Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 14-033

"Stochastic Games in Continuous Time: Persistent Actions in Long-Run Relationships", Second Version

by

J. Aislinn Bohren

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

J. Aislinn Bohren†

This version: August 2014
First version: November 2011

Abstract

This paper studies a class of continuous-time stochastic games in which the actions of a long-run player have a persistent effect on payoffs. For example, the quality of a firm's product depends on past as well as current effort, or the level of a policy instrument depends on a government's past choices. The long-run player faces a population of small players, and its actions are imperfectly observed. I establish the existence of Markov equilibria, characterize the Perfect Public Equilibria (PPE) payoff set as the convex hull of the Markov Equilibria payoff set, and identify conditions for the uniqueness of a Markov equilibrium in the class of *all* PPE. The existence proof is constructive: it characterizes the explicit form of Markov equilibria payoffs and actions, for any discount rate. Action persistence creates a crucial new channel to generate intertemporal incentives in a setting where traditional channels fail, and can provide permanent non-trivial incentives in many settings. These results offer a novel framework for thinking about reputational dynamics of firms, governments, and other long-run agents.

KEYWORDS: Continuous Time Games, Stochastic Games, Reputation
JEL: C73, L1

*I thank Matt Elliott, Jeff Ely, Ben Golub, Alex Imas, Bart Lipman, David Miller, George Mailath, Markus Mobius, Paul Niehaus, Yuliy Sannikov, Andy Skrzypacz, Joel Sobel, Jeroen Swinkels, Joel Watson, Alex Wolitzsky and especially Nageeb Ali for useful comments. I also thank numerous seminar participants for helpful feedback.

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

Persistence and rigidities are fundamental features of many economic settings. Key characteristics of these environments depend not only on actions taken today, but are also directly influenced by choices made in the past. Examples abound of such settings. 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 does not 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 the impact of persistence in a class of continuous-time stochastic games with imperfect public monitoring between a long-run player and a population of small players. Past actions influence a state variable, which captures their persistent effect on the information and payoff structure of the game. Small players learn about the long-run player's action through a noisy public signal, whose precision may depend on the state variable. The observable state variable and public signal both evolve stochastically, perturbed by Brownian information. A stochastic game introduces a second channel for intertemporal incentives: in addition to conditioning future equilibrium play on past outcomes, as is possible in standard repeated games, players' actions now directly impact feasible payoffs and the information structure in future periods.

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 payoff set, and determine when a Markov equilibrium is unique. The existence proof is constructive, and specifies the explicit form of behavior and payoffs in Markov equilibria, for any discount rate, as a function of the state variable. Equilibrium continuation values are specified as the solution to a *nonstochastic* differential equation, defined over the state space, while the long-run player's action is determined by the sensitivity of its continuation value to changes in the state variable (the first derivative of this solution).

The key contributions of this paper are along three dimensions. The innovative technical work lies in deriving the main results for an unbounded state space and an unbounded flow payoff (the instantaneous payoff of the long-run player). With an unbounded flow payoff, the continuation value may also be unbounded; the first challenge lies in establishing that the continuation value has linear growth with respect to the state under the proper assumption on the growth rate of flow payoffs and the state. This is a necessary step to characterize conditions for sequential rationality and the evolution of the continuation value in any PPE of the stochastic game. To establish existence of a Markov equilibrium, I need to find upper and lower solutions for an unbounded ordinary differential equation and show that these solutions have linear growth. Characterization of the PPE payoff set requires bounding the difference between continuation values in any two PPE, despite the fact that continuation values may be unbounded. Finally, deriving sufficient conditions for uniqueness of a Markov equilibrium requires establishing that the slope of the continuation value with respect to the state converges to the same value k in any Markov equilibrium. Given this and the linear growth condition, we show that there is also a unique intercept b such that the continuation value in any Markov equilibrium converges to $kX+b$. Note that when $k \neq 0$, incentives can be maintained near the boundary of the state space, a contribution in itself. Unique boundary conditions yield a unique Markov equilibrium across the entire state space. The general state and payoff space can capture many applications, including models where equilibrium payoffs are not monotonic with respect to the state variable and models where persistence generates permanent (as opposed to temporary) incentives. Thus, I demonstrate that the insights first presented in Faingold and Sannikov (2011) apply to a very broad set of games.

Conceptually, I demonstrate that action persistence creates a crucial effective channel to generate intertemporal incentives in a setting where traditional channels fail. Faingold and Sannikov (2011) and Fudenberg and Levine (2007) show that in continuous time repeated games with a single long-run player and Brownian information, the long-run player cannot attain payoffs beyond those of the highest static Nash equilibrium payoff.[1] This paper establishes that when actions have a persistent effect on payoffs, it is possible to create effective intertemporal incentives and the long-run player can earn payoffs higher than the static Nash. At first glance, there appears to be a tension between my result and the rich literature examining incentives in continuous-time repeated games with imperfect monitoring (initiated by Abreu, Milgrom, and Pearce (1991)). Sannikov and Skrzypacz (2010) show that Brownian information must be used linearly; burning value through punishments that affect *all players* does not provide effective incentives. Additionally, in many settings, including those between long-run and small, anonymous players, tangential transfers along the boundary of the equilibrium payoff set are not possible.[2] In a repeated game, using Brownian information linearly and non-tangentially will result in the continuation value escaping its upper bound with positive probability, a contradiction. Thus, it is not possible to structure incentives

[1]This insight generalizes to continuous time stochastic games *without* persistent actions (see Corollary 1 in Section 3.2.

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

at the long-run player's highest continuation payoff, and intertemporal incentives collapse. However, a stochastic game introduces the possibility of using linear, non-tangential transfers for some values of the state variable. This important feature introduces the possibility of a non-trivial linear incentive structure.

