is confirmed by the fact that these equilibria are not purifiable, and that the only purifiable equilibrium is the repeated one-shot equilibrium. Requiring quasi-Markov perfection rules out similarly unrealistic equilibria at positive priors.

2.5.2 Product Choice Game with Long Records (Stationary)

The analysis of the Section 2.5.1 shows that with 1-period memory, a focus on stationary equilibria is restrictive. Nevertheless, as K increases, so do the dimensions of the HBP space. Given the complexity of just 1-period memory, studying non-stationary equilibria for long memory may require numerical methods similar to the techniques developed by Judd, Yeltekin, and Conklin (2003) for the APS environment.

With that in mind, this section studies stationary, purifiable equilibrium payoffs for long records in the product choice game. Taking advantage of the simplicity of the stationary environment, I also relax the assumption of constant cost of effort by player 1 with respect to player 2's action, using the more general stage game in Figure 2.5.8.

I first start with a relatively straightforward result that the continuation payoff at the clean history is the Stackelberg payoff (2) for long memory, similar to the bound of Theorem 2 in Liu and Skrzypacz (2014b).[27] Recall that Proposition 2.3.6 shows any

[27] Liu and Skrzypacz's result for their continuous product choice game also applies to arbitrary priors on

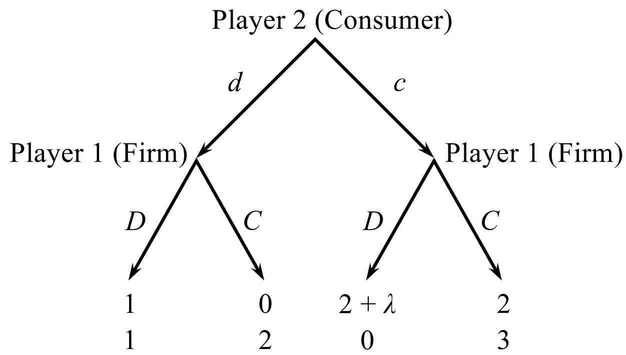

Figure 2.5.8: The product choice game, with player 1 payoffs on top and player 2 payoffs on bottom, where $\lambda > 0$.

stationary equilibrium corresponds to a self-generating HBP (ϕ, μ, γ). The intuition is as follows. A continuation payoff $\gamma(I_K) < 2$ requires that player 2 sometimes plays d at the clean history. This requires both that the belief $\mu(C^K)$ at the clean history be low (so $\phi(I_K)$ must be sufficiently high), and that player 1's strategy at the clean history C^K must play D sufficiently frequently (so $\phi(I_0)$ must be also be sufficiently large). Playing D frequently at the clean history I_K means the "yesterday-dirty" histories I_0 must be reached sufficiently frequently. Thus, the probability mass at $\phi(I_K)$ must be high at the same time that enough mass is "flowing out to I_0" to make player 2 willing to play d at I_K. To keep the clean history probability $\phi(I_K)$ sufficiently high requires that enough mass is flowing in from I_{K-1} (the history one period away from being clean), which also requires enough from I_{K-2}, and so on. When K is large, this "stretches" the probability distribution $\phi(\cdot|\theta_0)$ over all the histories $I_0, I_1, ..., I_{K-1}$, until there is a contradiction because the sum of the probabilities must be greater than 1.

Proposition 2.5.2. *Suppose $\mu^0 > 0$. There exists $K^*(\mu^0)$ such that for all $K > K^*$, $V(C^K) \geq 2$.*

Let $\underline{e}$ denote the minimum equilibrium payoff. This result shows that for large K, player 1 gets close to the Stackelberg payoff when sufficiently patient, i.e. $\underline{e}$ is bounded by

the calendar date in their model, while I assume the improper uniform prior.

$(1-\delta)\cdot 0+\delta^K V(C^K)$, following the standard Fudenberg and Levine (1989) argument of bounding the payoff of deviating to the commitment action C. Thus, $\lim_{\delta\to 1}\lim_{K\to\infty}\underline{e}\geq 2$. However, the order of limits is critical; reversing them gives this lower bound no bite: $\lim_{K\to\infty}\lim_{\delta\to 1}\underline{e}\geq 0$. This makes the welfare impact (on player 1) of increasing memory unclear. Longer memory ensures that the payoff of the clean history is high, but it also makes it harder to clean the history.

The following result shows that under purifiability, there is no tradeoff: as long as the discount factor is above a threshold dependent only on the stage game payoffs (not the prior), player 1 gets exactly the Stackelberg payoff when memory is sufficiently long.

Proposition 2.5.3. *Suppose $\mu^0>0$ and $\delta>\frac{\max\{\lambda,1\}}{1+\max\{\lambda,1\}}$. There exists $K^*(\mu^0)$ such that for all $K>K^*$, the player 1 payoff is exactly 2 for any stationary, purifiable PBE.*

Note that if $\delta<\frac{\max\{\lambda,1\}}{1+\max\{\lambda,1\}}$, then after player 2 plays c, player 1's impatience makes it infeasible for future incentives to outweigh the myopic incentive of playing D. For such low δ, for all priors below a threshold (dependent only on stage payoffs) the only equilibrium outcome is the repeated static Nash (even without purifiability), as was the case in (2.5.2).

I discuss the intuition of the proof for $\lambda=1$, referring the interested reader to the proof for the more complicated general case.[28] This is simpler because $\lambda=1$ means c and d are incentive-equivalent for player 1, so player 1 does not condition on player 2's action. First, it must be that C is a best response at every history or the continuation payoff of I_0 (where D has just been played) is 1. To see why, if D is a strict best response at some I_k, then player 2 will always play d at history I_k, which will be followed by the continuation payoff $V(I_0)$ because player 1 plays D. This continuation payoff $V(I_k)=(1-\delta)+\delta V(I_0)$ cannot be better than $V(I_0)$, so there is no intertemporal incentive to play C at history I_{k-1}. By backward induction, D is a best response at I_0 and so $V(I_0)=1$ (the minmax).

[28] $\lambda>1$ means the short-run incentive to exploit is greater when the short-run player is more trusting. Liu and Skrzypacz (2014b) make the analogous assumption in their continuous stage game to find reputation bubble behavior, but the proof of Proposition 2.5.3 indicates this behavior does not exist in the game studied here; specifically, the strict preference to "prick" the bubble at the clean history is ruled out — contrast Case 2 on page 177 with Lemma 6 and Corollary 1 of Liu and Skrzypacz.

The next result is that for high enough K, the continuation payoff for the clean history is exactly $V(I_K) = 2$. If $V(I_K) < 2$, then player 2 would have to play d with positive probability at the clean history ($\alpha_2(I_K) < 1$), which means player 1 would sometimes have to play D with at least probability $\frac{1}{2}$. Recall from Proposition 2.3.6 there exists a self-generating HBP $w \equiv (\phi, \mu, \gamma)$. The fact that d is a best response at I_K implies that $\mu(I_K) \leq \frac{1}{2}$, which requires the normal player spend sufficient time at I_K, i.e. $\phi(I_K|\theta_0) \geq \frac{1-\mu^0}{\mu^0}$. However, playing D with at least probability $\frac{1}{2}$ means player 1 is also spending time at the freshly dirty history I_0. The long-run player must also spend time at history classes $I_1, ..., I_{K-1}$ "cleaning" the history to arrive at I_K frequently enough. When memory is long, satisfying all these requirements becomes impossible as the history distribution gets "stretched out," giving a contradiction.

If player 1 strictly prefers C at the clean history C^K, then she also prefers it at initial history C^{K-1}, so player 2 plays c at C^{K-1} and by backward induction the equilibrium payoff is 2. If player 1 strictly preferred D at the clean history, she would also prefer it at I_{K-1}, and so by the backward induction argument given above, the freshly dirty history continuation payoff is $V(I_0) = 1$, which would make the long-run player strictly prefer d at C^K, a contradiction.

Supposing that player 1 is indifferent at C^K, it must be that for each $k \in \{1, ..., K - 1\}$, the continuation payoffs satisfy $V(I_k) \geq V(I_0) + \frac{1-\delta}{\delta} = V(I_K) = 2$ because C must be a best response. This similarly implies that $V(C^k) \geq 2$ for each $k \in \{1, ..., K-1\}$. Because the incentive to play C cannot be less, purifiability requires that the long-run player play C at least as much in the initial histories $\varnothing, C^1, ..., C^{K-1}$ as in the dirty histories $I_0, I_1, ..., I_{K-1}$. The short-run player must therefore play c at least as much as well. However, because the belief at a clean initial history is strictly positive (at least μ^0), the incentive to play c is actually strictly greater. Then if player 2 is indifferent at dirty history I_k, they must always play c at clean history C^k. It must be that player 2 mixes at the freshly dirty history I_0; always playing c at I_0 would make player 1 strictly prefer D at the clean history C^K, and always playing d would make player 1 strictly prefer C at the clean history — either is a contradiction. Since player 2 mixes at I_0, she must strictly prefer c at the initial history $\varnothing$. Thus, the equilibrium payoff is $V(\varnothing) = (1-\delta)u_1(c, C) + \delta V(I_1) \geq 2$, which holds with equality since a stationary equilibrium with payoff greater than 2 is impossible.

2.6 Conclusion

This chapter lays out two theoretical tools for the study of bounded memory reputation games with sequential-move stage games where the short-run player moves first. Extending the self-generation methods of APS to bounded memory reputation, I derive a recursive characterization of the set of equilibrium payoffs, which allows study of non-stationary equilibria and gives a simple interpretation for stationary equilibria as self-generating "points." I also define a simplifying equilibrium refinement, quasi-Markov perfection, and extend the results of BMM to show that this is a necessary condition of purifiable equilibria.

These tools are applied to a product choice game example with a "honest" Stackelberg commitment type, where only the most recent long-run player actions are observed. In the 1-memory case, I obtain a complete characterization of the minimum and maximum quasi-Markov equilibrium payoffs for all priors and almost all discount factors, giving insight into how restrictive the assumption of stationarity is. I also show that for long memory, even when not especially patient, the long-run player obtains exactly the Stackelberg payoff in all purifiable stationary equilibria even for very low priors. Thus, for bounded memory in this environment, long memory is important for guaranteeing payoffs, rather than patience.

Although the recursive framework is specified for the applications examined in this chapter, it is worth noting that it is straightforward to extend it to simultaneous move stage games, imperfect monitoring of player 2's action by player 1, and even multiple long-run players who have bounded memory (with beliefs on each other's types). An application of the latter case might be firms whose employees and managers turn over at regular intervals, giving them bounded institutional memory. An interesting direction for future research is the development of a computational implementation of the algorithm presented here — as was done by Judd, Yeltekin, and Conklin (2003) for the original complete information, full memory APS algorithm — which would allow study of games too complicated for analytical solutions.

Chapter 3

Drug Wars and Government Policy

3.1 Introduction

On December 1, 2006, Felipe Calderon took office as president of Mexico. Within two weeks, he deployed 6500 federal troops to combat drug trafficking organizations (DTOs) in his home state of Michoacán, which had experienced a wave of execution style killings (Grillo, 2006). This marked the beginning of the massive crackdown on DTOs that would become a hallmark of his administration, involving 45,000 troops by the time he left office in 2012 (Dell, 2011). Before Calderon's presidency, the Mexican National Human Rights Commission (Comisión Nacional de los Derechos Humanos) reports a total of 8901 drug-related homicides from 2001 to 2006, an average of 1484 per year . As the government ramped up its war on the cartels over subsequent years, drug war violence exploded, reaching over 10,000 per year by 2010. By the end of Calderon's presidency in 2012, at least 60,000 people were killed, though some reports suggest the number is much higher (Molzahn, Rodriguez, and Shirk, 2013).

Although it is fairly straightforward to see why government-vs-cartel violence would increase following a crackdown — a crackdown by definition increases government interaction with DTOs — it is less obvious why government intervention would increase violence between the cartels. In fact, the vast majority of drug-related homicides are due to violence between the cartels. The most reliable data on drug-related homicides come from Mexico's National Security Council (Consejo de Seguridad Nacional, CSN), who publicly reported drug-related homicides in Mexico from December 2006 (when the crackdown began) until September 2011. In CSN's data, 89.3% of casualties are "targeted executions linked to drug-trafficking operations" (Rios, 2013).

This chapter develops a model that offers a theoretical explanation for why DTOs engage in violence with each other, and why violence increased during the Calderon's crackdown between 2006 to 2012 but not during previous law enforcement operations against DTOs. In the model, the threat of violence is used to enforce collusion between the drug cartels in order to increase profits. Since cartels are not able to observe defection, but instead only observe a noisy signal of actions, punishments and therefore violence occur in equilibrium. The government has the power to arrest traffickers, which makes drug trafficking more expensive. If the government arrests traffickers only when cartels are punishing each other, punishments become harsher and therefore allow more collusive behavior with less frequent violence. One can think of this policy as "corrupt" because it effectively helps traffickers cooperate and maximizes profits, which could be desirable if the government takes some fixed percentage of profits as a bribe. This policy also minimizes violence. If instead the government cracks down indiscriminately, always arresting traffickers, this cooperative incentive goes away, and so long as the government's crackdown is sufficiently bounded (by its capacity to arrest), violence increases even though drug trafficking has become less profitable.

The model is based on the Green and Porter (1984) model of collusion under imperfect monitoring. Two firms play a repeated modified Cournot duopoly game, where they choose both a quantity of drug traffickers to hire and whether or not to attack each other, thereby killing some of the other firm's traffickers. Cartels find it profitable to attack in the short-run because killing opposing traffickers reduces the quantity of competing drugs delivered to market, thereby raising the price and the firm's profits. It is possible to achieve higher payoffs by enforcing collusive behavior through the threat of punishments, which include violence. However, since quantities are only imperfectly monitored, punishments and violence occur on the equilibrium path. The firms play the optimal equilibrium among a class of equilibria similar to those constructed by Green and Porter. If the government arrests traffickers (which is modeled as having the same effect on profits as killing them, except that cartels do not have to hire assassins) during punishments, then punishments become harsher and are able to enforce smaller quantities and higher profits. This allows punishments to be triggered less frequently, resulting in less violence. By contrast, if the government arrests every period no matter the state, the incentive to cooperate is reduced and so punishments are triggered more frequently.

The most technically difficult task here is characterizing the optimal equilibrium. I restrict attention to a class of equilibria I call "Green-Porter equilibria" (GPE) because of their similarity to the equilibria constructed in Green and Porter (1984). A GPE is in one of two states: the reward state or the punishment state. The reward state has firms producing smaller quantities than the static equilibrium, and transitions to the punishment state upon observation of either killing (which is perfectly monitored) or low prices. In the punishment state, firms play the static Nash equilibrium, which involves killing and Cournot quantities.[1] The firms observe a public correlation device and randomly choose to return to the reward state after some given realizations of the device, rather than the price, such that there are no intertemporal incentives and the static Nash equilibrium is incentive compatible. I use a simplified version of Judd, Yeltekin, and Conklin's (2003) implementation of the recursive Abreu, Pearce, and Stacchetti (1990) method to numerically characterize these optimal equilibria.

The fact that an explosion in violence occurred just after Calderon's announcement and grew worse as he pursued this policy raises the question of whether the crackdown caused the increase in violence, or if the crackdown simply pre-empted an inflammation of an existing conflict between cartels. Indeed, the latter case suggests that the violence might have been even worse in the absence of this policy. However, empirical evidence indicates that the Calderon crackdown did actually cause increases in violence. Looking at municipalities holding mayoral elections in 2007 and 2008 where the margin of victory was within 5%, Dell (2011) finds that municipalities which elected a candidate from Calderon's conservative PAN party were more likely to subsequently experience drug violence than those electing non-PAN candidates. Since the margin of victory is so close, Dell argues the

[1] To clarify, these are Cournot quantities of drugs *delivered* to market, equal to the number of traffickers who survive both attacks from the other cartel and arrests from the government. This requires firms to *hire* a larger number of traffickers than in the Cournot game without attacks or arrests.

outcomes of these elections provide exogenous variation in law enforcement activity, since PAN mayors were politically aligned with PAN President Calderon.

One may be tempted to think that crackdowns would always spark conflict through the "hydra effect" — that removing the head of a cartel opens up competition between his lieutenants — yet this does not always appear to be the case. Rios (2014) points to law enforcement operations against DTOs before 2000, when Mexico was a one-party state ruled by the liberal PRI party, "which resulted not in violent confrontations, but in the maintenance of a highly disciplined group of oligopolistic criminal organizations that operated without fighting each other." When the leader of the then most-powerful Guadalajara cartel, Felix Gallardo, was arrested in 1989 following the capture, torture, and murder of DEA agent Enrique Camarena, his cartel split peacefully into some of the forerunners of today's major cartels, dividing territories as reportedly agreed upon at a conference in Acapulco (Rios, 2014). Rios particularly points to the end of one-party-rule in 2000 as preventing the government from coordinating in a way that could enforce cooperation between cartels, as PAN and PRI controlled different parts of the Mexican government. By contrast, PRI-controlled Mexico featured explicit corruption agreements between the government and cartels, requiring that cartels only traffic within their territories, not kill each other on the streets, and that they not sell drugs to Mexicans (instead only transporting them to the US). Rios shows that as PRI lost its grip on the Mexican government, markets for consumer cocaine *within* Mexico began to open up, suggesting a collapse of these agreements. Although the model presented here cannot, at least in its present simple form, capture all of these rich dynamics, it provides a starting point for analyzing how some government enforcement operations can make violence less prevalent, while others can lead to more war.

3.2 Model

My model of the drug market is based on the repeated Cournot game with imperfect monitoring of Green and Porter (1984). The stage game is modified by giving firms the option to attack each other at a cost, thereby removing competing drugs from the market.

These attacks are what I refer to as "violence."[2] This is individually profitable for the firm because it raises the price and therefore the firm's own revenue. I also allow an exogenous government to arrest traffickers, which similarly removes drugs from the market.

3.2.1 Stage Game

There are two firms, 1 and 2. The stage game is as follows. Each firm i simultaneously chooses to hire some quantity $q_i \in \mathcal{Q}$ of traffickers, where $\mathcal{Q}$ is a fine grid on an interval $[0, Q]$, and whether to attack $s_i \in \{0, 1\}$, where $s_i = 1$ indicates attacking and $s_i = 0$ indicates not. I assume that $Q > \frac{1}{3}(r - c)$ so that this upper bound is never binding. Let $A_i \equiv \mathcal{Q} \times \{0, 1\}$ denote the action space of firm i, $A \equiv A_1 \times A_2$ the set of action profiles, and ΔX the set of probability distributions over X. The exogenous government chooses to arrest some number $g \in \mathcal{G}$ traffickers of each cartel, where $\mathcal{G}$ is a fine grid on an interval $[0, G]$.

Firm i pays some constant marginal cost c to hire each unit of traffickers, who each carry 1 unit of drugs. If firm i attacks (i.e., $s_i = 1$), then κ traffickers from firm $-i$ are killed and do not deliver their drugs to market, and firm i pays some cost η to hire the assassins. If the government arrests g traffickers, those traffickers also do not deliver their drugs. Let $\hat{q}_i = \max\{q_i - \kappa s_{-i} - g, 0\}$ be the quantity of drugs firm i delivers to market. The firms face a demand curve

$$p(a) = r - (\hat{q}_1(a) + \hat{q}_2(a))$$

where $a \equiv ((q_i, q_{-i}), (s_i, s_{-i}))$ is a strategy profile and $r > 0$ is some constant. Firms receive expected profit[3]

$$u_i(a) = p(a)\hat{q}_i(a) - cq_i - \eta s_i.$$

The stage game profits are the revenues from drugs delivered minus the cost of hiring all

[2] Violence is defined formally in (3.4.2).

[3] I use the term *expected* profit because I later introduce imperfect monitoring of the choice of quantities q_i, q_{-i} by making the publicly observable price p subject to some mean-preserving shocks. Since firms are risk neutral, this has no effect on the one-shot game.

traffickers (including those arrested or killed) and the cost of assassinations:

$$u_i(a) = p(a)\hat{q}_i(a) - cq_i - \eta s_i$$

It is straightforward to show that the maximum feasible payoff is obtained by the firms not attacking ($s_i = 0$) and each choosing quantity $q_i \equiv q^m = \frac{1}{4}(r-c)$, which I refer to the "monopoly quantity."[4] For simplicity, I assume that the cost η of attacking is sufficiently small such that attacking is a best response so long as firm i delivers at least the monopoly quantity:

$$\eta < \kappa q^m. \tag{3.2.1}$$

I also assume that feasible killings are bounded by

$$\kappa < \tfrac{1}{2}(r-c), \tag{3.2.2}$$

since otherwise a profitable one-shot equilibrium with violence is impossible.

3.2.2 Repeated Game

I assume that the firms play the infinitely repeated stage game, maximizing their discounted average payoffs given discount factor $\delta \in (0,1)$. Attacks (s_1, s_2) are publicly observed, but the quantities (q_1, q_2) are not. Instead they observe the price $p(\theta, a) = \theta p(a)$ subject to some shock θ distributed log-normally, with mean one and variance $\exp(\zeta^2) - 1$, according to cdf

$$F(\theta) = \Phi\left(\frac{\log\theta + \frac{1}{2}\zeta^2}{\zeta}\right)$$

and pdf

$$f(\theta) = \frac{1}{\theta\zeta\sqrt{2\pi}}\exp\left(-\frac{(\log\theta + \frac{1}{2}\zeta^2)^2}{2\zeta^2}\right) = \frac{1}{\zeta\theta}\phi\left(\frac{\log\theta + \frac{1}{2}\zeta^2}{\zeta}\right),$$

[4] This is, of course, obtainable by any action profile yielding total output $\frac{1}{2}(r-c)$, but it will be more natural to use per-firm values since I focus on symmetric equilibria.

where Φ, ϕ are the cdf and pdf, respectively, of the standard normal distribution. Since $E[\theta] = 1$, it does not affect expected stage game payoffs. Players also observe a public correlation device ω, which is uniformly distributed on $[0, 1]$.

I restrict attention to a class of equilibria that I call *Green-Porter equilibria* (GPE), based on the equilibria described in Green and Porter (1984). These are equilibria which are either in a "reward" state or a "punishment" state. The government chooses a policy $(\bar{g}, \tilde{g}) \in \mathcal{G}^2$, where $\bar{g}$ is the number of arrests in the reward state and $\tilde{g}$ is the number of arrests in the punishment state. A GPE begins in the reward state and remains there until either a price below some threshold $\bar{p}$ or an attack are observed. In the reward state, the firms play an action profile $\bar{a}$, and the government plays $\bar{g}$. In the punishment state, the firms play the static Nash equilibrium $\tilde{a}$ given the government's policy $\tilde{g}$. The punishment state ends when the public correlation device is realized below some threshold $\tilde{\omega}$.

A GPE where punishment state has mean duration $T = \frac{1}{\tilde{\omega}}$ will be referred to as a T-GPE. The arguments in Porter (1983) show that choosing $T = \infty$ (permanent punishment) yields the maximum possible payoff. Since never-ending war between drug traffickers may not be realistic, perhaps due to renegotiation, I let $T \in [1, \infty]$ be exogenous and pick $T = 20$ for the example presented in Section 3.4.[5] I assume the firms play the optimal such equilibrium.

Assumption 3. *Given $T \in [1, \infty]$ and policy $(\bar{g}, \tilde{g})$, the firms play the T-GPE yielding the maximum payoff.*

3.3 Numerical Solution

This section presents a numerical method for characterizing the optimal T-GPE when the grid of quantities $\mathcal{Q}$ is "sufficiently fine," meaning that the results are not changed by adding points to the grid. I define this formally as follows: a statement X is true when

[5] Choosing $T = \infty$ does not qualitatively change the results presented in Section 3.4.

$\mathcal{Q}$ is a *sufficiently fine grid* on $[0, Q]$ if and only if there exists a finite subset $\mathcal{F} \subset [0, Q]$ such that $\mathcal{F} \subset \mathcal{Q}$ implies that X is true.

I first characterize the symmetric static Nash equilibria for all government policies $g \in \mathcal{G}$.

Lemma 3.3.1. *Define* $\mathfrak{q}(\lambda) \equiv \frac{1}{2}(r - c - \lambda)$*, and implicitly define function* $\check{L}(g, \check{s})$ *by*

$$\check{L}(g, \check{s}) = r - c - 2\sqrt{\left(c\kappa \frac{\check{L}(g, \check{s})}{\mathfrak{q}(\check{L}(g, \check{s}))} + \eta\right) \check{s} + cg}. \tag{3.3.1}$$

$\check{L}(g, \check{s})$ *is strictly decreasing in* $\check{s}$.

Proposition 3.3.1. *For a policy* $g \in \mathcal{G}$*, let a symmetric Nash equilibrium of the one-shot game be given. Let* $E[q_i^*], E[s_i^*]$ *denote the expected actions in the equilibrium. When* $\mathcal{Q}$ *is a sufficiently fine grid on* $[0, Q]$,

$$E[q_i^*] = \begin{cases} \frac{1}{3}(r - c) + \eta s_i^* + g & g \leq \frac{1}{c}\left[\frac{1}{9}(r - c)^2 + \eta - c\kappa\right] \\ r - c - 2\sqrt{(c\kappa + \eta)E[s_i^*] + cg} + \eta E[s_i^*] + g & \text{otherwise} \end{cases} \tag{3.3.2}$$

$$E[s_i^*] = \begin{cases} 1 & g \leq \frac{1}{c}\left[\frac{1}{9}(r - c)^2 + \eta - c\kappa\right] \\ \xi(g) & \frac{1}{c}\left[\frac{1}{9}(r - c)^2 + \eta - c\kappa\right] < g \leq \frac{1}{c}\left(\frac{1}{4}\left[(r - c) - \frac{\eta}{\kappa}\right]^2 - (c\kappa + \eta)\right) \\ \chi(g) & \frac{1}{c}\left(\frac{1}{4}\left[(r - c) - \frac{\eta}{\kappa}\right]^2 - (c\kappa + \eta)\right) < g < \frac{1}{4c}\left[(r - c) - \frac{\eta}{\kappa}\right]^2 \\ 0 & \text{otherwise,} \end{cases} \tag{3.3.3}$$

where $\xi(g)$ *is the (unique) nonnegative solution to the cubic equation*

$$0 = c\kappa\xi^3 + 4(c\kappa + \eta + cg)\xi^2 + 4(c\kappa + \eta + cg)\xi + 4(\eta + cg) - (r - c)^2, \tag{3.3.4}$$

and

$$\chi(g) = \begin{cases} \xi(g) & \check{L}(g, 1) > B_0 \\ 0 & \check{L}(g, 0) < B_0 \\ \frac{\check{L}(g, \check{s}^*)}{\mathfrak{q}(\check{L}(g, \check{s}^*))}\check{s}^* & \text{otherwise,} \end{cases} \tag{3.3.5}$$

where $B_0 \equiv \frac{1}{2}\left[(r - c - \sqrt{(r - c)^2 - 8\eta}\right]$ *and* $\check{s}^*$ *solves* $\check{L}(g, \check{s}^*) = B_0$.

Note that the symmetric equilibrium outcomes are unique (when $\mathcal{Q}$ is sufficiently fine).

Remark 3.3.1. For $g \leq \frac{1}{c}\left[\frac{1}{9}(r-c)^2 + \eta - c\kappa\right]$, the one-shot equilibrium is pure (when $\mathcal{Q}$ is sufficiently fine). For g above this threshold, the one-shot equilibrium involves mixing between "shutting down" ($q_i = 0, s_i = 0$) and an action with generally positive values $(\check{q}_i, \check{s}_i)$. These mixing equilibria have value zero, since shutting down is a best response, even though drugs and violence happen in equilibrium (in expectation).

Remark 3.3.2. Proposition 3.3.1 gives simple closed form solutions for the equilibrium outcomes except for a range of policies $g \in \left(\frac{1}{c}\left[\frac{1}{9}(r-c)^2 + \eta - c\kappa\right], \frac{1}{4c}\left[(r-c) - \frac{\eta}{\kappa}\right]^2\right)$, which are defined implicitly. For these policies, I calculate $E[s_i^*]$ numerically. Note that, although $\xi(g)$ has a closed form solution, I use the faster and simpler approach of using Newton's method. For $\chi(g)$, it is possible solve $\check{L}(g, \check{s}^*) = B_0$ through bisection, evaluating $\check{L}(g, \check{s})$ at each iteration by solving (3.3.1) via Newton's method.

For the purposes of solving the optimal T-GPE and make use of existing results, it will be useful to introduce a slightly more general equilibrium definition, the *non-stationary Green-Porter equilibrium* (NGPE) (a T-NGPE is analogously defined). An T-NGPE is the same as a T-GPE, except that the firms (but not the government) may condition on the calendar date t during the reward state. Thus, the T-GPEs are a subset of the T-NGPEs. I will show how to solve for the optimal T-NGPE payoff, which turns out to be a T-GPE payoff as well (in Corollary 3.3.2).

Let $\mathscr{E} \subset \mathbb{R}$ denote the set of T-NGPE equilibrium payoffs (NGPEs are symmetric so the set of equilibrium payoffs is one dimensional), and let $\bar{\mathscr{E}} \equiv \max \mathscr{E}$ be the payoff of the optimal T-NGPE. I characterize $\bar{\mathscr{E}}$ through a simplified version of the method of Abreu, Pearce, and Stacchetti (1986; 1990) (borrowing notation from Mailath and Samuelson (2006b)) and the numerical implementation developed by Judd, Yeltekin, and Conklin (2003).

Definition 3.3.1. Let some set $W \subset \mathbb{R}$ be given. A mixed action profile α is *enforceable* on W if there exists $\bar{p}$ and $V \in W$ such that

$$
\begin{aligned}
& \mathsf{V}(\alpha, \bar{p}, V) \\
\equiv\ & (1-\delta)u_i(\alpha) + \delta\left[\left(1 - F\left(\frac{\bar{p}}{p(\alpha)}\right)\right)V + F\left(\frac{\bar{p}}{p(\alpha)}\right)((1-\delta^T)u^N + \delta^T V)\right] \quad (3.3.6)\\
\geq\ & (1-\delta)u_i(\alpha_i', \alpha_{-i}) \\
& +\delta\left[\left(1 - F\left(\frac{\bar{p}}{p(\alpha_i', \alpha_{-i})}\right)\right)V + F\left(\frac{\bar{p}}{p(\alpha_i', \alpha_{-i})}\right)((1-\delta^T)u^N + \delta^T V)\right] \quad (3.3.7)
\end{aligned}
$$

for all $\alpha_i \in \Delta A_i$. The price $\bar{p}$ and reward payoff V *enforce* α (on W).

Similarly I adapt the notion of decomposition and the "generating operator" $\mathscr{B}(\cdot)$.

Definition 3.3.2. A payoff $V \in \mathbb{R}$ is *decomposable* on $W \subset \mathbb{R}$ if there exists action profile α enforced by some price $\bar{p}$ and $V' \in W$ such that $V = \mathsf{V}(\alpha, \bar{p}, V')$. V is *decomposed* by $\alpha, \bar{p}, V'$ (on W). Define

$$\mathscr{B}(W) \equiv \{V \in \mathbb{R} : V = \mathsf{V}(\alpha, \bar{p}, V') \text{ for some } \alpha \text{ enforced by } \bar{p} \text{ and } V' \in W\}.$$

A straightforward, much simpler version of the arguments in Abreu, Pearce, and Stacchetti (1990) show the following. (Since the arguments are nearly identical to those in Appendices 2.1.2 and 2.1.3, replacing terms with their NGPE analogues, I omit them.)

Proposition 3.3.2. *The following hold:*

1. *If $W \subset \mathscr{B}(W)$, i.e. W is a "self-generating set," then W is the payoff of a T-NGPE.*

2. *$\mathscr{B}(\mathscr{E}) = \mathscr{E}$.*

3. *Let $\mathscr{F}^\dagger$ denote the set of feasible payoffs. Then $\lim_{m\to\infty} \bigcap_m \mathscr{B}^m(\mathscr{F}^\dagger) = \mathscr{E}$.*

I now characterize Q, U in order to compute $\bar{\mathscr{B}}(\cdot)$. The following proposition characterizes the interior solution, if it exists.

Proposition 3.3.3. *Let some payoff $\bar{V}$ be given. Let $\tilde{V} = (1-\delta^T)u^N(\tilde{g}) + \delta^T\bar{V}$ be the value of the punishment phase, and define $\Delta V \equiv \bar{V} - \tilde{V}$. Define $\bar{\alpha} \equiv (\bar{q}, \bar{s})$ and $\bar{p}$ which solve*

$$\begin{aligned} \max_{\alpha,\breve{p}} \quad & \mathsf{V}(\alpha, \breve{p}) \\ \textit{such that} \quad & \alpha \textit{ is enforced by } \breve{p} \textit{ and } \bar{V}. \end{aligned}$$

Define $\bar{\theta} \equiv \bar{p}/p(\bar{\alpha})$ and

$$\psi(\bar{\theta}) \equiv C_0 - \sqrt{C_1 + \gamma(\bar{\theta})\Delta V} \tag{3.3.8}$$

where $C_0 \equiv \frac{1}{12}(5r - 2c), C_1 \equiv \frac{1}{144}(r + 2c)^2, \gamma(\bar{\theta}) \equiv \frac{1}{6}\frac{\delta}{1-\delta}f(\bar{\theta})\bar{\theta}$. Define $\bar{\beta} \equiv \kappa\bar{s}_i + \bar{g}$.

Suppose that the solution is interior, i.e., $\bar{\theta} \in (0, \infty)$. Then when Ω is sufficiently fine,

$$\bar{q}_i = \psi(\bar{\theta}) + \bar{\beta}, \tag{3.3.9}$$

$$0 = \big(4\psi(\bar{\theta}) - (r - c)\big)\frac{-(\log\bar{\theta} + \frac{1}{2}\zeta^2)}{12\zeta^2\sqrt{C_1 + \gamma(\bar{\theta})\Delta V}} - 1, \tag{3.3.10}$$

which has a unique solution such that $\bar{\theta} \leq \exp(-\frac{1}{2}\zeta^2)$. Also, $\bar{s}_i = 0$ if and only if $0 \geq (1-\delta)[\kappa\psi(\bar{\theta}) - \eta] - \delta\Delta V$.

Although the equation (3.3.10) does not appear analytically tractable, the fact that the solution is unique and known to be within the interval $(0, \exp(-\frac{1}{2}\zeta^2)]$ allows the use of a root-finding algorithm like Newton's method. The following corollary gives a test for whether the interior solution above is in fact optimal, and if it is not, gives the corner solution that is optimal.

Corollary 3.3.1. *Let $\bar{v}^N$ be the static Nash equilibrium payoff given government arrests $\bar{g}$. If*

$$(1-\delta)u_i(\bar{q}_i, \bar{s}_i) + \delta[(1 - F(\bar{\theta}))\bar{V} + F(\bar{\theta})\tilde{V}] < (1-\delta)\bar{v}^N + \delta\max\{\bar{V}, \tilde{V}\}, \tag{3.3.11}$$

then

$$\bar{\theta} = \begin{cases} 0 & \bar{V} \geq \tilde{V} \\ \infty & \bar{V} < \tilde{V}. \end{cases}$$

and $E[(\bar{q}, \bar{s})]$ is characterized by the one-shot equilibrium in Proposition 3.3.1. Otherwise, the globally optimum $\bar{\theta}$ is the interior solution characterized by Proposition 3.3.3.

The proof is straightforward. First, if $\tilde{V} > \bar{V}$, an interior solution can only enforce quantities higher than the static equilibrium (see (3.3.8)), yielding a lower current period payoff, when the static equilibrium always followed by continuation payoff $\tilde{V}$ is enforceable and decomposes a higher payoff. A corner solution $\bar{\theta} \in \{0, \infty\}$ can only enforce the one-shot equilibrium, so the decomposed payoff must be the right hand side of (3.3.11). If the best interior $\bar{\theta}$ decomposes a higher payoff, the solution is interior. Otherwise, it is best to simply always choose the highest continuation payoff.

Since we are interested only in the maximum payoff $\bar{\mathscr{E}}$, I show that further simplification is possible by focusing on the maximum payoff at each iteration of the algorithm in part (3) of Proposition 3.3.2.

Definition 3.3.3. Define

$$\bar{\mathscr{B}}(W) \equiv \max \mathscr{B}(W), \tag{3.3.12}$$

and define the pair $\mathsf{Q}(V) = \alpha, \mathsf{U}(V) = \bar{p}$ as an action profile α and price $\bar{p}$ which, with some $V \in W$, decompose $\bar{\mathscr{B}}(W)$.

Lemma 3.3.2. *Let closed set $W \subset \mathbb{R}$ be given, and define $\bar{W} = \max W$. Then $\bar{\mathscr{B}}(W)$ is decomposable on the singleton $\{\bar{W}\}$, i.e. $\bar{\mathscr{B}}(W) \in \mathscr{B}(\{\bar{W}\})$, and $\mathsf{Q}(\bar{V}) = \bar{\alpha}, \mathsf{U}(\bar{V}) = \bar{p}$ where $\bar{\alpha}$ and $\bar{p} = \bar{\theta} p(\bar{\alpha})$ are characterized by Proposition 3.3.3 and Corollary 3.3.1.*

This immediately gives the following algorithm for computing a tight upper bound on $\bar{\mathscr{E}}$, and also shows that $\bar{\mathscr{E}}$ is the optimal T-GPE (not just T-NGPE) payoff.

Corollary 3.3.2. *Let $\bar{\mathscr{F}} \equiv \max \mathscr{F}^{\dagger}$ denote the maximum feasible payoff. Then*

$$\lim_{m \to \infty} \bar{\mathscr{B}}^m(\bar{\mathscr{F}}) = \bar{\mathscr{E}}, \tag{3.3.13}$$

and $\bar{\mathscr{E}}$ is the payoff of a T-GPE.

Thus, repeated application of $\bar{\mathscr{B}}(\cdot)$ to the payoff yielded by the monopoly quantity converges to the optimal T-GPE payoff $\bar{\mathscr{E}}$. The sequence $\{\bar{\mathscr{B}}^m(\bar{\mathscr{F}})\}_m$ yields a decreasing sequence of upper bounds on $\bar{\mathscr{E}}$. To prove that $\bar{\mathscr{B}}^m(\bar{\mathscr{F}})$ is within some precision $\varepsilon > 0$ of $\bar{\mathscr{E}}$, I use the following result.

Lemma 3.3.3. *Let some $V \geq u^N$ be given. If $\bar{\mathscr{B}}(\{V\}) \geq V$, then $V \leq \bar{\mathscr{E}}$.*

Following the example of the inner-bound method of Judd, Yeltekin, and Conklin (2003), I use Lemma 3.3.3 to establish a lower bound for $\bar{\mathscr{E}}$. For any m, if $(1-\varepsilon)\bar{\mathscr{B}}^m(\bar{\mathscr{F}}) \leq \bar{\mathscr{B}}((1-\varepsilon)\bar{\mathscr{B}}^m(\bar{\mathscr{F}}))$, then $(1-\varepsilon)\bar{\mathscr{B}}^m(\bar{\mathscr{F}})$ is a lower bound on $\bar{\mathscr{E}}$. Thus, this test establishes that

$$(1-\varepsilon)\bar{\mathscr{B}}^m(\bar{\mathscr{F}}) \leq \bar{\mathscr{E}} \leq \bar{\mathscr{B}}^m(\bar{\mathscr{F}}).$$

Once $\bar{\mathscr{E}}$ is established with sufficient precision, I can back out the strategy from $\mathsf{Q}(\cdot), \mathsf{U}(\cdot)$.

I conclude this section by summarizing the algorithm. The upper-bound algorithm is given by Steps 1 - 4, and Steps 5 - 6 establish the lower bound.

1. Set $\bar{V} \leftarrow u^m$, the maximum feasible payoff u^m (the monopoly payoff).
2. Use Proposition 3.3.3 and Corollary 3.3.1 to compute $\mathsf{Q}(\bar{V}), \mathsf{U}(\bar{V})$.
   (a) Compute interior solution $\bar{\theta}$ for (3.3.10) using Newton's method. Check for corner solution using (3.3.11).
   (b) If solution is interior, set $\mathsf{Q}(\bar{V}) \leftarrow \bar{\alpha}, \mathsf{U}(\bar{V}) = \bar{\theta}p(\bar{q}, \bar{s})$ according to Proposition 3.3.3. Otherwise, set $\mathsf{Q}(\bar{V})$ to the static Nash equilibrium according to Proposition 3.3.1, computed numerically when necessary according to Remark 3.3.2, and
   $$\mathsf{U}(\bar{V}) \leftarrow \begin{cases} 0 & \bar{V} \geq \tilde{V} \\ \infty & \bar{V} < \tilde{V}. \end{cases}$$
3. Set $\bar{V}' \leftarrow \bar{\mathscr{B}}(\bar{V}) = \mathsf{V}(\mathsf{Q}(\bar{V}), \mathsf{U}(\bar{V}), \bar{V})$. Set $\Delta\bar{V} \leftarrow |\bar{V}' - \bar{V}|$. Set $\bar{V} \leftarrow \bar{V}'$.

