Tulane Economics Working Paper Series

Bounded Memory, Reputation, and Impatience

Benjamin Sperisen
Department of Economics
Tulane University
bsperise@tulane.edu

Working Paper 1602
March 2016

Abstract

Reputation models typically assume players have full memory, yet in many applications this does not hold. This paper studies incomplete information games where players observe only finitely many recent periods, deriving a recursive characterization of the equilibrium payoff set that captures both stationary and previously unexplored non-stationary equilibria, as well as tools for studying purifiable (i.e. robust to payoff perturbations) equilibria. These tools are applied to a product choice game. For 1-period memory, I obtain the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on an "honest" firm type. For long memory, I characterize the minimum purifiable non-stationary equilibrium payoff and unique stationary payoff. In both cases, incomplete information and non-stationary behavior qualitatively change the equilibrium payoff set. These results hold for fixed discount factors independent of prior beliefs, and so do not require extreme patience.

Keywords: Reputation, bounded memory, purifiability, product choice game
JEL codes: C73, D82, D83, L14

Bounded Memory, Reputation, and Impatience

Benjamin Sperisen*

March 6, 2016

Abstract

Reputation models typically assume players have full memory, yet in many applications this does not hold. This paper studies incomplete information games where players observe only finitely many recent periods, deriving a recursive characterization of the equilibrium payoff set that captures both stationary and previously unexplored non-stationary equilibria, as well as tools for studying purifiable (i.e. robust to payoff perturbations) equilibria. These tools are applied to a product choice game. For 1-period memory, I obtain the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on an "honest" firm type. For long memory, I characterize the minimum purifiable non-stationary equilibrium payoff and unique stationary payoff. In both cases, incomplete information and non-stationary behavior qualitatively change the equilibrium payoff set. These results hold for fixed discount factors independent of prior beliefs, and so do not require extreme patience.

JEL Classification: C73; D82; D83; L14
Keywords: Reputation; Bounded Memory; Purifiability; Product Choice Game

1 Introduction

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 in many countries, as are infractions on driving records, and workers typically provide only recent references to prospective employers when applying for jobs. Many online markets display only recent reviews of sellers, or even if the full list is available, make recent reviews ones the most prominent (e.g. eBay); it may be safer to assume buyers simply glance at a few of the latest reviews instead of reading them all. The Court of Justice of the European Union ruled in *Google v. Costeja* that individuals have a "right to be forgotten" and may demand search engines remove links to old information "in light of the time that has elapsed." What happens when agents know that today's behavior will some day be forgotten?

*Department of Economics, Tulane University. Email address: bsperise@tulane.edu. I am especially grateful to Tom Wiseman for his continuous advice and guidance on this paper, as well as to Stefano Barbieri, V. Bhaskar, Uli Doraszelski, Neal Ghosh, Melinda Petre, Dale Stahl, Ina Taneva, Caroline Thomas, Sevin Yeltekin, and seminar participants at Texas, Tulane, and the 11th World Congress of the Econometric Society for their helpful comments.

The first half of this paper introduces general tools to study such environments where players observe only finitely many recent periods. The primary contribution is a recursive characterization of the set of equilibrium payoffs for a general class of bounded memory games with incomplete information, presented in Section 3. This dynamic programming method has two main advantages. First, it allows analysis of stationary as well as previously unexplored non-stationary equilibria. Second, it works for a fixed discount factor, and so does not rely on the classic Fudenberg and Levine (1989, 1992) bounds for very patient players. Section 4 introduces an equilibrium refinement I call "quasi-Markov perfection" (extending 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 paper (Section 5) applies these tools to a product choice game, where a sequence of customers face a firm who is either a normal strategic type or an "honest" Stackelberg commitment type. When customers have 1-period memory of the firm's actions, I obtain a complete characterization of the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on the commitment type. For any discount factor above a fixed threshold, the introduction of a positive prior belief on the commitment type dramatically expands the equilibrium set; non-stationary equilibria also expand the set of equilibrium payoffs for a range of priors. I also study the case of long memory. For any discount factor above a fixed threshold, the Stackelberg payoff is the unique purifiable equilibrium payoff when memory is sufficiently long. However, the minimum non-stationary equilibrium payoff is strictly less, and approaches the minmax payoff as the prior belief approaches zero. These results show that both incomplete information and non-stationary behavior have sharp impact on the set of purifiable equilibria when the firm is not extremely 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, APS characterize the equilibrium set for full memory, while DE characterize it for bounded memory; in both cases, the recursive structure of these games allows the set of equilibrium payoffs to be calculated as the largest fixed point of a "generating operator" that transforms sets of objects containing payoffs, called the largest "self-generating set." However, the assumption of complete information can be restrictive, as even a slight relaxation can dramatically change the equilibria in the product choice game.

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.[1] I show how these games also have a

[1]Note 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; i.e., for complete information, all bounded memory equilibria are also full memory equilibria, but not so for incomplete information.

recursive structure, allowing the construction of an analogous generating operator that, roughly speaking, transforms sets of objects containing both payoffs and probability distributions over beliefs.[2] 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 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] Although this makes the analysis much simpler, there are a variety of real-world applications 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. 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.

To simplify application of the recursive framework, I focus attention on equilibria that are *purifiable* in the sense of Harsanyi (1973), meaning they survive the addition of small, independent (across actions and time) private shocks to the payoffs. For this purpose I define a refinement called *quasi-Markov perfection*. For complete information dynamic games, attention is often restricted to Markov perfect equilibria, where strategies condition only on the payoff relevant "Markov state" instead of payoff-irrelevant history. Quasi-Markov perfection naturally extends this notion to the incomplete information environment,[4] and particularly simplifies games with one commitment type and perfect monitoring of the long-run player's actions. I show that all purifiable equilibria in this class of games (in particular, with sequential-moves) are quasi-Markov, i.e. even very tiny private payoff information destroys all non-quasi-Markov equilibria. This result is an extension of Bhaskar, Mailath, and Morris (2013) (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 argue, purifiability can be motivated by the notion that games are only approximations of reality and so real payoffs are generally not exactly those in the model.

In applying these methods to the repeated sequential-move product choice game, I first consider the case of 1-period memory, which is appealing because it is both the most restrictive possible limit on memory and yields a simple state space in which to apply the recursive algorithm, making

[2] Unlike the full memory case, beliefs themselves are not sufficient.

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

[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 (2006) for an example). Using beliefs as states is too coarse for the results presented here, as discussed in Section 4.

it possible to solve for the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on the honest type.

What is new about these results relative to the literature? The complete information case is well-understood: under full memory cooperation is simple to achieve,[5] yet BMM show that any bound K on memory (no matter how long) is disastrous: all cooperative equilibria are fragile (non-purifiable), leaving the low one-shot equilibrium payoff as the unique purifiable equilibrium payoff.

The analysis of 1-period memory with incomplete information yields several insights. First, an arbitrarily small but positive prior belief on the honest type resurrects non-fragile cooperation even with the shortest possible memory, allowing a purifiable equilibrium with the Stackelberg payoff. Second, allowing non-stationary behavior expands the set of equilibrium payoffs: for a range of priors, the extreme equilibrium payoffs are given by purifiable equilibria where players and beliefs condition on the time in periodic cycles. Third, 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).

There are similarities in the results for long-memory (the limit as $K \to \infty$). Reputation has a particularly sharp effect for stationary equilibria: the firm receives *exactly* the Stackelberg payoff for discount factors above a threshold independent of the prior belief in all purifiable stationary equilibria. This is found under the stationary behavior assumption by solving for the unique self-generating point. Instead, I show how to adapt the recursive machinery to inductively construct self-generating sets in nearby perturbed games to prove the existence of a purifiable equilibrium whose payoff is arbitrarily close to the minimum equilibrium payoff as the memory grows large. The result shows that the stationary assumption is restrictive: the minimum non-stationary purifiable equilibrium is strictly less than the Stackelberg payoff, approaching the one-shot payoff as the prior approaches zero (and then K is made sufficiently large).

For long memory, the arguments from Fudenberg and Levine (1989) give a bound showing that a patient firm receives close to the Stackelberg payoff, by making memory long and then making the firm very patient (i.e., "$\lim_{\delta \to 1} \lim_{K \to \infty}$"). Under stationarity, the Stackelberg payoff is the unique purifiable equilibrium payoff, meaning the Fudenberg and Levine is very slack for a fixed discount factor as the prior approaches zero. By contrast, the bound has much less slack with non-stationary equilibria, as I find that the actual minimum payoff approaches the minmax.

This work relates to a variety of other papers on repeated games and reputation. Besides DE, a number of other papers have extended APS to a variety of settings, e.g. see Atkeson (1991); Phelan and Stacchetti (2001); and Ely, Hörner, and Olszewski (2005). Doraszelski and Escobar (2010) show that stationary Markov perfect equilibria of almost all complete information dynamic stochastic games are purifiable. Other work on bounded memory and complete information under perfect monitoring includes Barlo, Carmona, and Sabourian (2009), who prove a folk theorem for 1-

[5] This is done with grim trigger strategies using arguments from Fudenberg, Kreps, and Maskin (1990).

period memory with rich action sets, and Mailath and Olszewski (2011), who prove a folk theorem for bounded memory strategies. Monte (2013) uses the term "bounded memory" in a different sense, but also shows how limits on learning can lead to qualitatively different equilibria.[6] Ekmekci (2011) shows how restricting learning through a finite rating system can change behavior to avoid the "temporary reputation" result of Cripps, Mailath, and Samuelson (2004). Under full-memory, Faingold and Sannikov (2011) study continuous time reputation games where a large player faces a population of small players, characterizing equilibria for fixed discount factors.

A number of papers study bounded memory with incomplete information focusing on stationary equilibria (and so calendar dates are unobserved). The product choice game application of Section 5 is related to that of Liu (2011) and is closest to Liu and Skrzypacz (2014). Liu (2011) studies behavior in a model where monitoring of past firm actions is endogenous as a costly action available to consumers (instead of a fixed bound), yielding "random auditing" and reputation cycles. Liu and Skrzypacz (2014) have similarly bounded memory with a continuous stage game, finding that all stationary equilibria have interesting dynamics where consumers "ride reputation bubbles" by helping a firm that they know is not honest build reputation to exploit future consumers, and also give a time-independent, patience-based bound on payoffs which has bite for the "$\lim_{\delta\to 1}\lim_{K\to\infty}$" limit discussed above. Sperisen (2015) studies bounded memory in the context of Ely and Välimäki's (2003) mechanic game where reputation is bad.

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 drawn from 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 paragraph). Player 2 observes the K most recent public signals but no other information about past play. Player i receives ex-post payoff $u_i^*(a, y)$, with ex-ante payoffs given by $u_i(a) = \sum_{y\in Y} u_i^*(a, y)\rho(y|a)$. I abuse notation by letting $u_i(\alpha)$ denote the ex ante payoff of mixed action profile $\alpha \in \Delta A$. Figure 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

[6] 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.

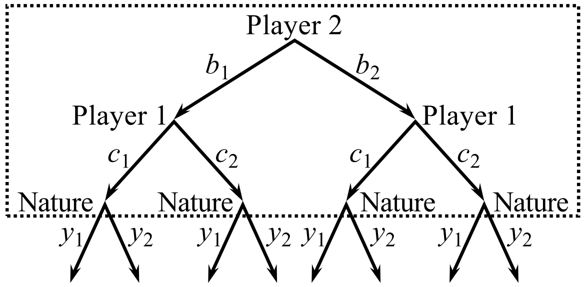

Figure 1: An example stage game with two player 2 actions $A_2 \equiv \{b_1, b_2\}$, two player 1 actions $A_1 \equiv \{c_1, c_2\}$, and two signals $Y \equiv \{y_1, y_2\}$. The dotted box is used for reference in Figure 3, enclosing the part of the game where players move rather than nature.

$\mathcal{H} \equiv \bigcup_{t=0}^{\infty} \mathcal{H}^t$. 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.[7] Player 2 observes only the last K periods, which I call the *public history*. For period t, the set of public histories is $Y^{\min\{t,K\}}$. 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. The focus is primarily on public histories instead of full histories, 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 on the equilibrium path.

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

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

I denote the value of a strategy profile $\sigma \equiv (\sigma_1, \sigma_2)$ to player 1 as $V(\sigma)$, while $V_\sigma(t, h)$ denotes the value (to player 1) of σ at some date-history (t, h) (I sometimes omit the subscript σ when the

[7] I discuss stationary equilibria when the date is hidden in an alternative specification in Section 3.4, where player 2 instead has the improper uniform prior on the date when she enters.

context is clear). Player 2 has beliefs $\mu^* : H \to \Delta\Theta$ where $\mu^*(\theta|t,h)$ is the belief on type $\theta \in \Theta$ at date-history $(t,h) \in H$.

Definition 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.

Because of the focus on payoffs I can and will restrict attention to only public strategies $\sigma_1 : \Theta \times H \times A_2 \to \Delta A_1$ for player 1 as well. 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 $\hat{\Sigma}^*$. Assuming public strategy profiles is not restrictive with respect to payoffs because for any wPBE with strategy profile $\tilde{\sigma} \in \Sigma^*$, there exists public wPBE strategy profile $\bar{\sigma} \in \hat{\Sigma}^*$ such that $V(\bar{\sigma}) = V(\tilde{\sigma})$.[8] For brevity, I define the constant $\phi^* \equiv \frac{\mu^0}{1-\mu^0}$ frequently used in the proofs.

3 Recursive Characterization of Equilibrium Payoffs

While the recursive character of repeated games with full memory and complete information used by APS is more clear, this section shows that bounded memory games with reputation have a more subtle recursive structure that can be similarly exploited. The framework 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 summary statistics of hypothetical future equilibrium play (continuation payoffs) be given.

2. *Enforcement:* Given the hypothetical futures, what current play (action profiles) is possible?

3. *Decomposition:* 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).

If the output of step 3 yields the hypothetical input in step 1, called a "self-generating set," APS show that input describes actual (non-hypothetical) equilibria.

Consider the example model depicted in Figure 1 satisfying the specification in Section 2, omitting reputation — i.e., θ_0 is known — until Section 3.1. The left-hand side of Figure 2 illustrates the recursive structure of the game under full memory. The subtrees in each gray bracket are strategically equivalent, due simply to the fact that the continuation subgame is the same as the full game. Suppose two payoffs v_1, v_2 are known to be in the equilibrium payoff set $\mathscr{E}$. It is possible to construct an equilibrium σ so that $V_\sigma(y_1) = v_1, V_\sigma(y_2) = v_2$ are the continuation payoffs at the

[8] Since player 2's strategy is necessarily public, this follows from a standard argument that if player 2's strategy is public, player 1 has a best response which is also public and yields the same expected payoffs to player 2 for each action. Thus, any equilibrium where player 1's strategy conditions on history not observed by player 2, there exists a public equilibrium with the same payoff.

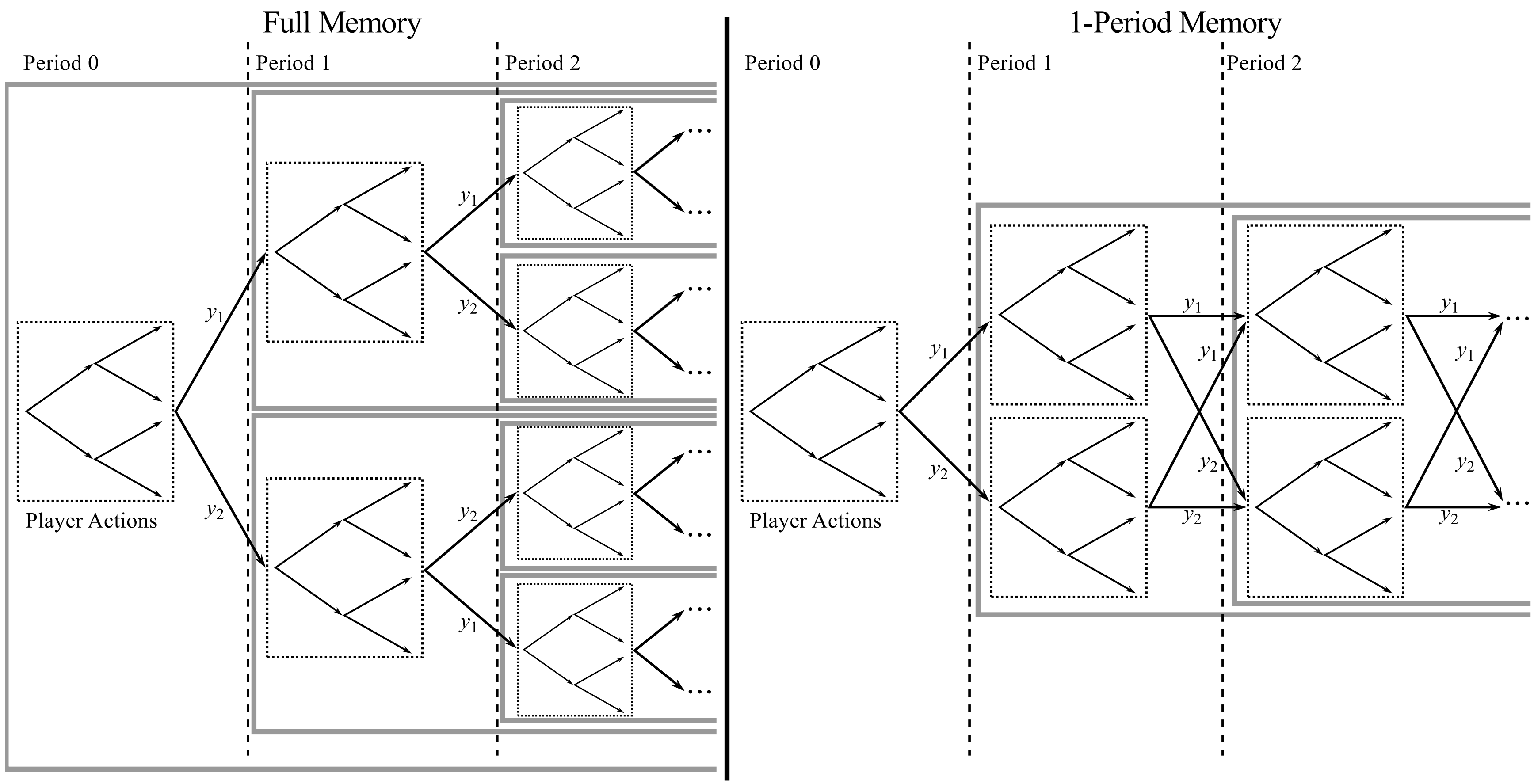

Figure 2: Depiction of the repeated stage game of Figure 1 under full and 1-period memory.

period 1 histories y_1 and y_2, and then finding an action profile α to play in period 0 that is incentive compatible with (*enforced by*) $V_\sigma(y_1), V_\sigma(y_2)$. The average of the payoff of α and the continuation payoffs v_1, v_2 then gives a new equilibrium payoff, $V(\sigma)$.

Bounded memory breaks this strategic equivalence of the full game to the subtree at each full history. The right hand side of Figure 2, depicting 1-period memory, shows how drawing a gray bracket around a subtree starting at a specific full history is no longer possible. In period 1, the path of possible future play starting from history y_1 is interwoven with that starting from history y_2, so constructing equilibria by arbitrarily assigning equilibrium payoffs at each history is no longer possible; even if v_1, v_2 are both equilibrium payoffs, this does not mean an equilibrium σ' exists with these continuation payoffs in the same period, i.e. $V_{\sigma'}(y_1) = v_1, V_{\sigma'}(y_2) = v_2$. Thus, individual payoffs are not sufficient summary statistics of future play to determine enforceable current behavior. For bounded memory strategies, DE give the insight that *payoff functions*, mapping from public histories to payoffs, are sufficient summaries of future play to determine what current play is enforceable,[9] which can then decompose new payoff functions. To get an intuition for this, note that we can draw gray brackets in a recursive way that intersects the tree at each *public history* on the right side of Figure 2. A self-generating set of payoff functions therefore describes actual bounded memory equilibria.

3.1 Recursive Structure

When incomplete information is introduced, bounded memory is no longer simply a restriction of the strategy set. Since the bounded memory affects learning, the game is actually different and the equilibrium set is not simply a subset of the full memory equilibrium payoffs. A primary challenge is to effectively summarize future play in a way that allows consistent updating of beliefs. In particular, unlike full memory, the period $t+1$ short-run player cannot simply update by using period t's posterior because it is forgotten: the period t player saw the outcome of period $t-K$, which is erased upon reaching period $t+1$.

The central insight used by this paper'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.

Construct a modified version of the full game, which I call a *variant game*, where nature randomly picks a fictitious initial history of length K before play begins, but is thereafter identical to the full game (starting at period K). With complete information, this initial history is payoff irrelevant but players may condition on it. Let some equilibrium σ^* of the full game be given, and construct a payoff function $\gamma : Y^K \to \mathbb{R}$ where it gives the continuation payoff for each history h at

[9] DE treat the period 0 empty history as another state to which payoff functions must assign a payoff. In this paper I treat "initial histories" (just period 0 in the 1-period memory case) separately because with reputation, learning is different once memory becomes bounded after the initial histories.

some period $t \geq K$, i.e. $\gamma(h) = V_{\sigma^*}(t,h)$. We can then construct a strategy profile $\tilde{\sigma}$ for the variant game which is identical to σ^* starting at period t, i.e. $\tilde{\sigma}|_{(t',h)} = \sigma^*|_{(t+t',h)}$. Because the variant game after nature's initial choice is strategically equivalent to the full game upon arriving at some history in period t, $\tilde{\sigma}$ is an equilibrium of the variant game with payoffs $V_{\tilde{\sigma}}(0,h) = \gamma(h)$ for each history h. Under complete information, the equilibrium distribution of play at period t in the full game is payoff irrelevant, as is the probability distribution from which nature chooses the fictitious history in the variant game. But when the long-run player's type is unknown, these probabilities are critical because they convey information about the long-run player.

Suppose that the variant game is defined so that nature picks the initial history according the equilibrium distribution of play at some period $t \geq K$ in some equilibrium σ^* of the full game. More formally, let $p_{h^0} \equiv P_{\sigma^*}((t,h^t) = (t,h^0)|\theta)$ denote the probability that the period t public history is h^0 under strategy profile σ^*, conditional on player 1 type θ. In the variant game, nature picks initial history h^0 conditional on type θ with probability p_{h^0}. Then the equilibrium distribution of subsequent play in $\tilde{\sigma}$ will be identical to that of σ^*, shifted t periods earlier. This means that beliefs consistent with $\tilde{\sigma}$ are also consistent with σ^* (again, shifted by t periods). Thus, 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.

3.2 Preliminaries and Variant Games

This section constructs objects that are *hypothetical* summary statistics of an equilibrium — specifically hypothetical probability distributions over public histories, hypothetical beliefs on public histories, and hypothetical continuation payoff functions. I begin by introducing these three main hypothetical "primitives," which are subsequently combined.

Definition 3.1. 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 3.2. 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 $\mu(\theta|h) = \frac{\mu^0(\theta)\phi(h|\theta)}{\sum_{\theta' \in \Theta} \mu^0(\theta)\phi(h|\theta)}$ for all $\theta \in \Theta$. 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 3.3. 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 3.4. 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 3.5. 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 3.3) 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 3.6. 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.

This strategic equivalence is between an equilibrium of full game and one of the variant game is illustrated in Figure 3. Focus first on the left half of the figure. Let some wPBE (σ^*, μ^*) be given

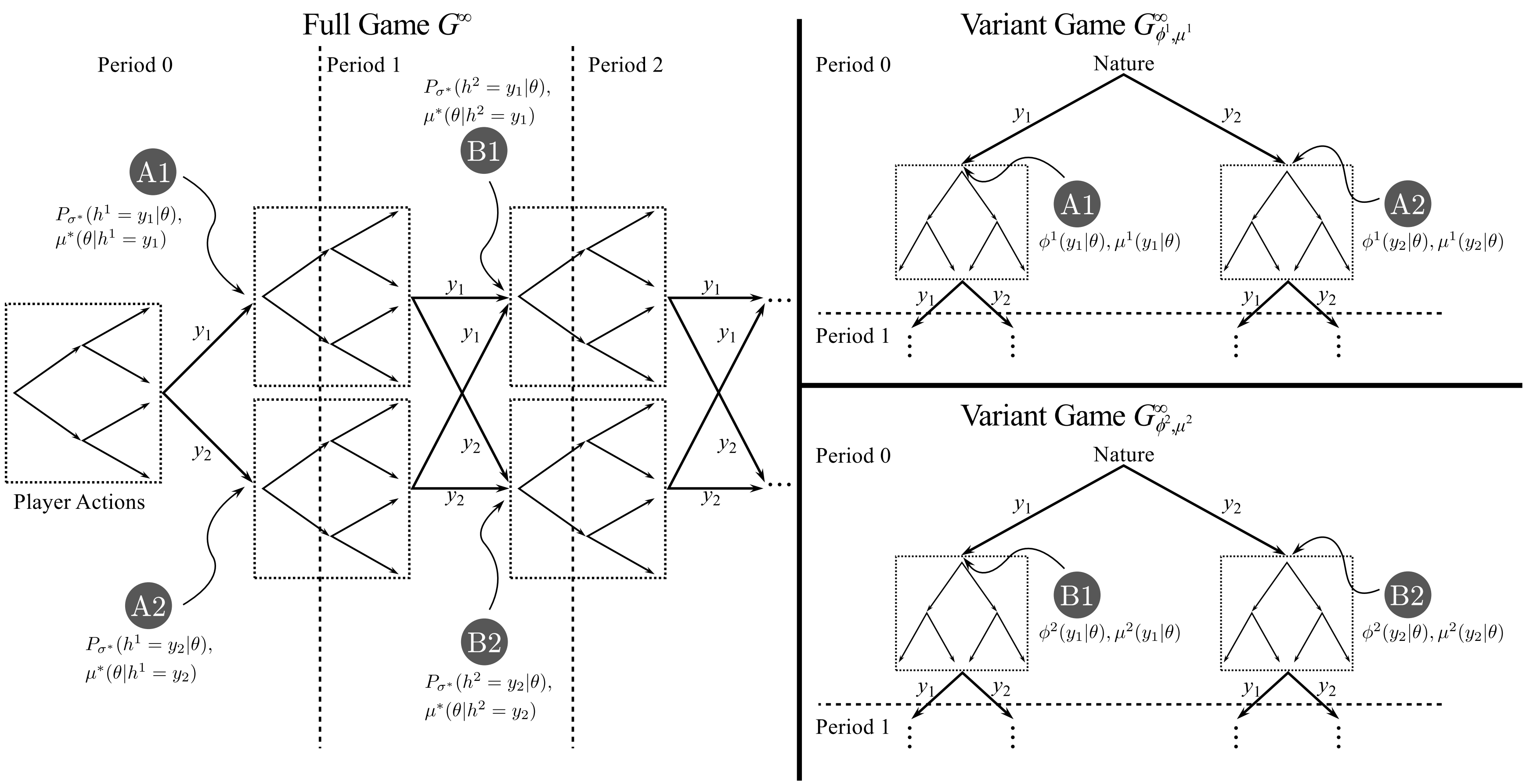

Figure 3: Depiction of the strategic equivalence between the full game and the variance games for 1-period memory ($K = 1$).

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.

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 wPBE 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.[10]

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 for variant game $G^\infty_{\phi,\mu}$, and denote $\Sigma^*_{\phi,\mu}$ as the set of wPBE strategy profiles fore $G^\infty_{\phi,\mu}$ (*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.

Definition 3.7. 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.

[10] 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.

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.

3.3 Self-Generation

This section defines the notions of enforceability, decomposition and self-generation, followed by the main results for Section 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 3.8. 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 3.9. 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 3.10. 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 $\theta \in \Theta$,

$$\tilde{\phi}(h'y|\theta) = \sum_{h' \in \tau(h)} \sum_{a_2 \in A_2} \sum_{y \in Y} \phi(h|\theta)\alpha_2(a_2|h)\alpha_1(a_1|ha_2, \theta)\rho(y|a_2, a_1) \tag{1}$$

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 3.2 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 3.11. 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 3.10).

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

$$\sum_{\theta \in \Theta} \sum_{a_2 \in A_2} u_2(\alpha_1(ha_2, \theta), a_2)\alpha_2(a_2|h)\mu(\theta|h) \geq \sum_{\theta \in \Theta} u_2(\alpha_1(ha'_2, \theta), a'_2)\mu(\theta|h) \tag{2}$$

   for all $a'_2 \in A_2$.

3. **Long-run player incentive compatibility:** player 1 has no profitable deviations: for all

$h \in Y^K$, $a_2 \in A_2$, $a_1' \in A_1$,

$$\begin{aligned} \mathsf{V}_{a_2}(x,\tilde{w})(h) &\equiv (1-\delta)u_1(a_2,\alpha_1(ha_2,\theta_0)) + \delta \sum_{a_1 \in A} \sum_{y \in Y} \alpha_1(a_1|ha_2,\theta_0)\rho(y|a_1,a_2)\tilde{\gamma}(hy) \quad (3) \\ &\geq (1-\delta)u_1(a_1',a_2) + \delta \sum_{y \in Y} \rho(y|a_2,a_1)\tilde{\gamma}(hy) \quad (4) \end{aligned}$$

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 (3) 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: $\mathsf{V}(x,\tilde{w})(h) = \sum_{a_2 \in A_2} \alpha_2(a_2|h)\mathsf{V}_{a_2}(x,\tilde{w})(h)$.

Note. It will also be convenient to use the abuse of notation $\mathsf{V}(v_1,v_2) \equiv (1-\delta)v_1 + \delta v_2$ where $v_1, v_2 \in \mathbb{R}$. For brevity I also use $\mathsf{V}^j(v_1,v_2) \equiv (1-\delta^j)v_1 + \delta^j v_2$.

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 3.12. 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 set operator and "self-generation."

Definition 3.13. 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 3.14. 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, showing that $\mathcal{D}$ is the largest self-generating set. Recall from Section 3.2 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.

Proposition 3.1. *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:*** *$\mathcal{D} = \mathcal{B}(\mathcal{D})$.*

Proofs for this and all other results are in the Appendix.

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 $\mathcal{D}$. This algorithm is effectively carried out by hand for the 1-period memory product choice game in Section 5.1.

Proposition 3.2. *For each $m \in \mathbb{N}$, $\mathcal{B}^{m+1}(\mathcal{F}^\dagger) \subset \mathcal{B}^m(\mathcal{F}^\dagger)$ and $\mathcal{D} = \bigcap_{m=0}^{\infty} \mathcal{B}^m(\mathcal{F}^\dagger)$.*

The difficult part of the analysis is behind us, and can now turn to the ultimate objective of calculating the set $\mathcal{E}$ of wPBE payoffs of the full game. Given $\mathcal{D}$, 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. I call such a finitely repeated game an "antegame" since it represents the first K periods of play in the full game.

Definition 3.15. 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

$$\mathcal{E}_{(\phi,\mu,\gamma)} \equiv \{V(\tilde{\sigma}) : \tilde{\sigma} \in \Sigma_\gamma^*; \forall h \in Y^K, P_{\tilde{\sigma}}(h|\theta) = \phi(h|\theta)\} \tag{5}$$

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 3.16. 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 3.3. *For any set of HBPs $W \subset \mathcal{D}$, $\mathcal{V}(W) \subset \mathcal{E}$. Furthermore, $\mathcal{V}(\mathcal{D}) = \mathcal{E}$.*

The full procedure of solving for the equilibrium payoffs can be summarized as follows. 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 3.2). Second, apply the $\mathcal{V}(\cdot)$ operator to $\mathcal{D}$, yielding the set $\mathcal{E}$ of equilibrium payoffs (as in Proposition 3.3).

3.4 Stationary Equilibria

Previous analysis of bounded memory reputation environments has generally restricted attention to equilibria with strategies that are stationary, i.e. not conditioning 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 $h^t \simeq \tilde{h}^{t'}$. 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 $\tilde{h}^{t'} \simeq h^t$ 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 3.17. 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 $h^t \simeq \tilde{h}^{t'}$, $\sigma_2(h^t) = \sigma_2(\tilde{h}^{t'})$ 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:

$$\mu(h|\theta) = \frac{\mu^0(\theta) \lim_{t\to\infty} \left\{ \frac{1}{t-K} \sum_{s=K}^{t} P^t_\sigma(h|\theta) \right\}}{\sum_{\theta' \in \Theta} \mu^0(\theta') \lim_{t\to\infty} \left\{ \frac{1}{t-K} \sum_{s=K}^{t} P^t_\sigma(h|\theta') \right\}} \tag{6}$$

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. Combining the "time-average" history distribution ϕ (given by the limit terms in (6)) 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 3.4. *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* $\theta \in \Theta$. *Furthermore, if HBP* $w' \equiv (\phi', \mu', \gamma')$ *satisfies* $w' \in \mathcal{B}(\{w'\})$, *then each* $v \in \mathcal{V}(\{w'\})$ *is the payoff of some stationary wPBE.*

This gives stationary equilibria in the game with unobserved calendar dates a very simple interpretation within the framework outlined in Section 3.3. 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*.

4 Purifiability and Quasi-Markov Equilibria

Even when the HBP space has few dimensions, carrying out the algorithm presented above (Propositions 3.2 and 3.3) 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 $\tilde{\phi}$.[11] Given even a singleton set of HBPs $\tilde{W} \equiv \{\tilde{w}\}$, finding the set of decomposed HBPs $\mathcal{B}(\tilde{W})$ means solving the above problem for every possible HB (ϕ, μ) (of which there are infinitely many).[12]

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 applications like that of Section 5, where the long-run player's actions are perfectly monitored and a single commitment type. For 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), i.e. they 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

[11] 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 $\tilde{\phi}(hy|\theta)$.

[12] This means solving for every possible probability distribution for nature's move in step 1 of Footnote 11.

bounded memory, all purifiable equilibria are Markov perfect. As BMM point out, since models cannot hope to describe reality perfectly and so at least some private payoff information is always present in reality, 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 5) that the long-run player has two actions $A_1 \equiv \{C, D\}$, 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 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,[13] consider the earlier example. 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, the belief tomorrow (and in fact the whole history) will

[13] I am grateful to V. Bhaskar for sharing private notes demonstrating how purifiability allows a similar simplification in the context of a chain store game.

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 $\{CCC\}$. 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 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 4.1 for almost all shocks.

Definition 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 B.2). The intuition is that the continuity of ψ ensures that a player being indifferent occurs with probability zero.

Proposition 4.1. *Every wPBE of the perturbed game is quasi-Markov perfect.*

The following condition is what BMM call "weak purifiability," which is weaker than the pu-

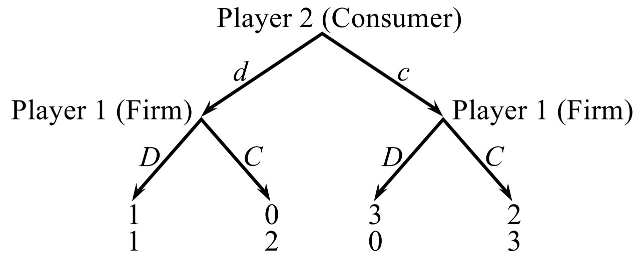

Figure 4: The product choice game, with player 1 (firm) payoffs on top and player 2 (consumer) payoffs on bottom.

rifiability notion of Harsanyi (1973).[14] A sequence of current shock strategies $(\tilde{\sigma}^k)_k$ converges in outcomes to a strategy σ in the unperturbed (full) game if

$$\lim_{k\to\infty} \int \tilde{\sigma}_2^k(a_2|h, z_2)\, d\psi^k(z_2) = \sigma_2(a_2|h) \qquad \text{and} \qquad \lim_{k\to\infty} \int \tilde{\sigma}_1^k(a_1|ha_2', z_1)\, d\psi^k(z_1) = \sigma_1(a_1|ha_2') \tag{7}$$

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

Definition 4.3. A wPBE (σ^*, μ^*) of the full game G^∞ is *purifiable* if there exists a sequence $(\psi^k, \varepsilon^k)_k$ 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.

I can now state the main purifiability result, the incomplete information version of Proposition 2 from BMM, showing that quasi-Markov perfection is implied by purifiability.

Proposition 4.2. *Every purifiable wPBE is quasi-Markov perfect.*

One may wonder to what extent the converse holds: are quasi-Markov equilibria purifiable? Although I do not attempt a general answer, in the 1-period memory application in Section 5, I find the minimum and maximum payoff quasi-Markov equilibria are purifiable for almost all priors.

5 Product Choice Game

This section applies the theoretical machinery of Sections 3 and 4 to the product-choice game[15] in Figure 4. I assume short-run players (consumers) have perfect monitoring of the K most recent long-run player (firm) actions but no monitoring of past short-run player actions.[16] The firm is either normal type θ_0 with payoffs in Figure 4 or "honest" commitment type $\hat{\theta}$ who always plays C ("high quality"). Consumers have prior belief $\mu^0(\hat{\theta})$ on the honest type, abbreviated simply as μ^0.

[14] BMM say 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" 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.

[15] This is the sequential-move version of the product-choice game used throughout Mailath and Samuelson (2006), e.g. Figure 15.1.1.

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

I assume that upon observing any play of D ("low quality") off the equilibrium path, the consumer has belief 0 on the honest type (Bayes' rule clearly implies this on the equilibrium path).

Assumption 1. *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.*

Quasi-Markov perfection then implies that players condition only on the period of the most recent D in the public history and the date.[17] For any history $h \in Y^K$, let

$$\iota(h) \equiv K + \min\{k : h_{1-k} = D\} \tag{8}$$

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". I also extend this notation to the initial histories: for each period $t \in \{0, ..., K-1\}$ and each $k \in \{0, ..., t\}$, let

$$I_k^t \equiv \{h \in H^t : h_{-(k+1)} = D \text{ and } h_{-k'} = C \text{ for all } k' \in \{k, ..., 1\}\}, \tag{9}$$

i.e. I_k^t is the set of all (initial) period t histories such that the k most recent outcomes are all 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. For each k, all the histories in I_k are in the same quasi-Markov state by Definition 4.1. Furthermore, the fact that the cost of effort is constant across the consumer's actions implies that both consumer actions are incentive equivalent; hence, the firm player does not condition on the consumer's action.

Since attention is restricted to quasi-Markov strategies, the recursive tools of Section 3 must be adapted to this restriction, which is straightforward. "Enforceability" as defined in Definition 3.11 is adapted as follows.

