

Stochastic Games in Continuous Time: Persistent Actions in Long-Run Relationships*

J. Aislinn Bohren[†]

June 15, 2013

Abstract

This paper studies a new class of stochastic games in which the actions of a long-run player have a persistent effect on the payoff and information structure. For example, the quality of a firm's product is determined by past as well as current effort or a policy variable persistently depends on choices made by a governing board. The setting is a continuous time game of imperfect monitoring between a long-run and a representative myopic player. I establish general conditions for the existence of Markov equilibria and conditions for the uniqueness of a Markov equilibrium in the class of *all* Perfect Public Equilibria . The existence proof is constructive and characterizes, for any discount rate, the explicit form of equilibrium payoffs, continuation values, and actions in Markov equilibria as a solution to a second order ODE. Action persistence creates a new channel to provide intertemporal incentives in a setting where traditional channels are ineffective, and offers a novel and different framework for thinking about the reputations of firms, governments, and other long-run agents.

*I thank David Miller, George Mailath, Paul Niehaus, Yuiy Sannikov, Andy Skrzypacz, Joel Sobel, Jeroen Swinkels, Joel Watson and especially Nageeb Ali for useful comments. I also thank numerous seminar participants for helpful feedback.

[†]Email: abohren@sas.upenn.edu; University of Pennsylvania

1 Introduction

Persistence and rigidities are fundamental features of many economic settings. Often, payoffs depend not only on actions today, but are also determined by choices made further in the past. I study a class of stochastic games between a long-run player and a population of small players in which actions have a persistent effect on the information and payoff structure of the game. This departure from a standard repeated game introduces a new channel for intertemporal incentives: in addition to conditioning future equilibrium play on past outcomes, as is possible in repeated games, players' actions now directly impact the future set of feasible payoffs and information structure.

Examples abound of situations in which past actions influence key characteristics of the long-run player's environment. A firm's ability to make high quality products is a function of not only its effort today but also its past investments in developing technology and training its workforce. A government's ability to offer efficient and effective public services to its citizens depends on its past investments in improving infrastructure and building competent institutions. A doctor doesn't cure patients through instantaneous effort alone, but needs to undertake costly training to develop skills and learn techniques. In all of these settings, and many others, past choices play a central role in determining current and future profitability.

This paper studies a continuous-time model between a long-run player and a continuum of small anonymous players, in which actions have a persistent effect on the structure of the game and the action of the long-run player is imperfectly observed. Small players learn about the long-run player's action through a Brownian public signal. A payoff-relevant state variable captures the impact of past actions, and its evolution is also driven by Brownian information. This setting has two potential channels through which intertemporal incentives can be used to guide behavior. First, as in repeated games, players can be rewarded or punished based on the public signal: actions today can affect, in equilibrium, how others behave in the future. Second, the game varies across time in response to players' actions, and thus a long-run player's actions directly impact payoffs and information in future periods as well as the current period.

Faingold and Sannikov (2011) and Fudenberg and Levine (2007) show that in standard continuous time games with a single long-run player and Brownian information, intertemporal incentives fail: the long-run player cannot attain payoffs beyond those of the highest static Nash equilibrium payoff. I show that allowing actions to persistently impact payoffs creates a crucial additional channel to generate intertemporal incentives, and allows the long-run player to earn payoffs higher than those earned by playing a myopic best response. Allowing actions to have a persistent effect could reintroduce the possibility of effectively coordinating equilibrium play, in addition to creating a new channel for intertemporal incentives. I show that this is not the case; non-trivial incentives stem entirely from the direct impact of past actions on future payoffs. In equilibrium, players disregard the public signal and the value of the state variable solely determines equilibrium actions and continuation values.

Markov equilibria play a key role in this setting. The main results of this paper establish conditions on the structure of the game that guarantee existence of Markov equilibria, characterize the Perfect Public Equilibria (PPE) payoff set as the convex hull of the Markov

equilibria correspondence, and determine when a Markov equilibrium is unique. The existence proof is constructive, and specifies the explicit form of Markov equilibria continuation values and actions, for any discount rate, as a function of the state variable. A *nonstochastic* differential equation defined over the state space determines equilibrium continuation values, while the long-run player's action is determined by the sensitivity of its future payoffs to changes in the state variable (the first derivative of this solution). Since the PPE payoff set is bounded by the highest and lowest Markov payoff, no equilibria use the public signal to generate intertemporal incentives and the sensitivity of the continuation value to the public signal is zero.

At first glance, there appears to be a tension between my result and other continuous-time repeated games papers. Say SS rule out effective incentives in this setting, blah blah. But actually, the results are consistent blah blah. Don't use linearly.... Sannikov and Skrzypacz (2010) show that burning value through punishments that affect *all players* is not effective for incentives in settings with Brownian signals; Brownian information must be used linearly. In some settings, it is possible to structure punishments that linearly transfer value between players. But in many settings, including those between long-run and myopic players, tangential transfers are not possible.¹ In a standard repeated game, using Brownian information linearly in a direction that is non-tangential to the boundary of the equilibrium payoff set will result in the continuation value escaping its upper bound with positive probability, a contradiction. Thus, it is not possible to structure incentives at the long-run player's highest continuation payoff, and intertemporal incentives collapse. However, in a stochastic game, it is possible for players to play a static Nash at the value of the state variable that yields the highest continuation payoff and to linearly use Brownian motion to structure non-trivial incentives at other values of the state variable. Therefore, the intertemporal incentives that persistent actions induce could not emerge with standard continuous-time repeated games.

This characterization develops a new understanding of reputational dynamics. Since Kreps, Milgrom, Roberts, and Wilson (1982), the canonical framework has modeled a long-run player's reputation as the belief that others have that the firm is a behavioral type and takes a fixed action in each period. Behavioral types connect behavior across time, and enable a long-run player to overcome binding moral hazard. In fact, behavioral types can be viewed as one form of action persistence, in which the state variable depends on small players' beliefs about the long-run player's past actions, as well as directly depending on past actions. However, there are many other ways that persistent actions can strengthen the link between past behavior and future payoffs without using incentives that are driven exclusively by the possibility that players may be non-strategic. This paper returns to a notion of reputation as a type of capital, stemming from actions that have a prolonged effect on the structure of the game (Klein and Leffler 1981) The effectiveness of this alternative form of persistence shows that it is action persistence generally, rather than a particular form of persistence as a behavioral type, that drives the power of reputation effects.

Viewing reputation as an asset connects reputation to the aspects of a firm's choices

¹Sannikov and Skrzypacz (2007) show how this issue also arises in games between multiple long-run players in which deviations between individual players are indistinguishable.

that are empirically identifiable. The tractable characterization of equilibrium actions and payoffs ties behavior to the parameters of the model such as the cost and depreciation rate of investment, the volatility of quality or the value consumers place on cheap, low-quality products compared to expensive high-end products. It also provides insights on the dynamics of reputation formation, including: when does a firm or government build its reputation and when does it allow it to decay; when do reputation effects persist in the long-run, and when are they temporary; what drives a firm to specialize in different quality products or an expert to seek verifiable certifications? The value of a reputation depends on the nature of persistence and market demand - the structure of the underlying game determines the shape of equilibrium continuation values.

These results of this paper have practical implications for equilibrium analysis in a wide range of applied settings known to exhibit persistence and rigidities, ranging from industrial organization to political economy to macroeconomics. Markov equilibria are a popular concept in applied work. Advantages of Markov equilibria include their simplicity and their dependence on payoff relevant variables to specify incentives. Establishing that non-Markov equilibria do not exist offers a strong justification for focusing on this more tractable class of equilibria. Additionally, this paper derives a tractable expression to construct Markov equilibria, which provides a tool for equilibrium analysis in applications. Once functional forms are specified for the underlying game, it is straightforward to derive the Markov equilibrium, calibrate it with realistic parameters, and use numerical methods to estimate its solution. This can be used to formulate empirically testable predictions about equilibrium behavior.

To fix ideas, consider the canonical product choice game. Consider a long-run firm interacting with a sequence of short-run consumers. The firm has a dominant strategy to choose low effort, but would have greater payoffs if it could somehow commit to high quality (its “Stackelberg payoff”). Repeated interaction in discrete time with imperfect monitoring generates a folk theorem (Fudenberg and Levine 1994), but the striking implication from Faingold and Sannikov (2011) and Fudenberg and Levine (2007) is that such intertemporal incentives disappear as the period length becomes small. Since Fudenberg and Levine (1992), we know that if the firm could build a reputation for being a *commitment type* that produces only high quality products, a patient normal firm can approach these payoffs in every equilibrium. Faingold and Sannikov (2011) shows that this logic remains in continuous-time games, but that as in discrete-time, these reputation effects are temporary: eventually, consumers learn the firm’s type, and reputation effects disappear in the long-run (Cripps, Mailath, and Samuelson 2007).²

Departing from standard repeated and reputational games, consider a simple and realistic modification in which the firm’s product quality is a noisy function of current and past effort. Applying the results of this paper to the product choice setting shows that there is a unique Perfect Public Equilibrium, which is Markov in the past investment of the firm. The firm’s reputation for quality will follow a cyclical pattern, characterized by phases of reputation building and decay. Importantly, this cyclical pattern does not dissipate with time

²Mailath and Samuelson (2001) show that reputational incentives can also come from a firm’s desire to *separate* itself from an incompetent type. Yet, these reputation effects are also temporary unless the type of the firm is replaced over time.

and reputation effects are permanent; this contrasts with the temporary reputation effects observed in behavioral types models. The product choice game is just one of the many settings that can utilize the tools of this paper to shed light on the relationship between persistent actions and equilibrium behavior.

The product choice application is similar in spirit to [Board and Meyer-ter-Vehn \(2011\)](#), who also look at persistent quality as an alternative way to generate reputation effects. In their setting, product quality is binary and consumers learn about quality through noisy signals. Reputation is defined as the consumers' belief that the current product quality is high. Quality is periodically replaced via a Poisson arrival process; when a replacement occurs, the firm's current effort determines the new quality value. Realized product quality in their setting is therefore discontinuous (jumping between low and high), and this discontinuity plays an important role in determining intertemporal incentives. In the product choice application of my setting, the quality of a firm's product is a smooth function of past investments and its investment today, and thus, the analysis is very different.

This paper builds on tools developed by [Faingold and Sannikov \(2011\)](#). Their setting can be viewed as a stochastic game in which the state variable is the belief that the firm is a commitment type, and the transition function follows Bayes rule. [Faingold and Sannikov \(2011\)](#) characterize the unique Markov equilibrium of this incomplete information game using an ordinary differential equation. This paper extends these tools to establish conditions for the existence and uniqueness of Markov equilibria in a setting with an arbitrary transition function between states, which may be independent of the public signal, where the long-run player's payoffs may also depend on the state variable and a general state space that may be bounded or unbounded and have multiple interior absorbing states. Allowing the state variable to be independent of the public signal isolates the role of persistence on generating incentives. The general state space allows for richer dynamics, including models where equilibrium payoffs are not monotonic with respect to the state variable and models where persistence generates long-run incentives.

The organization of this paper proceeds as follows. [Section 2](#) sets up the model. [Section 3](#) establishes the existence of a Markov equilibrium, characterizes equilibrium behavior and payoffs and determines conditions for a Markov equilibrium to be unique in the class of all perfect public equilibria. Three simple examples are introduced at the end of [2](#), and used to illustrate the main results of the model at the end of [4](#). A final section relates properties of equilibrium payoffs to the underlying stage game. All proofs are in the Appendix.

2 Model

I study a stochastic game of imperfect monitoring between a single long-run player and a continuum of small, anonymous short-run players, $I = [0, 1]$, with each individual indexed by i . Time $t \in [0, \infty)$ is continuous.

Actions and States: At each instant t , long-run and short-run players simultaneously choose actions a_t from A and b_t^i from B , respectively, where A and B are compact sets of a Euclidean space. Individual actions are privately observed, while the aggregate distribution

of short-run players' actions, $\bar{b}_t \in \Delta B$ and a public signal of the long-run player's action, Y_t , are publicly observed. The stage game varies across time through its dependence on a publicly observed state variable $X_t \in \Xi \subset \mathbb{R}$, which determines the payoff and information structure. The state variable evolves stochastically as a function of the current state and players' actions. Represent the public signal profile and state variable by a system of stochastic differential equations,

$$\begin{bmatrix} dY_t \\ dX_t \end{bmatrix} = \begin{bmatrix} \mu_y(a_t, \bar{b}_t, X_t) \\ \mu_x(a_t, \bar{b}_t, X_t) \end{bmatrix} dt + \begin{bmatrix} \sigma_y(\bar{b}_t, X_t) & \sigma_{yx}(\bar{b}_t, X_t) \\ \sigma_{xy}(\bar{b}_t, X_t) & \sigma_{xx}(\bar{b}_t, X_t) \end{bmatrix} \cdot \begin{bmatrix} dZ_t^y \\ dZ_t^x \end{bmatrix}$$

where $(Z_t^y, Z_t^x)_{t \geq 0}$ is a d -dimensional Brownian motion, $\mu_y : A \times B \times \Xi \rightarrow \mathbb{R}^{d-1}$ is the drift of the public signal, $\mu_x : A \times B \times \Xi \rightarrow \mathbb{R}$ is the drift of the state variable and

$$\sigma = \begin{bmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{bmatrix} : B \times \Xi \rightarrow \mathbb{R}^{d \times d}$$

is the volatility matrix, with each function linearly extended to $A \times \Delta B$. Assume μ_y , μ_x and σ are Lipschitz continuous functions.³

The drift of the public signal, μ_y provides a signal of the long-run player's action and can also depend on the aggregate action of the short-run players and the state, but is independent of individual actions b_t^i to preserve anonymity. Volatility is independent of the long-run player's action to maintain the assumption of imperfect monitoring.⁴ The state variable influences the information structure through the drift and volatility of the public signal profile, and may also provide an additional signal of the long-run player's action through its own drift, μ_x . Let $\{F_t\}_{t \geq 0}$ represent the filtration generated by the public information, $(Y_t, X_t)_{t \geq 0}$.

Define a state \tilde{X} as an *absorbing state* if the drift and volatility of the transition function, dX_t , are both zero.

Definition 1. $\tilde{X} \in \Xi$ is an absorbing state if $\mu_x(a, b, \tilde{X}) = 0$ and $(\sigma_{xy}(b, \tilde{X}), \sigma_{xx}(b, \tilde{X})) = 0$ for all $(a, \bar{b}) \in A \times B$.

I make several assumptions on the structure of $(Y_t, X_t)_{t \geq 0}$. First, the volatility of the state variable is positive at all but a finite number of points of the state space. This ensures that the future path of the state variable is almost always stochastic.

Assumption 1. The state space can be represented as the union of a finite number of compact subsets, $\Xi = \cup_{i=1}^n \Xi_i$ such that for any compact proper subset $I \subset \Xi_i$ for some $i = 1, \dots, n$, there exists a c_I such that

$$\sigma_I = \inf_{\bar{b} \in B, X \in I} \left[|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right] > c_I$$

³The state space may or may not be bounded. It is bounded above if there exists an upper bound \bar{X} at which the volatility is zero, $(\sigma_{xy}(\bar{X}), \sigma_{xx}(\bar{X})) = (0, 0)$, and the drift is weakly negative, $\mu_x(a, \bar{b}, \bar{X}) \leq 0$ for all $(a, \bar{b}) \in A \times B$; it is bounded below if there exists a lower bound \underline{X} such that the volatility is zero and the drift is weakly positive.

⁴The volatility must be independent of the long run player's action since it is essentially observable.

Second, in order to maintain imperfect monitoring, the long-run player's action cannot be inferred from the path of the public signal and state variable. This requires all drift terms that depend on a_t to be unobservable: the volatility of each public signal is positive, and the rows of the public signal volatility matrix, σ_y , are linearly independent. Additionally, when the volatility of the state variable is positive, the rows of the total volatility matrix, σ , must be linearly independent. At any point where the volatility of the state variable is zero, and the drift of the state variable must be independent of the long-run player's action.

Assumption 2. 1. *There exists a constant c_1 such that $|\sigma_{yy}(b, X) \cdot y| \geq c_1 |y|$ for all $y \in \mathbb{R}^d, b \in B$ and $X \in \Xi$.*

2. *If $X \in \Xi$ is not an endpoint of the state space, then there exists a constant c_2 such that $|\sigma(b, X) \cdot y| \geq c_2 |y|$ for all $y \in \mathbb{R}^d, b \in B$.*⁵

3. *If $(\sigma_{xy}(\tilde{b}, \tilde{X}), \sigma_{xx}(\tilde{b}, \tilde{X})) = 0$, then there exists a function $f(b, X)$ such that $\mu_x(a, \tilde{b}, \tilde{X}) = f(\tilde{b}, \tilde{X})$ at (\tilde{b}, \tilde{X}) .*

The game defined here includes several subclasses of games. If $d = 1$, then the state variable contains all relevant information about the long-run player's action and there is no additional public signal. If $\mu_x(a, \bar{b}, \tilde{X})$ is independent of a , then the state evolves independently of the long-run player's action. If $\sigma_{xy} = 0$ and $\sigma_{yx} = 0$, then the state variable and the public signal are independent.

Payoffs: The state variable determines the set of feasible payoffs in a given instant. Given an action profile (a, \bar{b}) and a state X , the long-run player receives an expected flow payoff of $g(a, \bar{b}, X)$. It seeks to maximize the expected normalized discounted payoff,

$$r \int_0^\infty e^{-rt} g(a_t, \bar{b}_t, X_t) dt$$

where r is the discount rate. Assume g is Lipschitz continuous and bounded for all $a \in A, \bar{b} \in \Delta B$ and $X \in \Xi$.

Short-run players have identical preferences: they each seek to maximize the expected flow payoff at time t ,

$$h(a_t, b_t^i, \bar{b}_t, X_t)$$

which is a continuous function. Ex post payoffs can only depend on a_t through public information, as is standard in games of imperfect monitoring.⁶

The dependence of payoffs on the state variable creates a form of action persistence for the firm, since the state variable is a function of prior actions. It is possible that the long-run player's payoffs only depend on the state indirectly through the action of the short-run

⁵This last assumption stems from the fact that if μ_x is independent of a , then $\sigma_{xy} \cdot dZ_t^y + \sigma_{xx} dZ_t^x$ is observable. Therefore, if the volatility of any of the public signals is a scalar multiple of the volatility of the state, this would render the long run player's action as observable.

⁶Alternatively, one could consider situation with a sequence of short-run players, in which payoffs depend directly on a but the short-run players are not able to pass on knowledge of a to subsequent players.

players ($g(a, \bar{b}, X) = g(a, \bar{b})$). If both g and h are independent of X , then the game reduces to a standard repeated game.

The timing of the stochastic game is as follows: at each instant t , players observe the current state X_t , choose actions, and then nature stochastically determines payoffs, the public signal and next state as a function of the current state and action profile.

Strategies: A public strategy for the long-run player is a stochastic process $(a_t)_{t \geq 0}$ with values $a_t \in A$ and progressively measurable with respect to $\{F_t\}_{t \geq 0}$. Likewise, a public strategy for a short-run player is an action $b_t^i \in B$ progressively measurable with respect to $\{F_t\}_{t \geq 0}$.

2.1 Equilibrium Structure

Perfect Public Equilibria: I restrict attention to pure strategy perfect public equilibria (PPE). A public strategy profile is a PPE if after any public history and for all t , no player wants to deviate given the strategy profile of its opponents.

In any PPE, short-run players choose b_t^i to myopically optimize expected flow payoffs each instant.⁷ Let $\mathcal{B} : A \times \Delta B \times \Xi \rightrightarrows B$ represent the best response correspondence that maps an action profile and a state to the set of short-run player actions that maximize payoffs in the current stage game, and $\bar{\mathcal{B}} : A \times \Xi \rightrightarrows \Delta B$ represent the aggregate best response function. In many applications, it will be sufficient to specify the aggregate best response function as a reduced form for short-run players' behavior.

Define the long-run player's continuation value as the expected discounted payoff at time t , given the public information contained in $\{F_t\}_{t \geq 0}$ and strategy profile $S = (a_t, b_t^i)_{t \geq 0}$:

$$W_t(S) := E_t \left[r \int_t^\infty e^{-r(s-t)} g(a_s, \bar{b}_s, X_s) ds \right]$$

The long-run player's action at time t can impact its continuation value through two channels: (1) future equilibrium play and (2) the set of future feasible flow payoffs. It is well known that the public signal can be used to punish or reward the long-run player in future periods by allowing continuation play to depend on the realization of the public signal. A stochastic game adds a second link between current play and future payoffs: the long-run player's action affects the evolution of the state variable, which in turn determines the set of future feasible stage payoffs. Each channel provides a potential source of intertemporal incentives.

This paper applies recursive techniques for continuous time games with imperfect monitoring to characterize the evolution of the continuation value and the long-run player's incentive constraint in a PPE. Fix an initial value for the state variable, X_0 .