4. If $|\Delta \bar{V}| \geq \varepsilon$, go to Step 2. Otherwise set $\bar{V}^U \leftarrow \bar{V}$.

5. Set $\bar{V}^L \leftarrow (1-\varepsilon)\bar{V}^U$. Use the method in Step 2 to compute $\mathsf{Q}(\bar{V}^L), \mathsf{U}(\bar{V}^L)$.

6. If $\bar{\mathscr{B}}(\bar{V}^L) = \mathsf{V}(\mathsf{Q}(\bar{V}^L), \mathsf{U}(\bar{V}^L), \bar{V}^L) < \bar{V}^L$, go to Step 2. Otherwise, stop.

Upon a successful conclusion at the end of Step 6, $\bar{\mathscr{E}}$ is guaranteed to satisfy $\bar{V}^U \geq \bar{\mathscr{E}} \geq \bar{V}^L = (1-\varepsilon)\bar{V}^U$, thereby establishing the optimal T-GPE equilibrium payoff within a precision specified by ε.

3.4 Results

I use the computational method described in Section 3.3 to find the outcomes under different government policies. Of particular interest are violence, drug consumption, and profit of the firms. Define the discounted average drug consumption as the following sum over the drugs delivered each period:

$$D = E\left[(1-\delta)\sum_{t=0}^{\infty} \delta^t (\hat{q}_1^t + \hat{q}_2^t)\right]. \tag{3.4.1}$$

Discounted average violence is similarly defined as

$$M = E\left[(1-\delta)\sum_{t=0}^{\infty} \delta^t (s_1^t + s_2^t)\right]. \tag{3.4.2}$$

The parameters used in the example are as follows: $r = 10$, $c = 1$, $\kappa = 0.3$, $\eta = 0.02$, $\zeta = 0.1$, $\delta = 0.95$, and $\tilde{\omega} = 0.05$ (so $T = 20$). I apply the method on a grid $\mathcal{G}$ over $[0, 22]^2$ for different parameters of $(\bar{g}, \tilde{g})$.

The per-firm profit $\bar{\mathscr{E}}(\bar{g}, \tilde{g})$ is plotted in Figure 3.4.1. The plot shows that equilibrium payoffs are always decreasing in $\bar{g}$: arresting during the reward phase always reduces profits. There are also clearly regions where increasing $\bar{\mathscr{E}}$ is decreasing in $\tilde{g}$. For example, at $\bar{g} = 10, \tilde{g} = 0$, the government arrests so much during the reward phase that the optimal equilibrium immediately switches to the punishment phase ($\bar{p} = \infty$), which actually has a

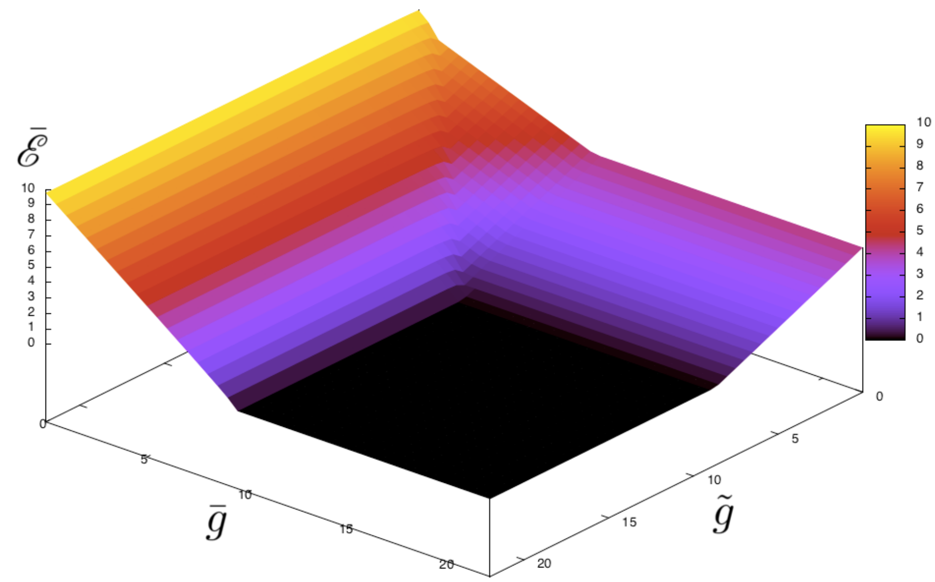

Figure 3.4.1: Discounted average payoffs $\bar{\mathscr{E}}$ over different policies $(\bar{g}, \tilde{g})$.

higher value. Increasing $\tilde{g}$ simply reduces the value of the punishment phase and thus the entire equilibrium.

What is more subtle (and not very visible in Figure 3.4.1) is that profits are actually increasing in $\tilde{g}$ for the region $\tilde{g} \geq \bar{g}, \bar{g} < 8.71$. For example, at $\bar{g} = \tilde{g} = 0$, profit is $\bar{\mathscr{E}} \approx 9.68$, while at $\bar{g} = 0, \tilde{g} = 22$, profit is $\bar{\mathscr{E}} \approx 9.78$. Making punishments harsher allows the enforcement of more collusive behavior without needing to trigger punishments as frequently, yielding higher profits.

Suppose that the government is "corrupt" and receives some fraction of the firm profits as a bribe. If the government seeks to maximize the bribe and therefore profit, the optimal policy is $\bar{g} = 0$, $\tilde{g} = G$. By increasing $\tilde{g}$ and making punishments harsher, more collusive behavior is enforceable during the reward phase and punishments need are not triggered as frequently, so profits actually increase. Thus, a corrupt government will always choose the "leftmost" point in Figure 3.4.1 (this point will be the "frontmost" point, pointing out of the page, in Figures 3.4.1 and 3.4.3).

Note that for region $\bar{g} > 8.72, \tilde{g} > 8.72$, profits are reduced to zero. For these

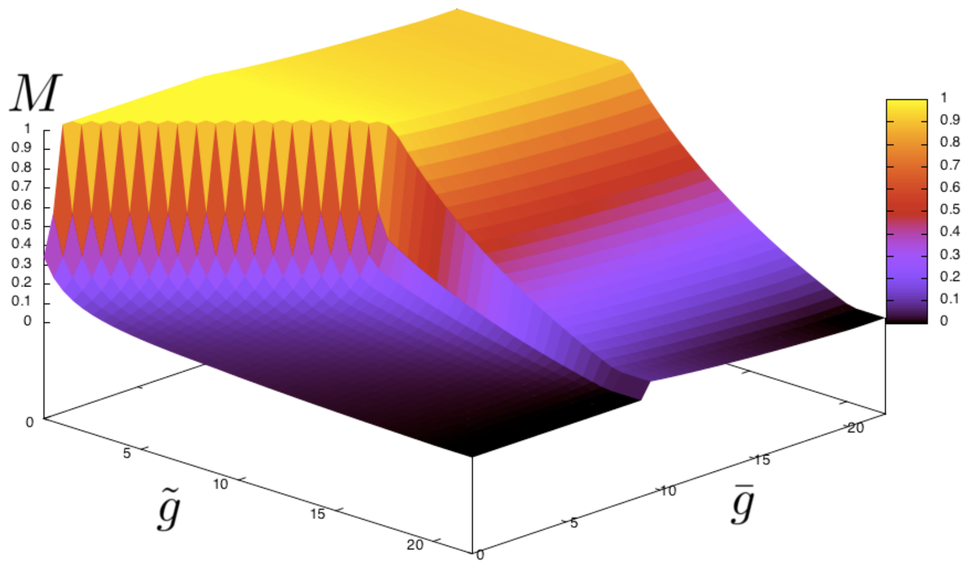

Figure 3.4.2: Discounted average violence M over different policies $(\bar{g}, \tilde{g})$.

policies, the government arrests so frequently that only mixing equilibria exist, where firms mix between delivering a positive quantity and zero drugs to market. Since shutting down and delivering zero drugs is always a best response, the equilibrium payoff is zero. Note, however, that firms still generally deliver positive quantities of drugs in expectation (as seen in Figure 3.4.3).

Violence M is plotted in Figure 3.4.2 (note that the policy grid is rotated approximately 90° counter-clockwise relative to Figure 3.4.1). When $\tilde{g} < 19.76$, starting from $\bar{g} = 0$, violence is strictly increasing in $\bar{g}$ until collusion completely breaks down. Arresting during the reward state reduces the difference between the values of the reward and punishment states, so punishments are less harsh *relative* to the reward state. Enforcing collusive behavior requires that the punishment state be triggered more frequently, and so violence is higher. Eventually collusion is no longer possible, resulting in static Nash behavior in both states and so the firms continuously attack. When arrests in the reward state are sufficiently severe ($\bar{g} > 8.72$), the cartels play the mixed static Nash equilibrium, and increases in $\bar{g}$ lead to producing drugs and violence with lower probability, leading to the slight downward slope seen.

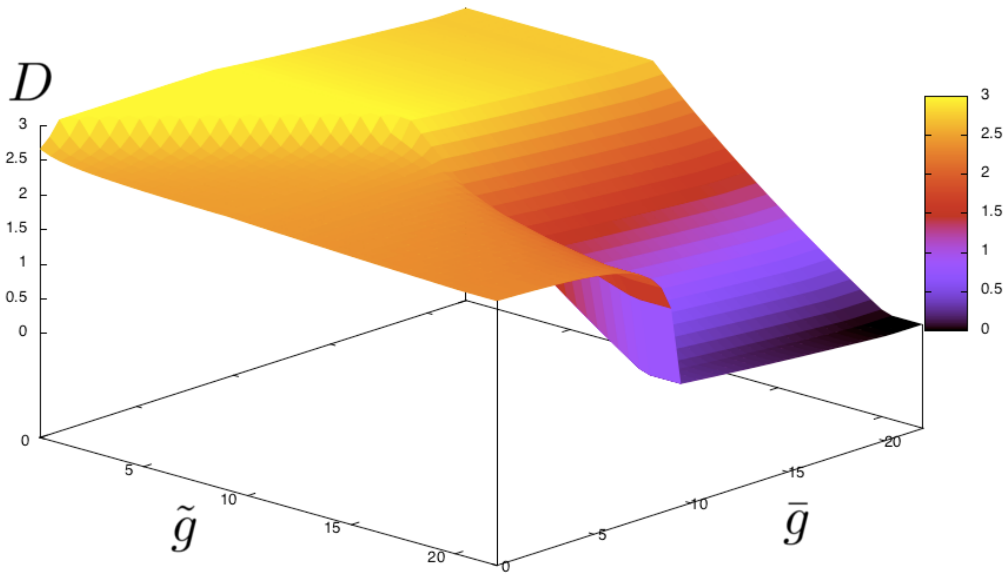

Figure 3.4.3: Discounted average drugs delivered D over different policies $(\bar{g}, \tilde{g})$.

By contrast, M is always weakly decreasing in $\tilde{g}$, and strictly so for the region $\tilde{g} \in [\bar{g}, 19.76), \bar{g} < 8.71$. In this region, increasing $\tilde{g}$ reduces the necessary frequency of punishments to sustain collusion, so violence happens less frequently as punishments become harsher. Eventually, as in the case of $\bar{g}$ discussed above, arrests become so great that attacks are not part of the static Nash equilibrium, resulting in zero violence altogether.

Figure 3.4.3 plots the average discounted quantity of drugs D delivered to consumers (the policy grid is oriented identically to Figure 3.4.2 and rotated approximately 90° counter-clockwise relative to Figure 3.4.1). D is always weakly (and usually strictly) decreasing in $\tilde{g}$. This is because arrests during the punishment state either incentivize collusion in the reward state, which reduces the quantity in the reward state, or reduce drugs delivered in the punishment state — both of which affect D negatively.

The effect of $\bar{g}$ is more complicated. For the region $\tilde{g} \in [\bar{g}, 8.71), \bar{g} < 8.71$, drugs are *increasing* in $\bar{g}$. In this region, arrests in the reward state inhibit collusion, increasing drugs delivered in the reward state, while the punishment static equilibrium quantity is unchanged. For the region $\tilde{g} > 8.72, \bar{g} < 8.71$, collusion is also inhibited by reward state arrests, but this is dominated by the negative effect on drugs delivered during the punishment state, which is also triggered more frequently in order to sustain collusion. For $\bar{g} > 8.72$, collusion is unsustainable and reward state arrests reduce the equilibrium quantities. In the rightmost corner ($\tilde{g} \geq 20.5, \bar{g} \geq 20.5$), arrests are sufficiently high in both states that the market effectively shuts down: firms repeatedly choose to hire zero traffickers and not attack.

3.5 Discussion and Conclusion

Depending on the government's objectives, the results above have differing policy implications. If the government wishes to maximize cartel profits, perhaps because of corruption, the optimal policy is clear: never arrest during the reward state ($\bar{g} = 0$) and arrest as much as possible in the punishment state ($\tilde{g} = G$). Interestingly, Figure 3.4.2 shows that this policy also minimizes violence, uniquely so if the government's power is sufficiently limited.

However, governments that outlaw drugs presumably also wish to reduce drug consumption. Figure 3.4.3 shows that for a very powerful government ($G \geq 20.5$), it is possible to reduce both drugs and violence to zero through a total crackdown. This works because the government arrests so many traffickers that shutting down is a dominant strategy.

This changes as the government becomes more constrained. A government with capacity $8.72 < G < 19.76$ faces a tradeoff between reducing violence and reducing drugs. Arrests during the punishment state serve both goals because arrests are sufficiently harsh to reduce quantities delivered. Arrests during the reward state reduce cooperative incentives, resulting in punishments being triggered more frequently, which reduces drug quantities but increases violence.

An even more constrained government ($G < 8.71$) faces no tradeoff: the corrupt policy minimizes drugs, minimizes violence, and maximizes firm profits (because prices are high). As above, arrests in the punishment state allow greater collusion, which reduces both drugs and violence. However, punishment quantities are greater than the reward quantities. By triggering more frequent punishments, arresting during the reward state actually increases equilibrium drug quantities.

The results show that switching from a corrupt policy to a crackdown policy of always arresting as much as possible can trigger a spike in violence so long as the government's power is limited. Since it may be quite difficult to assess a government's capacity to arrest traffickers (at least ex-ante), a policymaker facing such uncertainty over G who cracks down, aiming to reduce drug consumption but also wanting to prevent violence, may be unpleasantly surprised by the outcome. This is one way of thinking about Calderon's crackdown, which often involved replacing entire local police forces accused of corruption. By contrast, for a corrupt government or one that simply aims to minimize violence without regard to drug consumption (some combination of which may be thought of as similar to the PRI government before 2000), the decision problem is easy: never arrest during the reward state and always arrest during punishments, no matter what G is. Thus, law enforcement operations can actually prevent inter-cartel competition, so long as they are used to punish signals of defection.

There are a number of interesting avenues for future research. A natural extension is to more than two firms, resulting in a more competitive market. I currently restrict the government to policies conditioning only on the reward/punishment state, which is necessary given my current numerical method. If the government chooses arbitrary policies based on the history, this effectively transforms the model into a dynamic game. Studying policies like "never arrest until violence is observed, and always arrest thereafter" may be possible through Yeltekin, Cai, and Judd's (2015) numerical implementation of the Abreu, Pearce, and Stacchetti (1990) framework for dynamic games. Studying asymmetric policies and strategies would also be of interest to evaluate the effects of favoring one cartel over another. Also, a generalization to include multiple governments may shed greater light on the effects of political decentralization, highlighted by Rios (2014) as being a key driver of increased competition between cartels.

Appendix

Appendix 1

Proofs for Chapter 1

1.1 Proof of Theorem 1.2.2

Suppose by contradiction that for any $\delta^* \in (0,1)$, there always exists $\delta \in (\delta^*, 1)$ such that a sequential equilibrium exists with positive probability of hiring on the equilibrium path. Let such an equilibrium be given. I begin by showing that if the mechanic performs a tune-up, then she is known to be good by all future motorists (even those who do not observe a tune-up).

Lemma 1.1.1. *If the mechanic performs a tune-up at any history at period t' on the equilibrium path, then motorists at any future period $t'' > t'$ will know that the mechanic is good (regardless of the subsequent actions played by the mechanic) on the equilibrium path.*

Proof. For convenience I use a subscript zero on any period $\tilde{t}$ to denote the earliest period $\tilde{t}_0 \equiv \max\{0, \tilde{t} - T\}$ observed by motorist $\tilde{t}$ (as given in Definition 1.2.1). Let any history $h^{\tilde{t}}$ at period $\tilde{t}$ on the equilibrium path be given, and let $\hat{h}^{\tilde{t}}_{\tilde{t}_0,t} \equiv (h^{\tilde{t}}_{\tilde{t}_0}, ..., h^{\tilde{t}}_{t-1})$ denote the *partial observable subhistory* at $\tilde{t}$ up to $t \in \{\tilde{t}_0, ..., \tilde{t}\}$, which contains the events of the periods before (but not including) t that are observed by motorist $\tilde{t}$. Let $C(\hat{h}^{\tilde{t}})$ be an indicator function equal to 1 if $\hat{h}^{\tilde{t}}$ has a tune-up or, for $\tilde{t} > T$, has the mechanic hired every period, and equal to 0 otherwise:

$$C(\hat{h}^{\tilde{t}}) \equiv \begin{cases} 1 & \exists t \in \{\tilde{t}_0, ..., \tilde{t}-1\} \text{ such that } \hat{h}^{\tilde{t}}_t = c \\ 1 & \tilde{t} > T \text{ and } \forall t \in \{\tilde{t}_0, ..., \tilde{t}-1\}, \hat{h}^{\tilde{t}}_t \neq \varnothing \\ 0 & \text{otherwise.} \end{cases}$$

I prove by induction the following claims to be true:

Claim 1.1.1. The full history $h^{\tilde{t}}$ contains a tune-up if and only if observable subhistory $\hat{h}^{\tilde{t}}$ contains a tune-up or, for $\tilde{t} > T$, has the mechanic hired in every observed period $\tilde{t}_0, ..., \tilde{t}-1$ (i.e. $C(\hat{h}^{\tilde{t}}) = 1$).

Claim 1.1.2. If $h^{\tilde{t}}$ does not contain a tune-up, then the observable subhistory $\hat{h}^{\tilde{t}}$ has the mechanic being hired at most L times.

Claim 1.1.3. Suppose that for observable subhistory $\hat{h}^{\tilde{t}}$, $C(\hat{h}^{\tilde{t}}) = 0$. For any $t \in \{\tilde{t}_0, ..., \tilde{t}-1\}$, the probability that the mechanic is hired at period t conditional on the partial observable subhistory $\hat{h}^{\tilde{t}}_{\tilde{t}_0,t} \equiv (h^{\tilde{t}}_{\tilde{t}_0}, ..., h^{\tilde{t}}_{t-1})$ and $C(\hat{h}^{\tilde{t}}) = 0$ is independent of the mechanic's type $s \in \{b, g\}$; that is,

$$P(\eta_t \neq \varnothing | s = g, \hat{h}^{\tilde{t}}_{\tilde{t}_0,t}, C(\hat{h}^{\tilde{t}}) = 0) = P(\eta_t \neq \varnothing | s = b, \hat{h}^{\tilde{t}}_{\tilde{t}_0,t}, C(\hat{h}^{\tilde{t}}) = 0).$$

Remark 1.1.1. For any history $h^{\tilde{t}}$ on the equilibrium path, if Claim 1.1.1 is true, then Claim 1.1.3 is true. By Claim 1.1.1, if motorist $\tilde{t}$ has an observable subhistory such that $C(\hat{h}^{\tilde{t}}) = 0$, then he knows that the full history $h^{\tilde{t}}$ has no tune-ups and therefore the mechanic has played indistinguishably from the bad type thus far, so conditioning on the mechanic's true state cannot change the probability that motorist $t \in \{\tilde{t}_0, ..., \tilde{t}-1\}$ hires, so Claim 1.1.3 is true.

First, suppose that $\tilde{t} \leq T$. Since motorist $\tilde{t}$ observes the full history $h^{\tilde{t}}$, Claim 1.1.1 is clearly true. If the mechanic performed more than L engine replacements without any tune-ups in $h^{\tilde{t}}$, then the $(L+1)$th hiring motorist would have a posterior greater than or equal to $\Upsilon^L(\mu^0) > p^*$, so hiring could not be a best response. Thus, Claim 1.1.2 is true. Claim 1.1.3 is implied by Claim 1.1.1 as stated in Remark 1.1.1.

Now suppose that for some $\tilde{t} \geq T$, Claims 1.1.1, 1.1.2 and 1.1.3 are true for any history h^t at any $t \leq \tilde{t}$ on the equilibrium path. I show that this implies they must also hold at any following history $h^{\tilde{t}+1}$ on the equilibrium path. First, suppose that $h^{\tilde{t}}$ contains a tune-up. By Claim 1.1.1, observable subhistory $\hat{h}^{\tilde{t}}$ satisfies $C(\hat{h}^{\tilde{t}}) = 1$, and motorist $\tilde{t}$ knows the mechanic is good. If $\hat{h}^{\tilde{t}}$ has a tune-up at some period $t \in \{\tilde{t}_0+1, ..., \tilde{t}-1\}$, then the next motorist will see the same tune-up in his observable subhistory $\hat{h}^{\tilde{t}+1}$ and thus $C(\hat{h}^{\tilde{t}+1}) = 1$.

If the tune-up is at period $\tilde{t}_0$, then by Assumption 1 motorists in all periods $\tilde{t}_0+1, ..., \tilde{t}-1$ hired, and motorist $\tilde{t}$ also hires, giving an observable subhistory $\hat{h}^{\tilde{t}+1}$ next period with the mechanic hired every period, so $C(\hat{h}^{\tilde{t}+1}) = 1$.

The previous paragraph shows that if $h^{\tilde{t}+1}$ contains a tune-up, then $C(\hat{h}^{\tilde{t}+1}) = 1$, but proving Claim 1.1.1 still requires the converse to be true, whose contrapositive is proven in this paragraph. Suppose $h^{\tilde{t}+1}$ does not contain a tune-up. Since Claim 1.1.2 holds at $h^{\tilde{t}}$, then $\hat{h}^{\tilde{t}}$ has the mechanic hired at most L times. If $\hat{h}^{\tilde{t}}$ has the mechanic hired strictly less than L times, then Claim 1.1.1 clearly holds because even if motorist $\tilde{t}$ hires, $\hat{h}^{\tilde{t}+1}$ will have at most $L < T$ hirings. Suppose $\hat{h}^{\tilde{t}}$ has exactly L hirings. Because $P(\eta_t \neq \varnothing | s = g, \hat{h}^{\tilde{t}}_{\tilde{t}_0,t}, C(\hat{h}^{\tilde{t}}) = 0) = P(\eta_t \neq \varnothing | s = b, \hat{h}^{\tilde{t}}_{\tilde{t}_0,t}, C(\hat{h}^{\tilde{t}}) = 0)$ for all $t \in \{\tilde{t}_0, ..., \tilde{t}-1\}$ (due to Claim 1.1.3 holds at $h^{\tilde{t}}$), I can give a lower bound for the posterior belief at $\tilde{t}$. Suppose that an observable period t has the mechanic not being hired ($\hat{h}^{\tilde{t}}_t = \varnothing$); then the partial posterior at $t+1$ is unchanged from period t: $\mu^{\tilde{t}}_{t+1}(\hat{h}^{\tilde{t}}) = \mu^{\tilde{t}}_t(\hat{h}^{\tilde{t}})$. Suppose that it instead has the mechanic doing an engine replacement ($\hat{h}^{\tilde{t}}_t = e$); then $\Upsilon(\cdot)$ bounds the partial posterior from below: $\mu^{\tilde{t}}_{t+1}(\hat{h}^{\tilde{t}}) \geq \Upsilon(\mu^{\tilde{t}}_t(\hat{h}^{\tilde{t}}))$. Since $\hat{h}^{\tilde{t}}$ contains L engine replacements and $T-L$ no-hire events, motorist $\tilde{t}$ has a posterior $\mu^{\tilde{t}}(\hat{h}^{\tilde{t}}) \geq \Upsilon^L(\mu^0) > p^*$, so he does not hire, giving $L < T$ engine replacements in $\hat{h}^{\tilde{t}+1}$. Thus, if $h^{\tilde{t}+1}$ does not have a tune-up, then the observable subhistory $\hat{h}^{\tilde{t}+1}$ has the mechanic hired at most $L < T$ times and $C(\hat{h}^{\tilde{t}+1}) = 0$, so both Claims 1.1.1 and 1.1.2 are proven true. Finally, Remark 1.1.1 shows that Claim 1.1.1 implies Claim 1.1.3.

Having proven Claims 1.1.1, 1.1.2 and 1.1.3 for all histories on the equilibrium path, Claim 1.1.1 implies Lemma 1.1.1 because the set of observable subhistories possible at histories on the equilibrium path where a tune-up has ever occurred (which can only happen if the mechanic is good) is disjoint from the set of observable subhistories possible on the equilibrium path when a tune-up has not ever occurred. □

The arguments in the proof of Lemma 1.1.1 show that more than L engine-replacements cannot occur within the first T periods unless a tune-up is performed.

Corollary 1.1.1. *Let any sequential equilibrium be given, and let h^T be a history at T on the equilibrium path. If no tune-ups occurred in h^T, then h^T contains at most L engine replacements.*

Proof. Claim 1.1.2 is shown in the proof of Lemma 1.1.1 to be true at all histories on the equilibrium path. □

Without loss of generality, let period 0 be the first period at which the mechanic is hired with positive probability. Lemma 1.1.1 implies that the continuation payoff of doing a tune-up at any period is u (the mechanic is hired forever after) given Assumption 1. Let $l \leq L$ be the maximum number of engine replacements without any tune-ups in any history at period T on the equilibrium path, and let h^T be such a history with l engine replacements. Let t_j denote the jth period in h^T containing an engine replacement.

I show that the last engine replacement (within the first $T+1$ periods) at t_l must occur sufficiently late. If the mechanic does an engine replacement at period t_l, then she is certainly not hired for periods $t_l+1, ..., T$, so the continuation payoff of an engine replacement is bounded from above by $\delta^{T-t_l}u$. Since the mechanic is hired she must be willing to perform an engine replacement when needed with positive probability, so a necessary condition is

$$(1-\delta)u + \delta^{T+1-t_l}u \geq -(1-\delta)w + \delta u$$

$$\delta^{T+1-t_l} \geq -(1-\delta)(1+w/u) + \delta$$

$$(T+1-t_l)\ln\delta \geq \ln(\delta - (1-\delta)(1+w/u))$$

$$t_l \geq T+1 - \ln(\delta - (1-\delta)(1+w/u))/\ln\delta.$$

By l'Hôpital's rule,

$$\begin{aligned} \lim_{\delta\to 1} \frac{\ln(\delta-(1-\delta)(1+w/u))}{\ln\delta} &= \lim_{\delta\to 1} \frac{\frac{1+(1+w/u)}{\delta-(1-\delta)(1+w/u)}}{\frac{1}{\delta}} \\ &= 2 + w/u. \end{aligned} \tag{1.1.1}$$

Let any $\varepsilon > 0$ be given. Then

$$t_l \geq T - 1 - w/u - \varepsilon$$

for δ close enough to one. Define $\underline{t}_l \equiv T - 1 - w/u - \varepsilon$ as a lower bound for t_l.

I define $\underline{t}_j$ for all $j \in \{1, ..., l\}$ by (backward) induction. Consider period t_j for some $j < l$, where $\underline{t}_{j'}$ is defined for all $j' \in \{j+1, ..., l\}$. The continuation payoff of doing an engine replacement is less than or equal to $\delta^{\underline{t}_{j+1} - t_j - 1} u$ (because $\underline{t}_{j+1}$ is a lower bound for the $(j+1)$th period with an engine replacement). Incentive compatibility gives the necessary condition

$$(1-\delta)u + \delta^{\underline{t}_{j+1} - t_j} u \geq -(1-\delta)w + \delta u$$

$$\delta^{\underline{t}_{j+1} - t_j} \geq \delta - (1-\delta)(1 + w/u)$$

$$(\underline{t}_{j+1} - t_j) \ln \delta \geq \ln\left(\delta - (1-\delta)(1+w/u)\right)$$

$$t_j \geq \underline{t}_{j+1} - \frac{\ln\left(\delta - (1-\delta)(1+w/u)\right)}{\ln \delta}. \tag{1.1.2}$$

The limit as $\delta \to 1$ of the second term on the right hand side of (1.1.2) is given by (1.1.1). For δ close enough to one, substituting into (1.1.2) gives

$$t_j \geq \underline{t}_{j+1} - 2 - w/u - \varepsilon \equiv \underline{t}_j.$$

Therefore

$$\begin{aligned} \underline{t}_1 &= -2 - \frac{w}{u} - \varepsilon + \underline{t}_2 = -2\left(2 + \frac{w}{u} + \varepsilon\right) + \underline{t}_3 = \cdots = T + 1 - \left(2 + \frac{w}{u} + \varepsilon\right) l \\ &\geq T + 1 - \left(2 + \frac{w}{u} + \varepsilon\right) L. \end{aligned}$$

Since the choice of $\varepsilon > 0$ is arbitrary, suppose we pick some $\varepsilon < (T+1)/L - (2 + w/u)$. Then $\underline{t}_1 > 0$ for δ close enough to one due to the lower bound (1.2.4) on T. Since the first period in which the mechanic is hired with positive probability is period 0, this is a contradiction.

1.2 Proofs of Fading History Results

1.2.1 Proof of Theorem 1.2.4

Let any sequential equilibrium and history h^t be given at which player 1 faces a decision node. Let $V_{a_1,a_2}(h^t)$ denote player 1's continuation payoff from the action profile (a_1, a_2). Player 1 plays a_d with certainty if

$$E_{a_2}[(1-\delta)u_1(a_d, a_2) + \delta V_{a_d,a_2}(h^t)] > E_{a_2}[(1-\delta)u_1(a_1', a_2) + \delta V_{a_1',a_2}(h^t)] \tag{1.2.1}$$

for all $a_1' \in A_1 \backslash \{a_d\}$, where $E_{a_2}[\cdot]$ is the expectation over player 2's actions $a_2 \in \tilde{A}_2$ given player 1's beliefs (in the mechanic game, of course, player 1 knows a_2 (hiring) because $\tilde{A}_2$ is a singleton). (1.2.1) can be rearranged as

$$E_{a_2}[V_{a_d,a_2}(h^t) - V_{a_1',a_2}(h^t)] < \frac{1-\delta}{\delta} E_{a_2}[u_1(a_d, a_2) - u_1(a_1', a_2)]. \tag{1.2.2}$$

The left hand side of (1.2.2) is equal to the discounted sum of the expected differences in stage payoffs at every future period. Denoting player 1's stage payoff at some period $\hat{t} > t$ as $v(\hat{t})$, let $\bar{v}_{a_1,a_2}(h^t, \hat{t}) \equiv E[v(\hat{t})|h^t, a_1, a_2]$ be the expected stage payoff at $\hat{t}$ conditional on h^t and action profile (a_1, a_2) at h^t. The maximum change in the expected stage payoff at $\hat{t}$ due to choosing an action different from a_d at period t is

$$\Delta \bar{v}(h^t, \hat{t}) = \max_{(a_1', a_2) \in A_1 \times \tilde{A}_2} \{\bar{v}_{a_1',a_2}(h^t, \hat{t}) - \bar{v}_{a_d,a_2}(h^t, \hat{t})\} \leq z.$$

The action at t can only affect player 1's payoff at $\hat{t}$ if period $\hat{t}$ observes t directly, or observes some period $t' \in \{t+1, ..., \hat{t}-1\}$ that observes t, etc.; otherwise, player 2's action at period $\hat{t}$ is necessarily independent of the events of period t. This notion of an "observation chain" is formalized as "t reaches $\hat{t}$" in the following definition.

Definition 1.2.1. Let two periods t' and $t'' > t'$ be given. Inductively define the relation "t' *k-reaches* t''" as follows. If period t'' observes period t', then t' is said to 0-reach t''. If period t'' observes some period $\tilde{t} \in \{t'+1, ..., t''-1\}$ and $\tilde{t}$ k-reaches t', then t' is said to

$(k+1)$-reach t''. More simply, if (and only if) period t' k-reaches t'' for some $k \in \{0, 1, ...\}$, then t' is said to *reach* t''.

Let $\phi(t, \hat{t})$ denote the probability that t reaches $\hat{t}$, which gives the upper bound $\Delta\bar{v}(h^t, \hat{t}) \leq \phi(t, \hat{t})z$. The following lemma gives an upper bound for $\phi(t, \hat{t})$.

Lemma 1.2.1. *For any two periods t and $\hat{t} > t$, $\phi(t, \hat{t}) \leq 2^{\hat{t}-t-1}\lambda^{\hat{t}-t}$.*

Proof. The proof is by induction. For $\hat{t} = t+1$, $\phi(t, \hat{t}) = \lambda$ is trivially true. Now suppose that for some $\hat{t} > t$, $\phi(t, t') \leq 2^{t'-t-1}\lambda^{t'-t}$ for all $t' \in \{t+1, ..., \hat{t}\}$. The probability that t reaches $\hat{t}+1$ is the probability that motorist $\hat{t}+1$ observes either t or some period $t' \in \{t+1, ..., \hat{t}-1\}$ such that t reaches t'. Then Boole's inequality gives

$$\begin{aligned}
\phi(t, \hat{t}+1) &\leq \lambda^{\hat{t}+1-t} + \sum_{t'=t+1}^{\hat{t}} \lambda^{\hat{t}+1-t'}\phi(t, t') \\
&\leq \lambda^{\hat{t}+1-t} + \sum_{t'=t+1}^{\hat{t}} \lambda^{\hat{t}+1-t'}(2^{t'-t-1}\lambda^{t'-t}) \\
&= \lambda^{\hat{t}+1-t}\left(1 + \sum_{k=0}^{\hat{t}-t-1} 2^k\right) \\
&= \lambda^{\hat{t}+1-t}\left(1 + \frac{1-2^{\hat{t}-t}}{1-2}\right) \\
&= 2^{\hat{t}-t}\lambda^{\hat{t}+1-t}.
\end{aligned}$$

□

I can now write an upper bound for the left hand side of (1.2.2):

$$\begin{aligned}
E_{a_2}[V_{a_d,a_2}(h^t) - V_{a_1',a_2}(h^t)] &\leq (1-\delta)\sum_{k=1}^{\infty}\delta^{k-1}\Delta\bar{v}(h^t, t+k) \\
&\leq (1-\delta)\sum_{k=1}^{\infty}\delta^{k-1}\phi(t, t+k)z \\
&\leq (1-\delta)z\sum_{k=1}^{\infty}\delta^{k-1}(2^{k-1}\lambda^k).
\end{aligned}$$

Since $\delta\lambda < \frac{1}{2}$,

$$\begin{aligned} E_{a_2}[V_{a_d,a_2}(h^t) - V_{a'_1,a_2}(h^t)] &\leq (1-\delta)\lambda z \sum_{k=0}^{\infty}(2\delta\lambda)^k \\ &= \frac{(1-\delta)\lambda z}{1-2\delta\lambda}. \end{aligned} \tag{1.2.3}$$

The right hand side of (1.2.3) is a strictly increasing function of λ for $\lambda \in (0, 1/(2\delta))$. Substituting (1.2.5) into λ gives

$$\begin{aligned} E_{a_2}[V_{a_d,a_2}(h^t) - V_{a'_1,a_2}(h^t)] &< \frac{(1-\delta)z\left(\frac{z_d}{\delta(z+2z_d)}\right)}{1-2\delta\left(\frac{z_d}{\delta(z+2z_d)}\right)} \\ &= \frac{(1-\delta)}{\delta}z_d \\ &\leq \frac{1-\delta}{\delta}E_{a_2}[u_1(a_d,a_2) - u_1(a'_1,a_2)] \end{aligned}$$

for any $a'_1 \in A_1, a_2 \in \tilde{A}_2$, so (1.2.1) is true.

1.2.2 A Higher Upper Bound for λ for Myopic Equilibria in the Mechanic Game

Theorem 1.2.4 assumes that the "worst case" when an "observation chain" reaches a future period is the stage payoff decreasing by the maximum feasible amount z; in the mechanic game, this difference is $u+w$. A tighter bound that seems natural is the difference between the highest feasible payoff and the minmax payoff (u). The following corollary uses that bound on the stage payoff difference to give a higher upper bound on λ, using Assumption 1 and Criterion 1. For δ close to one, as w/u approaches 1 the bound (1.2.4) approaches $\frac{2}{5}$ (corresponding to motorists talking to an average of $\frac{2}{3}$ future motorists) and as w/u approaches ∞, (1.2.4) approaches $\frac{1}{2}$ (corresponding to an average of 1 future motorist).

Corollary 1.2.1. *Consider the fading history mechanic game with*

$$\lambda < \frac{1}{\delta\left(2+\frac{u}{u+w}\right)}. \tag{1.2.4}$$

Then the action outcome of any sequential equilibrium satisfying Assumption 1 and Criterion 1 has the good mechanic doing the correct repair when hired.

Proof. Let σ_g^* denote the equilibrium strategy of the good mechanic, and let $\bar{\sigma}_g$ be the strategy identical to σ_g^* except that at any history containing a tune-up, the mechanic does the right repair with certainty (it may be that $\sigma_g^* = \bar{\sigma}_g$). The following result allows a simplification of the continuation payoffs for a tune-up. Note that μ_t (with a subscript instead of superscript t) denotes motorist t's beliefs about the mechanic's type and the history (as opposed to μ^t, which is simply the belief on the type).

Lemma 1.2.2. *Let a sequential equilibrium $(\sigma_g^*, (\sigma_t^*)_t, (\mu_t)_t)$ under fading history given λ satisfying Assumption 1 and Criterion 1 be given. At any history h^t on the equilibrium path containing a tune-up, it is a best response for the mechanic to perform the correct repair.*

Proof. Any motorist observing the entire history (which occurs with positive probability at every history) must hire due to Assumption 1. This is only possible if the mechanic performs the correct repair with at least positive probability β^* no matter the car's state, so it must be a best response. □

Thus, deviating to $\bar{\sigma}_g$ must be a best response at any history containing a tune-up. Calculation of the expected stage payoffs following a tune-up (simply the probability of being hired times u) is simpler for $\bar{\sigma}_g$ and allows them to be used as upper bounds on the expected stage payoffs following an engine replacement because of Criterion 1.

For the remainder of the proof, suppose that the mechanic deviates to $\bar{\sigma}_g$, which has the same continuation payoffs as σ_g^* at every history, an implication of Lemma 1.2.2. Let the notation and arguments in the proof of Theorem 1.2.4 (Appendix 1.2.1) up to and including Lemma 1.2.1 be given, except that all notation is with respect to the strategy $\bar{\sigma}_g$ (not σ_g^*) and a_2 is omitted from subscripts (because in the mechanic game, at a mechanic's decision node, a_2 is known to be "hire"). At any history h^t, let $\rho_a^{\hat{t}}(h^t)$ be the probability that the mechanic is hired at period $\hat{t} > t$ conditional on repair a at h^t.

Criterion 1 implies that the continuation payoff for a tune-up is greater than or equal to that of an engine replacement because $\bar{v}_c(h^t, \hat{t}) = \rho_c^{\hat{t}}(h^t)u \geq \rho_e^{\hat{t}}(h^t)u \geq \bar{v}_e(h^t, \hat{t})$. Therefore, when the motorist at h^t needs a tune-up, performing a tune-up strictly dominates an engine replacement.