Definition 5.1. Let an HBA $x \equiv (\phi, \mu, \alpha)$ enforced by some HBP $\tilde{w}$ be given. x is *quasi-Markov enforced* by $\tilde{w}$ if $\alpha_2(h) = \alpha_2(h')$ and $\alpha_1(hc) = \alpha_1(hd)$ for any two histories $h, h' \in I_k$ for all k.

The analogous definitions of *quasi-Markov decompose* and *quasi-Markov self-generating*, as well as the analogues of Propositions 3.1 and 3.2, 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 $0, ..., K$. Similarly

[17] In their environment, Liu and Skrzypacz (2014) similarly show that strategies in all stationary PBEs, including (possibly) non-purifiable ones, only depend on the time since the last non-commitment action. Proposition 4.2 shows that imposing purifiability also allows such a simplification for all finite stage games with bounded memory and a single long-run player commitment type, as well as to non-stationary equilibria.

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 $\mu \equiv \mu(\hat{\theta}|I_K)$ (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).

5.1 1-Period Memory

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 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 consumers actions $A_2 \equiv \{c, d\}$, I abbreviate $\sigma_2(t, h) \equiv \sigma_2(c|t, h) = 1 - \sigma_2(d|t, h)$. For player 1 we can abbreviate further because quasi-Markov perfection requires that the firm condition only on the 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 the consumer's actions c and d are incentive-equivalent. Thus, 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 $a_2 \in A_2$. I use similar abbreviation for action profile mappings $\alpha \in \mathcal{A}$: $\alpha_2(h) \equiv \alpha_2(c|h), \alpha_1 \equiv \alpha_1(C|ha_2)$. I also use real numbers in $[0, 1]$ interchangeably with these mixed actions.

The results are presented first, followed by discussion of how the algorithm is carried out.

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 $\delta \neq \frac{1}{2}$. The results are more easily understood with the plot in Figure 5.

Proposition 5.1. *Let $\mathcal{E}(\delta, \mu^0)$ be the set of quasi-Markov equilibrium player 1 payoffs of the 1-period*

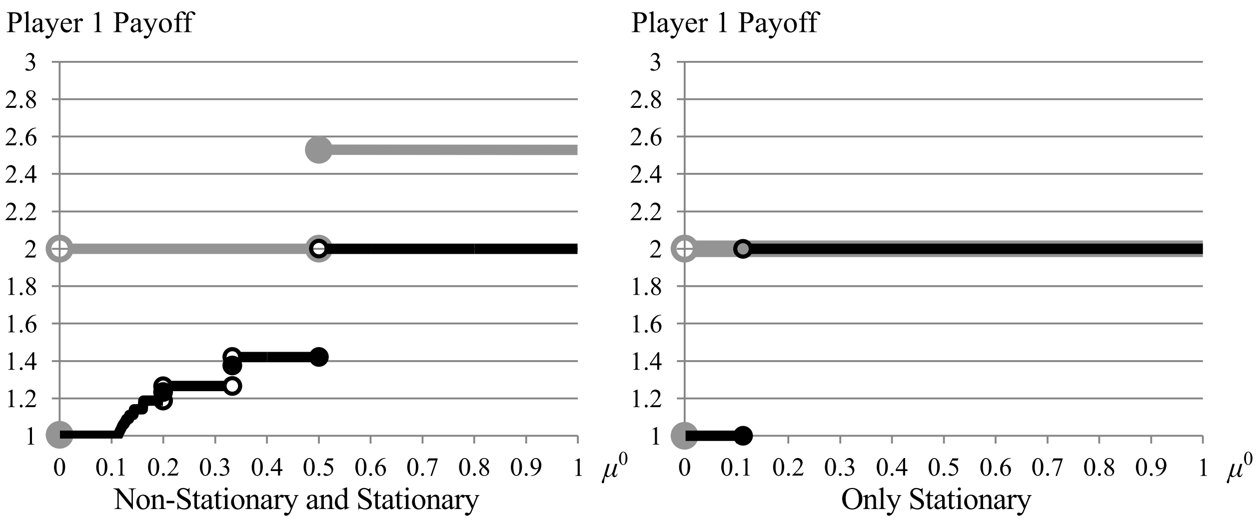

Figure 5: On the left, the minimum (black) and maximum (gray) quasi-Markov equilibrium payoffs are plotted for $\delta = 0.9$. On the right, the minimum and maximum stationary payoffs are plotted.

memory product choice game. For $\delta < \frac{1}{2}$,

$$\min \mathcal{E}(\delta, \mu^0) = \begin{cases} 1 & \mu^0 \leq \frac{1}{2} \\ 3(1-\delta) + \delta & \mu^0 > \frac{1}{2} \end{cases} \qquad \max \mathcal{E}(\delta, \mu^0) = \begin{cases} 1 & \mu^0 < \frac{1}{2} \\ 3(1-\delta) + \delta & \mu^0 \geq \frac{1}{2}. \end{cases} \tag{10}$$

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

$$\min \mathcal{E}(\delta, \mu^0) = \begin{cases} 1 & 0 \leq \mu^0 \leq \frac{1}{9} \\ \lambda^*(\mu^0) & \frac{1}{9} < \mu^0 \leq \frac{1}{2} \\ 2 & \frac{1}{2} < \mu^0 \leq 1 \end{cases} \qquad \max \mathcal{E}(\delta, \mu^0) = \begin{cases} 1 & \mu^0 = 0 \\ 2 & 0 < \mu^0 < \frac{1}{2} \\ \frac{1-\delta}{1-\delta^2}(3+2\delta) & \frac{1}{2} \leq \mu^0 \leq 1, \end{cases} \tag{11}$$

where $\lambda^*(\mu^0)$ *is defined as follows. Let* $q : \mathbb{R} \times [0,1] \to \mathbb{R}$ *be defined by*[18] $q(\phi, \mu^0) = [\mu^0/(1-\mu^0)]/[1-2\phi]$. *Let* $L(\mu^0) = \min_k q^k(0, \mu^0)$ *such that* $q^k(0,\mu^0) \geq \frac{1}{2}$. *Define* $\lambda_1^k \equiv \frac{1}{1-\delta^{k+1}}[(1-\delta^k) + 3\delta^k(1-\delta) - \delta^{k-1}(1-\delta)]$ *for* $k \in \{L-1, L\}$. *Finally,*

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

Recall that Proposition 4.2 proves that quasi-Markov perfection is only a necessary condition of purifiability. In Appendix C.2 I prove that there is a purifiable equilibrium with the minimum and maximum payoffs given in Proposition 5.1 for almost all priors.[19]

Two features of the results are striking. First, the impact of reputation is sharp even when the

[18] Recall that real numbers (in $[0,1]$) and history distributions (in Φ) are used interchangeably.

[19] The Lebesgue measure zero set of priors ignored in Appendix C.2 correspond to the priors with discontinuities in Figure 5.

firm is not very patient. In fact, the discontinuity from $\mu^0 = 0$ to $\mu^0 > 0$ exists everywhere that persuasive intertemporal incentives are even plausible, i.e. any discount factor $\delta > \frac{1}{2}$ (at $\delta < \frac{1}{2}$, even the strongest possible intertemporal incentives are discounted too heavily to affect firm behavior). Second, we can see that non-stationary behavior expands the equilibrium payoff set beyond what is possible with stationary behavior.

In the complete information case ($\mu^0 = 0$), all histories are in the same quasi-Markov state. BMM show that the only purifiable equilibrium is the repeated one-shot 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 a stationary equilibrium where the firm plays C with some probability $\alpha_1 \in [0, \frac{1}{2})$. The lower α_1 is, the more the consumer prefers d conditional on the normal type θ_0; however, a smaller α_1 also implies a higher belief at history C the next period. The minimum payoff stationary equilibrium has the firm mixing such that these countervailing effects balance and the consumer is indifferent upon observing history C. Since the consumer is indifferent at history C (where she has positive belief on the honest type), she strictly prefers d at history D (when the belief is zero and the firm still plays the same mixed action). The consumer plays $\alpha_2(D) = 0$ and mixes with positive probability $\alpha_2(C)$ at history C to keep the firm indifferent.

The maximum payoff stationary equilibrium (payoff of 2) is qualitatively similar, with the firm mixing with probability $\alpha_1 = \frac{1}{2}$ so that the consumer is indifferent at history D, while strictly preferring c at history C. At history D, the consumer mixes so that the firm is indifferent.

For prior $\mu^0 > \frac{1}{9}$, the payoff 1 stationary equilibria are no longer sustainable. If the firm mixes today (period t) with probability $\alpha_1(t)$ such that the consumer 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 consumer indifferent at history C requires the firm 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 the firm 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 the consumer strictly prefers c, at which point the firm plays C with high probability, yielding a low belief next period and restarting the cycle. As the prior increases, this cycle becomes shorter, reaching a high belief more quickly and thus yielding a higher equilibrium payoff. Each discontinuity in the graph is from a step in the cycle being eliminated by a higher prior. By contrast, the high payoff stationary equilibrium survives because the consumer 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 consumer is exploited every other period upon observing history C. The firm always plays D in even periods and C in odd periods. The odd-period consumers clearly have c as a strict best response. The even-period consumers also have c as a strict best response, because the prior is so high that c gives a higher payoff despite the firm playing D with certainty and the fact that the

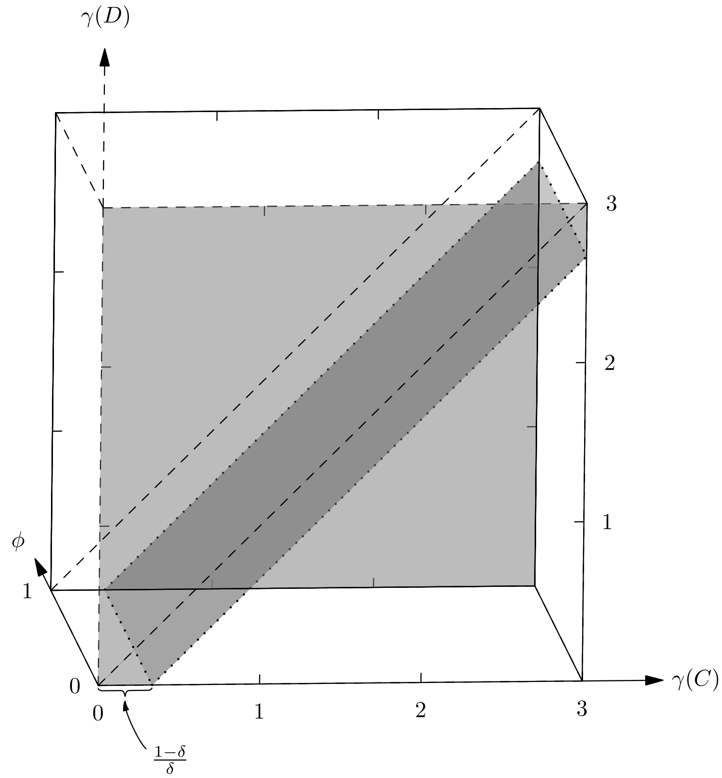

Figure 6: The set $\mathcal{F}^\dagger$ of feasible HBPs (the "cube") and the set of "useful" points $\bar{\mathcal{F}}$ (the gray set).

history is uninformative ($\mu(C) = \mu^0$). This equilibrium gives the maximum payoff for these high priors.

Remark 5.1. 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, Appendix C.3 constructs a set 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 the firm 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. These equilibria feature relatively complicated and arbitrary mixing by both players in order to keep each other indifferent, which is indeed not purifiable.

5.1.2 Algorithm

The starting point of the algorithm of Proposition 3.2 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 6 as a "cube." The algorithm involves repeatedly applying the $\mathcal{B}(\cdot)$ operator to all the points in this 3-dimensional set.

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 6 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 firm is indifferent between C and D.

To the "northwest" of the indifference plane, the firm strictly prefers D. This includes the 45° plane (a line in payoff function space) outlined by the diagonal dashed lines in Figure 6, where there is no intertemporal incentive because the continuation payoffs for both 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 the firm always playing D: $\alpha_1 = 0$. Since the firm'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 $\tilde{w}$: $\mathcal{B}(\{\tilde{w}\}) = \varnothing$. In the northwest, only HBPs on the floor need to be tracked as the algorithm is carried out.

Exactly the opposite occurs in the "southeast" of the indifference plane, where the firm strictly prefers C. Letting $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ in the southeast be given, any enforceable HBA (ϕ, α) must have the firm 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 in the southeast, only the ceiling needs to be tracked.

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 the firm 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 6, called the "useful points." 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 then possible to instead carry out the algorithm with $\bar{\mathcal{B}}(\cdot)$.

What can the useful points actually generate? Figure 7 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" $\tilde{w}^j \equiv (\tilde{\phi}^j, \tilde{\gamma}^j)$, referring to $\tilde{\phi}^j$ as the "enforcing history distribution," and so on. For an HBA (ϕ^j, α^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. Which HBAs $x \equiv (\phi^1, \mu^1, \alpha^1)$ are enforced by $\tilde{w}$? Inducibility requires that the firm always play D (i.e., $\alpha_1^1 = 0$), so if the consumer knows she faces the normal type (like at history D), she will strictly prefer d. At history C, when her beliefs are low (ϕ^1 is high), she also strictly prefers d; when her belief is low (ϕ^1 is high) she strictly prefers c (i.e. $\alpha_1^1 = 1$). Only at a certain threshold belief (and a corresponding threshold for ϕ^1) is the consumer is indifferent so that we may choose any action $\alpha_2^1(C) \in [0, 1]$.

Given an HBA x enforced by $\tilde{w}$, what HBPs $w \equiv (\phi^1, \mu^1, \gamma^1)$ are decomposed by x and $\tilde{w}$? As Figure 7 indicates, if ϕ^1 is off the "consumer indifference threshold" and so the consumer has strict preferences at history C, the resulting HBP either has no intertemporal incentives ($\gamma^1(C) = \gamma^1(D)$) or too much ($\gamma^1(C) > \gamma^1(D) + \frac{1-\delta}{\delta}$), thereby generating HBPs above the floor of the northwest or

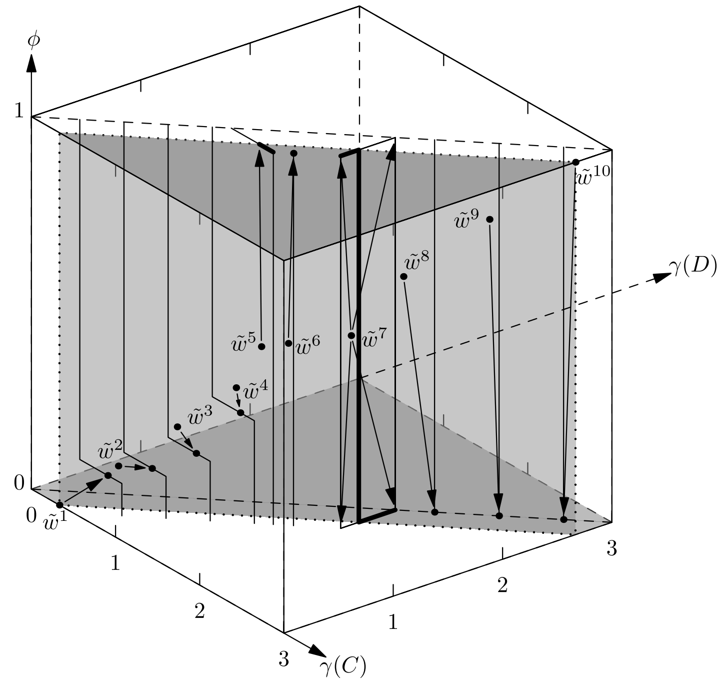

Figure 7: Ten example points $\tilde{w}^1, ..., \tilde{w}^{10}$ on the indifference plane and the sets of points $\mathcal{B}(\{\tilde{w}^1\}), ..., \mathcal{B}(\{\tilde{w}^{10}\})$ they each generate. The useful points $\bar{\mathcal{B}}(\{\tilde{w}^1\}), ..., \mathcal{B}(\{\tilde{w}^{10}\})$ generated by each respectively is bolded. All points at the same elevation generate qualitatively similar sets, so the examples are purposely selected at distinct elevations.

below the ceiling of the southeast — which are useless. Only when ϕ^1 is exactly at this "consumer indifference threshold" is it possible to decompose HBPs with "moderate" intertemporal incentives — including on the indifference plane. This is crucial because the point generated on the indifference plane is the only useful one, meaning we can forget about the rest in the next iteration of the algorithm. Except for HBPs at the "half-way mark" like $\tilde{w}^7$ (see below), all other HBPs generate useful points only at a unique "elevation." I describe why for the other cases shown in Figure 7.

Case 1. The HBPs $\tilde{w}^2, \tilde{w}^3, \tilde{w}^4$ have positive but still low enforcing history distributions: $0 < \tilde{\phi}^2 < \tilde{\phi}^3 < \tilde{\phi}^4$. The set of generated points for each $\tilde{w}^j$ is qualitatively similar but "shifted up" because the firm is playing C with increasing (current) probability, making the consumer indifference threshold ever higher. Again, each of these generates a unique HBP, which is all we track for the next iteration. Given enforcing distribution $\tilde{\phi}$, the proof shows this generated HBP's elevation is given 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 8).

Case 2. The HBP $\tilde{w}^5$ has a high enough enforcing distribution $\tilde{\phi}^5 = q^{-1}(1)$ where player 2 is indifferent at current distribution $\phi = 1$ (i.e. belief μ^0). Because the "lap" of the "wall-sit" figure is in the ceiling of the southeast, the line segment of useful points (bolded) is decomposed. For enforcing distributions strictly between $\tilde{\phi}^5$ and $\frac{1}{2}$ (e.g. $\tilde{w}^6$), a vertical line is decomposed in the southeast whose only useful point is on the ceiling.

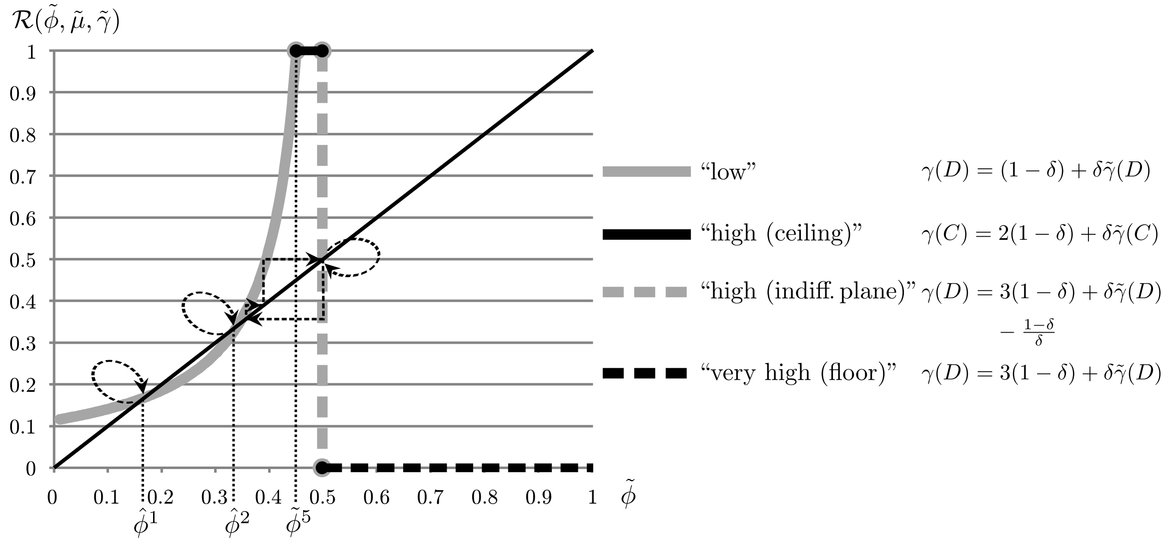

Figure 8: Given an enforcing HBP $(\tilde{\phi}, \tilde{\gamma})$ with history distribution $\tilde{\phi}$ (horizontal axis), I plot the set of history distributions ϕ (vertical axis) for which a point (ϕ, γ) is generated by $\tilde{w}$. The black line is the 45° line. The different line styles describe the generated payoff function γ. The prior is $\mu^0 = \frac{1}{10}$.

Case 3. The enforcing HBP $\tilde{w}^7 = (\tilde{\phi}^7, \tilde{\gamma}^7)$ is at the half-way mark (i.e. $\tilde{\phi}^7 = \frac{1}{2}$), which makes the consumer is indifferent at history D when she knows she is facing the normal type, and so she strictly prefers c when she has positive belief on the honest type (at history C). The consumer's indifference at D allows this HBP to generate the rectangle depicted in Figure 7, whose useful points (intersecting the gray set $\bar{\mathcal{F}}$) are bolded.

Case 4. For enforcing HBPs above the halfway mark like $\tilde{w}^8, \tilde{w}^9, \tilde{w}^{10}$ the consumer strictly prefers c at both histories, which means decomposed HBPs have no intertemporal incentives, so the unique useful generated points are on the 45° line on the floor.

I summarize the set of useful points generated by different HBPs with Figure 8 summarizes the mapping from "enforcing elevation" to "generated elevation" for prior $\mu^0 = \frac{1}{10}$. Roughly speaking, Figure 8 "collapses" the 3-dimensional Figure 7 into two dimensions, showing the enforcing elevation on the horizontal axis and the generated elevation on the vertical axis. This mapping is the correspondence $\mathcal{R}(\tilde{\phi}, \tilde{\gamma}) \equiv \{\phi : (\phi, \gamma) \in \bar{\mathcal{B}}(\{(\tilde{\phi}, \tilde{\gamma})\})\}$. The correspondence is labeled by line style to indicate the payoffs that are decomposed at each elevation.

Proposition 5.1 gives the minimum and maximum equilibrium payoffs. Because $\mathcal{R}(\cdot)$ determines which elevations "plug into each other," it becomes possible to determine the minimum and maximum payoffs generated at each elevation at each iteration. The proof works by then taking the limit and applying the $\mathcal{V}(\cdot)$ operator to the resulting HBPs with the highest and lowest payoffs.

Note that $\mathcal{R}$ crosses the 45° line at three points: $\hat{\phi}^1$, $\hat{\phi}^2$ and $\frac{1}{2}$. These correspond to three stationary equilibria, since stationary equilibria are given by to self-generating singleton HBPs (recall Proposition 3.4). There are two "low" payoff and one "high" payoff stationary equilibria.

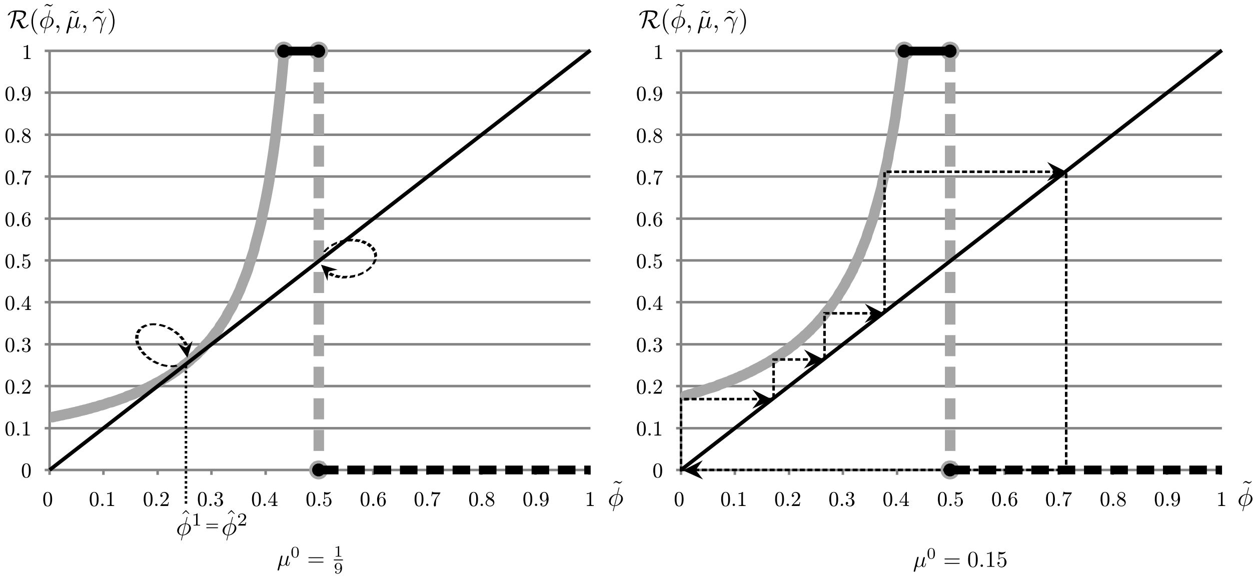

Figure 9: Left, $\mathcal{R}(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ is plotted for $\mu^0 = \frac{1}{9}$, where there is just a single low payoff stationary equilibrium. Right, $\mathcal{R}(\tilde{\phi}, \tilde{\mu}, \tilde{\gamma})$ is plotted for $\mu^0 = 0.15$, where the minimum equilibrium payoff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

Figure 9 shows the effect of the prior μ^0 increasing to $\frac{1}{9}$ and just above. At $\mu^0 = \frac{1}{9}$, the two stationary low payoff equilibria seen in Figure 8 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.

As the prior increases, the number of low payoff steps in the cycle necessarily decreases as the solid gray curve rises, as depicted in Figure 10. At $\mu^0 = \frac{1}{2}$, the solid gray curve has been reduced to a single point at the upper right corner, because $q(0) = 1$. This allows 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 the consumer plays c) and very high payoff of 3 (playing D after the consumer plays c). In this case, the lowest payoff equilibrium is the stationary with payoff 2.

5.2 Long Memory

Two aspects of the 1-period memory results in Section 5.1 stand out. First, incomplete information causes a qualitative change in the (purifiable) equilibrium payoff set even without extreme patience. Second, assuming stationary equilibria is restrictive — relaxing the assumption expands the equilibrium set. How do these results extend if memory is longer?

This section considers the case of long memory, the limit as $K \to \infty$. In fact, both features hold in this limit. The effect of incomplete information is especially stark under stationary strategies: for $\delta > \frac{1}{2}$, the *unique* purifiable equilibrium payoff is 2 (the Stackelberg payoff). However, allowing non-stationary behavior expands the purifiable equilibrium payoff set, allowing payoffs less than the

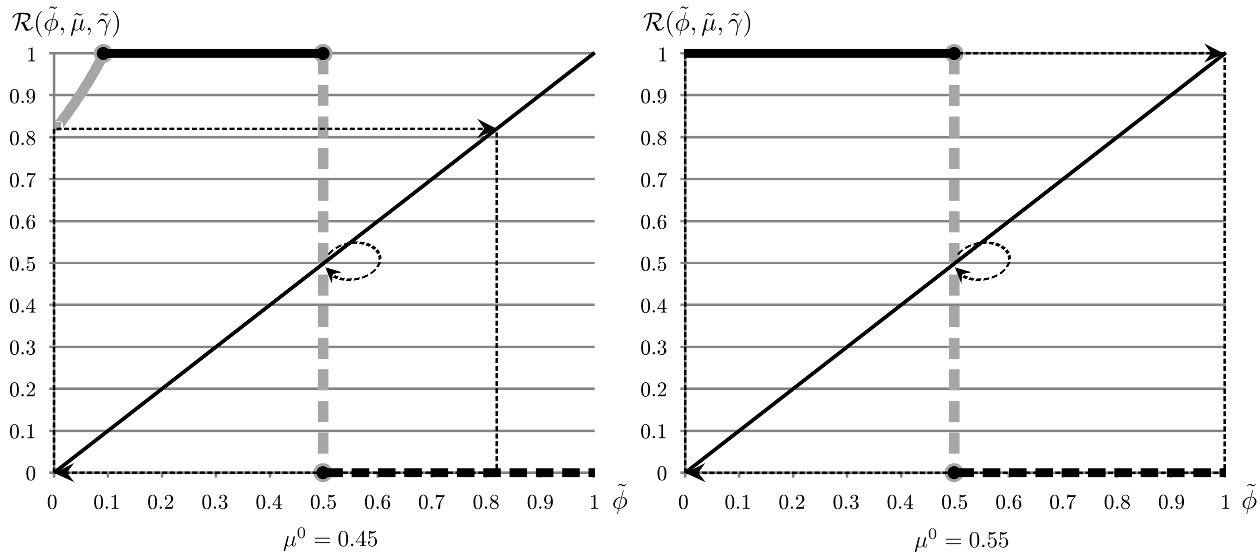

Figure 10: Left, the graph of $\mathcal{R}$ for $\mu^0 = 0.45$, with the minimum payoff non-stationary equilibrium traced out. Right, the graph of $\mathcal{R}$ for $\mu^0 = 0.55$, with the maximum payoff non-stationary equilibrium traced out. In both cases, the stationary equilibrium with payoff 2 is depicted, which corresponds to the maximum equilibrium payoff on the left and the minimum on the right.

Stackelberg payoff: the non-stationary minimum purifiable equilibrium payoff approaches 1 (the one-shot payoff) as the prior belief becomes small.

5.2.1 Stationary Equilibria

The following proposition shows that for stationary equilibria, the effect of reputation is particularly strong: for discount factors above a threshold independent of the prior, the firm gets exactly the Stackelberg payoff when memory is sufficiently long.

Proposition 5.2. *Suppose $\mu^0 > 0$ and $\delta > \frac{1}{2}$. There exists $K^*(\mu^0)$ such that for all $K > K^*$, the player 1 payoff is exactly 2 for any stationary, purifiable wPBE.*

The proof uses Proposition 3.4's simple interpretation of a stationary equilibrium as a singleton self-generating HBP (ϕ, μ, γ). The proof begins with a lemma characterizing this HBP, showing that the payoff at the clean history $\gamma(C^K)$ is at least the Stackelberg payoff (2) for long memory, similar to the bound of Theorem 2 in Liu and Skrzypacz (2014). The intuition is as follows. A payoff $\gamma(I_K) < 2$ requires that the consumer 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 the firm's strategy at the clean history C^K must play D sufficiently frequently. However, playing D with at least probability $\frac{1}{2}$ means the firm is also spending time at the freshly dirty history I_0. The firm must also spend time at histories in $I_1, ..., I_{K-1}$ "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.

It must also be that C is a best response at I_k for every k. To see why, if D is a strict best response at some I_k, then the consumer will always play d at history I_k, which will be followed by the continuation payoff $\gamma(I_0)$ because the firm plays D. This continuation payoff $\gamma(I_k) = (1-\delta)+\delta\gamma(I_0)$ cannot be better than $\gamma(I_0)$, 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 $\gamma(I_0) = 1$ (the one-shot payoff). Yet the fact that $\gamma(I_K) \geq 2$ and $\delta > \frac{1}{2}$ then imply $\gamma(I_{K-1}) \geq 2$; an inductive argument then gives $\gamma(I_k) \geq 2$, a contradiction. Since C must be a best response at every non-initial history, it is easy to then check that $\gamma(I_K) = 2$. These arguments also show that the consumer must weakly prefer c at such histories.

I then use purifiability to characterize play in the initial periods $0, ..., K-1$ (i.e., the antegame). In an equilibrium $\tilde{\sigma}$ of the perturbed game, at any initial history I_k^{K-1} where $k < K-1$, the firm is known to be normal and faces the same decision problem as at non-initial history I_k, and so has the same strict best response at both I_k^{K-1} and I_k for almost all shocks. Similarly, the consumer also has the same strict best responses for almost all shocks, and so the continuation payoffs must be identical: $V_{\tilde{\sigma}}(I_k^{K-1}) = V_{\tilde{\sigma}}(I_k)$. However, at the initial clean history C^{K-1}, although the firm's decision problem is the same as at I_{K-1} where she weakly prefers c, the consumer now has at least the prior belief on the honest type. When the perturbations of payoffs are sufficiently small, the consumer at C^{K-1} strictly prefers c for all shocks, yielding a continuation payoff $V_{\tilde{\sigma}}(C^{K-1}) \geq V_{\tilde{\sigma}}(I_K)$. A backward induction argument shows that $V_{\tilde{\sigma}}(C^0) \geq V_{\tilde{\sigma}}(C^1) \geq \cdots \geq V_{\tilde{\sigma}}(C^{K-1})$ and that the firm plays C with positive probability at each of these clean histories. Making the perturbations small thus shows that any purifiable equilibrium σ must have value $V(\sigma) = 2$.

5.3 Non-Stationary Equilibria

For non-stationary equilibria, as memory grows long the dimensions of the HBP space explodes, making the application of the algorithm performed for 1-period memory difficult. Instead, I establish the existence of a purifiable equilibrium with a low payoff, shown to be approximately the minimum (non-stationary) purifiable equilibrium payoff, in the limit as $K \to \infty$. In particular, for memory long enough the minimum equilibrium payoff is arbitrarily close to a value corresponding to some number $\bar{L}(\mu^0)$ periods of receiving 1 (the minimax payoff), and thereafter receiving continuation payoff $2 - \frac{1-\delta}{\delta}$ (close to the Stackelberg payoff for a patient firm).

Proposition 5.3. *Let $\delta > \frac{1}{2}$ and any $\beta > 0$ be given. For almost all priors such that $0 < \mu^0 < \frac{1}{2}$, there exists positive integer $\bar{L}(\mu^0)$ and K^* such that for all $K > K^*$, $\left|\min \mathcal{E}(\delta, \mu^0, K) - \underline{v}^*(\delta, \mu^0)\right| < \beta$ where $\underline{v}^*(\delta, \mu^0) \equiv (1 - \delta^{\bar{L}(\mu^0)}) + \delta^{\bar{L}(\mu^0)}\left(2 - \frac{1-\delta}{\delta}\right)$ and $\lim_{\mu^0 \to 0} \bar{L}(\mu^0) = \infty$.*

The proof proceeds through the construction of series of lemmas. The first shows that $\underline{v}^*(\delta, \mu^0)$ is a lower bound for $\lim_{K\to\infty} \min \mathcal{E}(\delta, \mu^0, K)$ by characterizing the maximum number of "low payoff" periods possible in the antegame given the prior μ^0.

The recursive machinery of Section 3 is naturally extended to the (ψ, ε)-perturbed game, essentially by adding shocks to the definition of an action profile mapping and then adapting the definitions of (ψ, ε)-enforceable, (ψ, ε)-decomposition, and so on. Thus, a (ψ, ε)-self-generating set of HBPs containing (ϕ, μ, γ) with a "matching" equilibrium $\tilde{\sigma}$ of the (ψ, ε)-perturbed γ-antegame implies the existence of an equilibrium of the full (ψ, ε)-perturbed game.

Another lemma shows that the existence of a *purifiable* equilibrium with some payoff v is implied by the existence of a "special" sequence of (ψ^l, ε^l)-self-generating sets $(W^l)_l$ and matching antegame equilibria $(\tilde{\sigma}^l)_l$ for each (ψ^l, ε^l)-perturbed game whose payoffs $V(\tilde{\sigma}^l) \to v$ as $\varepsilon^l \to 0$. What is special about this sequence is that the antegame game equilibria converge in outcomes, that the payoff functions of the self-generating HBPs converge, and that the HBPs are in a particular set $\tilde{W}$. The $\tilde{W}$ is the set of HBPs (ϕ, μ, γ) such that the payoff function $\gamma(I_k)$ is strictly increasing in k are sufficiently close together, i.e. that $\gamma(I_K) - \gamma(I_0) < 2(1-\delta) + 2\varepsilon$.

The rest of the proof shows that there exists a purifiable equilibrium with the payoff $\underline{v}^*(\delta, \mu^0)$. It is shown that for the (ψ, ε)-perturbed game, an HBP w in the special set $\tilde{W}$ can "reverse-generate" another HBP w' also in $\tilde{W}$, i.e. there exists some $w' \in \tilde{W}$ which (ψ, ε)-decomposes w. Furthermore, the continuity of this reverse-generation function makes it possible to pick $w \equiv (\phi, \gamma)$ so that this infinite sequence of reverse-generated HBPs is bounded. Given the HBP w at the beginning of this infinite sequence, I construct a matching γ-antegame equilibrium with a payoff within an order of ε of $\underline{v}^*(\delta, \mu^0)$. With this bounded self-generating set of HBPs plus a matching antegame equilibrium in hand, the existence of an equilibrium in the perturbed game with the above payoff is established. As constructed, as $\varepsilon^l \to 0$ the antegame equilibria converge in outcomes, and the self-generating set of HBPs have convergent payoff functions, so the payoffs of these perturbed equilibria converge to the payoff of a purifiable equilibrium.

5.3.1 Discussion

By making memory long and then making the firm very patient, the classic Fudenberg and Levine (1989) bound converges to the Stackelberg payoff:

$$\lim_{\delta \to 1} \lim_{\mu^0 \to 0} \lim_{K \to \infty} \min \mathcal{E}(\delta, \mu^0, K) = 2.$$

This works because because the periods where memory is bounded are pushed far into the future where they are discounted away. When the limits for δ and μ^0 are reversed, the Fudenberg and Levine bound has no bite.

For long memory with stationary behavior and a fixed discount factor, reputation has exactly the same effect for a very small prior as it does for a larger one: the purifiable equilibrium payoff is exactly 2. Thus, reversing the limits for δ and μ^0 makes no difference, and there exists a discontinuity from complete information to positive μ^0. By contrast, allowing non-stationary behavior gives a very different result:

$$\lim_{\mu^0 \to 0} \lim_{\delta \to 1} \lim_{K \to \infty} \min \mathcal{E}(\delta, \mu^0, K) = 1.$$

Thus, with non-stationary behavior there is no longer a discontinuity in the minimum purifiable equilibrium payoff going from $\mu^0 = 0$ to $\mu^0 > 0$. (However, the discontinuity in the maximum purifiable equilibrium payoff remains.)

Why does the argument for ruling out low payoff stationary equilibria not apply to the stationary case? The proof for the existence of a low payoff non-stationary equilibrium constructs a special HBP $w \equiv (\phi, \gamma)$ for the perturbed game, corresponding to the distribution of play and continuation payoffs at period K. The payoff function γ is carefully to induce a strict preference for D by the firm at all histories and shocks in period $\bar{L} - 1$ of the perturbed γ-antegame, which in turn allows construction of strategies for the periods $0, ..., \bar{L} - 1$ where the firm chooses D with sufficient frequency to have L periods of low payoffs. An HBP like w cannot generate itself because several iterations of the generating operator leads to a "strict" (over all shocks) preference for D, which the proof of Proposition 5.2 rules out. By contrast, the proof of Proposition 5.3 inductively constructs a sequence of *distinct* HBPs for the perturbed game that constitute a self-generating set, and so represent a non-stationary equilibrium of the perturbed game.