⁷The individual action of a short run player, b_t^i , has a negligible impact on the aggregate action \bar{b}_t (and therefore X_t) and is not observable by the long run player. Therefore, the model could also allow for long run small, anonymous players.

Lemma 1. *A public strategy profile $S = (a_t, b_t^i)_{t \geq 0}$ is a PPE with continuation values $(W_t)_{t \geq 0}$ if and only if for some $\{F_t\}$ – measurable process $(\beta_t)_{t \geq 0}$ in \mathcal{L}*

1. $(W_t)_{t \geq 0}$ is a bounded process and satisfies:

$$\begin{aligned} dW_t(S) &= r(W_t(S) - g(a_t, \bar{b}_t, X_t)) dt \\ &\quad + r\beta_{yt} [dY_t - \mu_y(a_t, \bar{b}_t, X_t)dt] \\ &\quad + r\beta_{xt} [dX_t - \mu_x(a_t, \bar{b}_t, X_t)dt] \end{aligned}$$

given $(\beta_t)_{t \geq 0}$

2. Strategies $(a_t, b_t^i)_{t \geq 0}$ are sequentially rational given $(\beta_t)_{t \geq 0}$. For all t , (a_t, b_t^i) satisfy:

$$\begin{aligned} a_t &\in \arg \max g(a', \bar{b}_t, X_t) + \beta_{yt}\mu_y(a_t, \bar{b}_t, X_t) + \beta_{xt}\mu_x(a_t, \bar{b}_t, X_t) \\ b_t^i &\in \mathcal{B}(a_t, X_t) \end{aligned}$$

The continuation value of the long-run player is a stochastic process that is measurable with respect to public information, $\{F_t\}_{t \geq 0}$. Two components govern the motion of the continuation value, a drift term that captures the difference between the current continuation value and the current flow payoff. This is the expected change in the continuation value. A volatility term β_{yt} determines the sensitivity of the continuation value to the public signal: future payoffs are more sensitive to good or bad signal realizations when this volatility is larger. A second volatility term β_{xt} captures the sensitivity of the continuation value to the state variable.

The condition for sequential rationality depends on the process $(\beta_t)_{t \geq 0}$, which specifies how the continuation value changes with respect to public information. Today's action impacts future payoffs through the drift of the public signal and state variable, μ_Y and μ_X , weighted by the sensitivity of the continuation value, β_x and β_y , while it impacts current payoffs through the flow payoff, $g(a, \bar{b}, X)$. A strategy for the long-run player is sequentially rational if it maximizes the sum of flow payoffs today and the expected impact of today's action on future payoffs. This condition is analogous to the one-shot deviation principle in discrete time.

A key feature of this characterization is the linearity of the continuation value and incentive constraint with respect to the Brownian information. Brownian information can only be used linearly to provide effective incentives in continuous time (Sannikov and Skrzypacz 2010). Therefore, the long-run player's incentive constraint takes a very tractable linear form, in which the process $(\beta_t)_{t \geq 0}$ captures all potential channels through which the long-run player's current action may impact future payoffs, including coordination of equilibrium play, the set of future feasible payoffs and the information structure of future public signals.

Remark 1. *The key aspect of this model that allows for this tractable characterization of the long-run player's incentive constraint is the assumption that the volatility of the state variable is almost always positive, which ensures that deviations do not alter the set of feasible paths for the state variable. Consider a deviation from a_t to \tilde{a}_t at time t . Given that the same*

paths are possible under a_t and \tilde{a}_t , the continuation value under both strategies is a non-degenerate expectation with respect to the future path of the state variable. Thus, the change in the continuation value when the long-run player deviates depends solely on the different probability measure the deviation induces over the evolution of state variable, but doesn't affect the support of the future paths. Given the linear incentive structure of Brownian information, this change is linear with respect to the difference in the drift of the public signal and state variable, $\mu_Y(\tilde{a}_t, \bar{b}_t, X_t) - \mu_Y(a_t, \bar{b}_t, X_t)$ and $\mu_X(\tilde{a}_t, \bar{b}_t, X_t) - \mu_X(a_t, \bar{b}_t, X_t)$.

Remark 2. *It is of interest to note that it is precisely this linear structure with respect to the Brownian information, coupled with the ineffectiveness of transferring continuation payoffs between players, that precludes the effective provision of intertemporal incentives in a standard repeated game with a long-run and short-run player. Myopic short-run players preclude incentive structures with tangential transfers between players on the boundary of the payoff set. Using Brownian information to create linear, non-tangential incentive structures results in the continuation value escaping the boundary of the payoff set with positive probability. Brownian information cannot be used in a non-linear manner, such as value-burning, to structure effective incentives.*

This paper will illustrate that effective intertemporal incentives are provided in a stochastic game by linearly (and non-tangentially) using Brownian information for some values of the state variable. In particular, incentives are structured so that the volatility of the continuation value depends on the state variable. When the state variable is at the value(s) that yields the maximum equilibrium payoff across all states, the continuation value is independent of the Brownian information to prevent it from escaping the boundary of the payoff set. In these periods, the long-run player acts myopically, as is the case in a standard repeated game. However, at other values of the state variable, the continuation value linearly depends on Brownian information, which creates effective intertemporal incentives.

Next, I use the sequential rationality condition to specify an auxiliary static game parameterized by the state variable and the volatility of the continuation value. Let $S^*(X, \beta_y, \beta_x) = \{(a, \bar{b})\}$ represent the correspondence of static Nash equilibrium action profiles in this auxiliary game, defined as:

Definition 2. *Define $S^*(X, \beta_y, \beta_x) = \Xi \times \mathbb{R}^d \rightrightarrows A \times \Delta B$ as the correspondence that describes the Nash equilibrium of the auxiliary static game parameterized by $(X, \beta_y, \beta_x) \in \Xi \times \mathbb{R}^d$:*

$$S^*(X, \beta_y, \beta_x) = \left\{ \begin{array}{l} a \in \arg \max_{a'} g(a', \bar{b}, X) + \beta_y \cdot \mu_y(a', \bar{b}, X) + \beta_x \mu_x(a', \bar{b}, X) \\ \bar{b} \in \bar{\mathcal{B}}(a, X) \end{array} \right\}$$

In any PPE strategy profile $(a_t, \bar{b}_t)_{t \geq 0}$ of the stochastic game, given some processes $(X_t)_{t > 0}$ and $(\beta_t)_{t > 0}$, the action profile at each instant must be a static Nash equilibrium of the auxiliary game i.e. $(a_t, \bar{b}_t) \in S^*(X_t, \beta_{yt}, \beta_{xt})$ for all t . Note that $S^*(X, 0, 0)$ corresponds to the static Nash equilibrium of the original game with the state variable equal to X . I assume that this auxiliary stage game has a unique static Nash equilibrium with an atomic distribution over small players' actions.

Assumption 3. Assume $S^*(X, \beta_y, \beta_x)$ is non-empty and single-valued for all $(X, \beta_y, \beta_x) \in \Xi \times \mathbb{R}^d$, Lipschitz continuous on any subset of $\Xi \times \mathbb{R}^d$, and the small players choose identical actions $b^i = \bar{b}$.

Remark 3. While this assumption is somewhat restrictive, it still allows for a broad class of games, including those discussed in the previous examples. It does not rule out games in which non-trivial equilibria of the repeated game are possible.

Static Equilibria Payoffs: The feasible payoffs of the stage game depend on the state variable, as do static Nash equilibrium payoffs. Define $v : \Xi \rightarrow \mathbb{R}$ as the payoff to the long-run player in a static Nash equilibrium, as a function of the state variable, where $v(X) := g(S^*(X, 0, 0), X)$. The assumption that the Nash equilibrium correspondence of the auxiliary stage game is Lipschitz continuous, non-empty and single-valued guarantees $v(X)$ is a Lipschitz continuous function and is well defined for all $X \in \Xi$. Represent the highest and lowest stage Nash equilibrium payoffs across all states as:

$$\begin{aligned}\bar{v}^* &= \sup_{X \in \Xi} v(X) \\ \underline{v}^* &= \inf_{X \in \Xi} v(X)\end{aligned}$$

These values are well-defined, given that g is bounded.

2.2 Interlude: Examples

This interlude presents four brief examples to illustrate the model. The first example uses a variation of the canonical product choice setting to demonstrate the model in a familiar framework. The remaining examples illustrate the breadth of the model by applying it to three diverse settings. At the end of Section 3, I will return to these examples to characterize payoffs and actions in the unique Markov equilibrium.

2.2.1 Persistent Effort as a Source of Reputation

Suppose a single long-run firm seeks to provide a sequence of short-run consumers with a product. At each instant t , the firm chooses an unobservable effort level $a_t \in [0, \bar{a}]$ and consumers simultaneously choose a purchase level $\bar{b}_t \in [0, 20]$. Effort is costly for the firm, but increases the likelihood of producing a high quality product both today and in the future.

The firm's quality depends on current and past effort. The stock component of quality, X_t , captures the persistent impact of past effort. This stock evolves according to a mean-reverting stochastic process

$$dX_t = \theta(a_t - X_t)dt + \sigma dZ_t.$$

with drift $\mu_x = \theta(a - X)$ proportional to the difference between current effort and stock quality, constant volatility σ , and Brownian noise dZ .⁸ Stock quality is increasing in expectation when effort exceeds the stock, and decreasing when effort is below the stock.

⁸This stochastic process is the Ornstein-Uhlenbeck process.

Parameter θ embodies the rate at which past effort decays; as θ increases, the impact of recent effort increases relative to effort further in the past.

The closed form of dX_t lends insight into the structure of persistence. Given a history of effort levels $(a_s)_{s \leq t}$ and initial quality stock X_0 , the current value of the stock is

$$X_t = X_0 e^{-\theta t} + \theta \int_0^t e^{-\theta(t-s)} a_s ds + \sigma \int_0^t e^{-\theta(t-s)} dZ_s$$

As shown in this expression, the impact of past effort decays at a rate proportional to θ and the time that has elapsed since the effort was made. Consumers observe the stock quality, which also provides a noisy public signal of the firm's effort each instant. Overall quality is a weighted average of the stock component and current effort,

$$q(a, X) = (1 - \lambda)a + \lambda X$$

where λ captures the relative importance of current and past effort in determining current quality.

Consumers enjoy high quality products: utility is increasing in expected quality. Marginal utility is decreasing in the purchase level, while the purchase price is fixed at one. Represent a consumer's payoff in instant t as:

$$\sqrt{b^i} q(a, X) - b^i$$

The firm's profit depends on the revenue from purchases and the cost of investment. Its average payoff is represented as

$$r \int_0^\infty e^{-rt} (\bar{b}_t - ca_t^2) dt$$

where r is the common discount rate and $c < 1$ captures the cost of investment. Note that this example satisfies the necessary continuity, positive volatility and imperfect monitoring assumptions from Section 2.

The firm is subject to binding moral hazard in that it would like to commit to a higher level of effort in order to entice consumers to choose a higher purchase level. However, in the absence of such a commitment device, the firm is tempted to deviate to low effort. In the static game, the firm always chooses the lowest possible effort level, $a^N = 0$, regardless of the stock quality X . This example seeks to characterize when intertemporal incentives, particularly incentives created by quality persistence, can provide the firm with endogenous incentives to choose a positive level of effort.

2.2.2 Building a Specialization with Oscillating Demand

Consider a variation of the previous example where a firm faces a demand curve that oscillates with quality. For example, a doctor chooses how much effort to spend specializing and gaining experience. Demand for doctors is non-monotonic in their skill levels: general practitioners and highly specialized doctors are in high demand, but intermediate levels of specialization

are less useful. Or a store decides whether to specialize in either high quality, expensive items, or low quality, cheap items; the market for intermediate levels of quality is thin. The set-up here mirrors the previous example, with a different utility function for the consumers. Consumers choose an employment level $b \in [0, 8]$ to maximize:

$$\sqrt{b^i (\sin 2q(a, X) + q/2)} - \frac{1}{2}b^i$$

The oscillating demand curve introduces a non-monotonicity into the instantaneous value of quality. It is possible for a firm to become too high quality for one market, but not yet high enough for the next. This variation seeks to characterize whether this non-monotonicity is also present in the long-run value of quality (the PPE payoff), and how this influences a firm's incentives to invest in quality.

2.2.3 Policy Targeting

Elected officials and governing bodies often play a role in formulating and implementing policy targets. For example, the Federal Reserve targets interest rates, a board of directors sets growth and return targets for its company, and the housing authority targets home ownership rates. Achieving such targets requires costly effort on behalf of officials; often, the policy level will depend on both current and past policy efforts. Moral hazard issues arise when the preferences of the officials are not aligned with the population they serve.

Consider a setting where constituents elect a board to implement a policy target. The policy X_t takes on values between $[0, 2]$. Constituents want to target $X_t = 1$, but in the absence of intervention, the policy drifts towards its natural level d . The board can undertake costly action $a_t \in [-1, 1]$ to alter the level of the policy variable; a negative action decreases X_t while a positive action increases X_t . The policy evolves according to

$$dX_t = X_t(2 - X_t) [a_t + \theta(d - X_t)] dt + X_t(2 - X_t)dZ_t$$

where θ captures the persistence of past effort. Effort has the largest impact on X_t for intermediate values of the policy; the policy is also most volatile at intermediate levels. Note that the process has two absorbing states, $\tilde{X} \in \{0, 2\}$

Constituents choose action each period, which represents their campaign contributions or support for the board. I omit specifying an underlying utility function for constituents, and represent the reduced form of the aggregate best response as

$$\bar{b}(a, X) = 1 + \lambda a^2 - (1 - X)^2$$

They pledge higher support when the policy is closer to the target and when the board exerts higher effort, where λ captures the marginal value of effort.

The board has no direct preference over the policy target; its payoffs are increasing in the support it receives from the constituents, and decreasing in effort.

$$g(a, \bar{b}, X) = \bar{b} - ca^2$$

In the unique static Nash equilibrium, the board chooses not to intervene, $a^N = 0$, and constituents pledge support for the board proportional to the difference between the target and current policy level, $b^N = 1 - (1 - X)^2$. This yields a stage game Nash equilibrium payoff of

$$v(X) = 1 - (1 - X)^2$$

for the board. Similar to the quality example, the board faces moral hazard when tasked with achieving its' constituents' policy target. In the next section, I will see whether persistent actions can allow the board to commit to a higher level of effort.

3 Equilibrium Analysis

This section presents the main results of the paper. First, I construct a Markov equilibrium, which simultaneously establishes existence and characterizes equilibrium behavior and payoffs. I show that the long run player's payoff in any PPE is bound above and below by the payoff in the best and worst Markov equilibria; therefore, the PPE payoff set is the convex hull of the payoffs in the Markov equilibria I construct. Next, I establish conditions for a Markov equilibrium to be the unique equilibrium in the class of all Perfect Public Equilibria. I return to the examples introduced in Section 2 to illustrate these results. The section concludes with an application of the existence and uniqueness results to a stochastic game *without* action persistence.

3.1 Existence of Markov Perfect Equilibria

The first main result of the paper establishes the existence of a Markov equilibrium, in which actions and payoffs are specified as a function of the state variable. The existence proof is constructive, and characterizes the explicit form of equilibrium continuation values and actions in Markov equilibria. This result applies to a general setting in which the state space may be bounded or unbounded, and there may or may not be absorbing states within the state space.

Theorem 1. *Suppose Assumptions 1, 2 and 3 hold. Given any initial state X_0 and action profile $(a^*, \bar{b}^*) = S(X, 0, U'(X))$, iff $U(X)$ is a bounded solution to the optimality equation:*

$$U''(X) = \frac{2r \left(U(X) - g(a^*, \bar{b}^*, X) \right)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} - \frac{2\mu_x(a^*, \bar{b}^*, X)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} U'(X)$$

then $U(X)$ is a Markov equilibrium. This equilibrium is characterized by:

1. *Equilibrium payoffs $U(X_0)$*
2. *Continuation values $(W_t)_{t \geq 0} = (U(X_t))_{t \geq 0}$*

3. Equilibrium actions $(a_t^*, \bar{b}_t^*)_{t \geq 0}$ uniquely specified as a function of $(X_t, U'(X_t))$:

$$S(X_t, 0, U'(X_t)) = \left\{ \begin{array}{l} a_t^* = \arg \max_{a'} rg(a', \bar{b}_t^*, X_t) + U'(X_t)\mu_x(a', \bar{b}_t^*, X_t) \\ \bar{b}_t^* = \bar{B}(a_t^*, X_t) \end{array} \right\}$$

The optimality equation has at least one solution $U \in C^2(\mathbb{R})$ that lies in the range of feasible payoffs for the long-run player $U(X) \in [\underline{g}, \bar{g}]$ for all states $X \in \Xi$. Thus, there exists at least one Markov equilibrium.

Theorem 1 shows that the stochastic game has at least one Markov equilibrium. Continuation values in this equilibrium are characterized by the solution to a second order ordinary differential equation, which specifies continuation payoffs as a function of the state variable. Rearranging the optimality equation as:

$$U(X) = g(a^*, \bar{b}^*, X) + \frac{1}{r}\mu_x(a^*, \bar{b}^*, X)U'(X) + \frac{1}{2r}U''(X) \left[\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X) \right]$$

lends insight into the relationship between the continuation value and the transition of state variable. The continuation value is equal to the sum of the equilibrium flow payoff today, $g(a^*, \bar{b}^*, X)$, and the expected change in the continuation value, weighted by the discount rate. The second term captures how the continuation value changes with respect to the equilibrium drift of the state variable. For example, if the state variable has positive drift ($\mu_x > 0$), and the continuation value is increasing in the state variable ($U' > 0$), then this increases the expected change in the continuation value. The third term captures how the continuation value changes with respect to the equilibrium volatility of the state variable. If U is concave ($U'' < 0$), it is more sensitive to negative shocks than positive shocks. Positive and negative shocks are equally likely, and therefore, the continuation value is decreasing in the volatility of the state variable. If U is linear ($U'' = 0$), then the continuation value is equally sensitive to positive and negative shocks, and the volatility of the state variable does not impact the continuation value. Now consider a value of the state variable that yields a local maximum $U(X^*)$ (note this implies $U' = 0$). Since the continuation value is at a local maximum, it must be decreasing as X moves away from X^* in either direction. This is captured by the fact that $U''(X) < 0$. Greater volatility of X or a more concave function U lead to a larger expected decrease in the continuation value.

The condition for sequential rationality takes an intuitive form. The current action affects future payoffs through its impact on the state variable, μ_x , scaled by the slope of the continuation value with respect to the state variable, $U'(X_t)$. The continuation value and equilibrium actions are independent of the public signal, as should be the case in a Markov equilibrium; this is born out mathematically by the condition $\beta_{yt} = 0$.

Following is an outline of the intuition behind the proof of Theorem 1. The first step in proving existence is to show that if a Markov equilibrium exists, then continuation values must be characterized by the solution to the optimality equation. In a Markov equilibrium, continuation values take the form of $W_t = U(X_t)$, for some function U . Using Ito's formula to differentiate $U(X_t)$ with respect to X_t yields an expression for the law of motion of the

continuation value in any Markov equilibrium $dW_t = dU(X_t)$, as a function of the law of motion for the state variable:

$$\begin{aligned} dU(X_t) &= U'(X_t)\mu_x(a, \bar{b}, X)dt \\ &+ \frac{1}{2}U''(X_t) \left[|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right] dt \\ &+ U'(X_t) \left[\sigma_{xy}(\bar{b}_t, X_t) dZ_t^y + \sigma_{xx}(\bar{b}_t, X_t) dZ_t^x \right] \end{aligned}$$

In order for this to be an equilibrium, continuation values must also follow the law of motion specified in Lemma 1, with drift

$$r \left(U(X_t) - g(a_t, \bar{b}_t, X_t) \right) dt$$

Matching the drifts of these two laws of motion yields the optimality equation.

The next step in the existence proof is to show that this ODE has at least one solution that lies in the range of feasible payoffs for the long-run player. Lipschitz continuity of the drift μ_x and flow payoff g ensures that they are bounded on any bounded interval of the state space, while Assumption 1 ensures the volatility of the state variable is positive on any bounded interval of the state space. These conditions are sufficient to guarantee the optimality equation meets the technical conditions required for the existence of a solution.

The final step is to show that the continuation value and actions characterized above do in fact constitute a Markov equilibrium. Lemma 1 establishes that the volatility of the continuation value must be

$$\beta_{yt} [\sigma_{yy} \cdot dZ_t^y + \sigma_{yx} dZ_t^x] + \beta_{xt} [\sigma_{xy} \cdot dZ_t^y + \sigma_{xx} dZ_t^x]$$

in any PPE. Combining this with the expression for the volatility for the continuation value as a function of the state variable, $dU(X_t)$, characterizes $(\beta_t)_{t \geq 0}$:

$$\begin{aligned} r\beta_{yt} &= 0 \\ r\beta_{xt} &= U'(X_t) \end{aligned}$$

This establishes the condition for sequential rationality, as a function of the state variable and the solution to the optimality equation. Given an action profile $(a^*, \bar{b}^*) = S^*(X, 0, U'(X))$, the state variable evolves uniquely and the continuation value $dU(X_t)$ satisfies the condition for a PPE established in Lemma 1.