What remains to be shown is that performing a needed engine replacement strictly dominates doing an incorrect tune-up. Let $\breve{\sigma}_g$ be the strategy identical to $\bar{\sigma}_g$, except that any history following (h^t, e) (i.e. any history that begins with h^t followed by e at period t) the mechanic always does the right repair. Let $\breve{V}_a, \breve{\rho}_a^{\hat{t}}, \breve{v}_a$ be the analogues of $V_a, \rho_a^{\hat{t}}, \bar{v}_a$ (which are defined for $\bar{\sigma}_g$) for a deviation to $\breve{\sigma}_g$. Note that $\breve{v}_e(h^t, \hat{t}) = \breve{\rho}_e^{\hat{t}}(h^t)u$, $\breve{\rho}_c^{\hat{t}}(h^t) = \rho_c^{\hat{t}}(h^t)$, $\breve{V}_c(h^t) = V_c(h^t)$, and $\breve{V}_e(h^t) \leq V_e(h^t)$. The fact that $\breve{v}_c(h^t, \hat{t}) - \breve{v}_e(h^t, \hat{t}) = (\breve{\rho}_c^{\hat{t}}(h^t) - \breve{\rho}_e^{\hat{t}}(h^t))u \leq \phi(t, \hat{t})u$ yields

$$
\begin{aligned}
V_c(h^t) - V_e(h^t) &\leq \breve{V}_c(h^t) - \breve{V}_e(h^t) \\
&= (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1}(\breve{v}_c(h^t, \hat{t}) - \breve{v}_e(h^t, \hat{t})) \\
&\leq (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1}\phi(t, t+k)u.
\end{aligned}
$$

By Lemma 1.2.1,

$$
\begin{aligned}
V_c(h^t) - V_e(h^t) &\leq (1-\delta)u\sum_{k=1}^{\infty} \delta^{k-1}(2^{k-1}\lambda^k) \\
&= \frac{(1-\delta)\lambda u}{1-2\delta\lambda}.
\end{aligned}
$$

Then substituting (1.2.4) into λ gives

$$
\begin{aligned}
V_c(h^t) - V_e(h^t) &< \frac{(1-\delta)u\left(\frac{1}{\delta(2+u/(u+w))}\right)}{1-2\delta\left(\frac{1}{\delta(2+u/(u+w))}\right)} \\
&= \frac{(1-\delta)u}{\delta\left(2+u/(u+w)\right) - 2\delta} \\
&= \frac{1-\delta}{\delta}(u+w).
\end{aligned}
$$

Therefore, doing an incorrect tune-up is not a best response. □

1.2.3 Proof of Theorem 1.2.5

Suppose by contradiction that for any $\lambda^* \in (0,1)$, there always exists $\lambda \in (\lambda^*, 1)$ such that a sequential equilibrium exists with positive probability of hiring on the equilibrium path. Let such an equilibrium be given. Without loss of generality, let period zero be the first period at which the mechanic is hired with positive probability. The following lemma establishes that if the mechanic is hired at some history in equilibrium, she must be hired again sufficiently soon (or else the temptation to do a tune-up will be too great).

Lemma 1.2.3. *Suppose the mechanic is hired at some history h^t on the equilibrium path at period t with positive probability, such that*

- $t = 0$*, or*
- *the mechanic is hired with probability greater than $\lambda^{t(t+1)/2}$.*

Suppose the mechanic chooses e at h^t, and let $t' > t$ be the earliest future period at which the mechanic is again hired with probability greater than $\lambda^{t'(t'+1)/2}$. Define

$$K(\delta, u, w) \equiv \frac{\ln(\delta - (1-\delta)(1 + w/u))}{\ln \delta}.$$

Then $t' \leq t + K(\delta, u, w)$ for λ close enough to one.

Proof. For any period t, if the mechanic is hired at period 0 or at h^t with probability greater than $\lambda^{t(t+1)/2}$, then the mechanic must perform a needed engine replacement with positive probability; otherwise, the motorist who sees the full history h^t would not hire, since the probability that the full history is observed is $\prod_{k=1}^{t} \lambda^k = \lambda^{t(t+1)/2}$. Incentive compatibility gives

$$(1-\delta)u + \delta V_e(h^t) \geq -(1-\delta)w + \delta V_c(h^t) \tag{1.2.5}$$

where $V_a(h^t)$ is the continuation payoff of action a. By definition, t' is the earliest period such that the mechanic is hired with probability greater than $\lambda^{t'(t'+1)/2}$ if she chooses e at period t, so an upper bound on her continuation payoff for e is

$$\begin{aligned} V_e(h^t) &\leq (1-\delta)\left[\sum_{k=t+1}^{t'-1} \delta^{k-t-1}(1-\lambda^{k(k+1)/2})u + \sum_{k=t'}^{\infty} \delta^{k-t-1}u\right] \\ &= (1-\delta)\sum_{k=t+1}^{t'-1} \delta^{k-t-1}(1-\lambda^{k(k+1)/2})u + \delta^{t'-t-1}u. \end{aligned}$$

Assumption 1 gives the following lower bound for the continuation payoff of c:

$$\begin{aligned} V_c(h^t) &\geq (1-\delta)\sum_{k=0}^{\infty} \delta^k \lambda^{k+1} u \\ &= \frac{(1-\delta)\lambda u}{1-\delta\lambda}. \end{aligned}$$

Substituting these bounds into (1.2.5) gives

$$(1-\delta)u + (1-\delta)\sum_{k=t+1}^{t'-1} \delta^{k-t}(1-\lambda^{k(k+1)/2})u + \delta^{t'-t}u \geq -(1-\delta)w + \delta\frac{(1-\delta)\lambda u}{1-\delta\lambda}$$

$$\delta^{t'-t} \geq (1-\delta)\left[\frac{\delta\lambda}{1-\delta\lambda} - (1+w/u) - \sum_{k=t+1}^{t'-1} \delta^{k-t}(1-\lambda^{k(k+1)/2})\right]. \tag{1.2.6}$$

Let any $\varepsilon > 0$ be given. Taking the limit of the right hand side of (1.2.6) as $\lambda \to 1$, there exists λ^* such that for all $\lambda \in (\lambda^*, 1)$,

$$\delta^{t'-t} \geq \frac{\delta - (1-\delta)(1+w/u)}{\exp(\varepsilon/(-\ln\delta))}$$

since $\exp(\varepsilon/(-\ln\delta)) > 1$. Solving for t' gives

$$(t'-t)\ln\delta \geq \ln(\delta - (1-\delta)(1+w/u)) - \frac{\varepsilon}{-\ln\delta}$$

$$t' \leq t + \frac{\ln(\delta - (1-\delta)(1+w/u))}{\ln\delta} + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we can pick $\varepsilon < \max\{1, \lceil K(\delta,u,w)\rceil - K(\delta,u,w)\}$. In that case, because t' is an integer, it must be that $t' \leq t + K(\delta, u, w)$ for λ close enough to one. $\square$

Lemma 1.2.3 implies that for λ close enough to one, if the mechanic is hired at period 0, with positive probability she must be hired in at least $L+1$ periods (with greater than probability $\lambda^{t'(t'+1)/2}$ at each such period t') in the first $KL+1$ periods on the equilibrium path, which means there must exist history $h^{\tilde{t}}$ at $\tilde{t} \leq KL+1$ on the equilibrium path that includes $L+1$ engine replacements and no tune-ups. This also implies that at each of these hirings, the mechanic must have performed a tune-up with at least probability β^* (see (1.2.2)). Yet this implies that the posterior of the motorist receiving the $(L+1)$th engine replacement at period t_{L+1} if he observes the full history must have been at least $\Upsilon^L(\mu^0) > p^*$. Thus, hiring was not a best response for that motorist with at least probability $\lambda^{t_{L+1}(t_{L+1}+1)/2}$, a contradiction.

1.2.4 Proof of Theorem 1.2.6

I begin by characterizing λ^*. For $t \in \{0, 1, ...\}$, $n \in \{2, 3, ...\}$, and $\lambda \in (0, 1)$, define

$$f(t, n; \lambda) \equiv \sum_{k=t+1}^{t+n-1} [\lambda^{k-t} - (1 - \lambda^{k(k+1)/2})].$$

Note the following useful properties about the function f.

Fact 1.2.1. *$f(t, n; \lambda)$ is strictly decreasing in t, and strictly increasing in λ.*

Fact 1.2.2. *Let any $\varepsilon > 0$ be given. For any $t \in \{0, 1, ...\}, n \in \{2, 3, ...\}$, there exists $\lambda' \in (0, 1)$ such that for any $\lambda \in (\lambda', 1)$, $f(t, n; \lambda) > n - 1 - \varepsilon$.*

Let n^* be an integer strictly greater than $1 + w/u$. Pick $\lambda^* \in (0, 1)$ such that

$$f(Ln^*, n^* + 1; \lambda^*) \geq 1 + w/u. \tag{1.2.7}$$

Let $\lambda \in (\lambda^*, 1)$ be given. Suppose by contradiction that for any $\delta^* \in (0, 1)$, there always exists $\delta \in (\delta^*, 1)$ such that there exists a sequential equilibrium $(\sigma_g^*, (\sigma_t^*)_t, (\mu_t)_t)$

(note that μ_t (with a subscript instead of superscript t) denotes motorist t's beliefs about the mechanic's type and the history, instead of μ^t, which is simply the belief on the type), where the mechanic is hired with positive probability on the equilibrium path. Without loss of generality, let the first such period be 0.

Let $\bar{\sigma}_g$ be the strategy identical to σ_g^* except that at any history containing a tune-up, the mechanic does the right repair with certainty (it may be that $\sigma_g^* = \bar{\sigma}_g$). Lemma 1.2.2, reproduced here as Lemma 1.2.4 for convenience, shows that deviating to $\bar{\sigma}_g$ is a best response at any history (the only histories at which $\bar{\sigma}_g$ may differ from σ_g^* are those with tune-ups, and for those histories doing the right repair is always a best response).

Lemma 1.2.4. *Let a sequential equilibrium $(\sigma_g^*, (\sigma_t^*)_t, (\mu_t)_t)$ satisfying Assumption 1 and Criterion 1 be given. At any history h^t containing a tune-up, it is a best response for the mechanic to perform the correct repair.*

I use a technique here similar to the proof of Corollary 1.2.1 to simplify calculation of continuation payoffs. For $\bar{\sigma}_g$, calculation of the expected stage payoffs following a tune-up is simple (due to Lemma 1.2.4, it is the probability of being hired times u) and due to Criterion 1, they can be used as upper bounds on the expected stage payoffs following an engine replacement (shown below).

For the remainder of the proof, suppose that the mechanic deviates to $\bar{\sigma}_g$; since by Lemma 1.2.4 such a deviation is a best response at any history, the continuation payoffs are identical at all histories. At any history h^t, let $\rho_a^k(h^t)$ be the probability that the mechanic is hired at period $k > t$ conditional on doing repair $a \in \{c, e\}$ at h^t, and let $\bar{v}_a^k(h^t)$ denote the expected stage payoff at period k conditional on a. Criterion 1 requires that $\rho_c^k(h^t) \geq \rho_e^k(h^t)$. Since the mechanic performs all correct repairs following a tune-up, $\bar{v}_c^k(h^t) = \rho_c^k(h^t)u \geq \rho_e^k(h^t)u \geq \bar{v}_e^k(h^t)$.

Lemma 1.2.5. *Let the assumptions of Theorem 1.2.6 and $\lambda \in (\lambda^*, 1)$ be given. For δ close enough to one, if there exists a sequential equilibrium where the mechanic is hired with positive probability at period 0, then there exists a history $h^{\tilde{t}}$ on the equilibrium path at some* *period $\tilde{t} \leq Ln^*$ where the mechanic is hired with probability greater than $1 - \lambda^{\tilde{t}(\tilde{t}+1)/2}$ and $h^{\tilde{t}}$ contains L engine replacements and no tune-ups such that the posterior after observing the full history is $\mu^{\tilde{t}}(h^{\tilde{t}}) \geq \Upsilon^L(\mu^0)$.*

Proof. The proof is by induction. Let $t_1 > 0$ be the first period after 0 at which the mechanic is hired with greater than probability $1 - \lambda^{t_1(t_1+1)/2}$, conditional on the mechanic doing an engine replacement in period 0.

I now show that $t_1 \leq n^*$ for δ close enough to one. The continuation payoff of a tune-up at 0 has a lower bound due to Assumption 1 given by

$$\frac{\delta}{(1-\delta)} V_c(h^0) = \sum_{k=1}^{\infty} \delta^k \bar{v}_c^k(h^0) \geq \sum_{k=1}^{t_1-1} \delta^k \lambda^k u + \sum_{k=t_1}^{\infty} \delta^k \bar{v}_e^k(h^0)$$

where h^0 is the empty history at period 0. Since the mechanic is hired, the incentive constraint

$$-(1-\delta)w + \delta V_c(h^0) \leq (1-\delta)u + \delta V_e(h^0)$$

must hold (when an engine replacement is needed). A necessary condition for this incentive constraint is

$$-w + \sum_{k=1}^{n-1} \delta^k \lambda^k u + \sum_{k=n}^{\infty} \delta^k \bar{v}_e^k(h^0) \leq u + \sum_{k=1}^{n-1} \delta^k (1 - \lambda^{k(k+1)/2}) u + \sum_{k=n}^{\infty} \delta^k \bar{v}_e^k(h^0) \qquad (1.2.8)$$

for any $n \leq t_1$. Suppose by contradiction that $t_1 > n^*$. After some rearrangement of (1.2.8), picking $n = n^* + 1$ gives

$$\sum_{k=1}^{n^*} \delta^k [\lambda^k - (1 - \lambda^{k(k+1)/2})] u \leq u + w. \qquad (1.2.9)$$

Dividing by u and taking the limit of δ gives

$$\lim_{\delta \to 1} \sum_{k=1}^{n^*} \delta^k [\lambda^k - (1 - \lambda^{k(k+1)/2})] = f(0, n^* + 1) < 1 + w/u,$$

so (1.2.9) contradicts (1.2.7) for $\lambda > \lambda^*$ and δ close enough to one. Thus, the mechanic must be hired at period t_1 with probability greater than $1-\lambda^{t_1(t_1+1)/2}$ and $t_1 \leq n^*$ at a history h^{t_1} with one engine replacement and no tune-ups.

Now for some $j \geq 1$, let h^{t_j} be a history at $t_j \leq jn^*$ on the equilibrium path where the mechanic is hired with probability greater than $1-\lambda^{t_j(t_j+1)/2}$, such that h^{t_j} has j engine replacements and no tune-ups. I show that there exists period $t_{j+1} \leq t_j + n^*$ such that the mechanic is hired with probability greater than $1-\lambda^{t_{j+1}(t_{j+1}+1)/2}$. Since the mechanic is hired at h^{t_j} with greater than probability $1-\lambda^{t_j(t_j+1)/2}$, the mechanic must perform an engine replacement with positive probability when it is needed (by the same argument as above), and incentive compatibility gives the necessary condition

$$\begin{aligned} & -w + \sum_{k=t_j+1}^{t_j+n-1} \delta^{k-t_j}\lambda^{k-t_j}u + \sum_{k=t_j+n}^{\infty} \delta^{k-t_j}\bar{v}_e^k(h^{t_j}) \\ \leq \quad & u + \sum_{k=t_j+1}^{t_j+n-1} \delta^{k-t_j}(1-\lambda^{k(k+1)/2})u + \sum_{k=t_j+n}^{\infty} \delta^{k-t_j}\bar{v}_e^k(h^{t_j}) \end{aligned}$$

for $n \leq t_{j+1} - t_j$. Suppose by contradiction that $t_{j+1} > t_j + n^*$. Picking $n = n^* + 1$ gives

$$\sum_{k=t_j+1}^{t_j+n^*} \delta^{k-t_j}[\lambda^{k-t_j} - (1-\lambda^{k(k+1)/2})]u \leq u + w. \tag{1.2.10}$$

Dividing by u and taking the limit of δ gives

$$\lim_{\delta \to 1} \sum_{k=t_j+1}^{t_j+n^*} \delta^{k-t_j}[\lambda^{k-t_j} - (1-\lambda^{k(k+1)/2})] = f(t_j, n^*+1) < 1 + w/u,$$

so (1.2.10) contradicts (1.2.7) for $\lambda > \lambda^*$ and δ close enough to one. Then there exists a history $h^{t_{j+1}}$ following h^{t_j} on the equilibrium path for some $t_{j+1} \leq t_j + n^* \leq n^*(j+1)$ where the mechanic is hired with probability greater than $1-\lambda^{t_{j+1}(t_{j+1}+1)/2}$ with $j+1$ engine replacements (where the good mechanic must have performed a tune-up with at least probability β^*) and no tune-ups. For δ close enough to one, this induction proves the existence of such a history on the equilibrium path containing any number of engine replacements $j \leq L$ such the posterior upon observing the full history is at least $\Upsilon^j(\mu^0)$, if motorist 0 hires with positive probability. $\square$

Lemma 1.2.5 shows the existence of some history $h^{\tilde{t}}$ at some $\tilde{t} \leq Ln^*$ on the equilibrium path whose full observation yields posterior $\mu^{\tilde{t}}(h^{\tilde{t}}) \geq \Upsilon^L(\mu^0) > p^*$ and the mechanic is hired with probability greater than $1 - \lambda^{\tilde{t}(\tilde{t}+1)/2}$, which requires that the motorist hire even if he observes the full history. Yet if he observes the full history, hiring cannot be a best response, a contradiction.

1.3 Proofs of Chain Store Game Results

1.3.1 An Equilibrium for the Limited History Chain Store Game with $T = 1$

I construct a sequential equilibrium (σ^*, μ^*). Let any history h^t be given. The strategies are as follows. The incumbent always fights when $h^t_{t-1} \in \{\varnothing, h^0\}$ (where h^0 is the empty history at period 0), always acquiesces when $h^t_{t-1} = A$, and acquiesces with probability $\beta_F \equiv (1-b)/(1-\mu^0)$ when $h^t_{t-1} = F$. For any t, competitor t never enters when $\hat{h}^t_{t-1} \in \{\varnothing, h^0\}$, always enters when $\hat{h}^t_{t-1} = A$, and enters with probability $\alpha_F \equiv 1 - (1-\delta)/(\delta c)$ when $\hat{h}^t_{t-1} = F$. Competitor t has belief μ^0 that the incumbent is tough when $\hat{h}^t \in \{h^0, \varnothing, F\}$, and of course knows that the incumbent is weak when $\hat{h}^t = A$.

The competitor's strategy is clearly a best response at $\hat{h}^t \in \{h^0, \varnothing, A\}$. When $\hat{h}^t = F$, the competitor's payoff for entering is

$$\begin{aligned}
& \mu^0(b-1) + (1-\mu^0)(\beta_F b + (1-\beta_F)(b-1)) \\
= \; & \mu^0(b-1) + (1-\mu^0)\left(\frac{1-b}{1-\mu^0}b + \frac{1-\mu^0-(1-b)}{1-\mu^0}(b-1)\right) \\
= \; & \mu^0(b-1) + (1-b)(b-(b-\mu^0)) \\
= \; & 0,
\end{aligned}$$

so both In and $\varnothing$ are best responses.

Since strategies and beliefs only depend on the last period, continuation payoffs also depend only on the last period. For the incumbent, let V_{In} be the continuation payoff at some period after the competitor enters, and let V_a be the continuation payoff following action A by the incumbent. Since V_{In} must simply be the value of the best action,

$$\begin{aligned} V_{In} &= \max\{\delta V_A, -(1-\delta) + \delta V_F\} \\ &= \max\{\delta V_{In}, -(1-\delta) + \delta(\alpha_F V_{In} + (1-\alpha_F)c)\} \\ &= \max\left\{\delta V_{In}, -(1-\delta) + \delta\left(\left(1 - \frac{1-\delta}{\delta c}\right)V_{In} + \frac{1-\delta}{\delta c}c\right)\right\} \\ &= \max\left\{\delta V_{In}, \delta\left(1 - \frac{1-\delta}{\delta c}\right)V_{In}\right\} \qquad (1.3.1) \\ &= \delta V_{In}. \end{aligned}$$

Since $\delta \in (0,1)$, it must be that $V_{In} = 0$. This means that the payoffs for playing A and F are equal to 0 (see (1.3.1)), so any strategy by the incumbent is a best response.

The beliefs at $\hat{h}^t \in \{A, F\}$ are off the equilibrium path, so it remains to checked that they are consistent with small perturbations. Let some $\varepsilon \in (0,1)$ be given. Let $(\sigma_k)_k$ be a sequence of strategy profiles,[1] where under σ_k, competitor t plays In with probability ε^k when $\hat{h}^t \in \{\varnothing, h^0\}$ and plays $\varnothing$ with probability ε^k when $\hat{h}^t = A$, and the incumbent plays A with probability ε^k when $\hat{h}^t = \varnothing$ and plays F with probability ε^k when $\hat{h}^k = A$. Otherwise σ_k is the same as σ^*, and beliefs μ_k are entirely determined by Bayes' rule. Then $\lim_{k\to\infty}(\sigma_k, \mu_k) = (\sigma^*, \mu^*)$.

1.3.2 Proof of Theorem 1.3.1

Let any sequential equilibrium be given. I bound the payoff for the following (possibly deviation) incumbent strategy: for some K, fight for every period $0, ..., K$, then acquiesce for every period thereafter. The probability that every competitor $1, ..., K$ sees their full history is $\prod_{k=1}^{K}\prod_{k'=1}^{k}\lambda^{k'} = \lambda^{\tilde{\zeta}(K)}$ (note that competitor 0's history is always empty, so

[1] This is not to be confused with the sequence of competitor strategies for all periods in a single strategy profile.

there is nothing to see) where

$$\tilde{\zeta}(K) \equiv \sum_{k=1}^{K} \tfrac{1}{2}k(k+1).$$

Suppose by contradiction that conditional on all competitors $1, ..., K$ seeing their full history, there are greater than L periods with entry with positive probability on the equilibrium path. That means the incumbent's *equilibrium* strategy has her acquiescing with at least probability b at each of those histories (otherwise the competitors would not enter when they observe those full histories). The $(L+1)$th competitor entering observes the full history and must have a posterior greater than or equal to $\Upsilon^L(\mu^0) > p^*$ and therefore will not enter, a contradiction.

Thus, conditional on all competitors $1, ..., K$ seeing their full history, the average discounted payoff for periods $0, ..., K$ for playing this strategy is greater than or equal to

$$\begin{aligned}
& (1-\delta)\left[\lambda^{\tilde{\zeta}(K)}\left(\sum_{k=0}^{L-1}\delta^k(-1)+\sum_{k=L}^{K}\delta^k c\right)+(1-\lambda^{\tilde{\zeta}(K)})\sum_{k=0}^{K}\delta^k(-1)\right] \\
= \; & (1-\delta)\left[\lambda^{\tilde{\zeta}(K)}\left(\frac{\delta^L(1-\delta^{K+1})}{1-\delta}c\right)-\frac{1-\delta^L}{1-\delta}-(1-\lambda^{\tilde{\zeta}(K)})\frac{\delta^L(1-\delta^{K+1})}{1-\delta}\right] \\
= \; & [\lambda^{\tilde{\zeta}(K)}c-(1-\lambda^{\tilde{\zeta}(K)})]\delta^L(1-\delta^{K+1})-(1-\delta^L) \\
= \; & [\lambda^{\tilde{\zeta}(K)}\delta^L(1-\delta^{K+1})c-(1-\lambda^{\tilde{\zeta}(K)})\delta^L(1-\delta^{K+1})]-(1-\delta^L). \qquad (1.3.2)
\end{aligned}$$

The incumbent receives at least 0 for acquiescing in periods beyond K, so (1.3.2) is a lower bound for the average discounted payoff for all periods, and therefore also $\underline{v}_I(\mu^0, \delta, \lambda)$. Let any $\varepsilon' \in (0,1)$ and any $\varepsilon'' > 0$ be given. There exists K^* such that for all $K > K^*$, $\delta^{L+K+1}c < \varepsilon''$. Since $\lim_{\lambda \to 1} \lambda^{\tilde{\zeta}(K)} = 1$, there exists $\lambda^* \in (0,1)$ such that for all $K > K^*$ and $\lambda \in (\lambda^*, 1)$,

$$\begin{aligned}
\underline{v}_I(\mu^0, \delta, \lambda) \; & \geq \; [(1-\varepsilon')\delta^L(1-\delta^{K+1})c-\varepsilon']-(1-\delta^L) \\
& = \; [\delta^L c-\delta^{L+K+1}c-\varepsilon'\delta^L(1-\delta^{K+1})c-\varepsilon']-(1-\delta^L). \\
& > \; [\delta^L c-\varepsilon''-\varepsilon'\delta^L(1-\delta^{K+1})c-\varepsilon']-(1-\delta^L).
\end{aligned}$$

Pick $\varepsilon', \varepsilon''$ such that $\varepsilon' + \varepsilon' \delta^L (1 - \delta^{K+1}) c + \varepsilon'' = \varepsilon$. Then for $\lambda \in (\lambda^*, 1)$,

$$\underline{v}_I(\mu^0, \delta, \lambda) \geq \delta^L c - (1 - \delta^L) - \varepsilon.$$

Appendix 2

Proofs for Chapter 2

2.1 Proofs for Sections 2.2 and 2.3

2.1.1 Proof of Lemma 2.2.1

Given that the short-run players are playing public strategies given by vector $\tilde{\sigma}_2$, player 1's payoffs of playing any particular action does not depend on player 1's full (private) history. Let

$$V_{a_1}(\tilde{\sigma}|\mathbf{h}a_2) \equiv (1-\delta)u_1(a_2, a_1) + \delta \sum_{y \in Y} \rho(y|a_2, a_1) \max_{\sigma_1 \in \Sigma_1} V(\sigma_1, \tilde{\sigma}_2|\mathbf{h}a_2 a_1 y)$$

denote the value of choosing a_1 at full history $\mathbf{h}a_2$ and then choosing a best response at all subsequent histories, where $V(\sigma_1, \tilde{\sigma}_2|\mathbf{h}a_2 a_1 y)$ is the value of playing strategy σ_1 given player 2 strategies $\tilde{\sigma}_2$ at full history $\mathbf{h}a_2 a_1 y$. Define $\bar{\mathbf{H}}^t : H^t \to 2^{\mathbf{H}^t}$ so that $\bar{\mathbf{H}}^t(h^t)$ is the set of full histories at period t whose public component (those visible to player 2 at period t) is equal to h^t. Consider any two histories $\mathbf{h}, \check{\mathbf{h}} \in \bar{\mathbf{H}}^t(h^t)$ for some public history h^t. Because future player 2s cannot condition on the events of period $t-K$ or earlier,

$$\begin{aligned}
& V_{a_1}(\tilde{\sigma}_2|\mathbf{h}a_2) \\
= \; & (1-\delta)u_1(a_2, a_1) + \delta \sum_{y \in Y} \rho(y|a_2, a_1) \max_{\sigma_1 \in \Sigma_1} V(\sigma_1, \tilde{\sigma}_2|\mathbf{h}a_2 a_1 y) && (2.1.1) \\
= \; & (1-\delta)u_1(a_2, a_1) + \delta \sum_{y \in Y} \rho(y|a_2, a_1) \max_{\sigma_1 \in \Sigma_1} V(\sigma_1, \tilde{\sigma}_2|h^t) \\
= \; & (1-\delta)u_1(a_2, a_1) + \delta \sum_{y \in Y} \rho(y|a_2, a_1) \max_{\sigma_1 \in \Sigma_1} V(\sigma_1, \tilde{\sigma}_2|\check{\mathbf{h}}a_2 a_1 y) = V_{a_1}(\tilde{\sigma}_2|\check{\mathbf{h}}a_2) && (2.1.2)
\end{aligned}$$

for all $a_1 \in A_1$.

Define public strategy profile $\bar{\sigma}_1(a_1|h^t a_2) \equiv \sum_{\mathbf{h} \in \bar{\mathbf{H}}^t(h^t)} P_{\tilde{\sigma}}(\mathbf{h}a_2)\tilde{\sigma}_1(a_1|\mathbf{h}a_2)$ where

$P_{\tilde{\sigma}}(\mathbf{h}a_2)$ is the probability of $\mathbf{h}a_2$ being realized in equilibrium given $\tilde{\sigma}$. By (2.1.2), $\bar{\sigma}_1$ is a best response to $\tilde{\sigma}_2$. Because the equilibrium conditional probability of each public history is the same between $(\bar{\sigma}_1, \tilde{\sigma}_2)$ and $\tilde{\sigma}$, i.e. $P_{\bar{\sigma}_1,\tilde{\sigma}_2}(h^t|\theta) = P_{\tilde{\sigma}}(h^t|\theta)$ for each $h^t \in H, \theta \in \Theta$, player 2 has the beliefs $\tilde{\mu}$ are also consistent with $(\bar{\sigma}_1, \tilde{\sigma}_2)$. Furthermore, the expected (from the perspective of player 2) play by player 1 is also identical. Thus, $\tilde{\sigma}_2$ is a best response to $\bar{\sigma}_1$.

2.1.2 Proof of Proposition 2.3.3

Proof of Part 1

The proof is constructive, following the example of Proposition 7.3.1 in Mailath and Samuelson (2006b) where possible. Let any $w \equiv (\phi, \mu, \gamma) \in \mathcal{B}(W) \subset W$ be given. I construct a wPBE strategy profile σ for the variant game $G^{\infty}_{\phi,\mu}$. Define functions $\mathsf{Q} : \mathcal{B}(W) \to \mathcal{A}, \mathsf{U} : \mathcal{B}(W) \to W$ so that $(\phi, \mu, \mathsf{Q}(w))$ is enforced by $\mathsf{U}(w)$.

Specify the wPBE strategy profile σ as follows. Recursively define $\mathsf{U}^t(w) \in W$ as follows: $\mathsf{U}^0(w) = w$, $\mathsf{U}^t(w) = \mathsf{U}(\mathsf{U}^{t-1}(w))$. Let any full semipublic history $\mathfrak{h}^t \in Y^{K+t}$ be given — note that $\mathfrak{h}^t$ contains the initial history $h^0 \in Y^K$. Denote $(\phi^t, \mu^t, \gamma^t) = \mathsf{U}^t(w)$ and $\mathsf{Q}^t(w) = \mathsf{Q}(\mathsf{U}^t(w)) = \alpha^t$ for all t. Define the strategy profile $\sigma_2(h^t) = \alpha_2^t(h^t)$ and $\sigma_1(h^t a_2) = \alpha_1^t(h^t a_2)$ for each $t, h^t \in H^t, a_2 \in A_2$.

First I show that $V(\sigma|h^0) = \gamma(h^0)$ for initial history h^0. Then

$$\begin{aligned}
\gamma(h^0) &= \mathsf{V}((\phi, \mu, \mathsf{Q}(w)), \mathsf{U}(w))(h^0) = \mathsf{V}((\phi, \mu, \mathsf{Q}^0(w)), \mathsf{U}^1(w))(h^0) && (2.1.3)\\
&= (1-\delta)u_1(\alpha^0(h^0, \theta_0)) + \delta \sum_{a \in A} \sum_{y^0 \in Y} \alpha^0(a|h^0, \theta_0)\rho(y^0|a)\gamma^1(h^0 y^0) && (2.1.4)\\
&= (1-\delta)u_1(\sigma(h^0, \theta_0)) + \delta \sum_{a \in A} \sum_{y^0 \in Y} \alpha^0(a|h^0, \theta_0)\rho(y^0|a) && (2.1.5)\\
&\qquad \cdot \mathsf{V}((\phi^1, \mu^1, \alpha^1), \mathsf{U}^2(w))(h^0 y^0) && (2.1.6)\\
&= (1-\delta)u_1(\sigma(h^0, \theta_0)) + \delta \sum_{a \in A} \sum_{y^0 \in Y} \alpha^0(a|h^0, \theta_0)\rho(y^0|a) && (2.1.7)\\
&\qquad \cdot \Big[(1-\delta)u_1(\alpha^1(h^0 y^0, \theta_0)) && (2.1.8)
\end{aligned}$$

$$+\delta \sum_{a\in A}\sum_{y^1\in Y} \alpha^1(a|h^0y^0,\theta_0)\rho(y^1|a)\gamma^2(h^0y^0y^1)\Bigg] \quad (2.1.9)$$

$$= (1-\delta)\sum_{s=0}^{t-1}\delta^s \sum_{\mathfrak{h}^s\in Y^s} P_\sigma(h^0\mathfrak{h}^s|h_0)u_1(\sigma(h^s,\theta_0)) \quad (2.1.10)$$

$$+\delta^t \sum_{\mathfrak{h}^t\in Y^t} P_\sigma(h^0\mathfrak{h}^t|h^0,\theta_0)\gamma^t(\mathfrak{h}^t_{t-K},...,\mathfrak{h}^t_{t-1})), \quad (2.1.11)$$

where $P_\sigma(h^s|h^0,\theta_0)$ is the probability of history h^s occurring under strategy profile σ in game $G_{\phi,\mu}$ conditional on initial history h^0 and the normal type θ_0. Taking the limit $t\to\infty$ gives

$$v = (1-\delta)\sum_{s=0}^{\infty}\delta^s \sum_{h^s\in Y^s} P_\sigma(h^0\mathfrak{h}^s|h^0,\theta_0)u_1(\sigma(h^0\mathfrak{h}^s,\theta_0)) = V(\sigma|h^0,\theta_0). \quad (2.1.12)$$

To show that σ is a wPBE strategy profile, beliefs must be consistent with σ. Short-run players do not know the full history nor do they know player 1's type; thus, beliefs map from public histories to probability distributions over elements of the set $\Theta\times\mathbf{H}$. However, since player 1 is playing a public strategy, beliefs on the full history do not affect player 2's payoffs, so they can be ignored aside from knowing the public history; the only belief that matters is the belief on the type θ. Together, a belief on θ and the public history h^t characterize player 2's maximization problem. I show that beliefs $(\mu^t)_{t=0}^\infty$ on the type are consistent with σ. The conditional probability $P_\sigma(h^0|\theta)$ of initial history h^0 at period 0 is given by $\phi(h^0|\theta)$. Suppose that for some t the conditional (given θ) probability of public history h^t satisfies $P_\sigma(h^t|\theta)=\phi^t(h^t|\theta)$. Then the conditional probability of history $h^{t+1}=h^ty$ at period $t+1$ is

$$\begin{aligned} P_\sigma(h^{t+1}|\theta) &= \sum_{h^t\in\tau(h^{t+1})} P_\sigma(h^t|\theta)\sum_{a\in A}\sigma(a|h^t,\theta)\rho(y|a) \\ &= \sum_{h^t\in\tau(h^{t+1})} \phi^t(h^t|\theta)\sum_{a\in A}\alpha^t(a|h^t,\theta)\rho(y|a) = \phi^{t+1}(h^{t+1}|\theta), \end{aligned}$$

where $\tau(h^{t+1})\equiv\{h^t\in H^t : \forall k\in\{K-1,...,1\}, h^{t+1}_{-k-1}=h^t_{-k}\}$ is the set of period t histories that match the oldest $K-1$ periods of h^{t+1}. Then by induction $P_\sigma(h^t|\theta)=\phi^t(h^t|\theta)$ for all

t. The definition of wPBE requires that the belief of player 2 at t for histories h^t on the equilibrium path is

$$\mu(\theta|h^t) = \frac{P_\sigma(h^t|\theta)\dot{\mu}^0(\theta)}{\sum_{\theta'\in\Theta} P(h^t|\theta')\dot{\mu}^0(\theta')} = \frac{\phi^t(h^t|\theta)\dot{\mu}^0(\theta)}{\sum_{\theta'\in\Theta}\phi^t(h^t|\theta')\dot{\mu}^0(\theta')} = \mu^t(\theta|h^t), \tag{2.1.13}$$

so the consistency requirement is satisfied. For histories not on the equilibrium path, wPBE does not impose any requirement beyond that a belief is defined, which is provided by μ^t.

Finally, I show that there are no profitable one-shot deviations. For short-run players this follows from (2.3.3) in Definition 2.3.13. For all h^t, play at period t is given by $\sigma^t(h^t,\theta) = \alpha^t(h^t,\theta)$ and beliefs by μ^t; substituting α^t for α and μ^t for μ, (2.3.3) immediately implies there is no profitable deviation for the short-run player. The long-run player also has no profitable one-shot deviations due (2.3.5) in Definition 2.3.13, as is now shown. For any period t, (ϕ^t,μ^t,α^t) is enforced by $(\phi^{t+1},\mu^{t+1},\gamma^{t+1})$. Then for all $h^t \in H^t$ and $a_2^t \in A_2$,

$$\begin{aligned} & (1-\delta)u_1(a_2^t,\alpha_1(h^ta_2^t,\theta_0)) + \delta\sum_{a_1\in A}\sum_{y^t\in Y}\alpha_1(a_1|h^ta_2^t,\theta_0)\rho(y^t|a_2^t,a_1)\tilde{\gamma}(h^ty^t) \\ \geq \;& (1-\delta)u_1(a_2,a_1') + \delta\sum_{y^t\in Y}\rho(y^t|a_2^t,a_1')\gamma^{t+1}(h^ty^t) \end{aligned}$$

for all $a_1' \in A_1$. By the same argument as (2.1.3) through (2.1.12), it can be shown that $\gamma^{t+1}(a)$ is the continuation payoff of action profile a being played in period t. Thus, there are no profitable one-shot deviations for player 1.

For all $w \equiv (\phi,\mu,\gamma) \in W$, I have shown that $\gamma(h^0)$ is the value (to player 1) of a wPBE strategy profile for variant game $G^\infty_{\phi,\mu}$ conditional on initial history h^0. Thus, $w \in \mathcal{D}_{\phi,\mu} \subset \mathcal{D}$, and so $\mathcal{B}(W) \subset \mathcal{D}$.

Proof of Part 2

I prove that $\mathcal{D} \subset \mathcal{B}(\mathcal{D})$ because Part 1 implies that if $\mathcal{D}$ is self-generating, $\mathcal{B}(\mathcal{D}) \subset \mathcal{D}$ and so $\mathcal{B}(\mathcal{D}) = \mathcal{D}$.

Let any $w^0 \equiv (\phi^0,\mu^0,\gamma^0) \in \mathcal{D}$ be given. Then by Definition 2.3.9 and Lemma 2.2.1, there exists a public wPBE (σ, μ) for variant game $G^{\infty}_{\phi^0,\mu^0}$ such that $V(\sigma|h^0) = \gamma(h^0)$ for initial history h^0. Define history distribution ϕ^t so that $\phi^t(h^t|\theta) = P_\sigma(h^t|\theta)$. Because this is a wPBE, for any $h^t \in H^t$ on the equilibrium path, player 2 beliefs are given by Bayes' rule:

$$\mu(\theta|h^t) = \frac{\mu^0(\theta)P_\sigma(h^t|\theta)}{\sum_{\theta'\in\Theta}\mu^0(\theta')P_\sigma(h^t|\theta')} = \frac{\mu^0(\theta)\phi^t(h^t|\theta)}{\sum_{\theta'\in\Theta}\mu^0(\theta')\phi^t(h^t|\theta')}. \tag{2.1.14}$$

Define belief mapping μ^t such that $\mu^t(\theta|h^t) = \mu(\theta|h^t)$; (2.1.14) shows that μ^t is consistent with ϕ^t. Because σ is public, I can define action mapping α^t such that $\alpha_2^t(h^t) = \sigma_2(h^t)$ and $\alpha_1^t(h^t a_2, \theta) = \sigma(h^t a_2, \theta)$ for each $h^t \in H^t, a_2 \in A_2, \theta \in \Theta$. Also define payoff functions $\gamma^t(h^t) \equiv V(\sigma|h^t)$.

I now show that $x^t \equiv (\phi^t, \mu^t, \alpha^t)$ is enforced by $w^{t+1} \equiv (\phi^{t+1}, \mu^{t+1}, \gamma^{t+1})$. First,

$$\begin{aligned}
\phi^{t+1}(h^{t+1}|\theta) &= P_\sigma(h^{t+1}|\theta) \\
&= \sum_{h^t\in\tau(h^{t+1})}\sum_{a_2\in A_2}\sum_{y\in Y} P_\sigma(h^t|\theta)\sigma_2(a_2|h^t)\sigma_1(a_1|h^t a_2,\theta)\rho(y|a_2,a_1) \\
&= \sum_{h^t\in\tau(h^{t+1})}\sum_{a_2\in A_2}\sum_{y\in Y} \phi^t(h^t|\theta)\alpha_2(a_2|h^t)\alpha_1(a_1|h^t a_2,\theta)\rho(y|a_2,a_1),
\end{aligned}$$

so inducibility is satisfied: $(\phi^{t+1}, \mu^{t+1}) \in \Upsilon(\phi^t, \mu^t, \alpha^t)$. Incentive compatibility for the long-run and short-run players is a straightforward implication of σ not having profitable one-shot deviations. Thus, $x^t \equiv (\phi^t, \mu^t, \alpha^t)$ is enforced by $w^{t+1} \equiv (\phi^{t+1}, \mu^{t+1}, \gamma^{t+1})$. Furthermore,

$$\begin{aligned}
\gamma(h^t) &= V(\sigma|h^t) \\
&= \sum_{a_2\in A_2}\alpha_2(a_2|h^t)\sum_{a_1\in A_1}\alpha_1(a_1|h^t,\theta_0)\left[(1-\delta)u_1(a_2,a_1) + \sum_{y\in Y}\rho(y|a_2,a_1)V(\sigma|h^t y)\right] \\
&= \sum_{a_2\in A_2}\alpha_2(a_2|h^t)\sum_{a_1\in A_1}\alpha_1(a_1|h^t,\theta_0)\left[(1-\delta)u_1(a_2,a_1) + \sum_{y\in Y}\rho(y|a_2,a_1)\gamma(h^t y)\right] \\
&= \sum_{a_2\in A_2}\alpha_2(a_2|h^t)\mathsf{V}_{a_2}(x^t, w^{t+1})(h^t),
\end{aligned}$$

so $(\phi^t, \mu^t, \gamma^t)$ is decomposed by x^t and w^{t+1} (see Definition 2.3.14). Since the period $t+1$ continuation game is strategically equivalent to the $G^{\infty}_{\phi^{t+1},\mu^{t+1}}$ variant game, $w^{t+1} \in \mathcal{D}$.