6 Conclusion

This paper lays out theoretical tools for the study of bounded memory reputation games. Extending the self-generation methods of APS and DE to bounded memory reputation, I derive a recursive characterization of the set of equilibrium payoffs, allowing characterization of the entire equilibrium set including non-stationary equilibria, with a simple interpretation for stationary equilibria as self-generating "points" rather than sets. 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 purifiability.

Using these methods to study a product-choice game, for 1-period memory I solve for the exact minimum and maximum purifiable equilibrium payoffs for almost all discount factors and prior beliefs on a Stackelberg commitment type. For long memory, I find that the unique stationary purifiable equilibrium payoff is exactly the Stackelberg payoff. By contrast, I find the approximate non-stationary minimum payoff is lower, showing that it approaches the one-shot payoff as the prior belief approaches zero. The results find that the introduction of the commitment type causes a qualitative change in the purifiable equilibrium set even when the long-run player is not especially patient, as does allowing non-stationary behavior. The product-choice game is a classic but simple setting; an interesting direction for future research is generalizing these results to a broader class of stage games.

Although the recursive framework is specified for the applications examined in this paper, 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 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.

Another avenue for future work is a computational implementation of the algorithm presented here — as done by Judd, Yeltekin, and Conklin (2003) for the original APS algorithm — which would allow study of games too complicated for analytical solutions.

References

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, pp. 1041–1063.

Atkeson, A. (1991): "International Lending with Moral Hazard and Risk of Repudiation," *Econometrica*, 59(4), 1069–89.

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 (2013): "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.

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.

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.

Faingold, E., and Y. Sannikov (2011): "Reputation in Continuous-Time Games," *Econometrica*, 79(3), 773–876.

Feldman, R. M., and C. Valdez-Flores (1996): *Applied Probability and Stochastic Processes*. PWS Publishing Company.

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), pp. 759–778.

——— (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed," *The Review of Economic Studies*, 59(3), pp. 561–579.

Furman, A. (2011): "What is ... a Stationary Measure?," *Notices of the AMS*, 58(9).

Harsanyi, J. C. (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.

Liu, Q. (2011): "Information Acquisition and Reputation Dynamics," *The Review of Economic Studies*, 78(4), 1400–1425.

Liu, Q., and A. Skrzypacz (2014): "Limited records and reputation bubbles," *Journal of Economic Theory*, 151(0), 2 – 29.

Mailath, G. J., and W. Olszewski (2011): "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 (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York.

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.

Sperisen, B. (2015): "Bad Reputation under Bounded and Fading Memory," Working Paper 1527, Tulane University, Department of Economics.

Appendix

A Proofs for Sections 2 and 3

A.1 Proof of Proposition 3.1

Proof of Part 1

The proof is constructive, where possible following the approach of Mailath and Samuelson's (2006) proof (Proposition 7.3.1) of the APS results. 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 $\mathsf{U}(w)$. 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 $\sigma_2(h^t) = \alpha^t_2(h^t)$ and $\sigma_1(h^t a_2) = \alpha^t_1(h^t a_2)$ 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. From above,

$$\begin{aligned} \gamma(h^0) &= \mathsf{V}((\phi, \mu, \mathsf{Q}(w)), \mathsf{U}(w))(h^0) = \mathsf{V}((\phi, \mu, \mathsf{Q}^0(w)), \mathsf{U}^1(w))(h^0) && (12) \\ &= (1-\delta) u_1(\alpha^0(h^0, \theta_0)) + \delta \sum_{a \in A} \sum_{y^0 \in Y} \alpha^0(a|h^0, \theta_0) \rho(y^0|a) \gamma^1(h^0 y^0). && (13) \end{aligned}$$

Substituting $\gamma^1(h^0 y^0) = \mathsf{V}((\phi^1, \mu^1, \alpha^1), \mathsf{U}^2(w))(h^0 y^0)$, from (3) I have

$$\begin{aligned} \gamma(h^0) \;\; = \;\; & (1-\delta) u_1(\sigma(h^0, \theta_0)) + \delta \sum_{a \in A} \sum_{y^0 \in Y} \alpha^0(a|h^0, \theta_0) \rho(y^0|a) \\ & \cdot \left[(1-\delta) u_1(\alpha^1(h^0 y^0, \theta_0)) + \delta \sum_{a \in A} \sum_{y^1 \in Y} \alpha^1(a|h^0 y^0, \theta_0) \rho(y^1|a) \gamma^2(h^0 y^0 y^1) \right] \end{aligned}$$

and continuing to substitute gives

$$\gamma(h^0) = (1-\delta) \sum_{s=0}^{t-1} \delta^s \sum_{\mathfrak{h}^s \in Y^s} P_\sigma(h^0 \mathfrak{h}^s | h_0) u_1(\sigma(h^s, \theta_0)) + \delta^t \sum_{\mathfrak{h}^t \in Y^t} P_\sigma(h^0 \mathfrak{h}^t | h^0, \theta_0) \gamma^t(\mathfrak{h}^t_{t-K}, ..., \mathfrak{h}^t_{t-1})), \tag{14}$$

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

$$\gamma(h^0) = (1-\delta) \sum_{s=0}^{\infty} \delta^s \sum_{h^s \in Y^s} P_\sigma(h^0 \mathfrak{h}^s | h^0, \theta_0) u_1(\sigma(h^0 \mathfrak{h}^s, \theta_0)) = V(\sigma | h^0, \theta_0). \tag{15}$$

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 beliefs that matter are beliefs on types θ. Together, beliefs on each 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 $P_\sigma(h^{t+1}|\theta)$ of history $h^{t+1} = h^t y$ at period $t+1$ is equal to

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

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

$$\mu(\theta|h^t) = \frac{P_\sigma(h^t|\theta)\dot{\mu}^0(\theta)}{\sum_{\theta' \in \Theta} P(h^t|\theta')\dot{\mu}^0(\theta')} = \frac{\phi^t(h^t|\theta)\dot{\mu}^0(\theta)}{\sum_{\theta' \in \Theta} \phi^t(h^t|\theta')\dot{\mu}^0(\theta')} = \mu^t(\theta|h^t), \tag{16}$$

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) in Definition 3.11. 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) immediately implies there is no profitable deviation for the short-run player. The long-run player also has no profitable one-shot deviations due to (4) in Definition 3.11, as is now shown. For any period t, $(\phi^t, \mu^t, \alpha^t)$ is enforced by $(\phi^{t+1}, \mu^{t+1}, \gamma^{t+1})$. Then for all $h^t \in H^t$ and $a_2^t \in A_2$,

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

for all $a_1' \in A_1$. By the same argument as (12) through (15), 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.

Then for all $w \equiv (\phi, \mu, \gamma) \in W$, $\gamma(h^0)$ 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 3.7, 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:

$$\mu^*(\theta|h^t) = \frac{\mu^0(\theta)P_\sigma(h^t|\theta)}{\sum_{\theta'\in\Theta}\mu^0(\theta')P_\sigma(h^t|\theta')} = \frac{\mu^0(\theta)\phi^t(h^t|\theta)}{\sum_{\theta'\in\Theta}\mu^0(\theta')\phi^t(h^t|\theta')}. \tag{17}$$

Define belief mapping μ^t such that $\mu^t(\theta|h^t) = \mu^*(\theta|h^t)$; (17) 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 3.12). Since the period $t+1$ continuation game is strategically equivalent to the $G^{\infty}_{\phi^{t+1},\mu^{t+1}}$ variant game, it is straightforward to show that $w^{t+1} \in \mathcal{D}_{\phi^{t+1},\mu^{t+1}} \subset \mathcal{D}$. Since $w^1 \in \mathcal{D}$ and w^0 is decomposed by w^1, $w^0 \in \mathcal{B}(\mathcal{D})$ and so $\mathcal{D} \subset \mathcal{B}(\mathcal{D})$.

A.2 Proof of Proposition 3.2

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 A.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. 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, 2006).

Lemma A.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 3.1 and Lemma A.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 (2006) to my setting, showing that $\mathcal{F}^\dagger_\infty \subset \mathcal{B}(\mathcal{F}^\dagger_\infty)$, which by Proposition 3.1 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 A.2, $\mathcal{F}^\dagger_\infty$ is closed, so there exists $\varepsilon > 0$ such that $\bar{B}^{\mathcal{W}}_\varepsilon(\tilde{w}^*) \cap \mathcal{F}^\dagger_\infty = \varnothing$, where $\bar{B}^{\mathcal{W}}_\varepsilon(\tilde{w})$ is defined as follows:[20]

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

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 $m > m'$, $\tilde{w}^m \in \bar{B}^{\mathcal{W}}_\varepsilon(\tilde{w}^*)$ and because $\{\mathcal{B}^m(\mathcal{F}^\dagger)\}_{m=0}^{\infty}$ is a decreasing sequence, for any $m'' > m'$,

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

The collection $\{\bar{B}^{\mathcal{W}}_\varepsilon(\tilde{w}^*)\} \bigcup_{m=1}^{\infty} \{\mathcal{B}^m(\mathcal{F}^\dagger)\}$ 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): $\bar{B}^{\mathcal{W}}_\varepsilon(\tilde{w}^*) \cap \mathcal{F}^\dagger_\infty \neq \varnothing$, a contradiction.

[20] 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.

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 3.1, $\mathcal{F}^\dagger_\infty = \mathcal{D}$.

A.3 Proof of Proposition 3.3

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 3.16, there exists HBP $(\phi, \mu, \gamma) \in \mathcal{D}$ such that $v \in \mathcal{E}_{(\phi,\mu,\gamma)}$ and 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 (14)), 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 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 $\sigma^*_2(t, h) = \check{\sigma}^t_2(h)$ 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 $\sigma^*_2(t + K, h) = \alpha^{t+K}_2(h)$ 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 $h' \in Y^K, y \in Y$,

$$\begin{aligned}\phi^{t+1}(h'y|\theta) &= \sum_{h\in\tau(h'y)} \sum_{a_2\in A_2} \sum_{y\in Y} \phi^t(h|\theta)\alpha^t_2(a_2|h)\alpha^t_1(a_1|ha_2, \theta)\rho(y|a_2, a_1)\\ = &\sum_{h\in\tau(h'y)} \sum_{a_2\in A_2} \sum_{y\in Y} P_{\sigma^*}((t, h)|\theta)\sigma^*_2(a_2|t, h)\sigma_1(a_1|t, ha_2, \theta)\rho(y|a_2, a_1) = P_{\sigma^*}((t + 1, h'y)|\theta).\end{aligned}$$

Then by induction, $P_{\sigma^*}((t, h)|\theta) = \phi^t(h|\theta)$ holds true at all $t \geq K$. The direct analogue of (16) shows that $\mu^*(\theta|h)$ is consistent at $t \geq K$, and the analogous steps of (12) through (15) show that $V_{\sigma^*}(t, h) = \gamma^t(h)$. The incentive compatibility requirements of Definition 3.11 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},

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

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 $\mathcal{V}(\mathcal{D}) = \mathcal{E}$ by showing that $\mathcal{E} \subset \mathcal{V}(\mathcal{D})$ (since the above arguments prove $\mathcal{V}(\mathcal{D}) \subset \mathcal{E}$). 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 3.7, $w \in \mathcal{D}_{\phi^K, \mu^K} \subset \mathcal{D}$.

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 $(\check{\sigma}, \check{\mu})$ 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 $v \in \mathcal{V}(\mathcal{D})$.

A.4 Proof of Proposition 3.4

Let any stationary wPBE (σ^*, μ^*) be given. Note that this is a Markov chain on the state space of histories.

Lemma A.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.*

$$\pi_i^\infty \equiv \lim_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t-1} \pi^0 Q^s. \tag{18}$$

for each $i \in \{1, \ldots, n\}$. 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 (18) exists (see Property 2.17 of Feldman and Valdez-Flores, 1996); for periodic states, the limit also clearly exists. Finally,[21] letting $r^t \equiv \frac{1}{t}\sum_{s=0}^{t-1} \pi^0 Q^s$, 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 $\pi^\infty Q = \left(\lim_{t\to\infty} r^t\right) Q = \lim_{t\to\infty} r^t Q = \lim_{t\to\infty} \frac{1}{t}\left(\sum_{s=0}^{t-1} \pi^0 Q^s + \pi^0 Q^t - \pi^0\right) = \lim_{t\to\infty} \frac{1}{t}\sum_{s=0}^{t-1} \pi^0 Q^s = \pi^\infty$. ■

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 (18); 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 A.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 $\gamma(h) \equiv V_{\sigma^*}(h)$, the lack of profitable one-shot deviations in σ^* immediately implies incentive compatibility for both players, so x is enforced by w. Finally,

$$\gamma(h) = \sum_{a_2\in A_2}\sum_{a_1\in A_1} \sigma_2^*(a_2|h)\sigma_1^*(a_1|h)\left[(1-\delta)u_1(a_2,a_1) + \delta\sum_{y\in Y}\rho(y|a_2,a_1)V_{\sigma^*}(hy)\right] = \mathsf{V}(x, w),$$

so w is decomposed by x and itself.

To prove the converse, let some HBP $w' \equiv (\phi', \mu', \gamma')$ and $v \in \mathcal{V}(\{w'\})$ where $w' \in \mathcal{B}(\{w'\})$ be given. By Definition 3.16, there exists γ'-antegame wPBE $\tilde{\sigma}$ where $V(\tilde{\sigma}) = v \in \mathcal{E}_{(\phi',\mu',\gamma')}$; then $P_{\tilde{\sigma}}(h|\theta) = \phi'(h|\theta)$ for all $h \in Y^K, \theta \in \Theta$. There also exists HBA $x \equiv (\phi', \mu', \alpha')$ such that x, w' decompose w'. Construct strategy profile σ for the full game as follows. For each initial period $t < K$ and all $h \in H^t$, let $\sigma_2(h) \equiv \tilde{\sigma}_2(h), \sigma_1(ha_2) \equiv \tilde{\sigma}_1(ha_2)$ for all $a_2 \in A_2$. For all public histories $h \in Y^K$ in later periods, define $\sigma_2(h) \equiv \alpha_2'(h), \sigma_1(h) \equiv \alpha_1'(h)$. It is straightforward that because $\tilde{\sigma}$ is a wPBE of the γ'-antegame, σ is incentive compatible under the same beliefs, and that those beliefs are consistent with play under σ, for initial periods $t < K$. Essentially performing the argument

[21] 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`.

above in reverse, the definition of decomposition implies that ϕ' is a stationary distribution over the set of histories under the Markov chain implied by σ, and so the player 2 beliefs (specified by μ') are consistent with player 1's behavior at periods $t \geq K$. The definition of decomposition easily yields $V_\sigma(h) = \gamma'(h)$ for all $h \in Y^K$, and the fact that x is enforced by w' directly shows there are no profitable deviations in σ for $t \geq K$. Hence, σ is a wPBE of the full game.

B Proofs for Section 4

B.1 Proof of Proposition 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 B.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 B.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 $V_{\sigma_1,\tilde{\sigma}_2}^{t+1}(hya_2^{t+1}z_1^{t+1})$ 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 $\hat{\sigma}_1 \equiv \arg\max_{\sigma_1} V_{\sigma_1,\tilde{\sigma}_2}^{t+1}(hya_2^{t+1}, z_1^{t+1})$ be the best response. The payoff (at period t) of playing action a_1 after player 2 action a_2^t is

$$\begin{aligned}
\tilde{V}^*_{a_1,\tilde{\sigma}_2}(ha_2, z_1^t) \quad \equiv \quad & \sum_{y \in Y} \rho(y|a_2, a_1)\Bigg[(1-\delta)(u_1(a_2, a_1) + \varepsilon z_1^{t,a_1}) && (19)\\
& + \delta \int\int \sum_{a_2} \tilde{\sigma}_2^{t+1}(a_2^{t+1}|hyz_2^{t+1}) V^{t+1}_{\hat{\sigma}_1,\tilde{\sigma}_2}(hya_2^{t+1}, z_1^{t+1})\, d\psi(z_1^{t+1})\, d\psi(z_2^{t+1})\Bigg]. && (20)
\end{aligned}$$

The decision problem is independent of all shocks before period t. For any 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. Then for almost all z_1^t, player 1 has a unique best response and $\tilde{\sigma}_1$ does not condition on $z_1^0, \ldots, z_1^{t-1}$. ■

The following result is an extension of Lemma 1 of BMM.

Lemma B.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 show that sequential strictness implies the same behavior at both 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 $V(\mathfrak{h}^t y^k)$ does not depend on $\mathfrak{h}^t$.

Now, let a $(K-1)$-length sequence $y^{K-1} \in Y^{K-1}$ be given. 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 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 $\mathfrak{h}_{t-(K-k)}^t \cdots \mathfrak{h}_{t-1}^t y^k$ and $\bar{\mathfrak{h}}_{t-(K-k)}^t \cdots \bar{\mathfrak{h}}_{t-1}^t y^k$, yielding the same unique maximizing action at both histories for almost all shocks $z_2^{t+k} \in Z_2$, giving $V(\mathfrak{h}_{t-(K-k)}^t \cdots \mathfrak{h}_{t-1}^t y^k) = V(\bar{\mathfrak{h}}_{t-(K-k)}^t \cdots \bar{\mathfrak{h}}_{t-1}^t y^k)$. By backwards induction, I have $\sigma_2(\mathfrak{h}_{t-K}^t \cdots \mathfrak{h}_{t-1}^t z_2^{t+k}) = \sigma_2(\bar{\mathfrak{h}}_{t-K}^t \cdots \bar{\mathfrak{h}}_{t-1}^t z_2^{t+k})$ and $\sigma_1(\mathfrak{h}^t a_2 z_1^{t+k}) = \sigma_1(\bar{\mathfrak{h}}^t a_2 z_1^{t+k})$ 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 B.1 (specifically (20)), I can write this as

$$\tilde{V}^*_{a_1, \tilde{\sigma}_2}(h a_2, z_1^t) = \tilde{V}^*_{a_1', \tilde{\sigma}_2}(h a_2, z_1^t). \tag{21}$$

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 (20). Substituting into (21) and rearranging yields

$$\varepsilon(z_1^{a_1} - z_1^{a_1'}) = (1-\delta)[u_1(a_2, a_1') - u_1(a_2, a_1)] + \delta \sum_{y \in Y} [\rho(y | a_2, a_1) - \rho(y | a_2, a_1')] V^*(\tilde{\sigma}_2 | hy). \tag{22}$$

This implies that the set of shocks z_1 such that (22) holds is Lebesgue measure zero, so the profile is essentially sequentially strict. Lemma B.2 shows that every essentially strict equilibrium is quasi-Markov perfect, so the wPBE is quasi-Markov.

B.2 Proof of Proposition 4.2

Proposition 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.

C Proofs for Section 5

C.1 Proof of Proposition 5.1

Let $\underline{e}(\delta,\mu^0) \equiv \min \mathcal{E}(\delta,\mu^0), \bar{e}(\delta,\mu^0) \equiv \max \mathcal{E}(\delta,\mu^0)$ denote the minimum and maximum quasi-Markov equilibrium payoffs, respectively.

First consider the simpler case $\delta < \frac{1}{2}$. Let any quasi-Markov equilibrium σ 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 where $V^{t+1}(\cdot)$ is the continuation payoff at period $t+1$. 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 $V^t(h^t c) = V^t(h^t d) + 2(1-\delta)$. Since $V^t(h^t a_2)$ is independent of h^t conditional on a_2,

$$\begin{aligned} V^t(C) - V^t(D) &= (\sigma_2^t(C) - \sigma_2^t(D))V^t(h^t c) + ((1-\sigma_2^t(C)) - (1-\sigma_2^t(D)))V^t(h^t d) \\ &= (\sigma_2^t(C) - \sigma_2^t(D))[V^t(h^t c) - V^t(h^t d)] = 2(1-\delta)(\sigma_2^t(C) - \sigma_2^t(D)). \end{aligned} \tag{23}$$

Note that if C is a best response at any period t following player 2 action a_2,

$$(1-\delta)u_1(a_2,C) + \delta V^{t+1}(C) \geq (1-\delta)u_1(a_2,D) + \delta V^{t+1}(D)$$

$$\delta(V^{t+1}(C) - V^{t+1}(D)) \geq (1-\delta)(u_1(a_2,D) - u_1(a_2,C)) = 1-\delta, \tag{24}$$

yet from (23) the left hand side of (24) is less than $2\delta(1-\delta)$, a contradiction for $\delta < \frac{1}{2}$. Thus, it must be that player 1 always plays D: $\sigma_1(t,h) = D$ for all t,h. For history D at any period, the belief on the commitment type is 0 so player 2's best response is d. For $\mu^0 < \frac{1}{2}$, at period 0 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, (10) has been proven.

For the rest of the proof, suppose $\delta > \frac{1}{2}$. 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 C.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. Let some HBA $x \equiv (\phi, \mu, \alpha)$ enforced by $\tilde{w}$ be given. First, suppose $\tilde{\phi} = 0$. 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 $\tilde{\phi} \in (0,1)$. 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 hold. ■

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 Propositions 3.2 and 3.3 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 decompose anything and for similar reasons, are not consistent with any antegame equilibria. Sometimes I abuse notation by letting $\bar{\mathcal{B}}(w) \equiv \bar{\mathcal{B}}(\{w\})$ denote the set of "useful" HBPs decomposed by the singleton $\{w\}$.

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 belief mappings are redundant, I often omit them when writing HBPs and HBAs (e.g. write "(ϕ, γ)" instead of "(ϕ, μ, γ)"). Define $\phi^* \equiv \frac{\mu^0}{1-\mu^0}$. I rewrite $q(\phi)$ (defined in the statement of Proposition 5.1) 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 $\mu^0 \in (0, \frac{1}{2}]$). Define $r(\cdot)$ as the inverse of $q(\cdot)$: $r(\phi) \equiv q^{-1}(\phi) = \frac{1}{2}\left(1 - \frac{\phi^*}{\phi}\right)$. 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. Recall the abuse of notation $\mathsf{V}(v_1, v_2) \equiv (1-\delta)v_1 + \delta v_2$ where $v_1, v_2 \in \mathbb{R}$, as well as $\mathsf{V}^j(v_1, v_2) \equiv (1-\delta^j)v_1 + \delta^j v_2$, used here for brevity.

The following lemma characterizes the set $\bar{\mathcal{B}}(\tilde{w})$ of useful HBPs generated by a single HBP $\tilde{w}$.

Lemma C.2. *Suppose $\mu^0 \in (0, \frac{1}{2}]$. Let any HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ be given. Then*

$$\bar{\mathcal{B}}(\tilde{w}) = \begin{cases} \{(q(\tilde{\phi}), g_{\mathcal{I}}(\mathsf{V}(1, \tilde{\gamma}(D))))\} & \tilde{\phi} \in [0, \hat{\phi}_3) \\ \{(1, \gamma) : \gamma(D) = \mathsf{V}(1, \tilde{\gamma}(D)), \gamma(C) \in [\gamma(D) + \eta, \mathsf{V}(2, \tilde{\gamma}(C))]\} & \tilde{\phi} = \hat{\phi}_3 \\ \{(1, g(\mathsf{V}(1, \tilde{\gamma}(D)), \mathsf{V}(3, \tilde{\gamma}(D))))\} & \tilde{\phi} \in (\hat{\phi}_3, \frac{1}{2}) \\ \{(\phi, \gamma) : \gamma(D) \in [\mathsf{V}(1, \tilde{\gamma}(D)), \mathsf{V}(3, \tilde{\gamma}(D))], \gamma(C) = \mathsf{V}(3, \tilde{\gamma}(D))\} \cap \bar{\mathcal{F}} & \tilde{\phi} = \frac{1}{2} \\ \{(0, \gamma) : \gamma(C) = \gamma(D) = \mathsf{V}(2, \tilde{\gamma}(C))\} & \tilde{\phi} \in (\frac{1}{2}, 1]. \end{cases}$$

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 $\gamma(D) = (1 - \delta) + \delta\tilde{\gamma}(D) = \mathsf{V}(1, \tilde{\gamma}(D))$. At history C, player 2 is indifferent if

$$E[u_2(c, \alpha_1)|C] = 3\mu + (1 - \mu)[\tilde{\phi} \cdot 3 + (1 - \tilde{\phi}) \cdot 0] = E[u_2(d, \alpha_1)|C] = 2\mu + (1 - \mu)[\tilde{\phi} \cdot 2 + (1 - \tilde{\phi}) \cdot 1]$$

$$\mu[3 - 2] + (1 - \mu)[3\tilde{\phi} - 2\tilde{\phi} - (1 - \tilde{\phi})] = \mu + (1 - \mu)[2\tilde{\phi} - 1)] = 0$$

$$[2 - 2\tilde{\phi}]\mu = 1 - 2\tilde{\phi}. \tag{25}$$

Substituting Bayes' rule for μ gives

$$[2 - 2\tilde{\phi}]\frac{\mu^0}{\mu^0 + (1 - \mu^0)\phi} = 1 - 2\tilde{\phi}$$

and rearranging yields $(1 - 2\tilde{\phi})\phi = \frac{\mu^0}{1 - \mu^0} = \phi^*$, so I can write

$$\phi = \frac{\phi^*}{1 - 2\tilde{\phi}} = q(\tilde{\phi}).$$

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 $w \in \mathcal{F}_\varnothing$ by Lemma C.1 (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 $\gamma(D) \in [\mathsf{V}(1, \tilde{\gamma}(D)), \mathsf{V}(3, \tilde{\gamma}(D))]$.

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 (this corresponds to the plane at a particular "elevation" in Figure 7). The following corollary restates some of the results of Lemma C.2 in a more directly useful way, characterizing the "elevations" from which useful HBPs can be decomposed, and vice versa.

Corollary C.1. *Suppose* $\mu^0 \in (0, \frac{1}{2})$.

1. *Let any HBP* $w \equiv (\phi, \gamma) \in \mathcal{I}$ *be given. If 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* $\tilde{\phi} = \frac{1}{2}$.

2. *Conversely, for any* $\phi \geq q(0)$*, any HBP* $\tilde{w} \equiv (r(\phi), \tilde{\gamma}) \in \mathcal{W}_{r(\phi)}$ *generates an HBP in* $\mathcal{W}_\phi$*:* $(\phi, g_{\mathcal{I}}((1-\delta) + \delta\tilde{\gamma}(D)) \in \mathcal{B}(\tilde{w})$.

3. *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[22] at each history distribution following the mth iteration of the $\bar{\mathcal{B}}(\cdot)$ operator (recall that the minima and maxima exist because $\bar{\mathcal{B}}^m(\mathcal{F}^\dagger)$ is closed by Lemma A.2):

$$\underline{v}_\phi^m \equiv \begin{cases} \min\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi \in [0, 1) \\ \min\{\gamma(C) : (1, \gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi = 1 \end{cases} \quad \bar{v}_\phi^m \equiv \begin{cases} \max\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi \in [0, 1) \\ \max\{\gamma(C) : (1, \gamma) \in \bar{\mathcal{B}}^m(\mathcal{F}^\dagger)\} & \phi = 1. \end{cases}$$

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 $\tilde{\phi} \in (0, 1]$. Because $\{\bar{\mathcal{B}}^m(\bar{\mathcal{F}})\}_m$ is a decreasing sequence (in the sense of set inclusion), $\{\underline{v}_\phi^m\}_m$ ($\{\bar{v}_\phi^m\}_m$) is weakly increasing (weakly decreasing) and bounded; it follows that

$$\underline{v}_\phi^\infty = \begin{cases} \min\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{D}}\} & \phi \in [0, 1) \\ \min\{\gamma(C) : (1, \gamma) \in \bar{\mathcal{D}}\} & \phi = 1 \end{cases} \qquad \bar{v}_\phi^\infty = \begin{cases} \max\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{D}}\} & \phi \in [0, 1) \\ \max\{\gamma(C) : (1, \gamma) \in \bar{\mathcal{D}}\} & \phi = 1. \end{cases}$$

The following result characterizes the output of the "initialization" $\mathcal{V}(\cdot)$ operator.

Lemma C.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*

$$\mathcal{V}(\{(\phi, \gamma)\}) = \begin{cases} 3(1-\delta) + \delta\gamma(D) & \phi = 0 \\ 2(1-\delta) + \delta\gamma(C) & \phi > 0. \end{cases} \tag{26}$$

[22]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 $\gamma(D)$ "pins down" the other.

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: $\sigma_1^*(\varnothing) = \phi$.

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 period 0 for playing c is greater than or equal to the payoff for d:

$$3\mu^0 + (1-\mu^0)[3\phi + 0\cdot(1-\phi)] \geq 2\mu^0 + (1-\mu^0)[\phi \cdot 2 + (1-\phi)\cdot 1] \tag{27}$$

$$\frac{\mu^0}{1-\mu^0} = \phi^* \geq 1 - 2\phi \qquad \Longrightarrow \qquad \frac{\phi^*}{1-2\phi} = q(\phi) \geq 1, \tag{28}$$

which is a contradiction because $\phi < \hat{\phi}_3$ implies $q(\phi) < q(\hat{\phi}_3) = 1$, due to the monotonicity of $q(\cdot)$. 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 inequalities in (28) with "=" and working backwards, doing the same until (27), 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 $\mathcal{V}(\{(\phi,\gamma)\}) = \{V(\sigma^*) : \sigma_2^*(\varnothing) \in [0,1], \sigma_1^*(\varnothing) = \hat{\phi}_3\} = [(1-\delta) + \delta\gamma(D), 3(1-\delta) + \delta\gamma(D)]$.

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 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 (aside from $\mu^0 = 0$) listed in (11): suppose $\mu^0 \in (0, \frac{1}{9}]$. I show that the function $q(\cdot)$ has two fixed points for $\mu^0 < \frac{1}{9}$, one for $\mu^0 = \frac{1}{9}$, and none for $\mu^0 > \frac{1}{9}$ (this turns out to be why $\mu^0 > \frac{1}{9}$ has a unique stationary equilibrium payoffs instead of two). Setting $\phi = q(\phi)$ yields $\phi = \frac{\phi^*}{1-2\phi}$ which can be rearranged to get $0 = 2\phi^2 - \phi + \phi^*$. The quadratic formula gives

$$\phi = \tfrac{1}{4}(1 \pm \sqrt{1 - 4\cdot 2\phi^*}) = \tfrac{1}{4} \pm \tfrac{1}{4}\sqrt{1 - 8\phi^*}; \tag{29}$$

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 C.4. *Suppose $\mu^0 \in (0, \frac{1}{9}]$.*

1. *For all $\phi \in [\hat{\phi}_1, 1)$, $\underline{v}_\phi^\infty = 1$. Furthermore, $\underline{v}_1^\infty = 1/\delta$, and $\underline{v}_{\phi'}^\infty \geq 1$ for all $\phi' \in [0, \hat{\phi}_1)$.*

2. *For all $\phi \in [\hat{\phi}_1, 1)$, $\bar{v}_\phi^\infty = 2 - \eta$. Furthermore, $\bar{v}_1^\infty = 2$ and $\bar{v}_{\phi'}^\infty \leq 2$ for all $\phi' \in [0, \hat{\phi}_1)$.*

Proof. Recall that $r(\cdot)$ is strictly increasing, and note that $r(\hat{\phi}_1) = \hat{\phi}_1$, $r(1) = \hat{\phi}_3 < 1$. Then for all

$\phi' \in [\hat{\phi}_1, 1)$, $r(\phi') \in [\hat{\phi}_1, 1) \subset [q(0), 1)$. Corollary C.1 implies that for every $\phi \in [\hat{\phi}_1, 1)$,

$$\begin{aligned}
\underline{v}_\phi^{m+1} &= \min\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{B}}^{m+1}(\bar{\mathcal{F}})\} \\
&= \min\{\min\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{B}}(\bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \mathcal{W}_{r(\phi_k)})\}, \min\{\gamma(D) : (\phi, \gamma) \in \bar{\mathcal{B}}(\bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \mathcal{W}_{\frac{1}{2}})\}\} \\
&= \min\{\min\{\mathsf{V}(1, \gamma(D)) : (r(\phi), \gamma) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}))\}, \min\{\mathsf{V}(3, \gamma(D)) - \eta : (\tfrac{1}{2}, \gamma) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}))\}\} \\
&= \min\{\mathsf{V}(1, \underline{v}_{r(\phi)}^m), \mathsf{V}(3, \underline{v}_{\frac{1}{2}}^m) - \eta\}.
\end{aligned}$$

Note that $\underline{v}_\phi^0 = 0$ for all $\phi \in [0, 1)$. Suppose that for some m and for all $\phi \in [\hat{\phi}_1, 1)$, $\underline{v}_\phi^m = \underline{v}_{\frac{1}{2}}^m = \underline{\zeta}^m$ for some constant $\underline{\zeta}^m$. Then

$$\underline{v}_\phi^{m+1} = \min\{(1-\delta) + \delta \underline{v}_{r(\phi)}^m, 3(1-\delta) + \delta \underline{v}_{\frac{1}{2}}^m - \eta\} = (1-\delta) + \delta \underline{\zeta}^m.$$

By induction, for all m and $\phi \in [\hat{\phi}_1, 1)$, then $(1-\delta) + \delta\underline{\zeta}^m = \underline{v}_\phi^{m+1} \equiv \underline{\zeta}^{m+1}$ for constant $\underline{\zeta}^{m+1}$. 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 and for all $\phi \in [\hat{\phi}_1, 1)$, $\bar{v}_\phi^m = \bar{\zeta}^m$ for some constant $\bar{\zeta}^m$. Then

$$\bar{v}_\phi^{m+1} = \max\{(1-\delta) + \delta \bar{v}_{r(\phi_k)}^m, 3(1-\delta) + \delta \bar{v}_{\frac{1}{2}}^m - \eta\} = 3(1-\delta) + \delta\bar{\zeta}^m - \eta.$$

By induction, for each m and $\phi \in [\hat{\phi}_1, 1)$, $\bar{v}_\phi^{m+1} = 3(1-\delta) + \delta\bar{\zeta}^m - \eta \equiv \bar{\zeta}^{m+1}$. 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 C.2 and the above result give

$$\begin{aligned}
\underline{v}_1^{m+1} &= \min\{\gamma(C) : (1, \gamma) \in \bar{\mathcal{B}}^{m+1}(\bar{\mathcal{F}})\} = \min\left\{\gamma(C) : (1, \gamma) \in \bar{\mathcal{B}}\left(\bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \left(\cup_{\tilde{\phi} \in [\hat{\phi}_3, \frac{1}{2}]} \mathcal{W}_{\tilde{\phi}}\right)\right)\right\} \\
&= \min\left\{\begin{array}{l} \min\{(1-\delta) + \delta\tilde{\gamma}(D) + \eta : (\hat{\phi}_3, \tilde{\gamma}) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \mathcal{W}_{\hat{\phi}_3}\}, \\ \qquad \min\{3(1-\delta) + \delta\tilde{\gamma}(D) : (\tilde{\phi}, \tilde{\gamma}) \in \bar{\mathcal{B}}^m(\bar{\mathcal{F}}) \cap \left(\cup_{\tilde{\phi} \in (\hat{\phi}_3, \frac{1}{2}]} \mathcal{W}_{\tilde{\phi}}\right) \end{array}\right\} \\
&= \min\{(1-\delta) + \delta\underline{\zeta}^m + \eta, 3(1-\delta) + \delta\underline{\zeta}^m\} = (1-\delta) + \delta\underline{\zeta}^m + \eta,
\end{aligned}$$

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 C.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 $\lim_{m\to\infty} \underline{\mathsf{v}}^m \geq 1$.[23] 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

$$\bar{v}_0^\infty = \max\{3(1-\delta) + \delta(2-\eta), 2(1-\delta) + 2\delta\} = 2$$

$$\bar{v}_\phi^\infty \leq \max\{(1-\delta) + \delta\bar{\mathsf{v}}^\infty, 3(1-\delta) + \delta(2-\eta) - \eta\} = \max\{(1-\delta) + \delta\bar{\mathsf{v}}^\infty, 2-\eta\} \qquad \forall\phi \in (0,\hat{\phi}_1).$$

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 C.3 and C.4 imply that

$$\begin{aligned}
\min \mathcal{V}(\bar{\mathcal{D}}) &= \min\{\min\{(1-\delta) + \delta\gamma(D) : (\phi,\gamma) \in \bar{\mathcal{D}}, \phi \leq \hat{\phi}^3\}, \\
&\qquad \min\{2(1-\delta) + \delta\gamma(C) : (\phi,\gamma) \in \bar{\mathcal{D}}, \phi > \hat{\phi}^3\}\} \\
&= \min\{\min\{\mathsf{V}(1,\underline{v}_\phi^\infty) : \phi \leq \hat{\phi}^3\}, \min\{\mathsf{V}(2,\underline{v}_\phi^\infty + \eta) : \phi \in (\hat{\phi}^3,1)\}, \mathsf{V}(2,\underline{v}_1^\infty - \eta)\} \\
&= \min\{\mathsf{V}(1,1), \mathsf{V}(2,1+\eta)\} = 1.
\end{aligned}$$

Thus, the first minimum listed in (11) 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 $q(\cdot)$ (this means the only stationary equilibrium is at $\phi = \frac{1}{2}$). I still define $\hat{\phi}_3$ so that $q(\hat{\phi}_3) = 1$.

Lemma C.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 $q^{\bar{k}}(\phi) \geq \frac{1}{2}$.*

Proof. Note that $\mu^0 \in (\frac{1}{9},\frac{1}{2}) \implies \phi^* \in (\frac{1}{8},1)$. Since $q^0(0) = \phi^*$, for $\phi^* \geq \frac{1}{2}$ the lemma is trivial.

[23] This simply confirms the obvious fact that the continuation payoff must not be less than the minmax payoff.