Theorem 1 also establishes that *each* solution U to the optimality equation characterizes a single Markov equilibrium. This is a direct consequence of Assumption 3, which guarantees that equilibrium actions are a unique function of U . If there are multiple solutions to the optimality equation, then each solution characterizes a single Markov equilibrium. The formal proof of Theorem 1 is presented in the Appendix.

Markov equilibria have an intuitive appeal in stochastic games. Advantages of Markov equilibria include their simplicity and their dependence on payoff relevant variables to specify incentives. Theorem 1 yields a tractable expression that can be used to construct equilibrium behavior and payoffs in a Markov equilibrium. As such, this result provides a tool to

analyze equilibrium behavior in a broad range of applied settings. Once functional forms are specified for the long-run player’s payoffs and the transition function of the state variable, it is straightforward to use Theorem 1 to characterize the optimality equation and incentive constraint for the long-run player, as a function of the state variable. Numerical methods for ordinary differential equations can then be used to estimate a solution to the optimality equation and explicitly calculate equilibrium payoffs and actions. These calculations yield empirically testable predictions about equilibrium behavior. Note that numerical methods are used for estimation in an equilibrium that has been characterized analytically, and not to simulate an approximate equilibrium. This is an important distinction.

3.2 PPE Payoff Set

This section characterizes bounds on the PPE payoff set. First, I need an intermediate result to establish that the highest and lowest Markov equilibria are well-defined. Define $\Upsilon : \Xi \rightrightarrows \mathbb{R}$ as the set of bounded solutions to the optimality equation.

Lemma 2. *Suppose Assumptions 1, 2 and 3. For all $U, V \in \Upsilon$, either $U(X) > V(X)$ for all $X \in \Xi$ or $U(X) < V(X)$ for all $X \in \Xi$ or $U(X) = V(X)$ for all $X \in \Xi$. In other words, two distinct solutions never cross.*

Lemma 2 establishes the existence of a Markov equilibrium which yields the highest continuation value across all states, U_1 and a Markov equilibrium which yields the lowest continuation value across all states, U_2 . When the optimality equation has a unique bounded solution, then $U_1 = U_2$.

This leads to the second main result, which shows that the PPE payoff set is the convex hull of the Markov equilibrium correspondence. It is not possible for the long-run player to achieve a payoff above the highest Markov equilibrium payoff or below the lowest Markov equilibrium payoff in any PPE.

Theorem 2. *Suppose Assumptions 1, 2 and 3 hold and let Υ be the set of bounded solutions to the optimality equation. The correspondence of perfect public equilibrium payoffs $\xi : \Xi \rightrightarrows \mathbb{R}$ is characterized by the convex hull of the highest and lowest Markov equilibrium payoff, $co(U_1, U_2)$.*

The impossibility of achieving PPE payoffs above the highest Markov payoff yields insight into the role action persistence plays in generating intertemporal incentives. In a stochastic game with imperfect monitoring, intertemporal incentives can be generated through two potential channels: (1) conditioning future equilibrium play on the public signal and (2) how the current action impacts future feasible payoffs through the state variable. Equilibrium play in a Markov equilibrium is completely specified by the state variable – the public signal is ignored. As such, intertemporal incentives are generated solely from the second channel. When this Markov equilibrium yields the highest payoff, it precludes the existence of any equilibria that use the public signal to build stronger incentives. *As such, the ability to generate effective intertemporal incentives in a stochastic game with imperfect monitoring stems entirely from the effect of the current action on the set of future feasible payoffs through the state variable.*

This insight relates to equilibrium degeneracy results from the continuous time repeated games literature, including Faingold and Sannikov (2011), Sannikov and Skrzypacz (2010) and Fudenberg and Levine (2007), which show that Brownian information cannot be used to provide effective intertemporal incentives via value burning. In a standard repeated game between a long-run and short-run player, punishing all players when there is a low signal (value burning) is the only potential channel for intertemporal incentives.⁹ Since this is not an effective channel, it is impossible to generate incentives. Likewise, in a stochastic game without action persistence, there is no link between the long-run player’s action and future feasible payoffs, and incentives collapse. Therefore, the introduction of action persistence creates an essential channel for creating intertemporal incentives.

A stochastic game has an important feature that allows Brownian information to be used effectively to provide intertemporal incentives. As discussed following lemma 1, Brownian information must be used linearly. Suppose the long-run player’s continuation value is at its upper bound. Using Brownian information linearly in a direction that is non-tangential to the boundary of the equilibrium payoff set will result in the continuation value escaping its upper bound with positive probability, a contradiction. Using Brownian information linearly and tangentially is precluded by the presence of myopic short-run players. Thus, it is not possible to linearly use Brownian information to structure incentives at the long-run player’s highest continuation payoff, and both the long-run player and short-run player will play a myopic best response at this point. But this is precisely the definition of a static Nash equilibrium, and therefore long-run player’s highest continuation payoff is bounded above by the highest static Nash equilibrium payoff. In a repeated game with a unique static Nash, this implies that the static Nash payoff is also the unique PPE payoff. However, in a stochastic game, this implies that the static Nash payoff is the unique PPE continuation value at the state that yields the highest continuation value. At other values of the state variable, it is possible for the continuation value to exceed the static Nash payoff.

Mathematically speaking, the volatility of the continuation value must be zero at its upper bound ($\beta \cdot \sigma(b, X) = 0$) to prevent it from escaping. In a stochastic game, it is possible to characterize β in a manner that depends on the state variable – the continuation value can have zero volatility at states that yield the highest continuation payoff and positive volatility at other states. Recall the characterization of β in Theorem 1:

$$\begin{aligned}\beta_y &= 0 \\ \beta_x &= U'(X)\end{aligned}$$

If an interior state X^* yields the highest continuation value, then $U'(X^*) = 0$, which ensures the volatility of the continuation value is zero at its upper bound. If a boundary of the state space yields the highest continuation value, then $\sigma(b, X) = 0$ and the continuation value has zero volatility at its upper bound. The long-run player must be playing a myopic best response at the state that yields the highest continuation payoff. However, for other states, it is possible for $|\beta_x| > 0$ and one can structure incentives so that the long-run player plays a non-myopic action. This is not possible in a repeated game – trivially, a repeated game is

⁹Conditioning future equilibrium play on the public signal by tangentially transferring continuation values is not possible in a game between a long-run and short-run player.

a stochastic game with a single state, and so the volatility of the continuation value must always be zero.

3.3 Uniqueness of Markov Equilibrium

The final main result establishes conditions under which there is a unique Markov equilibrium, which is also the *unique equilibrium* in the class of *all Perfect Public Equilibria*. The main step of this result is to establish when the optimality equation has a unique bounded solution. Any Markov equilibrium must have continuation values and equilibrium actions specified by a solution to the optimality equation. Thus, when the optimality equation has a unique solution, there is a unique Markov equilibrium.

The optimality equation will have a unique solution when its solution satisfies certain boundary conditions. These boundary conditions depend on the limiting behavior of the flow payoffs and state variable; Assumption 4 outlines sufficient conditions on g and μ_x to guarantee uniqueness.

Assumption 4. 1. *If the state space is unbounded, $\Xi = \mathbb{R}$, then, then there exists a δ such that for $|X| > \delta$, $v(X)$ is monotonic in X .*

2. *If the state space is bounded, $\Xi = [\underline{X}, \overline{X}]$, then the boundary points are absorbing states: $\mu_x(a, b, \overline{X}) = \mu_x(a, b, \underline{X}) = 0$ for all $(a, b) \in A \times B$.*

Namely, the flow payoff must have a well-defined limit and the impact of the firm's action on its continuation value must converge to zero. When the state space is bounded, Lipschitz continuity guarantees that $v(X)$ has a well-defined limit at \underline{X} and \overline{X} . If the state space is unbounded, an additional assumption is necessary to guarantee that $v(X)$ doesn't perpetually oscillate. Additionally, when the state space is bounded, the upper and lower bounds of the state space must be absorbing points.

Remark 4. *When the endpoints of the state space are absorbing points, whether the state variable actually converges to one of its absorbing points with positive probability will depend on the relationship between the drift and the volatility as the state variable approaches its boundary points. It is possible that the state variable converges to an absorbing point with probability zero.*

The following theorem establishes the uniqueness of a Markov equilibrium in the class of all Perfect Public Equilibria.

Theorem 3. *Suppose Assumptions 1, 2, 3 and 4 hold. Then, for each initial value of the state variable $X_0 \in \Xi$, there exists a unique perfect public equilibrium, which is Markov. The unique bounded solution U of the optimality equation characterizes the equilibrium payoff $U(X_0)$ and continuation values.*

1. *When the state space is unbounded, $\Xi = \mathbb{R}$, then the solution satisfies the following boundary conditions:*

$$\begin{aligned} \lim_{X \rightarrow \infty} U(X) &= v_\infty \text{ and } \lim_{X \rightarrow -\infty} U(X) = v_{-\infty} \\ \lim_{X \rightarrow \infty} \mu_x(a^*, \bar{b}^*, X)U'(X) &= \lim_{X \rightarrow -\infty} \mu_x(a^*, \bar{b}^*, X)U'(X) = 0 \end{aligned}$$

where $v_\infty = \lim_{X \rightarrow \infty} v(X)$ and $v_{-\infty} = \lim_{X \rightarrow -\infty} v(X)$.

2. When the state space is bounded, $\Xi = [\underline{X}, \overline{X}]$, then the solution satisfies the following boundary conditions:

$$\begin{aligned} \lim_{X \rightarrow \overline{X}} U(X) &= v(\overline{X}) \text{ and } \lim_{X \rightarrow \underline{X}} U(X) = v(\underline{X}) \\ \lim_{X \rightarrow \overline{X}} \mu_x(a^*, \overline{b}^*, X)U'(X) &= \lim_{X \rightarrow \underline{X}} \mu_x(a^*, \overline{b}^*, X)U'(X) = 0 \end{aligned}$$

The boundary conditions characterized in Theorem 3 have several implications for equilibrium play and payoffs. Recall the incentive constraint for the long-run player from Theorem 1. The link between the long-run player's action and future payoffs is proportional to the slope of the continuation value and the drift of the state variable. In the limit, $\mu_x(a^*, \overline{b}^*, X)U'(X)$ converges to zero and the long-run player's incentive constraint is reduced to maximizing its instantaneous flow payoff. Therefore, the long-run player's equilibrium action converges to the static Nash action. Additionally, the continuation value converges to the limit of the static Nash payoff. This happens in a bounded state space because the boundary points are absorbing states. In an unbounded state space, the drift of the state variable can't pull the state back to a region with non-negligible incentives fast enough when the state variable becomes very large.

The proof establishing that the optimality equation has a unique solution has two parts: (i) show that any solution to the optimality equation must satisfy the same boundary conditions, and (ii) show that it is not possible for two different solutions to the optimality equation to satisfy the same boundary conditions. Suppose the state space is unbounded and U is a bounded solution to the optimality equation. By Assumption 4 and the boundedness of g , the static Nash payoff function v has a well-defined limit. This ensures that U also has a well-defined limit and that U' converges to zero at rate $O(1/X)$.¹⁰ When μ_x is Lipschitz continuous, it grows at or slower than linearly, and $U'(X)\mu_x(a^*, b^*, X)$ converges to zero, establishing one of the boundary conditions. Given that $U'(X)\mu_x(a^*, b^*, X)$ converges to zero, the long-run player's action converges to a myopic best response, which yields a flow payoff equal to the static Nash equilibrium payoff, $\lim_{X \rightarrow \infty} g(a^*, \overline{b}^*, X) = v_\infty$. The boundedness of U also ensures that the second derivative, U'' , doesn't converge to a constant.¹¹ From the optimality equation,

$$\begin{aligned} \lim_{X \rightarrow \infty} & \left[\left| \sigma_{xy}(\overline{b}^*, X) \right|^2 + \sigma_{xx}^2(\overline{b}^*, X) \right] U''(X) \\ &= \lim_{X \rightarrow \infty} 2r \left[U(X) - g(a^*, \overline{b}^*, X) \right] - 2\mu_x(a^*, \overline{b}^*, X)U'(X) \\ &= 2r \left[\lim_{X \rightarrow \infty} U(X) - v_\infty \right] \end{aligned}$$

¹⁰The monotonicity assumption on $v(X)$, plays a key role in ensuring the limit of U' exists, as it is possible for a bounded function to converge to a finite limit, but have a derivative that oscillates. This assumption guarantees that U is monotonic for large X , and prevents U' from oscillating. A similar assumption is not necessary in the bounded state space case, as the Lipschitz continuity of v is sufficient to ensure the limit of U' is well-defined.

¹¹The boundedness of U ensures that U'' either converges to zero, or oscillates around zero.

Since $|\sigma_{xy}|^2 + \sigma_{xx}^2 > 0$ and U'' doesn't converge to a constant, it must be that U converges to v_∞ so that the right hand side converges to zero. This establishes the remaining boundary condition, $\lim_{X \rightarrow \infty} U(X) = v_\infty$. Showing that it is not possible for two different solutions U_1 and U_2 to both satisfy these boundary conditions concludes the proof. The case of a bounded state space is similar.

3.4 Interlude: Return to Examples

I return to the examples introduced in Section to apply the results of Theorems 1 and 3.

3.4.1 Persistent Effort as a Source of Reputation

Recall I am interested in determining whether persistent effort can provide a firm with endogenous incentives to build its quality. This product choice game can be viewed as a stochastic game with stock quality X as the state variable and the change in stock quality dX as the transition function, which depends on the firm's action. Let $U(X)$ represent the continuation value of the firm in a Markov perfect equilibrium when $X_t = X$. Then, given equilibrium action profile $(a(X), \bar{b}(X))$, the continuation value can be expressed as an ordinary differential equation,

$$rU(X) = r(\bar{b}(X) - ca(X)^2) + \theta(a(X) - X)U'(X) + \frac{1}{2}\sigma^2U''(X)$$

where the first term, $\bar{b} - ca^2$, is the payoff that the firm earns today, and the second term, $\theta(a - X)U'(X) + \frac{1}{2}\sigma^2U''(X)$, is the expected change in the continuation value. The firm's payoff is increasing in stock quality ($U' > 0$). Therefore, the expected change in the continuation value is increasing with the drift of the stock, $\theta(a - X)$.

The volatility of the stock determines how the shape of the continuation value relates to its expected change. If the value of quality is concave ($U'' < 0$), the firm is "risk averse" in quality. It is more sensitive to negative quality shocks than positive shocks. With Brownian noise, positive and negative shocks are equally likely, so volatility hurts the firm. On the other hand, if the value of quality is convex, then the firm benefits more from positive shocks than it is hurt by negative shocks and volatility is beneficial. For example, if quality is research and there is a big payoff once a certain threshold is crossed, then volatility is good for the firm.

Figure 1 illustrates $U(X)$ for several discount rates. There is an interesting nonmonotonicity of the continuation payoff with respect to the discount rate, which is driven by two competing factors. A firm with a low discount rate places a greater weight on the future, which gives it a stronger incentive to choose high effort today and build up its quality. On the other hand, a low discount rate means that transitory positive shocks to quality have a lower value relative to the long-run expected quality. When stock quality is low, the first effect dominates and low discount rates yield higher payoffs; this relationship flips when the stock quality becomes large.

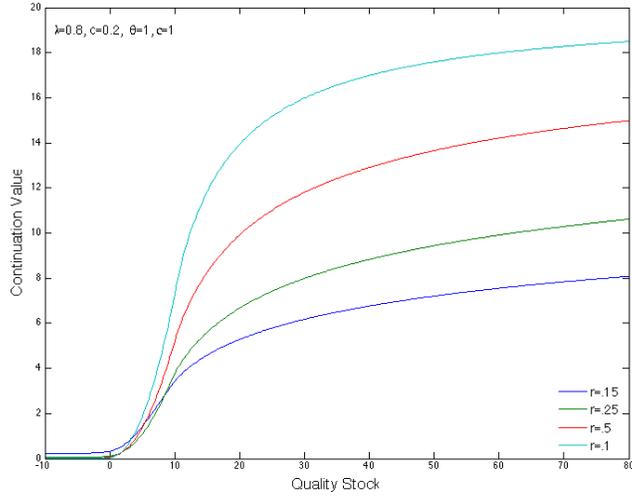


FIGURE 1. Equilibrium Payoffs

In equilibrium, consumers myopically optimize flow payoffs by choosing a purchase level such that the marginal utility of an additional unit of product is zero:

$$\bar{b}(a(X), X) = \begin{cases} 0 & \text{if } q(a(X), X) \leq 0 \\ \min \left\{ \frac{1}{4}q(a(X), X)^2, 20 \right\} & \text{if } q(a(X), X) > 0 \end{cases}$$

They are willing to purchase more when they expect higher quality.

The firm faces a trade-off when choosing its investment level: the cost of investment is borne in the current period, but yields a benefit in future periods through higher purchase levels by consumers. The slope of the continuation value captures the impact of effort on future payoffs. In equilibrium, effort is chosen to equate the marginal cost of effort with its expected marginal benefit,

$$a(X) = \min \left\{ \frac{\theta}{2cr} U'(X), \bar{a} \right\}.$$

In a Markov equilibrium, the expected marginal benefit depends on the sensitivity of the continuation value to changes in stock quality, $U'(X)$, and the marginal impact of current effort on the stock quality, θ . When the continuation value is more sensitive to changes in stock quality (captured by a steeper slope) or current effort has a larger immediate impact on quality (captured by higher θ) or patience rises, the firm chooses a higher effort level. It is interesting to note the trade-off between persistence and the discount rate. Only the ratio of these two parameters is relevant for determining investment and the continuation value; therefore, doubling θ has the same impact as halving r . As θ approaches 0, stock quality is almost entirely determined by its initial level and the intertemporal link between effort and quality is very small.

Figures 3 and 2 graph equilibrium actions for the firm and consumers, respectively. The firm's effort level peaks in a region of negative drift in the figure. When the firm receives

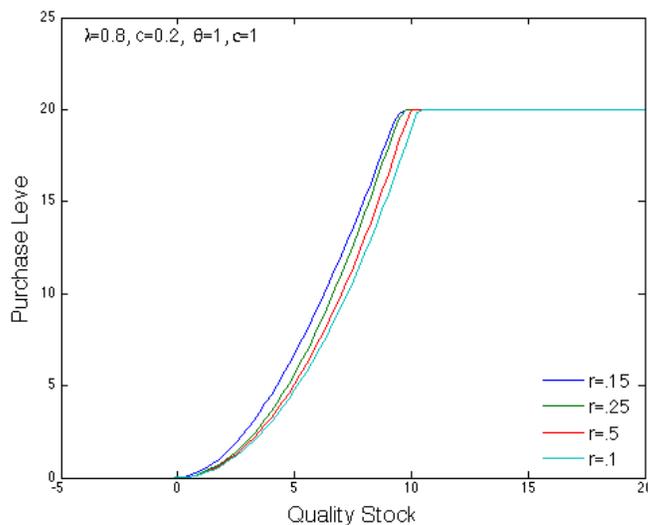


FIGURE 2. Consumer Equilibrium Behavior

a positive quality shock, it exerts effort to maintain this high quality, but also enjoys high payoffs today at the expense of allowing the quality to slowly drift down. Note that these action correspondences satisfy Assumption 3.

The firm has the strongest incentive to invest at intermediate quality levels - a “reputation building” phase characterized by high effort and rising quality. The slope of the continuation value converges to 0 as quality becomes very high or low; the firm’s continuation value is not very sensitive to changes in quality and the firm has weak incentives to choose high effort. When quality is very high, the firm in effect “rides” its good reputation by choosing low effort and allowing quality to drift downward. Very negative shocks lead to periods where the firm chooses low effort and allows quality to recover - “reputation recovery” - before beginning to rebuild.

These reputation effects are present in the long-run.¹² Product quality is cyclical, and periods of lower quality are followed by periods of higher quality.¹³ This contrasts with models in which reputation is derived from behavioral types: as Cripps et al. (2007) and Faingold and Sannikov (2011) show, reputation effects are temporary insofar as consumers eventually learn the firm’s type, and so asymptotically, a firm’s incentives to build a reputation disappear. Additionally, conditional on the firm being strategic, reputation in types models has negative drift.

Figure 4 illustrates the equilibrium dynamics of the stock quality. When the equilibrium drift is positive, stock quality increases in expectation, whereas the opposite holds when

¹²Mathematically, reputation effects are present in the long-run when there is positive probability that the state variable doesn’t converge to an absorbing state. In this example, there are no absorbing states, and the requirement is satisfied trivially.