With an unbounded flow payoff, it is possible to provide effective incentives as the state variable becomes large; the same holds true for a bounded flow payoff if the static Nash payoff perpetually oscillates. In both cases, the slope of the continuation value doesn't converge to zero and non-trivial incentives can be sustained. On the other hand, when the flow payoff is bounded and the static Nash payoff is monotonic, the slope of the continuation value converges to zero and incentives dissipate as the state grows large. Similarly, when the state space is bounded, it is possible to sustain non-trivial incentives as the state variable approaches the boundary if the state is not absorbed. The ability to sustain non-trivial intertemporal incentives across the entire state space is an important and novel insight of this paper.

Finally, these results make a practical contribution for equilibrium analysis. Many dynamic interactions depend on some type of state variable, and Markov equilibria are a popular concept in applied work; advantages include their simplicity and dependence on payoff relevant variables to specify incentives. I develop a common framework and a rich set of tools that these applied models can utilize to shed light on equilibrium behavior. The characterization of equilibrium actions and payoffs connects the value of a firm, government or other long-run player to choices that are empirically identifiable. The value of past actions depends on the nature of persistence and the structure of the market, including observable parameters such as the cost and depreciation rate of investment, the volatility of quality or the relative value that consumers derive from cheap, low-quality products compared to expensive high-end products. After stipulating functional forms to apply the model to a specific economic setting, it is straightforward to derive the Markov equilibrium, calibrate it with realistic parameters, and use numerical methods to estimate the solution. This will generate empirically testable predictions about the dynamics of equilibrium behavior, including: when does a firm or government build its quality and when does it allow it to decay; when does a firm maintain high quality in the long-run; what drives a firm to specialize in different quality products or an expert to seek the next level of a verifiable certification? Establishing that it is not possible to achieve payoffs above the best Markov equilibrium payoff offers a strong justification for focusing on this more tractable class of equilibria.

To fix ideas, consider the canonical product choice game. A long-run firm interacts 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 (1989) and 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 2004).[3] Commitment types connect behavior across time, and enable the player to overcome binding moral hazard, at least in the short-run. In fact, these models can be viewed as a type of stochastic game, since the belief about the player's type depends on its past actions (as well as beliefs about these 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 by the possibility that players may be non-strategic. This paper returns to a notion of reputation as a type of capital (Klein and Leffler 1981), "a word which denotes the persistence of quality and [...] commands a price (or exacts a penalty) because it economizes on search" (Stigler 1961). Consider a simple and realistic modification of the product choice setting in which the firm's product quality is a noisy function of current and past effort. Applying my results to this application establishes that the firm also overcomes moral hazard; importantly, this effect can remain in the long-run and incentives are *permanent*. Therefore, it is the persistent effect of actions, rather than the particular form of persistence as a commitment type, that drives the power of reputation models. Additionally, the general framework of this paper applies to a broader set of environments, including settings in which the value of quality is non-monotonic (firms specialize in high or low quality, governments formulate different policy targets), a firm has different price-taking behavior (monopolist or fixed prices), or a student invests in a series of verifiable certifications.

This paper builds on tools developed by Faingold and Sannikov (2011). They characterize the unique Markov equilibrium in an incomplete information game in which the state variable is the belief that the firm is a commitment type, the transition function for the state variable

[3]Since Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982) and Milgrom and Roberts (1982), a player's reputation has often been modeled as the belief that others have that it is a commitment type and takes a fixed action in each period. Mailath 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.

follows Bayes rule, the state variable evolves as a function of the public signal, and the firm's payoffs do not directly depend on the state. I look at a complete information setting in which both the long-run and small players may have state-dependent preferences, the state space may be bounded or unbounded and have multiple interior absorbing states, the flow payoff of the long-run player may be bounded or unbounded, the transition rule guiding the evolution of the state takes a general form and the state may be independent of the public signal (this isolates the impact of persistence). Unlike in Faingold and Sannikov (2011), the state variable in this paper does not depend on beliefs about the long-run player's equilibrium action. Cisternas (2014) also builds on Faingold and Sannikov (2011) in a game where the long-run player can either manipulate a hidden state itself or the public signal about the state. He characterizes equilibrium incentives in the general setting, and obtains a closed-form solution for these incentives in a subclass of linear-quadratic games.

Board and Meyer-ter vehn (2013) also look at persistent quality as an alternative way to generate reputation effects in a setting similar to the product choice example. Product quality is binary and reputation is defined as the consumers' belief that the current product quality is high. Consumers learn about quality through noisy signals, which is periodically replaced; when a replacement occurs, the monopolist's current effort determines the new quality. The value of reputation stems from the higher price that the monopolist can charge when the consumer has higher beliefs about quality. In contrast, in the product choice application of this paper, quality is observable and the instantaneous return to quality need not be linear. Technically speaking, realized quality in Board and Meyer-ter vehn (2013) is discontinuous (jumping between low and high), which plays an important role in determining intertemporal incentives, while quality in the product choice application of this paper is a smooth function of past investments. Thus, the analysis uses different techniques.

The paper relates to a broad literature on stochastic games. Sobel (1973), examine equilibrium existence in continuous-time stochastic games with perfect monitoring, while Flesch, Thuijsman, and Vrieze (2003) study existence in discrete time stochastic games with imperfect monitoring. Dutta and Sundaram (1992) establish the existence of Markov equilibria in a discrete-time stochastic growth model with two long-run players. Several folk theorems exist for discrete time irreducible stochastic games with an observable state, beginning with a perfect monitoring setting in Dutta (1995), and extending to imperfect monitoring environments in Fudenberg and Yamamoto (2011) and Hörner, Sugaya, Takahashi, and Vieille (2011). The technique in this paper also relates to methods established by Strulovici and Szydlowski (2014) to ensure a well-behaved value function for optimal control problems with a diffusion process.

Many applied papers analyze equilibria in continuous time games with a state variable.