Suppose $\phi^* \in (\frac{1}{8}, \frac{1}{2})$Let $\Delta q(\phi) \equiv q(\phi) - \phi = \frac{\phi^*}{1-2\phi} - \phi = \frac{\phi^* - \phi + 2\phi^2}{1-2\phi}$. I show that $\Delta q(\phi) > 0$ for $\phi \in [0, \frac{1}{2})$. Suppose not. Note that $\Delta q(0) = \phi^* > 0$ and $\lim_{\phi \to \frac{1}{2}} \Delta q(\phi) = \infty$. Since $\Delta q(\cdot)$ is continuous, there must exist $\phi' \in [0, \frac{1}{2})$ such that $0 = \phi^* - \phi + 2\phi^2 \implies \phi = \frac{1}{4}(1 \pm \sqrt{1 - 8\phi^*})$. But $\phi^* > \frac{1}{8}$ because $\mu^0 > \frac{1}{9}$, so the discriminant is negative, a contradiction.

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 $2\phi^*(1-2\phi)^{-2} - 1 = 0$ which can be solved for $\phi = \frac{1}{2}(1 \pm \sqrt{2\phi^*})$; for $\phi^* \in (\frac{1}{8}, \frac{1}{2})$ a solution $\phi \in [0, \frac{1}{2})$ exists. Thus, $\Delta q(\phi)$ has some minimum value $\Delta \underline{q} > 0$ for $\phi \in [0, \frac{1}{2})$ and so for any k,

$$q^k(\phi) = \phi + \Delta q(\phi) + \Delta q(q(\phi)) + \Delta q(q^2(\phi)) + \cdots + \Delta q(q^{k-1}(\phi)) \geq \phi + k\Delta \underline{q},$$

proving that there exists finite $\bar{k}$ such that $q^{\bar{k}}(\phi) \geq \frac{1}{2}$. ■

Having established its finite existence with Lemma C.5, recall the definition of $L(\mu^0) \equiv \min_k q^k(0)$ such that $q^k(0) \geq \frac{1}{2}$; recall also the definitions of λ_1^k for $k \in \{L-1, L\}$ in the statement of the proposition. Define $p^k \equiv q^k(0)$ for $k \in \{0, ..., L\}$, and define intervals $\mathcal{S}^k \equiv (p^k, p^{k+1}) \subset \mathbb{R}$ for $k \in \{0, ..., L-1\}$ and $\mathcal{S}^L \equiv (p^L, 1)$.

Lemma C.6. *Suppose $\mu^0 \in (\frac{1}{9}, \frac{1}{2}]$. If $p^L = \frac{1}{2}$, then $\underline{v}^\infty_{p^L} = \lambda_1^L$; otherwise, $\underline{v}^\infty_{p^{L-1}} = \lambda_1^{L-1}$. Also,*

$$\underline{v}^\infty_{p^L} = \begin{cases} (1-\delta) + \delta \underline{v}^\infty_{p^{L-1}} & p^L < 1 \\ (1-\delta) + \delta \underline{v}^\infty_{p^{L-1}} + \eta & p^L = 1. \end{cases}$$

Finally, $\min\{\underline{v}^\infty_\phi : \phi \in [0, \hat{\phi}_3]\} = \underline{v}^\infty_{p^{L-1}}$ *and*

$$\min\{\underline{v}^\infty_\phi : \phi \in (\hat{\phi}_3, 1)\} = \begin{cases} \underline{v}^\infty_{p^L} & p^L < 1 \\ \underline{v}^\infty_{p^{L-1}} & p^L = 1 \end{cases} \qquad \underline{v}^\infty_1 = \begin{cases} \underline{v}^\infty_{p^L} + \eta & p^L < 1 \\ \underline{v}^\infty_{p^L} & p^L = 1. \end{cases}$$

Proof. Note that λ_1^k as defined in the statement of Proposition 5.1 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 $\lambda_2^L = \frac{1}{1-\delta^{L+1}}[1 - \delta^L + 3\delta^L(1-\delta)]$. I abuse notation by letting $\underline{v}^\infty_\mathcal{S} \equiv \min\{\underline{v}^\infty_\phi : \phi \in \mathcal{S}\}$ for a set $\mathcal{S}$.

First, Corollary C.1 implies that for any m, $\underline{v}^{m+1}_{\mathcal{S}^0} = 3(1-\delta) + \delta \underline{v}^m_{1/2} - \eta$. To see why, recall that $r(\cdot)$ is strictly increasing and note that for all $\phi \in \mathcal{S}^0$, $p^0 < \phi < p^1$ and then $r(\phi) < r(p^1) = p^0 = 0$, so there does not exist ϕ' such that $q(\phi') = \phi$. Hence only HBPs (ϕ', γ') where $\phi' = \frac{1}{2}$ can decompose to HBPs (ϕ, γ) where $\phi \in \mathcal{S}^0$. Taking the limit as $m \to \infty$ gives $\underline{v}^\infty_{\mathcal{S}^0} = 3(1-\delta) + \delta \underline{v}^\infty_{1/2} - \eta$. In the rest of the proof, for brevity I skip ahead to the limit as $m \to \infty$, e.g. omitting the intermediate step of "$\underline{v}^{m+1}_{\mathcal{S}^0} = 3(1-\delta) + \delta \underline{v}^m_{1/2} - \eta$" above.

For each $k \in \{1, ..., L\}$ and all $\phi \in \mathcal{S}^k$, $r(\phi) \in \mathcal{S}^{k-1}$. To see this, note that $p^k < \phi < p^{k+1}$ for $k < L$ (this is easily adapted for $k = L$), so $r(p^k) = p^{k-1} < r(\phi) < p^k = r(p^{k+1})$. Corollary C.1 shows that $\underline{v}^\infty_{\mathcal{S}^k} = \min\{\mathsf{V}(1, \underline{v}^\infty_{\mathcal{S}^{k-1}}), \mathsf{V}(3, \underline{v}^\infty_{1/2}) - \eta\}$. Suppose (by contradiction) that $\mathsf{V}(3, \underline{v}^\infty_{1/2}) - \eta <$

$\mathsf{V}(1, \underline{v}^\infty_{\mathcal{S}^{k-1}})$. Then

$$\mathsf{V}(3, \underline{v}^\infty_{1/2}) - \eta < (1-\delta) + \delta \underline{v}^\infty_{\mathcal{S}^{k-1}} \leq (1-\delta) + \delta \left[\mathsf{V}(3, \underline{v}^\infty_{1/2}) - \frac{1-\delta}{\delta}\right] = \delta[\mathsf{V}(3, \underline{v}^\infty_{1/2})].$$

$$(1-\delta)\mathsf{V}(3, \underline{v}^\infty_{1/2}) < \eta = \frac{1-\delta}{\delta} \qquad \Longrightarrow \qquad 3\delta(1-\delta) + \delta^2 \underline{v}^\infty_{1/2} < 1 \tag{30}$$

For $\delta > \frac{1}{2}$, the left hand side of (30) is less than $\frac{3}{4} + \frac{1}{4}\underline{v}^\infty_{1/2} < 1 \implies \underline{v}^\infty_{1/2} < 1$, which is easily shown to be contradicted by Lemma C.2.[24] Thus, $\underline{v}^\infty_{\mathcal{S}^k} = \mathsf{V}(1, \underline{v}^\infty_{\mathcal{S}^{k-1}})$, and in fact the argument above shows that for all $\phi^{k-1} \in \mathcal{S}^{k-1}, \phi^k \in \mathcal{S}^k$, I have $\underline{v}^\infty_{\phi^k} = \mathsf{V}(1, \underline{v}^\infty_{\phi^{k-1}})$. An almost identical argument shows that $\underline{v}^\infty_{p^k} = \mathsf{V}(1, \underline{v}^\infty_{p^{k-1}})$. Note that because $\underline{v}^\infty_{p^0} \leq \mathsf{V}(3, \underline{v}^\infty_{1/2}) - \eta = \underline{v}^\infty_{\mathcal{S}^0}$, $\underline{v}^\infty_{p^k} \leq \underline{v}^\infty_{\mathcal{S}^k}$ for each k.

Suppose $p^L < 1$. I show that $\underline{v}^\infty_1 = \underline{v}^\infty_{\mathcal{S}^L} + \eta$. 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 C.2,

$$\underline{v}^\infty_1 = \min\{\mathsf{V}(1, \underline{v}^\infty_{\hat{\phi}_3}) + \eta, \min\{\mathsf{V}(3, \underline{v}^\infty_\phi) : \phi \in (\hat{\phi}_3, \tfrac{1}{2})\}, \mathsf{V}(3, \underline{v}^\infty_{1/2})\} = \min\{\mathsf{V}(1, \underline{v}^\infty_{\mathcal{S}^{L-1}}) + \eta, \mathsf{V}(3, \underline{v}^\infty_{1/2})\}. \tag{31}$$

If $p^L > \frac{1}{2}$, then $\frac{1}{2} \in \mathcal{S}^{L-1}$ and $\underline{v}^\infty_{1/2} = \underline{v}^\infty_{\mathcal{S}^{L-1}}$, so

$$\underline{v}^\infty_1 = (1-\delta) + \delta \underline{v}^\infty_{\mathcal{S}^{L-1}} + \eta = \underline{v}^\infty_{\mathcal{S}^L} + \eta. \tag{32}$$

Suppose $p^L = \frac{1}{2}$. By Lemma C.2,

$$\underline{v}^\infty_{p^0} = \min\{\mathsf{V}(3, \underline{v}^\infty_{1/2}) - \eta, \min\{\mathsf{V}(3, \underline{v}_\phi) : \phi \in (\tfrac{1}{2}, 1)\}, \mathsf{V}(2, \underline{v}^\infty_1)\} = \min\{\mathsf{V}(3, \underline{v}^\infty_{1/2}) - \eta, \mathsf{V}(2, \underline{v}^\infty_1)\}, \tag{33}$$

because $(\frac{1}{2}, 1) = \mathcal{S}^L$ and $\underline{v}^\infty_{\mathcal{S}^L} \geq \underline{v}^\infty_{p^L} = \frac{1}{2}$. Recall (31); suppose by contradiction that

$$\underline{v}^\infty_1 = 3(1-\delta) + \delta \underline{v}^\infty_{1/2} < (1-\delta) + \delta \underline{v}^\infty_{\mathcal{S}^{L-1}} + \eta. \tag{34}$$

Since $\underline{v}^\infty_{1/2} \leq 2 - \eta$ by Corollary C.1, I have

$$\mathsf{V}(2, \underline{v}^\infty_1) = 2(1-\delta) + \delta[3(1-\delta) + \delta \underline{v}^\infty_{1/2}] > 3(1-\delta) + \delta \underline{v}^\infty_{1/2} - \eta = \underline{v}^\infty_{p^0} = \underline{v}^\infty_{\mathcal{S}^0}.$$

By induction that means $\underline{v}^\infty_{\mathcal{S}^k} = \underline{v}^\infty_{p^k}$ for each k, and $\underline{v}^\infty_{\frac{1}{2}} = \underline{v}^\infty_{p^L} = \mathsf{V}^L(\underline{v}^\infty_{p^0}) = (1-\delta^L) + \delta^L[3(1-\delta) + \delta \underline{v}^\infty_{p^L} - \eta] = \lambda^L_1$, since p^L matches the characterization for λ^L_1 at the beginning of the proof. Returning to (34) I have

$$3(1-\delta) + \delta \underline{v}^\infty_{\frac{1}{2}} = 3(1-\delta) + \delta \underline{v}^\infty_{p^L} < (1-\delta) + \delta \underline{v}^\infty_{\mathcal{S}^{L-1}} + \eta = (1-\delta) + \delta \underline{v}^\infty_{p^{L-1}} + \eta = \underline{v}^\infty_{p^L} + \eta$$

$$3(1-\delta) - \eta < (1-\delta)\underline{v}^\infty_{p^L} \qquad \Longrightarrow \qquad 3 - \frac{1}{\delta} = 2 - \eta < \underline{v}^\infty_{p^L},$$

[24] Besides the fact that this implies a continuation payoff less than the minmax, this can be seen by simply letting $\underline{\mathsf{v}}^m \equiv \min\{\min\{\underline{v}^m_\phi : \phi \in [0,1)\}, \underline{v}^m_1\}$ and generating the lowest payoff given by Lemma C.2 at any ϕ as the new lower bound $\underline{\mathsf{v}}^{m+1} \geq (1-\delta) + \delta \underline{\mathsf{v}}^m$, which converges to 1 as $m \to \infty$.

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 C.1. To summarize, for $p^L = \frac{1}{2}$, it has been shown that $\underline{v}^{\infty}_1 = (1-\delta) + \delta \underline{v}^{\infty}_{\mathcal{S}^{L-1}} + \eta = \underline{v}^{\infty}_{\mathcal{S}^L} + \eta$, 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 C.2, $\underline{v}^{\infty}_{p^0} = \min\{\mathsf{V}(3, \underline{v}^{\infty}_{1/2}) - \eta, \min\{\mathsf{V}(3, \underline{v}_{\phi}) : \phi \in (\frac{1}{2}, 1)\}, \mathsf{V}(2, \underline{v}^{\infty}_1)\}$.

Suppose by contradiction that

$$\underline{v}^{\infty}_{p^0} < 3(1-\delta) + \delta \underline{v}^{\infty}_{1/2} - \eta. \tag{35}$$

Then by Lemma C.2 and (32) it must be that $\underline{v}^{\infty}_{p^0} = 3(1-\delta) + \delta \underline{v}^{\infty}_{p^L}$. Since $\underline{v}^{\infty}_{p^L} = (1-\delta^L) + \delta^L \underline{v}^{\infty}_{p^0} = (1-\delta^L) + \delta^L[3(1-\delta) + \delta \underline{v}^{\infty}_{p^L}]$, it then matches the characterization of λ^L_2 at the beginning of the proof: $\underline{v}^{\infty}_{p^L} = \lambda^L_2$. Recall that $\underline{v}^{\infty}_{\mathcal{S}^0} = 3(1-\delta) + \delta \underline{v}^{\infty}_{1/2} - \eta$; since $\frac{1}{2} \subset \mathcal{S}^{k-1}$,

$$\underline{v}^{\infty}_{\frac{1}{2}} = \underline{v}^{\infty}_{\mathcal{S}^{L-1}} = (1-\delta^{L-1}) + \delta^{L-1} \underline{v}^{\infty}_{\mathcal{S}^0} = (1-\delta^{L-1}) + \delta^{L-1}[3(1-\delta) + \delta \underline{v}^{\infty}_{\mathcal{S}^{L-1}} - \eta].$$

Therefore $\underline{v}^{\infty}_{1/2}$ matches the characterization of λ^{L-1}_1 above: $\underline{v}^{\infty}_{1/2} = \lambda^{L-1}_1$. I can then write (35) as

$$\underline{v}^{\infty}_{p^0} = 3(1-\delta) + \delta \underline{v}^{\infty}_{p^L} = 3(1-\delta) + \delta \lambda^L_2 < 3(1-\delta) + \delta \underline{v}^{\infty}_{1/2} - \eta = 3(1-\delta) + \delta \lambda^{L-1}_1 - \eta$$

Rearranging obtains $\lambda^L_2 < \lambda^{L-1}_1 - \frac{1-\delta}{\delta^2}$. Substituting the definitions of $\lambda^L_2, \lambda^{L-1}_1$ gives

$$\frac{\delta^2}{1-\delta^{L+1}}[1 - \delta^L + 3\delta^L(1-\delta)] < \frac{\delta^2}{1-\delta^L}[(1-\delta^{L-1}) + 3\delta^{L-1}(1-\delta) - \delta^{L-1}\eta] - (1-\delta).$$

I spare the reader the tedious algebra that yields

$$1 - 3\delta^{L+1} + 2\delta^{L+2} < 0. \tag{36}$$

At $\delta = 1$ the left hand side is 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 negative for $\delta \in (0,1)$, so the left hand side of (36) is positive for $\delta \in (0,1)$, a contradiction.

Thus, $\underline{v}^{\infty}_{p^0} = 3(1-\delta) + \delta \underline{v}^{\infty}_{1/2} - \eta = \underline{v}^{\infty}_{\mathcal{S}^0}$, implying 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 rearrangement gives $\underline{v}^{\infty}_{p^{L-1}} = \lambda^{L-1}_2$. 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 clear that $\min\{\underline{v}^{\infty}_{\phi} : \phi \in (\hat{\phi}_3, 1)\} = \underline{v}^{\infty}_{p^L} = \underline{v}^{\infty}_{\mathcal{S}^L}$. Above it was shown that $\underline{v}^{\infty}_1 = \underline{v}^{\infty}_{p^L} + \eta$, proving the lemma for $p^L < 1$. ■

By Lemmas C.3 and C.6, for $p^L < 1$,

$$\min \mathcal{V}(\bar{\mathcal{D}}) = \min\{\min\{\mathsf{V}(1, \underline{v}^{\infty}_{\phi}) : \phi \leq \hat{\phi}_3\}, \min\{\mathsf{V}(2, \underline{v}^{\infty}_{\phi} + \eta) : \phi \in (\hat{\phi}_3, 1)\}, \mathsf{V}(2, \underline{v}^{\infty}_1)\};$$

noting that $\underline{v}^\infty_{p^{L-1}} = \min\{\underline{v}^\infty_\phi : \phi \le \hat{\phi}_3\}$, $\underline{v}^\infty_{p^L} = \min\{\underline{v}^\infty_\phi : \phi \in (\hat{\phi}_3, 1)\}$, and $\underline{v}^\infty_1 = \underline{v}^\infty_{p^L} + \eta$, I have

$$\begin{aligned} \min \mathcal{V}(\bar{\mathcal{D}}) &= \min\{\mathsf{V}(1, \underline{v}^\infty_{p^{L-1}}), \mathsf{V}(2, \underline{v}^\infty_{p^L})\} = \min\{\mathsf{V}(1, \underline{v}^\infty_{p^{L-1}}), 2(1-\delta) + \delta((1-\delta) + \delta \underline{v}^\infty_{p^{L-1}} + \eta)\} \\ &= \min\{\mathsf{V}(1, \underline{v}^\infty_{p^{L-1}}), (1-\delta)(3+\delta) + \delta^2 \underline{v}^\infty_{p^{L-1}}\}. \end{aligned} \tag{37}$$

If $p^L = 1$, via similar substitutions I have

$$\begin{aligned} \min \mathcal{V}(\bar{\mathcal{D}}) &= \min\{\min\{\mathsf{V}(1, \underline{v}^\infty_\phi) : \phi \le \hat{\phi}_3\}, \min\{\mathsf{V}(2, \underline{v}^\infty_\phi + \eta) : \phi \in (\hat{\phi}_3, 1)\}, \mathsf{V}(2, \underline{v}^\infty_1)\} \\ &= \min\{\mathsf{V}(1, \underline{v}^\infty_{p^{L-1}}), \mathsf{V}(2, \underline{v}^\infty_{p^{L-1}} + \eta), \mathsf{V}(2, \underline{v}^\infty_{p^L})\} \\ &= \min\{\mathsf{V}(1, \underline{v}^\infty_{p^{L-1}}), (1-\delta)(3+\delta) + \delta^2 \underline{v}^\infty_{p^{L-1}}\}, \end{aligned} \tag{38}$$

noting that (38) matches (37). 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 $\delta(1-\delta)\underline{v}^\infty_{p^{L-1}} > (1-\delta)(2+\delta) \implies \underline{v}^\infty_{p^{L-1}} > \frac{2+\delta}{\delta} > 3$, an infeasible payoff, and so a contradiction. Rearranging $(1-\delta) + \delta \underline{v}^\infty_{p^{L-1}} = \underline{v}^\infty_{p^L}$ yields $\underline{v}^\infty_{p^{L-1}} = \frac{1}{\delta}\underline{v}^\infty_{p^L} - \eta$. Continuing,

$$\begin{aligned} \min \mathcal{V}(\bar{\mathcal{D}}) &= \mathsf{V}(1, \underline{v}^\infty_{p^{L-1}}) = \begin{cases} \mathsf{V}(1, \lambda_1^{L-1}) & p^L > \frac{1}{2} \\ \mathsf{V}\left(1, \frac{1}{\delta}\underline{v}^\infty_{p^L} - \eta\right) & p^L = \frac{1}{2} \end{cases} = \begin{cases} (1-\delta) + \delta \lambda_1^{L-1} & p^L > \frac{1}{2} \\ (1-\delta) + \delta\left[\frac{1}{\delta}\lambda_1^L - \eta\right] & p^L = \frac{1}{2} \end{cases} \\ &= \begin{cases} (1-\delta) + \delta \lambda_1^{L-1} & p^L > \frac{1}{2} \\ \lambda_1^L & p^L = \frac{1}{2}. \end{cases} \end{aligned}$$

Turning to the maxima $\bar{v}^\infty_\phi$, the conclusions of Lemma C.4 hold for the $\mu^0 \in (\frac{1}{9}, \frac{1}{2})$ case using almost identical arguments to the proof thereof. This then implies $\max\{\bar{v}^\infty_\phi : \phi \in [0, \hat{\phi}_3)\} = 2$, $\max\{\bar{v}^\infty_\phi : \phi \in [\hat{\phi}_3, 1)\} = 2 - \eta$ and $\bar{v}^\infty_1 = 2$. Thus,

$$\begin{aligned} \max \mathcal{V}(\bar{\mathcal{D}}) &= \max\{\mathsf{V}(1, \max\{\bar{v}^\infty_\phi : \phi \in [0, \hat{\phi}_3)\}), \mathsf{V}(3, \max\{\bar{v}^\infty_\phi : \phi \in [\hat{\phi}_3, 1)\}), \mathsf{V}(2, \bar{v}^\infty_1)\} \\ &= \max\{(1-\delta) + 2\delta, 3(1-\delta) + \delta(2-\eta), 2(1-\delta) + 2\delta\} = 2. \end{aligned}$$

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 start with the following analogue of Lemma C.2.

Lemma C.7. *Suppose $\mu^0 \in (\frac{1}{2}, 1]$. Let any HBP $\tilde{w} \equiv (\tilde{\phi}, \tilde{\gamma})$ be given. Then*

$$\bar{\mathcal{B}}(\tilde{w}) = \begin{cases} \{(1, g(\mathsf{V}(1, \tilde{\gamma}(D)), \mathsf{V}(3, \tilde{\gamma}(D))))\} & \tilde{\phi} \in [0, \frac{1}{2}) \\ \{(\phi, \gamma) : \gamma(D) \in [\mathsf{V}(1, \tilde{\gamma}(D)), \mathsf{V}(3, \tilde{\gamma}(D))], \gamma(C) = \mathsf{V}(3, \tilde{\gamma}(D))\} \cap \bar{\mathcal{F}} & \tilde{\phi} = \frac{1}{2} \\ \{(0, \gamma) : \gamma(D) = \gamma(C), \gamma(C) = 2(1-\delta) + \delta \tilde{\gamma}(C)\} & \tilde{\phi} \in (\frac{1}{2}, 1]. \end{cases}$$

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 $\gamma(D) = \gamma(C) = (1-\delta)u_1(c, C) + \delta\tilde{\gamma}(C) = 2(1-\delta) + \delta\tilde{\gamma}(C)$. ■

Lemma C.7 gives

$$\underline{v}_0^{m+1} = \min\{\min\{3(1-\delta) + \delta\underline{v}_\phi^m - \eta : \phi \in [\tfrac{1}{2}, 1)\}, 2(1-\delta) + \delta\underline{v}_1^m\}$$

$$\underline{v}_\phi^{m+1} = 3(1-\delta) + \delta\underline{v}_{1/2}^m - \eta \qquad \forall\phi \in (0, 1)$$

$$\underline{v}_1^{m+1} = \min\{3(1-\delta) + \delta\underline{v}_\phi^m : \phi \in [0, \tfrac{1}{2}]\}.$$

Then $\underline{v}_{1/2}^\infty = 3(1-\delta) + \delta\underline{v}_{1/2}^\infty - \eta$, which yields $\underline{v}_{1/2}^\infty = 3 - 1/\delta = 2 - \eta$. So for all $\phi \in (0, 1)$,

$$\underline{v}_\phi^\infty = 3(1-\delta) + \delta\,(3 - 1/\delta) - (1-\delta)/\delta = 3(1-\delta) + 3\delta - 1 - (1-\delta)/\delta = 2 - \eta.$$

I can then write

$$\underline{v}_0^\infty = \min\{3(1-\delta) + \delta(2-\eta) - \eta, 2(1-\delta) + \delta\underline{v}_1^\infty\} = \min\{2-\eta, 2(1-\delta) + \delta\underline{v}_1^\infty\}$$

$$\underline{v}_1^\infty = \min\{3(1-\delta) + \delta\underline{v}_0^\infty, 3(1-\delta) + \delta(2-\eta)\} = \min\{3(1-\delta) + \delta\underline{v}_0^\infty, 2\}.$$

Suppose by contradiction $\underline{v}_1^\infty < 2$. Then

$$\underline{v}_0^\infty = \min\{2-\eta, 2(1-\delta) + \delta[3(1-\delta) + \delta\underline{v}_0^\infty]\} = \min\{2-\eta, 2 + \delta - 3\delta^2 + \delta^2\underline{v}_0^\infty\}.$$

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 $\underline{v}_0^\infty = [2+\delta(1-3\delta)]/[1-\delta^2] > 2$, 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 (26) in Lemma C.3 shows that $\min \mathcal{V}(\bar{\mathcal{D}}) = 2$.

Turning to the maxima, Lemma C.7 implies

$$\bar{v}_0^{m+1} = \max\{\max\{3(1-\delta) + \delta\bar{v}_\phi^m - \eta : \phi \in [\tfrac{1}{2}, 1)\}, 2(1-\delta) + \delta\bar{v}_1^m\}$$

$$\bar{v}_\phi^{m+1} = 3(1-\delta) + \delta\bar{v}_{1/2}^m - \eta \qquad \forall\phi \in (0, 1)$$

$$\bar{v}_1^{m+1} = \max\{3(1-\delta) + \delta \bar{v}_\phi^m : \phi \in [0, \tfrac{1}{2}]\}$$

Then for all $\phi \in (0,1)$, $\underline{v}_{1/2}^{m+1} = \bar{v}_\phi^{m+1} = 3(1-\delta) + \delta \bar{v}_{1/2}^m - \eta$, which in the limit gives $\bar{v}_\phi^\infty = 3(1-\delta) + \delta \bar{v}_\phi^\infty - (1-\delta)/\delta \implies \bar{v}_\phi^\infty = 3 - 1/\delta = 2 - \eta$. Note that $\bar{v}_0^0 = \bar{v}_\phi^0 = 3 - \eta$ and $\bar{v}_1^0 = 3$. Suppose for some m that $\underline{v}_{1/2}^m \leq \bar{v}_1^m - \eta$ and $\underline{v}_{1/2}^m \leq \bar{v}_0^m$. 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

$$\max\{3(1-\delta) + \delta \bar{v}_\phi^m - \eta : \phi \in [\tfrac{1}{2}, 1)\} = 3(1-\delta) + \delta \underline{v}_{1/2}^m - \eta > 2(1-\delta) + \delta \bar{v}_1^m \tag{39}$$

$$(1-\delta) - \eta > \delta(\bar{v}_1^m - \underline{v}_{1/2}^m) \qquad \implies \qquad -\eta > \delta(\bar{v}_1^m - \eta - \underline{v}_{1/2}^m),$$

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 (39) 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, $\bar{v}_1^{m+1} = 3(1-\delta) + \delta \bar{v}_0^m \geq 3(1-\delta) + \delta \underline{v}_{1/2}^m$. 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

$$\bar{v}_0^\infty = 2(1-\delta) + \delta[3(1-\delta) + \delta \bar{v}_0^\infty] = 2(1-\delta) + 3\delta(1-\delta) + \delta^2 \bar{v}_0^\infty \implies \bar{v}_0^\infty = \frac{1-\delta}{1-\delta^2}(2+3\delta)$$

$$\bar{v}_1^\infty = 3(1-\delta) + \delta(1-\delta)\frac{2+3\delta}{1-\delta^2} = (1-\delta)\left[3 + \frac{2\delta + 3\delta^2}{1-\delta^2}\right] = \frac{1-\delta}{1-\delta^2}(3+2\delta).$$

Finally, by Lemma C.3,

$$\begin{aligned} \max \mathcal{V}(\bar{\mathcal{D}}) &= \max\{\mathsf{V}(3, \bar{v}_0^\infty), \max\{\mathsf{V}(2, \bar{v}_\phi^\infty + \eta) : \phi \in (0,1)\}, \mathsf{V}(2, \bar{v}_1^\infty)\} \\ &= \max\left\{\mathsf{V}\left(3, \frac{1-\delta}{1-\delta^2}(2+3\delta)\right), \mathsf{V}(2, \bar{v}_{1/2}^\infty + \eta), \mathsf{V}\left(2, \frac{1-\delta}{1-\delta^2}(3+2\delta)\right)\right\}. \end{aligned}$$

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 C.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 $\max \mathcal{V}(\bar{\mathcal{D}}) = \frac{1-\delta}{1-\delta^2}(3+2\delta)$.

C.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 5.1.[25] For each case, I construct the unperturbed equilibrium strategy profile σ^* and a sequence of perturbed game

[25] The $\delta < \frac{1}{2}$ cases are straightforward (with player 1 always playing D) and so I omit them.

strategy profiles $(\tilde{\sigma}^k)_k$, each corresponding to the (ψ, ε^k)-perturbed game where ψ is a distribution with strictly positive density on $[0,1]^{|A|}$ and $(\varepsilon^k)_k$ is a sequence where $\varepsilon^k > 0$ and $\varepsilon^k \to 0$. I show purifiability (according to Definition 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, as they are essentially simpler versions of the proofs for the perturbed equilibria.) The following lemma will be helpful in showing the existence of the equilibria constructed.

Lemma C.8. *Let two continuous functions $\tilde{\mathfrak{z}}_2 : [-1,1] \times (0,\infty) \to \mathbb{R}$ and $\tilde{\mathfrak{z}}_1 : [-1,1] \times (0,\infty) \to \mathbb{R}$ be given. Suppose that there exists $\varepsilon^* > 0$ such that for all $\varepsilon \in (0, \varepsilon^*)$, there exist numbers $\underline{\zeta}_{-i}, \bar{\zeta}_{-i} \in [-1,1]$ such that $\tilde{\mathfrak{z}}_i(\underline{\zeta}_{-i}|\varepsilon) = -1, \tilde{\mathfrak{z}}_i(\bar{\zeta}_{-i}|\varepsilon) = 1$ for both $i \in \{1,2\}$. Then there exist $\zeta_1^*, \zeta_2^* \in [-1,1]$ such that $\tilde{\mathfrak{z}}_2(\zeta_1^*|\varepsilon) = \zeta_2^*, \tilde{\mathfrak{z}}_1(\zeta_2^*|\varepsilon) = \zeta_1^*$.*

Proof. For ε sufficiently small, there exist $\underline{\zeta}_1, \bar{\zeta}_1, \underline{\zeta}_2, \bar{\zeta}_2 \in (-1,1)$ such that $\tilde{\mathfrak{z}}_2(\underline{\zeta}_1|\varepsilon) = -1, \tilde{\mathfrak{z}}_2(\bar{\zeta}_1|\varepsilon) = 1, \tilde{\mathfrak{z}}_1(\underline{\zeta}_1|\varepsilon) = -1, \tilde{\mathfrak{z}}_1(\bar{\zeta}_2|\varepsilon) = 1$, so I can define the inverse functions $\tilde{\mathfrak{z}}_2^{-1} : [-1,1] \times (0,\infty) \to [-1,1], \tilde{\mathfrak{z}}_1^{-1} : [-1,1] \times (0,\infty) \to [-1,1]$. Since these functions are continuous, so is the vector-valued $\tilde{\mathfrak{z}}^{-1} : [-1,1]^2 \times (0,\infty) \to [-1,1]^2$ where $\tilde{\mathfrak{z}}^{-1}(\zeta_1, \zeta_2|\varepsilon) = (\tilde{\mathfrak{z}}_2^{-1}(\zeta_1|\varepsilon), \tilde{\mathfrak{z}}_1^{-1}(\zeta_2|\varepsilon))$. Then $\tilde{\mathfrak{z}}^{-1}(\cdot|\varepsilon)$ is a continuous function from $[-1,1]^2$ into itself, so by Brouwer's fixed point theorem, $\tilde{\mathfrak{z}}^{-1}(\cdot|\varepsilon)$ has a fixed point $(\zeta_1^*, \zeta_2^*) \in [-1,1]$. Thus, $\tilde{\mathfrak{z}}_i^{-1}(\zeta_{-i}^*) = \zeta_i^*$ implies that $\tilde{\mathfrak{z}}_2(\zeta_1^*|\varepsilon) = \zeta_2^*$ and $\tilde{\mathfrak{z}}_1(\zeta_2^*|\varepsilon) = \zeta_1^*$. ■

C.2.1 Stationary Equilibrium with Payoff 2 ($0 < \mu^0 \leq 1$)

The equilibrium for the unperturbed game is defined as follows:

$$\sigma_2^*(t,h) = \begin{cases} 1 & h = \varnothing \\ 1 & h = C \\ 1 - \frac{1}{2\delta} & h = D \end{cases} \qquad \sigma_1^*(t) = \frac{1}{2}.$$

Define the sequence $(\tilde{\sigma}^k)_k$ of equilibria in the sequence of perturbed games:

$$\tilde{\sigma}_2^k(\varnothing, z_2) \equiv \tilde{\sigma}_2^k(t, C, z_2) \equiv 1 \qquad \tilde{\sigma}_2^k(t, D, z_2) \equiv \begin{cases} 0 & \Delta z_2 \leq \zeta_2^k \\ 1 & \Delta z_2 > \zeta_2^k \end{cases} \qquad \tilde{\sigma}_1^k(t, z_1) \equiv \begin{cases} 0 & \Delta z_1 \leq \zeta_1^k \\ 1 & \Delta z_1 > \zeta_1^k \end{cases}$$

where ζ_1^k, ζ_2^k are defined as the solution to the system of equations

$$\zeta_2^k \equiv \mathfrak{z}_2(\zeta_1^k|\varepsilon^k) \equiv \frac{1}{\varepsilon^k}(2\Delta\Psi_1(\zeta_1^k) - 1), \qquad \zeta_1^k \equiv \mathfrak{z}_1(\zeta_2^k|\varepsilon^k) \equiv \frac{1}{\varepsilon^k}(1-\delta)(1 - 2\delta\Delta\Psi_2(\zeta_2^k)). \quad (40)$$

Note that for large enough k (and hence small enough ε^k), $\mathfrak{z}_2(-1|\varepsilon^k) < -1, \mathfrak{z}_2(1|\varepsilon^k) > 1, \mathfrak{z}_1(-1|\varepsilon^k) > 1, \mathfrak{z}_1(1|\varepsilon^k) < -1$. Since $\mathfrak{z}_2(\cdot), \mathfrak{z}_1(\cdot)$ are continuous, there exist $\underline{\zeta}_{-i}, \bar{\zeta}_{-i} \in [-1,1]$ such that $\tilde{\mathfrak{z}}_i(\underline{\zeta}_{-i}|\varepsilon) = -1, \tilde{\mathfrak{z}}_i(\bar{\zeta}_{-i}|\varepsilon) = 1$ for both $i \in \{1,2\}$. Thus, $\mathfrak{z}_2(\cdot), \mathfrak{z}_1(\cdot)$ satisfy the assumptions of Lemma C.8, so a solution $\zeta_1^k, \zeta_2^k \in [-1,1]$ to this system exists for small enough ε^k. For convenience, define the expected outcomes

$$\tilde{\alpha}_2^k(h) \equiv \int \tilde{\sigma}_2^k(t,h)\, d\psi(z_2) = \begin{cases} 1 & h = C \\ 1 - \Delta\Psi_2(\zeta_2^k) & h = D \end{cases} \qquad \tilde{\alpha}_1^k \equiv \int \tilde{\sigma}_1^k(t)\, d\psi(z_1) = 1 - \Delta\Psi_1(\zeta_1^k)$$

for any $t \geq 1$. Also define the beliefs at history C at any period t: $\tilde{\mu}^k \equiv \mu^0/[\mu^0 + (1-\mu^0)\tilde{\alpha}_1^k]$.

I now show that these strategies are mutual best responses. First, 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 action $\bar{\alpha}_1$, and shock z_2:

$$V_2^c(p, \bar{\alpha}_1, z_2) = 3p + (1-p)[3\bar{\alpha}_1] + \varepsilon^k z_2^c, \qquad V_2^d(p, \bar{\alpha}_1, z_2) = 2p + (1-p)[2\bar{\alpha}_1 + (1-\bar{\alpha}_1)] + \varepsilon^k z_2^d.$$

Define ΔV_2 as the benefit of playing c over d: $\Delta V_2(p, \bar{\alpha}_1, z_2) \equiv V_2^c(p, \bar{\alpha}_1, z_2) - V_2^d(p, \bar{\alpha}_1, z_2)$. Then

$$\begin{aligned} \Delta V_2(p, \bar{\alpha}_1, z_2) &= 3p + (1-p)[3\bar{\alpha}_1] - 2p - (1-p)[2\bar{\alpha}_1 + (1-\bar{\alpha}_1)] + \varepsilon^k(z_2^c - z_2^d) \\ &= p + (1-p)[2\bar{\alpha}_1 - 1] + \varepsilon^k \Delta z_2. \end{aligned} \tag{41}$$

At history D, the belief is $p = 0$, giving $\Delta V_2(0, \tilde{\alpha}_1^k, z_2) = 2\tilde{\alpha}_1^k - 1 + \varepsilon^k \Delta z_2 = 2[1 - \Delta\Psi_1(\zeta_1^k)] - 1 + \varepsilon^k \Delta z_2 = 1 - 2\Delta\Psi_1(\zeta_1^k) + \varepsilon^k \Delta z_2$. Since ζ_2^k satisfies $0 = 1 - 2\Delta\Psi_1(\zeta_1^k) + \varepsilon^k \zeta_2^k$, $\tilde{\sigma}_2^k(t, D, z_2)$ 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 $\tilde{\sigma}_2^k(t, C, z_2)$ 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 $\Delta V_1(t, z_1) \equiv V_1^C(t, z_1) - V_1^D(t, z_1)$, giving

$$\begin{aligned} \Delta V_1(t, z_1) &= -(1-\delta) + \delta(V(t+1, C) - V(t+1, D)) + \varepsilon^k(z_1^C - z_1^D) \\ &= -(1-\delta) + 2(1-\delta)\delta(\bar{\alpha}_2^{t+1}(C) - \bar{\alpha}_2^{t+1}(D)) + \varepsilon^k \Delta z_1 \end{aligned} \tag{42}$$

where $\bar{\alpha}_2^{t+1}(h)$ is the probability of next period's player 2 choosing c at history h. Then

$$\begin{aligned} \Delta V_1(t, z_1) &= -(1-\delta) + 2(1-\delta)\delta(\tilde{\alpha}_2^k(C) - \tilde{\alpha}_2^k(D)) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + 2(1-\delta)\delta(1 - (1 - \Delta\Psi_2(\zeta_2^k))) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta)(1 - 2\delta\Delta\Psi_2(\zeta_2^k)) + \varepsilon^k \Delta z_1 = -\varepsilon^k \zeta_1^k + \varepsilon^k \Delta z_1, \end{aligned}$$

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 (7). 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. I show ζ_1^k converges to $\Delta\Psi_1^{-1}(\frac{1}{2})$. Suppose by contradiction that there exists some $\beta > 0$ so that for every k^* there existed $k > k^*$ such that $|\zeta_1^k - \Delta\Psi_1^{-1}(\frac{1}{2})| > \beta \implies |\Delta\Psi_1(\zeta_1^k) - \frac{1}{2}| > \beta'$ for some $\beta' > 0$. Then $|\zeta_2^k| = |\mathfrak{z}_2(\zeta_1^k|\varepsilon^k)| > \frac{1}{\varepsilon^k}\left|2\Delta\Psi_1(\zeta_1^k) - 1\right| > \frac{2}{\varepsilon^k}\beta'$, yet for ε^k small enough this contradicts the fact that a solution $\zeta_1^k, \zeta_2^k \in [-1,1]$ for the system of equations (40) exists.