¹³One could also model a situation in which the firm exits once quality hits a lower bound \underline{X} by making \underline{X} an absorbing state and setting the state space equal to $\Xi = [\underline{X}, \infty)$. This would not necessarily create short-term reputation effects; whether quality converges to \underline{X} depends on the equilibrium dynamics of X_t .

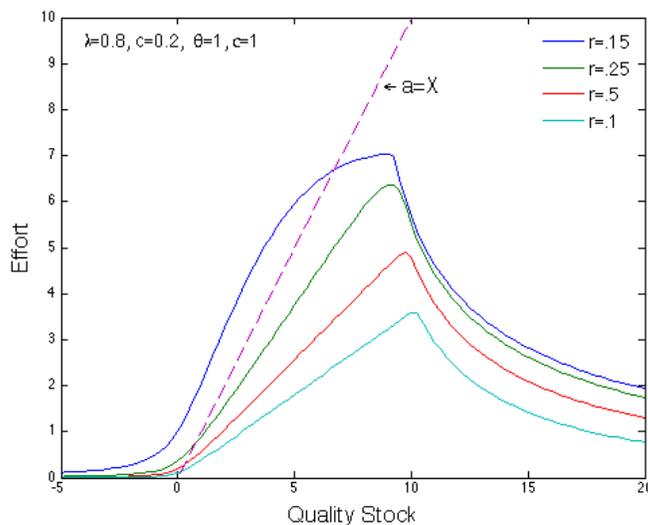


FIGURE 3. Firm Equilibrium Behavior

the equilibrium drift is negative. Quality is stable when effort exactly offsets decay, or mathematically, when the drift is zero. As the firm becomes more patient, a higher level of quality is stable.

The Markov equilibrium described above is the unique Perfect Public Equilibrium, by Theorem 3. When the state space is unbounded, a sufficient condition for uniqueness is that the static Nash payoff is monotonic in quality above a certain quality level. In this example, the static Nash payoff of the firm depends on the consumers' purchase levels, which is determined by the value of the stock:

$$v(X) = \begin{cases} 0 & \text{if } X \leq 0 \\ \min \left\{ \frac{1}{4}(\lambda X)^2, 20 \right\} & \text{if } X > 0 \end{cases}$$

This payoff is monotonically increasing in X .

Next I compare the firm's payoffs in the game with action persistence to the game without action persistence (this corresponds to $\theta = 0$). When effort is transitory, the unique equilibrium of the stochastic game is to choose zero effort each period, which yields an equilibrium payoff of

$$V(X_0) = E_0 \left[r \int_0^\infty e^{-rs} v(X_s) ds \right]$$

Persistence enhances the firm's payoffs through two complimentary channels. First, the firm chooses an effort level that equates the marginal cost of investment today with the marginal future benefit. In order for the firm to be willing to choose a positive level of effort, the future benefit of doing so must exceed the future benefit of choosing zero effort. Second, the link with future payoffs allows the firm to commit to a positive level of effort in the current period, which increases the equilibrium purchase level of consumers in the current period. I

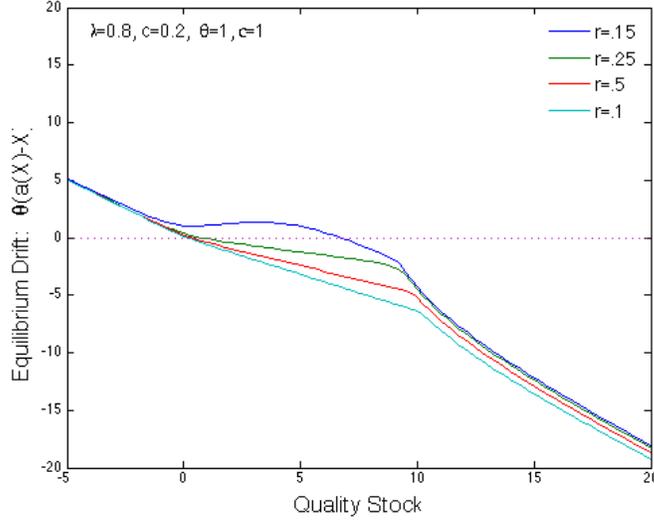


FIGURE 4. Equilibrium Drift of Stock Quality

will show in Theorem 5 that the firm achieves higher equilibrium payoffs when its actions are persistent, $U(X_0) \geq V(X_0)$ for all X_0 . Therefore, action persistence creates nontrivial intertemporal incentives for the firm to exert effort.

3.4.2 Building a Specialization with Oscillating Demand

The expressions for equilibrium actions and continuation values in this variation mirrors those characterized above, with the exception of the equilibrium demand curve,

$$\bar{b}(a(X), X) = \begin{cases} 0 & \text{if } q(a(X), X) \leq 0 \\ \min(\sin 2q(a, X) + q/2), 8 & \text{if } q(a(X), X) > 0 \end{cases}$$

However, the equilibrium dynamics are quite different. Investment oscillates as a function of quality. When a firm is in a region where demand is increasing with quality, the firm continues to build its quality. Once it crosses over to a region of decreasing demand, it slacks off for a while and lets its quality decay. Overall, the value of quality trends upward. However, at low levels of quality, this value oscillates. A low quality firm may be better off remaining a low quality firm, rather than trying and failing to move up market. Figures 5 and 6 illustrate equilibrium behavior.

3.4.3 Policy Targeting

This example seeks to determine when persistent actions create incentives for a governing board to implement its' constituents policy target. Let $U(X)$ represent the continuation value in a Markov equilibrium. From Section 3, the continuation value can be represented as

$$U(X) = r(\bar{b} - ca^2) + X(2 - X)[a + \theta(d - X)]U'(X) + \frac{1}{2}X(2 - X)U''(X)$$

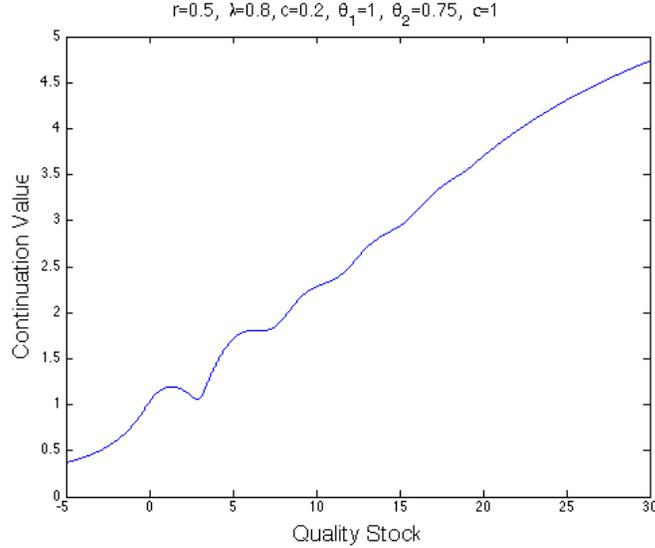


FIGURE 5. Equilibrium Continuation Value

which is plotted in Figure 7.

The continuation value is concave in the policy level, and skewed in the direction of the natural drift. The natural drift d increases the board's payoff when it pushes the policy variable in the direction of the target; namely, when $d - X$ and U' have the same sign. If the natural drift lies far below the target, then at high levels, the policy will naturally move toward the target, which benefits the board. This skewness is illustrated in Figure 7 for several levels of d .

Volatility hurts the board, given that the continuation value is concave. The policy variable is most volatile at intermediate values; this puts a damper on the continuation value at the target. In fact, the highest Markov payoff, which occurs at the target, is strictly less than the static Nash at $X = 1$, $U(1) < v(1) = 1$.

The board's incentives are driven by the slope of the continuation value and the impact of its efforts on the policy level,

$$a_t(X) = \frac{X_t(2 - X_t)}{2rc} U'(X_t)$$

When the current policy level is very far from its optimal target, the board's effort has a smaller impact on the policy level, and the board has a lower incentive to undertake costly effort. When the policy level is close to its target, the continuation value approaches its maximum, the slope of the continuation value approaches zero and the board also has a lower incentive to exert effort. Therefore, the board has the strongest incentives to exert effort when the policy variable is an intermediate distance from its target. Figure 8 plots the equilibrium effort choice of the board, and Figure 9 shows the equilibrium constituent support.

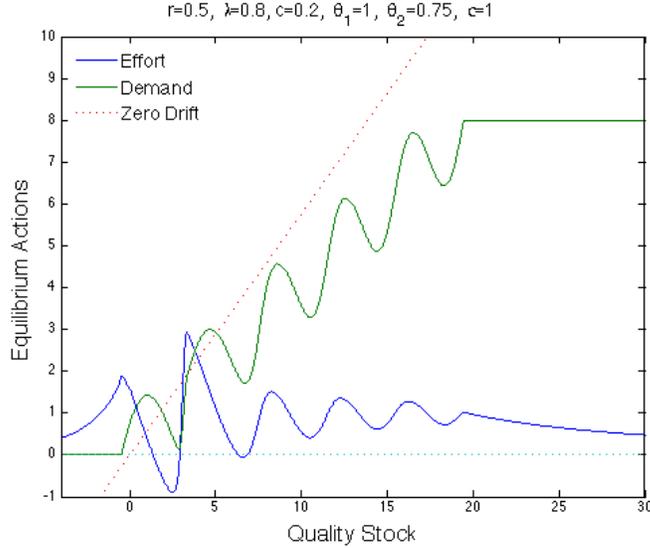


FIGURE 6. Equilibrium Actions

3.5 Stochastic Games without Action Persistence

Suppose that the state variable evolves independently of the long-run player's and there is a stochastic game with no action persistence. This section establishes that the unique PPE in this setting is one in which the long-run player acts myopically and plays the static Nash equilibrium in each state.

Given an initial state X_0 , define the average discounted payoff from playing the static Nash equilibrium action profile $(S^*(X_t, 0, 0))_{t \geq 0}$ in each state as:

$$V_{NE}(X_0) = E \left[r \int_0^{\infty} e^{-rt} v(X_t) dt \right]$$

and the expected continuation payoff from playing a static Nash equilibrium action profile as:

$$W_{NE}(X_t) = E_t \left[r \int_t^{\infty} e^{-rs} v(X_s) dt \right]$$

where these expectations are taken with respect to the state variable, given that the state evolves according to the measure generated by $(S^*(X_t, 0, 0))_{t \geq 0}$. This expression defines the stochastic game payoff that the long-run player will earn if it myopically optimizes flow payoffs each instant. Note that repeated play of the static Nash action profile is not necessarily an equilibrium strategy profile of the stochastic game.¹⁴

The following Corollary establishes that without action persistence, repeated play of the static Nash action profile is an equilibrium strategy profile; in fact, it is the unique equilibrium strategy profile. The corollary also directly characterizes equilibrium payoff $V_{NE}(X_0)$ as the solution to the optimality equation.

¹⁴This is a general property of stochastic games.

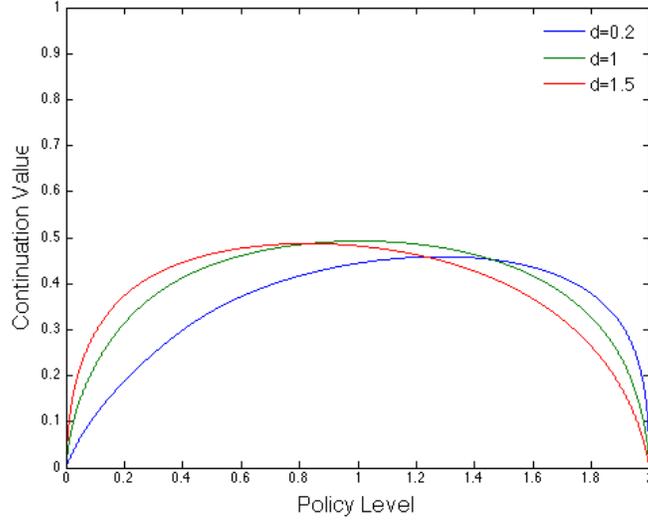


FIGURE 7. Equilibrium Payoffs

Corollary 1. *Suppose that the drift of the state variable is independent of the long-run player's action for all X , and suppose Assumptions 1, 2 and 3 hold. Then, given an initial state X_0 , there is a unique perfect public equilibrium characterized by the unique bounded $U(X)$ solution to the optimality equation with:*

1. *Equilibrium payoff $U(X_0)$*
2. *Continuation values $(W_t)_{t \geq 0} = (U(X_t))_{t \geq 0}$*
3. *Equilibrium actions $(a_t^*, \bar{b}_t^*)_{t \geq 0}$ uniquely specified by the static Nash equilibrium action profile, for each X :*

$$S^*(X, 0, 0) = \left\{ \begin{array}{l} a = \arg \max_{a'} g(a', \bar{b}, X) \\ \bar{b} = \bar{B}(a, X) \end{array} \right\}$$

Additionally, this equilibrium payoff and the continuation values correspond to the expected payoff from playing the static Nash equilibrium action profile, $U(X_0) = V_{NE}(X_0)$ and $(U(X_t))_{t \geq 0} = (W_{NE}(X_t))_{t \geq 0}$.

This result is a direct application of Theorem 1. In the absence of a link between the long-run player's action and the state variable, the incentive constraint is independent of the continuation value and the long-run player plays a static best response.

Uniqueness in a game without action persistence stems directly from the existence characterization, and does not require Assumptions 4. The incentive constraint is independent of the solution U to the optimality equation, so all solutions to the optimality equation yields the same equilibrium action profile. When a static best response is played each period, the

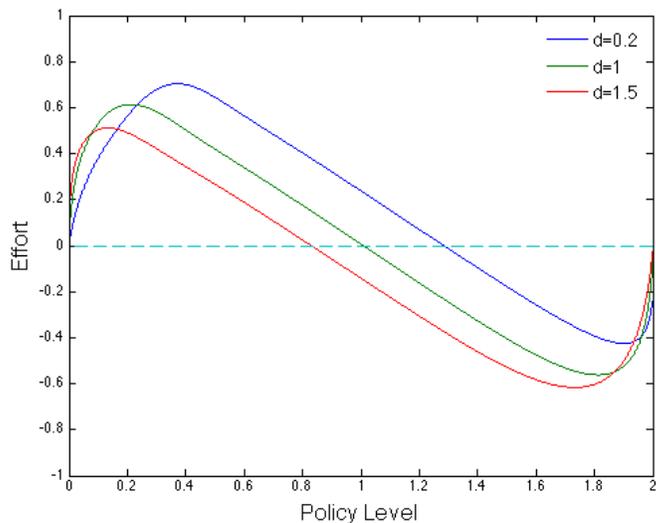


FIGURE 8. Government Equilibrium Behavior

continuation value in any Markov equilibrium evolves according to the expected payoff from playing the static Nash equilibrium action profile in each state:

$$\begin{aligned}
 U(X_t) &= E_t \left[r \int_t^\infty e^{-rs} v(X_s) ds \right] \\
 &= W_{NE}(X_t)
 \end{aligned}$$

where the expectation is taken with respect to the measure over the state variable. Given that this measure over the state variable is the same in any Markov equilibrium, any solution to the optimality equation must yield the same continuation values $(U(X_t))_{t \geq 0} = (W_{NE}(X_t))_{t \geq 0}$. Therefore, this solution must be unique. The solution to the optimality equation in Corollary 1 can be used to explicitly characterize the expectation $V_{NE}(X_0)$ and $W_{NE}(X_t)$.

4 Properties of Equilibrium Payoffs

This section uses describes several interesting properties relating the long-run player's equilibrium payoffs to the structure of the underlying stage game. The main results of this section are to provide an upper and lower bound on the PPE payoffs of the stochastic game across all states and characterize how the PPE payoff of the long-run player varies with the state variable

First, I examine properties of the highest and lowest PPE payoffs across all states, represented by \bar{W} and \underline{W} , respectively. Incentives for long-run player to choose a non-myopic action in the current period are provided through the link between the long-run player's action, the transition of the state variable and future feasible payoffs. When the continuation value is at its upper or lower bound, then the continuation value must have zero volatility

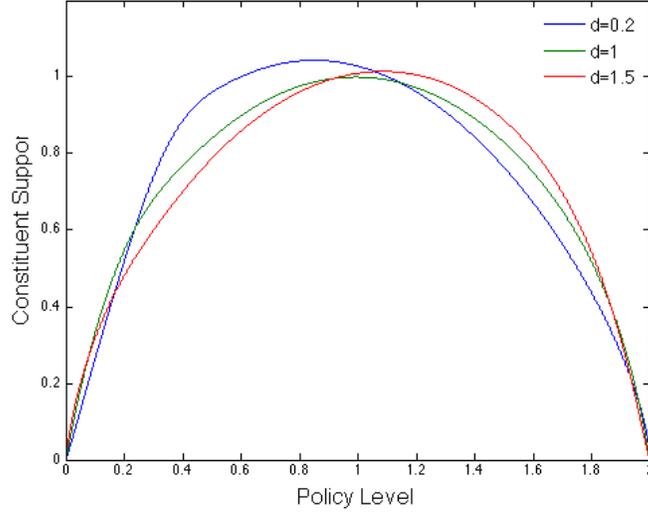


FIGURE 9. Constituent Equilibrium Behavior

so as not to escape its bound. The volatility of the continuation value is proportional to β_t , which also determines the incentive constraint for the long-run player. If the continuation value doesn't respond to the public signal, then the long-run player will myopically best respond by choosing the action that maximizes current flow payoffs. Therefore, the action profile at the set of states that yield the highest and lowest PPE payoffs across all states must be a Nash equilibrium of the stage game at that state.

At its upper bound, the drift of the continuation value must be negative. Using the law of motion for the continuation value characterized above, this means that the current flow payoff must exceed the continuation value. The current flow payoff in any stage Nash equilibrium is bounded above by \bar{v}^* , the highest stage Nash equilibrium payoff across all states. Thus, the highest PPE payoff across all states is bounded above by the highest static Nash equilibrium payoff across all states. Similar reasoning applies to showing that the lowest PPE payoff across all states is bounded below by the lowest static Nash equilibrium payoff.

Theorem 4. $\bar{W} \leq \bar{v}^*$ and $\underline{W} \geq \underline{v}^*$

If there is an absorbing state X in the set of states that yields the highest stage Nash equilibrium payoff, then it is possible to remain in this state once it is reached. Thus, the highest PPE payoff across all states arises from repeated play of this stage Nash equilibrium, and yields a continuation value of $\bar{W} = \bar{v}^*$. By construction, this continuation value occurs at the state X that yields the highest static payoff. An analogous result holds for the lowest PPE payoff.

This result has an intuitive relation to reputation models of incomplete information. Recall that the state variable in this model is the short-run players' beliefs about the long-run player's type. Therefore, $X = 0$ and $X = 1$ are absorbing states. When $X = 0$,

short-run players place probability one on the long-run player being a normal type, and it is not possible for the long-run player to earn payoffs above the static Nash equilibrium payoff of the game with complete information. Provided the long-run player's payoffs are increasing in the belief it is a behavioral type, the lowest PPE payoff occurs at $X = 0$ and is equal to the Nash equilibrium payoff of the complete information game, v^* . Conditional on the long-run player being a normal type, the transition function governing beliefs has negative drift, and beliefs converge to the absorbing state $X = 0$. This captures the temporary reputation phenomenon associated with reputation models of incomplete information. Once short-run players learn the long-run player's type, it is not possible to return to a state $X > 0$. Note that although $X = 1$ is also an absorbing state, but conditional on the long-run player being the normal type, the state variable never converges to $X = 1$.

In the current model, if either endpoint is an absorbing state, *and* the state variable converges to this endpoint, then the intertemporal incentives created by the stochastic game will be temporary. Once this absorbing state is reached, the dynamic game is reduced to a standard repeated game and the unique equilibrium involves repeated play of the static Nash equilibrium. On the other hand, if neither endpoint is an absorbing state, or if the state variable doesn't converge to its absorbing states with positive probability, then the intertemporal incentives created by the stochastic game are permanent. As noted in the above discussion, it is possible to have an absorbing state that the state variable converges to with probability zero.

The second main result on PPE payoffs relates how the continuation value of the long-run player changes with the state variable to how the stage Nash equilibrium of the underlying stage game varies with the state variable.

Theorem 5. *Assume 1, 2, 3 and 4. The following properties show how PPE payoffs change with the state variable:*

1. *Suppose $v(X)$ is increasing (decreasing) in X for all $X \in \Xi$. Then $U(X)$ is also increasing (decreasing) in X for all $X \in \Xi$. The states that yield the highest and lowest PPE payoffs are boundary points.*
2. *Suppose $v(X)$ is concave and has a unique interior maximum. Then $U(X)$ is concave and has a unique interior maximum.*
3. *Suppose $v(X)$ is convex and has a unique interior minimum X^* . Then $U(X)$ is convex and has a unique interior minimum.*

If the stage Nash equilibrium payoff of the long-run player is increasing in the state variable, then the PPE payoff to the long-run player is also increasing in the the state variable. The state that yields the highest PPE payoff to the long-run player corresponds to the state that yields the highest stage Nash equilibrium payoff, and the state that yields the lowest PPE payoff to the long-run player corresponds to the state that yields the lowest PPE payoff. Stochastic games differ from standard repeated games in that it is not necessarily possible to achieve an equilibrium payoff of the stochastic game that is equal to the stage Nash equilibrium payoff. Thus, for intermediate values of the state variable, the PPE payoff

may lie above or below the static Nash equilibrium payoff. A symmetric result holds if the stage Nash equilibrium payoff of the long-run player is decreasing in the state variable.