Thus, since $w^1 \in \mathcal{D}$ and w^0 is generated by w^1, $w^0 \in \mathcal{B}(\mathcal{D})$ so $\mathcal{D} \subset \mathcal{B}(\mathcal{D})$.

2.1.3 Proof of Proposition 2.3.4

I start with a useful lemma that establishes $\mathcal{B}$ is additive in the sense of set unions, and therefore monotonic in the sense of set inclusion.

Lemma 2.1.1. *Let any two sets W, W' of HBPs be given. Then $\mathcal{B}(W \cup W') = \mathcal{B}(W) \cup \mathcal{B}(W')$.*

Proof. I first show that $\mathcal{B}(W \cup W') \subset (\mathcal{B}(W) \cup \mathcal{B}(W'))$. Let any $w \in \mathcal{B}(W \cup W')$ be given. Then there exist HBA x and HBP $\tilde{w} \in (W \cup W')$ which decompose w. Without loss of generality, suppose $\tilde{w} \subset W$. Since w is decomposed by x and $\tilde{w}$, then $w \in \mathcal{B}(W) \subset (\mathcal{B}(W) \cup \mathcal{B}(W'))$. Thus $\mathcal{B}(W \cup W') \subset (\mathcal{B}(W) \cup \mathcal{B}(W'))$.

Now let $w \in (\mathcal{B}(W) \cup \mathcal{B}(W'))$ be given. Without loss of generality, suppose $w \in \mathcal{B}(W)$, so then there exist HBA x and HBP $\tilde{w} \in W$ which decompose w. Since $\tilde{w} \in (W \cup W')$, it is also true that $w \in \mathcal{B}(W \cup W')$, so $(\mathcal{B}(W) \cup \mathcal{B}(W')) \subset \mathcal{B}(W \cup W')$. □

The following lemma is a virtually identical adaptation from the APS version, so the proof is omitted (see Lemma 7.3.2 of Mailath and Samuelson (2006b)).

Lemma 2.1.2. *If W is compact, $\mathcal{B}(W)$ is closed.*

$\mathcal{F}^\dagger$ is compact and the set of HBPs decomposable on $\mathcal{F}^\dagger$ is also feasible: $\mathcal{B}(\mathcal{F}^\dagger) \subset \mathcal{F}^\dagger$. Proposition 2.3.3 and Lemma 2.1.1 imply that, for any m, $\mathcal{D}(\delta) \subset \mathcal{B}^m(\mathcal{F}^\dagger) \subset \mathcal{F}^\dagger$. Repeatedly applying $\mathcal{B}$ therefore gives a decreasing sequence $\{\mathcal{B}^m(\mathcal{F}^\dagger)\}_{m=0}^{\infty}$. Then

$$\mathcal{D} \subset \mathcal{F}^\dagger_\infty \equiv \bigcap_{m=0}^{\infty} \mathcal{B}^m(\mathcal{F}^\dagger).$$

I adapt the proof of Proposition 7.3.3 of Mailath and Samuelson (2006b) to my setting, showing that $\mathcal{F}^\dagger_\infty \subset \mathcal{B}(\mathcal{F}^\dagger_\infty)$, which by Proposition 2.3.3 proves the result. For any

$w \equiv (\phi, \mu, \gamma) \in \mathcal{F}^\dagger_\infty$, $w \in \mathcal{B}^m(\mathcal{F}^\dagger)$ for all $m \in \{1, 2, ...\}$, and there exists $(x^m, \tilde{w}^m)$ such that $x^m \equiv (\phi, \mu, \alpha^m) \in \mathcal{X}$, $\tilde{w}^m \in \mathcal{B}^{m-1}(\mathcal{F}^\dagger)$, and w is decomposed by x^m and $\tilde{w}^m$.

Because the sequence $(x^m, \tilde{w}^m)_{m=0}^\infty$ is bounded, without loss of generality assume that it converges to a limit $(x^*, \tilde{w}^*) \equiv ((\phi, \mu, \alpha^*), (\tilde{\phi}^*, \tilde{\mu}^*, \tilde{\gamma}^*))$, using a convergent subsequence if necessary by the Bolzano-Weierstrass theorem. I show that $\tilde{w}^* \in \mathcal{F}^\dagger_\infty$, that x^* is enforced by $\tilde{w}^*$ and that w is decomposed by x^* and $\tilde{w}^*$.

Suppose by contradiction that $\tilde{w}^* \notin \mathcal{F}^\dagger_\infty$. By Lemma 2.1.2, $\mathcal{F}^\dagger_\infty$ is closed, so there exists $\varepsilon > 0$ such that $\bar{B}_\varepsilon(\tilde{w}^*) \cap \mathcal{F}^\dagger_\infty = \varnothing$, where $\bar{B}_\varepsilon(\tilde{w})$ is defined as follows:[1]

$$\begin{aligned} \bar{B}_\varepsilon((\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})) \quad \equiv \quad & \{(\phi', \mu', \gamma') : \forall h \in Y^K, \forall \theta \in \Theta, \phi'(h|\theta) \in [\tilde{\phi}(h|\theta) - \varepsilon, \tilde{\phi}(h|\theta) + \varepsilon], \\ & \mu' \in M_{\phi'}, \mu'(\theta|h) \in [\tilde{\mu}(\theta|h) - \varepsilon, \tilde{\mu}(\theta|h) + \varepsilon], \gamma' \in \bar{B}^\Gamma_\varepsilon(\tilde{\gamma})\} \end{aligned}$$

where $\bar{B}^\Gamma_\varepsilon(\tilde{\gamma})$ is the closed ball in Γ (which is a $|Y|^K$-dimensional Euclidean space) of radius ε with center $\tilde{\gamma}$. There exists m' such that for all $m > m'$, $\tilde{w}^m \in \bar{B}_\varepsilon(\tilde{w}^*)$ and because $\{\mathcal{B}^m(\mathcal{F}^\dagger)\}_{m=0}^\infty$ is a decreasing sequence, for any $m'' > m'$,

$$\bar{B}_\varepsilon(\tilde{w}^*) \cap \left(\bigcap_{m \le m''} \mathcal{B}^m(\mathcal{F}^\dagger) \right) \neq \varnothing.$$

The collection $\{\bar{B}_\varepsilon(\tilde{w}^*)\} \bigcup_{m=1}^\infty \{\mathcal{B}^m(\mathcal{F}^\dagger)\}$ has the finite intersection property and $\bar{B}_\varepsilon(\tilde{w}^*) \cup \mathcal{F}^\dagger$ is compact, so the aforementioned collection has a non-empty intersection (by Theorem 4.7.15 of Corbae, Stinchcombe, and Zeman (2009)): $\bar{B}_\varepsilon(\tilde{w}^*) \cap \mathcal{F}^\dagger_\infty \neq \varnothing$, a contradiction.

It is easy to see that $x^* \in \mathcal{X}$ because $\mathcal{X}$ is closed and all $x^m \in \mathcal{X}$. Since x^m is enforced by $\tilde{w}^m$ for all m, taking the limit it is straightforward to show that x^* is enforced by $\tilde{w}^*$. Similarly it is clear that $\lim_{m \to \infty} \mathsf{V}(x^m, \tilde{w}^m) = \gamma = \mathsf{V}(x^*, \tilde{w}^*)$, w is decomposed by x^* and $\tilde{w}^*$. Thus, $\mathcal{F}^\dagger_\infty$ is self generating and bounded, so $\mathcal{F}^\dagger_\infty \subset \mathcal{D}$, and by Proposition 2.3.3, $\mathcal{F}^\dagger_\infty = \mathcal{D}$.

[1] The "closed ball with center $\tilde{w}$ and radius ε" is not well defined in the space $\mathcal{W} \equiv \mathcal{M} \times \Gamma$ because $\mathcal{M}$ is not a Euclidean space.

2.1.4 Proof of Proposition 2.3.5

I prove that each element of $\mathcal{V}(W)$ is a wPBE player 1 payoff for the full game G^∞. Let any $v \in \mathcal{V}(W)$ be given. By Definition 2.3.18, there exists HBP $(\phi, \mu, \gamma) \in \mathcal{D}$ such that $v \in \mathcal{E}_{(\phi,\mu,\gamma)}$. This means there exists a strategy profile $\check{\sigma} \in \Sigma^*_\gamma$ of the γ-antegame such that the distribution of outcomes matches ϕ, i.e. $P_{\check{\sigma}}(h|\theta) = \phi(h|\theta)$ for each $h \in Y^K$ (where $P_{\check{\sigma}}(\cdot)$ is defined as it was just after (2.1.10)), and with value $V(\check{\sigma}) = v$. Let $\check{\mu}$ be the associated beliefs of the equilibrium with strategy profile $\check{\sigma}$.

As in the proof of part 1 of Proposition 2.3.1, define functions $\mathsf{Q} : \mathcal{D} \to \mathcal{A}, \mathsf{U} : \mathcal{D} \to \mathcal{D}$ so that $(\phi, \mu, \mathsf{Q}(w))$ is enforced by $\mathsf{U}(w)$. Recursively define $\mathsf{U}^t(w) \in W$ for $t \in \{K, K+1, ...\}$ as follows: $\mathsf{U}^K(w) = w$, $\mathsf{U}^{t+1}(w) = \mathsf{U}(\mathsf{U}^t(w))$. Denote $(\phi^t, \mu^t, \gamma^t) = \mathsf{U}^t(w)$ and $\mathsf{Q}^t(w) = \mathsf{Q}(\mathsf{U}^t(w)) = \alpha^t$ for all t.

For the full game G^∞, define public strategy profile $\sigma \in \hat{\Sigma}$ such that $\sigma_2^t(h) = \check{\sigma}_2^t(h)$ and $\sigma_1(t, ha_2)$ for each date-history $(t, h) \in H$ and each $a_2 \in A_2$. For periods $t \geq K$, define $\sigma_2(t+K, h) = \alpha_2^{t+K}(h)$ and $\sigma_1(t+K, ha_2) = \alpha_1(t+K, ha_2, \theta_0)$ for each $(t+K, h) \in H, a_2 \in A_2$. Define $(\phi^{t+K}, \mu^{t+K}, \gamma^{t+K}) = \mathsf{U}^t(w)$. Define beliefs μ^* such that $\mu^*(\theta|t, h) = \mu^t(\theta|h)$ for all $t \geq K$.

Note that $P_{\check{\sigma}}(h|\theta) = P_\sigma((K, h)|\theta) = \phi^K(h|\theta)$. Suppose that for some some $t \geq K$, $P_\sigma((t, h)|\theta) = \phi^t(h|\theta)$. Since $(\phi^{t+1}, \mu^{t+1}, \gamma^{t+1}) = \mathsf{U}((\phi^t, \mu^t, \gamma^t)) \in \mathcal{B}(\{(\phi^t, \mu^t, \gamma^t)\})$, the HB (ϕ^{t+1}, μ^{t+1}) is induced by $(\phi^t, \mu^t, \alpha^t)$. Then for all $h' \in Y^K, y \in Y$,

$$\begin{aligned}
\phi^{t+1}(h'y|\theta) &= \sum_{h \in \tau(h'y)} \sum_{a_2 \in A_2} \sum_{y \in Y} \phi^t(h|\theta)\alpha_2^t(a_2|h)\alpha_1^t(a_1|ha_2, \theta)\rho(y|a_2, a_1) \\
&= \sum_{h \in \tau(h'y)} \sum_{a_2 \in A_2} \sum_{y \in Y} P_\sigma((t, h)|\theta)\sigma_2(a_2|t, h)\sigma_1(a_1|t, ha_2, \theta)\rho(y|a_2, a_1) \\
&= P_\sigma((t+1, h'y)|\theta).
\end{aligned}$$

Then by induction, $P_\sigma((t, h)|\theta) = \phi^t(h|\theta)$ holds true at all $t \geq K$. The direct analogue of (2.1.13) shows that $\mu^*(\theta|h)$ is consistent at $t \geq K$, and the analogous steps of (2.1.3) through (2.1.12) show that $V(\sigma|t, h) = \gamma^t(h)$. The incentive compatibility requirements of

Definition 2.3.13 directly imply the absence of profitable one-shot deviations in (σ, μ^*) for $t \geq K$.

The arguments above also show that continuation payoffs $V(\sigma|t, h^K)$ for play beginning at period K with public history h is given by $\gamma(h^K)$. The fact that $\check{\sigma}$ is a wPBE of the γ-antegame means that by incentive compatibility at for player 1 at period $K-1$ with history h^{K-1},

$$\begin{aligned} & (1-\delta)u_1(a_2, \check{\sigma}_1(h^{K-1}a_2)) + \delta \sum_{a_1 \in A_1} \sum_{y^{K-1} \in Y} \check{\sigma}_1(a_1|h^{K-1}a_2)\rho(y|a_2, a_1)\gamma(h^{K-1}y^{K-1}) \\ \geq \quad & (1-\delta)u_1(a_2, a_1') + \delta \sum_{y^{K-1} \in Y} \rho(y|a_2, a_1')\gamma(h^{K-1}y^{K-1}) \end{aligned}$$

for each $a_1' \in A_1$, so substituting $\sigma_1(h^{K-1}a_2)$ for $\check{\sigma}_1(h^{K-1}a_2)$ and $V(\sigma|h^{K-1}y^{K-1})$ for $\gamma(h^{K-1}y^{K-1})$ proves that σ is incentive compatible for player 1 at period $K-1$. By backward induction, it is straightforward to show that $\check{\sigma}$ being a wPBE of the γ-antegame implies incentive compatibility for σ in the full game and $V(\check{\sigma}|h^t) = V(\sigma|h^t)$ for each $t \in \{0, ..., K-1\}$ and $h^t \in H^t$. Because player 2 observes the full history for periods $0, \ldots, K-1$, beliefs carry over from the antegame to the full game without modification ($\check{\mu}(\theta|h) = \mu^*(\theta|h)$ for all $t \in \{0, ..., K-1\}, h \in H^t, \theta \in \Theta$). Thus, σ is a wPBE of the full game and $v = V(\check{\sigma}|\varnothing) = V(\sigma|\varnothing)$, so $v \in \mathcal{E}$.

I now prove $\mathcal{V}(\mathcal{D}) = \mathcal{E}$, and since the above arguments have shown $\mathcal{V}(\mathcal{D}) \subset \mathcal{E}$, I show $\mathcal{E} \subset \mathcal{V}(\mathcal{D})$. Let $v \in \mathcal{E}$ and the associated public wPBE (σ, μ^*) of the full game G^∞ be given (so that $v = V(\sigma)$). At period K, the equilibrium distribution of histories conditional on type θ is $P_\sigma((K, h^K)|\theta)$. Define history distribution ϕ^K so that $\phi^K(h^K|\theta) = P_\sigma((K, h^K)|\theta)$, define belief mapping μ^K so that $\mu^K(\theta|h^K) = \mu^*(\theta|(K, h^K))$, and payoff function $\gamma^K(h^K) = V(\sigma|(K, h^K))$.

I show that $w \equiv (\phi^K, \mu^K, \gamma^K) \in \mathcal{D}_{\phi^K, \mu^K}$. Define strategy profile $\tilde{\sigma} \in \Sigma^*_{\phi^K, \mu^K}$ of the (ϕ^K, μ^K)-variant game as follows. Set $\tilde{\sigma}_2(t, h) = \sigma_2(t+K, h)$ and $\tilde{\sigma}_1(t, ha_2) = \sigma_1(t+K, ha_2)$ for all $h \in Y^K, a_2 \in A_2$. Define beliefs $\tilde{\mu}^*(t, h) = \mu(t+K, h)$. Since beliefs and strategies are the same (shifted by K periods), the lack of profitable deviations in (σ, μ) for periods

$K, K+1, ...$ implies a lack of profitable deviations for $(\tilde{\sigma}, \tilde{\mu}^*)$ for periods $0, 1, ...$ Thus, $(\tilde{\sigma}, \tilde{\mu}^*)$ is a wPBE of the (ϕ^K, μ^K)-variant game. Since $\sigma|_{(K,h)} = \tilde{\sigma}|_{(0,h)}$, $V(\tilde{\sigma}|(0,h)) = V(\sigma|(t,h)) = \gamma^K(h)$. Then by Definition 2.3.9, $w \in \mathcal{D}_{\phi^K, \mu^K} \subset \mathcal{D}$.

For the γ^K-antegame, define strategy profile $\check{\sigma} \in \Sigma_{\gamma^K}$ so that $\check{\sigma}(h^t) = \sigma(h^t)$. For the antegame, define beliefs $\check{\mu}^*$ so that $\check{\mu}^*(h) = \mu^*(h)$ for all $t \in \{0, ..., K-1\}, h \in H^t$. Since the short-run players in periods $0, ..., K-1$ observe the full history in the full game G^∞, the beliefs $\check{\mu}$ in the γ^K-antegame are consistent with $\check{\sigma}$. Since there are no profitable deviations in (σ, μ) and the maximization problems are identical in the antegame at every history $t \in \{0, ..., K-1\}, h \in H^t$, there are also no profitable one-shot deviations in $\check{\sigma}$, so $(\check{\sigma}, \mu)$ is a wPBE of the γ-antegame with distribution of period K histories ϕ^K. Thus, $v \in \mathcal{E}_{(\phi^K, \mu^K, \gamma^K)}$, and so $v \in \mathcal{V}(\mathcal{D})$.

2.1.5 Proof of Proposition 2.3.6

Let any stationary wPBE (σ^*, μ^*) be given. Note that this is a Markov chain on the state space of histories.

Lemma 2.1.3. *Let some finite state space $\Omega \equiv \{\omega_1, \ldots, \omega_n\}$ and a Markov chain $\{X_t\}$ on that state space with transition matrix Q be given. Let some initial probability distribution on the states $\pi^0 \equiv (\pi_1^0, \ldots, \pi_n^0)$ be given, and let π^∞ be the* time average distribution *of X_t, i.e. for each $i \in \{1, \ldots, n\}$*

$$\pi^\infty \equiv \lim_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t-1} \pi^0 Q^s. \tag{2.1.15}$$

Then π^∞ is stationary, i.e. $\pi^\infty = \pi^\infty Q$.

Proof. Every Markov chain on a finite state space has at least one stationary distribution (Furman, 2011). For aperiodic states, the limit (2.1.15) exists (see Property 2.17 of Gallager (2012)); for periodic states, the limit also clearly exists. Finally,[2] letting $r^t \equiv \frac{1}{t}\sum_{s=0}^{t-1} \pi^0 Q^s$,

[2] The conclusion of this proof is based on an answer at Mathematics Stack Exchange: `http://math.stackexchange.com/questions/897667/for-finite-markov-chain-time-average-distribution-is-always-a-stationary-distri`.

I have

$$r^t Q = \frac{1}{t}\sum_{s=0}^{t-1} \pi^0 Q^{s+1} = \frac{1}{t}\left(\sum_{s=0}^{t-1} \pi^0 Q^s + \pi^0 Q^t - \pi^0\right),$$

so taking the limit gives

$$\begin{aligned}
\pi^\infty Q &= \left(\lim_{t\to\infty} r^t\right) Q = \lim_{t\to\infty} r^t Q = \lim_{t\to\infty} \frac{1}{t}\left(\sum_{s=0}^{t-1} \pi^0 Q^s + \pi^0 Q^t - \pi^0\right) \\
&= \lim_{t\to\infty} \frac{1}{t}\sum_{s=0}^{t-1} \pi^0 Q^s = \pi^\infty,
\end{aligned}$$

proving the lemma. □

Define

$$\pi^\infty(h|\theta) \equiv \lim_{t\to\infty} \frac{1}{t-K}\sum_{t=K}^{\infty} P_\sigma^t(h|\theta)$$

as the time average probability of history $h \in Y^K$ generated by the Markov chain specified by σ^* conditional on type θ as given in (2.1.15); note that the initial distribution of this Markov chain is $\pi^0(h|\theta) = P_\sigma^K(h|\theta)$. Define ϕ so that $\phi(h|\theta) \equiv \pi^\infty(h|\theta)$, and define belief mapping μ so that $\mu(\theta|h) \equiv \mu^*(\theta|h)$. Note that $\mu \in M_\phi$ because μ^* satisfies Bayes' rule on the equilibrium path. Define action mapping α so that $\alpha_2(h) \equiv \sigma_2^*(h)$ and $\alpha_1(ha_2, \theta) \equiv \sigma_1^*(ha_2, \theta)$. I show that $x \equiv (\phi, \mu, \alpha)$ is enforced by $w \equiv (\phi, \mu, \gamma)$. Lemma 2.1.3 shows that $\pi^\infty(\cdot|\theta)$ is stationary, so for any history $h'y \in Y^K$,

$$\begin{aligned}
\phi(h'y|\theta) &= \pi^\infty(h|\theta) \\
&= \sum_{h\in\tau(h'y)} \sum_{a_2\in A_2} \sum_{a_1\in A_1} \pi^\infty(h|\theta)\sigma_2(a_2|h)\sigma_1(a_1|h,\theta)\rho(y|a_2,a_1) \\
&= \sum_{h\in\tau(h'y)} \sum_{a_2\in A_2} \sum_{a_1\in A_1} \phi(h|\theta)\alpha_2(a_2|h)\alpha_1(a_1|h,\theta)\rho(y|a_2,a_1),
\end{aligned}$$

so inducibility is satisfied: $(\phi, \mu) \in \Upsilon(\phi, \mu, \alpha)$. Defining payoff function γ so that $\gamma(h) \equiv V(h)$, the lack of profitable one-shot deviations in σ^* immediately implies incentive compatibility for both players, so x is enforced by w. Finally,

$$\gamma(h) = V(h)$$

$$
\begin{aligned}
&= \sum_{a_2 \in A_2} \sum_{a_1 \in A_1} \sigma_2(a_2|h)\sigma_1(a_1|h) \left[(1-\delta)u_1(a_2,a_1) + \delta \sum_{y \in Y} \rho(y|a_2,a_1)V(hy)\right] \\
&= \mathsf{V}(x,w),
\end{aligned}
$$

so w is decomposed by x and itself.

2.2 Proofs for Section 2.4

2.2.1 Proof of Proposition 2.4.1

Although player 1 observes all her past shocks, all of them except the current shock can be ignored in equilibrium. The following definition and lemma (essentially the same as Definition 4 and Lemma 2 in BMM) show this formally.

Definition 2.2.1. A strategy $\tilde{\sigma}_1$ is a *current shock strategy* if for all $\tilde{\mathbf{h}} \in \tilde{\mathbf{H}}_1$, containing non-shock history $\mathbf{h}a_2 \in \mathbf{H} \times A_2$ and shocks $z_1^0, \ldots, z_1^t$,

$$
\sigma_1(\mathbf{h}a_2, z_1^0, ..., z_1^{t-1}, z_1^t) = \sigma_1(\mathbf{h}a_2, \bar{z}_1^{t-K}, ..., \bar{z}_1^{t-1}, z_1^t)
$$

for almost all $z_1^t \in Z_1$ and any $\bar{z}_1^0, \ldots, \bar{z}_1^{t-1} \in Z_1$.

Lemma 2.2.1. *If $\tilde{\sigma}_1$ is a best response to $\tilde{\sigma}_2$, then $\tilde{\sigma}_1$ is a current shock strategy.*

Proof. Let a player 1 history $\tilde{\mathbf{h}} \in \tilde{\mathbf{H}}_1^t$, containing non-shock history $\mathbf{h}a_2 \in \mathbf{H}^t \times A_2$ and shocks $z_1^0, \ldots, z_1^t$, be given. Denote the public history at period t as h. Let $V^{t+1}(\sigma_1, \tilde{\sigma}_2|hya_2^{t+1}z_1^{t+1})$ be the value of playing the strategy σ_1 starting at period $t+1$, following the realizations of a_2^{t+1} and z_1^{t+1}, and let $\hat{\sigma}_1 \equiv \arg\max_{\sigma_1} V^{t+1}(\sigma_1, \tilde{\sigma}_2|hya_2^{t+1}, z_1^{t+1})$ be the best response. The payoff (at period t) of playing action a_1 after player 2 action a_2^t is

$$
\begin{aligned}
&\tilde{V}^*(a_1, \tilde{\sigma}_2|ha_2, z_1^t) \\
\equiv\ & \sum_{y \in Y} \rho(y|a_2,a_1) \Bigg[(1-\delta)(u_1(a_2,a_1) + \varepsilon z_1^{t,a_1}) \\
& \qquad + \delta \int\int \sum_{a_2} \tilde{\sigma}_2^{t+1}(a_2^{t+1}|hyz_2^{t+1}) V^{t+1}(\hat{\sigma}_1, \tilde{\sigma}_2|hya_2^{t+1}, z_1^{t+1})
\end{aligned}
$$

$$d\psi(z_1^{t+1})\, d\psi(z_2^{t+1})\Bigg]. \tag{2.2.1}$$

Note that the decision problem is independent of all shocks before period t. For any two actions $a_1, a_1' \in A_1$, player 1 can only be indifferent between the two actions if $z_1^{t,a_1} - z_1^{t,a_1'} = \zeta$ for some constant ζ. Thus, for almost all realizations of z_1^t, player 1 has a unique best response and thus $\tilde{\sigma}_1$ does not condition on $z_1^0, \ldots, z_1^{t-1}$. □

The following result is an extension of Lemma 1 of BMM.

Lemma 2.2.2. *Every essentially sequentially strict wPBE is quasi-Markov perfect.*

Proof. Let any two full semipublic histories $\mathfrak{h}^t, \bar{\mathfrak{h}}^t$ which lead to the same quasi-Markov state for player 2 be given. I will show that sequential strictness implies the same behavior at the two histories.

Let some $k \geq K$ and k-length sequence $y^k \in Y^k$ be given. Since players at periods $t+K+1, t+K+2, ...$ cannot observe period t, it is clear that the value function $V(\mathfrak{h}^t y^k)$ does not depend on $\mathfrak{h}^t$.

Now, let a $(K-1)$-length sequence $y^{K-1} \in Y^{K-1}$. For each $a_2^{t+K-1} \in A_2$, the decision problem facing player 1 at $\mathfrak{h}^t y^{K-1} a_2^{t+K-1}$ (at period $t+K-1$) is independent of $\mathfrak{h}^t$, and because the equilibrium is essentially sequentially strict, the set of maximizing actions is the same singleton for almost all shocks $z_1^{t+K-1} \in Z_1$. Then I have $\sigma_1(\mathfrak{h}^t y^{K-1} a_2, z_1^{t+K-1}) = \sigma_1(\bar{\mathfrak{h}}^t y^{K-1} a_2, z_1^{t+K-1})$ for almost all z_1^{t+K-1}. By Definition 2.4.1, player 2 has the same beliefs $\mu(\mathfrak{h}_{t-1}^t y^{K-1}) = \mu(\bar{\mathfrak{h}}_{t-1}^t y^{K-1})$ at both public histories (the shock z_2^{t+K-1} does not affect the belief since it is independently drawn). Since player 1's subsequent action is identical at both histories with probability one, player 2's decision problem is the same at both histories, giving $\sigma_2(\mathfrak{h}^t y^{K-1}, z_2^{t+K-1}) = \sigma_1(\bar{\mathfrak{h}}^t y^{K-1},, z_2^{t+K-1})$ for almost all shocks $z_2^{t+K-1} \in Z_2^{t+K-1}$.

Now suppose that for any $k \in \{0, ..., K-2\}$, $V(\mathfrak{h}^t y^{k+1}) = V(\bar{\mathfrak{h}}^t y^{k+1})$, where $y^k \in Y^k$. Then as above, the decision problem and unique maximizer for player 1 is the same at $\mathfrak{h}^t y^k a_2 z_1^{t+k}$ and $\bar{\mathfrak{h}}^t y^k a_2 z_1^{t+k}$ for almost all $z_1^{t+k} \in Z_1$; similarly, player 2 has the same beliefs at $\mathfrak{h}^t_{t-(K-k)} \cdots \mathfrak{h}^t_{t-1} y^k$ and $\bar{\mathfrak{h}}^t_{t-(K-k)} \cdots \bar{\mathfrak{h}}^t_{t-1} y^k$, yielding the same unique maximizing action at both histories for almost all shocks $z_2^{t+k} \in Z_2$, giving $V(\mathfrak{h}^t_{t-(K-k)} \cdots \mathfrak{h}^t_{t-1} y^k) = V(\bar{\mathfrak{h}}^t_{t-(K-k)} \cdots \bar{\mathfrak{h}}^t_{t-1} y^k)$. By backwards induction, I have

$$\sigma_2(\mathfrak{h}^t_{t-K} \cdots \mathfrak{h}^t_{t-1} z_2^{t+k}) = \sigma_2(\bar{\mathfrak{h}}^t_{t-K} \cdots \bar{\mathfrak{h}}^t_{t-1} z_2^{t+k}) \qquad \text{and} \qquad \sigma_1(\mathfrak{h}^t a_2 z_1^{t+k}) = \sigma_1(\bar{\mathfrak{h}}^t a_2 z_1^{t+k})$$

for almost all $z_2^{t+k} \in Z_2, z_1^{t+k} \in Z_1$. Thus, strategies are the same within a quasi-Markov state for player 2. An almost identical argument shows the same for quasi-Markov states of player 1, and so the equilibrium is quasi-Markov perfect. □

Consider player 1's decision at some period t. Suppose player 1 is indifferent between distinct actions a_1 and a_1' at public history $h \in H^t$, player 2 action a_2 and shock realization z_1; borrowing notation from the proof of Lemma 2.2.1 (specifically (2.2.1)), I can write this as

$$\tilde{V}^*(a_1, \tilde{\sigma}_2 | h a_2, z_1^t) = \tilde{V}^*(a_1', \tilde{\sigma}_2 | h a_2, z_1^t). \tag{2.2.2}$$

Define

$$V^*(\tilde{\sigma}_2 | hy) \equiv \int \int \sum_{a_2 \in A_2} \tilde{\sigma}_2^{t+1}(a_2^{t+1} | hy z_2^{t+1}) V^{t+1}(\hat{\sigma}_1, \tilde{\sigma}_2 | hy a_2^{t+1}, z_1^{t+1})\, d\psi(z_1^{t+1})\, d\psi(z_2^{t+1})$$

as the integral term in (2.2.1). Substituting into (2.2.2) and rearranging yields

$$\varepsilon(z_1^{a_1} - z_1^{a_1'}) = (1-\delta)[u_1(a_2, a_1') - u_1(a_2, a_1)] + \delta \sum_{y \in Y} [\rho(y | a_2, a_1) - \rho(y | a_2, a_1')] V^*(\tilde{\sigma}_2 | hy). \tag{2.2.3}$$

This implies that the set of shocks z_1 such that (2.2.3) holds is Lebesgue measure zero, so the profile is essentially sequentially strict. Lemma 2.2.2 shows that every essentially strict equilibrium is quasi-Markov perfect, so the wPBE is quasi-Markov.

2.2.2 Proof of Proposition 2.4.2

Proposition 2.4.1 shows that every wPBE in any perturbed game is quasi-Markov. Thus, for any sequence $(\psi^k, \varepsilon^k)_k$ where $\lim_{k\to\infty} \varepsilon^k = 0$, the limit of any sequence of wPBEs of the (ψ^k, ε^k)-perturbed game must converge to a quasi-Markov equilibrium.

2.3 Proofs for Section 2.5

2.3.1 Proof of Proposition 2.5.1

First, suppose $\delta < \frac{1}{2}$ and let any quasi-Markov equilibrium be given. For any period t, let

$$\hat{a}_1^t = \begin{cases} C & (1-\delta)u_1(c, C) + \delta V^{t+1}(C) \geq (1-\delta)u_1(c, D) + \delta V^{t+1}(D) \\ D & \text{otherwise} \end{cases}$$

be a best response at period t following player 2 playing c (note that the best response is independent of the history), where $V^{t+1}(\cdot)$ is the continuation payoff at period $t+1$. Note that $\hat{a}_1^t$ is also a best response after player 2 plays d. Note that

$$V^t(h^t c) = (1-\delta)u_1(c, \hat{a}_1^t) + \delta V^{t+1}(\hat{a}_1^t), \qquad V^t(h^t d) = (1-\delta)u_1(d, \hat{a}_1^t) + \delta V^{t+1}(\hat{a}_1^t)$$

so $V^t(h^t c) = V^t(h^t d) + (1-\delta)$. Since $V^t(h^t a_2)$ is independent of h^t,

$$\begin{aligned} V^t(C) - V^t(D) &= (\sigma_2^t(C) - \sigma_2^t(D))V^t(h^t c) + ((1-\sigma_2^t(C)) - (1-\sigma_2^t(D)))V^t(h^t d) \\ &= (\sigma_2^t(C) - \sigma_2^t(D))V^t(h^t c) + (\sigma_2^t(D) - \sigma_2^t(C))V^t(h^t d) \\ &= (\sigma_2^t(C) - \sigma_2^t(D))[V^t(h^t c) - V^t(h^t d)] \\ &= (1-\delta)(\sigma_2^t(C) - \sigma_2^t(D)). \end{aligned}$$

Note that if C is a best response at any period t following player 2 action a_2,

$$(1-\delta)u_1(a_2, C) + \delta V^{t+1}(C) \geq (1-\delta)u_1(a_2, D) + \delta V^{t+1}(D)$$

$$\delta(V^{t+1}(C) - V^{t+1}(D)) \geq (1-\delta)(u_1(a_2, D) - u_1(a_2, C)) = 1-\delta$$

$$\delta(1-\delta) \geq \delta(V^{t+1}(C) - V^{t+1}(D)) \geq 1-\delta,$$

a contradiction, so player 1's strict best response is always D. For history D at any period, the belief on the commitment type is 0 so player 2's best response is d. Similarly, at period 0, for $\mu^0 < \frac{1}{2}$, player 2's best response is d, so (d, D) is played every period and the equilibrium payoff is 1. For $\mu^0 > \frac{1}{2}$, player 2's best response at period 0 is c, so (c, D) is played in period 0 and then (d, D) every period for every period after that. For $\mu^0 = \frac{1}{2}$, player 2 is indifferent at period 0, any mixture for player 2 is possible in period 0, and (d, D) is played after that. Thus, (2.5.2) has been proven.

I begin by introducing notation and some useful preliminary lemmas. For brevity, for any history distribution ϕ, denote $\phi(C|\theta_0)$ simply as ϕ; for any action mapping α, denote $\alpha_2(C|h)$ as $\alpha_2(h)$, and denote $\alpha_1(C|h, \theta_0)$ as α_1 (since player 1 does not condition on history due to quasi-Markov perfection). Define $\mathcal{I} \equiv \{(\phi, \mu, \gamma) \in \mathcal{F}^\dagger : \gamma(C) = \gamma(D) + \eta\}$, where $\eta \equiv \frac{1-\delta}{\delta}$, as the "indifference plane" (i.e. at each HBP in $\mathcal{I}$, player 1 is indifferent between C and D). Let $\mathcal{F}_\varnothing \equiv \{w \in \mathcal{F}^\dagger : \mathcal{B}(\{w\}) = \varnothing\}$ be the set of "useless points" in $\mathcal{F}^\dagger$, meaning those which can only generate empty sets. Ignoring those points greatly simplifies the analysis.

Lemma 2.3.1. *Define the set $\bar{\mathcal{F}} \equiv \mathcal{F}^\dagger \backslash \mathcal{F}_\varnothing$. An HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ is in the set $\bar{\mathcal{F}}$ only if one of the following conditions is true:*

1. *$\tilde{\phi} = 0$ and $\tilde{\gamma}(C) - \tilde{\gamma}(D) \geq \eta$;*
2. *$\tilde{w} \in \mathcal{I}$; or*
3. *$\tilde{\phi} = 1$ and $\tilde{\gamma}(C) - \tilde{\gamma}(D) \leq \eta$.*

Proof. First, suppose $\tilde{\phi} = 0$, and let some HBA $x \equiv (\phi, \mu, \alpha)$ enforced by $\tilde{w}$ be given. Since inducibility requires $\alpha_1 = \tilde{\phi} = 0$, D must be a best response for player 1, and therefore $\tilde{\gamma}(C) - \tilde{\gamma}(D) \geq \eta$; otherwise, I have a contradiction, so $\mathcal{B}(\tilde{w}) = \varnothing$. Second, suppose $\tilde{\phi} \in (0, 1)$, again letting some enforced HBA $x \equiv (\phi, \mu, \alpha)$ be given. Inducibility requires $\alpha_1 = \tilde{\phi} \in (0, 1)$, so both C and D must be best responses for player 1, and hence

$\tilde{\gamma}(C) - \tilde{\gamma}(D) = \eta$, which means $\tilde{w} \in \mathcal{I}$. Third, suppose $\tilde{\phi} = 1$. Inducibility requires that C be a best response for player 1, so Condition 3 must be satisfied. □

Define $\bar{\mathcal{B}}(W) \equiv \mathcal{B}(W) \cap \bar{\mathcal{F}}$ and $\bar{\mathcal{D}} \equiv \mathcal{D} \cap \bar{\mathcal{F}}$. It is straightforward to adapt Proposition 2.3.4 to show $\bigcap_m \bar{\mathcal{B}}^m(\bar{\mathcal{F}}) = \bar{\mathcal{D}}$ and that $\mathcal{V}(\bar{\mathcal{D}}) = \mathcal{V}(\mathcal{D}) = \mathcal{E}$, since the HBPs in $\mathcal{F}_\varnothing$ cannot generate anything.

I present some results characterizing the set of useful HBPs that can be generated by any particular HBP. For convenience, define $g : \mathbb{R}^2 \to \Gamma$ so that $g(v_D, v_C) = \gamma$ such that $\gamma(D) = v_D, \gamma(C) = v_C$; for further brevity, let $g_{\mathcal{I}}(v_D) \equiv g(v_D, v_D + \eta)$. Since beliefs are pinned down for each history distribution and hence redundant, I often omit belief mappings in HBPs and HBAs (i.e. write "(ϕ, γ)" instead of "(ϕ, μ, γ)"). Define $\phi^* \equiv \frac{\mu^0}{1-\mu^0}$. I rewrite (2.5.4) as $q(\phi) = \frac{\phi^*}{1-2\phi}$, omitting μ^0 since it is taken as given. Define $\hat{\phi}_3$ so that $q(\hat{\phi}_3) = 1$ (this is well-defined for $\mu^0 \in (0, \frac{1}{2}]$). Also define $r : (0, \infty) \to \mathbb{R}$ as the inverse of $q(\cdot)$:

$$r(\phi) \equiv q^{-1}(\phi) = \frac{1}{2}\left(1 - \frac{\phi^*}{\phi}\right). \tag{2.3.1}$$

It is straightforward to show that $q(\cdot)$ and $r(\cdot)$ are strictly increasing for $\phi \in [0, \frac{1}{2})$ and $\phi \in (0, \infty)$, respectively.