A similar argument shows that ζ_2^k converges to $\Delta\Psi_2^{-1}\left(\frac{1}{2\delta}\right)$. Suppose by contradiction that there exists some $\beta > 0$ such that for every k^* there exists $k > k^*$ so that $|\zeta_2^k - \Delta\Psi_2^{-1}\left(\frac{1}{2\delta}\right)| > \beta \implies |\Delta\Psi_2(\zeta_2^k) - \frac{1}{2\delta}| > \beta'$. Then $|\zeta_1^k| = |\mathfrak{z}_1(\zeta_2^k|\varepsilon^k)| > \frac{1}{\varepsilon^k}(1-\delta)|1 - 2\delta\Delta\Psi_2(\zeta_2^k)| > \frac{1}{\varepsilon^k}(1-\delta)\beta'$, yet for ε^k small enough this contradicts $\zeta_1^k, \zeta_2^k \in [-1,1]$. Thus, as $k \to \infty$, $\tilde{\alpha}_2^k(D) \to \sigma_2^*(t,D)$ and $\tilde{\alpha}_1^k \to \sigma_1^*(t)$ completing the proof that $(\tilde{\sigma}^k)_k$ converge in outcomes to σ^*.

C.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

$$\sigma_2^*(t,h) = \begin{cases} 0 & h = \varnothing \\ \frac{1}{2\delta} & h = C \\ 0 & h = D \end{cases} \qquad \sigma_1^*(t) = \frac{1}{4} \pm \frac{1}{4}\sqrt{1 - 8\phi^*}.$$

Note that $\sigma_1^*(t)$ is a fixed point of $r(\cdot) \equiv q^{-1}(\cdot)$ defined in the proof of Proposition 5.1; see (29) 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 ζ_1^k, ζ_2^k is defined by the system of equations

$$\zeta_2^k \equiv \mathfrak{z}_2(\zeta_1^k, \varepsilon^k) \equiv \frac{1}{\varepsilon^k}\left[1 - 2\frac{(1-\Delta\Psi_1(\zeta_1^k))\Delta\Psi_1(\zeta_1^k)}{\phi^* + (1-\Delta\Psi_1(\zeta_1^k))}\right],\ \zeta_1^k \equiv \mathfrak{z}_1(\zeta_2^k, \varepsilon^k) \equiv \frac{1-\delta}{\varepsilon^k}[1 - 2\delta(1-\Delta\Psi_2(\zeta_2^k))]. \tag{43}$$

Note that because $\mu^0 < \frac{1}{9}$ implies $\phi^* < \frac{1}{8}$,

$$\mathfrak{z}_2(\Delta\Psi_1^{-1}(\tfrac{2}{3}), \varepsilon^k) = \frac{1}{\varepsilon^l}\left[1 - 2\frac{(1-\frac{2}{3})\cdot\frac{2}{3}}{\phi^* + \frac{1}{4}}\right] < \frac{1}{\varepsilon^k}\left[1 - \frac{\frac{4}{9}}{\frac{1}{8}+\frac{1}{4}}\right] = \frac{1}{\varepsilon^k}\left[1 - \frac{32}{27}\right] < -\frac{1}{\varepsilon^k}.$$

Furthermore, $\mathfrak{z}_2(1,\varepsilon^k) = \frac{1}{\varepsilon^k}, \mathfrak{z}_1(-1,\varepsilon^k) = \frac{1-\delta}{\varepsilon^k}, \frac{1}{\varepsilon^k}, \mathfrak{z}_1(1,\varepsilon^k) = \frac{1}{\varepsilon^k}(1-\delta)(1-2\delta)$, and so by Lemma C.8 a solution exists for small enough ε^l. For any $t \geq 1$ the expected outcomes are

$$\tilde{\alpha}_2^k(h) \equiv \int \tilde{\sigma}_2^k(t,h)\, d\psi(z_2) = \begin{cases} 1 - \Delta\Psi_2(\zeta_2^k) & h = C \\ 0 & h = D \end{cases} \qquad \tilde{\alpha}_1^k \equiv \int \tilde{\sigma}_1^k(t)\, d\psi(z_1) = 1 - \Delta\Psi_1(\zeta_1^k).$$

I show that $\tilde{\sigma}^k$ are mutual best responses. Note that $\tilde{\mu}^k = \mu^0/[\mu^0 + (1-\mu^0)(1-\Delta\Psi_1(\zeta_1^k))]$. Reusing the notation above in (41), at history C the belief is $p = \tilde{\mu}^k$, so

$$\begin{aligned}\Delta V_2(\tilde{\mu}^k, \tilde{\alpha}_1^k(C), z_2) &= \tilde{\mu}^k + (1-\tilde{\mu}^k)[\tilde{\alpha}_1^k(C) - (1-\tilde{\alpha}_1^k(C))] + \varepsilon^k \Delta z_2 \\ &= \tilde{\mu}^k + (1-\tilde{\mu}^k)[2(1-\Delta\Psi_1(\zeta_1^k)) - 1] + \varepsilon^k \Delta z_2.\end{aligned}$$

Rearranging and substituting $\tilde{\mu}^k$ gives

$$\begin{aligned}\Delta V_2(\tilde{\mu}^k, \tilde{\alpha}_1^k, z_2) &= 1 - 2(1-\tilde{\mu}^k)\Delta\Psi_1(\zeta_1^k) + \varepsilon^k \Delta z_2 \\ &= 1 - 2\left(1 - \frac{\mu^0}{\mu^0 + (1-\mu^0)(1-\Delta\Psi_1(\zeta_1^l))}\right)\Delta\Psi_1(\zeta_1^k) + \varepsilon^k \Delta z_2 \\ &= 1 - 2\frac{(1-\mu^0)(1-\Delta\Psi_1(\zeta_1^k))}{\mu^0 + (1-\mu^0)(1-\Delta\Psi_1(\zeta_1^k))}\Delta\Psi_1(\zeta_1^k) + \varepsilon^k \Delta z_2 \\ &= 1 - 2\frac{(1-\Delta\Psi_1(\zeta_1^k))\Delta\Psi_1(\zeta_1^k)}{\phi^* + (1-\Delta\Psi_1(\zeta_1^k))} + \varepsilon^k \Delta z_2 = -\varepsilon^k \zeta_2^k + \varepsilon^k \Delta z_2,\end{aligned}$$

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 $\tilde{\sigma}_2^k(\varnothing, z_2), \tilde{\sigma}_2^k(t, D, z_2)$ are best responses. Reusing the notation in (42),

$$\begin{aligned}0 &= -(1-\delta) + 2(1-\delta)\delta(\tilde{\alpha}_2^k(C) - \tilde{\alpha}_2^k(D)) + \varepsilon^k \Delta z_1 \\ &= -(1-\delta) + 2(1-\delta)\delta(1 - \Delta\Psi_2(\zeta_2^k)) + \varepsilon^k \Delta z_1 = -\varepsilon^k \zeta_1^k + \varepsilon^k \Delta z_1,\end{aligned}$$

so player 1's strategy $\tilde{\sigma}_1^k$ 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. I show ζ_1^k converges to $\Delta\Psi_1^{-1}\left(\frac{3}{4} \pm \sqrt{1-8\phi^*}\right)$. The fact that a solution $\zeta_1^k, \zeta_2^k \in [-1, 1]$ for the system (43) requires that

$$|\zeta_2^k| = |\mathfrak{z}_2(\zeta_1^k|\varepsilon^k)| = \frac{1}{\varepsilon^l}\left|1 - 2\frac{(1-\Delta\Psi_1(\zeta_1^k))\Delta\Psi_1(\zeta_1^k)}{\phi^* + (1-\Delta\Psi_1(\zeta_1^k))}\right| \leq 1 \implies \lim_{k\to\infty} \frac{(1-\Delta\Psi_1(\zeta_1^k))\Delta\Psi_1(\zeta_1^k)}{\phi^* + (1-\Delta\Psi_1(\zeta_1^k))} = \frac{1}{2},$$

so $(\Delta\Psi_1(\zeta_1^k))_k$ converges; let $\bar{p}_1 \equiv \lim_k \Delta\Psi_1(\zeta_1^k)$. Then

$$\frac{(1-\bar{p})\bar{p}}{\phi^* + (1-\bar{p})} = \frac{1}{2} \quad\implies\quad 2\bar{p} - 2\bar{p}^2 = \phi^* + 1 - \bar{p} \quad\implies\quad -2\bar{p}^2 + 3\bar{p} - (\phi^* + 1) = 0$$

so $\bar{p} = \frac{1}{-4}\left(-3 \pm \sqrt{9 - 4\cdot 2(\phi^*+1)}\right) = \frac{3}{4} \pm \frac{1}{4}\sqrt{9 - 4\cdot 2(\phi^*+1)} = \frac{3}{4} \pm \frac{1}{4}\sqrt{1-8\phi^*}$. Thus, $\Delta\Psi_1(\zeta_1^k) \to \frac{3}{4} \pm \frac{1}{4}\sqrt{1-8\phi^*} \implies \tilde{\alpha}_1^k = 1 - \Delta\Psi_1(\zeta_1^k) \to \frac{1}{4} \pm \frac{1}{4}\sqrt{1-8\phi^*} = \sigma_1^*(t)$. A very similar argument to that at the end of Section C.2.1 proves that $\tilde{\alpha}_2(C) \to \sigma_2^*(t, C)$. Thus, $(\tilde{\sigma}^k)_k$ converge in outcomes to σ^*.

C.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 5.1 (I will usually omit the "μ^0" argument in both for brevity). Define $q^0(0) = 0$ for convenience. I restrict attention to priors μ^0 such that $\frac{1}{2} < q^L(0, \mu^0) < 1$ (the set of priors $\mu^0 \in (\frac{1}{9}, \frac{1}{2}]$ where $L(\mu^0) \in \{\frac{1}{2}, 1\}$ is Lebesgue measure zero). Define $\bar{l}(t) \equiv (t+1) \bmod (L+1)$. The equilibrium strategies are

$$\sigma_2^*(t,h) = \begin{cases} 0 & h = \varnothing \quad \text{or} \quad \bar{l}(t) \neq 0, h = D \\ \frac{1}{2\delta} & \bar{l}(t) \neq 0, h = C \\ 1 & \bar{l}(t) = 0 \end{cases} \qquad\qquad \sigma_1^*(t) = q^{L-\bar{l}(t)}(0).$$

Define the sequence $(\tilde{\sigma}^k)_k$ of equilibria in the sequence of perturbed games:

$$\tilde{\sigma}_2^k(t,h,z_2) \equiv \begin{cases} 0 & \varepsilon^l \Delta z_2 \leq \zeta_2^{k,\bar{l}(t)}(h) \\ 1 & \varepsilon^l \Delta z_2 > \zeta_2^{k,\bar{l}(t)}(h) \end{cases} \qquad\qquad \tilde{\sigma}_1^k(t,z_1) \equiv \begin{cases} 0 & \varepsilon^k \Delta z_1 \leq \zeta_1^{k,\bar{l}(t)} \\ 1 & \varepsilon^k \Delta z_1 > \zeta_1^{k,\bar{l}(t)} \end{cases}$$

where I define the thresholds $\zeta_2^{k,0}(C), \zeta_2^{k,0}(D), \zeta_1^{k,0}, ..., \zeta_2^{k,L}(C), \zeta_2^{k,L}(D), \zeta_1^{k,L}$ below; for the empty history, $\zeta_2^{k,1}(\varnothing) \equiv 1$. For convenience and clarity, define expected outcomes

$$\tilde{\alpha}_1^{k,l} \equiv \int \tilde{\sigma}_1^k(t,z_1)\, d\psi(z_1) = 1 - \Delta\Psi_1(\zeta_1^{k,l}) \qquad\qquad \tilde{\alpha}_2^{k,l}(h) \equiv \int \tilde{\sigma}_2^k(t,h,z_2)\, d\psi(z_2) = 1 - \Delta\Psi_1(\zeta_2^{k,l}(h))$$

and beliefs $\tilde{\mu}^{k,l} \equiv \mu^0/[\mu^0 + (1-\mu^0)\tilde{\alpha}_1^{k,(l-1) \bmod (L+1)}]$ for some t such that $\bar{l}(t) = l$. Define the thresholds as follows. Let $\zeta_2^{k,l}(D) \equiv 1$ for all $l \geq 1$, and let $\zeta_1^{k,L} \equiv 1$. Define $\zeta_1^{k,L-1}, \zeta_2^{k,L}(C)$ as solutions to the system of equations

$$\zeta_2^{k,L}(C) \equiv \mathfrak{z}_2^L(\zeta_1^{k,L-1}|\varepsilon^k) \equiv \frac{1}{\varepsilon^k}[1 - 2\tilde{\mu}^{k,L}] \tag{44}$$

$$\zeta_1^{k,L-1} \equiv \mathfrak{z}_1(\zeta_2^{k,L}|\varepsilon^k) \equiv \frac{1-\delta}{\varepsilon^k}[1 - 2\delta(1 - \Delta\Psi_2(\zeta_2^{k,L}))]. \tag{45}$$

(note that $\tilde{\mu}^{k,l}$ is a function of $\zeta_1^{k,l-1}$ for each $l \geq 1$). Then for each $l \in \{1, ..., L-1\}$, inductively define $\zeta_1^{k,l-1}, \zeta_2^{k,l}$, taking $\tilde{\alpha}_1^{k,l}$ as given, as solutions to the equations

$$\zeta_2^{k,l}(C) \equiv \mathfrak{z}_2^l(\zeta_1^{k,l-1}|\varepsilon^k) \equiv -\frac{1}{\varepsilon^k}[\tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})(2\tilde{\alpha}_1^{k,l} - 1)] \tag{46}$$

$$\zeta_1^{k,l-1} \equiv \mathfrak{z}_1(\zeta_2^{k,l}|\varepsilon^k) = \frac{1-\delta}{\varepsilon^k}[1 - 2\delta(1 - \Delta\Psi_2(\zeta_2^{k,l}(C)))]. \tag{47}$$

I show the existence of a solution below. Define $\zeta_2^{k,0}(C) \equiv \zeta_2^{k,0}(D) \equiv -2$ (i.e. defined so the consumer chooses c for all shocks $z_2 \in Z_2$).

Before continuing, the following result will be useful proving that the perturbed equilibria converge in outcomes to σ^*.

Lemma C.9. *There exists finite $J > 0$ such that there exists $\bar{\varepsilon} > 0$ so that for all $\varepsilon \in (0, \bar{\varepsilon})$, $l < L$,*

$$q^{L-l}(0) - J^{L-l}\varepsilon < \frac{(1 + J^{L-l-1}\varepsilon)\phi^*}{1 - 2q^{L-l-1}(0) + J^{L-l-1}\varepsilon} < \frac{(1 - J^{L-l-1}\varepsilon)\phi^*}{1 - 2q^{L-l-1}(0) - J^{L-l-1}\varepsilon} < q^{L-l}(0) + J^{L-l}\varepsilon. \quad (48)$$

Proof. Define functions $\mathfrak{f}_{\pm}(\varepsilon, l) \equiv \frac{1 \mp J^{L-l-1}\varepsilon}{1-2q^{L-l-1}(0) \mp J^{L-l-1}\varepsilon}$. Differentiating gives

$$\mathfrak{g}_{\pm}(\varepsilon, l) \equiv \frac{\partial \mathfrak{f}_{\pm}(\varepsilon, l)}{\partial \varepsilon} = \frac{\pm 2J^{L-l-1}q^{L-l-1}(0)}{(1 - 2q^{L-l-1}(0) \mp J^{L-l-1}\varepsilon)^2} \implies \mathfrak{g}_{\pm}(0, l) = \pm J^{L-l-1}\frac{2q^{L-l-1}(0)}{(1 - 2q^{L-l-1}(0))^2}$$

for all l. Choosing some $J > \frac{2q^{L-l-1}(0)}{(1-2q^{L-l-1}(0))^2}$ yields $J^{L-l} > \mathfrak{g}_+(0, l)$ and $-J^{L-l} < \mathfrak{g}_-(0, l)$. Recall that $q^{L-l}(0) = \frac{\phi^*}{1-2q^{L-l-1}(0)}$. Then for small enough ε,

$$\mathfrak{f}_-(0, l) - J^{L-l}\varepsilon < \mathfrak{f}_-(\varepsilon, l) < \mathfrak{f}_+(\varepsilon, l) < \mathfrak{f}_+(0, l) + J^{L-l}\varepsilon,$$

and since $\mathfrak{f}_-(0, l) = \mathfrak{f}_+(0, l) = q^{L-l}(0) = \frac{\phi^*}{1-2q^{L-l-1}(0)}$, it has been shown that (48) holds for small enough ε. ■

I show that for small enough ε^k, a solution to the system (44) - (47) exists. Note that $\mathfrak{z}_2^L(-1, \varepsilon^k) > 1$ for small enough ε^k (since $\zeta_1^{k,L-1} = -1$ implies $\tilde{\mu}^{k,L} = \mu^0 < \frac{1}{2}$) and $\mathfrak{z}_2^L(1, \varepsilon^k) = -\frac{1}{\varepsilon^k}$. Further note that $\mathfrak{z}_1(\zeta_2, \varepsilon^k)$ is defined identically to (43), and so it is straightforward to check that $\mathfrak{z}_2^L(\cdot, \varepsilon^k), \mathfrak{z}_1(\cdot)$ satisfy the assumptions of Lemma C.8. Thus, the system (44) and (45) has a solution $\zeta_2^{k,L-1}(C), \zeta_1^{k,L-2} \in [-1, 1]$. This implies that

$$-1 \leq \zeta_2^{k,L}(C) = \frac{1}{\varepsilon^k}[1 - 2\tilde{\mu}^{k,L}] \leq 1 \qquad \implies \qquad -\varepsilon^k \leq 1 - 2\tilde{\mu}^{k,L} \leq \varepsilon^k$$

$$1 + \varepsilon^k \geq 2\tilde{\mu}^{k,L} = \frac{2\mu^0}{\mu^0 + (1 - \mu^0)\tilde{\alpha}_1^{k,L-1}} = \frac{2\phi^*}{\phi^* + \tilde{\alpha}_1^{k,L-1}} \geq 1 - \varepsilon^k$$

which can be rearranged to find

$$(1 + \varepsilon^k)\phi^* + (1 + \varepsilon^k)\tilde{\alpha}_1^{k,L-1} \geq 2\phi^* \geq (1 - \varepsilon^k)\phi^* + (1 - \varepsilon^k)\tilde{\alpha}_1^{k,L-1}$$

$$\varepsilon^k\phi^* + (1 + \varepsilon^k)\tilde{\alpha}_1^{k,L-1} \geq \phi^* \geq -\varepsilon^k\phi^* + (1 - \varepsilon^k)\tilde{\alpha}_1^{k,L-1}$$

$$\frac{1 - \varepsilon^k}{1 + \varepsilon^k}\phi^* \leq \tilde{\alpha}_1^{k,L-1} \leq \frac{1 + \varepsilon^k}{1 - \varepsilon^k}\phi^*.$$

Recall that $q^0(0) = 0$ and $q^1(0) = \phi^*$. Then Lemma C.9 shows that there exists finite $J > 0$ such that $\tilde{\alpha}_1^{k,L-1} \in B_{J\varepsilon}(q^1(0))$ where $B_\beta(v)$ is the open ball of radius β centered at v.

Suppose that for some l and some finite $J > 0$, $\tilde{\alpha}_1^{k,l} \in B_{J^{L-l-1}\varepsilon}(q^{L-l-1}(0))$. Note that

$$\begin{aligned} \mathfrak{z}_2^l(\zeta_1^{k,l}|\varepsilon^k) &= -\frac{1}{\varepsilon^k}\left[\tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})(2\tilde{\alpha}_1^{k,l} - 1)\right] \\ &> -\frac{1}{\varepsilon^k}\left[\tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})(2(q^{L-l-1}(0) + J(L-l-1)\varepsilon^k) - 1)\right]. \end{aligned}$$

Recall that $r(\phi) \equiv q^{-1}(\phi) = \frac{1}{2}\left(1 - \frac{\phi^*}{\phi}\right)$. Note that since $q^L(0) < 1$, $q^{L-1}(0) < r(1) = \frac{1}{2}(1 - \phi^*)$; thus, for small enough ε^k, $q^{L-l-1}(0) \leq q^{L-1}(0) < \frac{1}{2}(1-\phi^*) - \left(J(L-l-1) + \frac{1}{1-\mu^0}\right)\varepsilon^k$. Furthermore, choosing $\zeta_1^{k,l-1} = -1$ (i.e. the firm chooses C with probability 1) yields $\tilde{\mu}^{k,l} = \mu^0$ and so

$$\begin{aligned} &-\varepsilon^k \mathfrak{z}_2^l(-1|\varepsilon^k) \\ <\ & \mu^0 + (1-\mu^0)\left(2\left(\frac{1}{2}\left(1 - \frac{\mu^0}{1-\mu^0}\right) - \left(J(L-l-1) + \frac{1}{1-\mu^0}\right)\varepsilon^k + J(L-l-1)\varepsilon^k\right) - 1\right) \\ =\ & \mu^0 + (1-\mu^0)\left(\frac{-\mu^0 - \varepsilon^k}{1-\mu^0}\right) = -\varepsilon^k, \end{aligned}$$

i.e. $\mathfrak{z}_2^l(-1|\varepsilon^k) > 1$. Note that $\mathfrak{z}_2^l(1|\varepsilon^k) = -\frac{1}{\varepsilon^k}$. It is then straightforward to check that $\mathfrak{z}_2^l(\cdot|\varepsilon^k), \mathfrak{z}_1(\cdot|\varepsilon^k)$ satisfy the assumptions of Lemma C.8, which shows that for l the system of equations (46) and (47) have a solution $\zeta_2^{k,l}, \zeta_1^{k,l} \in [-1, 1]$. This implies that

$$-1 < \zeta_2^{k,l}(C) = \frac{1}{\varepsilon^k}[-\tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})(1 - 2\tilde{\alpha}_1^{k,l})] < 1$$

$$-\varepsilon^k < -\tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l}) - 2(1-\tilde{\mu}^{k,l})\tilde{\alpha}_1^{k,l} = 1 - 2\tilde{\mu}^{k,l} - 2(1-\tilde{\mu}^{k,l})\tilde{\alpha}_1^{k,l} < \varepsilon^k$$

Rearranging and using the fact that $\tilde{\alpha}_1^{k,l} \in B_{J^{L-l-1}\varepsilon}(q^{L-l-1}(0))$ yields

$$2(1-\tilde{\mu}^{k,l})\tilde{\alpha}_1^{k,l} - \varepsilon^k < 1 - 2\tilde{\mu}^{k,l} < 2(1-\tilde{\mu}^{k,l})\tilde{\alpha}_1^{k,l} + \varepsilon^k$$

$$2(1-\tilde{\mu}^{k,l})(q^{L-l-1}(0) - J^{L-l-1}\varepsilon^k) - \varepsilon^k < 1 - 2\tilde{\mu}^{k,l} < 2(1-\tilde{\mu}^{k,l})(q^{L-l-1}(0)) + J^{L-l-1}\varepsilon^k) + \varepsilon^k$$

$$2(1-\tilde{\mu}^{k,l})q^{L-l-1}(0) - J^{L-l}\varepsilon^k < 1 - 2\tilde{\mu}^{k,l} < 2(1-\tilde{\mu}^{k,l})q^{L-l-1}(0) + J^{L-l}\varepsilon^k$$

which can be further rearranged to

$$2q^{L-l-1}(0) - J^{L-l}\varepsilon^k < 1 - 2(1 - q^{L-l-1}(0))\tilde{\mu}^{k,l} < 2q^{L-l-1}(0) + J^{L-l}\varepsilon^k.$$

Substituting Bayes' rule for $\tilde{\mu}^{k,l}$ gives

$$-1 + 2q^{L-l-1}(0) - J^{L-l}\varepsilon^k < \frac{-2(1 - q^{L-l-1}(0))\mu^0}{\mu^0 + (1-\mu^0)\tilde{\alpha}_1^{k,l-1}} < -1 + 2q^{L-l-1}(0) + J^{L-l}\varepsilon^k$$

$$\frac{2(1-q^{L-l-1})\phi^*}{1-2q^{L-l-1}(0)-J^{L-l}\varepsilon^k} > \phi^* + \tilde{\alpha}_1^{k,l-1} > \frac{2(1-q^{L-l-1}(0))\phi^*}{1-2q^{L-l-1}(0)+J^{L-l}\varepsilon^k}$$
$$\frac{(1-J^{L-l}\varepsilon^k)\phi^*}{1-2q^{L-l-1}(0)-J^{L-l}\varepsilon^k} > \tilde{\alpha}_1^{k,l-1} > \frac{(1+J^{L-l}\varepsilon^k)\phi^*}{1-2q^{L-l-1}(0)+J^{L-l}\varepsilon^k}$$

Finally, Lemma 48 yields $\tilde{\alpha}_1^{k,l-1} \in B_{J^{L-l}\varepsilon^k}\left(q^{L-l}(0)\right)$. By induction, $\tilde{\alpha}_1^{k,0} = B_{J^L\varepsilon^k}\left(q^L(0)\right)$.

I show that these are mutual best responses. For any t, let $l \equiv \bar{l}(t)$. I sometimes abuse notation by letting "l" take the place of "t," e.g. $\sigma_1^*(l)$ instead of $\sigma_1^*(t)$, to mean some period t satisfying $\bar{l}(t) = l$. For all $l \geq 1$, the benefit to player 2 of playing c over that of d at history C is

$$\Delta V_2(\tilde{\mu}^{k,l}, \tilde{\alpha}_1^{k,l}, z_2) \quad = \quad \tilde{\mu}^{k,l} + (1-\tilde{\mu}^{k,l})[2\tilde{\alpha}_1^{k,l} - 1] + \varepsilon^k \Delta z_2 = -\varepsilon^k \zeta_2^{k,l}(C) + \varepsilon^k \Delta z_2$$

so $\tilde{\sigma}_2^k(l, C, z_2)$ is a best response. For both the empty history $\varnothing$ and history D, the belief is 0, so for all $l \in \{1, ..., L\}$ and small enough ε^k, $\Delta V_2(0, \tilde{\alpha}_1^{k,l}, z_2) < 0$ 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 $l = 0$, $\tilde{\alpha}_1^{k,0} = B_{J^L\varepsilon^k}\left(q^L(0)\right)$. Since $q^L(0) > \frac{1}{2}$, for small enough ε^k, $\Delta V_2(0, \tilde{\alpha}_1^{k,0}, z_2) > 0$ for all shocks $z_2 \in Z_2$ and so $\tilde{\sigma}_2^k(l = 0, C, z_2) = \tilde{\sigma}_2^k(l = 0, D, z_2) = c$ is a best response. Turning to player 1, for $0 \leq l < L$, I have

$$\begin{aligned}
\Delta V_1(l, z_1) \quad &= \quad -(1-\delta) + 2(1-\delta)\delta(\tilde{\alpha}_2^{k,l+1}(C) - \tilde{\alpha}_2^{k,l+1}(D)) + \varepsilon^k \Delta z_1 \\
&= \quad (1-\delta)[-1 + 2\delta(1 - \Delta\Psi_2(\zeta_2^{k,l+1}))] + \varepsilon^k \Delta z_1 = -\varepsilon^k \zeta_1^{k,l} + \varepsilon^k \Delta z_1
\end{aligned}$$

so $\tilde{\sigma}_1^k(l, z_1)$ is a best response. For $l = L$, $\Delta V_1(l, z_1) = -(1-\delta) + 2(1-\delta)\delta(\tilde{\alpha}_2^{k,0}(C) - \tilde{\alpha}_2^{k,0}(D)) + \varepsilon^k \Delta z_1 = -(1-\delta) + \varepsilon^k \Delta z_1$ and so for small enough ε^k, $\tilde{\sigma}_1^k(l = L-1, z_1) = D$ for all $z_1 \in Z_1$ is a best response.

The arguments above showed that $\tilde{\alpha}_1^{k,l} \in B_{J^{L-l-1}\varepsilon}(p^{L-l-1})$ for each l, which implies that $\lim_{k\to\infty} \tilde{\alpha}_1^{k,l} = \sigma_1^*(l)$. The fact that $\zeta_1^{k,l} \in [-1, 1]$ for $l < L$ implies that as $\varepsilon^k \to 0$, $(1-\delta)[1 - 2\delta\tilde{\alpha}_2^{k,l}(C)] \to 0$ and hence $\tilde{\alpha}_k^{k,l}(C) \to \frac{1}{2\delta}$. Convergence for $\tilde{\sigma}_2^{k,l}(l, D, z_2) = 0$ for $l \geq 1$ and $\tilde{\sigma}_2(l = 0, C, z_2) = \tilde{\sigma}_2(l = 0, D, z_2) = 1$ is trivial. Thus, $(\tilde{\sigma}^k)_k$ converge in outcomes to σ^*.

C.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 $p(t, h)$ where $p(t, C) = p(t, \varnothing) = \mu^0$ for even t, $p(t, C) = \mu^0$ for odd t, and $p(t, h) = 0$ for $h = D$. The benefit of playing c over d is given by

$$\begin{aligned}\Delta V_2(t,h,\bar{\alpha}_1,z_2) &= p(t,h)+(1-p(t,h))[2\bar{\alpha}_1-1]+\varepsilon^k\Delta z_2\\ &= \begin{cases}\mu^0-(1-\mu^0)+\varepsilon^k\Delta z_2 & \text{even } t, h=C\\ -1+\varepsilon^k\Delta z_2 & \text{even } t, h=D\\ 1+\varepsilon^k\Delta z_2 & \text{odd } t, h=C\\ \mu^0+(1-\mu^0)+\varepsilon^k\Delta z_2 & \text{odd } t, h=D;\end{cases}\end{aligned}$$

so for small enough ε^k, $\tilde{\sigma}_2^k$ is a best response for all z_2 (recall $\mu^0>\frac{1}{2}$). For player 1,

$$\Delta V_1(t,z_1)=\begin{cases}-(1-\delta)+\varepsilon^k\Delta z_1 & \text{even } t\\ (1-\delta)[-1+2\delta]+\varepsilon^k\Delta z_1 & \text{odd } t,\end{cases}$$

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 clear.

C.3 Non-Quasi-Markov Equilibria

Define the firm'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 the consumer indifferent at all histories $\varnothing,C,D$ (at all periods). To keep the firm 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 the firm's incentives. The equilibrium payoff can be calculated by evaluating the strategy of playing D every period (a best response):

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

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 in $[1,2]$.

C.4 Proof of Proposition 5.2

By Proposition 3.4, for any stationary public wPBE (σ,μ), there exists a self-generating HBP (ϕ,μ,γ) enforcing an HBA (ϕ,μ,α) such that $\sigma_2(h)=\alpha_2(h)$, $\sigma_1(h)=\alpha_1(h)$, 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 begin with a lemma bounding the "clean history" payoff from below by exactly the Stackelberg payoff (recall that $I_K=\{C^K\}$).

Lemma C.10. *Suppose $\mu^0>0$. There exists $K^*(\mu^0)$ such that for all $K>K^*$, $\gamma(I_K)\geq 2$.*

Proof. Pick $K^* > 1/2\phi^*$. Suppose by contradiction that $\gamma(I_K) < 2$. This implies that $\alpha_2(I_K) < 1$, for otherwise player 2 would always play c at I_K, yielding $\gamma(I_K) \geq (1-\delta)u(c,C) + \delta V(I_K) = 2(1-\delta) + \delta\gamma(I_K) \implies \gamma(I_K) \geq 2$, a contradiction. Thus d is a best reply for player 2 at I_K, so I have

$$2[\mu(I_K) + (1-\mu(I_K))\alpha_1(I_K)] + (1-\mu(I_K))(1-\alpha_1(I_K)) \geq 3[\mu(I_K) + (1-\mu(I_K))\alpha_1(I_K)].$$

Solving for $\alpha_1(I_K)$ gives $\alpha_1(I_K) \leq \frac{1-2\mu(I_K)}{2(1-\mu(I_K))}$. Substituting Bayes' rule yields

$$\alpha_1(I_K) \leq \frac{1 - \frac{2\mu^0}{\mu^0+(1-\mu^0)\phi(I_K)}}{2\left(1 - \frac{\mu^0}{\mu^0+(1-\mu^0)\phi(I_K)}\right)} = \frac{(1-\mu^0)\phi(I_K) - \mu^0}{2(1-\mu^0)\phi(I_K)} = \frac{1}{2}\left(1 - \frac{\phi^*}{\phi(I_K)}\right). \tag{49}$$

Since $\alpha_1(I_K) \geq 0$, we can bound $\phi(I_K)$ from below: $0 \leq 1 - \frac{\phi^*}{\phi(I_K)} \implies \phi(I_K) \geq \phi^*$. Inducibility (see (1)) requires

$$\phi(I_K) = \phi(I_K)\alpha_1(I_K) + \phi(I_{K-1})\alpha_1(I_{K-1}) \leq [\phi(I_K) + \phi(I_{K-1})] \cdot \frac{1}{2}\left(1 - \frac{\phi^*}{\phi(I_K)}\right) \tag{50}$$

which can be rearranged to $\phi(I_{K-1}) \geq \frac{\phi(I_K)+\phi^*}{1-\phi^*/\phi(I_K)} > 2\phi^*$. Inducibility also requires that $2\phi^* \leq \phi(I_{K-1}) \leq \phi(I_{K-2}) \leq \cdots \leq \phi(I_0)$. Then $\sum_{k=0}^{K}\phi(I_k) \geq 2K\phi^* > 1$ for all $K \geq K^*$, which violates Definition 3.1 for history distribution ϕ, a contradiction. ■

Lemma C.11. *$\gamma(I_0) = 1$ or $\gamma(I_k) \geq \gamma(I_0) + \eta$ for all $k \in \{1, ..., K\}$.*

Proof. Suppose by contradiction that $\gamma(I_0) > 1$ and there exists $k \in \{1, ..., K\}$ such that $\gamma(I_k) < \gamma(I_0) + \eta$. Then at histories $h \in I_{k-1}$, C is not a best reply so $\alpha_1(I_{k-1}) = 0$. At I_{k-1} the belief is $\mu(I_{k-1}) = 0$, so player 2's payoff for playing c is 0 while the payoff for d is 1, so d is the strict best response: $\alpha_2(I_{k-1}) = 0$. Thus, $\gamma(I_{k-1}) = (1-\delta)u_1(d,D) + \delta V(I_0) = (1-\delta) + \delta\gamma(I_0) \leq \gamma(I_0) < \gamma(I_0) + \eta$. 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. ■

For the rest of the proof, assume $K > K^*$, and so $\gamma(I_K) \geq 2$. Suppose $\gamma(I_K) > \gamma(I_0) + \eta$. Since player 1's best response at I_K 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} (at period $K-1$) is C, so player 2's best response at C^{K-1} is c: $V_\sigma(C^{K-1}) = (1-\delta)u(c,C) + \delta\gamma(I_K) = 2$. By backwards induction, $V_\sigma(C^k) = 2$ for all $k \in \{0, ..., K-1\}$, so $V(\sigma) = 2$.

Now suppose $\gamma(I_K) < \gamma(I_0) + \eta$. Lemma C.11 implies that $\gamma(I_0) = 1$. Since $\delta > \frac{1}{2}$, I have $\gamma(I_0) > \gamma(I_K) - \frac{1-\delta}{\delta} \geq 2 - \frac{1-\delta}{\delta} > 1$, a contradiction.

For the remainder of the proof suppose that $\gamma(I_K) = \gamma(I_0) + \eta$. Note that since C is a best response, $\gamma(I_K) = 2(1-\delta) + \delta\gamma(I_K)$, so then $\gamma(I_K) = 2 \implies \gamma(I_0) = 2 - \eta$. By Lemma C.11 this implies $\gamma(I_{K-1}) \geq 2$, but since C is a best response, $\gamma(I_{K-1}) = 2(1-\delta) + \delta\gamma(I_K) = 2$. Note that this requires $\sigma_2(I_K) = \sigma_2(I_{K-1}) = 1$.

Since (σ, μ) is purifiable, by Proposition 4.2 there exists a sequence $(\psi^l, \varepsilon^l)_l$, such that $\psi^l \in \Delta^*(Z)$ and $\varepsilon^l \to 0$, and a sequence of strategy profiles $(\tilde{\sigma}^l)_l$, such that $\tilde{\sigma}^l$ is a wPBE of the (ψ^l, ε^l)-perturbed game, which converges in outcomes to σ. By Proposition 4.1, each $\tilde{\sigma}^l$ is essentially sequentially strict and hence quasi-Markov. I prove the following claim:

Claim C.1. Suppose that for some $j \in \{1, ..., K-1\}$: there exists l^* such that for all $l > l^*$, $V_{\tilde{\sigma}^l}(I_j^j) \geq V_{\tilde{\sigma}^l}(I_K)$, $V_{\tilde{\sigma}^l}(I_k^j) = V_{\tilde{\sigma}^l}(I_k)$ for all $k < j$, and $\lim_{l \to \infty} V_{\tilde{\sigma}^l}(I_j^j) = 2$. Then $V_{\tilde{\sigma}^l}(I_{j-1}^{j-1}) \geq V_{\tilde{\sigma}^l}(I_K)$, $V_{\tilde{\sigma}^l}(I_k^{j-1}) = V_{\tilde{\sigma}^l}(I_k)$ for all $k < j-1$, and $\lim_{l \to \infty} V_{\tilde{\sigma}^l}(I_{j-1}^{j-1}) = 2$.