If the stage Nash equilibrium payoff of the long-run player is concave in the state variable, then the PPE payoff will be increasing over the region that the stage Nash equilibrium payoff is increasing, and decreasing over the region that the stage Nash equilibrium payoff is decreasing. The state that yields the lowest PPE payoff will occur at either endpoint of the state space. If the endpoint that yields the lowest stage Nash equilibrium payoff is an absorbing state, then this state also yields the lowest PPE payoff and $\underline{W} = \underline{v}^*$. Otherwise, the endpoint that yields the lowest stage Nash equilibrium payoff will depend on the transition function. A symmetric result holds if the stage Nash equilibrium payoff of the long-run player is convex in the state variable.

This result characterizes properties of the PPE payoff as a function of the state variable for several relevant classes of games. More generally, if the stage game Nash equilibrium payoff is not monotonic or single-peaked in the state variable, then the highest and lowest PPE payoffs of the stochastic game may not coincide with the states that yield the maximum or minimum stage game Nash equilibrium payoffs.

5 Conclusion

Persistence and rigidities are pervasive in economics. There are many situations in which a payoff-relevant stock variable is determined not only by actions chosen today, but also by the history of past actions. This paper shows that this realistic departure from a standard repeated game provides a new channel for intertemporal incentives. The long-run player realizes that the impact of the action it chooses today will continue to be felt tomorrow, and incorporates the future value of this action into its decision. Persistence is a particularly important source of intertemporal incentives in the class of games examined in this paper; in the absence of such persistence, the long-run player cannot earn payoffs higher than those earned by playing a myopic best response.

The main results of this paper are to establish conditions on the structure of the game that guarantee existence of Markov equilibria, and uniqueness of a perfect public equilibria, which is Markov. Markov equilibria have attractive features for use in applied work. These results not only provide a theoretical justification for restricting attention to such equilibria, but also develop a tractable method to characterize equilibrium behavior and payoffs in a Markov equilibrium. The equilibrium dynamics can be directly related to observable features of a long-run player, or other long-run player, and used to generate empirically testable predictions.

This paper leaves open several interesting avenues for future research. Continuous time provides a tractable framework for studying games of imperfect monitoring. Ideally, equilibria of the continuous time will be robust in the sense that nearby discrete time games will exhibit similar equilibrium properties, as the period length becomes small. [Faingold \(2008\)](#) establishes such a robustness property in the context of a reputation game with commitment types. Whether the current setting is robust to the period length remains an open question.

Often, multiple long-run players may compete for the support of a fixed population of

small players. For instance, rival long-run players may strive for a larger consumer base, political parties may contend for office, or universities may vie for the brightest students. These examples describe a setting in which each long-run player takes an action that persistently affects its state variable. Analyzing a setting with multiple state variables is technically challenging; if one could reduce such a game to a setting with a single payoff-relevant state, this simplification could yield a tractable characterization of equilibrium dynamics. For example, perhaps it is only the difference between two firms' product qualities that guide consumers' purchase behavior, or the difference between the platform of two political parties that determines constituents voting behavior.

Additionally, examining other classes of stochastic games, such as games between two long-run players whose actions jointly determine a stock variable, or games with different information structures governing the state transitions, remain unexplored.

References

- BOARD, S. AND M. MEYER-TER-VEHN (2011): "Reputation for Quality," Tech. rep., UCLA.
- CRIPPS, M. W., G. J. MAILATH, AND L. SAMUELSON (2007): "Disappearing private reputations in long-run relationships," *Journal of Economic Theory*, 134, 287 – 316.
- FAINGOLD, E. (2008): "Building a Reputation under Frequent Decisions," Tech. rep.
- FAINGOLD, E. AND Y. SANNIKOV (2011): "Reputation in Continuous-Time Games," *Econometrica*, 79, 773–876.
- FUDENBERG, D. AND D. LEVINE (2007): "Continuous Time Limits of Repeated Games with Imperfect Public Monitoring," *Review of Economic Dynamics*, 10, 173–192.
- FUDENBERG, D. AND D. K. LEVINE (1992): "Maintaining a Reputation When Strategies Are Imperfectly Observed," *Review of Economic Studies*, 59, 561–79.
- (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103–135.
- KLEIN, B. AND K. B. LEFFLER (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 615–41.
- KREPS, D. M., P. MILGROM, J. ROBERTS, AND R. WILSON (1982): "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory*, 27, 245–252.
- MAILATH, G. J. AND L. SAMUELSON (2001): "Who Wants a Good Reputation?" *Review of Economic Studies*, 68, 415–41.

SANNIKOV, Y. AND A. SKRZYPACZ (2007): “Impossibility of Collusion under Imperfect Monitoring with Flexible Production,” *American Economic Review*, 97, 1794–1823.

——— (2010): “The Role of Information in Repeated Games With Frequent Actions,” *Econometrica*, 78, 847–882.

6 Appendix

6.1 Proof of Lemma 1

This Lemma extends recursive techniques in continuous time games to the current setting of a stochastic game.

6.1.1 Evolution of the continuation value

Let $W_t(S)$ be the long-run player’s continuation value at time t , given $X_t = X$ and strategy profile $S = (a_t, \bar{b}_t)_{t \geq 0}$, and let $V_t(S)$ be the average discounted payoff conditional on info at time t .

$$\begin{aligned} V_t(S) & : = E_t \left[r \int_0^\infty e^{-rs} g(a_s, \bar{b}_s, X_s) ds \right] \\ & = r \int_0^t e^{-rs} g(a_s, \bar{b}_s, X_s) ds + e^{-rt} W_t(S) \end{aligned}$$

Lemma 3. *The average discounted payoff at time t , $V_t(S)$, is a martingale.*

$$\begin{aligned} E_t[V_{t+k}(S)] & = E_t \left[r \int_0^{t+k} e^{-rs} g(a_s, \bar{b}_s, X_s) ds + e^{-r(t+k)} W_{t+k}(S) \right] \\ & = r \int_0^t e^{-rs} g(a_s, \bar{b}_s, X_s) ds \\ & \quad + E_t \left[r \int_t^{t+k} e^{-rs} g(a_s, \bar{b}_s, X_s) ds + e^{-r(t+k)} E_{t+k} \left[r \int_{t+k}^\infty e^{-r(s-(t+k))} g(a_s, \bar{b}_s, X_s) ds \right] \right] \\ & = r \int_0^t e^{-rs} g(a_s, \bar{b}_s, X_s) ds \\ & \quad + e^{-rt} E_t \left[r \int_t^{t+k} e^{-r(s-t)} g(a_s, \bar{b}_s, X_s) ds + r \int_{t+k}^\infty e^{-r(s-t)} g(a_s, \bar{b}_s, X_s) ds \right] \\ & = r \int_0^t e^{-rs} g(a_s, \bar{b}_s, X_s) ds + e^{-rt} W_t(S) = V_t(S) \end{aligned}$$

Lemma 4. *In any PPE, the continuation value evolves according to the stochastic differential equation*

$$dW_t(S) = r (W_t(S) - g(a_t, \bar{b}_t, X_t)) dt + r \beta_{yt} [dY_t - \mu_y(a_t, \bar{b}_t, X_t) dt] + r \beta_{xt} (dX_t - \mu_x(a_t, \bar{b}_t, X_t) dt)$$

Take the derivative of $V_t(S)$ wrt t :

$$dV_t(S) = re^{-rt}g(a_t, \bar{b}_t, X_t)dt - re^{-rt}W_t(S)dt + e^{-rt}dW_t(S)$$

By the martingale representation theorem, there exists a progressively measurable process $(\beta_t)_{t \geq 0}$ such that V_t can be represented as:

$$dV_t(S) = re^{-rt}\beta_t^\top \sigma(\bar{b}_t, X_t) dZ_t$$

Combining these two expressions for $dV_t(S)$ yields the law of motion for the continuation value:

$$\begin{aligned} re^{-rt}\beta_t^\top \sigma(\bar{b}_t, X_t) dZ_t &= re^{-rt}g(a_t, \bar{b}_t, X_t)dt - re^{-rt}W_t(S)dt + e^{-rt}dW_t(S) \\ \Rightarrow dW_t(S) &= r(W_t(S) - g(a_t, \bar{b}_t, X_t))dt + r\beta_t^\top \sigma(\bar{b}_t, X_t) dZ_t \\ &= r(W_t(S) - g(a_t, \bar{b}_t, X_t))dt + r\beta_{yt} [dY_t - \mu_y(a_t, \bar{b}_t, X_t)dt] \\ &\quad + r\beta_{xt} [dX_t - \mu_x(a_t, \bar{b}_t, X_t)dt] \end{aligned}$$

where $\beta_t = (\beta_{yt}, \beta_{xt})$ is a vector of length d . The component β_{yt} captures the sensitivity of the continuation value to the public signal, while the component β_{xt} captures the sensitivity of the continuation value to the state variable.

Q.E.D.

6.1.2 Sequential Rationality

Lemma 5. *A strategy $(a_t)_{t \geq 0}$ is sequentially rational for the long-run player if, given $(\beta_t)_{t \geq 0}$, for all t :*

$$a_t \in \arg \max g(a', \bar{b}_t, X_t) + \beta_{yt}\mu_y(a', \bar{b}_t, X_t) + \beta_{xt}\mu_x(a', \bar{b}_t, X_t)$$

Consider strategy profile $(a_t, \bar{b}_t)_{t \geq 0}$ played from period τ onwards and alternative strategy $(\tilde{a}_t, \bar{b}_t)_{t \geq 0}$ played up to time τ . Recall that all values of X_t are possible under both strategies, but that each strategy induces a different measure over sample paths $(X_t)_{t \geq 0}$.

At time τ , the state variable is equal to X_τ . Action a_τ will induce

$$\begin{aligned} dY_\tau &= \mu_y(a_\tau, \bar{b}_\tau, X_\tau) dt + \sigma(\bar{b}_\tau, X_\tau) \cdot \frac{dZ_\tau^y}{dZ_\tau^x} \\ dX_\tau &= \mu_x(a_\tau, \bar{b}_\tau, X_\tau) dt + \sigma(\bar{b}_\tau, X_\tau) \cdot \frac{dZ_\tau^y}{dZ_\tau^x} \end{aligned}$$

whereas action \tilde{a}_τ will induce

$$\begin{aligned} dY_\tau &= \mu_y(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) dt + \sigma(\bar{b}_\tau, X_\tau) \cdot \frac{dZ_\tau^y}{dZ_\tau^x} \\ dX_\tau &= \mu_x(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) dt + \sigma(\bar{b}_\tau, X_\tau) \cdot \frac{dZ_\tau^y}{dZ_\tau^x} \end{aligned}$$

Let \tilde{V}_τ be the expected average payoff conditional on info at time τ when follows \tilde{a} up to τ and a afterwards, and let W_τ be the continuation value when the long-run player follows strategy $(a_t)_{t \geq 0}$ starting at time τ .

$$\tilde{V}_\tau = r \int_0^\tau e^{-rs} g(\tilde{a}_s, \bar{b}_s, X_s) ds + e^{-r\tau} W_\tau$$

Consider changing τ so that long-run player plays strategy (\tilde{a}_t, \bar{b}_t) for another instant: $d\tilde{V}_\tau$ is the change in average expected payoffs when the long-run player switches to $(a_t)_{t \geq 0}$ at $\tau + d\tau$ instead of τ . Note

$$\begin{aligned} dW_\tau &= r(W_\tau - g(a_\tau, \bar{b}_\tau, X_\tau)) d\tau \\ &\quad + r\beta_{y\tau} [dY_\tau - \mu_y(a_\tau, \bar{b}_\tau, X_\tau)d\tau] \\ &\quad + r\beta_{x\tau} [dX_\tau - \mu_x(a_\tau, \bar{b}_\tau, X_\tau)d\tau] \end{aligned}$$

when long-run player switches strategies at time τ .

$$\begin{aligned} d\tilde{V}_\tau &= re^{-r\tau} [g(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - W_\tau] d\tau + e^{-r\tau} dW_\tau \\ &= re^{-r\tau} [g(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - g(a_\tau, \bar{b}_\tau, X_\tau)] d\tau \\ &\quad + re^{-r\tau} \beta_{y\tau} [dY_\tau - \mu_y(a_\tau, \bar{b}_\tau, X_\tau)d\tau] \\ &\quad + re^{-r\tau} \beta_{x\tau} [dX_\tau - \mu_x(a_\tau, \bar{b}_\tau, X_\tau)d\tau] \\ &= re^{-r\tau} \left\{ \begin{array}{l} [g(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - g(a_\tau, \bar{b}_\tau, X_\tau)] d\tau \\ + \beta_{y\tau} [\mu_y(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - \mu_y(a_\tau, \bar{b}_\tau, X_\tau)] d\tau \\ + \beta_{x\tau} [\mu_x(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - \mu_x(a_\tau, \bar{b}_\tau, X_\tau)] d\tau \\ + \beta_\tau^\top \sigma(\bar{b}_\tau, X_\tau) dZ_\tau \end{array} \right\} \end{aligned}$$

There are two components to this strategy change: how it affects the immediate flow payoff and how it affects future public signal Y_t and state X_t , which impact the continuation value. The profile $(\tilde{a}_t, \bar{b}_t)_{t \geq 0}$ yields the long-run player a payoff of:

$$\begin{aligned} \tilde{W}_0 &= E_0 [\tilde{V}_\infty] = E_0 \left[\tilde{V}_0 + \int_0^\infty d\tilde{V}_t \right] \\ &= W_0 + E_0 \left[r \int_0^\infty e^{-rt} \left\{ \begin{array}{l} g(\tilde{a}_t, \bar{b}_t, X_t) + \beta_{yt}\mu_y(\tilde{a}_t, \bar{b}_t, X_t) + \beta_{xt}\mu_x(\tilde{a}_t, \bar{b}_t, X_t) \\ - g(a_t, \bar{b}_t, X_t) - \beta_{yt}\mu_y(a_t, \bar{b}_t, X_t) - \beta_{xt}\mu_x(a_t, \bar{b}_t, X_t) \end{array} \right\} dt \right] \end{aligned}$$

If

$$g(a_t, \bar{b}_t, X_t) + \beta_{yt}\mu_y(a_t, \bar{b}_t, X_t) + \beta_{xt}\mu_x(a_t, \bar{b}_t, X_t) \geq g(\tilde{a}_t, \bar{b}_t, X_t) + \beta_{yt}\mu_y(\tilde{a}_t, \bar{b}_t, X_t) + \beta_{xt}\mu_x(\tilde{a}_t, \bar{b}_t, X_t)$$

holds for all $t \geq 0$, Then $W_0 \geq \tilde{W}_0$ and deviating to $S = (\tilde{a}_t, \bar{b}_t)$ is not a profitable deviation. This yields the condition for sequential rationality for the long-run player.

Q.E.D.

6.2 Proof of Theorem 1: Characterization of Markov Equilibrium

Theorem 6. *Suppose Assumptions 1, 2 and 3 hold. Then given X_0 and equilibrium actions $(a^*, \bar{b}^*) = S(X, 0, U'(X))$, any solution $U(X)$ to the second order differential equation,*

$$U''(X) = \frac{2r(U(X) - g(a^*, \bar{b}^*, X))}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} - \frac{2\mu_x(a^*, \bar{b}^*, X)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} U'(X)$$

referred to as the optimality equation, characterizes a unique Markov equilibrium in the state variable $(X_t)_{t \geq 0}$ with

1. Equilibrium payoffs $U(X_0)$
2. Continuation values $(W_t)_{t \geq 0} = (U(X_t))_{t \geq 0}$
3. Equilibrium actions $(a_t^*, \bar{b}_t^*)_{t \geq 0}$ uniquely specified by

$$S(X_t, 0, U'(X_t)) = \left\{ \begin{array}{l} a_t^* = \arg \max_{a'} rg(a', \bar{b}_t^*, X_t) + U'(X_t) \mu_x(a', \bar{b}_t^*, X_t) \\ \bar{b}_t^* = \bar{B}(a_t^*, X_t) \end{array} \right\}$$

The optimality equation has at least one solution $U \in C^2(\mathbb{R})$ that lies in the range of feasible payoffs for the long-run player $U(X) \in [\underline{g}, \bar{g}]$ for all $X \in \Xi$. Thus, there exists at least one Markov equilibrium.

6.2.1 Proof of form of Optimality Equation:

Lemma 6. *If a Markov equilibrium exists, it takes the following form:*

1. Continuation values are characterized by a solution $U(X)$ to the optimality equation:

$$U''(X) = \frac{2r \left(U(X) - g(a^*, \bar{b}^*, X) \right)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} - \frac{2\mu_x(a^*, \bar{b}^*, X)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} U'(X)$$

given equilibrium actions $a^* = a(X)$ and $\bar{b}^* = \bar{b}(X)$.

2. Given solution $U(X)$, the process governing incentives for the long-run player is characterized by:

$$\begin{aligned} r\beta_y &= 0 \\ r\beta_x &= U'(X) \end{aligned}$$

Look for a Markov equilibrium in the state variable X_t . In a Markov equilibrium, the continuation value and equilibrium actions are characterized as a function of the state variable as $W_t = U(X_t)$, $a_t^* = a(X_t)$ and $\bar{b}_t^* = \bar{b}(X_t)$. By Ito's formula, in a Markov equilibrium, the continuation value will evolve according to:

$$\begin{aligned} dW_t &= dU(X_t) \\ &= U'(X_t) dX_t + \frac{1}{2} U''(X_t) \left[\left| \sigma_{xy}(\bar{b}_t^*, X_t) \right|^2 + \sigma_{xx}^2(\bar{b}_t^*, X_t) \right] dt \\ &= U'(X_t) \mu_x(a_t^*, \bar{b}_t^*, X_t) dt \\ &\quad + \frac{1}{2} U''(X_t) \left[\left| \sigma_{xy}(\bar{b}_t^*, X_t) \right|^2 + \sigma_{xx}^2(\bar{b}_t^*, X_t) \right] dt \\ &\quad + U'(X_t) \left[\sigma_{xy}(\bar{b}_t^*, X_t) dZ_t^y + \sigma_{xx}(\bar{b}_t^*, X_t) dZ_t^x \right] \end{aligned}$$

Also, given players are playing strategy $(a_t^*, \bar{b}_t^*)_{t \geq 0}$ and the state variable evolves according to the transition function dX_t , the continuation value evolves according to:

$$dW_t = r \left(W_t - g(a_t^*, \bar{b}_t^*, X_t) \right) dt + r \beta_t^\top \sigma \left(\bar{b}_t^*, X_t \right) dZ_t$$

I can match the drift of these two characterizations to obtain the optimality equation for strategy profile (a^*, \bar{b}^*) :

$$\begin{aligned} r \left(U(X) - g(a^*, \bar{b}^*, X) \right) &= U'(X) \mu_x(a^*, \bar{b}^*, X) + \frac{1}{2} U''(X) \left[\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X) \right] \\ \Rightarrow U''(X) &= \frac{2r \left(U(X) - g(a^*, \bar{b}^*, X) \right)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} - \frac{2\mu_x(a^*, \bar{b}^*, X)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} U'(X) \end{aligned}$$

which is a second order non-homogenous differential equation. Matching the volatility characterizes the process governing incentives. Note that

$$r \beta_t^\top \sigma dZ_t = [\beta_{yt} \sigma_{yy} + \beta_{xt} \sigma_{xy}] \cdot dZ_t^y + [\beta_{yt} \sigma_{yx} + \beta_{xt} \sigma_{xx}] dZ_t^x$$

Therefore,

$$\begin{aligned} r \beta_{yt} &= 0 \\ r \beta_{xt} &= U'(X_t) \end{aligned}$$

Intuitively, the continuation value and equilibrium actions are independent of the public signal in a Markov equilibrium; this is born out mathematically by the condition $\beta_{yt} = 0$.

Plugging these into the constraints for sequential rationality yields

$$\begin{aligned} S^*(X, 0, U'(X)) &= (a^*, \bar{b}^*) \text{ s.t.} \\ a^* &= \arg \max_a r g(a, \bar{b}^*, X) + U'(X) \mu_x(a, \bar{b}^*, X) \\ \bar{b}^* &= \bar{\mathcal{B}}(a^*, X) \end{aligned}$$

which are unique by Assumption 3.

Q.E.D.