Lemma 2.3.2. *Suppose $\mu^0 \in (0, \frac{1}{2}]$. Let any HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ be given. Then*

$$\bar{\mathcal{B}}(\tilde{w}) = \begin{cases} \{(q(\tilde{\phi}), g_{\mathcal{I}}((1-\delta) + \delta\tilde{\gamma}(D)))\} & \tilde{\phi} \in [0, \hat{\phi}_3) \\ \{(1, \gamma) : \gamma(D) = (1+\delta) + \delta\tilde{\gamma}(D), & \\ \qquad \gamma(C) \in [\gamma(D) + \eta, 2(1-\delta) + \delta\tilde{\gamma}(C)]\} & \tilde{\phi} = \hat{\phi}_3 \\ \{(1, g((1-\delta) + \delta\tilde{\gamma}(D), 3(1-\delta) + \delta\tilde{\gamma}(D)))\} & \tilde{\phi} \in (\hat{\phi}_3, \frac{1}{2}) \\ \{(\phi, \gamma) : \gamma(D) \in [(1-\delta) + \delta\tilde{\gamma}(D), 3(1-\delta) + \delta\tilde{\gamma}(D)], & \\ \qquad \gamma(C) = 3(1-\delta) + \delta\tilde{\gamma}(D)\} \cap \bar{\mathcal{F}} & \tilde{\phi} = \frac{1}{2} \\ \{(0, \gamma) : \gamma(C) = \gamma(D) = 2(1-\delta) + \delta\tilde{\gamma}(C)\} & \tilde{\phi} \in (\frac{1}{2}, 1]. \end{cases}$$

Proof. I consider each case sequentially. Let some HBA $x \equiv (\phi, \mu, \alpha)$ enforced by $\tilde{w}$ and the HBP $w \equiv (\phi, \mu, \gamma)$ decomposed by x and $\tilde{w}$ be given.

First, suppose $\tilde{\phi} \in [0, \hat{\phi}_3)$. Inducibility requires that $\alpha_1 = \tilde{\phi} < \frac{1}{2}$, so player 2's payoff for playing c at history D is

$$u_2(c, \alpha_1(\theta_0)) = 3\tilde{\phi} + 0 \cdot (1 - \tilde{\phi}) < 2\tilde{\phi} + 1 \cdot (1 - \tilde{\phi}) = u_2(d, \alpha_1(\theta_0));$$

thus, d is strict best response, so $\alpha_2(D) = 0$ and $\gamma(D) = (1 - \delta) + \delta\tilde{\gamma}(D)$. At history C, player 2 is indifferent if

$$E[u_2(c, \alpha_1)|C] = 3\mu + (1-\mu)[\tilde{\phi} \cdot 3 + (1-\tilde{\phi}) \cdot 0] = E[u_2(d, \alpha_1)|C] = 2\mu + (1-\mu)[\tilde{\phi} \cdot 2 + (1-\tilde{\phi}) \cdot 1]$$

$$\mu[3 - 2] + (1 - \mu)[3\tilde{\phi} - 2\tilde{\phi} - (1 - \tilde{\phi})] = \mu + (1 - \mu)[2\tilde{\phi} - 1)] = 0$$

$$\mu[2 - 2\tilde{\phi}] = 1 - 2\tilde{\phi} \tag{2.3.2}$$

Substituting Bayes' rule for μ gives

$$\frac{\mu^0}{\mu^0 + (1 - \mu^0)\phi}[2 - 2\tilde{\phi}] = 1 - 2\tilde{\phi}$$

$$\mu^0[2 - 2\tilde{\phi}] = (1 - 2\tilde{\phi})[\mu^0 + (1 - \mu^0)\phi]$$

$$\mu^0[2 - 2\tilde{\phi} - 1 + 2\tilde{\phi}] = (1 - 2\tilde{\phi})(1 - \mu^0)\phi$$

$$\frac{\mu^0}{1 - \mu^0} = (1 - 2\tilde{\phi})\phi$$

Substituting $\phi^* \equiv \frac{\mu^0}{1-\mu^0}$ gives

$$\phi = \frac{\frac{\mu^0}{1-\mu^0}}{1 - 2\tilde{\phi}} = \frac{\phi^*}{1 - 2\tilde{\phi}} = q(\tilde{\phi}).$$

Thus, if $\phi > q(\tilde{\phi})$, player 2 strictly prefers d, and if the inequality is reversed strictly prefers c. Hence, if $\phi > q(\tilde{\phi})$, $\gamma(C) = \gamma(D) = (1 - \delta) + \delta\tilde{\gamma}(D) < \gamma(D) + \eta$, and thus $w \in \mathcal{F}_\varnothing$ (note that $q(\tilde{\phi}) > \tilde{\phi} \geq 0$). Similarly, if $\phi < q(\tilde{\phi})$, $\gamma(C) = 2(1-\delta) + \delta\tilde{\gamma}(C) \geq 2(1-\delta) + \delta[\tilde{\gamma}(D) + \eta] = 3(1 - \delta) + \delta\tilde{\gamma}(D) > \gamma(D) + \eta$, so $w \in \mathcal{F}_\varnothing$ (note that $q(\tilde{\phi}) < 1$). Thus, $w \in \bar{\mathcal{F}}$ only if $\phi = q(\tilde{\phi})$ and $\gamma(C) = \gamma(D) + \eta$.

Suppose $\tilde{\phi} = \hat{\phi}_3$. By the same argument as the previous paragraph, $\gamma(D) = (1 -$

$\delta) + \delta\tilde{\gamma}(D)$. $q(\hat{\phi}_3) = 1$, so if $\phi < 1$ then $\gamma(C) = 2(1-\delta) + \delta\tilde{\gamma}(C) > \gamma(D) + \eta$ and again, $w \in \mathcal{F}_\varnothing$. If $\phi = 1$, then player 2 is indifferent so any $\alpha_2(C) \in [0,1]$ is enforceable, but if I want $w \in \bar{\mathcal{F}}$ I must choose $\alpha_2(C)$ sufficiently high such that $\gamma(C) \geq \gamma(D) + \eta$; thus, $\gamma(C) \in [\gamma(D) + \eta, 2(1-\delta) + \delta\tilde{\gamma}(C)]$.

Suppose $\tilde{\phi} \in (\hat{\phi}_3, \frac{1}{2})$. As before $\gamma(D) = (1-\delta) + \delta\tilde{\gamma}(D)$. Furthermore, player 2 strictly prefers c at history C, so $\gamma(C) = 2(1-\delta) + \delta\tilde{\gamma}(C) > \gamma(D) + \eta$. Hence if $w \in \bar{\mathcal{F}}$, $\phi = 1$.

Suppose $\tilde{\phi} = \frac{1}{2}$. Then player 2 strictly prefers c at history C so $\gamma(C) = 2(1-\delta) + \delta\tilde{\gamma}(C)$, but is indifferent at history D. Thus, any $\alpha_2(D)$ is enforceable, giving $\gamma(D) \in [(1-\delta) + \delta\tilde{\gamma}(D), 3(1-\delta) + \delta\tilde{\gamma}(D)]$.

Finally, suppose $\tilde{\phi} \in (\frac{1}{2}, 1]$. Player 2 strictly prefers c at either history, so $\gamma(D) = \gamma(C) = 2(1-\delta) + \delta\tilde{\gamma}(C)$, and $w \in \bar{\mathcal{F}}$ implies $\phi = 0$. □

Define $\mathcal{W}_\phi \equiv \{(\phi, \gamma) \in \bar{\mathcal{F}}\}$ as the set of "useful" HBPs with history distribution ϕ. The following corollary restates some of the results of Lemma 2.3.2 in a more directly useful way.

Corollary 2.3.1. *Suppose $\mu^0 \in (0, \frac{1}{2})$. Let any HBP $w \equiv (\phi, \gamma) \in \mathcal{I}$ be given. Suppose there exists $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ such that $w \in \bar{\mathcal{B}}(\{\tilde{w}\})$. Then either $\phi = q(\tilde{\phi})$ or $\tilde{\phi} = \frac{1}{2}$. Conversely, for any $\phi' \in q(0)$, any HBP $\tilde{w}' \equiv (r(\phi'), \tilde{\gamma}') \in \mathcal{W}_{r(\phi')}$ generates an HBP in $\mathcal{W}_{\phi'}$: $(\phi', g_{\mathcal{I}}((1-\delta) + \delta\tilde{\gamma}'(D)) \in \mathcal{B}(\tilde{w}')$. Also, for any $\phi' \in \Phi$, any HBP $\tilde{w}' \equiv (\frac{1}{2}, \tilde{\gamma}') \in \mathcal{W}_{\frac{1}{2}}$ generates an HBP in $\mathcal{W}_{\phi'}$: $(\phi', g_{\mathcal{I}}(3(1-\delta) + \delta\tilde{\gamma}(D) - \eta)) \in \mathcal{B}(\tilde{w}')$. For $\phi' \in (0,1)$, the sets $\mathcal{B}(\tilde{w}')$ above are singletons.*

Define the minimum and maximum "relevant" payoffs[3] at each history distribution following the mth iteration of the $\bar{\mathcal{B}}(\cdot)$ operator (recall that the minima and maxima exist

[3]By relevant, I mean the following for a given $(\phi, \gamma) \in \bar{\mathcal{F}}$. If $\phi = 0$, $\gamma(C)$ does not affect what can be generated; if $\phi = 1$, $\gamma(D)$ does not affect what can be generated; if $\phi \in (0,1)$, $\gamma(C) = \gamma(D) + \eta$ so player 1 is indifferent and either $\gamma(C)$ or $\gamma(D)$ is "binding."

because $\bar{\mathcal{B}}^m(\mathcal{F}^\dagger)$ is closed by Lemma 2.1.2):

$$\underline{v}_\phi^m \equiv \begin{cases} \min\{\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi \in [0,1) \\ \min\{\gamma(C) : (1,\gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi = 1 \end{cases}$$

$$\bar{v}_\phi^m \equiv \begin{cases} \max\{\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi \in [0,1) \\ \max\{\gamma(C) : (1,\gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi = 1. \end{cases}$$

Also, for all $\phi \in \Phi$, define $\underline{v}_\phi^\infty \equiv \lim_{m\to\infty} \underline{v}_\phi^m, \bar{v}_\phi^\infty \equiv \lim_{m\to\infty} \bar{v}_\phi^m$. Note that $\underline{v}_{\tilde{\phi}}^0 = 0, \bar{v}_{\tilde{\phi}}^0 = 3-\eta$ and $\underline{v}_1^0 = \eta, \bar{v}_0^0 = 3$ for all $\tilde{\phi} \in (0,1]$. Note that because $\{\bar{\mathcal{B}}^m(\bar{\mathcal{F}})\}_m$ is a decreasing sequence (in the sense of set inclusion), $\{\underline{v}_\phi^m\}_m$ is weakly increasing and bounded, and so it follows that

$$\underline{v}_\phi^\infty = \begin{cases} \min\{\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{D}}\} & \phi \in [0,1) \\ \min\{\gamma(D) : (1,\gamma) \in \bar{\mathcal{D}}\} & \phi = 1. \end{cases}$$

The following result characterizes the output of the $\mathcal{V}(\cdot)$ operator.

Lemma 2.3.3. *Let $\mu^0 \in (0, \frac{1}{2}]$ and $w \equiv (\phi,\gamma) \in \bar{\mathcal{D}}(\delta,\mu^0)$ be given. Then*

$$\mathcal{V}(\{(\phi,\gamma)\}) = \begin{cases} \{(1-\delta)+\delta\gamma(D)\} & \phi < \hat{\phi}_3 \\ [(1-\delta)+\delta\gamma(D), 3(1-\delta)+\delta\gamma(D)] & \phi = \hat{\phi}_3 \\ \{2(1-\delta)+\delta\gamma(C)\} & \phi > \hat{\phi}_3. \end{cases}$$

If $\mu^0 > \frac{1}{2}$, then

$$\mathcal{V}(\{(\phi,\gamma)\}) = \begin{cases} 3(1-\delta)+\delta\gamma(D) & \phi = 0 \\ 2(1-\delta)+\delta\gamma(C) & \phi > 0. \end{cases} \tag{2.3.3}$$

Proof. For any $v \in \mathcal{V}(\{(\phi,\gamma)\})$, there exists a PBE (σ^*,μ^*) of the γ-antegame such that $V(\sigma^*) = v$ and the distribution of outcomes matches ϕ: $\sigma_1^*(\varnothing) = \phi$ (recall that player 1 does not condition on player 2's action due to quasi-Markov perfection).

First consider the $\mu^0 \in (0, \frac{1}{2}]$ case. Suppose $\phi < \hat{\phi}_3$, and suppose by contradiction that player 2's expected payoff at the empty history for playing c is weakly greater than the payoff for d:

$$E[u_2(c,\sigma^*(\varnothing))|\varnothing] = 3\mu^0 + (1-\mu^0)[3\phi + 0\cdot(1-\phi)]$$
$$\geq E[u_2(d,\sigma^*(\varnothing))|\varnothing] = 2\mu^0 + (1-\mu^0)[\phi\cdot 2 + (1-\phi)\cdot 1] \tag{2.3.4}$$

$$\mu^0 + (1-\mu^0)[3\phi] \geq (1-\mu^0)[1+\phi] \tag{2.3.5}$$

$$\mu^0 \geq (1-\mu^0)[1-2\phi] \tag{2.3.6}$$

$$\frac{\mu^0}{1-\mu^0} = \phi^* \geq 1-2\phi \tag{2.3.7}$$

$$\frac{\phi^*}{1-2\phi} = q(\phi) \geq 1, \tag{2.3.8}$$

which is a contradiction because $\phi < \hat{\phi}_3$ implies $q(\phi) < q(\hat{\phi}_3) = 1$, due to the monotonicity of $q(\phi')$. Thus, player 2 strictly prefers d. Since $(\phi,\gamma) \in \bar{\mathcal{D}}$ and $\phi < 1$, $\gamma(D) \geq \gamma(C) - \eta$ and so D is a best response at period 0 for player 1. Thus, $V(\sigma^*) = (1-\delta)u_1(d,D) + \delta\gamma(D) = (1-\delta) + \delta\gamma(D)$.

Suppose $\phi = \hat{\phi}_3$. Then $q(\phi) = \frac{\phi^*}{1-2\phi} = 1$ so replacing the inequality in (2.3.8) with "=" and working backwards, doing the same to each equation until (2.3.4), shows that player 2 is indifferent between c and d. For player 1, D is a best response since $\hat{\phi}_3 < 1$. Thus, I can pick any $\sigma_2^*(\varnothing) \in [0,1]$, so $\mathcal{V}(\{(\phi,\gamma)\}) = \{V(\sigma^*) : \sigma_2^*(\varnothing) \in [0,1], \sigma_1^*(\varnothing) = \hat{\phi}_3\} = [0\cdot(1-\delta) + \delta\gamma(C), 2(1-\delta) + \delta\gamma(C)]$.

Suppose $\phi > \hat{\phi}_3$. Then $q(\phi) < 1$, and performing the analogous steps in the previous paragraph shows that player 2 strictly prefers c to d. Since $\sigma_1^*(\varnothing) = \phi > \hat{\phi}_3 \geq 0$, C is a best response, so $V(\sigma^*) = (1-\delta)u_1(c,C) + \delta\gamma(C) = 2(1-\delta) + \delta\gamma(C)$.

Now I turn to the simpler $\mu^0 > \frac{1}{2}$ case. Player 2 strictly prefers c no matter what player 1's strategy is. If $\phi = 0$, then D is a best response, so $V(\sigma^*) = (1-\delta)u_1(c,D) + \delta\gamma(D) = 3(1-\delta) + \delta\gamma(D)$. If $\phi > 0$, then C is a best response: $V(\sigma^*) = (1-\delta)u_1(c,C) + \delta\gamma(C) = 2(1-\delta) + \delta\gamma(C)$. □

Now consider the first case listed in (2.5.3): suppose $\mu^0 \in (0, \frac{1}{9}]$. The function q has two fixed points for $\mu^0 < \frac{1}{9}$, one fixed point for $\mu^0 = \frac{1}{9}$, and no fixed points for $\mu^0 > \frac{1}{9}$ (this turns out to be the reason why $\mu^0 > \frac{1}{9}$ has only one stationary equilibrium instead of two). Setting $\phi = q(\phi)$ and substituting into (2.5.4) yields

$$\phi = \frac{\phi^*}{1 - 2\phi}$$

which can be rearranged to get

$$0 = 2\phi^2 - \phi + \phi^*.$$

The quadratic formula gives

$$\phi = \frac{1 \pm \sqrt{1 - 4 \cdot 2\phi^*}}{4} = \frac{1}{4} \pm \frac{\sqrt{1 - 8\phi^*}}{4}; \tag{2.3.9}$$

since $\phi^* = \frac{1}{8}$ for $\mu^0 = \frac{1}{9}$, $\mu^0 \leq \frac{1}{9}$ is sufficient and necessary for the discriminant to be non-negative, and therefore for the existence of a fixed point. Define $\hat{\phi}_1, \hat{\phi}_2 \in \Phi$ as the fixed points of $q(\cdot)$ such that $\hat{\phi}_1 < \hat{\phi}_2$ for $\mu^0 < \frac{1}{9}$ and $\hat{\phi}_1 = \hat{\phi}_2$ for $\mu^0 = \frac{1}{9}$.

Lemma 2.3.4. *Suppose $\mu^0 \in (0, \frac{1}{9}]$. For any $\phi \in [\hat{\phi}_1, 1)$, $\underline{v}_\phi^\infty = 1$. Furthermore, $\underline{v}_1^\infty = 1/\delta$, and $\underline{v}_{\phi'}^\infty \geq 1$ for all $\phi' \in [0, \hat{\phi}_1)$. Finally, $\bar{v}_\phi^\infty = 2 - \eta$ for all $\phi \in [\hat{\phi}_1, 1)$. Furthermore, $\bar{v}_1^\infty = 2$ and $\bar{v}_{\phi'}^\infty \leq 2$ for all $\phi' \in [0, \hat{\phi}_1)$.*

Proof. I begin by recalling that $r(\cdot)$ is strictly increasing and noting that for any $\phi \in [\hat{\phi}_1, 1)$, the following holds:

$$r(\phi) \begin{cases} = \phi & \phi \in \{\hat{\phi}_1, \hat{\phi}_2\} \\ > \phi & \phi \in (\hat{\phi}_1, \hat{\phi}_2) \\ < \phi & \phi \in (\hat{\phi}_1, 1). \end{cases} \tag{2.3.10}$$

I show that for every $\phi' \in [\hat{\phi}_1, 1)$, $r(\phi') \in [\hat{\phi}_1, 1)$ and hence $r(\phi') \geq q(0)$, satisfying the corresponding condition in Corollary 2.3.1 that used below. First, (2.3.10) immediately shows this to be true for $\phi \in \{\hat{\phi}_1, \hat{\phi}_2\}$. Suppose that $\hat{\phi}_1 < \phi < \hat{\phi}_2$; the monotonicity of $r(\cdot)$ and (2.3.10) imply $r(\hat{\phi}_1) = \hat{\phi}_1 < \phi < r(\phi) < r(\hat{\phi}_2) = \hat{\phi}_2$. Finally, suppose that $\hat{\phi}_2 < \phi < 1$; again the monotonicity of $r(\cdot)$ and (2.3.10) imply $r(\hat{\phi}_2) = \hat{\phi}_2 < r(\phi) < \phi < 1$.

Building on the previous paragraph, Corollary 2.3.1 implies that for every $\phi \in [\hat{\phi}_1, 1)$,

$$\begin{aligned}
\underline{v}_\phi^{m+1} &= \min\{\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{B}}^{m+1}(\bar{\mathcal{F}})\} \\
&= \min\{\min\{\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{B}}(\bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \mathcal{W}_{r(\phi_k)})\}, \\
&\qquad \min\{\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{B}}(\bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \mathcal{W}_{\frac{1}{2}})\}\} \\
&= \min\{\min\{(1-\delta) + \delta\gamma(D) : (r(\phi),\gamma) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}))\}, \\
&\qquad \min\{3(1-\delta) + \delta\gamma(D) - \eta : (\tfrac{1}{2},\gamma) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}))\}\} \\
&= \min\{(1-\delta) + \delta\underline{v}_{r(\phi)}^m, 3(1-\delta) + \delta\underline{v}_{\frac{1}{2}}^m - \eta\}.
\end{aligned}$$

Note that $\underline{v}_\phi^0 = 0$ for all $\phi \in [0,1)$. Suppose that for some m and some $\underline{\zeta}^m \in \mathbb{R}$, $\underline{v}_\phi^m = \underline{v}_{\frac{1}{2}}^m = \underline{\zeta}^m$ for all $\phi \in [\hat{\phi}_1, 1)$. Then

$$\begin{aligned}
\underline{v}_\phi^{m+1} &= \min\{(1-\delta) + \delta\underline{v}_{r(\phi)}^m, 3(1-\delta) + \delta\underline{v}_{\frac{1}{2}}^m - \eta\} \\
&= \min\{(1-\delta) + \delta\underline{\zeta}^m, 3(1-\delta) + \delta\underline{\zeta}^m - \eta\} \\
&= (1-\delta) + \delta\underline{\zeta}^m.
\end{aligned}$$

By induction, for each m, there exists $\underline{\zeta}^{m+1} = (1-\delta) + \delta\underline{\zeta}^m$ such that $\underline{v}_\phi^m = \underline{\zeta}^m$ for all $\phi \in [\hat{\phi}_1, 1)$. Hence, $\underline{v}_\phi^\infty = \lim_{m\to\infty} \underline{\zeta}^m = 1$.

Similarly, note that $\bar{v}_\phi^0 = 3 - \eta$ for all $\phi \in [0,1)$. Suppose for some m, $\bar{v}_\phi^m = \bar{\zeta}^m$ for all $\phi \in [\hat{\phi}_1, 1)$. Then

$$\begin{aligned}
\bar{v}_\phi^{m+1} &= \max\{(1-\delta) + \delta\bar{v}_{r(\phi_k)}^m, 3(1-\delta) + \delta\bar{v}_{\frac{1}{2}}^m - \eta\} \\
&= \max\{(1-\delta) + \delta\bar{\zeta}^m, 3(1-\delta) + \delta\bar{\zeta}^m - \eta\} \\
&= 3(1-\delta) + \delta\bar{\zeta}^m - \eta.
\end{aligned}$$

By induction, for each m, there exists $\bar{\zeta}^{m+1} = 3(1-\delta) + \delta\bar{\zeta}^m - \eta$ such that $\bar{v}_\phi^m = \bar{\zeta}^m$ for all $\phi \in [\hat{\phi}_1, 1)$. Hence, $\bar{v}_\phi^\infty = \lim_{m\to\infty} \bar{\zeta}^{m+1} = 3(1-\delta) + \delta \lim_{m\to\infty} \bar{\zeta}^m - \eta$ which can be solved to get $\bar{v}_\phi^\infty = \lim_{m\to\infty} \bar{\zeta}^m = 3 - \frac{1}{\delta} = 2 - \eta$.

Next, Lemma 2.3.2 and the above result give

$$\begin{aligned}
\underline{v}_1^{m+1} &= \min\{\gamma(C) : (1,\gamma) \in \bar{\mathcal{B}}^{m+1}(\bar{\mathcal{F}})\} \\
&= \min\left\{\gamma(C) : (1,\gamma) \in \bar{\mathcal{B}}\left(\bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \left(\bigcup_{\tilde{\phi}\in[\hat{\phi}_3,\frac{1}{2}]} \mathcal{W}_{\tilde{\phi}}\right)\right)\right\} \\
&= \min\left\{ \min\{(1-\delta) + \delta\tilde{\gamma}(D) + \eta : (\hat{\phi}_3, \tilde{\gamma}) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \mathcal{W}_{\hat{\phi}_3}\}, \right. \\
&\qquad \left. \min\{3(1-\delta) + \delta\tilde{\gamma}(D) : (\tilde{\phi},\tilde{\gamma}) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \left(\bigcup_{\tilde{\phi}\in(\hat{\phi}_3,\frac{1}{2}]} \mathcal{W}_{\tilde{\phi}}\right) \right\} \\
&= \min\{(1-\delta) + \delta\underline{\zeta}^m + \eta, 3(1-\delta) + \delta\underline{\zeta}^m\} = (1-\delta) + \delta\underline{\zeta}^m + \eta,
\end{aligned}$$

so

$$\underline{v}_1^\infty = \lim_{m\to\infty} \underline{v}_1^{m+1} = (1-\delta) + \frac{1-\delta}{\delta} + \delta \lim_{m\to\infty} \underline{\zeta}^m = (1-\delta) + \frac{1-\delta}{\delta} + \delta = \frac{1}{\delta}.$$

Analogously,

$$\begin{aligned}
\bar{v}_1^{m+1} &= \max\{3(1-\delta) + \delta\tilde{\gamma}(D) : (\tilde{\phi},\tilde{\gamma}) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}), \tilde{\phi} \in [\hat{\phi}_3, \tfrac{1}{2}]\} \\
&= \max\{3(1-\delta) + \delta\bar{v}_{\tilde{\phi}}^m, \tilde{\phi} \in [\hat{\phi}_3, \tfrac{1}{2}]\} = 3(1-\delta) + \delta\bar{\zeta}^m
\end{aligned}$$

so $\bar{v}_1^\infty = 3(1-\delta) + \delta \lim_{m\to\infty} \bar{\zeta}^m = 3(1-\delta) + \delta(2-\eta) = 2$.

Finally, let $\underline{\mathsf{v}}^m \equiv \min\{\underline{v}_{\phi'}^m : \phi' \in [0,1)\}$ and $\bar{\mathsf{v}}^m \equiv \max\{\bar{v}_{\phi'}^m : \phi' \in [0, \hat{\phi}_1)\}$. Lemma 2.3.2 shows that for any $(\phi,\gamma) \in \bar{\mathcal{B}}^{m+1}(\bar{\mathcal{F}})$,

$$\begin{aligned}
\gamma(D) &\geq \min\{(1-\delta) + \delta\tilde{\gamma}(D) : (\tilde{\phi},\tilde{\gamma}) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}})\} \\
&\geq \min\{(1-\delta) + \delta\underline{\mathsf{v}}^m, 2(1-\delta) + \delta\underline{v}_1^m\}
\end{aligned}$$

so $\underline{\mathsf{v}}^{m+1} \geq (1-\delta) + \delta\underline{\mathsf{v}}^m$ which gives

$$\begin{aligned}
\lim_{m\to\infty} \underline{\mathsf{v}}^{m+1} &\geq \lim_{m\to\infty} \min\{(1-\delta) + \delta\underline{\mathsf{v}}^m, 2(1-\delta) + \delta\underline{v}_1^m\} \\
&= \min\left\{(1-\delta) + \delta \lim_{m\to\infty} \underline{\mathsf{v}}^m, 2(1-\delta) + 1\right\},
\end{aligned}$$

so $\lim_{m\to\infty} \underline{\mathsf{v}}^m \geq 1$, which proves the last part of the lemma.[4] Turning to the maxima,

$$\bar{v}_0^{m+1} = \max\{\max\{3(1-\delta)+\delta\bar{v}_\phi^m : \phi \in [\tfrac{1}{2},1)\}, 2(1-\delta)+\delta\bar{v}_1^m\}$$

$$\bar{v}_\phi^{m+1} \leq \max\{(1-\delta)+\delta\bar{\mathsf{v}}^m, 3(1-\delta)+\delta\bar{v}_{\frac{1}{2}}^m - \eta\} \qquad \forall \phi \in (0,\hat{\phi}_1)$$

so

$$\bar{v}_0^\infty = \max\{3(1-\delta)+\delta(2-\eta), 2(1-\delta)+2\delta\} = 2$$

$$\bar{v}_\phi^{m+1} \leq \max\{(1-\delta)+\delta\bar{\mathsf{v}}^\infty, 3(1-\delta)+\delta(2-\eta)-\eta\} = \max\{(1-\delta)+\delta\bar{\mathsf{v}}^\infty, 2-\eta\} \qquad \forall \phi \in (0,\hat{\phi}_1).$$

If $(1-\delta)+\delta\bar{\mathsf{v}}^\infty > 2-\eta$, then $\bar{\mathsf{v}}^\infty = \max\{\bar{v}_0^\infty, \max\{\bar{v}_\phi^{m+1} : \phi \in (0,\hat{\phi}_1)\} = \max\{2, (1-\delta)+\delta\bar{\mathsf{v}}^\infty\}$. If $\bar{\mathsf{v}}^\infty > 2$, then $\bar{\mathsf{v}}^\infty = (1-\delta)+\delta\bar{\mathsf{v}}^\infty \implies \bar{\mathsf{v}}^\infty = 1$, a contradiction. Thus, $\bar{\mathsf{v}}^\infty = 2$. □

Lemmas 2.3.3 and 2.3.4 imply that

$$\begin{aligned}
\min \mathcal{V}(\bar{\mathcal{D}}) &= \min\{\min\{(1-\delta)+\delta\gamma(D) : (\phi,\gamma)\in\bar{\mathcal{D}}, \phi \leq \hat{\phi}^3\}, \\
&\qquad \min\{2(1-\delta)+\delta\gamma(C) : (\phi,\gamma)\in\bar{\mathcal{D}}, \phi > \hat{\phi}^3\}\} \\
&= \min\{\min\{(1-\delta)+\delta\underline{v}_\phi^\infty, \phi \leq \hat{\phi}^3\}, \min\{2(1-\delta)+\delta(\underline{v}_\phi^\infty+\eta) : \phi \in (\hat{\phi}^3,1)\}, \\
&\qquad 2(1-\delta)+\delta(\underline{v}_1^\infty-\eta)\} \\
&= \min\{(1-\delta)+\delta\cdot 1, 2(1-\delta)+\delta(1+\eta)\} = 1.
\end{aligned}$$

Thus, the first minimum listed in (2.5.3) has been proven.

Now, suppose $\mu^0 \in (\frac{1}{9},\frac{1}{2})$. This eliminates the fixed points $\hat{\phi}_1, \hat{\phi}_2$ of $q(\cdot)$ (and so it turns out the only stationary equilibrium is at $\phi = \frac{1}{2}$). I still define $\hat{\phi}_3$ so that $q(\hat{\phi}_3) = 1$.

Lemma 2.3.5. *Let $\mu^0 \in (\frac{1}{9},\frac{1}{2})$. For any $\phi \in [0,\hat{\phi}_3)$, there exists finite $\bar{k}$ such that $q^{\bar{k}}(\phi) \geq \frac{1}{2}$.*

[4] This simply confirms the obvious fact that the continuation payoff must be greater than or equal to the minmax payoff.

Proof. Let

$$\Delta q(\phi) \equiv q(\phi) - \phi = \frac{\phi^*}{1-2\phi} - \phi = \frac{\phi^* - \phi(1-2\phi)}{1-2\phi} = \frac{\phi^* - \phi + 2\phi^2}{1-2\phi}.$$

It is clear that for $\phi \in [0, \frac{1}{2})$, because if $\Delta q(\phi) \leq 0$ for some ϕ, there would exist $\phi' \in [0, \frac{1}{2})$ such that

$$0 = \phi^* - \phi + 2\phi^2$$

$$\phi = \frac{1 \pm \sqrt{1 - 8\phi^*}}{4},$$

but $\phi^* > \frac{1}{8}$ because $\mu^0 > \frac{1}{9}$. The first and second derivatives of $\Delta q(\cdot)$ are

$$\Delta q'(\phi) = 2\phi^*(1-2\phi)^{-2} - 1$$

$$\Delta q''(\phi) = 8\phi^*(1-2\phi)^{-3} > 0$$

for $\phi < \frac{1}{2}$, so $\Delta q(\cdot)$ is strictly convex. Furthermore, setting $\Delta q'(\phi) = 0$ gives

$$2\phi^*(1-2\phi)^{-2} - 1 = 0$$

$$(1-2\phi)^{-2} = \frac{1}{2\phi^*}$$

$$1 - 2\phi = \pm\sqrt{2\phi^*}$$

$$\phi = \tfrac{1}{2}(1 \pm \sqrt{2\phi^*}), \tag{2.3.11}$$

and for $\phi^* \in (\frac{1}{8}, \frac{1}{2})$, such a solution $\phi \in [0, \frac{1}{2})$ to (2.3.11) exists. Thus, $\Delta q(\phi) \geq \Delta \underline{q}$ for some $\Delta \underline{q} > 0$, so for any k,

$$q^k(\phi) = \phi + \Delta q(\phi) + \Delta q(q(\phi)) + \Delta q(q^2(\phi)) + \cdots + \Delta q(q^{k-1}(\phi)) \geq \phi + k\Delta \underline{q},$$

so there exists finite $\bar{k}$ such that $q^{\bar{k}}(\phi) \geq \frac{1}{2}$. □

Having established its finite existence with Lemma 2.3.5, let

$$L(\mu^0) \equiv \min_k q^k(0) \text{ such that } q^k(0) \geq \tfrac{1}{2}.$$

Define $p^k \equiv q^k(0)$ for $k \in \{0, ..., L\}$. Define $\mathcal{S}^k \equiv (p^k, p^{k+1})$ for $k \in \{0, ..., L-1\}$ and $\mathcal{S}^L \equiv (p^L, 1)$.

Lemma 2.3.6. *Suppose* $\mu^0 \in (\frac{1}{9}, \frac{1}{2}]$. *If* $p^L = \frac{1}{2}$, *then* $\underline{v}_{p^L}^{\infty} = \lambda_1^L$; *otherwise,* $\underline{v}_{p^{L-1}}^{\infty} = \lambda_1^{L-1}$. *Also,*

$$\underline{v}_{p^L}^{\infty} = \begin{cases} (1-\delta) + \delta \underline{v}_{p^{L-1}}^{\infty} & p^L < 1 \\ (1-\delta) + \delta \underline{v}_{p^{L-1}}^{\infty} + \eta & p^L = 1. \end{cases}$$

Finally, $\min\{\underline{v}_{\phi}^{\infty} : \phi \in [0, \hat{\phi}_3]\} = \underline{v}_{p^{L-1}}^{\infty}$ *and*

$$\min\{\underline{v}_{\phi}^{\infty} : \phi \in (\hat{\phi}_3, 1)\} = \begin{cases} \underline{v}_{p^L}^{\infty} & p^L < 1 \\ \underline{v}_{p^{L-1}}^{\infty} & p^L = 1 \end{cases} \qquad \underline{v}_1^{\infty} = \begin{cases} \underline{v}_{p^L}^{\infty} + \eta & p^L < 1 \\ \underline{v}_{p^L}^{\infty} & p^L = 1. \end{cases}$$

Proof. Note that λ_1^k above in (2.5.5) is defined so that $\lambda_1^k = (1-\delta^k) + \delta^k[3(1-\delta) + \delta\lambda_1^k - \eta]$. Also define λ_2^L so that $\lambda_2^L = (1-\delta^L) + \delta^L[3(1-\delta) + \delta\lambda_1^L]$, which gives

$$\lambda_2^L = \frac{1}{1-\delta^{L+1}}[1 - \delta^L + 3\delta^L(1-\delta)].$$

For brevity in this proof, unlike that of Lemma 2.3.4 I skip ahead to the minimum generating limits, e.g. instead of writing "$\underline{v}_{\mathcal{S}^0}^{m+1} = 3(1-\delta) + \delta \underline{v}_{\frac{1}{2}}^{m} - \eta$," I take the limit as $m \to \infty$ as shown in (2.3.12) below.

First, by Corollary 2.3.1,

$$\underline{v}_{\mathcal{S}^0}^{\infty} = 3(1-\delta) + \delta \underline{v}_{\frac{1}{2}}^{\infty} - \eta; \tag{2.3.12}$$

to see why, note that $\mathcal{W}_\phi \subset \mathcal{I}$ and $r(\phi) < 0 = p^0 = r(p^1)$ for any $\phi \in \mathcal{S}^0$, so there does not exist ϕ' such that $q(\phi') = \phi$. For each $k \in \{1, ..., L\}$, for any $\phi \in \mathcal{S}^k$, $r(\phi) \in \mathcal{S}^{k-1}$. To see this, recall that $r(\cdot)$ is monotonic and that $p^k < \phi < p^{k+1}$ for $k < L$ (the argument here is easy to adapt for $k = L$), so $r(p^k) = p^{k-1} < r(\phi) < p^k = r(p^{k+1})$. Corollary 2.3.1 shows that $\underline{v}_{\mathcal{S}^k}^{\infty} = \min\{(1-\delta) + \delta \underline{v}_{\mathcal{S}^{k-1}}^{\infty}, 3(1-\delta) + \delta \underline{v}_{\frac{1}{2}}^{\infty} - \eta\}$. Suppose (by contradiction) that $3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}-\eta<(1-\delta)+\delta\underline{v}_{\mathcal{S}^{k-1}}^{\infty}$. Then

$$\begin{aligned} 3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}-\eta &< (1-\delta)+\delta\underline{v}_{\mathcal{S}^{k-1}}^{\infty}\leq(1-\delta)+\delta[3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}-\eta] \\ &= \delta[3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}]. \end{aligned}$$

$$(1-\delta)[3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}]<\eta=\frac{1-\delta}{\delta}$$

$$3\delta(1-\delta)+\delta^{2}\underline{v}_{\frac{1}{2}}^{\infty}<1 \qquad (2.3.13)$$

For $\delta>\frac{1}{2}$, the left hand side of (2.3.13) is less than $\frac{3}{4}+\frac{1}{4}\underline{v}_{\frac{1}{2}}^{\infty}<1 \implies \underline{v}_{\frac{1}{2}}^{m}<1$, which is easily shown to be contradicted by Lemma 2.3.2.[5] Thus, $\underline{v}_{\mathcal{S}^k}^{\infty}=(1-\delta)+\delta\underline{v}_{\mathcal{S}^{k-1}}^{\infty}$, and in fact the argument above shows that for any $\phi^{k-1}\in\mathcal{S}^{k-1}$, $\phi^k\in\mathcal{S}^k$, I have $\underline{v}_{\phi^k}^{\infty}=(1-\delta)+\delta\underline{v}_{\phi^{k-1}}^{\infty}$. An almost identical argument shows that $\underline{v}_{p^k}^{\infty}=(1-\delta)+\delta\underline{v}_{p^{k-1}}^{\infty}$. Note that because $\underline{v}_{p^0}^{\infty}\leq 3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}-\eta=\underline{v}_{\mathcal{S}^0}^{\infty}$, $\underline{v}_{p^k}^{\infty}\leq\underline{v}_{\mathcal{S}^k}^{\infty}$ for each k.

Suppose $p^L<1$. I show that $\underline{v}_{1}^{\infty}=\underline{v}_{\mathcal{S}^L}^{\infty}+\eta$. Note that $p^{L-1}=r(p^L)<r(1)=\hat{\phi}_3<\frac{1}{2}\leq p^L$, so $[\hat{\phi}_3,\frac{1}{2})\subset\mathcal{S}^{L-1}$. By Lemma 2.3.2,

$$\begin{aligned} \underline{v}_{1}^{\infty} &= \min\{(1-\delta)+\delta\underline{v}_{\hat{\phi}_3}^{\infty}+\eta,\min\{3(1-\delta)+\delta\underline{v}_{\phi}^{\infty}:\phi\in(\hat{\phi}_3,\tfrac{1}{2})\},3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}\} \\ &= \min\{(1-\delta)+\delta\underline{v}_{\mathcal{S}^{L-1}}^{\infty}+\eta,3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}\} \qquad (2.3.14) \end{aligned}$$

If $p^L>\frac{1}{2}$, then $\frac{1}{2}\in\mathcal{S}^{L-1}$ and $\underline{v}_{\frac{1}{2}}^{\infty}=\underline{v}_{\mathcal{S}^{L-1}}^{\infty}$, so

$$\underline{v}_{1}^{\infty}=(1-\delta)+\delta\underline{v}_{\mathcal{S}^{L-1}}^{\infty}+\eta=\underline{v}_{\mathcal{S}^L}^{\infty}+\eta. \qquad (2.3.15)$$

The next paragraph proves the result for $p^L=\frac{1}{2}$.