First, note that the decision problem for the firm at I_{k-1}^{j-1} is identical to that at I_k for all $k \in \{1, ..., j-1\}$. Thus, $\tilde{\sigma}_1^l(I_{k-1}^{j-1}) = \tilde{\sigma}_1^l(I_{k-1})$. For the consumer at I_{k-1}^{j-1}, the belief (zero) and the firm's expected behavior $\tilde{\sigma}_1^l(I_{k-1}^{j-1})$ — and therefore the decision problem — are identical at I_{k-1}. For almost all shocks $z_2 \in Z_2$, the consumer has the same strict best response $\tilde{\sigma}_2^l(I_{k-1}^{j-1}, z_2) = \tilde{\sigma}_2^l(I_{k-1}, z_2)$ and so $\int \tilde{\sigma}_2^l(I_{k-1}^{j-1}, z_2)\, d\psi(z_2) = \int \tilde{\sigma}_2^l(I_{k-1}, z_2)\, d\psi(z_2)$. Thus, $V_{\tilde{\sigma}^l}(I_{k-1}^{j-1}) = V_{\tilde{\sigma}^l}(I_k^j) = V_{\tilde{\sigma}^l}(I_k)$.

Consider the decision problem for the firm at the clean initial history I_{j-1}^{j-1}. Note that the intertemporal incentive is at least as strong at I_{j-1}^{j-1} as at I_{K-1}: $V_{\tilde{\sigma}^l}(I_j^j) - V_{\tilde{\sigma}^l}(I_0^j) \geq V_{\tilde{\sigma}^l}(I_{K-1}) - V_{\tilde{\sigma}^l}(I_0)$. Denote $E\tilde{\sigma}_i^l(h) \equiv \int \tilde{\sigma}_i^l(h, z_i)\, d\psi(z_i)$. I omit the straightforward argument that $E\tilde{\sigma}_1^l(I_{j-1}^{j-1}) \geq E\tilde{\sigma}_1^l(I_{K-1})$, i.e. in expectation the firm plays C at least as frequently (it is essentially a simpler version of the argument for the consumer immediately below). For the consumer at I_{j-1}^{j-1}, they have both belief $\mu(I_{j-1}^{j-1}) \geq \mu^0$ and are facing a firm playing C with probability $E\tilde{\sigma}_1^l(I_{j-1}^{j-1}) \geq E\tilde{\sigma}_1^l(I_{K-1})$. Let $\tilde{U}_2^l(a_2|h, z_2)$ be the consumer's expected payoff for playing a_2 at history h given shock z_2.Then

$$\begin{aligned}
\Delta\tilde{U}_2^l(h, z_2) &\equiv \tilde{U}_2^l(c|h, z_2) - \tilde{U}_2^l(d|h, z_2) \\
&= [[\mu(h) \cdot 3 + (1-\mu(h))(3E\tilde{\sigma}_1^l(h, z_1) + 0 \cdot \tilde{\sigma}_1^l(h, z_1)) + \varepsilon^l z_2^c] \\
&\qquad -[\mu(h) \cdot 2 + (1-\mu(h))(2E\tilde{\sigma}_1^l(h, z_1) + (1 - E\tilde{\sigma}_1^l(h, z_1)) + \varepsilon^l z_2^d]] \\
&= \mu(h) + (1-\mu(h))(2E\tilde{\sigma}_1^l(h, z_1) - 1) + \varepsilon^l \Delta z_2
\end{aligned}$$

where $\Delta z_2 \equiv z_2^c - z_2^d$. Note that

$$\begin{aligned}
\Delta\tilde{U}_2^l(I_{j-1}^{j-1}, z_2) &= \mu(I_{j-1}^{j-1}) + (1-\mu(I_{j-1}^{j-1}))(2E\tilde{\sigma}_1^l(I_{j-1}^{j-1}) - 1) + \varepsilon^l \Delta z_2 \\
&\geq \mu^0 + (1-\mu^0)(2E\tilde{\sigma}_1^l(I_{K-1}) - 1) + \varepsilon^l \Delta z_2]
\end{aligned}$$

and

$$\Delta\tilde{U}_2^l(I_{K-1}, z_2) = 2E\tilde{\sigma}_1^k(I_{K-1}) - 1 + \varepsilon^l \Delta z_2$$

for all z_2. Let $\mathfrak{z}_2^l(h)$ be the value of Δz_2 such that $\Delta\tilde{U}_2^l(h, z_2) = 0$. Then

$$0 \geq \mu^0 + (1-\mu^0)(2E\tilde{\sigma}_1^l(I_{K-1}) - 1) + \varepsilon^l \mathfrak{z}_2^l(I_{j-1}^{j-1}) \implies \mathfrak{z}_2^l(I_{j-1}^{j-1}) \leq -\frac{1}{\varepsilon^l}[\mu^0 + (1-\mu^0)(2E\tilde{\sigma}_1^l(I_{K-1}) - 1)] \tag{51}$$

$$0 = 2E\tilde{\sigma}_1^k(I_{K-1}) - 1 + \varepsilon^l \mathfrak{z}_2^l(I_{K-1}) \implies \mathfrak{z}_2^l(I_{K-1}) = -\frac{1}{\varepsilon^l}[2E\tilde{\sigma}_1^l(I_{K-1}) - 1] \tag{52}$$

Purifiability requires $\lim_l E\tilde{\sigma}_2^l(I_{K-1}) = \sigma_2(I_{K-1}) = 1 \implies \lim_l \mathfrak{z}_2^l(I_{K-1}) \leq -1$, which implies

$\lim_l E\tilde{\sigma}_1^l(I_{K-1}) \geq \frac{1}{2}(1+\varepsilon^l)$. Substituting this bound into (51) shows there exists l' such that for all $l > l'$, $\mathfrak{z}_2^l(I_{j-1}^{j-1}) < -1 \implies E\tilde{\sigma}_2^l(I_{j-1}^{j-1}) = 1$. Then

$$\begin{aligned}
& V_{\tilde{\sigma}^l}(I_{j-1}^{j-1}) \\
= \ & 2(1-\delta)E\tilde{\sigma}_2^l(I_{j-1}^{j-1}) + \int \max\{(1-\delta)\varepsilon z_1^C + \delta V_{\tilde{\sigma}^l}(I_j^j), (1-\delta)(1+\varepsilon z_1^D) + \delta V_{\tilde{\sigma}^l}(I_0^j)\}\, d\psi(z_1) \\
\geq \ & 2(1-\delta)E\tilde{\sigma}_2^l(I_K) + \int \max\{(1-\delta)\varepsilon z_1^C + \delta V_{\tilde{\sigma}^l}(I_K), (1-\delta)(1+\varepsilon z_1^D) + \delta V_{\tilde{\sigma}^l}(I_0)\}\, d\psi(z_1) \\
= \ & V_{\tilde{\sigma}^l}(I_K),
\end{aligned}$$

for large l since it must be that $1 = E\tilde{\sigma}_2^l(I_{j-1}^{j-1}) \geq E\tilde{\sigma}_2^l(I_K)$. Also, $\sigma_1(I_{j-1}^{j-1}) = \lim_l E\tilde{\sigma}_1^l(I_j^j) \geq \lim_l E\tilde{\sigma}_1^l(I_{K-1}) = \sigma_1(I_{K-1}) > 0$, it must be that $\lim_l V_{\tilde{\sigma}^l}(I_{j-1}^{j-1}) = 2(1-\delta) + \delta \lim_l V_{\tilde{\sigma}^l}(I_j^j) = 2$. Thus, Claim C.1 is proven.

Defining $I_k^K \equiv I_k$ for all k, we can see that the hypothesis of Claim C.1 is true for $j = K$, and by induction it is also true for all $j < K$, including $j = 0$. Thus, $V(\sigma) = \lim_l V_{\tilde{\sigma}^l}(I_0^0) = 2$.

C.5 Proof of Proposition 5.3

Define $\bar{q}(\xi_{t+1}, \mu^0)$ so that $\xi_{t+1} = \frac{1}{2}(\bar{q}(\xi_{t+1}, \mu^0) - \phi^*)$, which is solved to get $\bar{q}(\xi_{t+1}, \mu^0) = 2\xi_{t+1} + \phi^*$. Define $\bar{L}(\mu^0)$ as the maximum integer k such that $\bar{q}^k(0, \mu^0) \leq 1$. For brevity I generally omit the argument "μ^0," e.g. "$\bar{q}(\xi)$." Define $\bar{r}(\xi_t) \equiv \frac{1}{2}(\xi_t - \phi^*)$ as the inverse of $\bar{q}(\cdot)$. Recall the abuse of notation $\mathsf{V}^j(v_1, v_2) \equiv (1-\delta^j)v_1 + \delta^j v_2$, where $v_1, v_2 \in \mathbb{R}$.

The following lemma shows that $\underline{v}^*(\delta, \mu^0) - \beta = \mathsf{V}^{\bar{L}}(1, 2-\eta) - \beta$ is a lower bound on the set of purifiable equilibrium payoffs when memory is long.

Lemma C.12. *Let any $\beta > 0$ be given. There exists K^* such that for all $K > K^*$, every purifiable wPBE σ of the full game satisfies $V(\sigma) \geq V^{\bar{L}}(1, 2-\eta) - \beta$.*

Proof. Let $\mathcal{J}_\sigma \equiv \{t \in \{0, ..., K-1\} : V_\sigma(I_t^t) < \mathsf{V}(2, V_\sigma(I_{t+1}^{t+1}))\}$ denote the set of "low" initial periods. I show that the size of this set $|\mathcal{J}_\sigma| \leq \bar{L}$. For any $j \in \mathcal{J}_\sigma$, a necessary condition is that $\tilde{\sigma}_2(I_j^j) < 1$, and so d must be a best response for the consumer. Hence

$$3[\mu(I_j^j) + (1-\mu(I_j^j))\tilde{\sigma}_1(I_j^j)] \leq 2[\mu(I_j^j) + (1-\mu(I_j^j))\tilde{\sigma}_1(I_j^j)] + (1-\mu(I_j^j))(1-\tilde{\sigma}_1(I_j^j))$$

which can be solved for $\sigma_1(I_j^j) \leq \frac{1-2\mu(I_j^j)}{2(1-\mu(I_j^j))} = \frac{1-2\mu^0/[\mu^0+(1-\mu^0)P_{\tilde{\sigma}}(I_j^j|\theta_0)]}{2(1-\mu^0/[\mu^0+(1-\mu^0)P_{\tilde{\sigma}}(I_j^j|\theta_0)])} = \frac{1}{2}\left(1 - \frac{\phi^*}{P_\sigma(I_j^j|\theta_0)}\right)$. Since $P_\sigma(I_{j+1}^{j+1}|\theta_0) = P_\sigma(I_j^j|\theta_0)\sigma_1(I_j^j)$, we have

$$P_\sigma(I_{j+1}^{j+1}|\theta_0) \leq P_\sigma(I_j^j|\theta_0) \cdot \frac{1}{2}\left(1 - \frac{\phi^*}{P_\sigma(I_j^j|\theta_0)}\right) = \frac{1}{2}\left(P_\sigma(I_j^j|\theta_0) - \phi^*\right) = \bar{r}(P_\sigma(I_j^j|\theta_0)).$$

Suppose by contradiction that $L' \equiv |\mathcal{J}_\sigma| > \bar{L}$, denoting $\{j_1, j_2, ..., j_{L'}\} = \mathcal{J}_\sigma$ such that $j_1 < \cdots < j_{L'}$ without loss of generality. Then there would exist $t \in \{0, ..., K\}$ such that $P_\sigma(I_t^t|\theta_0) \leq$

$\bar{r}(P_\sigma(I_{L'}^{L'}|\theta_0)) \leq \cdots \leq \bar{r}^{L'}(P_\sigma(I_0^0|\theta_0))$. Yet by the definition of $\bar{L}$, $\bar{r}^{L'}(P_\sigma(I_0^0|\theta_0)) \leq \bar{r}^{L'}(1) < 0$, a contradiction.

Let $\tilde{L} \equiv |\mathcal{J}_\sigma| \leq \bar{L}$, and denote $\{j_1, j_2, ..., j_{\tilde{L}}\} = \mathcal{J}_\sigma$ such that $j_1 < \cdots < j_{\tilde{L}}$ without loss of generality. By definition of $\mathcal{J}_\sigma$ and $\tilde{L}$, it must be that $V_\sigma(I_{j_{\tilde{L}}+1}^{j_{\tilde{L}}+1}) \geq \mathsf{V}^{K-(j_{\tilde{L}}+1)}(2, V_\sigma(K, I_K))$ where $V_\sigma(K, I_K)$ is the payoff at the clean history at period K. For each $l < \tilde{L}$, $V_\sigma(I_{j_l+1}^{j_l+1}) \geq \mathsf{V}^{j_{l+1}-j_l}(2, V_\sigma(K, I_K))$.

For $j \in \mathcal{J}_\sigma$ such that $j \geq 1$, the fact that $V_\sigma(I_j^j) < \mathsf{V}(2, V_\sigma(I_{j+1}^{j+1}))$ requires that that D is a best response for the firm and that $P_\sigma(I_j^j|\theta_0) > 0$ (otherwise, the consumer's belief would be 1). These facts require that $V_\sigma(I_j^j) = (1-\delta)(1 + 2\sigma_2(I_j^j)) + \delta V_\sigma(I_0^{j+1})$ and that $V_\sigma(I_j^j) \geq V_\sigma(I_0^0) + \eta$. Note that $V_\sigma(I_0^j) \geq (1-\delta) + \delta V_\sigma(I_0^{j+1})$. I can then write

$$V_\sigma(I_j^j) = (1-\delta)(1 + 2\sigma_2(I_J^j)) + \delta V_\sigma(I_0^{j+1}) \geq V_\sigma(I_0^j) + \eta \geq (1-\delta) + \delta V_\sigma(I_0^{j+1}) + \eta$$

which can be solved to find that $\sigma_2(I_j^j) \geq \frac{1}{2\delta}$. Since $V_\sigma(I_j^j)$ is bounded from below by the payoff conditional on playing C, I can write

$$V_\sigma(I_j^j) \geq 2\sigma_2(I_j^j)(1-\delta) + \delta V_\sigma(I_{j+1}^{j+1}) \geq 2 \cdot \frac{1}{2\delta}(1-\delta) + \delta V_\sigma(I_{j+1}^{j+1}) = \mathsf{V}\left(\frac{1}{\delta}, V_\sigma(I_{j+1}^{j+1})\right).$$

Note that for $j = 0$, we can simply bound $V_\sigma(\varnothing) \geq \delta V_\sigma(I_1^1)$.

The above arguments imply that $V_\sigma(\varnothing) \geq \delta \mathsf{V}^{\tilde{L}-1}\left(\frac{1}{\delta}, \mathsf{V}^{K-\tilde{L}}(2, V_\sigma(K, I_K))\right)$. For any $\beta > 0$, there exists K^* such that for $K \geq K^*$, $V_\sigma(\varnothing) \geq \delta \mathsf{V}^{\tilde{L}-1}\left(\frac{1}{\delta}, 2\right) - \beta$ which can be rearranged to get $V_\sigma(\varnothing) \geq (1 - \delta^{\tilde{L}}) + \left(2 - \frac{1-\delta}{\delta}\right)\delta^{\tilde{L}} - \beta$. Since $\tilde{L} \leq \bar{L}$ the lemma is proven. ■

I now give the natural adaptation of the machinery of Section 3 to the (ψ, ε)-perturbed game. HBPs carry over directly, but HBAs must be slightly modified into objects I call "pHBAs" to account for the payoff shocks.

Definition C.1. A *perturbed action profile mapping* (pAM) $\alpha : Y^K \times \Theta \times Z \to \Delta A_2 \times (\Delta A_1)^{A_2}$ is a mapping from a public history h, type θ, and shocks z to mixed actions for each player $\alpha(\cdot|h, \theta, z) \equiv (\alpha_2(\cdot|h, z_2), \alpha_1(\cdot|ha_2, \theta, z_1))$, where $\alpha(a|h, \theta, z)$ denotes the probability of pure action profile a being played given h and θ. The set of all AMs is denoted $\tilde{\mathcal{A}}$.

Definition C.2. A history distribution, belief mapping, and perturbed action profile mapping object (pHBA) (ϕ, μ, α) is a triplet containing an HB (ϕ, μ) combined with an pAM α. The set of all pHBAs is denoted $\tilde{\mathcal{X}} \equiv \mathcal{M} \times \tilde{\mathcal{A}}$.

Definition C.3. The set of new HBs that are *ψ-inducible* by the pHBA $x \equiv (\phi, \mu, \alpha)$ is denoted by the correspondence $\Upsilon_\psi : \tilde{\mathcal{X}} \rightrightarrows \mathcal{M}$. $\Upsilon_\psi(x)$ is the set of HBs $(\tilde{\phi}, \tilde{\mu})$ such that for each $\theta \in \Theta$,

$$\tilde{\phi}(h'y|\theta) = \sum_{h' \in \tau(h)} \sum_{a_2 \in A_2} \sum_{y \in Y} \phi(h|\theta) \int \alpha_2(a_2|h, z_2)\alpha_1(a_1|ha_2, \theta, z_1)\rho(y|a_2, a_1)\, d\psi(z) \tag{53}$$

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

I can then construct perturbed version of enforceability.

Definition C.4. For any $W \subset \mathcal{W}$, a pHBA $x \equiv (\phi, \mu, \alpha) \in \tilde{\mathcal{X}}$ is (ψ, ε)-*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: $(\tilde{\phi}, \tilde{\mu}) \in \Upsilon_\psi(x)$ (see Definition C.3).

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

$$\begin{aligned} & \sum_{\theta\in\Theta} \int \sum_{a_2 \in A_2} \tilde{u}_2(\alpha_1(ha_2', \theta, z_1), \alpha_2(h, z_2), z_2)\alpha_2(a_2|h, z_2)\mu(\theta|h)\, d\psi(z_1) \\ \geq \; & \sum_{\theta\in\Theta} \int \tilde{u}_2(\alpha_1(ha_2', \theta, z_1), a_2', z_2)\mu(\theta|h)\, d\psi(z_1) \end{aligned}$$

   for all $z_2 \in Z_2$, $a_2' \in A_2$.

3. **Long-run player incentive compatibility:** player 1 has no profitable deviations: for all $h \in Y^K$, $a_2 \in A_2$, $a_1' \in A_1$, $z_1 \in Z_1$

$$\begin{aligned} \mathsf{V}_{a_2,z_1}(x, \tilde{w})(h) & \equiv (1-\delta)\tilde{u}_1(a_2, \alpha_1(ha_2, \theta_0), z_1) + \delta \sum_{a_1 \in A} \sum_{y \in Y} \alpha_1(a_1|ha_2, \theta_0, z_1)\rho(y|a_1, a_2)\tilde{\gamma}(hy) \\ & \geq (1-\delta)\tilde{u}_1(a_1', a_2, z_1) + \delta \sum_{y\in Y} \rho(y|a_2, a_1)\tilde{\gamma}(hy) \end{aligned}$$

   where $\mathsf{V}_{a_2,z_1}(x, \tilde{w}) \in \Gamma$ is a payoff function.

Given pHBA $x \equiv (\phi, \mu, \alpha)$ enforced by $\tilde{w}$, abuse notation by letting $\mathsf{V}(x, \tilde{w}) \in \Gamma$ be the expected player 1 payoff upon arriving at history h: $\mathsf{V}(x, \tilde{w})(h) = \int \sum_{a_2 \in A_2} \alpha_2(a_2|h, z_2)\mathsf{V}_{a_2,z_1}(x, \tilde{w})(h)\, d\psi(z)$. From this point, the analogous definitions of (ψ, ε)-*decompose* and (ψ, ε)-*self-generating* are given by inserting "(ψ, ε)-" in the appropriate places and are omitted. Let $\mathcal{B}_{\psi,\varepsilon}(W)$ denote the set of HBPs that are (ψ, ε)-decomposed by a set of HBPs W. Define the (ψ, ε)-perturbed (ϕ, μ)-variant game as a straightforward adaptation of the construction of the (ψ, ε)-perturbed game in Section 4 to the (unperturbed) (ϕ, μ)-variant game defined in Definition 3.6. Similarly, for a given payoff function γ, define the (ψ, ε)-perturbed γ-antegame as the adaptation of the (unperturbed) γ-antegame in Definition 3.15. Define $\Sigma_\gamma^{(\psi,\varepsilon)*}$ as the set of wPBEs of the (ψ, ε)-perturbed γ-antegame. For any HBP (ϕ, μ, γ), define the set of "matching" γ-antegame equilibria

$$\Sigma_{(\phi,\mu,\gamma)}^{(\psi,\varepsilon)*} \equiv \{\tilde{\sigma} \in \Sigma_\gamma^{(\psi,\varepsilon)*} : \forall h \in Y^K, \forall\theta \in \Theta, P_{\tilde{\sigma}}(h|\theta) = \phi(h|\theta)\}$$

and $\mathcal{E}_{(\phi,\mu,\gamma)}^{\psi,\varepsilon} \equiv \{V(\tilde{\sigma}) : \tilde{\sigma} \in \Sigma_{(\phi,\mu,\gamma)}^{(\psi,\varepsilon)*}\}$, allowing the "initialization operator" analogue of Definition 3.16.

Definition C.5. Let a set of HBPs W be given. Define $\mathcal{V}_{\psi,\varepsilon}(W) \equiv \{v \in \mathcal{E}_w^{\psi,\varepsilon} : w \in W\}$.

Define $\mathcal{E}^{\psi,\varepsilon}$ as the set of wPBE payoffs of the (ψ,ε)-perturbed game.

Proposition C.1. *Let any set of* (ψ,ε)-self-generating *HBPs W be given (that is, $W \subset \mathcal{B}_{\psi,\varepsilon}(W)$). Then $\mathcal{V}_{\psi,\varepsilon}(W) \subset \mathcal{E}^{\psi,\varepsilon}$.*

The proof is a straightforward adaptation of the proofs of Propositions 3.1 and 3.3.

Note. From this point forward, definitions pertain specifically to the product choice game (or its perturbed versions).

I introduce some useful notation and constructions for the lemmas below. Let any pHBA $x \equiv (\phi,\mu,\alpha)$ enforced by some HBP $\tilde{w}$ be given. Recall from Proposition 4.1 that wPBEs of the perturbed game are quasi-Markov perfect. The same logic shows that for almost all shocks $z \in Z$ and $h, h' \in I_k$ for all k, $\alpha_2(h,z_2) = \alpha_2(h',z_2)$ and $\alpha_1(ha_2, ha_2, z_1) = \alpha_1(ha_2, h'a_2', z_1)$ for all $a_2, a_2' \in \{c,d\}$. For brevity, I abbreviate values of payoff functions γ using a subscript k as follows: $\gamma_k \equiv \gamma(I_k) = \gamma(h)$ for all $h \in I_k$.

Let a sequence $(\psi^l, \varepsilon^l)_l$ such that $\psi^l \in \Delta^*(Z)$, $\varepsilon^l \to 0$, and $(\varepsilon^l)_l$ is strictly decreasing be given. I abbreviate $\mathcal{B}_l(\cdot) \equiv \mathcal{B}_{\psi^l,\varepsilon^l}(\cdot)$. The *Chebyshev distance* between two payoff functions γ, γ' is $\|\gamma - \gamma'\|_\infty = \max_k\{|\gamma_k - \gamma_k'|\}$.

Definition C.6. Define $\tilde{\Gamma}$ as the set of payoff functions γ satisfying

$$\gamma_K > \gamma_{K-1} > \cdots > \gamma_1 > \gamma_0 > \gamma_K - 2(1-\delta) + 2\varepsilon \tag{54}$$

and define $\tilde{W} \equiv \{(\phi,\gamma) : \gamma \in \tilde{\Gamma}\}$.

I construct several functions that capture the incentive compatibility requirements in the (ψ,ε)-perturbed game. Define $\hat{\zeta}_1 : \mathbb{R} \to \mathbb{R}$ as follows. Suppose the firm is at some date-history t, I_k for $k < K$, and choosing C reaches I_{k+1} next period with continuation payoff v_{k+1}, while choosing D reaches I_0 with continuation payoff v_0. Then $\hat{\zeta}_1(v_{k+1} - v_0)$ gives the shock difference $\Delta z_1 = z_1^C - z_1^D$ at which the firm would be indifferent between C and D (note that if $|v_{k+1} - v_0| > \varepsilon$, then $|\hat{\zeta}_1(v_{k+1} - v_0)| > 1$ meaning a strict preference for all shocks z_1). More formally, $\hat{\zeta}_1(v_{k+1} - v_0)$ satisfies

$$(1-\delta)\varepsilon\hat{\zeta}_1(v_{k+1} - v_0) + \delta v_{k+1} = (1-\delta) + \delta v_0$$

$$\hat{\zeta}_1(v_{k+1} - v_0) \equiv z_1^C - z_1^D = \frac{1}{\varepsilon}\frac{1}{1-\delta}\left[(1-\delta) - \delta(v_{k+1} - v_0)\right] = -\frac{1}{\varepsilon}\frac{\delta}{1-\delta}(v_{k+1} - v_0 - \eta)).$$

Define $\hat{\pi}_1(\zeta_1)$ as the probability that the plays C given indifference at shock difference ζ_1: $\hat{\pi}_1(\zeta_1) = \int_0^{\zeta_1} \mathbf{1}\{\Delta z_1 > \zeta_1\}\, d\psi(z_1)$ where $\mathbf{1}\{\Delta z_1 > \zeta_1\}$ is an indicator function equal to 1 if $\Delta z_1 > \zeta_1$. (Note that $\hat{\pi}_1(\zeta_1) = 1$ for $\zeta_1 < -1$ and $\hat{\pi}_1(\zeta_1) = 0$ for $\zeta_1 > 1$.)

Define $\hat{\zeta}_2 : \mathbb{R} \to \mathbb{R}$ similarly for the consumer. Suppose the consumer is at some history I_k for $k < K$, and the firm will subsequently be indifferent between C and D when the shock difference Δz_1 is equal to some threshold ζ_1. $\hat{\zeta}_2(\zeta_1)$ gives the shock difference $\Delta z_2 = z_1^c - z_1^d$ at which player 2 is indifferent given player 1's indifference at shock difference ζ_1 (given belief 0 since $k < K$). That is, $\hat{\zeta}_2(\zeta_1)$ satisfies

$$3\hat{\pi}_1(\zeta_1) + \varepsilon\hat{\zeta}_2(\zeta_1) = 2\hat{\pi}_1(\zeta_1) + (1 - \hat{\pi}_1(\zeta_1)) \qquad \Longrightarrow \qquad \hat{\zeta}_2(\zeta_1) \equiv z_2^c - z_2^d = \frac{1}{\varepsilon}(1 - 2\hat{\pi}_1(\zeta_1)).$$

Also define $\hat{\pi}_2(\zeta_2) = \int_0^{\zeta_2} \mathbf{1}\{\Delta z_2 > \zeta_2\}\, d\psi(z_2)$. For convenience define $\hat{\pi}_1^*(\Delta v) \equiv \hat{\pi}_1(\hat{\zeta}_1(\Delta v)), \hat{\pi}_2^*(\Delta v) \equiv \hat{\pi}_2(\hat{\zeta}_2(\hat{\zeta}_1(\Delta v)))$.

Define $\hat{g} : \mathbb{R}^2 \to \mathbb{R}$ as follows. Suppose a history t, I_k for $k < K$ is reached, and the continuation payoffs for the firm are $v_{k+1} \equiv V(t+1, I_{k+1}), v_0 \equiv V(t+1, I_0)$. Define $\hat{g}(v_{k+1}, v_0)$ to give what must be the continuation payoff $V(t, I_k)$:

$$\begin{aligned} \hat{g}(v_{k+1}, v_0) &= 2(1-\delta)\hat{\pi}_2(\hat{\zeta}_2(\hat{\zeta}_1(v_{k+1} - v_0))) \\ &\quad + \int \max\{(1-\delta)\varepsilon z_1^C + \delta v_{k+1}, (1-\delta)(1 + \varepsilon z_1^D) + \delta v_0\}\, d\psi(z_1). \end{aligned} \tag{55}$$

The t, I_K case is handled separately here. Note that given continuation payoffs $v_K \equiv V(t + 1, I_K), v_0 \equiv V(t+1, I_0)$, the firm is still indifferent at shock difference $\hat{\zeta}_1(v_K - v_0)$. Define $\bar{\zeta}_2(\zeta_1, \xi)$ as the shock difference Δz_2 at which the consumer is indifferent given that the firm is indifferent at shock difference ζ_1, and where ξ is the probability that the normal type arrives at I_K at period t. Denote the belief at (t, I_K) as $\bar{\mu}(\xi) \equiv \frac{\mu^0}{\mu^0 + (1-\mu^0)\xi}$. The consumer is indifferent if

$$3\left[\bar{\mu}(\xi) + (1 - \bar{\mu}(\xi))\hat{\pi}_1(\zeta_1)\right] + \varepsilon z_2^c = 2\left[\bar{\mu}(\xi) + (1 - \bar{\mu}(\xi))\hat{\pi}_1(\zeta_1)\right] + (1 - \bar{\mu}(\xi))(1 - \hat{\pi}_1(\zeta_1)) + \varepsilon z_2^d,$$

so then

$$\begin{aligned} \bar{\zeta}_2(\zeta_1, \xi) &\equiv z_2^c - z_2^d = \frac{1}{\varepsilon}\left((1 - \bar{\mu}(\xi))(1 - \hat{\pi}_1(\zeta_1)) - [\bar{\mu}(\xi) + (1 - \bar{\mu}(\xi))\hat{\pi}_1(\zeta_1)]\right) \\ &= \frac{1}{\varepsilon}\left(1 - 2[\bar{\mu}(\xi) + (1 - \bar{\mu}(\xi))\hat{\pi}_1(\zeta_1)]\right). \end{aligned} \tag{56}$$

Similar to above, define

$$\begin{aligned} \bar{g}(v_K, v_0, \xi) &= 2(1-\delta)\hat{\pi}_2(\bar{\zeta}_2(\hat{\zeta}_1(v_K, v_0), \xi)) \\ &\quad + \int \max\{(1-\delta)\varepsilon z_1^C + \delta v_K, (1-\delta)(1 + \varepsilon z_1^D) + \delta v_0\}\, d\psi(z_1). \end{aligned} \tag{57}$$

Let $B_\varepsilon(v)$ denote the open ball of radius ε centered at v. Abbreviate $\bar{z}_i^{a_i} \equiv \int z_i^{a_i}\, d\psi(z_i)$ as the expected value of the shock to action a_i.

The following result shows that, roughly speaking, a sequence of bounded (ψ^l, ε^l)-self-generating subsets of $\tilde{W}$ that converge in payoffs, with corresponding perturbed antegame equilibria that also converge in outcomes, yield a purifiable equilibrium payoff in the limit.

Lemma C.13. *Let a sequence $(W^l, w^l, \tilde{\sigma}^l)_l$ be given where for each l, $W^l \subset \tilde{W}$ is a bounded set of (ψ^l, ε^l)-self-generating set of HBPs; $w^l \equiv (\phi^l, \gamma^l) \in W^l$; and $\tilde{\sigma}^l \in \Sigma_{(\phi^l,\gamma^l)}^{(\psi^l,\varepsilon^l)*}$ is a wPBE of the (ψ^l, ε^l)-perturbed γ^l-antegame consistent with ϕ^l. Suppose there exists l^* such that for all $l > l^*$:*

1. *For each pair of HBPs $\hat{w}^l \equiv (\hat{\phi}^l, \hat{\gamma}^l), \tilde{w}^l \equiv (\tilde{\phi}^l, \tilde{\gamma}^l) \in W^l$ such that $\hat{w}^l \in \mathcal{B}_l(\tilde{w}^l)$, there exists a sequence $(\hat{w}^{l'}, \tilde{w}^{l'})_{l'=l+1}^{\infty}$ such that:*

   (a) *for each l', $\hat{w}^{l'} \in \mathcal{B}_{l'}(\tilde{w}^{l'})$, and*

   (b) *for some finite $\kappa > 0$, $\|\hat{\gamma}^l - \hat{\gamma}^{l'}\|_\infty < \kappa\varepsilon^l$ and $\|\tilde{\gamma}^l - \tilde{\gamma}^{l'}\|_\infty < \kappa\varepsilon^l$.*

2. *For all $(\phi, \gamma) \in W^l$, $\gamma_K \in B_\varepsilon(2)$.*

3. *The antegame equilibria $(\tilde{\sigma}^l)_l$ converge in outcomes to some limit: for all $t \in \{0, ..., K-1\}, k, i$, $\left(\int \tilde{\sigma}_i^l(I_k^t, z_i)\, d\psi(z_i)\right)_l$ is a convergent sequence.*

Then there exists a purifiable wPBE σ of the unperturbed game G^∞ such that $V(\sigma) = \lim_{l\to\infty} V(\tilde{\sigma}^l)$.

Proof. For each l, Proposition C.1 shows that the existence of $(W^l, w^l, \tilde{\sigma}^l)$ — i.e., a bounded self-generating set W^l containing w^l and a wPBE $\tilde{\sigma}^l$ of the (ψ^l, ε^l)-perturbed γ^l-antegame — implies the existence of a wPBE $\bar{\sigma}^l$ of the (ψ^l, ε^l)-perturbed full game such that $V(\bar{\sigma}^l) = V(\tilde{\sigma}^l)$. With straightforward modifications to the proof of Proposition 3.3, the proof of Proposition C.1 constructs $\bar{\sigma}^l$ so that for all $t < K, k, i$, $\bar{\sigma}_i^l(t, I_k, z_i) = \tilde{\sigma}_i^l(t, I_k, z_i)$, and also constructs a sequence of HBPs $(w^{l,t} \equiv (\phi^{l,t}, \gamma^{l,t}))_{t=K}^\infty$ in W^l such that $w^{l,K} \equiv w^l$ and for all $t \in \{K, K+1, ...\}$, $w^{l,t} \in \mathcal{B}_l(w^{l,t+1})$ and $V_{\bar{\sigma}^l}(t, I_k) = \gamma_k^{l,t}$. Since $\int \bar{\sigma}_i^l(I_k^t, z_i)\, d\psi(z_i) = \int \tilde{\sigma}_i^l(I_k^t, z_i)\, d\psi(z_i)$ for all $t < K$, the convergence of the sequence $\left(\int \tilde{\sigma}_i^l(I_k^t, z_i)\, d\psi(z_i)\right)_l$ implies convergence in outcomes of $(\bar{\sigma}^l)_l$ for initial periods $t < K$.

The rest of the proof shows convergence in outcomes for $t \geq K$. The statement of the proposition implies that we can choose $(w^{l,t})_{t=K}^\infty$ so that for each $t \in \{K, K+1, ...\}$, $\|\gamma^{l,t} - \gamma^{l',t}\|_\infty < \kappa\varepsilon^l$ for all $l' > l$. Thus, the sequence $(\gamma^{l,t})_l$ is a Cauchy sequence with some limit γ^{*t}. Furthermore, because $\gamma_K^l \in B_\varepsilon(2)$ for all l, $\gamma_K^{*t} = 2$.

We can see that this implies convergence in outcomes to a wPBE of the unperturbed game as follows. Let $k < K$ and $t \geq K$ be given. If $\gamma_{k+1}^{*t+1} > \gamma_0^{*t+1} + \eta$, then for l large enough, it must be that $\bar{\sigma}_1^l(t, I_k, z_1) = C$ for all z_1 (since $\varepsilon^l \to 0$). Similarly, if $\gamma_{k+1}^{*t+1} < \gamma_0^{*t+1} + \eta$, then for large l, it must be that $\bar{\sigma}_1^l(t, I_k, z_1) = D$ for all z_1. For large l, the consumer's best response in each case is clearly $\bar{\sigma}_2^l(t, I_k, z_2) = c$ and $\bar{\sigma}_2^l(t, I_k, z_2) = d$, respectively.

For the rest of the proof consider the case where $\gamma_{k+1}^{*t+1} = \gamma_0^{*t+1} + \eta$. The fact that $\gamma^{l,t} \in \tilde{\Gamma}$ means

$$\gamma_K^{l,t} > \gamma_{K-1}^{l,t} > \cdots > \gamma_1^{l,t} > \gamma_0^{l,t} > \gamma_K^{l,t} - 2(1-\delta) + 2\varepsilon^l. \tag{58}$$

Let pHBA $x^{l,t} \equiv (\phi^{l,t}, \mu^{l,t}, \alpha^{l,t})$ be given such that $w^{l,t}$ is (ψ^l, ε^l)-decomposed by $x^{l,t}, w^{l,t+1}$ and $\alpha_i^{l,t}(I_k, z_i) = \bar{\sigma}^l(t, I_k, z_i)$ (this is possible due to Definition C.4).