Ericson and Pakes (1995) study Markov perfect equilibrium in a setting where uncertain investment has a persistent impact on profits, which is similar in spirit to this paper's quality example (Section 2.1). Back (1992) looks at an asset pricing model with an insider trader, Daley and Green (2012) study a dynamic asymmetric information setting where the value of a seller's asset is gradually revealed to buyers, Murto (2004) looks at firm exit decisions in a duopoly setting with uncertain profit streams and Papageorgiou (2013) study how workers learn about their types.

The organization of this paper proceeds as follows. Section 1 sets up the model and Section 2 introduces several simple examples to illustrate potential applications. Section 3 presents the three main results: (1) establishing the existence of a Markov equilibrium, (2) characterizing the PPE payoff set and (3) establishing when a Markov equilibrium is unique in the class of all PPE; the section concludes with a discussion of the structure of intertemporal incentives. In Section 4, the paper returns to the examples to illustrate the main results of the model. A final section relates the shape of equilibrium payoffs to the structure of the game. Most proofs are in the Appendix.

1 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, Information 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 is a compact subset of a Euclidean space and B is a closed subset 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 $X_t \in \Xi$, which determines the payoff and information structure. Assume that the state space Ξ is a closed subset of $\mathbb{R}$. The state variable begins at initial value X_0 and 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,

$$\begin{bmatrix} dY_t \\ dX_t \end{bmatrix} = \begin{bmatrix} \mu_y(a_t, \bar{b}_t, X_t) \\ \mu_x(a_t, \bar{b}_t, X_t) \end{bmatrix} dt + \begin{bmatrix} \sigma_y\left(\bar{b}_t, X_t\right) & \sigma_{yx}\left(\bar{b}_t, X_t\right) \\ \sigma_{xy}\left(\bar{b}_t, X_t\right) & \sigma_{xx}\left(\bar{b}_t, X_t\right) \end{bmatrix} \cdot \begin{bmatrix} dZ_t^y \\ dZ_t^x \end{bmatrix} \tag{1}$$

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

$$\sigma = \begin{bmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{bmatrix} = \begin{bmatrix} \sigma_y \\ \sigma_x \end{bmatrix},$$

with each function linearly extended to $A \times \Delta B$. Assume μ_y, μ_x and σ are Lipschitz continuous. The state space may or may not be bounded. It is bounded above if there exists an upper bound $\overline{X} < \infty$ at which the volatility is a vector of zeros, $\sigma_x(b, \overline{X}) = \mathbf{0}$ and the drift is weakly negative, $\mu_x\left(a, b, \overline{X}\right) \leq 0$, for all $(a, b) \in A \times B$; it is bounded below if there exists a lower bound $\underline{X} > -\infty$ such that the volatility is zero and the drift is weakly positive for all $(a, b) \in A \times B$.

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 public information, $(Y_t, X_t)_{t\geq 0}$.

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

Definition 1. *$\widetilde{X} \in \Xi$ is an absorbing state if $\mu_x(a, b, \widetilde{X}) = 0$ and $\sigma_x\left(b, \widetilde{X}\right) = \mathbf{0}$ for all $(a, b) \in A \times B$.*

Denote the set of absorbing states as $\widetilde{\Xi}$. I make several assumptions on the structure of public signals and the state. First, the volatility of the state variable is positive at all but a finite number of points of the state space, and any interior point of the state space that has zero volatility is an absorbing state. 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 subsets, $\Xi = \cup_{i=1}^n \Xi_i$ such that for any compact proper subset $I \subset \Xi_i$, there exists a c_I with:*

$$\sigma_I = \inf_{b\in B, X\in I} \left|\sigma_x(b, X)\right|^2 > c_I.$$

If $X \in int(\Xi)$ and $\left|\sigma_x(b, X)\right|^2 = 0$ for some $b \in B$, then X is an absorbing state.

[4] The 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. I assume that the rows of the public signal volatility matrix, σ_y, are linearly independent. Additionally, when the volatility of the state variable is a linear combination of the volatility of the public signals, then the drift of the state variable is the same linear combination of the drift of the public signals.

Assumption 2. *There exists a constant c such that*

$$|\sigma_{yy}(b,X)\cdot y| \geq c\,|y|$$

for all $y \in \mathbb{R}^{d-1}$, $X \in \Xi$ and $b \in B$. If there exists an $X \in \Xi$ and $b \in B$ and scalars $\alpha_1, ..., \alpha_{d-1}$ such that $\sigma_x(b,X) = \sum \alpha_i \sigma_y^i(b,X)$, then there exist functions f_1 and f_2 such that for all $a \in A$,

$$\begin{aligned} f_1(a,b,X) &= \sum \alpha_i \mu_y^i(a,b,X) \\ \mu_x(a,b,X) &= f_1(a,b,X) + f_2(b,X) \end{aligned}$$

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 $\mu_x(a,b,X)$ 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 value of its 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. The dependence of payoffs on the state variable creates a form of action persistence for the long-run player, 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 players ($g(a, \bar{b}, X) = \widetilde{g}(a, \bar{b})$).

Short-run players have identical preferences: they maximize the expected flow payoff, $h(a, b^i, \bar{b}, X)$, which is continuous.[5] Ex-post payoffs can only depend on a_t through public

[5] If both g and h are independent of X, then the game reduces to a standard repeated game.

information, as is standard in games of imperfect monitoring.[6] Let $\mathcal{B} : A \times \Delta B \times \Xi \rightrightarrows B$ represent the static best response correspondence for the short-run player, which maps an action profile and a state to the set of short-run player actions that maximize the expected flow payoff, and $\overline{\mathcal{B}} : A \times \Xi \rightrightarrows \Delta B$ represent the aggregate best response correspondence. Assume that these correspondences are well-defined in the sense that they are non-empty and if $b \in \mathcal{B}(a, \bar{b}, X)$ ($\bar{b} \in \overline{\mathcal{B}}(a, X)$), then $|b| < \infty$ ($|b| < \infty$ for each $b \in supp \bar{b}$).[7] In many applications, it will be sufficient to specify the aggregate best response correspondence as a reduced form for short-run players' behavior.