Suppose $p^L=\frac{1}{2}$. By Lemma 2.3.2,

$$\underline{v}_{p^0}^{\infty} = \min\{3(1-\delta)+\delta\underline{v}_{\frac{1}{2}}^{\infty}-\eta,\min\{3(1-\delta)+\delta\underline{v}_{\phi}:\phi\in(\tfrac{1}{2},1)\},2(1-\delta)+\delta\underline{v}_{1}^{\infty}\} \qquad (2.3.16)$$

[5] Besides the fact that this implies a continuation payoff less than the minmax, this can be seen by simply letting $\underline{\mathrm{v}}^m\equiv\min\{\min\{\underline{v}_{\phi}^{m}:\phi\in[0,1)\},\underline{v}_{1}^{m}\}$ and generating the lowest payoff given by Lemma 2.3.2 at any ϕ as the new lower bound $\underline{\mathrm{v}}^{m+1}\geq(1-\delta)+\delta\underline{\mathrm{v}}^{m}$, which converges to 1 as $m\to\infty$.

$$= \quad \min\{3(1-\delta)+\delta\underline{v}^{\infty}_{\frac{1}{2}}-\eta, 2(1-\delta)+\delta\underline{v}^{\infty}_{1}\},$$

where the last step is because $(\frac{1}{2},1)=\mathcal{S}^L$ and $\underline{v}^{\infty}_{\mathcal{S}^L}\geq\underline{v}^{\infty}_{p^L}=\frac{1}{2}$. Suppose by contradiction that

$$\underline{v}^{\infty}_{1}=3(1-\delta)+\delta\underline{v}^{\infty}_{\frac{1}{2}}<(1-\delta)+\delta\underline{v}^{\infty}_{\mathcal{S}^{L-1}}+\eta \tag{2.3.17}$$

(see (2.3.14)). It is straightforward to show that

$$2(1-\delta)+\delta\underline{v}^{\infty}_{1}=2(1-\delta)+\delta[3(1-\delta)+\delta\underline{v}^{\infty}_{\frac{1}{2}}]\geq 3(1-\delta)+\delta\underline{v}^{\infty}_{\frac{1}{2}}-\eta$$

since $\underline{v}^{\infty}_{\frac{1}{2}}\leq 2-\eta$ by Corollary 2.3.1, so $\underline{v}^{\infty}_{p^0}=3(1-\delta)+\delta\underline{v}^{\infty}_{\frac{1}{2}}-\eta=\underline{v}^{\infty}_{\mathcal{S}^0}$. By induction that means $\underline{v}^{\infty}_{\mathcal{S}^k}=\underline{v}^{\infty}_{p^k}$ for each k, and

$$\underline{v}^{\infty}_{\frac{1}{2}}=\underline{v}^{\infty}_{p^L}=(1-\delta^L)+\delta^L\underline{v}^{\infty}_{p^0}=(1-\delta^L)+\delta^L[3(1-\delta)+\delta\underline{v}^{\infty}_{p^L}-\eta]=\lambda^L_1,$$

where the last step is because the above equation matches the characterization for λ^L_1 at the beginning of the proof. Returning to (2.3.17) I have

$$3(1-\delta)+\delta\underline{v}^{\infty}_{\frac{1}{2}}=3(1-\delta)+\delta\underline{v}^{\infty}_{p^L}<(1-\delta)+\delta\underline{v}^{\infty}_{\mathcal{S}^{L-1}}+\eta=(1-\delta)+\delta\underline{v}^{\infty}_{p^{L-1}}+\eta=\underline{v}^{\infty}_{p^L}+\eta$$

$$3(1-\delta)-\eta<(1-\delta)\underline{v}^{\infty}_{p^L}$$

$$3-\frac{1}{\delta}=2-\eta<\underline{v}^{\infty}_{p^L},$$

which is a contradiction because $\underline{v}^{\infty}_{p^L}=\underline{v}^{\infty}_{\frac{1}{2}}\leq\bar{v}^{\infty}_{\frac{1}{2}}\leq 2-\eta$ by Corollary 2.3.1. Thus, to summarize, for $p^L=\frac{1}{2}$, I now also have $\underline{v}^{\infty}_{1}=(1-\delta)+\delta\underline{v}^{\infty}_{\mathcal{S}^{L-1}}+\eta=\underline{v}^{\infty}_{\mathcal{S}^L}+\eta$, and furthermore $\underline{v}^{\infty}_{\mathcal{S}^k}=\underline{v}^{\infty}_{p^k}$ for all $k\in\{0,...,L\}$ and $\underline{v}^{\infty}_{p^L}=\lambda^L_1$. It is then straightforward to check that the lemma has been proven for $p^L=\frac{1}{2}$.

For the rest of the proof suppose $\frac{1}{2}<p^L<1$. The $p^L=1$ case follows almost exactly the same argument, adding or subtracting "η" as appropriate where "$\underline{v}^{\infty}_{p^L}$" is mentioned and ignoring the consequently empty set $\mathcal{S}^L$. By Lemma 2.3.2, the first step of (2.3.16) holds here as well. Suppose (by contradiction) that

$$\underline{v}_{p^0}^{\infty} < 3(1-\delta) + \delta \underline{v}_{\frac{1}{2}}^{\infty} - \eta. \tag{2.3.18}$$

Then by Lemma 2.3.2 and (2.3.15) it must be that $\underline{v}_{p^0}^{\infty} = 3(1-\delta) + \delta \underline{v}_{p^L}^{\infty}$. Since $\underline{v}_{p^L}^{\infty} = (1-\delta^L) + \delta^L \underline{v}_{p^0}^{\infty} = (1-\delta^L) + \delta^L[3(1-\delta) + \delta \underline{v}_{p^L}^{\infty}]$, it then matches the characterization of λ_2^L at the beginning of the proof: $\underline{v}_{p^L}^{\infty} = \lambda_2^L$. Since $\frac{1}{2} \subset \mathcal{S}^{k-1}$ as shown near (2.3.14),

$$\underline{v}_{\frac{1}{2}}^{\infty} = \underline{v}_{\mathcal{S}^{L-1}}^{\infty} = (1-\delta^{L-1}) + \delta^{L-1}\underline{v}_{\mathcal{S}^0}^{\infty} = (1-\delta^{L-1}) + \delta^{L-1}[3(1-\delta) + \delta \underline{v}_{\mathcal{S}^{L-1}}^{\infty} - \eta]$$

due to (2.3.12); therefore $\underline{v}_{\frac{1}{2}}^{\infty}$ matches the characterization of λ_1^{L-1} above: $\underline{v}_{\frac{1}{2}}^{\infty} = \lambda_1^{L-1}$. Thus, I can write (2.3.18) as

$$\underline{v}_{p^0}^{\infty} = 3(1-\delta) + \delta \underline{v}_{p^L}^{\infty} = 3(1-\delta) + \delta \lambda_2^L < 3(1-\delta) + \delta \underline{v}_{\frac{1}{2}}^{\infty} - \eta = 3(1-\delta) + \delta \lambda_1^{L-1} - \eta$$

$$\lambda_2^L < \lambda_1^{L-1} - \frac{1-\delta}{\delta^2}. \tag{2.3.19}$$

The following lemma proves that (2.3.19) is a contradiction.

Lemma 2.3.7. $\lambda_2^L \geq \lambda_1^{L-1} - \frac{1-\delta}{\delta^2}$.

Proof. Suppose by contradiction the opposite:

$$\frac{\delta^2}{1-\delta^{L+1}}[1-\delta^L + 3\delta^L(1-\delta)] < \frac{\delta^2}{1-\delta^L}[(1-\delta^{L-1}) + 3\delta^{L-1}(1-\delta) - \delta^{L-1}\eta] - (1-\delta).$$

I spare the reader the tedious algebra that yields

$$1 - 3\delta^{L+1} + 2\delta^{L+2} < 0. \tag{2.3.20}$$

Note that at $\delta = 1$ the left hand side is equal to 0. Taking the derivative of the left hand side with respect to δ gives

$$-3(L+1)\delta^L + 2(L+2)\delta^{L+1} = -(3L+3)\delta^L + (2L+4)\delta^{L+1}.$$

Since $3L+3 \geq 2L+4$, the derivative is strictly negative for $\delta \in (0,1)$, so the left hand side of (2.3.20) is strictly positive for $\delta \in (0,1)$, a contradiction. $\square$

Thus, going back to (2.3.16) I have $\underline{v}^{\infty}_{p^0} = 3(1-\delta) + \delta \underline{v}^{\infty}_{\frac{1}{2}} - \eta = \underline{v}^{\infty}_{\mathcal{S}^0}$. This then implies that $\underline{v}^{\infty}_{p^k} = \underline{v}^{\infty}_{\mathcal{S}^k}$ for each $k \in \{0, ..., L\}$. Then $\underline{v}^{\infty}_{p^0} = 3(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}} - \eta$, and some rearrangement gives $\underline{v}^{\infty}_{p^{L-1}} = \lambda_2^{L-1}$. Since $\underline{v}^{\infty}_{p^0} = \underline{v}^{\infty}_{\mathcal{S}^0} > \underline{v}^{\infty}_{p^1} = \underline{v}^{\infty}_{\mathcal{S}^1} > \cdots > \underline{v}^{\infty}_{p^{L-1}} = \underline{v}^{\infty}_{\mathcal{S}^{L-1}}$, $\min\{\underline{v}^{\infty}_{\phi} : \phi \in [0, \hat{\phi}_3]\} = \underline{v}^{\infty}_{p^{L-1}}$. It is also easy to see that $\min\{\underline{v}^{\infty}_{\phi} : \phi \in (\hat{\phi}_3, 1)\} = \underline{v}^{\infty}_{p^L} = \underline{v}^{\infty}_{\mathcal{S}^L}$. Since I showed earlier that $\underline{v}^{\infty}_1 = \underline{v}^{\infty}_{p^L} + \eta$, the lemma is proven for $p^L < 1$. $\square$

By Lemmas 2.3.3 and 2.3.6, if $p^L < 1$ I have

$$\begin{aligned}
\min \mathcal{V}(\bar{\mathcal{D}}) &= \min\{\min\{(1-\delta) + \delta \underline{v}^{\infty}_{\phi} : \phi \leq \hat{\phi}_3\}, \\
&\qquad \min\{2(1-\delta) + \delta(\underline{v}^{\infty}_{\phi} + \eta) : \phi \in (\hat{\phi}_3, 1)\}, 2(1-\delta) + \delta \underline{v}^{\infty}_1\} \quad (2.3.21) \\
&= \min\{(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}}, 2(1-\delta) + \delta(\underline{v}^{\infty}_{p^L} + \eta), 2(1-\delta) + \delta(\underline{v}^{\infty}_{p^L} + \eta)\} \\
&= \min\{(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}}, 2(1-\delta) + \delta((1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}} + \eta)\} \\
&= \min\{(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}}, (1-\delta)(3+\delta) + \delta^2 \underline{v}^{\infty}_{p^{L-1}}\} \quad (2.3.22)
\end{aligned}$$

If $p^L = 1$, I have

$$\begin{aligned}
\min \mathcal{V}(\bar{\mathcal{D}}) &= \min\{\min\{(1-\delta) + \delta \underline{v}^{\infty}_{\phi} : \phi \leq \hat{\phi}_3\}, \\
&\qquad \min\{2(1-\delta) + \delta(\underline{v}^{\infty}_{\phi} + \eta) : \phi \in (\hat{\phi}_3, 1)\}, 2(1-\delta) + \delta \underline{v}^{\infty}_1\} \quad (2.3.23) \\
&= \min\{(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}}, 2(1-\delta) + \delta(\underline{v}^{\infty}_{p^{L-1}} + \eta), 2(1-\delta) + \delta \underline{v}^{\infty}_{p^L}\} \\
&= \min\{(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}}, 2(1-\delta) + \delta((1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}} + \eta)\} \\
&= \min\{(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}}, (1-\delta)(3+\delta) + \delta^2 \underline{v}^{\infty}_{p^{L-1}}\}, \quad (2.3.24)
\end{aligned}$$

where (2.3.24) matches (2.3.22). Suppose by contradiction that

$$(1-\delta) + \delta \underline{v}^{\infty}_{p^{L-1}} > (1-\delta)(3+\delta) + \delta^2 \underline{v}^{\infty}_{p^{L-1}};$$

rearrangement yields

$$\delta(1-\delta)\underline{v}^{\infty}_{p^{L-1}} > (1-\delta)(2+\delta)$$

$$\underline{v}^{\infty}_{p^{L-1}} > \frac{2+\delta}{\delta} > 3,$$

an infeasible payoff, and so a contradiction. Rearranging $(1-\delta)+\delta\underline{v}^{\infty}_{p^{L-1}} = \underline{v}^{\infty}_{p^L}$ yields $\underline{v}^{\infty}_{p^{L-1}} = \frac{\underline{v}^{\infty}_{p^L}}{\delta} - \eta$. Continuing,

$$\begin{aligned}
\min \mathcal{V}(\bar{\mathcal{D}}) &= (1-\delta)+\delta\underline{v}^{\infty}_{p^{L-1}} = \begin{cases} (1-\delta)+\delta\lambda_1^{L-1} & p^L > \frac{1}{2} \\ (1-\delta)+\delta\left[\frac{\underline{v}^{\infty}_{p^L}}{\delta}-\eta\right] & p^L = \frac{1}{2} \end{cases} \\
&= \begin{cases} (1-\delta)+\delta\lambda_1^{L-1} & p^L > \frac{1}{2} \\ (1-\delta)+\delta\left[\frac{\lambda_1^L}{\delta}-\eta\right] & p^L = \frac{1}{2} \end{cases} \\
&= \begin{cases} (1-\delta)+\delta\lambda_1^{L-1} & p^L > \frac{1}{2} \\ \lambda_1^L & p^L = \frac{1}{2}. \end{cases}
\end{aligned}$$

With respect to the maxima $\bar{v}^{\infty}_{\phi}$, the conclusions of Lemma 2.3.4 with almost identical arguments to the proof thereof, giving $\max\{\bar{v}^{\infty}_{\phi} : \phi \in [0,\hat{\phi}_3)\} = 2$, $\max\{\bar{v}^{\infty}_{\phi} : \phi \in [\hat{\phi}_3, 1)\} = 2-\eta$ and $\bar{v}^{\infty}_1 = 2$. Thus,

$$\begin{aligned}
\max \mathcal{V}(\bar{\mathcal{D}}) &= \max\{(1-\delta)+\delta\max\{\bar{v}^{\infty}_{\phi} : \phi \in [0,\hat{\phi}_3)\}, 3(1-\delta)+\delta\max\{\bar{v}^{\infty}_{\phi} : \phi \in [\hat{\phi}_3,1)\}, \\
&\qquad 2(1-\delta)+\delta\bar{v}^{\infty}_1\} \\
&= \max\{(1-\delta)+2\delta, 3(1-\delta)+\delta(2-\eta), 2(1-\delta)+2\delta\} = 2.
\end{aligned}$$

Now consider the $\mu^0 > \frac{1}{2}$ case (the $\mu^0 = \frac{1}{2}$ case is handled at the end of the proof). I use the following analogue of Lemma 2.3.2.

Lemma 2.3.8. *Suppose $\mu^0 \in (\frac{1}{2}, 1]$. Let any HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ be given. Then*

$$\bar{\mathcal{B}}(\tilde{w}) = \begin{cases} \{(1, g((1-\delta)+\delta\tilde{\gamma}(D), 3(1-\delta)+\delta\tilde{\gamma}(D))\} & \tilde{\phi} \in [0, \frac{1}{2}) \\ \{(\phi,\gamma) : \gamma(D) \in [(1-\delta)+\delta\tilde{\gamma}(D), 3(1-\delta)+\delta\tilde{\gamma}(D)], & \\ \qquad \gamma(C) = 3(1-\delta)+\delta\tilde{\gamma}(D)\} \cap \bar{\mathcal{F}} & \tilde{\phi} = \frac{1}{2} \\ \{(0,\gamma) : \gamma(D) = \gamma(C), \gamma(C) = 2(1-\delta)+\delta\tilde{\gamma}(C)\} & \tilde{\phi} \in (\frac{1}{2}, 1]. \end{cases}$$

Proof. I consider each case sequentially. Let some HBA $x \equiv (\phi, \mu, \alpha)$ enforced by $\tilde{w}$ and the HBP $w \equiv (\phi, \mu, \gamma)$ decomposed by x and $\tilde{w}$ be given.

Suppose $\tilde{\phi} \in [0, \frac{1}{2})$. Inducibility requires $\alpha_1 = \tilde{\phi} < \frac{1}{2}$, so player 2 has d as a strict best response by the same reasoning as in the proof of Lemma 5.2, and D must also be a best response for player 1. Thus, $\gamma(D) = (1-\delta)u_1(d, D) + \delta\tilde{\gamma}(D) = (1-\delta) + \delta\tilde{\gamma}(D)$. At history C, player 2 has a belief greater than $\frac{1}{2}$ for the commitment type, so arguments in the proof of Lemma 5.2 show that c is a strict best response. Then $\gamma(C) = (1-\delta)u_1(c, D) + \delta\tilde{\gamma}(D) = 3(1-\delta) + \delta\tilde{\gamma}(D)$.

Suppose $\tilde{\phi} = \frac{1}{2}$. Player 2 is now indifferent at history D, so can choose any $\alpha_2 \in [0, 1]$, which yields $\gamma(D) = (1-\delta)[\alpha_2 u_1(c, D) + (1-\alpha_2)u_2(d, D)] + \delta\tilde{\gamma}(D) = (1-\delta)[3\alpha_2 + (1-\alpha_2)] + \delta\tilde{\gamma}(D)$. The same reasoning as above gives $\gamma(C) = 3(1-\delta) + \delta\tilde{\gamma}(D)$.

Suppose $\tilde{\phi} \in (\frac{1}{2}, 1]$. Player 2 strictly prefers c at both histories C and D, while C is a best response for player 1. Thus $\gamma(D) = \gamma(C) = (1-\delta)u_1(c, C) + \delta\tilde{\gamma}(C) = 2(1-\delta) + \delta\tilde{\gamma}(C)$. $\square$

Lemma 2.3.8 gives

$$\underline{v}_0^{m+1} = \min\{\min\{3(1-\delta) + \delta\underline{v}_\phi^m - \eta : \phi \in [\tfrac{1}{2}, 1)\}, 2(1-\delta) + \delta\underline{v}_1^m\}$$

$$\underline{v}_\phi^{m+1} = 3(1-\delta) + \delta\underline{v}_{\frac{1}{2}}^m - \eta \qquad \forall\phi \in (0, 1)$$

$$\underline{v}_1^{m+1} = \min\{3(1-\delta) + \delta\underline{v}_\phi^m : \phi \in [0, \tfrac{1}{2}]\}.$$

Then $\underline{v}_{\frac{1}{2}}^\infty = 3(1-\delta) + \delta\underline{v}_{\frac{1}{2}}^\infty - \eta$, which yields $\underline{v}_{\frac{1}{2}}^\infty = 3 - \frac{1}{\delta} = 2 - \eta$. So for all $\phi \in (0, 1)$,

$$\underline{v}_\phi^\infty = 3(1-\delta) + \delta\left(3 - \frac{1}{\delta}\right) - \frac{1-\delta}{\delta} = 3(1-\delta) + 3\delta - 1 - \frac{1-\delta}{\delta} = 2 - \eta.$$

I can then write

$$\begin{aligned} \underline{v}_0^\infty &= \min\{3(1-\delta) + \delta(2-\eta) - \eta, 2(1-\delta) + \delta\underline{v}_1^\infty\} \\ &= \min\{2-\eta, 2(1-\delta) + \delta\underline{v}_1^\infty\} \end{aligned}$$

$$\begin{aligned}\underline{v}_1^\infty &= \min\{3(1-\delta)+\delta\underline{v}_0^\infty, 3(1-\delta)+\delta(2-\eta)\} \\ &= \min\{3(1-\delta)+\delta\underline{v}_0^\infty, 2\}\end{aligned}$$

Suppose by contradiction $\underline{v}_1^\infty < 2$. Then

$$\begin{aligned}\underline{v}_0^\infty &= \min\{2-\eta, 2(1-\delta)+\delta[3(1-\delta)+\delta\underline{v}_0^\infty]\} \\ &= \min\{2-\eta, 2(1-\delta)+3\delta-3\delta^2+\delta^2\underline{v}_0^\infty\} \\ &= \min\{2-\eta, 2+\delta-3\delta^2+\delta^2\underline{v}_0^\infty\}.\end{aligned}$$

If $\underline{v}_0^\infty < 2-\eta$, then $\underline{v}_0^\infty = 2+\delta-3\delta^2+\delta^2\underline{v}_0^\infty$ which can be solved for

$$\underline{v}_0^\infty = \frac{2+\delta(1-3\delta)}{1-\delta^2} > 2,$$

a contradiction; thus $\underline{v}_1^\infty < 2$ implies $\underline{v}_0^\infty = 2-\eta$. Yet $\underline{v}_1^\infty = 3(1-\delta)+\delta\underline{v}_0^\infty = 3(1-\delta)+\delta(2-\eta) = 2$, a contradiction. Thus, $\underline{v}_1^\infty = 2$. Also suppose by contradiction $\underline{v}_0^\infty < 2-\eta$; then $\underline{v}_0^\infty = 2(1-\delta)+\delta\bar{v}_1^\infty = 2$. Hence, $\underline{v}_0^\infty = 2-\eta$.

To summarize: $\underline{v}_\phi^\infty = 2-\eta$ for all $\phi \in [0,1)$ and $\underline{v}_1^\infty = 2$. Plugging this into (2.3.3) in Lemma 2.3.3 shows that $\min \mathcal{V}(\bar{\mathcal{D}}) = 2$.

Turning to the maxima, Lemma 2.3.8 implies

$$\bar{v}_0^{m+1} = \max\{\max\{3(1-\delta)+\delta\bar{v}_\phi^m - \eta : \phi \in [\tfrac{1}{2},1)\}, 2(1-\delta)+\delta\bar{v}_1^m\}$$

$$\bar{v}_\phi^{m+1} = 3(1-\delta)+\delta\bar{v}_{\frac{1}{2}}^m - \eta \qquad \forall \phi \in (0,1)$$

$$\bar{v}_1^{m+1} = \max\{3(1-\delta)+\delta\bar{v}_\phi^m : \phi \in [0,\tfrac{1}{2}]\}$$

Then for all $\phi \in (0,1)$, $\bar{v}_{\frac{1}{2}}^{m+1} = \bar{v}_\phi^{m+1} = 3(1-\delta)+\delta\bar{v}_{\frac{1}{2}}^m - \eta$, which in the limit gives

$$\bar{v}_\phi^\infty = 3(1-\delta)+\delta\bar{v}_\phi^\infty - \frac{1-\delta}{\delta} \implies \bar{v}_\phi^\infty = 3-\frac{1}{\delta} = 2-\eta.$$

Note that $\bar{v}_0^0 = \bar{v}_\phi^0 = 3-\eta$ and $\bar{v}_1^0 = 3$. Suppose for some m that $\bar{v}_{\frac{1}{2}}^m \leq \bar{v}_1^m - \eta$ and $\bar{v}_{\frac{1}{2}}^m \leq \bar{v}_0^m$.

Suppose by contradiction that $\max\{3(1-\delta)+\delta\bar{v}_\phi^m-\eta : \phi\in[\frac{1}{2},1)\} > 2(1-\delta)+\delta\bar{v}_1^m$. Then

$$\max\{3(1-\delta)+\delta\bar{v}_\phi^m-\eta : \phi\in[\tfrac{1}{2},1)\} = 3(1-\delta)+\delta\bar{v}_{\frac{1}{2}}^m-\eta > 2(1-\delta)+\delta\bar{v}_1^m \tag{2.3.25}$$

$$(1-\delta)-\eta > \delta(\bar{v}_1^m-\bar{v}_{\frac{1}{2}}^m)$$

$$-\eta > \delta(\bar{v}_1^m-\eta-\bar{v}_{\frac{1}{2}}^m),$$

but since the left hand side is strictly negative and the right hand side is non-negative, I reach a contradiction. Note that the left hand side of (2.3.25) is equal to $\bar{v}_{\frac{1}{2}}^{m+1}$. Thus by induction, for all m, $\bar{v}_0^{m+1}=2(1-\delta)+\delta\bar{v}_1^m\geq\bar{v}_{\frac{1}{2}}^{m+1}$. Similarly, $\bar{v}_1^{m+1}=3(1-\delta)+\delta\bar{v}_0^m\geq 3(1-\delta)+\delta\bar{v}_{\frac{1}{2}}^m$. Taking the limit,

$$\bar{v}_0^\infty=2(1-\delta)+\delta\bar{v}_1^\infty, \qquad \bar{v}_1^\infty=3(1-\delta)+\delta\bar{v}_0^\infty.$$

Solving these two equations gives

$$\bar{v}_0^\infty=2(1-\delta)+\delta[3(1-\delta)+\delta\bar{v}_0^\infty]=2(1-\delta)+3\delta(1-\delta)+\delta^2\bar{v}_0^\infty$$

$$\bar{v}_0^\infty=\frac{1-\delta}{1-\delta^2}(2+3\delta)$$

$$\bar{v}_1^\infty=3(1-\delta)+\delta(1-\delta)\frac{2+3\delta}{1-\delta^2}=(1-\delta)\left[3+\frac{2\delta+3\delta^2}{1-\delta^2}\right]=\frac{1-\delta}{1-\delta^2}(3+2\delta).$$

Finally, by Lemma 2.3.3,

$$\begin{aligned}
\max\mathcal{V}(\bar{\mathcal{D}}) &= \max\{3(1-\delta)+\delta\bar{v}_0^\infty,\\
&\qquad \max\{2(1-\delta)+\delta(\bar{v}_\phi^\infty+\eta):\phi\in(0,1)\},2(1-\delta)+\delta\bar{v}_1^\infty\}\\
&= \max\left\{3(1-\delta)+\delta\frac{1-\delta}{1-\delta^2}(2+3\delta),2(1-\delta)+\delta(\bar{v}_{\frac{1}{2}}^\infty+\eta),\right.\\
&\qquad \left.2(1-\delta)+\delta\frac{1-\delta}{1-\delta^2}(3+2\delta)\right\}\\
&= \max\left\{3(1-\delta)+\delta\frac{1-\delta}{1-\delta^2}(2+3\delta),2,2(1-\delta)+\delta\frac{1-\delta}{1-\delta^2}(3+2\delta)\right\}.
\end{aligned}$$

It is straightforward to show that the first term is strictly greater than the other two for

$\delta \in (\frac{1}{2}, 1)$, so

$$\max \mathcal{V}(\bar{\mathcal{D}}) = 3(1-\delta) + \frac{1-\delta}{1-\delta^2}(2\delta + 3\delta^2) = \frac{1-\delta}{1-\delta^2}(3+2\delta).$$

Now suppose $\mu^0 = \frac{1}{2}$. Note that $\phi^* = 1$, so $q(0) = \frac{1}{1-2\cdot 0} = 1$, $\hat{\phi}_3 = 0$, $L = 1$ and $p^L = 1$. Lemma 2.3.6 shows that

$$\underline{v}_0^\infty = \underline{v}_{\hat{\phi}_3}^\infty = \underline{v}_{p^{L-1}}^\infty = \lambda_1^{L-1} = \min\{\underline{v}_\phi^\infty : \phi \in (0,1)\}$$

and $\underline{v}_1^\infty = \underline{v}_{p^L}^\infty = (1-\delta) + \delta\lambda_1^{L-1}$. Using the same arguments as for the $\mu^0 \in (\frac{1}{9}, \frac{1}{2})$ case above, I have $\min \mathcal{V}(\bar{\mathcal{D}}) = (1-\delta) + \delta\lambda_1^{L-1}$. For the maxima, the same arguments as those above for $\mu^0 \in (\frac{1}{2}, 1]$ hold, so $\max \mathcal{V}(\bar{\mathcal{D}}) = \frac{1-\delta}{1-\delta^2}(3+2\delta)$.

2.3.2 Purifiability of Quasi-Markov Equilibria

This section shows that for almost all priors μ^0 and discount factors $\delta > \frac{1}{2}$, there exists a purifiable equilibrium giving the minimum and maximum payoffs given in Proposition 2.5.1.[6] For each case, I construct the unperturbed equilibrium strategy profile σ^* and a sequence of perturbed game strategy profiles $(\tilde{\sigma}^k)_k$, each corresponding to the (ψ, ε^k)-perturbed game where ψ is the uniform distribution on $[0,1]^{|A|}$ and $(\varepsilon^k)_k$ is some sequence where $\varepsilon^k > 0$ and $\varepsilon^k \to 0$. I show purifiability (according to Definition 2.4.3) by proving that for small enough ε^k, $\tilde{\sigma}^k$ is a wPBE of the (ψ, ε^k)-perturbed game, and that $(\tilde{\sigma}^k)_k$ converges in outcomes to σ^*. (I omit proofs that the unperturbed strategies σ^* are wPBEs of the unperturbed game since the proofs are essentially simpler versions of the proofs for the perturbed equilibria.)

2.3.2.1 Stationary Equilibrium with Payoff 2 ($0 < \mu^0 \le 1$)

The equilibrium for the unperturbed game is defined as follows:

[6] The $\delta < \frac{1}{2}$ cases are straightforward (with player 1 always playing D) and so I omit them.

$$\sigma_2^*(t,h) = \begin{cases} 1 & h = \varnothing \\ 1 & h = C \\ 1 - \frac{1-\delta}{\delta} & h = D \end{cases} \qquad \sigma_1^*(t) = \frac{1}{2}.$$

Define the sequence $(\tilde{\sigma}^k)_k$ of equilibria in the sequence of perturbed games:

$$\tilde{\sigma}_2^k(\varnothing, z_2) \equiv \tilde{\sigma}_2^k(t, C, z_2) \equiv 1 \qquad \tilde{\sigma}_2^k(t, D, z_2) \equiv \begin{cases} 0 & \Delta z_2 \leq \zeta_2^k \\ 1 & \Delta z_2 > \zeta_2^k \end{cases}$$

$$\tilde{\sigma}_1^k(t, z_1) \equiv \begin{cases} 0 & \Delta z_1 \leq \zeta_1^k \\ 1 & \Delta z_1 > \zeta_1^k \end{cases}$$

where I define ζ_1^k, ζ_2^k by the system of equations[7]

$$\zeta_2^k \equiv -1 + \frac{2(1-\delta)}{\delta} - \frac{2}{\delta}\varepsilon^k \zeta_1^k \qquad \zeta_1^k \equiv \varepsilon^k \zeta_2^k.$$

For convenience, define the expected outcomes

$$\tilde{\alpha}_2^k(h) \equiv \int \tilde{\sigma}_2^k(t, D)\, d\psi(z_2) = \begin{cases} 1 & h = C \\ \frac{1}{2}(1 - \zeta_2^k) & h = D \end{cases}$$

$$\tilde{\alpha}_1^k \equiv \int \tilde{\sigma}_1^k(t)\, d\psi(z_1) = \tfrac{1}{2}(1 - \zeta_1^k)$$

for any $t \geq 1$. Also define the beliefs at history C at any period t:

$$\tilde{\mu}^k \equiv \frac{\mu^0}{\mu^0 + (1-\mu^0)\tilde{\alpha}_1^k}. \tag{2.3.26}$$

I now show that these strategies are mutual best responses. Define $V_2^{a_2}(p, \bar{\alpha}_1, z_2)$ as the expected payoff for player 2 of action a_2 given posterior belief p, expected player 1

[7] Though these equations are clearly easy to solve, leaving them in this form makes it simpler to confirm that the strategies are best responses.

action $\bar{\alpha}_1$, and shock z_2:

$$V_2^c(p, \bar{\alpha}_1, z_2) = 3p + (1-p)[3\bar{\alpha}_1] + \varepsilon^k z_2^c$$

$$V_2^d(p, \bar{\alpha}_1, z_2) = 2p + (1-p)[2\bar{\alpha}_1 + (1-\bar{\alpha}_1)] + \varepsilon^k z_2^d.$$

Then define ΔV_2 as the benefit of playing c over d:

$$\begin{aligned} \Delta V_2(p, \bar{\alpha}_1, z_2) &\equiv V_2^c(p, \bar{\alpha}_1, z_2) - V_2^d(p, \bar{\alpha}_1, z_2) \\ &= 3p + (1-p)[3\bar{\alpha}_1] - 2p - (1-p)[2\bar{\alpha}_1 + (1-\bar{\alpha}_1)] + \varepsilon^k (z_2^c - z_2^d) \\ &= p + (1-p)[2\bar{\alpha}_1 - 1] + \varepsilon^k \Delta z_2. \end{aligned} \tag{2.3.27}$$

At history D, the belief is $p = 0$ so

$$\Delta V_2(0, \tilde{\alpha}_1^k, z_2) = 2\tilde{\alpha}_1^k - 1 + \varepsilon^k \Delta z_2 = 2 \cdot \tfrac{1}{2}(1 - \zeta_1^k) - 1 + \varepsilon^k \Delta z_2 = -\varepsilon^k \zeta_2^k + \varepsilon^k \Delta z_2,$$

which makes it clear that $\tilde{\sigma}_2^k(t, D, z_2)$ is a best response. At history C, the posterior is $\tilde{\mu}^k \geq \mu^0$, so for small enough ε^k, $\Delta V_2(\tilde{\mu}^k, \tilde{\alpha}_1^k, z_2)$ is positive for all z_2 and thus $\tilde{\sigma}_2^k(t, C, z_2)$ is a best response. The same is true at the empty history $\varnothing$ at period 0.

Define $V_1^{a_1}(t, z_1)$ as the payoff to player 1 of playing action a_1 at period t:

$$V_1^C(t, z_1) = (1-\delta) u_1(a_2, C) + \delta V(t+1, C) + \varepsilon^k z_1^C$$

$$V_1^D(t, z_1) = (1-\delta) u_1(a_2, D) + \delta V(t+1, D) + \varepsilon^k z_1^D$$

where $V(t+1, h)$ is the continuation payoff for the start of period $t+1$ with history h. The benefit $\Delta V_1(t, z_1)$ of playing C over D is

$$\begin{aligned} \Delta V_1(t, z_1) &\equiv V_1^C(t, z_1) - V_1^D(t, z_1) \\ &= -(1-\delta) + \delta(V(t+1, C) - V(t+1, D)) + \varepsilon^k (z_1^C - z_1^D) \\ &= -(1-\delta) + \delta(\bar{\alpha}_2^{t+1}(C) - \bar{\alpha}_2^{t+1}(D)) + \varepsilon^k \Delta z_1 \end{aligned} \tag{2.3.28}$$

where $\bar{\alpha}_2^{t+1}(h)$ is the strategy of next period's player 2 at history h. In this equilibrium, those strategies are $\tilde{\alpha}_2^k(C) = 1, \tilde{\alpha}_2^k(D) = \frac{1}{2}(1-\zeta_2^k)$ so

$$\begin{aligned}
\Delta V_1(t, z_1) &= -(1-\delta) + \delta(\tilde{\alpha}_2^k(C) - \tilde{\alpha}_2^k(D)) + \varepsilon^k \Delta z_1 \\
&= -(1-\delta) + \delta(1 - \tfrac{1}{2}(1-\zeta_2)) + \varepsilon^k \Delta z_1 \\
&= -(1-\delta) + \tfrac{1}{2}\delta\left(1 + \left[-1 + \frac{2(1-\delta)}{\delta} - \frac{2}{\delta}\varepsilon^k \zeta_1\right]\right) + \varepsilon^k \Delta z_1 \\
&= -\varepsilon^k \zeta_1 + \varepsilon^k \Delta z_1,
\end{aligned}$$

showing player 1's strategy is a best response.

Finally, I show that $(\tilde{\sigma}^k)_k$ converges in outcomes to σ^* as given in (2.4.1). For $\tilde{\sigma}_2^k(\varnothing, z_2) = \tilde{\sigma}_2^k(t, C, z_2) = 1 = \sigma_2^*(\varnothing) = \sigma_2^*(t, C)$, the convergence is trivial. For history D, I integrate over the shocks:

$$\begin{aligned}
\lim_{k\to\infty} \int \tilde{\sigma}_2^k(t, D, z_2)\, d\psi(z_2) &= \lim_{k\to\infty} \tfrac{1}{2}(1-\zeta_2^k) \\
&= \lim_{k\to\infty} \frac{1}{2}\left(1 - \left[-1 + \frac{2(1-\delta)}{\delta} - \frac{2}{\delta}\varepsilon^k \zeta_1^k\right]\right) \\
&= 1 - \frac{1-\delta}{\delta} = \sigma_2^*(t, D).
\end{aligned}$$

For player 1, I have

$$\lim_{k\to\infty} \int \tilde{\sigma}_1^k(t, z_1)\, d\psi(z_1) = \tfrac{1}{2}(1-\zeta_1^k) = \tfrac{1}{2}(1-\varepsilon^k\zeta_2^k) = \tfrac{1}{2} = \sigma_1^*(t).$$

2.3.2.2 Stationary Equilibrium with Payoff 1 ($0 < \mu^0 < \frac{1}{9}$)

For $0 < \mu^0 < \frac{1}{9}$, the minimum payoff equilibrium is defined as

$$\sigma_2^*(t,h) = \begin{cases} 0 & h = \varnothing \\ \frac{1-\delta}{\delta} & h = C \\ 0 & h = D \end{cases} \qquad \sigma_1^*(t) = \frac{1}{4} \pm \frac{1}{4}\sqrt{1 - \frac{8\mu^0}{1-\mu^0}}.$$

Note that $\sigma_1^*(t)$ is a fixed point of $r(\cdot) \equiv q^{-1}(\cdot)$ defined in (2.5.1) and (2.3.1), respectively; see (2.3.9) for the calculation of the fixed points. Define the sequence $(\tilde{\sigma}^k)_k$ of equilibria in the sequence of perturbed games:

$$\tilde{\sigma}_2^k(t,C,z_2) \equiv \begin{cases} 0 & \varepsilon^k \Delta z_2 \leq \zeta_2^k \\ 1 & \varepsilon^k \Delta z_2 > \zeta_2^k \end{cases} \qquad \tilde{\sigma}_2^k(t,D,z_2) \equiv 0 \qquad \tilde{\sigma}_1^k(t,z_1) \equiv \begin{cases} 0 & \varepsilon^k \Delta z_1 \leq \zeta_1^k \\ 1 & \varepsilon^k \Delta z_1 > \zeta_1^k \end{cases}$$

where I define ζ_1^k, ζ_2^k by the system of equations[8]

$$\zeta_2^k \equiv 1 - \frac{2(1-\delta)}{\delta} + \frac{2}{\delta}\varepsilon^k \zeta_1^k \qquad \zeta_1^k \equiv \frac{\tilde{\mu}^k + \varepsilon^k \zeta_2^k}{1-\tilde{\mu}^k} \tag{2.3.30}$$

where $\tilde{\mu}^k$ is defined as in (2.3.26).

I show that the strategies are mutual best responses. Reusing the notation above in (2.3.27), at history C the belief is $p = \tilde{\mu}^k$, so

$$\begin{aligned} \Delta V_2(\tilde{\mu}^k, \tilde{\alpha}_1^k, z_2) &= \tilde{\mu}^k + (1-\tilde{\mu}^k)[2\tilde{\alpha}_1^k - 1] + \varepsilon^k \Delta z_2 \\ &= \tilde{\mu}^k + (1-\tilde{\mu}^k)[2 \cdot \tfrac{1}{2}(1-\zeta_1^k) - 1] + \varepsilon^k \Delta z_2 \\ &= \tilde{\mu}^k - (1-\tilde{\mu}^k)\left[\frac{\tilde{\mu}^k + \varepsilon^k \zeta_2^k}{1-\tilde{\mu}^k}\right] + \varepsilon^k \Delta z_2 \\ &= -\varepsilon^k \zeta_2^k + \varepsilon^k \Delta z_2, \end{aligned}$$

which makes it clear that $\tilde{\sigma}_2^k(t,C,z_2)$ is a best response. At the initial history $\varnothing$, the belief is the prior $\mu^0 < \tilde{\mu}^k$, so for small enough ε^k, $\Delta V_2(\mu^0, \tilde{\alpha}_1^k, z_2)$ is negative for all z_2; similarly, at history D, the posterior is $0 < \tilde{\mu}^k$, giving the same result. Thus $\tilde{\sigma}_2^k(\varnothing, z_2), \tilde{\sigma}_2^k(t,D,z_2)$

[8] I leave them in this form for the same reason given in footnote 7. To show that a solution exists, I can write

$$\begin{aligned} \tilde{\alpha}_1^k &\equiv \frac{1}{2}\left(1-\zeta_1^k\right) = \frac{1}{2}\left(1 - \frac{\tilde{\mu}^k + \varepsilon^k \zeta_2^{k,t}}{1-\tilde{\mu}^k}\right) = \frac{1}{2}\left(1 - \frac{\frac{\mu^0}{\mu^0 + (1-\mu^0)\tilde{\alpha}_1^k} + \varepsilon^k \zeta_2^{k,t}}{1 - \frac{\mu^0}{\mu^0+(1-\mu^0)\tilde{\alpha}_1^k}}\right) \\ &= \frac{1}{2}\left(1 - \frac{\mu^0 + \varepsilon^k \zeta_2^{k,t}(\mu^0 + (1-\mu^0)\tilde{\alpha}_1^k)}{(1-\mu^0)\tilde{\alpha}_1^k}\right) = \frac{1}{2}\left(\frac{(1-\mu^0)\tilde{\alpha}_1^k - \mu^0 - \varepsilon^k \zeta_2^{k,t}(\mu^0 + (1-\mu^0)\tilde{\alpha}_1^k)}{(1-\mu^0)\tilde{\alpha}_1^k}\right) \\ &= \frac{1}{2}\left(1 - \frac{\frac{\mu^0}{1-\mu^0}}{\tilde{\alpha}_1^k} - \frac{\varepsilon^k \zeta_2^{k,t}(\mu^0 + (1-\mu^0)\tilde{\alpha}_1^k)}{(1-\mu^0)\tilde{\alpha}_1^k}\right) \end{aligned} \tag{2.3.29}$$

Note that $\mu^0 < \frac{1}{9}$ guarantees two fixed points of $r(\tilde{\alpha}_1^k) = \frac{1}{2}\left(1 - \frac{\mu^0/(1-\mu^0)}{\tilde{\alpha}_1^k}\right)$ (see (2.3.9) for the reasoning), and that $r(\cdot)$ is strictly concave. Since the limit of the right hand side of (2.3.29) as $\varepsilon^k \to 0$ is $r(\tilde{\alpha}_1^k)$, for small enough ε^k a solution to (2.3.30) exists.

are best responses. Reusing the notation in (2.3.28),

$$\begin{aligned}\Delta V_1(t, z_1) &= -(1-\delta) + \delta(\tilde{\alpha}_2^k(C) - \tilde{\alpha}_2^k(D)) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + \delta(\tfrac{1}{2}(1-\zeta_2^k) - 0) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + \tfrac{1}{2}\delta\left(1 - \left[1 - \frac{2(1-\delta)}{\delta} + \frac{2}{\delta}\varepsilon^k \zeta_1^k\right]\right) + \varepsilon^k \Delta z_1 \\ &= -\varepsilon^k \zeta_1^k + \varepsilon^k \Delta z_1,\end{aligned}$$

so player 1's strategy is a best response.