Denote $E\alpha_i^{l,t}(I_k) \equiv \int \alpha_i^{l,t}(I_k, z_i)\, d\psi(z_i)$. I show that $E\alpha_1^{l,t}(I_0) < E\alpha_1^{l,t}(I_1) < \cdots < E\alpha_1^{l,t}(I_{K-1})$. Suppose not, so there exists $k \in \{0, ..., K-2\}$ such that $E\alpha_1^{l,t}(I_k) = \hat{\pi}_1^*(\gamma_{k+1}^{l,t+1} - \gamma_0^{l,t+1}) \geq E\alpha_1^{l,t}(I_{k+1}) = \hat{\pi}_1^*(\gamma_{k+2}^{l,t+1} - \gamma_0^{l,t+1})$. Note that $\hat{\pi}_i^*(\cdot)$ is weakly increasing. Since $\gamma_{k+1}^{l,t+1} - \gamma_0^{l,t+1} < \gamma_{k+2}^{l,t+1} - \gamma_0^{l,t+1}$ it must be that $E\alpha_1^{l,t}(I_k) = E\alpha_1^{l,t}(I_{k+1}) \in \{0, 1\}$. If $E\alpha_1^{l,t}(I_k) = E\alpha_1^{l,t}(I_{k+1}) = 1$, the fact that $\gamma^{l,t+1}$ also satisfies (58) means $E\alpha_1^{l,t}(I_{k+2}) = \cdots = E\alpha_1^{l,t}(I_{K-1}) = E\alpha_1^{l,t}(I_K) = 1$, which implies

$$\gamma_{K-1}^{l,t} = \hat{g}(\gamma_{K-1}^{l,t+1}, \gamma_0^{l,t+1}) = (1-\delta)(2 + \varepsilon^l \bar{z}_1^C) + \delta\gamma_K^{l,t+1} = \bar{g}(\gamma_K^{l,t+1}, \gamma_0^{l,t+1}) = \gamma_K^{l,t},$$

a contradiction of (58). If $E\alpha_1^{l,t}(I_k) = E\alpha_1^{l,t}(I_{k+1}) = 0$, similar reasoning as above implies that $E\alpha_1^{l,t}(I_{k-1}) = \cdots = E\alpha_1^{l,t}(I_0) = 1$, which yields $\gamma_k^{l,t} = (1-\delta)(1+\varepsilon^l \bar{z}_1^D) + \delta\gamma_0^{l,t+1} = \gamma_{k-1}^{l,t} = \cdots = \gamma_0^{l,t}$, a contradiction of (58).

Because $\alpha_1^{l,t}(I_{K-1}, z_1) = \alpha_1^{l,t}(I_K, z_1)$ for almost all z_1, the fact that $\gamma_K^{l,t} > \gamma_{K-1}^{l,t}$ implies that $E\alpha_2^{l,t}(I_{K-1}) < E\alpha_2^{l,t}(I_K)$. The fact that $E\alpha_1^{l,t}(I_{K-1}) > 0$ for all l implies that $\gamma_K^{*t+1} \geq \gamma_0^{*t+1} + \eta$. Furthermore, it cannot be that $\gamma_K^{*t+1} > \gamma_0^{*t+1} + \eta$; otherwise for large l we would have $E\alpha_1^{l,t}(I_{K-1}) = E\alpha_1^{l,t}(I_K) = 1$ which implies $E\alpha_2^{l,t}(I_{K-1}) = E\alpha_2^{l,t}(I_K) = 1$ and so $\gamma_{K-1}^{l,t} = \gamma_K^{l,t}$, a contradiction.

I show that $\lim_l (E\alpha_2^{l,t}(I_K))_l = 1$. Suppose not; then for some $\beta > 0$, for any l' there exists $l > l'$ such that $E\alpha_2^{l,t}(I_K) \leq 1 - \beta$. Then

$$\gamma_K^{l,t} \leq 2(1-\delta)E\alpha_2^{t,l}(I_{K-1}) + \delta\gamma_K^{l,t+1} + \varepsilon^l \leq 2(1-\delta)(1-\beta) + \delta\gamma_K^{l,t+1} + \varepsilon^l. \tag{59}$$

Recall that $\gamma_K^{*t} = \lim_l \gamma_K^{l,t} = \lim_l \gamma_K^{l,t+1} = 2$, so (59) is a contradiction.

I also show $E\alpha_2^{l,t}(I_0) > 0$. Suppose not: $E\alpha_2^{l,t}(I_0) = 0$. Recall that $\gamma_K^{l,t} > 2 - \varepsilon$ and $\gamma_K^{l,t+1} > 2 - \varepsilon$, so $\gamma_K^{l,t} > (2-\varepsilon)(1-\delta) + \delta\gamma_K^{l,t+1} > 2(1-\delta) + \delta\gamma_K^{l,t+1} - \varepsilon$. Then $\gamma_0^{l,t} \leq 2(1-\delta)\cdot 0 + \delta\gamma_K^{l,t+1} + \varepsilon^l < \gamma_K^{l,t} - 2(1-\delta) + 2\varepsilon^l$, a contradiction of (58).

For large l, the fact that $0 < E\alpha_2^{l,t}(I_0), E\alpha_2^{l,t}(I_{K-1}) < 1$, imply that $0 < E\alpha_1^{l,t}(I_0), E\alpha_1^{l,t}(I_{K-1}) < 1$; since $E\alpha_1^{l,t}(I_0) < E\alpha_1^{l,t}(I_1) < \cdots < E\alpha_1^{l,t}(I_{K-1})$ and $E\alpha_2^{l,t}(I_k) = \hat{\pi}_2^*(\Delta v)$ is strictly increasing for values $\Delta v \in \{\Delta v' \in \mathbb{R} : 0 < \hat{\pi}_2^*(\Delta v') < 1\}$, I have $0 < E\alpha_2^{l,t}(I_0) < E\alpha_2^{l,t}(I_1) < \cdots < E\alpha_2^{l,t}(I_{K-1}) < 1$. Note that $\gamma_k^{l,t} \in B_{\varepsilon^l}(2E\alpha_2^{l,t}(I_k)(1-\delta) + \delta\gamma_{k+1}^{l,t+1})$. The convergence of $(\gamma_k^{l,t}, \gamma_{k+1}^{l,t+1}, \varepsilon^l)_l \to (\gamma_k^{*t}, \gamma_{k+1}^{*t+1}, 0)$ clearly implies that $(E\alpha_2^{l,t}(I_k))_l$ converges. Furthermore, for $k < K$ each $E\alpha_2^{l,t}(I_k) \in (0,1)$ implies a unique $E\alpha_1^{l,t}(I_k)$ (reasoning from the definition of $\hat{\pi}_2(\cdot)$), and so the convergence of $E\alpha_2^{l,t}(I_k)$ implies the convergence of $E\alpha_1^{l,t}(I_k)$.

Since $\int \bar{\sigma}_i^l(t, I_k, z_i)\, d\psi(z_i) = E\alpha_i^{l,t}(I_k)$ for all $t \geq K, k, i$, the above proves that $(\bar{\sigma}^l)_l$ converges in outcomes for all date-histories in periods $t \geq K$. Define the strategy profile $\bar{\sigma}^*$ of the (unperturbed) full game so that $\bar{\sigma}_i^*(t, I_k) \equiv \lim_l \int \bar{\sigma}_i^l(t, I_k, z_i)\, d\psi(z_i)$. It is then straightforward to check that $\bar{\sigma}^*$ constitutes a wPBE of the unperturbed game. ■

The rest of the proof establishes the existence of a purifiable wPBE σ such that $V(\sigma) = \mathsf{V}^{\bar{L}}(1, 2-\eta)$, constructing a sequence satisfying the assumptions of Lemma C.13. Let $\mu^0 \in (0, \frac{1}{2})$ be given such that $\bar{q}^{\bar{L}}(0, \mu^0) < 1$, which is true for almost all $\mu^0 \in (0, \frac{1}{2})$. The following lemma establishes that for any HBP in $\tilde{W}$, it is possible to construct an HBP $w' \in \tilde{W}$ that (ψ, ε)-generates it.

Lemma C.14. *For small enough ε, there exists a continuous function $f : \tilde{\Gamma} \to \tilde{\Gamma}$ such that for any HBP $w \equiv (\phi, \gamma) \in \tilde{W}$, there exists an HBP $w' \equiv (\phi', f(\gamma)) \in \tilde{W}$ such that $w \in \mathcal{B}_{\psi,\varepsilon}(w')$. Furthermore, $f_0(\gamma) + \eta(1-\varepsilon) < f_1(\gamma) < \cdots < f_K(\gamma) < f_0(\gamma) + \eta(1+\varepsilon)$.*

Proof. The proof is by construction. Note that $\hat{\zeta}_i(\cdot), \hat{\pi}_i(\cdot), \hat{\pi}_i^*(\cdot), \bar{\zeta}_2(\cdot)$ are continuous, and therefore

$\hat{g}(\cdot), \bar{g}(\cdot)$ are as well. Moving $v_0' \in \mathbb{R}$ through values "near" (within ε of) $v_{k+1}' - \eta$ lets $\hat{g}(v_{k+1}', v_0')$ span a range given by

$$\hat{g}(v_{k+1}', v_0') = \begin{cases} (1-\delta)(2+\varepsilon\bar{z}_1^C) + \delta(v_0'+\eta) & v_0' = v_{k+1}' - \eta - \varepsilon\eta \\ (1-\delta)\varepsilon\bar{z}_1^D + \delta(v_0'+\eta) & v_0' = v_{k+1}' - \eta + \varepsilon\eta. \end{cases}$$

Furthermore, if $(1-\delta)\varepsilon + \delta(v_0'+\eta) < \hat{g}(v_{k+1}', v_0') < (1-\delta)(2+\varepsilon\bar{z}_1^C) + \delta(v_0'+\eta)$, then (55) (the definition of $\hat{g}(\cdot)$) implies that $0 < \hat{\pi}_2^*(v_{k+1}' - v_0') \leq 1$, which yields

$$-1 \leq \hat{\zeta}_2(\hat{\zeta}_1(v_{k+1}' - v_0')) = \frac{1}{\varepsilon}(1 - 2\hat{\pi}_1^*(v_{k+1}' - v_0')) < 1 \implies \tfrac{1}{2}(1-\varepsilon) < \hat{\pi}_1^*(v_{k+1}' - v_0') \leq \tfrac{1}{2}(1+\varepsilon).$$

Therefore, for any v_K' satisfying

$$(1-\delta)\varepsilon + \delta v_K' < \gamma_{K-1} < (1-\delta)(2+\varepsilon\bar{z}_1^C) + \delta v_K' \tag{60}$$

there exists some $v_0' \in [v_K' - \eta - \varepsilon\eta, v_K' - \eta + \varepsilon\eta]$ such that $\hat{g}(v_K', v_0') = \gamma_{K-1}$ and

$$\tfrac{1}{2}(1-\varepsilon) < \hat{\pi}_1^*(v_K' - v_0') \leq \tfrac{1}{2}(1+\varepsilon). \tag{61}$$

Solving (60) for v_K' yields

$$\frac{1}{\delta}\left[\gamma_{K-1} - (1-\delta)(2+\varepsilon\bar{z}_1^C)\right] < v_K' < \frac{1}{\delta}\left[\gamma_{K-1} - (1-\delta)\varepsilon\right].$$

Define interval $\mathfrak{V}(v_{K-1}) \equiv \left(\frac{1}{\delta}\left[v_{K-1} - (1-\delta)(2+\varepsilon\bar{z}_1^C)\right], \frac{1}{\delta}\left[v_{K-1} - (1-\delta)\varepsilon\right]\right)$. Then I can say for any $v_K' \in \mathfrak{V}(\gamma_{K-1})$, there exists some v_0' such that $\hat{g}(v_K', v_0') = \gamma_{K-1}$ and (61) is true. Define the set $\mathfrak{S} \equiv \{(v_K', v_{K-1}) \in \mathbb{R}^2 : v_K' \in \mathfrak{V}(v_{K-1})\}$. The continuity of $\hat{g}(\cdot)$ implies the existence of continuous function $s : \mathfrak{S} \to \mathbb{R}$ such that $\gamma_{K-1} = \hat{g}(v_K', s(v_K'|\gamma_{K-1}))$, $s(v_K'|\gamma_{K-1}) \leq v_K' - \eta + \eta\varepsilon$ and[26]

$$\lim_{v_K' \to \inf \mathfrak{V}(\gamma_{K-1})} s(v_K'|\gamma_{K-1}) = v_K' - \eta - \varepsilon\eta. \tag{62}$$

From (56) (the definition of $\bar{\zeta}_2(\cdot)$) we have $\bar{\zeta}_2(\hat{\zeta}_1(v_K' - v_0'), \xi) = \frac{1}{\varepsilon}(1 - 2[\bar{\mu}(\xi) + (1-\bar{\mu}(\xi))\hat{\pi}_1^*(v_K' - v_0')])$, and substituting (61) yields

$$\begin{aligned} \bar{\zeta}_2(\hat{\zeta}_1(v_K' - s(v_K'|\gamma_{K-1})), \xi) &< \frac{1}{\varepsilon}(1 - 2[\bar{\mu}(\xi) + \tfrac{1}{2}(1-\bar{\mu}(\xi))(1-\varepsilon)]) \\ &= -\frac{1}{\varepsilon}\bar{\mu}(\xi) + (1-\bar{\mu}(\xi)) \leq 1 - \left(1 + \frac{1}{\varepsilon}\right)\mu^0 \end{aligned}$$

[26] (62) follows from the fact that if $v_K' \equiv \inf \mathfrak{V}(\gamma_{K-1}) = \frac{1}{\delta}\left[\gamma_{K-1} - (1-\delta)(2+\varepsilon\bar{z}_1^C)\right]$ and $v_0' \equiv v_K' - \eta - \varepsilon\eta$, then

$$\hat{\zeta}_1(v_K' - v_0') = -\frac{1}{\varepsilon}\eta(v_K' - [v_K' - \eta - \varepsilon\eta] - \eta) = -\frac{1}{\varepsilon}\eta(-\varepsilon\eta) = 1,$$

and so $\hat{\pi}_1^*(v_K' - v_0') = 1$ which implies $\hat{g}(v_K', v_0') = (1-\delta)(2+\varepsilon\bar{z}_1^C) + \delta \cdot \frac{1}{\delta}\left[\gamma_{K-1} - (1-\delta)(2+\varepsilon\bar{z}_1^C)\right] = \gamma_{K-1}$.

for all $v'_K \in \mathfrak{V}(\gamma_{K-1})$. Then for ε small enough, $\bar{\zeta}_2(\hat{\zeta}_1(v'_K - s(v'_K|\gamma_{K-1})), \xi) < -1$ (meaning the consumer strictly prefers c for all shocks), and from (57) we have

$$\bar{g}(v'_K, s(v'_K|\gamma_{K-1})), \phi_K) = 2(1-\delta) + E_{\psi(z_1)}[\max\{(1-\delta)\varepsilon z_1^C + \delta v'_K, (1-\delta)(1+\varepsilon z_1^D) + \delta s(v'_K|\gamma_{K-1})\}]. \tag{63}$$

Thus $\bar{g}(v'_K, s(v'_K|\gamma_{K-1}), \phi_K)$ is independent of ϕ_K.

From (55) (the definition of $\hat{g}(\cdot)$) we have

$$\begin{aligned}\gamma_{K-1} \quad = \quad & 2(1-\delta)\hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1})) \\ & + E_{\psi(z_1)}[\max\{(1-\delta)\varepsilon z_1^C + \delta v'_K, (1-\delta)(1+\varepsilon z_1^D) + \delta s(v'_K|\gamma_{K-1})\}].\end{aligned}$$

(62) implies that

$$\lim_{v'_K \to \inf \mathfrak{V}(\gamma_{K-1})} \hat{\pi}_1^*(v'_K - s(v'_K|\gamma_{K-1})) = 1 \qquad \Longrightarrow \qquad \lim_{v'_K \to \inf \mathfrak{V}(\gamma_{K-1})} \hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1})) = 1.$$

Note that

$$\bar{g}(v'_K, s(v'_K|\gamma_{K-1})), \phi_K) - \hat{g}(v'_K, s(v'_K|\gamma_{K-1})) = 2(1-\delta)[1 - \hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1}))], \tag{64}$$

and so

$$\lim_{v'_K \to \inf \mathfrak{V}(\gamma_{K-1})} \bar{g}(v'_K, s(v'_K|\gamma_{K-1})), \phi_K) = \gamma_{K-1} < \gamma_K. \tag{65}$$

Let any $\beta > 0$ be given. Suppose that $v'_K > \sup \mathfrak{V}(\gamma_{K-1}) - \beta$. Then

$$\begin{aligned}\gamma_{K-1} \quad & = \quad 2(1-\delta)\hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1})) \\ & \quad\quad + E_{\psi(z_1)}[\max\{(1-\delta)\varepsilon z_1^C + \delta v'_K, (1-\delta)(1+\varepsilon z_1^D) + \delta s(v'_K|\gamma_{K-1})\}] \\ & > \quad 2(1-\delta)\hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1})) + \delta\left(\frac{1}{\delta}\left[\gamma_{K-1} - (1-\delta)\varepsilon\right] - \beta\right) \\ & = \quad 2(1-\delta)\hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1})) + \gamma_{K-1} - \beta\delta - (1-\delta)\varepsilon\end{aligned}$$

which yields $\hat{\pi}_2^*(v'_K - s(v'_K|\gamma_{K-1})) < \frac{\beta\delta}{2(1-\delta)} + \frac{1}{2}\varepsilon$. From (64) this yields

$$\begin{aligned}\bar{g}(v'_K, s(v'_K|\gamma_{K-1})), \phi_K) - \hat{g}(v'_K, s(v'_K|\gamma_{K-1})) \quad & > \quad 2(1-\delta)\left[1 - \frac{\beta\delta}{2(1-\delta)} - \frac{1}{2}\varepsilon\right] \\ & = \quad (2-\varepsilon)(1-\delta) - \beta\delta.\end{aligned}$$

Thus, for $\beta < \frac{\varepsilon}{\delta}$ we have

$$\begin{aligned}\bar{g}(v'_K, s(v'_K|\gamma_{K-1})), \phi_K) \quad & > \quad \hat{g}(v'_K, s(v'_K|\gamma_{K-1})) + (2-\varepsilon)(1-\delta) - \beta\delta \\ & > \quad \hat{g}(v'_K, s(v'_K|\gamma_{K-1})) + (2-\varepsilon)(1-\delta) - \varepsilon \\ & > \quad \gamma_{K-1} + 2(1-\delta) - 2\varepsilon > \gamma_K\end{aligned} \tag{66}$$

where the last step is due to the fact that $\gamma \in \tilde{\Gamma}$ (see Definition C.6). The continuity of $\bar{g}(\cdot)$ together with (65) and (66) prove that there exists some $v'_K \in \mathfrak{V}(\gamma_{K-1})$ such that $\bar{g}(v'_K, s(v'_K|\gamma_{K-1})) = \gamma_K$.

Thus, there exist continuous functions $f_K : \tilde{\Gamma} \to \mathbb{R}$, $f_0 : \tilde{\Gamma} \to \mathbb{R}$ such that

$$\gamma_K = \bar{g}(f_K(\gamma), s(f_K(\gamma)|\gamma_{K-1}), \phi_K), \qquad f_0(\gamma) = s(f_K(\gamma)|\gamma_{K-1}).$$

From (63), we have $\gamma_K = 2(1-\delta) + E_{\psi(z_1)}[\max\{(1-\delta)\varepsilon z_1^C + \delta v'_K, (1-\delta)(1+\varepsilon z_1^D) + \delta s(v'_K|\gamma_{K-1})\}] > 3(1-\delta) + \delta f_0(\gamma)$. Then $\delta f_0(\gamma) < \gamma_K - 3(1-\delta)$, and so

$$\begin{aligned} \hat{g}(f_0(\gamma), f_0(\gamma)) &= (1-\delta)(1+\varepsilon\bar{z}_1^D) + \delta f_0(\gamma) < (1-\delta)(1+\varepsilon\bar{z}_1^D) + \gamma_K - 3(1-\delta) && (67) \\ &= \gamma_K - 2(1-\delta) + \varepsilon\bar{z}_1^D < \gamma_0. && (68) \end{aligned}$$

There exists some $v'_{K-1} \in (f_0(\gamma), f_K(\gamma))$ such that $\hat{g}(v'_{K-1}, f_0(\gamma)) = \gamma_{K-2}$ because $\hat{g}(f_0(\gamma), f_0(\gamma)) < \gamma_0 < \gamma_{K-2} < \gamma_{K-1} = \hat{g}(f_K(\gamma), f_0(\gamma))$. Similarly, for each $k \in \{1, ..., K-2\}$, there exists v'_k such that $\hat{g}(v'_k, f_0(\gamma)) = \gamma_{k-1}$. Thus, for each $k \in \{1, ..., K-1\}$ there exists continuous function $f_k : \tilde{\Gamma} \to \mathbb{R}$ satisfying $\hat{g}(f_k(\gamma), f_0(\gamma)) = \gamma_{k-1}$.

Define function $f : \tilde{\Gamma} \to \tilde{\Gamma}$ so that for $f(\gamma) = \gamma'$, $\gamma'_k = f_k(\gamma)$ for each $k \in \{0, ..., K\}$. To see that $f(\gamma) \in \tilde{\Gamma}$, note that

$$\gamma_0 = \hat{g}(f_1(\gamma), f_0(\gamma)) < \gamma_1 = \hat{g}(f_2(\gamma), f_0(\gamma)) < \cdots < \gamma_{K-1} = \hat{g}(f_K(\gamma), f_0(\gamma)),$$

and since $\hat{g}(v'_{k+1}, v'_0)$ is strictly increasing in v'_{k+1}, it must be that $f_1(\gamma) < f_2(\gamma) < \cdots < f_K(\gamma)$. Furthermore, for all $v'_{k+1} < f_0(\gamma) + \eta(1-\varepsilon)$, we can see from (67) - (68) that $\hat{g}(v'_{k+1}, f_0(\gamma)) = (1-\delta)(1+\varepsilon\bar{z}_1^D) + \delta f_0(\gamma) < \gamma_0$, and so it must be that $f_0(\gamma) + \eta(1-\varepsilon) < f_1(\gamma) < \cdots < f_K(\gamma)$.

The definition of $s(f_K(\gamma)|\gamma_{K-1})$ requires (61) which implies

$$\eta - \frac{1-\delta}{\delta}\varepsilon = \eta(1-\varepsilon) < f_K(\gamma) - f_0(\gamma) < \eta + \frac{1-\delta}{\delta}\varepsilon = \eta(1+\varepsilon) \qquad (69)$$

, and $f_1(\cdot), ..., f_{K-1}(\cdot)$ are defined so that $f_0(\gamma) < f_1(\gamma) < \cdots < f_{K-1}(\gamma)$. Since $\delta > \frac{1}{2}$ I have $\eta < 2(1-\delta)$, and thus for small enough ε, $f_0(\gamma) > f_K(\gamma) - (1+\varepsilon)\eta > f_K(\gamma) - 2(1-\delta) + 2\varepsilon$. Since $\gamma'_k \equiv f_k(\gamma)$ for all k, it has been shown that $\gamma' \in \tilde{\Gamma}$.

I complete the construction of HBP $w' \equiv (\phi', \gamma')$. Let pHBA $x \equiv (\phi, \mu, \alpha)$ so that for all $k < K$,

$$\alpha_1(I_k, z_1) = \begin{cases} C & \Delta z_1 \geq \hat{\zeta}_1(\gamma'_{k+1} - \gamma'_0) \\ D & \Delta z_1 < \hat{\zeta}_1(\gamma'_{k+1} - \gamma'_0) \end{cases} \qquad \alpha_2(I_k, z_2) = \begin{cases} c & \Delta z_2 \geq \hat{\zeta}_2(\hat{\zeta}_1(\gamma'_{k+1} - \gamma'_0)) \\ d & \Delta z_2 < \hat{\zeta}_2(\hat{\zeta}_1(\gamma'_{k+1} - \gamma'_0)) \end{cases}$$

$$\alpha_1(I_K, z_1) = \begin{cases} C & \Delta z_1 \geq \hat{\zeta}_1(\gamma'_K - \gamma'_0) \\ D & \Delta z_1 < \hat{\zeta}_1(\gamma'_K - \gamma'_0) \end{cases} \qquad \alpha_2(I_K, z_2) = \begin{cases} c & \Delta z_2 \geq \bar{\zeta}_2(\hat{\zeta}_1(\gamma'_K - \gamma'_0), \phi_K) \\ d & \Delta z_2 < \bar{\zeta}_2(\hat{\zeta}_1(\gamma'_K - \gamma'_0), \phi_K). \end{cases}$$

Then define $\phi'_0 \equiv \sum_k \left(1 - \int \alpha_1(I_k, z_1)\, d\psi(z_1)\right) \phi_k$, $\phi'_K \equiv \sum_{k=K-1}^{K} \phi_k \int \alpha_1(I_k, z_1)\, d\psi(z_1)$, and $\phi'_k \equiv \phi_{k-1} \int \alpha_1(I_{k-1}, z_1)\, d\psi(z_1)$ for all $k \in \{1, ..., K-1\}$. The constructions of $\hat{\zeta}_i(\cdot), \bar{\zeta}_2(\cdot), \hat{g}(\cdot), \bar{g}(\cdot)$ ensure that x is (ψ, ε)-enforced by w', and w is (ψ, ε)-decomposed by x, w'. ∎

The next lemma characterizes $f(\gamma)$ when γ_K is particularly large and small, i.e. outside $B_\varepsilon(2)$.

Lemma C.15. *Let $\bar{v} \equiv 2 + \varepsilon$ and $\underline{v} \equiv 2 - \varepsilon$. The following hold:*

1. *If $\gamma \in \tilde{\Gamma}$ and $\gamma_K \geq \bar{v}$, then $f_K(\gamma) > \gamma_K$.*
2. *If $\gamma \in \tilde{\Gamma}$ and $\gamma_K \leq \underline{v}$, then $f_K(\gamma) < \gamma_K$.*

Proof. This proof reuses the notation and constructions in the proof of Lemma C.14.

Proof of Part 1: Let $\gamma \in \tilde{\Gamma}$ where $\gamma_K \geq \bar{v}$ be given. Suppose by contradiction that $f_K(\gamma) \leq \gamma_K$. The proof of Lemma C.14 shows that

$$\gamma_K = 2(1-\delta) + E_{\psi(z_1)}[\max\{(1-\delta)\varepsilon z_1^C + \delta f_K(\gamma), (1-\delta)(1+\varepsilon z_1^D) + \delta f_0(\gamma)\}] \tag{70}$$

and due to (69) (i.e. because the firm is playing D with positive probability at I_K), we have $\gamma_K \leq 2(1-\delta) + (1-\delta)(1+\varepsilon) + \delta f_0(\gamma) = (1-\delta)(2+\varepsilon) + \delta(f_0(\gamma) + \eta)$ which can be solved for $f_0(\gamma) \geq \frac{1}{\delta}[\gamma_K - (1-\delta)(2+\varepsilon)] - \eta$. (69) implies that $f_0(\gamma) < f_K(\gamma) - \eta(1-\varepsilon)$, so $f_K(\gamma) - \eta(1-\varepsilon) \geq f_0(\gamma) \geq \frac{1}{\delta}[\gamma_K - (1-\delta)(2+\varepsilon)] - \eta$ which is solved to find $\gamma_K \leq (2+\varepsilon)(1-\delta) + \delta f_K(\gamma) \leq (1-\delta)\bar{v} + \delta\gamma_K$, which implies $\gamma_K < \bar{v}$, a contradiction.

Proof of Part 2: Let some $\gamma \in \tilde{\Gamma}$ where $\gamma_K \leq \underline{v}$ be given. Suppose by contradiction that $f_K(\gamma) \geq \gamma_K$. From (70) we have $\gamma_K \geq 2(1-\delta) + (1-\delta) + \delta f_0(\gamma) = 2(1-\delta) + \delta(f_0(\gamma) + \eta) \implies f_0(\gamma) \leq \frac{1}{\delta}\gamma_K - 3\eta$. From (69), $f_0(\gamma) \geq f_K(\gamma) - \eta(1+\varepsilon)$. Then $f_K(\gamma) - \eta(1+\varepsilon) \leq f_0(\gamma) \leq \frac{1}{\delta}\gamma_K - 3\eta$ which is solved to find $\gamma_K \geq (1-\delta)(2-\varepsilon) + \delta f_K(\gamma) \geq (1-\delta)\underline{v} + \delta\gamma_K$, which implies $\gamma_K > \underline{v}$, a contradiction. ∎

Inductively define $f^m(\gamma)$ so that $f^1(\gamma) \equiv f(\gamma)$ and $f^{m+1}(\gamma) \equiv f(f^m(\gamma))$. It will be convenient to separate a payoff function γ into two parts: the value at I_K (i.e. γ_K) and all other values (a K-length vector $\gamma_{-K} \equiv (\gamma_0, ..., \gamma_{K-1}) \in \mathbb{R}^K$). I use the notation $(\gamma_{-K}, \gamma_K) \equiv \gamma$ and $f^m(\gamma_{-K}, \gamma_K) \equiv f^m(\gamma)$.

The following lemma and its corollaries show it is possible to construct a self-generating set with an HBP $(\phi, \gamma \equiv (\gamma_{-K}, \gamma_K))$ with a particular γ_K and arbitrary γ_{-K} satisfying (71).

Lemma C.16. *Let any K-length vector $\gamma_{-K} \equiv (\gamma_0, ..., \gamma_{K-1}) \in \mathbb{R}^K$ such that*

$$2 - \varepsilon > \gamma_{K-1} > \cdots > \gamma_1 > \gamma_0 > 2 - 2(1-\delta) + 2\varepsilon = 2\delta + 2\varepsilon \tag{71}$$

be given. Let any $m \geq 1$ be given. Then there exist $\underline{v}^m, \bar{v}^m$ where $\underline{v} < \underline{v}^m < \bar{v}^m < \bar{v}$ such that $f_K^m(\gamma_{-K}, \bar{v}^m) = \bar{v}$, $f_K^m(\gamma_{-K}, \underline{v}^m) = \underline{v}$, and for all $v \in (\underline{v}^m, \bar{v}^m)$, $\underline{v} < f_K^m(\gamma_{-K}, v) < \bar{v}$. Furthermore, $\underline{v} < \underline{v}^1 < \cdots < \underline{v}^m < \bar{v}^m < \cdots < \bar{v}^1 < \bar{v}$ for all m.

Proof. The proof is by induction, first proving the $m = 1$ case. From (71), $(\gamma_{-K}, \underline{v})$ and $(\gamma_{-K}, \bar{v})$ are in the set $\tilde{\Gamma}$. Lemma C.15 shows that $f_K(\gamma_{-K}, \bar{v}) > \bar{v}$ and $f_K(\gamma_{-K}, \underline{v}) < \underline{v}$. Because $f(\cdot)$ is continuous, there exists some $\underline{v}^1, \bar{v}^1 \in (\underline{v}, \bar{v})$ where $\underline{v}^1 < \bar{v}^1$ such that $f_K(\gamma_{-K}, \bar{v}^1) = \underline{v}$, $f_K(\gamma_{-K}, \bar{v}^1) = \bar{v}$, and for all $v \in (\underline{v}^1, \bar{v}^1)$, $\underline{v} < f_K(\gamma_{-K}, v) < \bar{v}$.

Now suppose there exists some m such that for all $m' \in \{1, ..., m\}$, the lemma has been proven, i.e. there exist $\underline{v}^m, \bar{v}^m$ where $\underline{v} < \underline{v}^1 < \cdots < \underline{v}^m < \bar{v}^m < \cdots < \bar{v}^1 < \bar{v}$ such that $f_K^m(\gamma_{-K}, \underline{v}^m) = \underline{v}$, $f_K^m(\gamma_{-K}, \bar{v}^m) = \bar{v}$, and for all $v \in (\underline{v}^m, \bar{v}^m)$, $\underline{v} < f_K^m(\gamma_{-K}, v) < \bar{v}$. Lemma C.15 implies that $f_K^{m+1}(\gamma_{-K}, \underline{v}^m) < \underline{v}$ and $f_K^{m+1}(\gamma_{-K}, \bar{v}^m) > \bar{v}$. The continuity of $f(\cdot)$ implies that there exist $\underline{v}^{m+1}, \bar{v}^{m+1} \in (\underline{v}^m, \bar{v}^m)$ such that $f_K^{m+1}(\gamma_{-K}, \underline{v}^{m+1}) = \underline{v}$, $f_K^{m+1}(\gamma_{-K}, \bar{v}^{m+1}) = \bar{v}$, and for all $v \in (\underline{v}^{m+1}, \bar{v}^{m+1})$, $\underline{v} < f_K^{m+1}(\gamma_{-K}, v) < \bar{v}$. ■

Corollary C.2. *There exists v such that $\underline{v}^m < v < \bar{v}^m$ for all m.*

Proof. Lemma C.16 proves that $(\underline{v}^m)_m$ is an increasing sequence bounded from above by $\bar{v}$, so it has some limit $\underline{v}^* \equiv \lim_m \underline{v}^m$. Similar reasoning, plus the fact that $\underline{v}^m < \bar{v}^m$ for all m, shows that there exists limit $\bar{v}^* \equiv \lim_m \bar{v}^m \geq \underline{v}^*$. Then any $v \in [\underline{v}^*, \bar{v}^*]$ satisfies $\underline{v}^m < v < \bar{v}^m$ for all m. ■

Corollary C.3. *Let any $\gamma_{-K} \equiv (\gamma_0, ..., \gamma_{K-1}) \in \mathbb{R}^K$ such that it satisfies (71) be given. There exists payoff function $\gamma \equiv (\gamma_{-K}, \gamma_K)$ where $\gamma_K \in (2-\varepsilon, 2+\varepsilon)$ such that for any history distribution ϕ, there exists a bounded (ψ, ε)-self-generating set W^* containing (ϕ, γ). Furthermore, for every $(\phi', \gamma') \in W^*$, $\gamma_0' \in B_{\varepsilon(1+\eta)}(2-\eta)$; $\gamma_k' \in B_{2\varepsilon}(2-\varepsilon)$ for all $k \in \{1, ..., K-1\}$; and $\gamma_K' \in B_\varepsilon(2)$.*

Proof. From Corollary C.2, pick γ_K such that $2 - \varepsilon < \underline{v}^m < \gamma_K < \bar{v}^m < 2 + \varepsilon$ for all m, and let $\gamma \equiv (\gamma_{-K}, \gamma_K)$. Denote $w^0 \equiv (\phi^0, \gamma^0) \equiv (\phi, \gamma)$. Since γ^0 satisfies (54), $w^0 \in \tilde{W}$. Lemma C.14 states that there exists $w^1 \equiv (\phi^1, \gamma^1) \in \tilde{W}$ such that $w^0 \in \mathcal{B}_{\psi,\varepsilon}(w^1)$. Furthermore, $\underline{v}^1 < \gamma_K^0 < \bar{v}^1$ means $\underline{v} < \gamma_K^1 = f_K(\gamma^0) < \bar{v}$. Suppose that there exists some $w^m \equiv (\phi^m, \gamma^m) \in \tilde{W}$ where $w^0 \in \mathcal{B}_{\psi,\varepsilon}(w^1), w^1 \in \mathcal{B}_{\psi,\varepsilon}(w^2), ..., w^{m-1} \in \mathcal{B}_{\psi,\varepsilon}(w^m)$, and $\gamma^m = f^m(\gamma^0)$. Then Lemma C.14 states that there exists $w^{m+1} \equiv (\phi^{m+1}, \gamma^{m+1}) \in \tilde{W}$ such that $w^m \in \mathcal{B}_{\psi,\varepsilon}(w^{m+1})$ and $\gamma^{m+1} = f(\gamma^m) = f^{m+1}(\gamma)$; also, $\underline{v}^{m+1} < \gamma_K^0 < \bar{v}^{m+1}$ means $\underline{v} < \gamma_K^{m+1} < \bar{v}$. By induction, there exists set of HBPs $W^* \equiv \{w^0, w^1, w^2, ...\} \subset \tilde{W}$ such that for all m, $w^m \in \mathcal{B}_{\psi,\varepsilon}(w^{m+1})$; thus $W^* \subset \mathcal{B}_{\psi,\varepsilon}(W^*)$ and $(\phi, \gamma) \in W^*$. For all m, $\underline{v} < \gamma_K^m < \bar{v} \implies \gamma_K^m \in B_\varepsilon(2)$ and $\underline{v} - 2(1-\delta) + 2\varepsilon < \gamma_k^m < \bar{v}$ for all k, $W^*$is bounded.

For each w^m, Lemma C.14 shows that $\gamma_0^m + \eta(1-\varepsilon) < \gamma_1^m < \cdots < \gamma_K^m < \gamma_0^m + \eta(1+\varepsilon)$. Since $\gamma_K^m \in B_\varepsilon(2)$, $2 - \varepsilon < \gamma_K^m < \gamma_0^m + \eta(1+\varepsilon) \implies \gamma_0^m > 2 - \eta - (1+\eta)\varepsilon > 2 - \eta - 2\varepsilon$ and $\gamma_0^m + \eta(1-\varepsilon) < 2 + \varepsilon \implies \gamma_0^m < 2 - \eta + (1+\eta)\varepsilon < 2 - \eta + 2\varepsilon$. Then for each $k \in \{1, ..., K-1\}$, $\gamma_k^m > \gamma_0^m + \eta(1-\varepsilon) > 2 - \eta - 2\varepsilon + \eta(1-\varepsilon) > 2 - 3\varepsilon = (2-\varepsilon) - 2\varepsilon$, and $\gamma_k^m < \gamma_K^m < 2 + \varepsilon = (2-\varepsilon) + 2\varepsilon$. ■

The next lemma shows that constructing particular payoff functions γ can ensure a specific continuation value at histories I_0^L in any equilibrium $\tilde{\sigma}$ of the (ψ, ε)-perturbed γ-antegame.

Lemma C.17. *For small enough ε, there exists K-length vector $\gamma_{-K} \equiv (\gamma_0, ..., \gamma_{K-1}) \in \mathbb{R}^{\mathbb{K}}$ satisfying (71) such that for any $\gamma_K \in \mathbb{R}$ such that $\gamma \in \tilde{\Gamma}$ and any wPBE $\tilde{\sigma}$ of the (ψ, ε)-perturbed (γ_{-K}, γ_K)-antegame, $V_{\tilde{\sigma}}(I_0^{\bar{L}}) = v_{\bar{L}} \equiv 2 - \eta + 2\varepsilon$.*

Proof. Choose $\gamma_0 \equiv 2 - \eta - 2\varepsilon$. Inductively define $\hat{g}^1(\cdot|\gamma_0) \equiv \hat{g}(\cdot, \gamma_0)$, $\hat{g}^{m+1}(\cdot|\gamma_0) = \hat{g}(\hat{g}^m(\cdot|\gamma_0), \gamma_0)$.