The next assumption requires that, for any action of the long-run player and best response of the short-run players, either the state grows at a rate slower than the discount rate, or the flow payoff is bounded with respect to the state. This assumption will guarantee that the expected discounted payoff of the long-run player is well-behaved.

Assumption 3. *At least one of the following conditions hold:*

1. *There exists a $k \in [0, r)$ and $c > 0$ such that, for all $X \geq 0$, $\mu_x(a, b, X) \leq kX + c$ for all $a \in A, b \in B$ and for all $X \leq 0$, $\mu_x(a, b, X) \geq kX - c$ for all $a \in A, b \in B$.*

2. *$g(a, b, X)$ is bounded.*

This assumption is trivially satisfied when the state space is bounded, by Lipschitz continuity.

Strategies and equilibrium: A public pure strategy for the long-run player is a stochastic process $(a_t)_{t \geq 0}$ with $a_t \in A$ and progressively measurable with respect to $\{F_t\}_{t \geq 0}$. Likewise, a public pure strategy for a short-run player is an action $b_t^i \in B$ progressively measurable with respect to $\{F_t\}_{t \geq 0}$. Given that small players have identical preferences, it is without loss of generality to work with aggregate strategies $(\bar{b}_t)_{t \geq 0}$. I restrict attention to pure strategy perfect public equilibria (PPE).

Definition 2. *A perfect public equilibria consists of public strategy profile $S = (a_t, \bar{b}_t)_{t \geq 0}$ such that after any public history and for all $t \geq 0$, (i) $(a_t)_{t \geq 0}$ maximizes the discounted expected payoff of the long-run player, and (ii) $b \in \arg\max_{b' \in B} h(a_t, b', \bar{b}_t, X_t)$ for all $b \in supp\ \bar{b}_t$.*

Timing: 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.

[6] Alternatively, 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, or a continuum of long run small, anonymous players, since the individual action of a short run player has a negligible impact on the aggregate action (and therefore X_t) and is not observable by the long run player.

[7] Assumptions on the primitives of the game that guarantee a well-defined best responce correspondence include (i) B is compact; (ii) for each X, the short-run players are restricted to selecting actions from a compact set B_X. Note that we do not make the standard assumption that B is compact in order to allow for the possibility that the optimal action b^* is not uniformly bounded with respect to X.

1.1 Equilibrium Structure

This section extends a recursive characterization of PPE to continuous time stochastic games. In dynamic games, a long-run player's incentives are created through the evolution of his continuation value, which is the expected discounted payoff at time t given the public history. The long-run player's action impacts 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 game payoffs. Each channel provides a potential source of intertemporal incentives.

Formally, define the continuation value as the expected value of the future discounted payoff at time t, given the public information contained in $\{F_t\}_{t\geq 0}$ and strategy profile $S = (a_t, \bar{b}_t)_{t\geq 0}$, as:

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

and the expected value of the long-run player's average discounted payoff at time t as:

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

First, I establish that the continuation value is finite whenever X_t is finite, and the expected value of the average discounted payoff is a martingale. This implies that $E|V_t(S)| < \infty$ for all $t \geq 0$.

Lemma 1. *Assume Assumption 3. For any public strategy profile $S = (a_t, \bar{b}_t)_{t\geq 0}$, initial state $X_0 < \infty$ and path of the state variable $(X_t)_{t\geq 0}$ that, given S, evolves according to (1), (i) there exists a $K > 0$ such that $|W_t(S)| \leq K(1 + |X_t|)$ for all $t \geq 0$, and (ii) $V_t(S)$ is a martingale.*

Given Assumption 3, either the slope of the drift with respect to X is less than the discount rate for large $|X|$, or the flow payoff of the long-run player is bounded. This ensures that $W_t(S)$ doesn't diverge, either because X_t grows large at a slow enough rate relative to the discount rate, or because large X_t doesn't lead to unboundedly large payoffs. Note that no lower bound is necessary on the slope of μ_x, since a negatively sloped drift either eventually pulls the state variable back to zero or asymptotes to a constant drift. This result is similar in spirit to Lemma 1 in Strulovici and Szydlowski (2014), which establishes that the value function of an optimal control problem is finite and satisfies a linear growth condition with respect to the state, and is not required in models with a bounded continuation value.

The next lemma characterizes the evolution of the continuation value and the long-run player's incentive constraint in a PPE. This is similar in spirit to Theorem 2 in Faingold and Sannikov (2011), with the requisite changes to allow for an unbounded state space and flow payoff.

Lemma 2. *Assume Assumption 3. 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. *Given $(\beta_t)_{t\geq 0}$, $(W_t)_{t\geq 0}$ satisfies:*

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

   If g or Ξ is bounded, then $(W_t)_{t\geq 0}$ is bounded; otherwise, there exists a $K > 0$ such that $|W_t| \leq K(1 + |X_t|)$ for all $t \geq 0$.

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

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

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. The 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. The 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 the flow payoff 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.

Use the condition for sequential rationality, I specify an auxiliary static game parameterized by the state variable and the volatility of the continuation value.

Definition 3. *Let $S^* : \Xi \times \mathbb{R}^d \rightrightarrows A \times \Delta B$ denote the correspondence of static Nash equilibrium action profiles in the auxiliary game parameterized by (X, z_y, z_x):*

$$S^*(X, z_y, z_x) := \left\{ (a, \overline{b}) : \begin{array}{c} a \in \arg\max_{a'} g(a', \overline{b}, X) + z_y \cdot \mu_y(a', \overline{b}, X) + z_x \mu_x(a', \overline{b}, X) \\ \overline{b} \in \overline{\mathcal{B}}(a, X) \end{array} \right\}.$$

In any PPE strategy profile $(a_t, \overline{b}_t)_{t \geq 0}$ of the stochastic game, given some processes $(X_t)_{t>0}$ and $(\beta_t)_{t>0}$, the action profile at time t must be a static Nash equilibrium of the auxiliary game, $(a_t, \overline{b}_t) \in S^*(X_t, \beta_{yt}, \beta_{xt})$. I assume that this auxiliary game has a unique static Nash equilibrium with an atomic distribution over small players' actions.