6.2.2 Prove existence of bounded solution to optimality equation

Linear Growth

Lemma 7. *The optimality equation has linear growth. Suppose Assumption 1 holds. For all $M > 0$ and compact intervals $I \subset \Xi$, there exists a $K_I > 0$ such that for all $X \in I$, $(a, \bar{b}) \in A \times B$, $u \in [-M, M]$ and $u' \in \mathbb{R}$,*

$$u'' = \frac{2r [u - g(a, \bar{b}, X)] - 2\mu_x(a, \bar{b}, X) u'}{\left| \sigma_{xy}(\bar{b}, X) \right|^2 + \sigma_{xx}^2(\bar{b}, X)} \leq K (1 + |u'|)$$

Follows directly from the fact that $u \in [-M, M]$, $X \in I$, g and μ_x are Lipschitz continuous and the bound on $\left| \sigma_{xy}(\bar{b}, X) \right|^2 + \sigma_{xx}^2(\bar{b}, X)$.

Q.E.D.

Existence for Unbounded support

Theorem 7. *The optimality equation has at least one solution $U \in C^2(\mathbb{R})$ that lies in the range of feasible payoffs for the long-run player i.e. for all $X \in \mathbb{R}$*

$$\inf_{a,b,X} g(a,b,X) \leq U(X) \leq \sup_{a,b,x} g(a,b,X)$$

The existence proof uses the following theorem from Schmitt, which gives sufficient conditions for the existence of a bounded solution to a second order differential equation defined on \mathbb{R}^3 . The Theorem is reproduced below.

Theorem 8. *Let $\alpha, \beta \in \mathbb{R}$ be such that $\alpha \leq \beta$, $E = \{(t, u, v) \in \mathbb{R}^3 | \alpha \leq u \leq \beta\}$ and $f : E \rightarrow \mathbb{R}$ be continuous. Assume that α and β are such that for all $t \in \mathbb{R}$*

$$\begin{aligned} f(t, \alpha, 0) &\leq 0 \\ f(t, \beta, 0) &\geq 0 \end{aligned}$$

Assume that for any bounded interval I , there exists a positive continuous function $\phi_I : \mathbb{R}^+ \rightarrow \mathbb{R}$ that satisfies

$$\int_0^\infty \frac{s ds}{\phi_I(s)} = \infty$$

and for all $t \in I$, $(u, v) \in \mathbb{R}^2$ with $\alpha \leq u \leq \beta$,

$$|f(t, u, v)| \leq \phi_I(|v|)$$

Then the equation $u'' = f(t, u, v)$ has at least one solution on $u \in C^2(\mathbb{R})$ such that for all $t \in \mathbb{R}$,

$$\alpha \leq u(t) \leq \beta$$

Requires $\sigma_I = \inf_{\bar{b} \in B, X \in I} [\sigma^2] > 0$ and on any interval I (not uniformly), and optimality equation is continuous (Lipschitz continuity of fns)

Let $\bar{g} = \sup g(a, \bar{b}, X)$ and $\underline{g} = \inf g(a, \bar{b}, X)$, which are well defined since g is bounded. Applying the above theorem with $\alpha = \underline{g}$ and $\beta = \bar{g}$ to $h(X, U(X), U'(X))$ yields

$$\begin{aligned} h(X, \underline{g}, 0) &= \frac{2r}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} \left(\underline{g} - g(a^*, \bar{b}^*, X) \right) \leq 0 \\ h(X, \bar{g}, 0) &= \frac{2r}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} \left(\bar{g} - g(a^*, \bar{b}^*, X) \right) \geq 0 \end{aligned}$$

for all X . For any bounded interval I , define

$$\phi_I(v) = \frac{2r}{\sigma_I^2} (\bar{g} - \underline{g}) - \frac{2\mu_I}{\sigma_I^2} v$$

where $\sigma_I = \inf_{b \in B, X \in I} |\sigma_{xy}(b, X)|^2 + \sigma_{xx}^2(b, X)$ is positive by assumption and $\mu_I = \inf_{a \in A, b \in B, X \in I} \mu_x(a, \bar{b}, X)$. Both are well-defined given μ_x and σ are Lipschitz continuous and A and B are compact. Note

$$\int_0^\infty \frac{s ds}{\phi_I(s)} = \infty$$

and for all $X \in I, (u, v) \in \mathbb{R}^2$ with $\underline{g} \leq u \leq \bar{g}$

$$\begin{aligned} |h(X, u, v)| &= \left| \frac{\frac{2r}{|\sigma_{xy}(\bar{b}^*, X)|^2 + \sigma_{xx}^2(\bar{b}^*, X)} \left(u - g(a^*, \bar{b}^*, X) \right)}{-\frac{2\mu_x(a^*, \bar{b}^*, X)}{|\sigma_{xy}(\bar{b}^*, X)|^2 + \sigma_{xx}^2(\bar{b}^*, X)}} v \right| \\ &\leq \frac{2r}{\sigma_I^2} (\bar{g} - \underline{g}) - \frac{2\mu_I}{\sigma_I^2} v \\ &\leq \phi_I(|v|) \end{aligned}$$

Additionally, $h(X_t, U(X_t), U'(X_t))$ is continuous given that μ_x and σ are Lipschitz continuous and $g(a^*, \bar{b}^*, X)$ is the composite of Lipschitz continuous functions. Thus, $h(X, U(X), U'(X))$ has at least one solution on $U \in C^2(\mathbb{R})$ such that for all $X \in \mathbb{R}$,

$$\underline{g} \leq U(X) \leq \bar{g}$$

Q.E.D.

Existence for Bounded support

Theorem 9. *The optimality equation has at least one solution $U \in C^2(\mathbb{R})$ that lies in the range of feasible payoffs for the long-run player i.e. for all $X \in \Xi$*

$$\inf g(a, \bar{b}, X) \leq U(X) \leq \sup g(a, \bar{b}, X)$$

The existence proof utilizes standard existence results from de Coster and Habets (2006) and an extension in Faingold and Sannikov (2011), applied to the current setting. The optimality equation is undefined at the upper and lower bound of the state space, \underline{X} and \bar{X} , since the volatility of the state variable is zero. Therefore, an extension of standard existence results for second order ODEs is necessary. The main idea is to show that the boundary value problem has a solution U_n on $X \in [\underline{X} + 1/n, \bar{X} - 1/n]$ for every $n \in N$, and then to show that this sequence of solutions converges point-wise to a continuously differentiable function U defined on (\underline{X}, \bar{X}) .

Given the boundary value problem has a solution U_n for every $n \in N$, with $\underline{X} = 0$ and $\bar{X} = 1$, Faingold and Sannikov (2011) show that when the second derivative of the ODE has quadratic growth, then a subsequence of $(U_n)_{n \geq 0}$ converges point-wise to a continuously differentiable function U defined on $(0, 1)$.

In this model, the second order derivative has linear growth, and therefore a similar argument shows existence of a continuously differentiable function U defined on (\underline{X}, \bar{X}) .

The existence results that are relevant for the current context are reproduced below:

Lemma 8. Let $E = \{(t, u, v) \in \Xi \times \mathbb{R}^2\}$ and $f : E \rightarrow \mathbb{R}$ be continuous. Assume that for any interval $I \subset \Xi$, there exists a $K_I > 0$ such that for all $t \in I$, $(u, v) \in \mathbb{R}^2$ with $\alpha \leq u \leq \beta$,

$$|f(t, u, v)| \leq K_I(1 + |v|)$$

Then the equation $u'' = f(t, u, v)$ has at least one solution on $u \in C^2(\mathbb{R})$ such that for all $t \in \Xi$,

$$\alpha \leq u(t) \leq \beta$$

Consider the optimality equation $h(X, U(X), U'(X))$. Let $\bar{g} = \sup g(a, \bar{b}, X)$ and $\underline{g} = \inf g(a, \bar{b}, X)$, which are well defined since g is bounded. By 7, for any bounded interval I and $u \in [\underline{g}, \bar{g}]$, there exists a K_I such that

$$|f(t, u, v)| \leq K_I(1 + |v|)$$

Additionally, $h(X_t, U(X_t), U'(X_t))$ is continuous given that μ_x and σ are Lipschitz continuous and $g(a(X), \bar{b}(X), X)$ is continuous. Let $\alpha = \underline{g}$ and $\beta = \bar{g}$. Then $h(X, U(X), U'(X))$ has at least one solution on $U \in C^2(\mathbb{R})$ such that for all $X \in \Xi$,

$$\underline{g} \leq U(X) \leq \bar{g}$$

Q.E.D.

6.2.3 Construct a Markov equilibrium

Suppose the state variable initially starts at X_0 and U is a bounded solution to the optimality equation. The action profile satisfying $(a^*, \bar{b}^*) = S^*(X, 0, U'(X))$ is unique and Lipschitz continuous in X and U . Thus, given X_0, U and $(a_t^*, \bar{b}_t^*)_{t \geq 0} = [S^*(X_t, 0, U'(X_t))]_{t \geq 0}$, the state variable uniquely evolves according to the stochastic differential equation

$$dX_t = \mu_x(a_t^*, \bar{b}_t^*, X_t)dt + \sigma_{xy}(\bar{b}_t^*, X_t) dZ_t^y + \sigma_{xx}(\bar{b}_t^*, X_t) dZ_t^x$$

yielding unique path $(X_t)_{t \geq 0}$ given initial state X_0 . Given that $U(X_t)$ is a bounded process that satisfies

$$\begin{aligned} dU(X_t) &= U'(X_t)\mu_x(a_t^*, \bar{b}_t^*, X_t)dt \\ &\quad + \frac{1}{2}U''(X_t) \left[\left| \sigma_{xy}(\bar{b}_t^*, X_t) \right|^2 + \sigma_{xx}^2(\bar{b}_t^*, X_t) \right] dt \\ &\quad + U'(X_t) \left[\sigma_{xy}(\bar{b}_t^*, X_t) dZ_t^y + \sigma_{xx}(\bar{b}_t^*, X_t) dZ_t^x \right] \\ &= r \left(U(X_t) - g(a_t^*, \bar{b}_t^*, X_t) \right) + U'(X_t) \left[\sigma_{xy}(\bar{b}_t^*, X_t) dZ_t^y + \sigma_{xx}(\bar{b}_t^*, X_t) dZ_t^x \right] \end{aligned}$$

this process satisfies the conditions for the continuation value in a PPE characterized in Lemma 1. Additionally, $(a_t^*, \bar{b}_t^*)_{t \geq 0}$ satisfies the condition for sequential rationality given process $(\beta_t)_{t \geq 0} = (0, U'(X_t))_{t \geq 0}$. Thus, the strategy profile $(a_t^*, \bar{b}_t^*)_{t \geq 0}$ is a PPE yielding equilibrium payoff $U(X_0)$.

Q.E.D.

6.3 Proof of Lemma 2

Lemma 9. *For all $U, V \in \Upsilon$, either $U(X) > V(X)$ for all $X \in \Xi$ or $U(X) < V(X)$ for all $X \in \Xi$ or $U(X) = V(X)$ for all $X \in \Xi$. In other words, two distinct solutions never cross.*

6.4 Proof of Theorem 2: Characterize PPE Payoff Set

Theorem 10. *Suppose Assumptions 1, 2 and 3 hold and let Υ be the set of bounded solutions to the optimality equation. The correspondence of perfect public equilibrium payoffs $\xi : \Xi \Rightarrow \mathbb{R}$ is characterized by the convex hull of the highest and lowest Markov equilibrium payoff, $co(U_1, U_2)$.*

The proof proceeds by a series of lemmas.

Lemma 10. *Let U_2 be the greatest bounded solution to the optimality equation and let U_1 be the least bounded solution to the optimality equation. Then any PPE payoff W_0 is such that $U_1(X_o) \leq W_0 \leq U_2(X_o)$.*

Let X_0 be the initial state, and let U be the greatest bounded solution to the optimality equation (which exists by Lemma ??). Suppose there is a PPE $(a_t, \bar{b}_t)_{t \geq 0}$ that yields an equilibrium payoff $W_0 > U(X_o)$. The continuation value in this equilibrium must evolve according to

$$\begin{aligned} dW_t(S) &= r (W_t(S) - g(a_t, \bar{b}_t, X_t)) dt \\ &\quad + r \beta_{yt}^\top [dY_t - \mu_y(a_t, \bar{b}_t, X_t) dt] \\ &\quad + r \beta_{xt} [dX_t - \mu_x(a_t, \bar{b}_t, X_t) dt] \end{aligned}$$

for some process $(\beta_t)_{t \geq 0}$ and by sequential rationality, $(a_t, \bar{b}_t) = S^*(X_t, \beta_t)$ for all t . The process $U(X_t)$ evolves according to

$$\begin{aligned} dU(X_t) &= U'(X_t) \mu_x(a_t, \bar{b}_t, X_t) dt + \frac{1}{2} U''(X_t) \left(|\sigma_{xy}(\bar{b}_t, X_t)|^2 + \sigma_{xx}^2(\bar{b}_t, X_t) \right) dt \\ &\quad + U'(X_t) \left(\sigma_{xy}(\bar{b}_t, X_t) dZ_t^y + \sigma_{xx}(\bar{b}_t, X_t) dZ_t^x \right) \end{aligned}$$

Define a process $D_t = W_t - U(X_t)$ with initial condition $D_0 = W_0 - U(X_o) > 0$. Then dD_t evolves with drift

$$\begin{aligned} &r (W_t - g(a_t, \bar{b}_t, X_t)) - U'(X_t) \mu_x(a_t, \bar{b}_t, X_t) - \frac{1}{2} U''(X_t) \left(|\sigma_{xy}(\bar{b}_t, X_t)|^2 + \sigma_{xx}^2(\bar{b}_t, X_t) \right) \\ &= r D_t + r (U(X_t) - g(a_t, \bar{b}_t, X_t)) - U'(X_t) \mu_x(a_t, \bar{b}_t, X_t) - \frac{1}{2} U''(X_t) \left(|\sigma_{xy}(\bar{b}_t, X_t)|^2 + \sigma_{xx}^2(\bar{b}_t, X_t) \right) \\ &= r D_t + d(a_t, \bar{b}_t, X_t) \end{aligned}$$

and volatility

$$f(\bar{b}_t, X_t, \beta_t) = \begin{bmatrix} r \beta_{yt}^\top \sigma_{yy}(\bar{b}_t, X_t) & r \beta_{yt}^\top \sigma_{yx}(\bar{b}_t, X_t) \\ r \beta_{xt} \sigma_{xy}(\bar{b}_t, X_t) & r \beta_{xt} \sigma_{xx}(\bar{b}_t, X_t) \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ U'(X_t) \sigma_{xy}(\bar{b}_t, X_t) & U'(X_t) \sigma_{xx}(\bar{b}_t, X_t) \end{bmatrix}$$

Note that in a Markov action profile $(a_t^*, \bar{b}_t^*) = S^*(X_t, 0, U'(X_t))$,

$$\begin{aligned} d(a_t^*, \bar{b}_t^*, X_t) &= r \left(U(X_t) - g(a_t^*, \bar{b}_t^*, X_t) \right) - U'(X_t) \mu_x(a_t^*, \bar{b}_t^*, X_t) \\ &\quad - \frac{1}{2} U''(X_t) \left(\left| \sigma_{xy}(\bar{b}_t^*, X_t) \right|^2 + \sigma_{xx}^2(\bar{b}_t^*, X_t) \right) \\ &= 0 \end{aligned}$$

by the optimality equation.

Lemma 11.

Claim 1. *For every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all (a, \bar{b}, X, β) satisfying the condition for sequential rationality*

$$\begin{aligned} a &\in \arg \max r g(a', \bar{b}, X) + \beta_y \mu_y(a, \bar{b}, X) + \beta_x \mu_x(a, \bar{b}, X) \\ \bar{b} &\in \bar{\mathcal{B}}(a, X) \end{aligned}$$

either $d(a_t, \bar{b}_t, X_t) > -\varepsilon$ or $|f(\bar{b}, X, \beta)| > \delta$.

Suppose the state space is unbounded, $\Xi = \mathbb{R}$.

Step 1: Show that if $|f(\bar{b}, X, \beta)| = 0$, then $d(a, \bar{b}, X) = 0$ (i.e. when the volatility of D_t is 0, the Markov action profile is used in both equilibria).

Let $|f(\bar{b}, X, \beta_1)| = 0$. Then $r\beta_y = 0$ and $r\beta_x = U'(X)$, and for each X , there is a unique action (a, \bar{b}) profile that satisfies

$$\begin{aligned} a &\in \arg \max r g(a', \bar{b}, X) + \beta_x \mu_x(a', \bar{b}, X) \\ \bar{b} &\in \bar{\mathcal{B}}(a, X) \end{aligned}$$

This action profile corresponds to the actions played in a Markov equilibrium at state X , and therefore $d(a_t, \bar{b}_t, X_t) = 0$.

Step 2: Fix ε . Show if $d(a_t, \bar{b}_t, X_t) \leq -\varepsilon$, then there exists a $\delta > 0$ such that $|f(\bar{b}, X, \beta)| > \delta$ for all (a, \bar{b}, X, β) such that the sequential rationality condition is satisfied.

Step 2a: Show there exists an $M > 0$ such that this is true for

$$(a, \bar{b}, X, \beta) \in \{A \times B \times \Xi \times \mathbb{R} : |\beta| > M\}$$

$U'(X)$ is bounded and there exists a $c > 0$ such that $|\sigma \cdot y| \geq c|y|$ for all $(b, X) \in B \times \Xi$ and $y \in \mathbb{R}^d$, so there exists an $M > 0$ and $m > 0$ such that $|f(\bar{b}, X, \beta)| > m$ for all $|\beta| > M$, regardless of d .

Step 2b: Show that there exists an X^* such that this is true for

$$(a, \bar{b}, X, \beta) \in \{A \times B \times \Xi \times \mathbb{R} : |\beta| \leq M \text{ and } |X| > X^*\}$$

Let $\lim_{X \rightarrow \infty} g(a, \bar{b}, X) := g_\infty(a, \bar{b})$ be the limit flow payoff for actions (a, \bar{b}) , and $\lim_{X \rightarrow \infty} \bar{\mathcal{B}}(a, X) := \bar{\mathcal{B}}_\infty(a)$. These limits *exist*, since g and $\bar{\mathcal{B}}(a)$ are Lipschitz continuous and bounded. Consider the set $\Phi = (a, \bar{b}, X, \beta)$ that satisfies

$$\begin{aligned} a &\in \arg \max r g_\infty(a', \bar{b}) + \beta_y \mu_y(a, \bar{b}, X) + \beta_x \mu_x(a, \bar{b}, X) \\ \bar{b} &\in \bar{\mathcal{B}}_\infty(a) \end{aligned}$$

with $|\beta| < M$. Suppose $r(v_\infty - g_\infty(a, \bar{b})) \leq -\gamma$. Then there exists a $\eta_2 > 0$ such that $|\beta| > \eta_2$. Thus, $\lim_{X \rightarrow \infty} |f(\bar{b}, X, \beta)| = r|\beta| > r\eta_2$.

Note $\lim_{X \rightarrow \infty} d(a, \bar{b}, X) = r[v_\infty - g_\infty(a, \bar{b})]$. Then there exists an X_1 such that for $X > X_1$, $|d(a, \bar{b}, X) - r[g_\infty(a, \bar{b}) - g_\infty(a^N, \bar{b}^N)]| < \varepsilon/2$. Consider the set $\Phi = (a, \bar{b}, X, \beta)$ that satisfy the condition for sequential rationality, with $|\beta| \leq M$ and $|X| \geq X_1$, and $d(a, \bar{b}, X) \leq -\varepsilon$. For $X > X_1$, $|d(a, \bar{b}, X) - r[g_\infty(a, \bar{b}) - g_\infty(a^N, \bar{b}^N)]| < \varepsilon/2$. Then $r(v_\infty - g_\infty(a, \bar{b})) \leq -\varepsilon/2$. Then there exists a η_2 such that $|\beta| > \eta_2$. Thus, $\lim_{X \rightarrow \infty} |f(\bar{b}, X, \beta)| = r|\beta| > r\eta_2$. Then there exists an X_2 such that for $X > X_2$, $|f(\bar{b}, X, \beta) - r\beta| < r\eta_2/2$. Then $|f(\bar{b}, X, \beta)| > r\eta_2/2 := \delta_2$. Take $X^* = \max\{X_1, X_2\}$. Then on the set $(a, \bar{b}, X, \beta) \in \{A \times B \times \Xi \times \mathbb{R} : |\beta| \leq M \text{ and } |X| > X^*\}$, if $d(a_t, \bar{b}_t, X_t) \leq -\varepsilon$ then $|f(\bar{b}, X, \beta)| > \delta_2$.