Finally, I show that $(\tilde{\sigma}^k)_k$ converges in outcomes to σ^*. For

$$\tilde{\sigma}_2^k(\varnothing, z_2) = \tilde{\sigma}_2^k(t, D, z_2) = 0 = \sigma_2^*(\varnothing) = \sigma_2^*(t, D),$$

the convergence is trivial; at history C,

$$\begin{aligned}\lim_{k\to\infty} \int \tilde{\sigma}_2^k(t, C, z_2)\, d\psi(z_2) &= \lim_{k\to\infty} \tfrac{1}{2}(1-\zeta_2^k) \\ &= \lim_{k\to\infty} \frac{1}{2}\left(1 - \left[1 - \frac{2(1-\delta)}{\delta} + \frac{2}{\delta}\varepsilon^k \zeta_1^k\right]\right) = \frac{1-\delta}{\delta} = \sigma_2^*(t, C).\end{aligned}$$

Footnote 8 shows that $\int \tilde{\sigma}_1^k(t, z_1)\, d\psi(z_1) = \lim_{k\to\infty} \tilde{\alpha}_1^k$ converges to a fixed point of $r(\cdot)$; since $\sigma_1^*(t)$ is also a fixed point of $r(\cdot)$, I can pick $\tilde{\sigma}^k$ that converges to $\sigma_1^*(t)$.

2.3.2.3 Non-Stationary Minimum Payoff Equilibrium ($\frac{1}{9} < \mu^0 \leq \frac{1}{2}$)

For $\frac{1}{9} < \mu^0 \leq \frac{1}{2}$, the minimum payoff equilibrium is defined as follows. Define $L(\mu^0), q(\phi, \mu^0)$ as stated in Proposition 2.5.1 (I will usually omit the "μ^0" argument in both for brevity). I restrict attention to priors μ^0 such that $L(\mu^0) > \frac{1}{2}$ (the set of priors $\mu^0 \in (\frac{1}{9}, \frac{1}{2}]$ such that this is not true is Lebesgue measure zero). The equilibrium strategies are

$$\sigma_2^*(t, h) = \begin{cases} 0 & h = \varnothing \\ \frac{1-\delta}{\delta} & (t+1) \bmod L \neq 0, h = C \\ 0 & (t+1) \bmod L \neq 0, h = D \\ 1 & (t+1) \bmod L = 0 \end{cases} \qquad \sigma_1^*(t) = q^{L-1-[(t+1) \bmod L]}(0, \mu^0).$$

Define the sequence $(\tilde{\sigma}^k)_k$ of equilibria in the sequence of perturbed games:[9]

$$\tilde{\sigma}_2^k(t,h,z_2) \equiv \begin{cases} 0 & \varepsilon^k \Delta z_2 < \zeta_2^{k,(t+1) \bmod L}(h) \\ 1 & \varepsilon^k \Delta z_2 > \zeta_2^{k,(t+1) \bmod L}(h) \end{cases} \qquad \tilde{\sigma}_1^k(t,z_1) \equiv \begin{cases} 0 & \varepsilon^k \Delta z_1 < \zeta_1^{k,(t+1) \bmod L} \\ 1 & \varepsilon^k \Delta z_1 > \zeta_1^{k,(t+1) \bmod L} \end{cases}$$

where I define the thresholds $\zeta_2^{k,0}(C), \zeta_2^{k,0}(D), \zeta_1^{k,0}, ..., \zeta_2^{k,L-1}(C), \zeta_2^{k,L-1}(D), \zeta_1^{k,L-1}$ below; for the empty history, $\zeta_2^{k,1}(\varnothing) \equiv 1$. For convenience and clarity, define:

$$\tilde{\alpha}_1^{k,l} \equiv \int \tilde{\sigma}_1^k(t,z_1)\, d\psi(z_1) = \tfrac{1}{2}(1-\zeta_1^{k,l}) \qquad \tilde{\alpha}_2^{k,l}(h) \equiv \int \tilde{\sigma}_2^k(t,h,z_2)\, d\psi(z_2) = \tfrac{1}{2}(1-\zeta_2^{k,l}(h))$$

$$\tilde{\mu}^{k,l} \equiv \frac{\mu^0}{\mu^0 + (1-\mu^0)\tilde{\alpha}_1^{k,(l-1) \bmod L}} \tag{2.3.31}$$

for some t such that $l = (t+1) \bmod L$. Define the thresholds as follows. Let $\zeta_2^{k,l}(D) \equiv 1$ for all $l \geq 1$, and let $\zeta_1^{k,L-1} \equiv 1$. Define $\zeta_1^{k,L-2}, \zeta_2^{k,L-1}(C)$ as the solutions to the system of equations[10]

$$0 = 2\tilde{\mu}^{k,L-1} - 1 + \varepsilon^k \zeta_2^{k,L-1} \tag{2.3.32}$$

$$0 = -(1-\delta) + \tfrac{1}{2}\delta(1-\zeta_2^{k,L-1}(C)) + \varepsilon^k \zeta_1^{k,L-2}. \tag{2.3.33}$$

[9] For the sake of simpler notation, I can and do ignore the measure zero set of cases where $\Delta z_i = \zeta_i^{k,(t+1) \bmod L}$; it is straightforward to fill in best responses for these remaining cases.

[10] I do this for the same reason given in footnote 7. I omit the tedious algebra that gives the solution

$$\begin{aligned} \zeta_1^{k,L-2} \;=\; & \frac{1}{4\varepsilon^k(1-\mu^0)}\Big(\delta(1-3\varepsilon^k)(1-\mu^0) + 2\varepsilon^k(1-\varepsilon^k)(1-\mu^0) \\ & +((\delta(1-3\varepsilon^k)(1-\mu^0) + 2\varepsilon^k(1-\varepsilon^k)(1-\mu^0))^2 + 8(\varepsilon^k)^2(1-\mu^0)(-2\varepsilon^k(1+\mu^0) \\ & \quad +\delta(-1+3\mu^0+3\varepsilon^k(1+\mu^0))))^{1/2}\Big). \end{aligned}$$

Note that for small enough ε^k, the discriminant is non-negative and thus a real solution exists.

Then for each $l \in \{1, ..., L-2\}$, define (backward inductively) $\zeta_1^{k,l-1}, \zeta_2^{k,l}$ as solutions to the equations[11]

$$0 = \tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})(2\tilde{\alpha}_1^{k,l} - 1) + \varepsilon^k \zeta_2^{k,l} \tag{2.3.34}$$

$$0 = -(1-\delta) + \tfrac{1}{2}\delta(1-\zeta_2^{k,l}(C)) + \varepsilon^k \zeta_1^{k,l-1} \tag{2.3.35}$$

(note that $\tilde{\mu}^{k,l}$ is a function of $\zeta_1^{k,l-1}$). Define $\zeta_2^{k,0}(C) \equiv \zeta_2^{k,0}(D) \equiv -1$. It is convenient to rearrange (2.3.34) as

$$\zeta_1^{k,l} = \frac{\tilde{\mu}^{k,l} + \varepsilon^k \zeta_2^{k,l}}{1-\tilde{\mu}^{k,l}}$$

and (2.3.35) as

$$\zeta_2^{k,l}(C) = 1 - \frac{2(1-\delta)}{\delta} + \frac{2}{\delta}\varepsilon^k \zeta_1^{k,l-1}.$$

I show that these are mutual best responses. Abusing notation, let "l" mean any period $t \geq 1$ such that $(t+1) \bmod L = l$. For each $l \in \{0, ..., L-1\}$, the benefit to player 2 of playing c over that of d is

$$\begin{aligned}
\Delta V_2(\tilde{\mu}^{k,l}, \tilde{\alpha}_1^{k,l}, z_2) &= \tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})[2\tilde{\alpha}_1^{k,l} - 1] + \varepsilon^k \Delta z_2 \\
&= \tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})[2 \cdot \tfrac{1}{2}(1-\zeta_1^{k,l}) - 1] + \varepsilon^k \Delta z_2 \\
&= \tilde{\mu}^{k,l} - (1-\tilde{\mu}^{k,l})\zeta_1^{k,l} + \varepsilon^k \Delta z_2
\end{aligned}$$

For $l = L-1$ and history C, I have

$$\begin{aligned}
\Delta V_2(\tilde{\mu}^{k,L-1}, \tilde{\alpha}_1^{k,L-1}, z_2) &= \tilde{\mu}^{k,L-1} - (1-\tilde{\mu}^{k,L-1}) + \varepsilon^k \Delta z_2 \\
&= 2\tilde{\mu}^{k,L-1} - 1 + \varepsilon^k \Delta z_2;
\end{aligned}$$

[11] Again, I omit the steps that yield the solution

$$\begin{aligned}
\zeta_1^{k,l-1} = \frac{1}{2}\Big(&(2\varepsilon^k - 3\varepsilon^k\delta)(1-\mu^0) + (\varepsilon^k)^2\delta(1+\mu^0) + \delta\zeta_1^{k,l}(1-\mu^0) \\
&+ \Big((\varepsilon^k(2(1-\mu^0) + \delta(\varepsilon^k(1+\mu^0) - 3(1-\mu^0))) + \delta(1-\mu^0)\zeta_1^{k,l})^2 \\
&+ 4(\varepsilon^k)^2\delta(1-\mu^0)(2\delta\mu^0 + \varepsilon^k(3\delta-2)(1+\mu^0) - \delta(1-\mu^0)\zeta_1^{k,l}) \Big)^{1/2} \Big).
\end{aligned}$$

For small enough ε^k, the discriminant is non-negative.

solving (2.3.32) for $\tilde{\mu}^{k,L-1}$ and substituting gives

$$\begin{aligned}\Delta V_2(\tilde{\mu}^{k,L-1}, \tilde{\alpha}_1^{k,L-1}, z_2) &= 2\left[\tfrac{1}{2}(1-\varepsilon^k \zeta_2^{k,L-1})\right] - 1 + \varepsilon^k \Delta z_2 \\ &= -\varepsilon^k \zeta_2^{k,L-1} + \varepsilon^k \Delta z_2,\end{aligned}$$

so $\tilde{\sigma}_2^k(L-1, C, z_2)$ is a best response. For both the empty history $\varnothing$ and history D, the belief is less than $\tilde{\mu}^{k,l}$for all $l \in \{0, ..., L-1\}$, so for small enough ε^k, $\Delta V_2$is negative for all z_2 and so $\tilde{\sigma}_2^k(\varnothing, z_2) = \tilde{\sigma}_2^k(l, D, z_2) = 0$ is a best response. For $1 \leq l < L-1$ and history C, I have

$$\begin{aligned}\Delta V_2(\tilde{\mu}^{k,l}, \tilde{\alpha}_1^{k,l}, z_2) &= \tilde{\mu}^{k,l} - (1-\tilde{\mu}^{k,l}) \frac{\tilde{\mu}^{k,l} + \varepsilon^k \zeta_2^{k,l}}{1-\tilde{\mu}^{k,l}} + \varepsilon^k \Delta z_2 \\ &= -\varepsilon^k \zeta_2^{k,l} + \varepsilon^k \Delta z_2.\end{aligned}$$

Applying the same algebra given in footnote 8 (replacing $\zeta_1^k, \tilde{\mu}^k$ with $\zeta_1^{k,l}, \tilde{\mu}^{k,l}$, respectively) shows

$$\tilde{\alpha}_1^{k,l} = r(\tilde{\alpha}^{k,l-1}) - \varepsilon^k \frac{\zeta_2^{k,l}[\mu^0 + (1-\mu^0)\tilde{\alpha}^{k,l-1}]}{2(1-\mu^0)\tilde{\alpha}^{k,l-1}}, \tag{2.3.36}$$

where $r(\cdot) \equiv q^{-1}(\cdot)$ is defined in (2.3.1). Since $q^L(L) > \frac{1}{2}$, for small enough ε^k, player 1's strategy must be to play C with more than probability $\frac{1}{2}$ at $l = 0$:

$$\lim_{\varepsilon^k \to 0} \tilde{a}^{k,L-1} = \lim_{\varepsilon^k \to 0} r^L(\tilde{a}^{k,0}) \implies \lim_{\varepsilon^k \to 0} \tilde{a}^{k,0} = \lim_{\varepsilon^k \to 0} q^L(\tilde{a}^{k,L-1}) = q^L(0) > \tfrac{1}{2}.$$

Then for small enough ε^k, $\Delta V_2(0, \tilde{\alpha}_1^{k,0}, z_2) = 2\tilde{\alpha}_1^{k,0} - 1 + \varepsilon^k \Delta z_2 > 0$ for all z_2. Turning to player 1, for $0 \leq l < L-2$, I have

$$\begin{aligned}\Delta V_1(l, z_1) &= -(1-\delta) + \delta(\tilde{\alpha}_2^{k,l+1}(C) - \tilde{\alpha}_2^{k,l+1}(D)) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + \tfrac{1}{2}\delta(1-\zeta_2^{k,l+1}) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + \tfrac{1}{2}\delta\left(1 - \left[1 - \frac{2(1-\delta)}{\delta} + \frac{2}{\delta}\varepsilon^k \zeta_1^{k,l}\right]\right) + \varepsilon^k \Delta z_1 \\ &= -\varepsilon^k \zeta_1^{k,l} + \varepsilon^k \Delta z_1\end{aligned}$$

so $\tilde{\sigma}_1^k(l, z_1)$ is a best response. For $l = L-1$,

$$\begin{aligned}\Delta V_1(t, z_1) &= -(1-\delta) + \delta(\tilde{\alpha}_2^{k,0}(C) - \tilde{\alpha}_2^{k,0}(D)) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + \varepsilon^k \Delta z_1\end{aligned}$$

which is negative for all z_1 for small enough ε^k.

Finally, I show convergence in outcomes. (I continue to abuse notation by letting "l" denote any period t such that $l = (t+1) \bmod L$.) For $l = 0$, $\zeta_2^{k,0}(C) = \zeta_2^{k,0}(D) = 1$ so $\int \tilde{\sigma}_2^k(l=0, h, z_2)\, d\psi(z_2) = 0 = \sigma_2^*(l=0, h)$. For $l \geq 1$ at history $h \in \{\varnothing, D\}$, $\zeta_2^{k,l}(h) = 1$ and so

$$\int \tilde{\sigma}_2^k(l, h, z_2)\, d\psi(z_2) = 0 = \sigma_2^*(l, h).$$

For player 1 and $l = L-1$, $\zeta_1^{k,L-1} = 1$ so

$$\int \tilde{\sigma}_1^k(l = L-1, z_1)\, d\psi(z_1) = 0 = \sigma_1^*(l = L-1).$$

For player 2 and $l \geq 1$ at history C,

$$\begin{aligned}&\lim_{k\to\infty} \int \tilde{\sigma}_2^k(l, C, z_2)\, d\psi(z_2) = \lim_{k\to\infty} \tfrac{1}{2}(1 - \zeta_2^{k,l}(C)) \\ &= \lim_{k\to\infty} \frac{1}{2}\left(1 - \left[1 - \frac{2(1-\delta)}{\delta} + \frac{2}{\delta}\varepsilon^k \zeta_1^{k,l-1}\right]\right) = \frac{1-\delta}{\delta} = \sigma_2^*(l, C).\end{aligned}$$

For player 1 and $l < L-1$, (2.3.36) shows that $\lim_{k\to\infty} \tilde{\alpha}_1^{k,l-1} = \lim_{k\to\infty} q(\tilde{\alpha}_1^{k,l})$ and so

$$\lim_{k\to\infty} \int \tilde{\sigma}_1^k(l, z_1)\, d\psi(z_1) = q^{L-1-l}(\tilde{\alpha}_1^{L-1}) = q^{L-1-l}(0) = \sigma_1^*(l).$$

2.3.2.4 Non-Stationary Maximum Payoff Equilibrium ($\frac{1}{2} < \mu^0 \leq 1$)

The maximum payoff equilibrium is defined as follows:

$$\sigma_2^*(t, h) = \begin{cases} 1 & \text{even } t, h \in \{C, D\} \\ 0 & \text{even } t, h = D \\ 1 & \text{odd } t \end{cases} \qquad \sigma_1^*(t) = \begin{cases} 0 & \text{even } t \\ 1 & \text{odd } t. \end{cases}$$

Since σ_2^* is sequentially strict, purifying it is straightforward. Define the strategies the same way for the perturbed games, ignoring the shocks: $\tilde{\sigma}_i^k(t,h) \equiv \sigma_i^*(t,h)$ for all i, t, h.

For player 2, the posterior belief at each date-history is given by

$$p = \begin{cases} \mu^0 & \text{even } t, h = C \\ 1 & \text{odd } t, h = C \\ 0 & h = D \end{cases}$$

so the benefit of playing c over d is given by

$$\begin{aligned} \Delta V_2(p, \bar{\alpha}_1, z_2) &= p + (1-p)[2\bar{\alpha}_1 - 1] + \varepsilon^k \Delta z_2 \\ &= \begin{cases} \mu^0 - (1-\mu^0) + \varepsilon^k \Delta z_2 & \text{even } t, h = C \\ -1 + \varepsilon^k \Delta z_2 & \text{even } t, h = D \\ 1 + \varepsilon^k \Delta z_2 & \text{odd } t, h = C \\ \mu^0 + (1-\mu^0) + \varepsilon^k \Delta z_2 & \text{odd } t, h = D; \end{cases} \end{aligned}$$

so for small enough ε^k, $\tilde{\sigma}_2^k$ is a best response for all z_2 at each history (because $\mu^0 > \frac{1}{2}$). For player 1,

$$\Delta V_1(t, z_1) = \begin{cases} -(1-\delta) + \varepsilon^k \Delta z_1 & \text{even } t \\ -(1-\delta) + \delta + \varepsilon^k \Delta z_1 & \text{odd } t, \end{cases}$$

so because $\delta > \frac{1}{2}$, for small enough ε^k, $\tilde{\sigma}_1^k$ is a best response for all z_2 at each period. Since $\tilde{\sigma}^k$ gives identical outcomes to σ^*, convergence in outcomes is trivial.

2.3.3 Proof of Proposition 2.5.2

By Proposition 2.3.6, there exists self-generating HBP (ϕ, μ, γ) and enforced HBA (ϕ, μ, α) such that $\gamma(C^K) = V(C^K)$. Suppose by contradiction that $\gamma(I_K) < 2$ (recall that $I_K = \{C^K\}$). This implies that $\alpha_2(I_K) < 1$, for otherwise player 2 would always play c at I_K, yielding $\gamma(I_K) \geq (1-\delta)u(c, C) + \delta V(I_K) = 2(1-\delta) + \delta\gamma(I_K)$ which gives $\gamma(I_K) \geq 2$, a contradiction. Thus d is a best reply for player 2 at I_K, so the payoff of playing d is

$$U_2(d|I_K) \quad = \quad 2\mu(I_K) + (1-\mu(I_K))[2\alpha_1(I_K d) + (1-\alpha_1(I_K d))] \geq U_2(c|I_K)$$

$$= \quad 3\mu(I_K) + (1-\mu(I_K)) \cdot 3\alpha_1(I_K c),$$

where $U_2(c|I_K)$ is the payoff of playing c. Solving for $\alpha_1(I_K c)$ gives

$$\begin{aligned} -\mu(I_K) &\geq (1-\mu(I_K))[3\alpha_1(I_K c) - 2\alpha_1(I_K d) - (1-\alpha_1(I_K d))] \\ &= (1-\mu(I_K))[3\alpha_1(I_K c) - \alpha_1(I_K d) - 1] \end{aligned}$$

$$-\mu(I_K) + (1-\mu(I_K))[\alpha_1(I_K d) + 1] \geq (1-\mu(I_K))3\alpha_1(I_K c)$$

$$\begin{aligned} \alpha_1(I_K c) &\leq \frac{1}{3}\left(\alpha_1(I_K d) + 1 - \frac{\mu(I_K)}{1-\mu(I_K)}\right) \leq \frac{1}{3}\left(2 - \frac{\mu(I_K)}{1-\mu(I_K)}\right) \\ &= \frac{1}{3}\left(\frac{2(1-\mu(I_K)) - \mu(I_K)}{1-\mu(I_K)}\right) = \frac{2-3\mu(I_K)}{3(1-\mu(I_K))}. \end{aligned} \tag{2.3.37}$$

I now show that either $\alpha_1(I_K d) = 0$ or $\alpha_2(I_K) \geq \frac{1+\eta}{2+\lambda}$. Suppose by contradiction that $\alpha_1(I_K d) > 0$ and $\alpha_2(I_K) < \frac{1+\eta}{2+\lambda}$. Since C is a best response at $I_K d$, $\gamma(I_K) \geq \gamma(I_0) + \eta$. Then I have

$$\begin{aligned} \gamma(I_K) &= \alpha_2(I_K)\max\{(1-\delta)u_1(c,C) + \delta\gamma(I_K), (1-\delta)u_1(c,D) + \delta\gamma(I_0)\} \\ &\quad +(1-\alpha_2(I_K))[(1-\delta)u_1(d,C) + \delta\gamma(I_K)] \\ &\leq \alpha_2(I_K)[(1-\delta)u_1(c,D) + \delta\gamma(I_K)] + (1-\alpha_2(I_K))[(1-\delta)u_1(d,C) + \delta\gamma(I_K)] \\ &= (1-\delta)\alpha_2(I_K)(2+\lambda) + \delta\gamma(I_K) \end{aligned}$$

$$\gamma(I_K) \leq \alpha_2(I_K)(2+\lambda) < 1 + \eta \leq \gamma(I_0) + \eta,$$

where the last step is because 1 is the minmax value, so a contradiction.

Bayes' rule gives

$$\mu(I_K) = \frac{\mu^0}{\mu^0 + (1-\mu^0)\phi(I_K)},$$

and substituting this into (2.3.37) yields

$$\alpha_1(I_K c) \leq \frac{2 - 3\frac{\mu^0}{\mu^0 + (1-\mu^0)\phi(I_K)}}{3\left(1 - \frac{\mu^0}{\mu^0+(1-\mu^0)\phi(I_K)}\right)} = \frac{2[\mu^0 + (1-\mu^0)\phi(I_K)] - 3\mu^0}{3\left(\mu^0 + (1-\mu^0)\phi(I_K) - \mu^0\right)}$$

$$= \frac{2(1-\mu^0)\phi(I_K)-\mu^0}{3(1-\mu^0)\phi(I_K)} = \frac{1}{3}\left(2-\frac{\mu^0/(1-\mu^0)}{\phi(I_K)}\right). \tag{2.3.38}$$

Inducibility (see (2.3.2)) requires

$$\begin{aligned}
\phi(I_K) &= [\alpha_2(I_K)\alpha_1(I_K c)+(1-\alpha_2(I_K))\alpha_1(I_K d)]\phi(I_K) \\
&\quad +[\alpha_2(I_{K-1})\alpha_1(I_{K-1}c)+(1-\alpha_2(I_{K-1}))\alpha_1(I_{K-1}d)]\phi(I_{K-1}) \\
&\le [\alpha_2(I_K)\alpha_1(I_K c)+(1-\alpha_2(I_K))\alpha_1(I_K d)]\phi(I_K)+\phi(I_{K-1})
\end{aligned} \tag{2.3.39}$$

Since $\alpha_1(I_K d)=0$ or $\alpha_2(I_K) \ge \frac{1+\eta}{2+\lambda}$, the term in brackets is bounded from above by

$$\max\left\{\alpha_2(I_K)\alpha_1(I_K c), \frac{1+\eta}{2+\lambda}\alpha_1(I_K c)+\left(1-\frac{1+\eta}{2+\lambda}\right)\right\} \le \frac{1+\eta}{2+\lambda}\alpha_1(I_K c)+\left(1-\frac{1+\eta}{2+\lambda}\right). \tag{2.3.40}$$

Rearranging (2.3.39) and substituting (2.3.40) and (2.3.38) gives

$$\begin{aligned}
\phi(I_{K-1}) &\ge (1-[\alpha_2(I_K)\alpha_1(I_K c)+(1-\alpha_2(I_K)\alpha_1(I_K d)])\phi(I_K) \\
&\ge \left[1-\frac{1+\eta}{2+\lambda}\alpha_1(I_K c)-\left(1-\frac{1+\eta}{2+\lambda}\right)\right]\phi(I_K) \\
&= \frac{1+\eta}{2+\lambda}\left[1-\alpha_1(I_K c)\right]\phi(I_K) \\
&\ge \frac{1+\eta}{2+\lambda}\left[1-\frac{1}{3}\left(2-\frac{\mu^0/(1-\mu^0)}{\phi(I_K)}\right)\right]\phi(I_K) \\
&= \frac{1+\eta}{2+\lambda}\left[\phi(I_K)-\frac{1}{3}\left(2\phi(I_K)-\frac{\mu^0}{1-\mu^0}\right)\right] \\
&\ge \frac{1+\eta}{3(2+\lambda)}\frac{\mu^0}{1-\mu^0}.
\end{aligned}$$

Inducibility also requires that $\frac{1+\eta}{3(2+\lambda)}\frac{\mu^0}{1-\mu^0} \le \phi(I_{K-1}) \le \phi(I_{K-2}) \le \cdots \le \phi(I_0)$. Then I have

$$\sum_{k=0}^{K}\phi(I_k) \ge \frac{1+\eta}{3(2+\lambda)}\frac{\mu^0}{1-\mu^0}K, \tag{2.3.41}$$

so picking $K^* \equiv \left(\frac{1+\eta}{3(2+\lambda)}\frac{\mu^0}{1-\mu^0}\right)^{-1}$, for all $K > K^*$ the right hand side of (2.3.41) is greater than 1, a violation of Definition 2.3.3 and a contradiction.

2.3.4 Proof of Proposition 2.5.3

By Proposition 2.3.6, for any stationary public PBE (σ, μ), there exists a self-generating HBP (ϕ, μ, γ) and enforced HBA (ϕ, μ, α) such that $\sigma_2(h) = \alpha_2(h)$, $\sigma_1(ha_2) = \alpha_1(ha_2)$, and $V(\sigma|h) = \gamma(h)$ for all $h \in Y^K$ and $a_2 \in A_2$. Define $\eta \equiv \frac{1-\delta}{\delta}$. I start with the following useful result.

Lemma 2.3.9. *$\gamma(I_0) = 1$ or $\gamma(I_k) \geq \gamma(I_0) + \eta \min\{\lambda, 1\}$ for all $k \in \{1, ..., K\}$.*

Proof. I write this proof for the $\lambda \leq 1$ case; for the $\lambda > 1$ case, the arguments are the same after replacing "λ" with "1." Suppose by contradiction that $\gamma(I_0) > 1$ and there exists $k \in \{1, ..., K\}$ such that $\gamma(I_k) < \gamma(I_0) + \lambda\eta$. Then at histories ha_2 for $h \in I_{k-1}$ and $a_2 \in \{c, d\}$, C is not a best reply, so $\alpha_1(I_{k-1}a_2) = 0$. At I_{k-1} the belief is $\mu(I_{k-1}) = 0$ since the history contains D. Denote player 2's ex-ante payoff of playing a_2 at a history in $I_{k'}$ as $U_2^{a_2}(I_{k'})$. Then player 2's payoff of playing c is $U_2^c(I_{k-1}) = 0$ while the payoff for d is $U_2^d(I_{k-1}) = 1 > U_2^c(I_{k-1})$, so d is the strict best response: $\alpha_2(I_{k-1}) = 0$. Thus,

$$\gamma(I_{k-1}) = (1-\delta)u_1(d, D) + \delta V(I_0) = (1-\delta) + \delta\gamma(I_0) \leq \gamma(I_0) < \gamma(I_0) + \lambda\eta.$$

By induction $\gamma(I_{k'}) = (1-\delta) + \delta\gamma(I_0)$ for each $k' \in \{0, ..., k-1\}$. Thus, $\gamma(I_0) = (1-\delta) + \delta\gamma(I_0) \implies \gamma(I_0) = 1$, a contradiction. □

Proposition 2.5.2 shows that there exists some $\bar{K}$ such that for all $K > \bar{K}$, $\gamma(C^K) \geq 2$. For the rest of the proof, assume $K > \bar{K}$, and so $\gamma(C^K) \geq 2$.

Suppose $\gamma(I_K) > \gamma(I_0) + \lambda\eta$. Since player 1's best response at $I_K c$ is C, I have $\gamma(I_K) = (1-\delta)u_1(c, C) + \delta\gamma(I_K) = 2(1-\delta) + \delta\gamma(I_K)$, which implies $\gamma(I_K) = 2$. Then player 1's best response at history $C^{K-1}c$ (at period $K-1$) is C, so player 2's best response at C^{K-1} is c: $V(C^{K-1}) = (1-\delta)u(c, C) + \delta\gamma(I_K) = 2$. By backwards induction, $V(C^k) = 2$ for all $k \in \{0, ..., K-1\}$, so $V(\varnothing) = 2$.

Now suppose $\gamma(I_K) < \gamma(I_0) + \lambda\eta$. I consider the $\lambda \leq 1$ and $\lambda > 1$ cases separately, reaching contradictions in both for large enough K.

Case 1. $\lambda \leq 1$: Lemma 2.3.9 implies that $\gamma(I_0) = 1$. Since $\delta > \frac{\max\{\lambda,1\}}{1+\max\{\lambda,1\}} = \frac{1}{2}$, I have

$$\gamma(I_0) > \gamma(I_K) - \lambda\frac{1-\delta}{\delta} = 2 - \lambda\frac{1-\delta}{\delta} > 2 - \lambda\frac{1-\frac{1}{2}}{\frac{1}{2}} = 2 - \lambda \geq 1, \quad (2.3.42)$$

a contradiction.

Case 2. $\lambda > 1$: Since $\delta > \frac{\lambda}{1+\lambda}$,

$$\gamma(I_0) > \gamma(I_K) - \lambda\frac{1-\delta}{\delta} \geq 2 - \lambda\frac{1-\delta}{\delta} > 2 - \lambda\frac{1-\frac{\lambda}{1+\lambda}}{\frac{\lambda}{1+\lambda}} = 2 - 1 = 1;$$

then by Lemma 2.3.9, C is a best response for player 1 at $I_k d$ for each $k \in \{1, ..., K\}$. Since D is player 1's strict best response at $I_K c$, it is also the strict best response at $I_{K-1}c$; then player 2's strict best response at I_{K-1} is d: $\alpha_2(I_{K-1}) = 0$. Thus

$$\gamma(I_{K-1}) = (1-\delta)u_1(d, C) + \delta\gamma(I_K) = \delta(I_K). \quad (2.3.43)$$

Since $\gamma(I_{K-1}) < \gamma(I_K) < \gamma(I_0) + \lambda\eta$, $\alpha_1(I_{K-2}c) = 0$ and $\alpha_2(I_{K-2}) = 0$, so by the same reasoning $\gamma(I_{K-2}) = \delta\gamma(I_{K-1}) = \delta^2\gamma(I_K)$. Applying this argument backward yields

$$\gamma(I_0) = \delta^K\gamma(I_K) > \gamma(I_K) - \lambda\frac{1-\delta}{\delta}.$$

Rearranging gives

$$\gamma(I_K) < \lambda\frac{1-\delta}{\delta(1-\delta^K)} < \lambda\frac{1-\frac{\lambda}{1+\lambda}}{\frac{\lambda}{1+\lambda}\left(1-\left(\frac{\lambda}{1+\lambda}\right)^K\right)} = \frac{1}{1-\left(\frac{\lambda}{1+\lambda}\right)^K}.$$

The limit of the right hand side as $K \to \infty$ is 1, so for large enough K I have $\gamma(I_K) < 2$, a contradiction.

For the remainder of the proof suppose that $\gamma(I_K) = \gamma(I_0) + \lambda\eta$. I first show that $\gamma(I_K) = 2$. Suppose not: $\gamma(I_K) > 2$ (recall $\gamma(I_K) \geq 2$ because $K > \bar{K}$). Then

$$\gamma(I_K) \leq (1-\delta)u_1(c, C) + \delta\gamma(I_K) = 2(1-\delta) + \delta\gamma(I_K)$$

which gives $\gamma(I_K) \leq 2$, a contradiction.

Lemma 2.3.10. *Let any $j \in \{0, ..., K-1\}$ be given. Then the following hold:*

1. $\sigma_2(C^j) \geq \alpha_2(I_j)$. *If $\alpha_2(I_j) \in (0,1)$, then $\sigma_2(C^j) = 1$.*

2. $V(C^j) \geq \gamma(I_j)$.

Furthermore, if $\gamma(I_K) = 2$ and $\alpha_2(I_{K-1}) > 0$, then $V(\varnothing) = 2$.

Proof. Since (σ, μ) is purifiable, by Proposition 2.4.2 there exists a sequence $(\psi^k, \varepsilon^k)_k$, such that $\psi^k \in \Delta^*(Z)$ and $\varepsilon^k \to 0$, and a sequence of strategy profiles $(\tilde{\sigma}^k)_k$, such that $\tilde{\sigma}^k$ is a PBE of the (ψ^k, ε^k)-perturbed game, which converges in outcomes to σ. By Proposition 2.4.1, each $\tilde{\sigma}^k$ is essentially sequentially strict and hence quasi-Markov.

Note that the decision problem for player 1 is the same at histories $C^{K-1}a_2$ and $DC^{K-1}a_2$. Thus, for almost all $z_1 \in Z_1$, $\tilde{\sigma}_1^k(C^{K-1}c, z_1) = \tilde{\sigma}_1^k(DC^{K-1}c, z_1)$. I now show that this means player 2 must play c at least as much at C^{K-1} as at DC^{K-1}. Let $\tilde{U}_2(a_2, \tilde{\sigma}_1^k|h, z_2)$ denote player 2's payoff for playing a_2, given player 1 strategy $\tilde{\sigma}_1^k$ at history h and realized shock z_2. Then

$$\begin{aligned}
\Delta U_2^k(h, z_2) &\equiv \tilde{U}_2(c, \tilde{\sigma}_1^k|h, z_2) - \tilde{U}_2(d, \tilde{\sigma}_1^k|h, z_2) \\
&= \int [[\mu^k(h) \cdot 3 + (1-\mu^k(h))(3\tilde{\sigma}_1^k(hc, z_1) + 0 \cdot \tilde{\sigma}_1^k(hc, z_1)) + z_2^c] \\
&\qquad -[\mu^k(h) \cdot 2 + (1-\mu^k(h))(2\tilde{\sigma}_1^k(hd, z_1) + (1-\tilde{\sigma}_1^k(hd, z_1)) + z_2^d]]\, d\psi(z_1) \\
&= \int [\mu^k(h) + (1-\mu^k(h)) \cdot (3\tilde{\sigma}_1^k(hc, z_1) - \tilde{\sigma}_1^k(hd, z_1) - 1) + z_2^c - z_2^d]\, d\psi(z_1) \\
&= \int [\mu^k(h)(1 - (3\tilde{\sigma}_1^k(hc, z_1) - \tilde{\sigma}_1^k(hd, z_1) - 1)) \\
&\qquad +(1-\mu^k(h))(3\tilde{\sigma}_1^k(hc, z_1) - \tilde{\sigma}_1^k(hd, z_1) - 1) + z_2^c - z_2^d]\, d\psi(z_1) \\
&\quad \int [\mu^k(h)(2 - (3\tilde{\sigma}_1^k(hc, z_1) - \tilde{\sigma}_1^k(hd, z_1))) + (3\tilde{\sigma}_1^k(hc, z_1) - \tilde{\sigma}_1^k(hd, z_1) - 1) \\
&\qquad +z_2^c - z_2^d]\, d\psi(z_1).
\end{aligned}$$

Define $\Delta\hat{U}_2^k(h) \equiv \int(2-(3\tilde{\sigma}_1^k(hc,z_1)-\tilde{\sigma}_1^k(hd,z_1)))\mu^k(h)\,d\psi(z_1)$, $\Delta\check{U}_2^k(h) \equiv \int(3\tilde{\sigma}_1^k(hc,z_1) - \tilde{\sigma}_1^k(hd,z_1)-1)\,d\psi(z_1)$ and $\Delta z_2 = z_2^c - z_2^d$, so that

$$\Delta U_2^k(h,z_2) = \Delta\hat{U}_2^k(h) + \Delta\check{U}_2^k(h) + \Delta z_2.$$

Note that $\Delta\check{U}_2^k(C^{K-1}) = \Delta\check{U}_2^k(DC^{K-1})$. Then for any $z_2 \in Z_2$,

$$\begin{aligned}\Delta U_2^k(C^{K-1},z_2) &= \Delta\hat{U}_2^k(C^{K-1}) + \Delta\check{U}_2^k(C^{K-1}) + \Delta z_2 \\ &= \Delta\hat{U}_2^k(C^{K-1}) + \Delta\check{U}_2^k(DC^{K-1}) + \Delta z_2 \\ &= \Delta\hat{U}_2^k(C^{K-1}) + \Delta U_2^k(DC^{K-1},z_2),\end{aligned} \tag{2.3.44}$$

where the last step is because $\mu(DC^{K-1}) = 0 \implies \Delta\hat{U}_2^k(DC^{K-1}) = 0$. Since

$$\Delta U_2^k(C^{K-1},z_2) \geq \Delta U_2^k(DC^{K-1},z_2)$$

for every z_2, $\tilde{\sigma}_2^k(C^{K-1},z_2) \geq \tilde{\sigma}_2^k(DC^{K-1},z_2)$ for almost all z_2, and so $\sigma_2(C^{K-1}) \geq \alpha_2(I_{K-1})$.

Suppose that $\sigma_2(DC^{K-1}) \in (0,1)$, which implies $\sigma_1(DC^{K-1}c) \in (0,1)$ and therefore $\lim_{k\to\infty}\Delta\hat{U}_2^k(C^{K-1}) > 0$. Define

$$\underline{E}_2^k(DC^{K-1}) \equiv \min_{z_2\in Z_2(\varepsilon^k)} \Delta U_2^k(DC^{K-1},z_2) \geq \Delta\check{U}_2^k(DC^{K-1}) - \varepsilon^k$$

$$\bar{E}_2^k(DC^{K-1}) \equiv \max_{z_2\in Z_2(\varepsilon^k)} \Delta U_2^k(DC^{K-1},z_2) \leq \Delta\check{U}_2^k(DC^{K-1}) + \varepsilon^k.$$

There must exist some k^* such that $\underline{E}_2^k(DC^{K-1}) < 0 < \bar{E}_2^k(DC^{K-1})$ for all $k > k^*$, and $\lim_{k\to\infty}\underline{E}_2^k(DC^{K-1}) = \lim_{k\to\infty}\bar{E}_2^k(DC^{K-1}) = 0$; otherwise, $(\tilde{\sigma}^k)_k$ would not converge in outcomes to σ, where player 2 is mixing at DC^{K-1}. Then (2.3.44) gives

$$\lim_{k\to\infty}\underline{E}_2^k(C^{K-1}) \geq \lim_{k\to\infty}\{\Delta U_2^k(C^{K-1}) - \varepsilon^k\} = \lim_{k\to\infty}\{\Delta\hat{U}_2^k(C^{K-1}) + \underline{E}_2^k(DC^{K-1})\}. \tag{2.3.45}$$

Note that because $\sigma_1(C^{K-1}c) \geq \sigma_1(C^{K-1}d)$,

$$\lim_{k\to\infty}\Delta\hat{U}_2^k(C^{K-1}) \;=\; \lim_{k\to\infty}\int(2-(3\tilde{\sigma}_1^k(C^{K-1}c,z_1)-\tilde{\sigma}_1^k(C^{K-1}d,z_1)))\mu^k(C^{K-1})\,d\psi(z_1)$$

$$
\begin{aligned}
&= (2 - (3\sigma_1(C^{K-1}c) - \sigma_1(C^{K-1}d)))\mu(C^{K-1}) \\
&\geq (2 - 2\sigma_1(C^{K-1}c))\mu(C^{K-1}) > 0
\end{aligned}
$$

since $\sigma_1(C^{K-1}c) < 1$. Continuing from (2.3.45) I have

$$
\lim_{k\to\infty} \underline{E}_2^k(C^{K-1}) = \lim_{k\to\infty} \{\Delta\hat{U}_2^k(C^{K-1}) + \underline{E}_2^k(DC^{K-1})\} > \lim_{k\to\infty} \underline{E}_2^k(DC^{K-1}) = 0,
$$

so there exists κ such that for $k > \kappa$, player 2 strictly prefers c for all shocks $z_2 \in Z_2$. Since $\tilde{\sigma}^k \to \sigma$, $\sigma(C^{K-1}) = 1$.