Assumption 4. *Assume $S^*(X, z_y, z_x)$ is non-empty and single-valued for all $(X, z_y, z_x) \in \Xi \times \mathbb{R}^d$, Lipschitz continuous on any subset of $\Xi \times \mathbb{R}^d$, and the small players choose identical actions, $b^i = \overline{b}\ \forall i$.*

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.

1.1.1 Discussion

A key feature of the equilibrium structure 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 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 the coordination of equilibrium play, the set of feasible payoffs and the information structure of public signals.

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 $\widetilde{a}_t$ at time t. The continuation value under both strategies is a non-degenerate expectation with respect to the future path of the state variable. Given that the same set of paths are possible under a_t and $\widetilde{a}_t$, the change in the continuation value when the long-run player deviates depends on the different probability measure that the deviation induces over the the state variable. This change is linear with respect to the difference in the drift of the public signal and state variable, $\mu_Y\left(\widetilde{a}_t, \overline{b}_t, X_t\right) - \mu_Y\left(a_t, \overline{b}_t, X_t\right)$ and $\mu_X\left(\widetilde{a}_t, \overline{b}_t, X_t\right) - \mu_X\left(a_t, \overline{b}_t, X_t\right)$, due to the linear structure of Brownian information.

It is of interest to note that it is precisely this linear structure with respect to the Brownian information, coupled with the inability to transfer 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 (Faingold and Sannikov 2011; Fudenberg and Levine 2007; Sannikov and Skrzypacz 2010). Short-run players preclude incentive structures with transfers that are tangential to the boundary of the payoff set. Using Brownian information to structure incentives with non-tangential linear transfers results in the continuation value escaping the boundary of the payoff set with positive probability, a contradiction. And non-linear incentive structures, such as value-burning, are ineffective, because the expected losses from false punishment exceed the expected gains from cooperating.

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 repeated game. However, at other values of the state variable, the continuation value linearly depends on Brownian information, which creates effective intertemporal incentives.[8]

2 Examples

This interlude presents several 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 diverse settings. At the end of Section 3, I will return to these examples to characterize payoffs and actions in the unique Markov equilibrium.

2.1 Persistent Effort as a Source of Reputation

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 $a_t \in [0, \overline{a}]$ and consumers

[8] See Section 3.4.1 for a more detailed discussion.

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 persistent component of quality, X_t, captures the impact of past effort. This stock evolves according to a mean-reverting stochastic process,

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

with drift $\mu_x = \theta\left(a - X\right)$ proportional to the difference between current effort and stock quality, constant volatility σ, Brownian noise dZ and persistence $\theta < r$.[9] Stock quality is increasing in expectation when effort exceeds the stock, and decreasing when effort is below the stock. 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 $\left(a_s\right)_{s\leq t}$ 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,

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

with λ to capture the relative importance of current and past effort in determining current quality. Consumers enjoy high quality products, exhibit diminishing returns in the purchase level, and the price is fixed at unity: $h(a, b^i, \bar{b}, X) = \sqrt{b^i} q(a, X) - b^i$.[10] The firm's instantaneous profit is the difference between revenue and investment expenses, $g(a, \bar{b}, X) = \bar{b} - ca^2$ for $c < 1$. Note that this example satisfies the necessary continuity, positive volatility and imperfect monitoring assumptions from Section 1.

Applying Lemma 2 and Definition 3, the best response correspondence, given incentive

[9] This is an Ornstein-Uhlenbeck process.

[10] An alternative version of this example could allow the firm to set a quality-dependent price.

weight $z_x = z$, is:[11]

$$a(X,z) = \begin{cases} 0 & \text{if } z \leq 0 \\ \min\{\frac{z\theta}{2c}, \overline{a}\} & \text{if } z > 0 \end{cases}$$

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

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. In equilibrium, investment is chosen to equate the marginal cost of effort with its expected marginal benefit. Effort increases with the incentive weight, persistence and patience. 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; 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. Consumers myopically optimize flow payoffs by choosing a purchase level such that the marginal value of an additional unit of product is equal to the price. They are willing to purchase more when they expect higher quality. As Assumption 4 requires, this correspondence is single-valued for all (X,z) and Lipschitz continuous.

The firm is subject to binding moral hazard in that it would like to commit to a higher level of effort in order to increase demand, but in the absence of such a commitment, will deviate to low effort. In the static game, the firm always chooses the lowest possible effort level, $a(X,0) = 0$, regardless of the stock quality X, and this is also the unique equilibrium strategy of the repeated game with no persistence ($\theta = 0$). This example seeks to characterize when quality persistence can provide the firm with intertemporal incentives to choose a positive level of effort.

2.1.1 Unbounded Demand

In the preceding example, although quality is unbounded, consumer's have a saturation purchase level of $\overline{b} = 20$. The model can also accommodate unbounded demand. Suppose that the persistent component of quality and the firm's flow payoff are the same as Section

[11]Note z_y is irrelevant, since there is no public signal.

2.1, but now consumers have a different value for quality,

$$h(a, b^i, \bar{b}, X) = \begin{cases} -b^i & \text{if } q(a, X) < 0 \\ q(a, X) \ln b^i - b^i & \text{if } q(a, X) \geq 0 \end{cases}$$

and can purchase as much as they like, $b^i \in [0, \infty]$. The best response correspondence is the same as before for the firm, and is now

$$\bar{b}(X, z) = \max \{q(a(X, z), X), 0\}$$

for the consumer.

Another interpretation of this model is that the consumer demands one unit of the good, has a linear value for quality and b^i represents price, $h(a, b^i, \bar{b}, X) = \max \{q(a(X, z), X), 0\} - b^i$. The firm has monopoly power and sets a price equal to expected quality, capturing all of the surplus.