Let any $v, v' \in (1+\varepsilon, 2-\varepsilon)$ such that $v < v'$ be given. I show there exists $\tilde{v}, \tilde{v}' \in (1+\varepsilon, 2-\varepsilon)$ such that $\hat{g}(\tilde{v}, \gamma_0) = v$, $\hat{g}(\tilde{v}', \gamma_0) = v'$, and $\tilde{v}' > \tilde{v}$. Recall from arguments in the proof of Lemma

C.14 that $\hat{g}(\gamma_0 + \eta - 2\varepsilon, \gamma_0) < (1-\delta)(1+\varepsilon) + \delta\gamma_0 < \gamma_0$ for small ε, and since $(2-\varepsilon) - \gamma_0 = (2-\varepsilon) - (2-\eta-2\varepsilon) = \eta + \varepsilon$, $\hat{g}(2-\varepsilon, \gamma_0) = (1-\delta)(2+\varepsilon\bar{z}_1^C) + \delta(2-\varepsilon) > \tilde{v}$. Since $\hat{g}(\cdot, \gamma_0)$ is continuous and weakly increasing there exists $\tilde{v} \in (\gamma_0 + \eta - 2\varepsilon, 2-\varepsilon)$ such that $\hat{g}(\tilde{v}, \gamma_0) = \tilde{v}$ and $\tilde{v}' \in (\tilde{v}, 2-\varepsilon)$ such that $\hat{g}(\tilde{v}', 2-\varepsilon)$.

The previous paragraph shows it is possible to choose $\gamma_1, ..., \gamma_{K-\bar{L}-1}$ such that $\hat{g}^m(\gamma_m|\gamma_0) = \gamma_0$ for each $m \in \{1, ..., K-\bar{L}-1\}$, and $\gamma_0 < \gamma_1 < \gamma_2 < \cdots < \gamma_{K-\bar{L}-1} < 2-\varepsilon$. Similarly, I can choose $\gamma_{K-\bar{L}} < 2-\varepsilon$ such that $\hat{g}^{K-\bar{L}}(\gamma_{K-\bar{L}}|\gamma_0) = v_{\bar{L}}$. (The proof of this lemma requires no restrictions on $\gamma_{K-\bar{L}+1}, ..., \gamma_{K-1}, \gamma_K$ except $\gamma \in \tilde{\Gamma}$.)

Let $\tilde{\sigma} \in \Sigma^*_{\gamma,(\psi,\varepsilon)}$ be given. Working backward, the definitions of $\hat{\zeta}_1(\cdot), \hat{\zeta}_2(\cdot)$ imply

$$\tilde{\sigma}_1(I_k^{K-1}, z_1) = \begin{cases} C & \Delta z_1 > \hat{\zeta}_1(\gamma_k - \gamma_0) \\ D & \Delta z_1 < \hat{\zeta}_1(\gamma_k - \gamma_0) \end{cases} \qquad \tilde{\sigma}_2(I_k^{K-1}, z_2) = \begin{cases} c & \Delta z_2 > \hat{\zeta}_2(\hat{\zeta}_1(\gamma_k - \gamma_0)) \\ d & \Delta z_1 < \hat{\zeta}_1(\gamma_k - \gamma_0) \end{cases} \tag{72}$$

for all $k < K$, and the definition of $\hat{g}(\cdot)$ implies $V_{\tilde{\sigma}}(I_k^{K-1}) = \hat{g}(\gamma_{k+1}, \gamma_0)$, and for all $t < K - 1$, $V_{\tilde{\sigma}}(I_k^t) = \hat{g}(V_{\tilde{\sigma}}(I_{k+1}^{t+1}), V_{\tilde{\sigma}}(I_0^{t+1}))$. Then $V_{\tilde{\sigma}}(I_0^{K-1}) = \hat{g}(\gamma_1, \gamma_0) = \gamma_0$. Similarly $V_{\tilde{\sigma}}(I_0^{K-2}) = \hat{g}(V_{\tilde{\sigma}}(I_1^{K-1}), \gamma_0) = \hat{g}(\hat{g}(\gamma_2, \gamma_0), \gamma_0) = \hat{g}^2(\gamma_2|\gamma_0) = \gamma_0$. Similarly, for all $t \in \{\bar{L}+1, ..., \bar{L}-1\}$, $V_{\tilde{\sigma}}(I_0^t) = \hat{g}^{K-t}(\gamma_{K-t}|\gamma_0) = \gamma_0$. A similar argument shows $V_{\tilde{\sigma}}(I_0^{\bar{L}}) = \hat{g}^{K-\bar{L}}(\gamma_{K-\bar{L}}|\gamma_0) = v_{\bar{L}}$. ■

With $\gamma \in \tilde{\Gamma}$ constructed as above, I can construct a wPBE of the (ψ, ε)-perturbed γ-antegame.

Lemma C.18. *Let payoff function $\gamma \in \tilde{\Gamma}$ where $(\gamma_0, ..., \gamma_{K-1})$ are chosen as specified by Lemma C.17 and $\gamma_K \in B_\varepsilon(2)$ be given. For small enough ε, there exists a wPBE $\tilde{\sigma} \in \Sigma^*_{\gamma,(\psi,\varepsilon)}$ of the (ψ, ε)-perturbed γ-antegame such that $V(\tilde{\sigma}) \in B_{2\varepsilon}(\mathsf{V}^{\bar{L}}(1, 2-\eta))$.*

Proof. (72) constructs strategies for $\tilde{\sigma}_i(I_k^{K-1}, z_i)$ for all i, $k < t$, and almost all shocks $z_i \in Z_i$, which are best responses by the construction of $\hat{\zeta}_i(\cdot)$. Since players are indifferent at the measure zero set of shocks omitted by (72), the actions at such shocks can be chosen arbitrarily. By construction of $\hat{g}(\cdot)$, $V_{\tilde{\sigma}}(I_k^{K-1}) = \hat{g}(\gamma_{k+1}, \gamma_0)$. For each $t \in \{\bar{L}+1, ..., K-1\}, k \in \{1, ..., t-1\}$, inductively define

$$\tilde{\sigma}_1(I_{k-1}^{t-1}, z_1) = \begin{cases} C & \Delta z_1 > \hat{\zeta}_1(V_{\tilde{\sigma}}(I_k^t) - \gamma_0) \\ D & \Delta z_1 < \hat{\zeta}_1(V_{\tilde{\sigma}}(I_k^t) - \gamma_0) \end{cases} \quad \tilde{\sigma}_2(I_{k-1}^{t-1}, z_2) = \begin{cases} c & \Delta z_2 > \hat{\zeta}_2(\hat{\zeta}_1(V_{\tilde{\sigma}}(I_k^t) - \gamma_0)) \\ d & \Delta z_1 < \hat{\zeta}_1(V_{\tilde{\sigma}}(I_k^t) - \gamma_0) \end{cases} \tag{73}$$

noting that $V_{\tilde{\sigma}}(I_{k-1}^{t-1}) = \hat{g}(V_{\tilde{\sigma}}(I_k^t), \gamma_0)$ since $V_{\tilde{\sigma}}(I_0^{\bar{L}}) = \cdots = V_{\tilde{\sigma}}(I_0^{\bar{L}}) = \gamma_0$ (as shown in the proof of Lemma (C.17)). Again, the constructions of $\hat{\zeta}_i(\cdot), \hat{g}(\cdot)$ ensure these are all best responses.

For all $t \in \{\bar{L}, ..., K-1\}$, set $\tilde{\sigma}_1(I_t^t, z_1) = C, \tilde{\sigma}_2(I_t^t, z_2) = c$. To see that these are best responses, note that since $\gamma_K > 2-\varepsilon$, we have $\gamma_K - \gamma_0 > 2 - \varepsilon - [2-\eta-2\varepsilon] = \eta + \varepsilon$, and so at I_{K-1}^{K-1} the firm weakly (strictly) prefers C for all (almost all) shocks $z_1 \in Z_1$. Since the firm always plays C at I_{K-1}^{K-1}, the consumer strictly prefers c. Thus, $V_{\tilde{\sigma}}(I_{K-1}^{K-1}) = (1-\delta)(2+\varepsilon\bar{z}_1^C) + \delta\gamma_K \in (2-\varepsilon, 2+\varepsilon)$. Since $V_{\tilde{\sigma}}(I_0^{K-1}) = \cdots = V_{\tilde{\sigma}}(I_0^{\bar{L}+1}) = \gamma_0$, a backward induction argument shows that $\tilde{\sigma}_i(I_t^t, z_i)$ is also a best response for each $t \in \{\bar{L}, ..., K-1\}$, yielding some $V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}) \in (2-\varepsilon, 2+\varepsilon)$.

Having constructed best responses for periods $t \geq \bar{L}$ above, the rest of the proof constructs strategies for $t < \bar{L}$. First, for all $t < \bar{L}$ and $k < t$, set $\tilde{\sigma}_2(I_k^t, z_2) = d$ and $\tilde{\sigma}_1(I_k^t, z_1) = D$ for all $z \in Z$. Note that since $V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}) - V_{\tilde{\sigma}}(I_0^{\bar{L}}) < 2 + \varepsilon - [2 - \eta + 2\varepsilon] = \eta - \varepsilon$, the firm strictly prefers D for almost all shocks $z_1 \in Z_1$, and so for all $k \in \{0, ..., K-1\}$, the consumer strictly prefers d. Define $v_{\bar{L}-1} \equiv (1-\delta)(1 + \varepsilon \bar{z}_1^D) + \delta v_{\bar{L}}$. Note that $V_{\tilde{\sigma}}(I_k^{\bar{L}-1}) = v_{\bar{L}-1}$ for all $k < \bar{L} - 1$, and so at all $I_0^{\bar{L}-2}, ..., I_{\bar{L}-3}^{\bar{L}-2}$ the firm strictly prefers D for all $z_1 \in Z_1$. A backward induction argument then shows that for all $t < \bar{L}$ and $k < t$, $\tilde{\sigma}_2(I_k^t, z_2) = d$ and $\tilde{\sigma}_1(I_k^t, z_1) = D$ for all $z \in Z$. Inductively define $v_{k-1} \equiv (1-\delta)(1 + \varepsilon \bar{z}_1^D) + \delta v_k$ for each $k \in \{1, ..., \bar{L}\}$. Then $V_{\tilde{\sigma}}(I_k^t) = v_t$ for all $t < \bar{L}$, $k < t$.

To construct the strategies at I_t^t for $t < t$, it will be easier to first characterize the equilibrium probabilities $P_{\tilde{\sigma}}(I_t^t | \theta_0)$ of the normal type reaching these histories, and subsequently back out the strategies that give rise to such probabilities. I use the variables $\xi_1, ..., \xi_{\bar{L}-1} \in [0,1]$ to represent $P_{\tilde{\sigma}}(I_1^1), ..., P_{\tilde{\sigma}}(I_{\bar{L}-1}^{\bar{L}-1})$, inductively construct a system of equations for $\xi_1, ..., \xi_{\bar{L}-1}$ whose solution corresponds to an equilibrium, and then show that a solution indeed exists. Define function $\bar{\mu}(\xi_t) \equiv \frac{\mu^0}{\mu^0 + (1-\mu^0)\xi_t}$ to translate the ξ_t variables into beliefs.

First, define function $\lambda_{\bar{L}-1}(\xi_{\bar{L}-1}) \equiv \bar{g}(V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}), v_{\bar{L}}, \xi_{\bar{L}-1})$ (this corresponds to the value function $V_{\tilde{\sigma}}(I_{\bar{L}-1}^{\bar{L}-1})$), noting that it is continuous and weakly decreasing. Note that because $V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}) - v_{\bar{L}} < \eta - \varepsilon$ (i.e. the firm strictly prefers D at $I_{\bar{L}-1}^{\bar{L}-1}$) $\lambda_{\bar{L}-1}(\xi_{\bar{L}-1})$, $\lambda_{\bar{L}-1}(\cdot)$ satisfies $\lambda_{\bar{L}-1}(0) = \bar{\lambda}_{\bar{L}-1} \equiv \mathsf{V}(3 + \varepsilon \bar{z}_1^D, v_{\bar{L}})$ (corresponding to belief 1 on the honest type) and $\lambda_{\bar{L}-1}(1) = \underline{\lambda}_{\bar{L}-1} \equiv \mathsf{V}(1 + \varepsilon \bar{z}_1^D, v_{\bar{L}})$ (corresponding to belief $\mu^0 < \frac{1}{2}$ on the honest type). Furthermore, given that the firm always chooses D, the consumer at $I_{\bar{L}-1}^{\bar{L}-1}$ with shock z_2 has a strict preference for d if $3\bar{\mu}(\xi_{\bar{L}-1}) + \varepsilon z_2^c < 2\bar{\mu}(\xi_{\bar{L}-1}) + (1 - \bar{\mu}(\xi_{\bar{L}-1})) + \varepsilon z_2^d$, which can be solved to find $\Delta z_2 < \bar{\zeta}_2(V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}), v_{\bar{L}}, \xi_{\bar{L}-1}) = \frac{1}{\varepsilon}[1 - 2\bar{\mu}(\xi_{\bar{L}-1})]$. Choosing $\bar{\xi}_{L-1}$ so that $1 = \bar{\zeta}_2(V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}), v_{\bar{L}}, \bar{\xi}_{\bar{L}-1}) = \frac{1}{\varepsilon}\left[1 - 2\frac{\mu^0}{\mu^0 + (1-\mu^0)\bar{\xi}_{\bar{L}-1}}\right]$ yields $\bar{\xi}_{\bar{L}-1} = \left[\frac{2}{1-\varepsilon} - 1\right]\phi^*$. Choosing $\underline{\xi}_{L-1}$ so that $-1 = \bar{\zeta}_2(V_{\tilde{\sigma}}(I_{\bar{L}}^{\bar{L}}), v_{\bar{L}}, \underline{\xi}_{\bar{L}-1})$ similarly yields $\underline{\xi}_{\bar{L}-1} = \left[\frac{2}{1+\varepsilon} - 1\right]\phi^*$. For small enough ε, $1 - 3\varepsilon < \frac{2}{1+\varepsilon} - 1 < \frac{2}{1-\varepsilon} - 1 < 1 + 3\varepsilon$.[27] Recalling that $\bar{q}(0) = \phi^*$,

$$\lambda_{\bar{L}-1}(\xi_{\bar{L}-1}) = \begin{cases} \mathsf{V}(3 + \varepsilon \bar{z}_1^D, v_{\bar{L}}) & \xi_{\bar{L}-1} \leq \bar{q}(0) - 3\varepsilon < (1-3\varepsilon)\bar{q}(0) \\ \mathsf{V}(1 + \varepsilon \bar{z}_1^D, v_{\bar{L}}) & \xi_{\bar{L}-1} \geq \bar{q}(0) + 3\varepsilon > (1+3\varepsilon)\bar{q}(0). \end{cases}$$

Define $\pi_{L-2}(\xi_{\bar{L}-1}) = \hat{\pi}_1^*(\lambda_{\bar{L}-1}(\xi_{\bar{L}-1}), v_{\bar{L}-1})$. Then $\pi_{\bar{L}-2}(\xi_{\bar{L}-1})$ is continuous, weakly decreasing and satisfies

$$\pi_{\bar{L}-2}(\xi_{\bar{L}-1}) = \begin{cases} 1 & \xi_{\bar{L}-1} \leq \bar{q}(0) - 3\varepsilon \\ 0 & \xi_{\bar{L}-1} \geq \bar{q}(0) + 3\varepsilon. \end{cases}$$

The next step is to impose the requirement that for any given $\xi_{\bar{L}-2}$, $\xi_{\bar{L}-1} = \xi_{\bar{L}-2}\pi_{\bar{L}-2}(\xi_{\bar{L}-1})$ (which corresponds to the fact that $P_{\tilde{\sigma}}(I_{\bar{L}-1}^{\bar{L}-1} | \theta_0) = P_{\tilde{\sigma}}(I_{\bar{L}-2}^{\bar{L}-2} | \theta_0) \int \tilde{\sigma}_1(I_{\bar{L}-2}^{\bar{L}-2}, z_1)\, d\psi(z_1)$). Rearranging gives

$$\xi_{\bar{L}-2} = \frac{\xi_{\bar{L}-1}}{\pi_{\bar{L}-2}(\xi_{\bar{L}-1})} \tag{74}$$

[27] Suppose by contradiction that $1 - 3\varepsilon \geq \frac{2}{1+\varepsilon} - 1$. Then $1 + \varepsilon - 3\varepsilon - 3\varepsilon^2 \geq 2 - (1+\varepsilon) = 1 - \varepsilon$, yielding $\varepsilon \leq -3\varepsilon^2$, which is clearly false for small ε. That $\frac{2}{1-\varepsilon} - 1 < 1 + 3\varepsilon$ is shown similarly.

for $\pi_{\bar{L}-2}(\xi_{\bar{L}-1}) > 0$. The right hand side is strictly increasing in $\xi_{\bar{L}-1}$, starting at 0 (for $\xi_{\bar{L}-1} = 0$) and reaching 1 for some $\xi_{\bar{L}-1} \in B_{3\varepsilon}(\bar{q}(0))$. Define function $\chi_{\bar{L}-1} : [0,1] \to [0,1]$ as the value of $\xi_{\bar{L}-1}$ that satisfies (74) as a function of $\xi_{\bar{L}-2}$, i.e. $\xi_{\bar{L}-2} = \frac{\chi_{\bar{L}-1}(\xi_{\bar{L}-2})}{\pi_{\bar{L}-2}(\chi_{\bar{L}-1}(\xi_{\bar{L}-2}))}$. Then $\chi_{\bar{L}-1}(\cdot)$ is also strictly increasing, ranging from $\chi_{\bar{L}-1}(0) = 0$ to $\chi_{\bar{L}-1}(1) \in B_{3\varepsilon}(\bar{q}(0))$. Define $\lambda_{\bar{L}-2}(\xi_{\bar{L}-2}) \equiv \bar{g}(\lambda_{\bar{L}-1}(\chi_{\bar{L}-1}(\xi_{\bar{L}-2})), v_{\bar{L}-1}, \xi_{\bar{L}-2})$.

Define $\tau_t \equiv 3(2^{\bar{L}-(t+1)} - 1)$. Suppose that for a given $t \in \{1, ..., \bar{L}-2\}$, $\lambda_{t+1}(\xi_{t+1})$ is a weakly decreasing function such that

1. If $\xi_{t+1} < \bar{q}^{\bar{L}-(t+1)}(0) - \tau_{t+1}\varepsilon$, then $\lambda_{t+1}(\xi_{t+1}) > \mathsf{V}(1+\varepsilon, v_{t+2}) + \eta + \varepsilon$.

2. If $\xi_{t+1} > \bar{q}^{\bar{L}-(t+1)}(0) + \tau_{t+1}\varepsilon$, then $\lambda_{t+1}(\xi_{t+1}) < \mathsf{V}(1+\varepsilon, v_{t+2})$.

3. $\pi_t(\xi_{t+1})$ is a function defined so that $\pi_t(\xi_{t+1}) = \hat{\pi}_1^*(\lambda_{t+1}(\xi_{t+1}), v_{t+1})$.

Then $\pi_t(\xi_{t+1})$ is a weakly decreasing function such that

$$\pi_t(\xi_{t+1}) = \begin{cases} 1 & \xi_{t+1} < \bar{q}^{\bar{L}-(t+1)}(0) - \tau_{t+1}\varepsilon \\ 0 & \xi_{t+1} > \bar{q}^{\bar{L}-(t+1)}(0) + \tau_{t+1}\varepsilon \end{cases} \tag{75}$$

Define the function $\chi_{t+1}^{-1} : [0,1] \to [0,1]$ as $\chi_{t+1}^{-1}(\xi_{t+1}) \equiv \frac{\xi_{t+1}}{\pi_t(\xi_{t+1})}$. Note that $\chi_{t+1}^{-1}(\xi_{t+1})$ is strictly increasing, ranging from $\chi_{t+1}^{-1}(0) = 0$ to $\chi_{t+1}^{-1}(\xi'_{t+1}) = 1$ at some value $\xi'_{t+1} \in B_{\tau_{t+1}\varepsilon}(\bar{q}^{\bar{L}-(t+1)}(0))$. Thus, $\chi_{t+1}^{-1}(\xi_{t+1})$ is invertible; let $\chi_{t+1}(\xi_t)$ be its inverse, so that

$$\xi_t = \frac{\chi_{t+1}(\xi_t)}{\pi_t(\chi_{t+1}(\xi_t))}. \tag{76}$$

Then $\chi_{t+1}(\xi_t)$ is strictly increasing, $\chi_{t+1}(0) = 0$ and $\chi_{t+1}(1) \in B_{\tau_{t+1}}(\bar{q}^{\bar{L}-(t+1)}(0))$. Further note that for $\xi_{t+1} < \bar{q}^{\bar{L}-(t+1)}(0) - \tau_{t+1}\varepsilon$, $\chi_{t+1}(\xi_t) = \xi_t$.

Given the above, define function $\lambda_t(\xi_t) \equiv \bar{g}(\lambda_{t+1}(\chi_{t+1}(\xi_t)), v_{t+1}, \xi_t)$. Let $\bar{\xi}_t$ be chosen so that $\bar{\zeta}_2(\lambda_{t+1}(\chi_{t+1}(\bar{\xi}_t)), v_{t+1}, \bar{\xi}_t) = 1$, corresponding to the consumer at I_t^t having a strict preference for d at almost all shocks $z_2 \in Z_2$ but being indifferent when $\Delta z_2 = 1$. From (56),

$$\varepsilon = 1 - 2\bar{\mu}(\bar{\xi}_t) - 2(1-\bar{\mu}(\bar{\xi}_t))\pi_t(\chi_{t+1}(\bar{\xi}_t)) = 1 - 2\frac{\mu^0}{\mu^0 + (1-\mu^0)\bar{\xi}_t} - 2\left(1 - \frac{\mu^0}{\mu^0 + (1-\mu^0)\bar{\xi}_t}\right)\pi_t(\chi_{t+1}(\bar{\xi}_t))$$

which rearranges to $0 = (1-\varepsilon)(1-\mu^0)\bar{\xi}_t - (1+\varepsilon)\mu^0 - 2(1-\mu^0)\bar{\xi}_t\pi_t(\chi_{t+1}(\bar{\xi}_t))$. Recall from (76) that $\xi_t\pi_t(\chi_{t+1}(\bar{\xi}_t)) = \chi_{t+1}(\bar{\xi}_t)$, so

$$0 = (1-\varepsilon)(1-\mu^0)\bar{\xi}_t - (1+\varepsilon)\mu^0 - 2(1-\mu^0)\chi_{t+1}(\bar{\xi}_t) \implies \bar{\xi}_t = \frac{1}{1-\varepsilon}[(1+\varepsilon)\phi^* + 2\chi_{t+1}(\bar{\xi}_t)].$$

Choose $\underline{\xi}_t$ so that

$$-1 = \bar{\zeta}_2(\lambda_{t+1}(\chi_{t+1}(\underline{\xi}_t)), v_{t+1}, \underline{\xi}_t) = \frac{1}{\varepsilon}[1 - 2\bar{\mu}(\underline{\xi}_t) - 2(1-\bar{\mu}(\underline{\xi}_t))\pi_t(\chi_{t+1}(\underline{\xi}_t))]. \tag{77}$$

Algebra similar to that above yields $\underline{\xi}_t = \frac{1}{1+\varepsilon}[(1-\varepsilon)\phi^* + 2\chi_{t+1}(\underline{\xi}_t)]$. Note that for small enough ε, (77) requires that $\pi_t(\chi_{t+1}(\underline{\xi}_t)) < 1$, and from (75) it must be that $\chi_{t+1}(\underline{\xi}_t) \geq \bar{q}^{\bar{L}-(t+1)}(0) - \tau_{t+1}\varepsilon$. For $v \geq 0$, $v - 2\varepsilon < \frac{v}{1+\varepsilon} < \frac{v}{1-\varepsilon} < v + 2\varepsilon$ for small enough ε.[28] Then

$$\begin{aligned}
\underline{\xi}_t &> (1-\varepsilon)\phi^* + 2\chi_{t+1}(\underline{\xi}_t) - 2\varepsilon > \phi^* + 2(\bar{q}^{\bar{L}-(t+1)}(0) - \tau_{t+1}\varepsilon) - 3\varepsilon \\
&= \phi^* + 2(\bar{q}^{\bar{L}-(t+1)}(0) - 3(2^{\bar{L}-(t+1)} - 1)\varepsilon) - 3\varepsilon = \phi^* + 2\bar{q}^{\bar{L}-(t+1)}(0) - [3(2^{\bar{L}-t} - 2) + 3]\varepsilon \\
&= \bar{q}^{\bar{L}-t}(0) - \tau_t\varepsilon.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\bar{\xi}_t &< (1+\varepsilon)\phi^* + 2\chi_{t+1}(\bar{\xi}_t) + 2\varepsilon < \phi^* + 2(\bar{q}^{\bar{L}-(t+1)}(0) + \tau_{t+1}\varepsilon) + 3\varepsilon \\
&= \phi^* + 2(\bar{q}^{\bar{L}-(t+1)}(0) + 3(2^{\bar{L}-(t+1)} - 1)\varepsilon) + 3\varepsilon = \phi^* + 2\bar{q}^{\bar{L}-(t+1)}(0) + [3(2^{\bar{L}-t} - 2) + 3]\varepsilon \\
&= \bar{q}^{\bar{L}-t}(0) + \tau_t\varepsilon.
\end{aligned}$$

Then:

1. If $\xi_t < \bar{q}^{\bar{L}-t}(0) - \tau_t\varepsilon$, then $\lambda_t(\xi_t) > \mathsf{V}(1+\varepsilon, v_{t+1}) + \eta + \varepsilon$.
2. If $\xi_t > \bar{q}^{\bar{L}-t}(0) + \tau_t\varepsilon$, then $\lambda_t(\xi_t) < \mathsf{V}(1+\varepsilon, v_{t+1})$.
3. Define $\pi_{t-1}(\xi_t) \equiv \hat{\pi}_1^*(\lambda_t(\xi_t), v_t)$.

By induction, the above statements hold for each $t \in \{1, ..., \bar{L}-2\}$. From the $t = 1$ case, the continuity of $\pi_0(\xi_1^*) \equiv \hat{\pi}_1^*(\lambda_1(\xi_1), v_1)$ and $\lambda_1(\cdot)$ ensures there exists $\xi_1^* \in B_{\tau_1\varepsilon}(\bar{q}^{\bar{L}-1}(0))$ such that $\pi_0(\xi_1^*) = \xi_1^*$. Inductively pick $\xi_{t+1}^* \equiv \chi_{t+1}(\xi_t^*)$ for each $t \in \{1, ..., \bar{L}-2\}$.

With $(\xi_1^*, ..., \xi_{\bar{L}-1}^*)$ in hand, I complete the construction of $\tilde{\sigma}$. Define $\Delta\Psi_1(\Delta z_1)$ as the cdf of $\Delta z_1 = z_1^C - z_1^D$, and define $\Delta\Psi_1^{-1}(\cdot)$ as the inverse of $\Delta\Psi_1(\cdot)$ (which exists because ψ has strictly positive density). Define $\xi_0^* \equiv 1, \xi_{\bar{L}}^* \equiv 0$. For each $t \in \{0, ..., \bar{L}-1\}$, define

$$\tilde{\sigma}_2(I_t^t, z_2) = \begin{cases} d & \Delta z_2 \leq \bar{\zeta}_2(\Delta\Psi_1^{-1}(1 - \pi_t(\xi_{t+1}^*)), \xi_t^*) \\ c & \text{otherwise} \end{cases} \quad \tilde{\sigma}_1(I_t^t, z_1) = \begin{cases} D & \Delta z_1 \leq \Delta\Psi_1^{-1}(1 - \pi_t(\xi_{t+1}^*)) \\ C & \text{otherwise.} \end{cases}$$

It is straightforward to show that $P_{\tilde{\sigma}}(I_t^t|\theta_0) = \xi_t^*$ for each t. This yields beliefs $\bar{\mu}(\xi_t^*)$ at each $I_1^1, ..., I_{\bar{L}-1}^{\bar{L}-1}$ consistent with $P_{\tilde{\sigma}}(\cdot)$, and the construction of $\bar{\zeta}_2(\cdot)$ proves that the player 2 strategies are best responses. It is also straightforward to check that $V_{\tilde{\sigma}}(I_t^t) = \bar{g}(\lambda_{t+1}(\xi_{t+1}^*), v_{t+1}, \xi_t^*) = \lambda_t(\xi_t^*)$ for each $t \in \{1, ..., \bar{L}-2\}$. Finally, note that since $\pi_0(\xi_1^*) = \xi_1^* \in B_{\tau_1\varepsilon}(\bar{q}^{\bar{L}-1}(0))$ and $\bar{q}^{\bar{L}-1}(0) < \frac{1}{2}$, it must be that $\lambda_1(\xi_1^*) - v_1 < \eta + \varepsilon$. Furthermore,

$$\varepsilon\bar{\zeta}_2(\hat{\zeta}_1(\lambda_1(\xi_1^*) - v_1), 1) = 1 - 2\mu^0 - 2(1-\mu^0)\pi_0(\xi_1^*) > 1 - 2\mu^0 - 2(1-\mu^0)(\bar{q}^{\bar{L}-1}(0) + \tau_1\varepsilon)$$

[28] To see this, suppose not: $v - 2\varepsilon \geq \frac{v}{1+\varepsilon}$ or $\frac{v}{1-\varepsilon} \geq v + 2\varepsilon$. If $v - \varepsilon v - 2\varepsilon - 2\varepsilon^2 \geq v$, then $-v - 2 - 2\varepsilon \geq 0 \implies 2\varepsilon \leq -v - 2$ a contradiction. If $\frac{v}{1-\varepsilon} \geq v + 2\varepsilon$, then $v \geq v - \varepsilon + 2\varepsilon - 2\varepsilon^2 \implies 0 \geq 1 - 2\varepsilon$, a contradiction for small ε.

and rearranging the right hand side yields

$$\begin{aligned}
\varepsilon\bar{\zeta}_2(\hat{\zeta}_1(\lambda_1(\xi_1^*)-v_1),1) &> (\mu^0+(1-\mu^0))-\mu^0-(1-\mu^0)\left(\phi^*+2\bar{q}^{\bar{L}-1}(0)\right)-2(1-\mu^0)\tau_1\varepsilon\\
&= (\mu^0+(1-\mu^0))-\mu^0-(1-\mu^0)\bar{q}^{\bar{L}}(0)-2(1-\mu^0)\tau_1\varepsilon\\
&= (1-\mu^0)\left(1-\bar{q}^{\bar{L}}(0)\right)-2(1-\mu^0)\tau_1\varepsilon.
\end{aligned}$$

Since $\bar{q}^{\bar{L}}(0)<1$, for small enough ε it must be that $\bar{\zeta}_2(\hat{\zeta}_1(\lambda_1(\xi_1^*)-v_1),1)>1$, and so $\tilde{\sigma}_2(I_0^0,z_2)=d$ for all $z_2\in Z_2$. Thus, $(1-\delta)+\delta v_1<V(\tilde{\sigma})<(1-\delta)(1+2\varepsilon)+\delta v_1$, and since $v_1=\mathsf{V}^{\bar{L}-1}(1+\varepsilon\bar{z}_1^D,v_{\bar{L}})=\mathsf{V}^{\bar{L}-1}(1+\varepsilon\bar{z}_1^D,2-\eta+2\varepsilon)\in B_{2\varepsilon}(\mathsf{V}^{\bar{L}-1}(1,2-\eta))$, then $V(\tilde{\sigma})\in B_{2\varepsilon}(\mathsf{V}^{\bar{L}}(1,2-\eta))$. ■

For each l, Lemma C.17 shows there is a K-length vector γ_{-K}^l so that for any payoff function $\gamma^l\equiv(\gamma_{-K}^l,\gamma_K^l)$ for any γ_K^l, it is possible to construct a wPBE $\tilde{\sigma}^l$ for the (ψ^l,ε^l)-perturbed γ^l-antegame where $V(\tilde{\sigma}^l)\in B_{2\varepsilon^l}(\mathsf{V}^{\bar{L}}(1,2-\eta))$. Construct history distribution ϕ^l to match the equilibrium probabilities of reaching final histories under $\tilde{\sigma}^l$: $\phi_k^l\equiv P_{\tilde{\sigma}}(I_k)$. Thus, $\tilde{\sigma}^l\in\Sigma_{(\phi,\gamma)}^{(\psi,\varepsilon)*}$. Corollary C.3 shows we can pick $\gamma_K^l\in B_{\varepsilon^l}(2)$ so that there exists a bounded (ψ^l,ε^l)-self-generating set W^{*l} containing (ϕ^l,γ^l), such that for every element $w'\equiv(\phi',\gamma')\in W^{*l}$, $\gamma_0'\in B_{\varepsilon^l(1+\eta)}(2-\eta)$; $\gamma_k'\in B_{2\varepsilon^l}(2-\varepsilon^l)$ for all $k\in\{1,...,K-1\}$; and $\gamma_K'\in B_{\varepsilon^l}(2)$. Then for any l and any pair of HBPs $\tilde{w}^l\equiv(\tilde{\phi}^l,\tilde{\gamma}^l)\in W^{*l},\tilde{w}^{l+1}\equiv(\tilde{\phi}^{l+1},\tilde{\gamma}^{l+1})\in W^{*(l+1)}$, the Chebyshev distance between $\tilde{\gamma}^l$ and $\tilde{\gamma}^{l+1}$ is bounded by $\|\tilde{\gamma}^l-\tilde{\gamma}^{l+1}\|_\infty<[(2-\varepsilon^{l+1})+2\varepsilon^{l+1}]-[(2-\varepsilon^l)-2\varepsilon^l]=\varepsilon^{l+1}+3\varepsilon^l<4\varepsilon^l$. From the definition of self-generation, for each l and any pair of HBPs $w^{l,0}\in W^{*l},w^{l+1,0}\in W^{*l+1}$, there exist $w^{l,1}\in W^{*l},w^{l+1,1}\in W^{*l+1}$ such that $w^{l,0}\in\mathcal{B}_{\psi^l,\varepsilon^l}(w^{l,1}),w^{l+1,0}\in\mathcal{B}_{\psi^{l+1},\varepsilon^{l+1}}(w^{l+1,1})$.

Corollary C.4. *The sequence $\left(\int\tilde{\sigma}_i^l(I_k^t,z_i)\,d\psi(z_i)\right)_l$ converges as $l\to\infty$ for all i, t, k.*

Proof. I reuse and adapt notation from the proof of Lemma C.18. Let $\xi_t^{*l}=P_{\tilde{\sigma}^l}(I_t^t|\theta_0)$ for all $t<\bar{L}$ and $\xi_{\bar{L}}^{*l}\equiv 0$; define $v_{\bar{L}}^l\equiv 2-\eta+2\varepsilon^l$ and inductively define $v_{k-1}^l\equiv(1-\delta)(1+\varepsilon^l\bar{z}_1^D)+\delta v_k^l$ for each $k\in\{1,...,\bar{L}\}$. Then $V_{\tilde{\sigma}^l}(I_k^t)=v_t^l$ for all $t<\bar{L}$, $k<t$. Abbreviate $E\tilde{\sigma}_i^l(I_k^t)\equiv\int\tilde{\sigma}_i^l(I_k^t,z_i)\,d\psi(z_i)$. The proof of Lemma C.18 shows that $\xi_t^{*l}\in B_{\tau_1\varepsilon^l}(\bar{q}^{\bar{L}-t}(0))$, and so $\lim_l\xi_t^{*l}=\bar{q}^{\bar{L}-t}(0)$. Since $E\tilde{\sigma}_1^l(I_t^t)=\pi_t(\xi_{t+1}^{*l})$ and $E\tilde{\sigma}_2^l(I_t^t)=\bar{\zeta}_2(\lambda_{t+1}(\xi_{t+1}^{*l}),v_{t+1}^l,\xi_{t+1}^{*l})$, $(E\tilde{\sigma}_i^l(I_t^t))_l$ converges for all i, $t<\bar{L}$. Since $\tilde{\sigma}_1^l(I_k^t,z_1)=D$ and $\tilde{\sigma}_2^l(I_k^t,z_2)=d$ for all $t<\bar{L},k<t,z\in Z$, convergence is clear.

For periods $t\in\{\bar{L},...,K-1\}$, the proof of Lemma C.18 constructs strategies at I_k^t for all $k<t$ such that

$$V_{\tilde{\sigma}^l}(I_0^t)=2-\eta-2\varepsilon^l<2-4\varepsilon^l<V_{\tilde{\sigma}^l}(I_1^t)<\cdots<V_{\tilde{\sigma}^l}(I_{t-1}^t)<2-\varepsilon^l. \tag{78}$$

The proof of Lemma C.18 also shows that $V_{\tilde{\sigma}^l}(I_t^t)\in B_{\varepsilon^l}(2)$ for all $t\geq\bar{L}$. For all $t\geq\bar{L},k$, the fact that the continuation payoffs $(V_{\tilde{\sigma}^l}(I_k^t))_l$ converge as $l\to\infty$ and satisfy (78) allows a straightforward adaptation of the arguments in the proof of Lemma C.13 to show that $(E\tilde{\sigma}_i^l(I_k^t))_l$ converges.[29] ■

[29]The proof of Lemma C.13 shows that the sequence $(W^l,w^l,\tilde{\sigma}^l)_l$ satisfying its assumptions implies the existence of a sequence $(\bar{\sigma}^l)_l$ where $\bar{\sigma}^l$ is a wPBE of the (ψ^l,ε^l)-perturbed full game. For all periods $t\geq K$, the proof shows that the fact that the continuation payoffs satisfy

$$V_{\bar{\sigma}^l}(t,I_K)-2(1-\delta)+2\varepsilon^l<V_{\bar{\sigma}^l}(t,I_0)<V_{\bar{\sigma}^l}(t,I_0)<\cdots<V_{\bar{\sigma}^l}(t,I_K)\in B_{\varepsilon^l}(2)$$

plus the fact that $(V_{\bar{\sigma}^l}(t,I_k))_l$ converges for all k implies that the sequence$\left(\int\bar{\sigma}_i^l(t,I_k,z_i)\,d\psi(z_i)\right)_l$ converges.

Thus, the assumptions of Lemma C.13 are satisfied, so there exists a purifiable wPBE σ of the unperturbed full game G^∞ such that $V(\sigma) = \lim_l V(\tilde{\sigma}^l) = \mathsf{V}^{\bar{L}}(1, 2 - \eta)$.