The above arguments imply that $V(\tilde{\sigma}^k|C^{K-1}) \geq V(\tilde{\sigma}^k|I_{K-1})$. Applying the same argument inductively backward proves conditions 1 and 2 in the statement of the lemma for all $j \in \{0, ..., K-1\}$.

Finally, I prove the last statement of the lemma. Suppose $\gamma(I_K) = 2$ and $\alpha_2(I_{K-1}) > 0$. This implies that $\alpha_1(I_{K-1}c) > 0$, i.e. C is a player 1 best response at $I_{K-1}c$. Furthermore, the decision problem facing player 1 at history $C^{K-1}a_2$ is identical to the one at $I_{K-1}a_2$, so $\sigma_1(C^{K-1}a_2) = \alpha_1(I_{K-1}a_2)$. The arguments above prove that $\sigma_2(C^{K-1}) = 1$. Thus, the continuation payoff for C^{K-1} is $V(C^{K-1}) = (1-\delta)u_1(c, C) + \delta V(C^K) = 2(1-\delta) + 2\delta = 2$. This means that the decision problem facing player 1 at $C^{K-2}a_2$ is identical to that at $C^{K-1}a_2$, $C^K a_2$ and $I_{K-1}a_2$, so $\sigma_1(C^{K-2}a_2) = \alpha_1(I_{K-1}a_2)$. Applying essentially the same argument as above shows $\sigma_2(C^{K-2}) = 1$, so $V(C^{K-2}) = 2$. Continuing backwards shows that $V(\varnothing) = 2$. □

I now consider the $\lambda \leq 1$ and $\lambda > 1$ cases separately, concluding the proof.

Case 1. $\lambda \leq 1$: I now show that $\alpha_2(I_0) \in (0, 1)$ (i.e. player 2 mixes the period after a play of D). First, I show that $\alpha_2(I_0) < 1$. Otherwise, $\gamma(I_0) \geq (1-\delta)u_1(c, D) + \delta\gamma(I_0) = 3(1-\delta) + \delta\gamma(I_0)$, which implies $\gamma(I_0) = 3 > \gamma(I_k) - \lambda\eta$, a contradiction. Because $\alpha_2(I_0) < 1$, it must be that player 1 plays D at I_0c sometimes (otherwise d would not be a best response for player 2). If D is a player 1 best response at

I_0c, it is also a best response at I_0d; then $\gamma(I_0) = (1-\delta)[\alpha_2(I_0)u_1(c,D) + (1-\alpha_2(I_0))u_1(d,D)] + \delta\gamma(I_0)$ so

$$\gamma(I_0) = 3\alpha_2(I_0) + (1-\alpha_2(I_0)).$$

It must then be that $\alpha_2(I_0) > 0$ for otherwise $\gamma(I_0) = 1$; but above I have supposed $\gamma(I_0) = \gamma(I_K) - \lambda\eta = 2 - \lambda\eta > 1$, a contradiction. Recall that Lemma 2.3.9 shows C is a player 1 best response at I_0c (since $\gamma(I_0) > 1$). Lemma 2.3.10 implies that because C is a best response at I_0c, it is also a best response at $\varnothing c$ (in period 0): $V(C) \geq \gamma(I_1) \geq \gamma(I_0) + \lambda\eta = 2$. Lemma 2.3.10 also shows that because $\alpha_2(I_0) \in (0,1)$, $\sigma_2(\varnothing) = 1$. Thus, the equilibrium payoff for player 1 is $V(\varnothing) = (1-\delta)u(c,C) + \delta V(C) \geq 2$.

Case 2. $\lambda > 1$: I prove that $\alpha_2(I_{K-1}) > 0$ for large enough K. Suppose by contradiction that $\alpha_2(I_{K-1}) = 0$. Since player 1 is indifferent at $I_{K-1}c$, she strictly prefers C at $I_{K-1}d$, so

$$\gamma(I_{K-1}) = (1-\delta)u_1(d,C) + \delta\gamma(I_K) = \delta\gamma(I_K) < \gamma(I_K).$$

Then player 1 must strictly prefer D at I_{K-2}, so

$$\gamma(I_{K-2}) = (1-\delta)u_1(d,C) + \delta\gamma(I_{K-1}) = \delta\gamma(I_{K-1}) = \delta^2\gamma(I_K) < \gamma(I_K).$$

Applying the same argument backward gives $\gamma(I_0) = \delta^K\gamma(I_K)$, so for K large enough, $\gamma(I_0) < 1$, a contradiction. Second, suppose by contradiction that $\gamma(I_K) > 2$ (recall that Proposition 2.5.2 shows $\gamma(I_K) \geq 2$). Then

$$\gamma(I_{K-1}) \leq (1-\delta)u_1(c,C) + \delta\gamma(I_K) = 2(1-\delta) + \delta\gamma(I_K) < \gamma(I_K).$$

By the same argument leading to (2.3.43), D is then a strict best response at $I_{K-2}c$, so $\gamma(I_{K-2}) = \delta\gamma(I_{K+1}) < \gamma(I_K)$, so by backward induction $\gamma(I_0) = \delta^{K-1}\gamma(I_{K+1})$. For large enough K, $\gamma(I_0) < 1$, a contradiction. Since $\gamma(I_K) = 2$

and $\alpha_2(I_{K-1}) > 0$, Lemma 2.3.10 proves the proposition.

Appendix 3

Proofs for Chapter 3

3.1 Proof of Lemma 3.3.1

Suppose by contradiction that $\check{L}(g,\check{s})$ is nondecreasing in $\check{s}$. Then the term $\frac{\check{L}(g,\check{s})}{\mathfrak{q}(\check{L}(g,\check{s}))}$ is nondecreasing in $\check{s}$, so the square root term in (3.3.1) is strictly increasing. But then the right hand side of (3.3.1) is strictly decreasing, a contradiction.

3.2 Proof of Proposition 3.3.1

First, note that choosing any $q_i \in (0, \eta E[s_{-i}] + g)$ is strictly dominated by choosing q_i. Eliminating these actions, there is a one-to-one mapping between q_i and $\hat{q}_i$, so we can rewrite the profit function in terms of $\hat{q}_i$:

$$\begin{aligned} u_i(a) &= p(a)\hat{q}_i - c(\hat{q}_i + \kappa s_{-i} + g) - \eta s_i \\ &= (r - \hat{q}_i - \hat{q}_{-i})\hat{q}_i - c(\hat{q}_i + \kappa s_{-i} + g) - \eta s_i, \end{aligned}$$

so long as $\hat{q}_i > 0$. If $\hat{q}_i = 0$, then $u_i(a) = -\eta s_i$, so if $\hat{q}_i = 0$ is a best response, then $s_i = 0$ is also a best response.

Suppose that when $\mathcal{Q}$ is sufficiently fine, there is an interior solution (i.e., $\hat{q}_i > 0$ is a best response). $\hat{q}_i$ must satisfy the first order condition

$$0 = \frac{\partial u_i}{\partial \hat{q}_i} - \hat{q}_i + (r - \hat{q}_i - \hat{q}_{-i}) - c$$

$$\hat{q}_i = \tfrac{1}{2}(r - c - \hat{q}_{-i}).$$

Define the function $\mathfrak{q}(\hat{q}_{-i}) \equiv \frac{1}{2}(r - c - \hat{q}_{-i})$. A corner solution ($\hat{q}_i = 0$) is a best response if

$$0 \geq u_i(\mathfrak{q}(\hat{q}_{-i})) = (r - \mathfrak{q}(\hat{q}_{-i}) - \hat{q}_{-i})\mathfrak{q}(\hat{q}_{-i}) - c(\mathfrak{q}(\hat{q}_{-i}) + \kappa E[s_{-i}] + g) - \eta s_i^*$$

where I abuse notation by letting s_i^* denote the mixed attack action conditional on delivering positive quantity. Note that by symmetry, $E[s_{-i}] = E[s_i^*] = \rho s_i^*$ where ρ is the probability of delivering positive quantity. Solving for $\hat{q}_{-i}$ yields

$$0 \geq (r - \tfrac{1}{2}(r - c - \hat{q}_{-i}) - \hat{q}_{-i}) \cdot \tfrac{1}{2}(r - c - \hat{q}_{-i}) - c(\tfrac{1}{2}(r - c - \hat{q}_{-i}) + \kappa E[s_i^*] + g) - \eta s_i^*.$$

$$\hat{q}_{-i} \geq r - c - 2\sqrt{c\kappa E[s_i^*] + \eta s_i^* + cg}.$$

Thus, s^* and g are positive, there exists some threshold for the opposing delivered quantity $\hat{q}_i$ at which choosing $\hat{q}_i = 0$ is a best response. Denote this threshold as $L(g) \equiv r - c - 2\sqrt{c\kappa E[s_i^*] + \eta s_i^* + cg}$.

Define $\hat{q}^{\text{int}}$ as the unique fixed point of $\mathfrak{q}(\cdot)$, which can be solved for as follows:

$$\hat{q}^{\text{int}} = \tfrac{1}{2}(r - c - \hat{q}^{\text{int}})$$

$$\hat{q}^{\text{int}} = \tfrac{1}{3}(r - c).$$

Define $\Delta_s u_i$ as the change in the payoff from choosing $s_i = 1$ over $s_i = 0$:

$$\Delta_s u_i \equiv u_i(s_i = 1, (q, s_{-i})) - u_i(s_i = 0, (q, s_{-i})).$$

Because $E[q_{-i}] - g \geq \kappa$ in any equilibrium, I can write $\Delta_s u_i \equiv \kappa \hat{q}_i - \eta,$. Due to (3.2.1), $\Delta_s u_i > 0$ for $\hat{q}_i = \hat{q}^{\text{int}} > q^m$.

A pure strategy equilibrium is possible under two circumstances when $\mathcal{Q}$ is sufficiently fine:

1. $\hat{q}_i^{\text{int}} \geq L(g)$: Both firms choose quantity delivered $\hat{q}^{\text{int}}$ and to attack, which are mutual best responses by the arguments above.

2. $L(g) = 0$: Both firms choose quantity 0 and to not attack. By the definition of L, when the opposing firm delivers quantity $\hat{q}_{-i} = L$, $\hat{q}_i = 0$ is a best response. It is also clear that $\Delta_s u_i = -\eta < 0$, so $s_i = 0$ is a best response.

The rest of the proof considers the remaining case where $0 < L(g) < \hat{q}^{\text{int}}$. I start by solving for the values of g where this holds:

$$\tfrac{1}{3}(r-c) > r - c - 2\sqrt{c\kappa E[s_i^*] + \eta s_i^* + cg}.$$

Since both firms attack when $L(g) = \hat{q}^{\text{int}}$, I can write

$$\sqrt{c\kappa + \eta + cg} > \tfrac{1}{3}(r-c)$$

$$c\kappa + \eta + cg > \tfrac{1}{9}(r-c)^2$$

$$cg > \tfrac{1}{9}(r-c)^2 - c\kappa + \eta$$

$$g > \frac{1}{c}\left[\tfrac{1}{9}(r-c)^2 + \eta - c\kappa\right]. \tag{3.2.1}$$

For g satisfying (3.2.1) and $L(g) > 0$, there is no pure strategy equilibrium in terms of quantities, so they must be mixing. Since firm i only has multiple best responses in quantities when $E[\hat{q}_{-i}] = L(g)$, this must hold in equilibrium. Hence, (3.3.2) is proven.

Suppose that attacking is a best response for the firms (when producing drugs). Note that

$$\begin{aligned}
\Delta_s u_i &\equiv u_i(s_i = 1, (q, s_{-i})) - u_i(s_i = 0, (q, s_{-i})) \\
&= (r - \hat{q}_i - \hat{q}_{-i} + \min\{\kappa, \hat{q}_{-i}\})\hat{q}_i - c(\hat{q}_i + \kappa E[s_{-i}] + g) - \eta \\
&\quad -[(r - \hat{q}_i - \hat{q}_{-i})\hat{q}_i - c(\hat{q}_i + \kappa E[s_{-i}] + g)] \\
&= \min\{\kappa, \hat{q}_{-i}\}\hat{q}_i - \eta.
\end{aligned}$$

In equilibrium, $\hat{q}_{-i} = L(g)$, and firm i chooses to attack only when also choosing $\hat{q}_i =$

$\mathfrak{q}(L(g))$. Attacking is a best response if

$$0 \leq \Delta_s u_i = \min\{\kappa, L\}\mathfrak{q}(L) - \eta.$$

Consider the following two cases:

Case 1. $L(g) \geq \kappa$: I show that $0 \leq \kappa\mathfrak{q}(L) - \eta$. Suppose not. Then

$$\begin{aligned} 0 &> \kappa\mathfrak{q}(L) - \eta \\ &= \tfrac{1}{2}\kappa(r - c - L) - \eta \\ &\geq \tfrac{1}{2}\kappa(r - c - \kappa) - \eta \\ &> \tfrac{1}{2}\kappa(r - c - \kappa) - \tfrac{1}{4}\kappa(r - c) \\ &= \tfrac{1}{4}\kappa(r - c) - \tfrac{1}{2}\kappa^2 \\ &= \kappa(\tfrac{1}{4}(r - c) - \tfrac{1}{2}\kappa) \end{aligned}$$

so

$$0 > \tfrac{1}{4}(r - c) - \tfrac{1}{2}\kappa$$

$$\kappa > \tfrac{1}{2}(r - c),$$

a contradiction of (3.2.2).

Case 2. $L(g) < \kappa$: Then $0 \leq L\mathfrak{q}(L) - \eta$. Rearranging gives

$$0 \leq L \cdot \tfrac{1}{2}(r - c - L) - \eta$$

$$0 \leq -L^2 + (r - c)L - 2\eta,$$

which holds with equality at

$$\begin{aligned} L &= \frac{-(r - c) \pm \sqrt{(r - c)^2 - 4(-1)(-2\eta)}}{-2} \\ &= \tfrac{1}{2}\left[(r - c) \pm \sqrt{(r - c)^2 - 8\eta}\right]. \end{aligned}$$

The "+" expression is greater than $\frac{1}{3}(r-c) = \hat{q}^{\text{int}}$, so that upper bound is clearly not binding. Focusing on the lower bound, I have

$$L = r - c - 2\sqrt{c\kappa E[s_i^*] + \eta s_i^* + cg} \geq \tfrac{1}{2}\left[(r-c) - \sqrt{(r-c)^2 - 8\eta}\right] \tag{3.2.2}$$

From (3.2.2) we can see that for all g satisfying

$$r - c - 2\sqrt{c\kappa + \eta + cg} \geq \tfrac{1}{2}\left[(r-c) - \sqrt{(r-c)^2 - 8\eta}\right],$$

which can be rearranged to get

$$g \leq \frac{1}{c}\left(\frac{1}{4}\left[(r-c) - \frac{\eta}{\kappa}\right]^2 - (c\kappa + \eta)\right), \tag{3.2.3}$$

$s_i^* = 1$ (i.e. always attack when producing drugs). Thus,

$$\begin{aligned}
E[s_i^*] &= \frac{L}{\mathfrak{q}(L)} = \frac{r - c - 2\sqrt{c\kappa E[s_i^*] + \eta + cg}}{\frac{1}{2}\left(r - c - \left(r - c - 2\sqrt{c\kappa E[s_i^*] + \eta + cg}\right)\right)} \\
&= \frac{r - c - 2\sqrt{c\kappa E[s_i^*] + \eta + cg}}{\sqrt{c\kappa E[s_i^*] + \eta + cg}} \\
&= \frac{r-c}{\sqrt{c\kappa E[s_i^*] + \eta + cg}} - 2
\end{aligned}$$

$$\frac{r-c}{\sqrt{c\kappa E[s_i^*] + \eta + cg}} = 2 + E[s_i^*]$$

$$\frac{(r-c)^2}{c\kappa E[s_i^*] + \eta + cg} = 4 + 4E[s_i^*] + E[s_i^*]^2$$

$$\begin{aligned}
(r-c)^2 &= (4 + 4E[s_i^*] + E[s_i^*]^2)(c\kappa E[s_i^*] + \eta + cg) \\
&= 4c\kappa E[s_i^*] + 4c\kappa E[s_i^*]^2 + c\kappa E[s_i^*]^3 \\
&\quad + 4(\eta + cg) + 4(\eta + cg)E[s_i^*] + (\eta + cg)E[s_i^*]^2
\end{aligned}$$

$$0 = c\kappa E[s_i^*]^3 + 4(c\kappa + \eta + cg)E[s_i^*]^2 + 4(c\kappa + \eta + cg)E[s_i^*] + 4(\eta + cg) - (r-c)^2,$$

so $E[s_i^*] = \xi(g)$. Similar steps to those leading to (3.2.3) imply that for all

$$g \geq \frac{1}{c}\left(\frac{1}{4}\left[(r-c) - \frac{\eta}{\kappa}\right]^2\right), \tag{3.2.4}$$

firms strictly prefer to not attack in equilibrium, $E[s_i^*] = s_i^* = 0$.

I finally turn to the case of g in between (3.2.3) and (3.2.4). Note that $L = \check{L}(g, s_i^*)$. Define $B_0 \equiv \frac{1}{2}\left[(r-c) - \sqrt{(r-c)^2 - 8\eta}\right]$ as the right hand side of (3.2.2). Then (3.2.2) implies that attacking is a strict best response when $\check{L}(g, 1) > B_0$ and that not attacking is a strict best response when $\check{L}(g, 0) < B_0$. If both are best responses, then $\check{L}(g, s_i^*) = B_0$. Thus, (3.3.3) is proven.

3.3 Proof of Proposition 3.3.3

As noted at the beginning of Proposition 3.3.1, hiring any quantity $q_i \in (0, \kappa E[\bar{s}_{-i}] + g)$ is strictly dominated by choosing $q_i = 0$. Thus, I can work simply with delivered quantities $\hat{q}_i = q_i - (\kappa E[\bar{s}_{-i}] + g)$ without loss of generality. The payoff of firm i is

$$V_i = (1-\delta)u_i(\hat{q}, \bar{s}) + \delta\left[\left(1 - F(\bar{\theta})\right)\bar{V} + F(\bar{\theta})\tilde{V}\right].$$

Suppose the best response is an interior solution, i.e. $\hat{q}_i > 0$. Then $\hat{q}_i$ satisfies the first order condition

$$0 = \frac{\partial V_i}{\partial \hat{q}_i} = (1-\delta)\frac{\partial u_i}{\partial \hat{q}_i} + \delta f(\bar{\theta})\frac{\partial \bar{\theta}}{\partial \hat{q}_i}\Delta V.$$

Writing u_i in terms of delivered quantities gives

$$\begin{aligned} u_i &= (r - \hat{q}_i - \hat{q}_{-i})\hat{q}_i - c\hat{q}_i - c(\kappa E[\bar{s}] + g) - \eta E[\bar{s}] \\ &= (r - c - \hat{q}_{-i})\hat{q}_i - \hat{q}_i^2 - c(\kappa E[\bar{s}] + g) - \eta E[\bar{s}] \end{aligned}$$

yielding the derivative

$$\frac{\partial u_i}{\partial \hat{q}_i} = r - c - \hat{q}_{-i} - 2\hat{q}_i. \tag{3.3.1}$$

Note that

$$\bar{\theta} = \frac{\bar{p}}{p(\hat{q})} = \frac{\bar{p}}{r - \hat{q}_i - \hat{q}_{-i}},$$

so its derivative is

$$\frac{\partial \bar{\theta}}{\partial \hat{q}_i} = -\frac{\bar{p}}{p(\hat{q})^2} \frac{\partial \bar{\theta}}{\partial \hat{q}_i} = \frac{\bar{\theta}}{p(\hat{q})}. \tag{3.3.2}$$

Substituting (3.3.1) and (3.3.2) gives

$$0 = (1-\delta)(r - c - \hat{q}_{-i} - 2\hat{q}_i) - \delta f(\bar{\theta}) \frac{\bar{\theta}}{r - \hat{q}_i - \hat{q}_{-i}} \Delta V$$

$$0 = (r - c - \hat{q}_{-i} - 2\hat{q}_i)(r - \hat{q}_i - \hat{q}_{-i}) - \frac{\delta}{1-\delta} f(\bar{\theta}) \bar{\theta} \Delta V.$$

The restriction to symmetric strategies means that $\hat{q}_i = \hat{q}_{-i}$, yielding

$$0 = (r - c - 3\hat{q}_i)(r - 2\hat{q}_i) - \frac{\delta}{1-\delta} f(\bar{\theta}) \bar{\theta} \Delta V,$$

which can be rearranged to

$$0 = 6\bar{q}_i^2 - (5r - 2c)\hat{q}_i + r(r - c) - \frac{\delta}{1-\delta} f(\bar{\theta}) \bar{\theta} \Delta V.$$

Applying the quadratic formula gives

$$\hat{q}_i = \frac{5r - 2c \pm \sqrt{(r+2c)^2 + 24\frac{\delta}{1-\delta} f(\bar{\theta}) \bar{\theta} \Delta V}}{12}.$$

It is straightforward to verify that the "−" solution yields the one-shot equilibrium in Proposition 3.3.1 for $\Delta V = 0$, and so is correct. Thus, (3.3.8) and (3.3.9) are proven.

$\bar{\theta}$ must solve the following maximization problem:

$$\max_{\bar{\theta}} (1-\delta) u_i(\hat{q}(\bar{\theta})) + \delta[(1 - F(\bar{\theta}))\bar{V} + F(\bar{\theta})\tilde{V}] \tag{3.3.3}$$

where I omit the index i from the delivered quantity $\hat{q}(\bar{\theta})$ given by (3.3.9). If the solution is interior ($\bar{\theta} \in (0, \infty)$), it satisfies the first order condition

$$0 = (1 - \delta) \frac{\partial u_i}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial \bar{\theta}} - \delta f(\bar{\theta}) \Delta V \tag{3.3.4}$$

Note that if $\bar{\theta}$ is interior, then $\Delta V > 0$ (otherwise a corner solution would be optimal). Since $\gamma(\bar{\theta})$ is single peaked at $\ln \bar{\theta} = -\frac{1}{2}\zeta^2$, $\hat{q}(\bar{\theta})$ has a single minimum $\bar{\theta} = \exp(-\frac{1}{2}\zeta^2)$. Thus, for any $\bar{\theta}' > \exp(-\frac{1}{2}\zeta^2)$, there exists $\bar{\theta}'' \le \exp(-\frac{1}{2}\zeta^2)$ yielding a higher value for (3.3.3).

Applying symmetric strategies gives

$$u_i(\hat{q}) = (r - 2\hat{q})\hat{q} - c\hat{q} = (r - c - 2\hat{q})\hat{q} = -2\hat{q}^2 + (r - c)\hat{q},$$

which can be differentiated to obtain

$$\frac{\partial u_i}{\partial \hat{q}} = -4\hat{q} + (r - c).$$

Differentiating the delivered quantity function itself obtains $\frac{\partial \hat{q}}{\partial \bar{\theta}} = -\frac{1}{2}(C_1 + \gamma(\bar{\theta})\Delta V)^{-1/2} \Delta V$. Substituting into (3.3.4) gives

$$0 = -(1 - \delta)\left(r - 4\hat{q} - c\right) \frac{\frac{d\gamma}{d\bar{\theta}} \Delta V}{2\sqrt{C_1 + \gamma(\bar{\theta})\Delta V}} - \delta f(\bar{\theta}) \Delta V. \tag{3.3.5}$$

Expanding $d\gamma / d\bar{\theta}$ and simplifying gives

$$\begin{aligned}
\frac{d\gamma}{d\bar{\theta}} = \frac{1}{6}\frac{\delta}{1-\delta}\frac{d}{d\bar{\theta}}[f(\bar{\theta})\bar{\theta}] \quad &= \quad \frac{1}{6}\frac{\delta}{1-\delta}\frac{d}{d\bar{\theta}}\left[\frac{1}{\zeta\sqrt{2\pi}} \exp\left(-\frac{(\log\bar{\theta} + \frac{1}{2}\sigma^2)^2}{2\zeta^2}\right)\right] \\
&= \quad \frac{1}{6}\frac{\delta}{1-\delta}\frac{1}{\zeta\sqrt{2\pi}} \exp\left(-\frac{(\log\bar{\theta} + \frac{1}{2}\zeta^2)^2}{2\zeta^2}\right)\left(-2\frac{\log\bar{\theta} + \frac{1}{2}\zeta^2}{2\zeta^2}\frac{1}{\bar{\theta}}\right) \\
&= \quad -\frac{1}{6}\frac{\delta}{1-\delta}\frac{\log\bar{\theta} + \frac{1}{2}\zeta^2}{\zeta^2} f(\bar{\theta}).
\end{aligned} \tag{3.3.6}$$

Substituting (3.3.6) into (3.3.5) yields

$$0 = \delta\left(4\hat{q}(\bar{\theta}) - (r - c)\right) \frac{-(\log\bar{\theta} + \frac{1}{2}\zeta^2) f(\bar{\theta}) \Delta V}{12\zeta^2\sqrt{C_1 + \gamma(\bar{\theta})\Delta V}} - \delta f(\bar{\theta}) \Delta V$$

$$0 = \big(4\hat{q}(\bar{\theta}) - (r-c)\big)\,\frac{-(\log\bar{\theta} + \frac{1}{2}\zeta^2) f(\bar{\theta})\Delta V}{12\zeta^2\sqrt{C_1 + \gamma(\bar{\theta})\Delta V}} - f(\bar{\theta})\Delta V$$

$$0 = \big(4\hat{q}(\bar{\theta}) - (r-c)\big)\,\frac{-(\log\bar{\theta} + \frac{1}{2}\zeta^2)}{12\zeta^2\sqrt{C_1 + \gamma(\bar{\theta})\Delta V}} - 1, \qquad (3.3.7)$$

thereby proving (3.3.10). Rearranging (3.3.7) gives

$$-(\log\bar{\theta} + \tfrac{1}{2}\zeta^2) = \frac{12\zeta^2\sqrt{C_1 + \gamma(\bar{\theta})\Delta V}}{4\hat{q}(\bar{\theta}) - (r-c)}.$$

Note that because $\gamma(\bar{\theta})$ is single peaked at $\exp(-\frac{1}{2}\zeta^2)$, for all $\bar{\theta} < \exp(-\frac{1}{2}\zeta^2)$, the numerator of the right hand side is increasing in $\bar{\theta}$ while the denominator is decreasing, so the entire right hand side is increasing. The left hand side is clearly strictly decreasing in $\bar{\theta}$, so (3.3.10) has a unique solution.

3.4 Proof of Lemma 3.3.2

Suppose by contradiction that $\bar{\mathscr{B}}(W) \notin \mathscr{B}(\{\bar{W}\})$. Let $V' < \bar{W}$ be some element of W such that $\bar{\mathscr{B}}(W) \in \mathscr{B}(\{V'\})$. Define $\tilde{V}' \equiv (1-\delta^T)u^N + \delta^T V'$ and $\tilde{W} \equiv (1-\delta^T)u^N + \delta^T \bar{W}$. By Definition 3.3.1, there exists $\bar{\alpha} \in \Delta A$ enforced by $\bar{p}$ and V'. Denote $(\bar{q}, \bar{s})$ as the expected actions of $\bar{\alpha}$.

First, suppose that $\bar{p} \in \{0, \infty\}$ (a corner solution). From Corollary 3.3.1, it is straightforward to see that $(1-\delta)\bar{v}^N + \delta\max\{\bar{W}, \tilde{W}\}$ is decomposed by $\bar{\alpha}, \bar{p}, \bar{W}$ (since $\bar{\alpha}$ is the static Nash equilibrium given $\bar{g}$) and that

$$(1-\delta)\bar{v}^N + \delta\max\{\bar{W}, \tilde{W}\} > (1-\delta)\bar{v}^N + \delta\max\{V', \tilde{V}'\} = \bar{\mathscr{B}}(W),$$

a contradiction.

Second, suppose that $\bar{p} \in (0, \infty)$ (an interior solution). From Proposition 3.3.3,

$$\bar{q}_i = C_0 - \sqrt{C_1 + \gamma(\bar{\theta})(V' - \tilde{V})} + \kappa\bar{s}_i + \bar{g}.$$

Since $\bar{W} - \tilde{W} > V' - \tilde{V}'$, there exists $\bar{\theta}_{\bar{W}} < \bar{\theta}$ such that

$$\bar{q}_i = C_0 - \sqrt{C_1 + \gamma(\bar{\theta}_{\bar{W}})(\bar{W} - \tilde{W})} + \kappa \bar{s}_i + \bar{g}.$$

Since attacks are observable, $\bar{s}$ is also enforceable. Thus, $\bar{\alpha}$ is enforced by $\bar{p}_{\bar{W}} \equiv \bar{\theta}_{\bar{W}} p(\bar{\alpha})$ and $\bar{W}$, so they can decompose payoff

$$(1-\delta)u_i(\bar{\alpha})+\delta[(1-F(\bar{\theta}_{\bar{W}}))\bar{W}+F(\bar{\theta}_{\bar{W}})\tilde{W}] > (1-\delta)u_i(\bar{\alpha})+\delta[(1-F(\bar{\theta}))V'+F(\bar{\theta})\tilde{V}'] = \bar{\mathscr{B}}(W),$$

a contradiction.

3.5 Proof of Corollary 3.3.2

Lemma 3.3.2 and part 3 of Proposition 3.3.2 immediately imply (3.3.13). To see that $\bar{\mathscr{E}}$ is the payoff of a T-GPE, construct the following T-GPE: in the reward state, always play action profile $\mathsf{Q}(\bar{\mathscr{E}})$, switching to the punishment state for prices below $\mathsf{U}(\bar{\mathscr{E}})$. By Lemma 3.3.2, $\mathsf{Q}(\bar{\mathscr{E}})$ is enforced by $\mathsf{U}(\bar{\mathscr{E}})$ and $\bar{\mathscr{E}}$. Definition 3.3.1 implies that there are no profitable one-shot deviations in the reward state, and there are clearly no profitable one-shot deviations in the punishment state since firms play the static Nash equilibrium without intertemporal incentives. By the one-shot deviation principle, this T-GPE is an equilibrium, and since $\mathsf{Q}(\bar{\mathscr{E}}), \mathsf{U}(\bar{\mathscr{E}}), \bar{\mathscr{E}}$ decompose $\bar{\mathscr{E}}$, its payoff is $\bar{\mathscr{E}}$.

3.6 Proof of Lemma 3.3.3

Recursively define the sequence $\{V_m\}_m$ as follows: $V_0 \equiv V$ and $V_{m+1} = \bar{\mathscr{B}}(\{V_m\})$. First, suppose there exists m such that $V_{m+1} \leq V_m$. Pick the lowest such m. If $V_{m+1} = V_m$, then $\{V_m\} \subset \bar{\mathscr{B}}(\{V_m\})$ is a self-generating set, so $V_m \geq V_0 = V$ is an equilibrium payoff, and hence $\bar{\mathscr{E}} \geq V_m \geq V$. Now suppose that $V_{m+1} < V_m$. Then $\bar{\mathscr{B}}(\{V_0, ..., V_m\}) = V_m$, yet by Lemma 3.3.2, $\bar{\mathscr{B}}(\{V_0, ..., V_m\}) = \bar{\mathscr{B}}(\max\{V_0, ..., V_m\}) = \bar{\mathscr{B}}(\{V_m\}) = V_{m+1} < V_m$, a contradiction.

Now suppose that $V_{m+1} > V_m$ for all m. It is clear from Definitions 3.3.1 and 3.3.2

that $\{V_m\}_m$ is bounded (to see why, note that from (3.3.6), V_{m+1} is a linear combination of V_m and some feasible payoff $v \leq \bar{\mathscr{F}}$ for all m), so there exists some limit $V^* = \lim_{m\to\infty} V_m$. The following lemma and proof is virtually identical to its analogue in Abreu, Pearce, and Stacchetti (1990), so its proof is omitted (see Lemma 7.3.2 of Mailath and Samuelson (2006)).

Lemma 3.6.1. *If W is compact, then $\mathscr{B}(W)$ is closed.*

Define the set $W \equiv \{V^*, V_0, V_1, ...\}$. By Definition 3.3.2, $\{V_1, V_2, ...\} \subset \mathscr{B}(W)$. Since $\lim_{m\to\infty} V_m = V^*$ and W is compact, Lemma 3.6.1 implies $V^* \in \mathscr{B}(W)$. Since there exist no $\alpha, \bar{p}, V_m$ that decompose V^*, there must be $\alpha, \bar{p}, V^*$ that decompose V^*. Hence $\{V^*\} \subset \mathscr{B}(\{V^*\})$ and so $V^* \in \mathscr{E}$, implying that $V < \bar{\mathscr{E}}$.

Bibliography

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal cartel equilibria with imperfect monitoring," *Journal of Economic Theory*, 39(1), 251 – 269.

——— (1990): "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, pp. 1041–1063.

Aoyagi, M. (1996): "Reputation and Entry Deterrence under Short-Run Ownership of a Firm," *Journal of Economic Theory*, 69(2), 411 – 430.

Apostol, T. M. (1976): *Introduction to Analytic Number Theory*. Springer-Verlag, New York.

Atkeson, A. (1991): "International Lending with Moral Hazard and Risk of Repudiation," *Econometrica*, 59(4), 1069–89.

Banerjee, A., and D. Fudenberg (2004): "Word-of-mouth learning," *Games and Economic Behavior*, 46(1), 1 – 22.

Barlo, M., G. Carmona, and H. Sabourian (2009): "Repeated games with one-memory," *Journal of Economic Theory*, 144(1), 312 – 336.

Bhaskar, V., G. J. Mailath, and S. Morris (2013a): "A Foundation for Markov Equilibria in Sequential Games with Finite Social Memory," *The Review of Economic Studies*, 80(3), 925–948.

——— (2013b): "A Foundation for Markov Equilibria in Sequential Games with Finite Social Memory," *The Review of Economic Studies*, 80(3), 925–948.

Corbae, D., M. B. Stinchcombe, and J. Zeman (2009): *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*. Princeton University Press.

Cripps, M. W., G. J. Mailath, and L. Samuelson (2004): "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72(2), 407–432.

Dell, M. (2011): "Trafficking Networks and the Mexican Drug War," *American Economic Review*, p. Forthcoming.

Doraszelski, U., and J. F. Escobar (2010): "A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification," *Theoretical Economics*, 5(3), 369–402.

——— (2012): "Restricted feedback in long term relationships," *Journal of Economic Theory*, 147(1), 142 – 161.

Ekmekci, M. (2011): "Sustainable reputations with rating systems," *Journal of Economic Theory*, 146(2), 479 – 503.

Ellison, G., and D. Fudenberg (1995): "Word-of-Mouth Communication and Social Learning," *The Quarterly Journal of Economics*, 110(1), 93–125.

Ely, J. C., D. Fudenberg, and D. K. Levine (2008): "When is reputation bad?," *Games and Economic Behavior*, 63, 498–526.

Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73(2), pp. 377–415.

Ely, J. C., and J. Välimäki (2003): "Bad Reputation," *The Quarterly Journal of Economics*, 118, 785–814.

Fudenberg, D., D. M. Kreps, and E. S. Maskin (1990): "Repeated Games with Long-Run and Short-Run Players," *The Review of Economic Studies*, 57(4), pp. 555–573.

Fudenberg, D., and D. K. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57(4), 759–778.

——— (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed," *The Review of Economic Studies*, 59(3), 561–579.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. MIT Press, Cambridge, MA.

Furman, A. (2011): "What is ... a Stationary Measure?," *Notices of the AMS*, 58(9).

Gallager, R. (2012): *Discrete Stochastic Processes*, The Springer International Series in Engineering and Computer Science. Springer US.

Gallup (2013): "Honesty/Ethics in Professions," `http://www.gallup.com/poll/1654/honesty-ethics-professions.aspx` (Accessed Jan. 31, 2014).

Gossner, O. (2011): "Simple Bounds on the Value of a Reputation," *Econometrica*, 79(5), 1627–1641.

Green, E. J., and R. H. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52(1), pp. 87–100.

Grillo, I. (2006): "Mexico Cracking Down on Drug Violence," *Associated Press*.

Harsanyi, J. (1973): "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points," *International Journal of Game Theory*, 2(1), 1–23.

Judd, K. L., S. Yeltekin, and J. Conklin (2003): "Computing Supergame Equilibria," *Econometrica*, 71(4), 1239–1254.

Kreps, D. M., and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

Liu, Q. (2011a): "Information Acquisition and Reputation Dynamics," *The Review of Economic Studies*, 78(4), 1400–1425.

——— (2011b): "Information Acquisition and Reputation Dynamics," *The Review of Economic Studies*, 78(4), 1400–1425.

Liu, Q., and A. Skrzypacz (2014a): “Limited records and reputation bubbles,” *Journal of Economic Theory*, 151(0), 2 – 29.

——— (2014b): “Limited records and reputation bubbles,” *Journal of Economic Theory*, 151(0), 2 – 29.

Mailath, G. J., and W. Olszewski (2011a): “Folk theorems with bounded recall under (almost) perfect monitoring,” *Games and Economic Behavior*, 71(1), 174 – 192.

——— (2011b): “Folk theorems with bounded recall under (almost) perfect monitoring,” *Games and Economic Behavior*, 71(1), 174 – 192, Special Issue In Honor of John Nash.

Mailath, G. J., and L. Samuelson (2006a): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York.

——— (2006b): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York.

Milgrom, P., and J. Roberts (1982): “Predation, Reputation and Entry Deterrence,” *Journal of Economic Theory*, 27, 280–294.

Molzahn, C., O. Rodriguez, and D. A. Shirk (2013): “Drug Violence in Mexico: Data and Analysis Through 2012,” Discussion paper, Trans-Border Institute, University of San Diego.

Monte, D. (2013): “Bounded memory and permanent reputations,” *Journal of Mathematical Economics*, 49(5), 345 – 354.

Phelan, C., and E. Stacchetti (2001): “Sequential Equilibria in a Ramsey Tax Model,” *Econometrica*, 69(6), 1491–1518.

Pitchik, C. (1993): “Commitment, Reputation, and Entry Deterrence,” *Games and Economic Behavior*, 5(2), 268 – 287.

PORTER, R. H. (1983): "Optimal cartel trigger price strategies," *Journal of Economic Theory*, 29(2), 313 – 338.

RIOS, V. (2013): "Why did Mexico become so violent? A self-reinforcing violent equilibrium caused by competition and enforcement," *Trends in Organized Crime*, 16(2), 138–155.

——— (2014): "The Role of Government Coordination in Crime: Deterrence and Citizen Security in Mexico," *Journal of Conflict Resolution*, p. Forthcoming.

SABOURIAN, H. (1998): "Repeated games with M-period bounded memory (pure strategies)," *Journal of Mathematical Economics*, 30(1), 1 – 35.

SELTEN, R. (1978): "The Chain Store Paradox," *Theory and Decision*, 9(2), 127–159.

WISEMAN, T. (2008): "Reputation and impermanent types," *Games and Economic Behavior*, 62(1), 190 – 210.

YELTEKIN, S., Y. CAI, AND K. L. JUDD (2015): "Computing Equilibria of Dynamic Games," Working Paper.