2.1.2 Building a Specialization with Oscillating Demand

Consider another variation of the quality example where a firm faces a revenue curve that oscillates with quality. For example, there is a large market for cheap, low quality or expensive, high quality goods, but the market for intermediate levels of quality is thin. Consumers choose an amount to spend $b \in [0, 8]$ to maximize:

$$h(a, b^i, \bar{b}, X) = \sqrt{b^i \left(\sin 2q(a, X) + q(a, X)/2\right)} - \frac{1}{2} b^i.$$

The oscillating value for quality introduces a non-monotonicity into the firm's instantaneous return on quality through consumer spending:

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

It is possible for a firm to build a quality that is too high for the bargain market, but not high enough for the upscale market. 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 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

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

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, $\widetilde{X} \in \{0, 2\}$.

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

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

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 flow payoff is increasing in the support it receives from the constituents, and decreasing in effort: $g(a, \overline{b}, X) = \overline{b} - ca^2$.

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 static Nash payoff of $1-(1-X_t)^2$ 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 the set of Markov equilibria, which simultaneously establishes existence and characterizes equilibrium behavior and payoffs in any Markov equilibria. Next, I show that the long run player's payoff in any PPE is bound above and below by the Markov equilibria payoff set; therefore, the PPE payoff set is the convex hull of the correspondence of Markov equilibria payoffs. Third, I establish conditions for a Markov equilibrium to be the unique equilibrium in the class of all Perfect Public Equilibria. The section concludes with a discussion of incentives in stochastic games.

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.

Given state X and incentive weights $(z_y, x_x) = (0, z/r)$, define

$$\begin{aligned} g^*(X,z) &:= g(S^*(X,0,z/r),X) \\ \mu^*(X,z) &:= \mu_x(S^*(X,0,z/r),X) \\ \sigma^*(X,z) &:= \sigma_x(S^*(X,0,z/r),X) \end{aligned}$$

as the equilibrium value of the flow payoff, drift and volatility, respectively, and let

$$\psi(X,z) := g^*(X,z) + \frac{z}{r}\mu^*(X,z)$$

be the equilibrium value of the long-run player's incentive constraint. Note that $\psi(X,0)$ corresponds to the static Nash equilibrium payoff at state X.

In a Markov equilibrium, the continuation value and equilibrium actions are characterized as a function of the state variable as $W_t = U(X_t)$ and $(a_t^*, \overline{b}_t^*) = (a(X_t), \overline{b}(X_t))$.

Theorem 1. *Suppose Assumptions 1, 2, 3 and 4 hold. Given any initial state X_0, iff $U : \Xi \to \mathbb{R}$ is a solution to the optimality equation,*

$$U(X) = \psi(X, U'(X)) + \frac{1}{2r}U''(X)\left|\sigma^*\left(X, U'(X)\right)\right|^2$$

on $\Xi \backslash \widetilde{\Xi}$, $U(X) = \psi(X, 0)$ on $X \in \widetilde{\Xi}$, and U is bounded (has linear growth) given bounded (unbounded) flow payoff $g(a, b, X)$, then U characterizes a Markov equilibrium 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 $S^*(X_t, 0, U'(X_t)/r)$ for $X_t \in \Xi \backslash \widetilde{\Xi}$ and $S^*(X_t, 0, 0)$ for $X_t \in \widetilde{\Xi}$.*

The optimality equation has at least one continuous bounded (linear growth) solution that lies in the range of feasible payoffs for the long-run player. 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 the Bellman equation. The continuation value is equal to the sum of the equilibrium flow payoff today, $g^*(X, U')$, and the expected change in the continuation value,

$$\frac{1}{r}U'\mu^*(X, U') + \frac{1}{2r}U'' \left|\sigma^*(X, U')\right|^2$$

weighted by the discount rate. The term $U'\mu^*$ 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 ($\mu^* > 0$), and the continuation value is increasing in the state variable ($U' > 0$), then this increases the expected change in the continuation value. The term $U'' \left|\sigma^*\right|^2$ 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 X^* that yields a local maximum $U(X^*)$ (note this implies $U'(X^*) = 0$). Since the continuation value is at a local maximum, it must be decreasing as X moves away from X^* in either direction. The fact that $U''(X) < 0$ ensures that the continuation value is less than the current flow payoff. 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. Theorem 1 characterizes β_t as:

$$(\beta_{yt}, \beta_{xt}) = (0, U'(X_t)/r)$$

The current action affects future payoffs through its impact on the state variable, μ^*, scaled by the slope of the continuation value with respect to the state variable, $U'(X)$ and the discount rate. 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 $\beta_y = 0$ for all t.

3.1.1 Outline of Proof

This section presents intuition for the proof of Theorem 1. The first step in proving existence is to show that if a Markov equilibrium $(a_t^*, \overline{b}_t^*)$ 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:

$$dU(X_t) = U'(X_t)\mu_x(a_t^*, \overline{b}_t^*, X_t)dt + \frac{1}{2}U''(X_t)\left|\sigma_x\left(\overline{b}_t^*, X_t\right)\right|^2 dt + U'(X_t)\sigma_x\left(\overline{b}_t^*, X_t\right)dZ_t$$

In order for this to be an equilibrium, continuation values must also follow the law of motion specified in Lemma 2, with drift:

$$r\left(U(X_t) - g(a_t^*, \overline{b}_t^*, X_t)\right)$$

and volatility:

$$\beta_{yt}\left[\sigma_y \cdot dZ_t\right] + \beta_{xt}\left[\sigma_x \cdot dZ_t\right]$$

Matching the drifts of these two laws of motion yields the optimality equation, while matching the volatilities yields $r\beta_{yt} = 0$ and $r\beta_{xt} = U'(X_t)$. This establishes the condition for sequential rationality, as a function of the state variable.