Step 2c: Show this is true for $(a, \bar{b}, X, \beta) \in \{A \times B \times \Xi \times \mathbb{R} : |\beta| \leq M \text{ and } |X| \leq X^*\}$

Consider the set $\Phi = (a, \bar{b}, X, \beta)$ that satisfy the condition for sequential rationality with $|\beta| \leq M$ and $|X| \leq X^*$, and $d(a, \bar{b}, X) \leq -\varepsilon$. The function d is continuous and $\{A \times B \times \Xi \times \mathbb{R} : |\beta| \leq M \text{ and } |X| \leq X^*\}$ is compact, so Φ is compact. f is also continuous, and therefore achieves a minimum η_1 on Φ . This minimum $\eta_1 > 0$, because $\eta_1 = 0$ would imply that $d = 0$, and I assume $d(a_t, \bar{b}_t, X_t) < -\varepsilon$. Thus, $|f(\bar{b}, X, \beta)| > \eta_1$ for all $(a, \bar{b}, X, \beta) \in \Phi$.

Take $\delta = \min\{\eta_1, \delta_2, m\}$. Then when $d(a, \bar{b}, X) \leq -\varepsilon$, $|f(\bar{b}, X, \beta)| > \delta$.

The proof for a bounded state space is analogous, omitting step 2b.

This claim implies that whenever the drift of D_t is less than $rD_t - \varepsilon$, the volatility is greater than δ . Take $\varepsilon = rD_0/4$ and suppose $D_t \geq D_0/2$. Then whenever the drift is less than $rD_t - \varepsilon > rD_0/2 - rD_0/4 = rD_0/4 > 0$, there exists a δ such that $|f(\bar{b}, X, \beta)| > \delta$. Thus, whenever $D_t \geq D_0/2 > 0$, it has either positive drift or positive volatility, and grows arbitrarily large with positive probability. This is a contradiction, since D_t is the difference of two bounded processes. Thus, it cannot be that $D_0 > 0$ and it must be the case that $W_0 \leq U(X_o)$

Similarly, when U is the least bounded solution to the optimality equation, it is not possible to have $D_0 < 0$ and therefore it must be the case that $W_0 \geq U(X_o)$.

Q.E.D.

The second part of Theorem 2 follows directly from Lemma 10, and the fact that it is possible to achieve any equilibrium payoff in the convex hull of the highest and lowest Markov equilibrium payoff with randomization.

6.5 Proof of Theorem 3: Characterize Unique Markov Equilibrium

Theorem 11. *Suppose Assumptions 1, 2, 3 and 4 hold. Then, for each initial value of the state variable $X_0 \in \Xi$, there exists a unique perfect public equilibrium, which is Markov. The unique bounded solution U of the optimality equation characterizes the equilibrium payoff $U(X_0)$ and continuation values.*

1. *When the state space is unbounded, $\Xi = \mathbb{R}$, then the solution satisfies the following boundary conditions:*

$$\begin{aligned} \lim_{X \rightarrow \infty} U(X) = v_\infty \text{ and } \lim_{X \rightarrow -\infty} U(X) = v_{-\infty} \\ \lim_{X \rightarrow \infty} f(X)U'(X) = \lim_{X \rightarrow -\infty} f(X)U'(X) = 0 \\ \lim_{X \rightarrow \infty} f(X)^2U''(X) = \lim_{X \rightarrow -\infty} f(X)^2U''(X) = 0 \end{aligned}$$

for any function $f(X)$ that is Lipschitz continuous on \mathbb{R} , where $v_\infty = \lim_{X \rightarrow \infty} v(X)$ and $v_{-\infty} = \lim_{X \rightarrow -\infty} v(X)$.

2. *When the state space is bounded, $\Xi = [\underline{X}, \overline{X}]$, then the solution satisfies the following boundary conditions:*

$$\begin{aligned} \lim_{X \rightarrow \overline{X}} U(X) = v(\overline{X}) \text{ and } \lim_{X \rightarrow \underline{X}} U(X) = v(\underline{X}) \\ \lim_{X \rightarrow \overline{X}} f(X)U'(X) = \lim_{X \rightarrow \underline{X}} f(X)U'(X) = 0 \\ \lim_{X \rightarrow \overline{X}} f(X)^2U''(X) = \lim_{X \rightarrow \underline{X}} f(X)^2U''(X) = 0 \end{aligned}$$

for any function $f(X)$ that is Lipschitz continuous on $[\underline{X}, \overline{X}]$ with $f(\overline{X}) = f(\underline{X}) = 0$.

In particular, this implies that $\lim_{X \rightarrow \overline{X}} \mu_x(a^*, \overline{b}^*, X)U'(X) = \lim_{X \rightarrow \underline{X}} \mu_x(a^*, \overline{b}^*, X)U'(X) = 0$ and $\lim_{X \rightarrow \infty} \mu_x(a^*, \overline{b}^*, X)U'(X) = \lim_{X \rightarrow -\infty} \mu_x(a^*, \overline{b}^*, X)U'(X) = 0$, since $\mu_x(a^*, \overline{b}^*, X)$ is Lipschitz continuous and $\mu_x(a, b, \overline{X}) = \mu_x(a, b, \underline{X}) = 0$ for all $a, b \in A, B$ by Assumption 4.

6.5.1 Convergence Rate of Lipschitz Functions

Lemma 12. *1. Suppose a function $f(x)$ is Lipschitz continuous on \mathbb{R} . Then there exists a $K \in \mathbb{R}$ and a $\delta_1 \in \mathbb{R}$ such that when $|x| > \delta_1$, $|f(x)| \leq K|x|$. In other words, $f(x)$ is $O(x)$.*

2. *Suppose a function $f(x)$ is Lipschitz continuous on $[\underline{x}, \overline{x}] \subset \mathbb{R}$. Then for any $x^* \in [\underline{x}, \overline{x}]$, $f(x) - f(x^*)$ is $O(x - x^*)$ as $x \rightarrow x^*$.*

Suppose $f(x)$ is Lipschitz continuous on \mathbb{R} . Then there exists a $K_1 \in \mathbb{R}$ such that for all $x_1, x_2 \in \mathbb{R}$, $|f(x_2) - f(x_1)| \leq K_1 |x_2 - x_1|$. Take $x_1 = 0$. Given that $f(x)$ is Lipschitz continuous, $\exists K_2 \in \mathbb{R}$, $|f(0)| \leq K_2$. Then

$$\begin{aligned} & |f(x_2) - f(0)| \leq K_1 |x_2| \\ \Rightarrow & |f(x_2)| \leq \max \{K_1 |x_2| + f(0), K_1 |x_2| - f(0)\} \\ \Rightarrow & |f(x_2)| \leq K_1 |x_2| + K_2 \end{aligned}$$

Therefore, there exists a K and an x^* such that for $|x| > x^*$, $|f(x)| \leq K|x|$, which is the definition of $O(x)$.

Suppose $f(x)$ is Lipschitz continuous on $[\underline{x}, \bar{x}]$. Then there exists a $K \in \mathbb{R}$ such that for all $x_1, x_2 \in [\underline{x}, \bar{x}]$, $|f(x_2) - f(x_1)| \leq K |x_2 - x_1|$. Take $x_1 = \underline{x}$. Then $|f(x_2) - f(\underline{x})| \leq K |x_2 - \underline{x}|$. Note that $f(\underline{x})$ is a constant. Therefore, $f(x) - f(\underline{x})$ is $O(x - \underline{x})$ as $x \rightarrow \underline{x}$. As a special case, this implies that if $f(\underline{x}) = 0$, then $f(x)$ is $O(x - \underline{x})$.

6.5.2 Boundary Conditions

Boundary Conditions for Unbounded Support

Theorem 12. *Suppose Assumptions 1, 2, 3 and 4 hold. Any bounded solution U of the optimality equation satisfies the following boundary conditions*

$$\begin{aligned} \lim_{X \rightarrow \infty} U(X) &= v_\infty \text{ and } \lim_{X \rightarrow -\infty} U(X) = v_{-\infty} \\ \lim_{X \rightarrow \infty} f(X)U'(X) &= \lim_{X \rightarrow -\infty} f(X)U'(X) = 0 \\ \lim_{X \rightarrow \infty} f(X)^2U''(X) &= \lim_{X \rightarrow -\infty} f(X)^2U''(X) = 0 \end{aligned}$$

for any function $f(X)$ that is Lipschitz continuous on \mathbb{R} .

In particular, this implies that

$$\lim_{X \rightarrow \infty} \mu_x(a^*, \bar{b}^*, X)U'(X) = \lim_{X \rightarrow -\infty} \mu_x(a^*, \bar{b}^*, X)U'(X) = 0$$

since $\mu_x(a^*, \bar{b}^*, X)$ is Lipschitz continuous. The proof proceeds by a series of lemmas.

Lemma 13. *The limits $v_\infty = \lim_{X \rightarrow \infty} v(X)$ and $v_{-\infty} = \lim_{X \rightarrow -\infty} v(X)$ exist.*

The fact that g is bounded, $v(X) = g(S^*(X, 0, 0), X)$ is a composite of Lipschitz continuous functions, and is therefore Lipschitz continuous, and $v(X)$ is monotone for large enough X guarantees the existence of $\lim_{X \rightarrow \infty} v(X)$ and $\lim_{X \rightarrow -\infty} v(X)$. Q.E.D.

Lemma 14. *If $U(X)$ is a bounded solution of the optimality equation, then there exists a δ such that for $|X| > \delta$, $U(X)$ is monotonic. Additionally, $U_\infty = \lim_{X \rightarrow \infty} U(X)$ and $U_{-\infty} = \lim_{X \rightarrow -\infty} U(X)$ exist.*

Proof: Suppose that there does not exist a δ such that for $X > \delta$, U is monotonic. Then for all $\delta > 0$, there exists a $X_n > \delta$ that corresponds to a local maxima of U , so $U'(X_n) = 0$ and $U''(X_n) \leq 0$ and there exists a $X_m > \delta$ that corresponds to a local minima of U , so $U'(X_m) = 0$ and $U''(X_m) \geq 0$, by the continuity of U . Given the incentives for the long-run player, a stage Nash equilibria is played when $U'(X) = 0$, yielding flow payoff $v(X)$. From the optimality equation, this implies $v(X_n) \geq U(X_n)$ at the maximum and $v(X_m) \leq U(X_m)$ at the minimum. Thus, the oscillation of $v(X)$ is at least as large as the oscillation of $U(X)$. This is a contradiction, as there exists a δ such that for $X > \delta$, $v(X)$ is monotonic. The case of $-X > \delta$ is similar. The fact that U is bounded and monotone for large enough X guarantees the existence of $\lim_{X \rightarrow \infty} U(X)$ and $\lim_{X \rightarrow -\infty} U(X)$. Q.E.D.

Lemma 15. *Suppose a function $f(X)$ is Lipschitz continuous on \mathbb{R} . Then any bounded solution U of the optimality equation satisfies:*

1.

$$\begin{aligned} \lim_{X \rightarrow \infty} \inf f(X)U'(X) &\leq 0 \leq \lim_{X \rightarrow \infty} \sup f(X)U'(X) \\ \lim_{X \rightarrow -\infty} \inf f(X)U'(X) &\leq 0 \leq \lim_{X \rightarrow -\infty} \sup f(X)U'(X) \end{aligned}$$

2.

$$\begin{aligned} \lim_{X \rightarrow \infty} \inf f(X)^2U''(X) &\leq 0 \leq \lim_{X \rightarrow \infty} \sup f(X)^2U''(X) \\ \lim_{X \rightarrow -\infty} \inf f(X)^2U''(X) &\leq 0 \leq \lim_{X \rightarrow -\infty} \sup f(X)^2U''(X) \end{aligned}$$

1. Suppose $f(X)$ is Lipschitz continuous and $\lim_{X \rightarrow \infty} \inf |f(X)U'(X)| > 0$. By 12, f is $O(X)$, so there exists an $M \in \mathbb{R}$ and a $\delta_1 \in \mathbb{R}$ such that when $X > \delta_1$, $|f(X)| \leq M|X|$. Given $\lim_{X \rightarrow \infty} \inf |f(X)U'(X)| > 0$, there exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $X > \delta_2$, $|f(X)U'(X)| > \varepsilon$. Take $\delta = \max\{\delta_1, \delta_2\}$. Then for $X > \delta$, $|U'(X)| > \frac{\varepsilon}{|f(X)|} \geq \frac{\varepsilon}{MX}$. Then the antiderivative of $\frac{\varepsilon}{MX}$ is $\frac{\varepsilon}{M} \ln X$ which converges to ∞ as $X \rightarrow \infty$. This violates the boundedness of U . Therefore $\lim_{X \rightarrow \infty} \inf f(X)U'(X) \leq 0$. The proof is analogous for the other cases.
2. Suppose $f(X)$ is Lipschitz continuous and $\lim_{X \rightarrow \infty} \inf |f(X)^2U''(X)| > 0$. By 12, f is $O(X)$, so there exists an $M \in \mathbb{R}$ and a $\delta_1 \in \mathbb{R}$ such that when $X > \delta_1$, $|f(X)| \leq MX$ and therefore, $f(X)^2 \leq M^2X^2$. There also exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $X > \delta_2$, $|f(X)^2U''(X)| > \varepsilon$. Take $\delta = \max\{\delta_1, \delta_2\}$. Then for $X > \delta$, $|U''(X)| > \frac{\varepsilon}{f(X)^2} > \frac{\varepsilon}{M^2X^2}$. Then the antiderivative of $\frac{\varepsilon}{M^2X^2}$ is $\frac{-\varepsilon}{M^2} \ln X$ which converges to $-\infty$ as $X \rightarrow \infty$. This violates the boundedness of U . Therefore $\lim_{X \rightarrow \infty} \inf |f(X)^2U''(X)| \leq 0$. The proof is analogous for the other cases.

Q.E.D.

Lemma 16. *Suppose a function $f(X)$ is Lipschitz continuous on \mathbb{R} . Then any bounded solution U of the optimality equation satisfies*

$$\lim_{X \rightarrow \infty} f(X)U'(X) = \lim_{X \rightarrow -\infty} f(X)U'(X) = 0$$

By Lemma 15, $\lim_{X \rightarrow \infty} \inf XU'(X) \leq 0 \leq \lim_{X \rightarrow \infty} \sup XU'(X)$. Suppose, without loss of generality, that $\lim_{X \rightarrow \infty} \sup XU'(X) > 0$. By Lemma 14, there exists a $\delta > 0$ such that $U(X)$ is monotonic for $X > \delta$. Then for $X > \delta$, $U'(X)$ doesn't change sign and therefore, $XU'(X)$ doesn't change sign. Therefore, if $\lim_{X \rightarrow \infty} \sup XU'(X) > 0$, then $\lim_{X \rightarrow \infty} \inf XU'(X) > 0$. This is a contradiction. Thus, $\lim_{X \rightarrow \infty} \sup XU'(X) = 0$. By similar reasoning, $\lim_{X \rightarrow \infty} \inf XU'(X) = 0$, and therefore $\lim_{X \rightarrow \infty} XU'(X) = 0$. Suppose $f(X)$ is $O(X)$. Then there exists an $M \in \mathbb{R}$ and a $\delta_1 \in \mathbb{R}$ such that when $X > \delta_1$, $|f(X)| \leq M|X|$. Thus, for $X > \delta_1$, $|f(X)U'(X)| \leq M|XU'(X)| \rightarrow 0$. The case for $\lim_{X \rightarrow -\infty} f(X)U'(X) = 0$ is analogous. Note this result also implies that

$$\lim_{X \rightarrow \infty} U'(X) = \lim_{X \rightarrow -\infty} U'(X) = 0$$

Q.E.D.

Lemma 17. *Let U be a bounded solution of the optimality equation. Then the limit of U converges to the limit of the stage Nash equilibrium payoffs as $X \rightarrow \{-\infty, \infty\}$*

$$\begin{aligned} \lim_{X \rightarrow \infty} U(X) &= v_\infty \\ \lim_{X \rightarrow -\infty} U(X) &= v_{-\infty} \end{aligned}$$

Requires: S^ converges to stage NE, assumptions on functions converging*

Check chain rule for limits applies, might need condition on how S^ converges*

Proof: By 14, $\lim_{X \rightarrow \infty} U(X) = U_\infty$ exists. Suppose $U_\infty < v_\infty$, where v_∞ is the limit of the stage game Nash equilibrium payoff at infinity. The function $\mu_x(a^*, \bar{b}^*, X)$, with $(a^*, \bar{b}^*) = S^*(X, 0, U'(X))$, is a composite of Lipschitz continuous functions, and is therefore Lipschitz continuous. By Lemma 16,

$$\lim_{X \rightarrow \infty} U'(X)\mu_x(a^*, \bar{b}^*, X) = 0$$

Prove $\lim_{X \rightarrow \infty} g(a^*, \bar{b}^*, X) := \lim_{X \rightarrow \infty} g(S^*(X, 0, U'(X)), X) = v_\infty$

By the Lipschitz continuity of S^* , $\limsup_{X \rightarrow \infty} S^*(X, 0, U'(X)) = \limsup_{X \rightarrow \infty} S^*(X, 0, 0)$ and $\liminf_{X \rightarrow \infty} S^*(X, 0, U'(X)) = \liminf_{X \rightarrow \infty} S^*(X, 0, 0)$ (all of these values exist, given that A and B are compact. Therefore,

$$\limsup_{X \rightarrow \infty} g(S^*(X, 0, U'(X)), X) = \lim_{X \rightarrow \infty} g(S^*(X, 0, 0), X) = v_\infty$$

and $\liminf_{X \rightarrow \infty} g(S^*(X, 0, U'(X)), X) = \lim_{X \rightarrow \infty} g(S^*(X, 0, 0), X) = v_\infty$. Hence, $\lim_{X \rightarrow \infty} g(S^*(X, 0, U'(X)), X) = v_\infty$.

Plugging the above conditions in to the optimality equation yields

$$\begin{aligned}
& \limsup_{X \rightarrow \infty} \left(|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right) U''(X) \\
&= \limsup_{X \rightarrow \infty} \left(2r \left(U(X) - g(a^*, \bar{b}^*, X) \right) - 2\mu_x(a^*, \bar{b}^*, X) U'(X) \right) \\
&= 2r(U_\infty - v_\infty) < 0
\end{aligned}$$

which violates Lemma 15 since σ is Lipschitz continuous. This is a contradiction. Thus, $U_\infty = v_\infty$. The proof for the other cases is analogous.

Q.E.D.

Lemma 18. *Any bounded solution U of the optimality equation satisfies*

$$\begin{aligned}
\lim_{X \rightarrow \infty} \left| \left(|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right) U''(X) \right| &= 0 \\
\lim_{X \rightarrow -\infty} \left| \left(|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right) U''(X) \right| &= 0
\end{aligned}$$

Note this also implies $U''(X) \rightarrow 0$, since $\left(|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right) > 0$.

Proof: Applying the squeeze theorem to the optimality equation yields

$$\begin{aligned}
& \lim_{X \rightarrow \infty} \left| \left(|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X) \right) U''(X) \right| \\
&= \lim_{X \rightarrow \infty} \left| 2r \left(U(X) - g(a^*, \bar{b}^*, X) \right) - 2\mu_x(a^*, \bar{b}^*, X) U'(X) \right| \\
&= 0
\end{aligned}$$

by applying Lemmas 16 and 17.

Q.E.D.

Boundary Conditions for Bounded Support

Theorem 13. *Suppose Assumptions 1, 2, 3 and 4 hold. Any bounded solution U of the optimality equation satisfies the following boundary conditions:*

$$\begin{aligned}
\lim_{X \rightarrow \bar{X}} U(X) &= v(\bar{X}) \text{ and } \lim_{X \rightarrow \underline{X}} U(X) = v(\underline{X}) \\
\lim_{X \rightarrow \bar{X}} f(X)U'(X) &= \lim_{X \rightarrow \underline{X}} f(X)U'(X) = 0 \\
\lim_{X \rightarrow \bar{X}} f(X)^2U''(X) &= \lim_{X \rightarrow \underline{X}} f(X)^2U''(X) = 0
\end{aligned}$$

for any function $f(X)$ that is Lipschitz continuous on $[\underline{X}, \bar{X}]$ with $f(\bar{X}) = f(\underline{X}) = 0$.

The proof proceeds by a series of lemmas.

Lemma 19. *Any bounded solution U of the optimality equation has bounded variation.*

Suppose not. Then there exists a sequence $(X_n)_{n \in \mathbb{N}}$ that correspond to local maxima of U , so $U'(X_n) = 0$ and $U''(X_n) \leq 0$. Given the incentives for the long-run player, a stage Nash equilibria is played when $U'(X) = 0$, yielding flow payoff $v(X)$. From the optimality equation, this implies $v(X_n) \geq U(X_n)$. Likewise, there exists a sequence $(X_m)_{m \in \mathbb{N}}$ that correspond to local minima of U , so $U'(X_m) = 0$ and $U''(X_m) \geq 0$. This implies $v(X_m) \leq U(X_m)$. Thus, v also has unbounded variation. This is a contradiction, since $v(X) = g(S^*(X, 0), X)$ is the composite of Lipschitz continuous functions, and is therefore Lipschitz continuous.

Q.E.D.