The next step is to show that this ODE has at least one solution that lies in the range of feasible payoffs for the long-run player and meets the requisite growth conditions (bounded or linear). Establishing existence for a bounded state space is similar in technique to Faingold and Sannikov (2011). The innovative part of the proof lies in establishing existence for an unbounded state space, particularly in the case that the flow payoff of the long-run player is also unbounded. First I find lower and upper solutions, $\alpha(X)$ and $\beta(X)$, to the optimality equation. In the case of a bounded flow payoff, the lower and upper solutions must be bounded to establish existence of a bounded solution. For an unbounded flow payoff, lower and upper solutions must have linear growth to establish existence of a solution with linear growth; this is only possible when the growth rate of the drift is less than the discount rate. The proof derives explicit expressions for these lower and upper solutions, as a function of the growth rate of the flow payoff and drift with respect to X, and the discount rate. In turn, these expressions characterize a lower and upper bound to the solution to the optimality equation, $\alpha(X) \leq U(X) \leq \beta(X)$. Next I show that the Nagumo condition is satisfied on any compact subset of the state space; this holds because the bound of the optimality equation grows linearly with respect to U'. Finally, Lipschitz continuity of the equilibrium drift μ^* and flow payoff g^* and Assumption 1 ensure that the optimality equation is continuous. These three conditions establish existence of a twice continuously differentiable solution to the optimality equation.

The final step is to show that the continuation value and actions characterized above do in fact constitute a Markov equilibrium. Given an action profile $(a^*, \bar{b}^*) = S^*(X, 0, U'(X)/r)$ and solution U, the state variable evolves uniquely, the continuation value $dU(X_t)$ satisfies the condition for a PPE established in Lemma 2 and equilibrium actions satisfy sequential rationality.

Theorem 1 also establishes that *each* solution U to the optimality equation characterizes a single Markov equilibrium. This is a direct consequence of Assumption 4, 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.

3.2 Perfect Public Equilibria Payoff Set

This section characterizes the PPE payoff set. Define $\Upsilon : \Xi \rightrightarrows \mathbb{R}$ as the set of Markov equilibria payoffs, which depend on the initial state and are derived from each bounded solution to the optimality equation. Then the PPE payoff set is the convex hull of the correspondence of Markov equilibria payoffs. 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, 3 and 4 hold and let Υ be the correspondence of Markov equilibria payoffs. Then the correspondence of perfect public equilibrium payoffs $\xi : \Xi \rightrightarrows \mathbb{R}$ is characterized by the convex hull of the Markov equilibrium payoff set, $co(\Upsilon)$.*

The proof of Theorem 2 shows that if a PPE $\{W_t\}_{t\geq 0}$ yields a payoff higher than the maximum Markov equilibrium payoff $\{U(X_t)\}_{t\geq 0}$ at state X_0, then the difference $D_t = W_t - U(X_t)$ between these two payoffs will grow arbitrarily large with positive probability, irrespective of the state. Faingold and Sannikov (2011) establish a similar property on a bounded state space; their proof relies on the compactness of the state space to show that the volatility of D_t achieves a minimum that is bounded away from 0. An innovative part of this proof establishes that the volatility of D_t is bounded away from 0 on an unbounded state space.

The possibility that D_t grows arbitrarily large is obviously a contradiction in the case of a bounded state space or flow payoff, as the continuation value in any PPE must be bounded. The second innovative contribution of this proof is to establish that this is also a contradiction when the flow payoff, and therefore continuation value, is unbounded. By Lemma 1, there exists a $K > 0$ such that $|W_t(S)| \leq K(1 + |X_t|)$ for all $t \geq 0$. Thus, D_t can only grow arbitrarily large when X_t grows arbitrarily large. But I showed above that D_t will grow arbitrarily large independent of X_t. Therefore, as long as the model has a restriction on the rate at which the continuation value can grow with respect to the state, this reasoning will preclude the existence of PPE with payoffs outside the convex hull of the Markov payoff set.

The impossibility of achieving PPE payoffs above the highest Markov payoff yields insight into the role that the persistent effect of actions play 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 (Faingold and Sannikov 2011; Fudenberg and Levine 2007; Sannikov and Skrzypacz 2010). As in stochastic games with persistent actions, (1) is an ineffective channel for intertemporal incentives in standard repeated games. However, in the latter class of games, (1) is the only possible channel for intertemporal incentives. Stochastic games without persistent actions exhibit this same degeneracy result. The state variable evolves independently of the long-run player's action and the incentive constraint is independent of the continuation value, precluding channel (2). The unique PPE is one in which the long-run player acts myopically.[12]

[12]Note that repeated play of the static Nash action profile is not necessarily an equilibrium strategy profile

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, 3 and 4 hold. Then, given an initial state X_0, there is a unique perfect public equilibrium in which both players play the static Nash action profile $S^*(X, 0, 0)$ for all X.*

Proof. When μ_x is independent of a for all X, the incentive constraint collapses to maximizing the static flow payoff, and the long-run player plays the unique static Nash equilibrium action profile $(S^*(X_t, 0, 0))_{t\geq 0}$ in each state. Existence and uniqueness of a Markov equilibrium stems directly from Theorem 1.[13] The incentive constraint is independent of the solution to the optimality equation, which means all solutions to the optimality equation yield the same equilibrium action profile. The continuation value evolves according to the expected payoff from playing the static Nash profile,

$$U(X_t) = E_t \left[r \int_t^{\infty} e^{-rs} g^*(X_s, 0) dt \right].$$

Therefore, this solution must be unique. The solution to the optimality equation characterizes this expected payoff, as a function of X_t. □

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 feasible solution.[14] 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 feasible solution, there is a unique Markov equilibrium.

The limiting behavior of the flow payoffs and state variable, or the boundary conditions, play a key role in determining uniqueness. Assumption 5 outlines a set of sufficient conditions to guarantee that all solutions satisfy the same boundary conditions.

Assumption 5. .

1. *If the state space is bounded from below (above), then $\underline{X}$ ($\overline{X}$) is an absorbing state.*

of the stochastic game, as is the case in repeated games.