Lemma 20. *Suppose a function $f(X)$ is Lipschitz continuous on $[\underline{X}, \bar{X}]$ with $f(\bar{X}) = f(\underline{X}) = 0$. Then any bounded solution U of the optimality equation satisfies*

1.

$$\begin{aligned} \liminf_{X \rightarrow \bar{X}} f(X)U'(X) &\leq 0 \leq \limsup_{X \rightarrow \bar{X}} f(X)U'(X) \\ \liminf_{X \rightarrow \underline{X}} f(X)U'(X) &\leq 0 \leq \limsup_{X \rightarrow \underline{X}} f(X)U'(X) \end{aligned}$$

2.

$$\begin{aligned} \liminf_{X \rightarrow \bar{X}} f(X)^2U''(X) &\leq 0 \leq \limsup_{X \rightarrow \bar{X}} f(X)^2U''(X) \\ \liminf_{X \rightarrow \underline{X}} f(X)^2U''(X) &\leq 0 \leq \limsup_{X \rightarrow \underline{X}} f(X)^2U''(X) \end{aligned}$$

1. Let $f(X)$ be a Lipschitz continuous function on $[\underline{X}, \bar{X}]$ with $f(\bar{X}) = 0$ and suppose $\lim_{X \rightarrow \bar{X}} \inf |f(X)U'(X)| > 0$. $f(X)$ is $O(\bar{X} - X)$, so there exists an $M \in \mathbb{R}$ and a $\delta_1 > 0$ such that when $|\bar{X} - X| < \delta_1$, $|f(X)| \leq M|\bar{X} - X|$. Given $\lim_{X \rightarrow \bar{X}} \inf |f(X)U'(X)| > 0$, there exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $|\bar{X} - X| < \delta_2$, $|f(X)U'(X)| > \varepsilon$. Take $\delta = \min\{\delta_1, \delta_2\}$. Then for $|\bar{X} - X| < \delta$, $|U'(X)| > \frac{\varepsilon}{|f(X)|} \geq \frac{\varepsilon}{M|\bar{X} - X|}$. Then the antiderivative of $\frac{\varepsilon}{M|\bar{X} - X|}$ is $\frac{\varepsilon}{M} \ln|\bar{X} - X|$ which diverges to $-\infty$ as $X \rightarrow \bar{X}$. This violates the boundedness of U . Therefore $\lim_{X \rightarrow \bar{X}} \inf f(X)U'(X) \leq 0$. The proof is analogous for the other cases.
2. Let $f(X)$ be a Lipschitz continuous function on $[\underline{X}, \bar{X}]$ with $f(\bar{X}) = 0$ and suppose $\lim_{X \rightarrow \infty} \inf |f(X)^2U''(X)| > 0$. $f(X)$ is $O(\bar{X} - X)$, so there exists an $M \in \mathbb{R}$ and a $\delta_1 > 0$ such that when $|\bar{X} - X| < \delta_1$, $|f(X)| \leq M|\bar{X} - X|$, and therefore, $f(X)^2 \leq M^2(\bar{X} - X)^2$. There also exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $|\bar{X} - X| < \delta_2$, $|f(X)^2U''(X)| > \varepsilon$. Take $\delta = \min\{\delta_1, \delta_2\}$. Then for $|\bar{X} - X| < \delta$, $|U''(X)| > \frac{\varepsilon}{f(X)^2} > \frac{\varepsilon}{M^2(\bar{X} - X)^2}$. The second antiderivative of $\frac{\varepsilon}{M^2(\bar{X} - X)^2}$ is $\frac{-\varepsilon}{M^2} \ln(\bar{X} - X)$ which converges to ∞ as $X \rightarrow \bar{X}$. This violates the boundedness of U . Therefore $\lim_{X \rightarrow \infty} \inf |f(X)^2U''(X)| \leq 0$. The proof is analogous for the other cases.

Q.E.D.

Lemma 21. *Suppose a function $f(X)$ is Lipschitz continuous on $[\underline{X}, \bar{X}]$ with $f(\bar{X}) = f(\underline{X}) = 0$. Then any bounded solution U of the optimality equation satisfies*

$$\lim_{X \rightarrow \bar{X}} f(X)U'(X) = \lim_{X \rightarrow \underline{X}} f(X)U'(X) = 0$$

By Lemma 20, $\lim_{X \rightarrow \bar{X}} \inf(\bar{X} - X)U'(X) \leq 0 \leq \lim_{X \rightarrow \bar{X}} \sup(\bar{X} - X)U'(X)$. Suppose, without loss of generality, that $\lim_{X \rightarrow \bar{X}} \sup(\bar{X} - X)U'(X) > 0$. Then there exist constants k and K such that $(\bar{X} - X)U'$ crosses the interval (k, K) infinitely many times as X approaches \bar{X} . Additionally, there exists an $L > 0$ such that

$$\begin{aligned} |U''(X)| &= \left| \frac{2r \left[U(X) - g(a^*, \bar{b}^*, X) \right] - 2\mu_x(a^*, \bar{b}^*, X)U'(X)}{\left| \sigma_{xy}(\bar{b}^*, X) \right|^2 + \sigma_{xx}^2(\bar{b}^*, X)} \right| \\ &\leq \left| \frac{L_1 - L_2(\bar{X} - X)U'(X)}{(\bar{X} - X)^2} \right| \\ &\leq \left| \frac{L_1 - L_2k}{(\bar{X} - X)^2} \right| = \frac{L}{(\bar{X} - X)^2} \end{aligned}$$

Therefore,

$$\begin{aligned} \left| [(\bar{X} - X)U'(X)]' \right| &\leq |U'(X)| + |(\bar{X} - X)U''(X)| \\ &= \left(1 + \left| (\bar{X} - X) \frac{U''(X)}{U'(X)} \right| \right) |U'(X)| \\ &\leq \left(1 + \frac{L}{k} \right) |U'(X)| \end{aligned}$$

where the first line follows from differentiating $(\bar{X} - X)U'(X)$ and the subadditivity of the absolute value function, the next line follows from rearranging terms, the third line follows from the bound on $|U''(X)|$ and $(\bar{X} - X)U'(X) \in (k, K)$. Then

$$|U'(X)| \geq \frac{\left| [(\bar{X} - X)U'(X)]' \right|}{\left(1 + \frac{L}{k} \right)}$$

Therefore, the total variation of U is at least $\frac{K-k}{\left(1 + \frac{L}{k} \right)}$ on the interval $(\bar{X} - X)U'(X) \in (k, K)$, which implies that U has unbounded variation near \bar{X} . This is a contradiction. Thus, $\lim_{X \rightarrow \bar{X}} \sup(\bar{X} - X)U'(X) = 0$. Likewise, $\lim_{X \rightarrow \bar{X}} \inf(\bar{X} - X)U'(X) = 0$, and therefore $\lim_{X \rightarrow \bar{X}} (\bar{X} - X)U'(X) = 0$. Then for any function $f(X)$ that is $O(\bar{X} - X)$, $|f(X)U'(X)| \leq M_1 \left| (\bar{X} - X)U'(X) \right| \rightarrow 0$, and therefore $\lim_{X \rightarrow \bar{X}} f(X)U'(X) = 0$

Q.E.D.

Lemma 22. *Let U be a bounded solution of the optimality equation. Then the limit of U converges to the limit of the stage Nash equilibrium payoffs as $X \rightarrow \{\underline{X}, \overline{X}\}$*

$$\begin{aligned}\lim_{X \rightarrow \overline{X}} U(X) &= v(\overline{X}) \\ \lim_{X \rightarrow \underline{X}} U(X) &= v(\underline{X})\end{aligned}$$

Requires: Lemmas, S^ converges to stage NE, assumptions on functions converging
Check chain rule for limits applies*

Suppose not. By 19 and the continuity of U , $\lim_{X \rightarrow \overline{X}} U(X) = U(\overline{X})$ exists. Suppose $U(\overline{X}) < v(\overline{X})$, where $v(\overline{X}) = g(S^*(\overline{X}, 0), \overline{X})$ is the static Nash equilibrium payoff at \overline{X} . The function $\mu_x(a^*, \overline{b}^*, X)$, with $(a^*, \overline{b}^*) = S^*(X, 0, U'(X))$, is a composite of Lipschitz continuous functions, and is therefore Lipschitz continuous. By Lemma 21 and the assumption that $\mu_x(a, b, \overline{X}) = 0$ for all $a, b \in A, B$,

$$\lim_{X \rightarrow \overline{X}} \mu_x(a^*, \overline{b}^*, X)U'(X) = 0$$

Plugging the above conditions in to the optimality equation yields

$$\begin{aligned}& \limsup_{X \rightarrow \overline{X}} \left(|\sigma_{xy}(\overline{b}, X)|^2 + \sigma_{xx}^2(\overline{b}, X) \right) U''(X) \\ &= \limsup_{X \rightarrow \overline{X}} \left[2r \left(U(X) - g(a^*, \overline{b}^*, X) \right) - 2\mu_x(a^*, \overline{b}^*, X)U'(X) \right] \\ &= 2r \left(U(\overline{X}) - v(\overline{X}) \right) < 0\end{aligned}$$

which violates Lemma 15 since σ is Lipschitz continuous and $\sigma(b, \overline{X}) = 0$ for all $b \in B$. This is a contradiction. Thus, $U(\overline{X}) = v(\overline{X})$. The proof for the other cases is analogous.

Q.E.D.

Lemma 23. *Any bounded solution U of the optimality equation satisfies*

$$\begin{aligned}\lim_{X \rightarrow \overline{X}} \left| \left(|\sigma_{xy}(\overline{b}^*, X)|^2 + \sigma_{xx}^2(\overline{b}^*, X) \right) U''(X) \right| &= 0 \\ \lim_{X \rightarrow \underline{X}} \left| \left(|\sigma_{xy}(\overline{b}^*, X)|^2 + \sigma_{xx}^2(\overline{b}^*, X) \right) U''(X) \right| &= 0\end{aligned}$$

Applying the squeeze theorem to the optimality equation yields

$$\begin{aligned}& \lim_{X \rightarrow \overline{X}} \left| \left(|\sigma_{xy}(\overline{b}^*, X)|^2 + \sigma_{xx}^2(\overline{b}^*, X) \right) U''(X) \right| \\ &= \lim_{X \rightarrow \overline{X}} \left| 2r \left(U(X) - g(a^*, \overline{b}^*, X) \right) - 2\mu_x(a^*, \overline{b}^*, X)U'(X) \right| \\ &= 0\end{aligned}$$

by applying Lemmas 21 and 22.

Q.E.D.

6.5.3 Uniqueness of Solution to Optimality Equation

This proof builds on a result from [Faingold and Sannikov \(2011\)](#). They prove that the optimality equation characterizing a Markov equilibrium in a repeated game of incomplete information over the type of the long-run player has a unique solution. The key element of this proof is that all solutions have the same boundary conditions when beliefs place probability 1 on the long-run player being a normal or behavioral type. This result also applies to the optimality equation characterized in this paper, given that all solutions have the same boundary conditions. An extension of this result is necessary for the case of an unbounded state space. The proof proceeds by two lemmas.

The first lemma follows directly from Lemma C.7 in [Faingold and Sannikov \(2011\)](#).

Lemma 24. *If two bounded solutions of the optimality equation, U and V , satisfy $U(X_0) \leq V(X_0)$ and $U'(X_0) \leq V'(X_0)$, with at least one strict inequality, then $U(X) \leq V(X)$ and $U'(X) \leq V'(X)$ for all $X > X_0$. Similarly if $U(X_0) \leq V(X_0)$ and $U'(X_0) \geq V'(X_0)$, with at least one strict inequality, then $U(X) < V(X)$ and $U'(X) > V'(X)$ for all $X < X_0$.*

The proof is analogous to the proof in [Faingold and Sannikov \(2011\)](#), defining

$$X_1 = \inf \{X \in [X_0, \bar{X}] : U'(X) \geq V'(X)\}$$

for $\Xi = [\underline{X}, \bar{X}]$, and

$$X_1 = \inf \{X \in [X_0, \infty) : U'(X) \geq V'(X)\}$$

for $\Xi = \mathbb{R}$.

Q.E.D.

Lemma 25. *There exists a unique solution U to the optimality equation.*

Suppose U and V are both bounded solutions to the optimality equation, and assume $U(X) - V(X) > 0$ for some $X \in \Xi$.

First consider $\Xi = \mathbb{R}$. Given that $\lim_{X \rightarrow \infty} U(X) = \lim_{X \rightarrow \infty} V(X) = v_\infty$, for all $\varepsilon > 0$, there exists a δ such that for $X \geq \delta$, $|U(X) - v_\infty| < \varepsilon/2$ and $|V(X) - v_\infty| < \varepsilon/2$. Then for $X \geq \delta$, $|U(X) - V(X)| < \varepsilon$.

Take an interval $X \in [X_1, X_2]$, and suppose $U(X) > V(X)$ for some $X \in [X_1, X_2]$. Let X^* be the point where $U - V$ is maximized, which is well-defined given U and V are continuous functions on a compact interval. Suppose the maximum occurs at an interior point $X^* \in (X_1, X_2)$. Then $U'(X^*) = V'(X^*)$. By Lemma 24, $U'(X) \geq V'(X)$ for all $X > X^*$, and this difference is strictly increasing, a contradiction. Suppose the maximum occurs at an endpoint, $X^* = X_2$, and let $U(X_2) - V(X_2) = M > 0$. Then it must be the case that $U'(X_2) \geq V'(X_2)$. By Lemma 24, $U'(X) \geq V'(X)$ for all $X > X_2$, and this difference is strictly increasing for $X > X_2$. But then for $\varepsilon < M$, there does not exist a δ such that $|U(X) - V(X)| < \varepsilon$ when $X > \delta$. This violates the boundary condition. The argument is analogous if the maximum occurs at $X^* = X_1$. Thus, it is not possible to have $U(X) > V(X)$.

The proof for $\Xi = [\underline{X}, \bar{X}]$ is similar, using $[X_1, X_2] = [\underline{X}, \bar{X}]$, and the fact that the boundary conditions at $[\underline{X}, \bar{X}]$ ensure the point where $U - V$ is maximized is an interior point.

Q.E.D.

6.5.4 Uniqueness of PPE

Lemma 26. *When there is a unique solution U to the optimality equation, the continuation values and equilibrium actions specified by U characterize the unique PPE.*

When there is a unique solution to the optimality equation, it is obvious to see that there is a unique Markov equilibrium (Theorem 1 showed that in any Markov equilibrium, continuation values and equilibrium actions must be characterized by the optimality equation). It remains to show that there are no other PPE.

From Theorem 2, any PPE payoff must lie in the convex hull of the Markov equilibrium payoffs. When there is a unique Markov equilibrium, this implies that in any PPE with continuation values $(W_t)_{t \geq 0}$, it must be the case that $W_t = U(X_t)$ for all t . Therefore, it must be that the volatility of the two continuation values is equal ($|f(\bar{b}, X, \beta)| = 0$). and actions are uniquely specified by $S^*(X, 0, U'(X))$.

Q.E.D.

6.6 Proofs: Equilibrium Payoffs

Theorem 14. *The highest PPE payoff across all states is bounded above by the highest static Nash equilibrium payoff across states, $\bar{W} \leq \bar{v}^*$ and the lowest PPE payoff across all states is bounded below by the lowest static Nash equilibrium payoff across states $\underline{W} \geq \underline{v}^*$.*

Let $\bar{W} = \sup_{\Xi} U(X)$ be the highest PPE payoff across all states. Suppose $\bar{W} = U(X)$ occurs at an interior point. Then $U'(X) = 0$ and $U''(X) \leq 0$. From the optimality equation,

$$U''(X) = \frac{2r [\bar{W} - v(X)]}{|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X)} \leq 0$$

and therefore $\bar{W} \leq v(X) \leq \bar{v}^*$. Suppose the state space is bounded and $\bar{W} = U(X)$ occurs at an endpoint. Suppose, without loss of generality, that \bar{W} occurs at \bar{X} . Then by the boundary conditions, $\bar{W} = v(\bar{X}) \leq \bar{v}^*$. Suppose the state space is unbounded and there is no interior maximum with $U(X) = \bar{W}$. Then $U(X)$ must converge to \bar{W} at either ∞ or $-\infty$. Suppose $\lim_{X \rightarrow \infty} U(X) = \bar{W}$. Then $\bar{W} = v_{\infty} \leq \bar{v}^*$. The proof for $\underline{W} \geq \underline{v}^*$ is analogous.

Q.E.D.

Theorem 15. *Assume 1, 2, 3 and 4. The following properties show how PPE payoffs change with the state variable:*

1. *Suppose $v(X)$ is increasing (decreasing) in X . Then $U(X)$ is also increasing (decreasing) in X . The states that yield the highest and lowest PPE payoffs are boundary points.*
2. *Suppose $v(X)$ is concave and has a unique interior maximum X^* . Then $U(X)$ is concave and has a unique interior maximum. The state that yields the lowest PPE payoff is a boundary point.*

3. Suppose $v(X)$ is convex and has a unique interior minimum X^* . Then $U(X)$ is convex and has a unique interior minimum. The state that yields the highest PPE payoff is a boundary point.

1. Suppose $v(X)$ is increasing in X , but $U(X)$ is not increasing in X . Thus, $U'(X) < 0$ for some $X \in \Xi$. Let $(X_1, X_2) \subset \Xi$ be a maximal subinterval such that $U'(X) < 0$ for all $X \in (X_1, X_2)$. Note $\lim_{X \rightarrow -\infty} U(X) = \underline{v}^* \leq \lim_{X \rightarrow \infty} U(X) = \bar{v}^*$ since v is increasing in X , so $U'(X)$ is not strictly decreasing on Ξ . Without loss of generality, assume $U(X)$ is increasing on $(-\infty, X_1)$. Then X_1 is an interior local maximum with $U'(X_1) = 0$ and $U''(X_1) \leq 0$. Then by the optimality equation,

$$U''(X_1) = \frac{2r}{|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X)} (U(X_1) - v(X_1)) \leq 0$$

which implies $U(X_1) \leq v(X_1)$. Then

$$\lim_{X \rightarrow X_2} U(X_2) < U(X_1) \leq v(X_1) \leq \bar{v}^* = \lim_{X \rightarrow \infty} U(X)$$

Thus, since $U(X_2) < U(X_1)$ by definition, it must be that $X_2 < \infty$ and X_2 is a local minimum with $U'(X_2) = 0$ and $U''(X_2) \leq 0$. Then by the optimality equation,

$$\begin{aligned} U''(X_2) &= \frac{2r}{|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X)} (U(X_2) - v(X_2)) \\ &\leq \frac{2r}{|\sigma_{xy}(\bar{b}, X)|^2 + \sigma_{xx}^2(\bar{b}, X)} (U(X_2) - v(X_1)) < 0 \end{aligned}$$

which implies X_2 is a local maximum. This is a contradiction. The proof for $U(X)$ decreasing is analogous.

2. If $v(X)$ has a unique interior maximum \hat{X} such that $v'(X) > 0$ for $X < \hat{X}$ and $v'(X) < 0$ for $X > \hat{X}$. Assume $U'(X) < 0$ for some $X < \hat{X}$. Let $(X_1, X_2) \subset (-\infty, \hat{X})$ be a maximal subinterval such that $U'(X) < 0$ for all $X \in (X_1, X_2)$. First suppose $X_1 > -\infty$ and $X_2 < \hat{X}$. Then X_1 is a local maximum with $U'(X_1) = 0$ and $U''(X_1) \leq 0$, and by the optimality equation, $U(X_1) \leq v(X_1)$. Also, X_2 is a local minimum with $U'(X_2) = 0$ and $U''(X_2) \geq 0$, and by the optimality equation, $U(X_2) \geq v(X_2)$. This implies:

$$U(X_1) \leq v(X_1) \leq v(X_2) \leq U(X_2)$$

which is a contradiction. Thus, (X_1, X_2) must include a boundary point of $(-\infty, \hat{X})$. Next suppose U is decreasing over $(-\infty, X_2)$ and $X_2 < \hat{X}$. Thus, X_2 is a local minimum. Given that $\lim_{X \rightarrow -\infty} U(X) = \lim_{X \rightarrow -\infty} v(X)$, this implies:

$$\lim_{X \rightarrow -\infty} U(X) = \lim_{X \rightarrow -\infty} v(X) \leq v(X_2) \leq U(X_2)$$

which is a contradiction. Thus, it can only be that $(X_1, X_2) = (-\infty, \hat{X})$. Likewise, if $U'(X) > 0$ for some $X > \hat{X}$, then it must be the case that any maximal subinterval (X_1, X_2) on which $U'(X) > 0$ is $(X_1, X_2) = (\hat{X}, \infty)$.

Suppose $U'(X) < 0$ on $(-\infty, \hat{X})$ and $U'(X) > 0$ on $(X_1, X_2) = (\hat{X}, \infty)$. Then \hat{X} is a global minimum, which implies:

$$U(X) > U(\hat{X}) \geq v(\hat{X}) = \bar{v}^*$$

But this is a contradiction, since $U(X) \leq \bar{v}^*$ for all X .

3. The proof is analogous to part 2.

Q.E.D.