[13] As will be shown in Theorem 3, uniqueness of a Markov equilibrium in a game with persistent actions will require Assumption 5.

[14] If the flow payoff of the long-run player is bounded, then any bounded solution that lies in the range of feasible payoffs is a feasible solution, while if the flow payoff of the long-run player is unbounded, then any solution with linear growth is a feasible solution.

2. *If the state space is unbounded from below (above) and g is bounded, then there exists a δ such that for $|X| > \delta$, $g^*(X, 0)$ is monotonic in X.*

3. *If g is unbounded, then $|\sigma^*(X, z)|^2$ is Lipschitz continuous, $g^*(X, z)$ and $\mu^*(X, z)$ are separable in (X, z), and for all $z \in \mathbb{R}$, there exists a δ_z such that for $|X| > \delta_z$, $\psi'(X, z)$ is monotonic in X.*

Theorem 3. *Suppose Assumptions 1, 2, 3, 4 and 5 hold. Then, for each initial value of the state variable $X_0 \in \Xi$, there exists a unique perfect public equilibrium, which is Markov.*

1. *If g is bounded, the unique bounded solution U of the optimality equation satisfies the following boundary conditions:*

$$\lim_{X \to p} U(X) - g^*(X, 0) = 0$$
$$\lim_{X \to p} U'(X) = 0$$

for $p \in \{\underline{X}, \overline{X}\}$.

2. *If g is unbounded, the unique solution U of the optimality equation with linear growth satisfies the following boundary conditions:*

$$\lim_{X \to p} U(X) - y_L(X) = 0$$
$$\lim_{X \to p} U'(X) = k_p$$

for $p \in \{\underline{X}, \overline{X}\}$, where, given there exist g_1, g_2, μ_1, μ_2 such that $g^*(X, k) = g_1(X) + g_2(k)$ and $\mu^*(X, k) = \mu_1(X) + \mu_2(k)$,

$$y_L(x) = -\exp\left(\int \frac{r}{\mu_1(x)} dx\right) \int \exp\left(-\int \frac{r}{\mu_1(x)} dx\right) \frac{r g_1(x)}{\mu_1(x)} dx + g_2(k_p) + \frac{k_p}{r} \mu_2(k_p)$$

and

$$k_p = \lim_{X \to p} \frac{r g_1(X)}{rX - \mu_1(X)}.$$

The boundary conditions characterized in Theorem 3 have implications for equilibrium play and payoffs as X approaches $\{\underline{X}, \overline{X}\}$. Recall the incentive constraint for the long-run player. 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, $\mu^*(X, U')U'(X)$. In the case of a bounded flow payoff, Assumption 5 guarantees that $\mu^*(X, U')U'(X)$ converges to 0. This reduces the long-run player's incentive constraint to maximizing its instantaneous flow payoff and the continuation value converges to the static Nash payoff. If the flow payoff is unbounded, Assumption 5 guarantees that $U'(X)$ converges to a unique limit slope k_p, which is proportional to the ratio of the flow payoff and drift. Unlike in the bounded case, it is possible to sustain non-trivial incentives at the boundary, and the continuation value may remain strictly above (or below) the static Nash payoff as X becomes large.

3.3.1 Outline of Proof

The proof establishing that the optimality equation has a unique feasible solution has two main parts: (i) show that any feasible solution to the optimality equation must satisfy the same boundary conditions, and (ii) show that it is not possible for two different feasible solutions to the optimality equation to satisfy the same boundary conditions. The innovative part of this proof is to establish the boundary conditions for an unbounded state space, particularly when the flow payoff is also unbounded. I briefly describe this part of the proof.

In the case of an unbounded state space, the boundary conditions describe the limiting behavior of the optimality equation as it approaches ∞. Consider the incentive constraint $\psi(X,k)$. Given the monotonicity assumption and the Lipschitz continuity of μ^* and g^*, there exists a function $\psi_\infty(k)$ such that

$$\lim_{X\to\infty} \psi(X,k)/X = \lim_{X\to\infty} \psi'(X,k) := \psi_\infty(k).$$

exist and are equal, where $\psi' = d\psi/dX$. Let U be a solution to the optimality equation with linear growth. For any k such that $U'(X) = k$ infinitely often as $X \to \infty$, the shape of the optimality equation will alternate between convex and concave at slope k. By the optimality equation, $\psi(X,k)$ will lie above $U(X)$ when it is concave at slope k, and below $U(X)$ when it is convex at slope k. This property guarantees that

$$\lim_{X\to\infty} U/X = \lim_{X\to\infty} U' := U_p$$

exist and are equal. The linear growth of U and the Lipschitz continuity of $|\sigma^*|^2$ guarantee that $\lim_{X\to\infty} |\sigma^*(X,U')|^2 U''(X) = 0$. Therefore, dividing the optimality equation by X and taking limits, $\lim_{X\to\infty} U/X - \psi(X,U')/X = 0$. Combining these results yields $U_p = \psi_\infty(U_p)$. Any limit slope of a solution to the optimality equation with linear growth must be a fixed point of ψ_∞. Assumption 5 requires that g^* and μ^* are separable in (X,U'), which is sufficient to ensure that ψ_∞ has a unique fixed point.

From the optimality equation and the uniqueness of the limit slope, there exists a ϕ such that any solution to the optimality equation with linear growth satisfies $\lim_{X\to\infty} U -$

$U'\mu_1(X)/r - g_1(X) = \phi$. Consider the linear first order differential equation $y - \mu_1(X)y'/r - g_1(X) = \phi$. When the growth rate of μ_1 is in $[0, r)$, then there is a unique solution y^L that has linear growth, and when the growth rate of μ_1 is less than zero, then any two solutions y_1^L and y_2^L with linear growth have the same boundary condition, $\lim_{X\to\infty} y_1^L - y_2^L = 0$. Given the growth rate of μ is less than r by Assumption 3, we can show that any solution to the optimality equation with linear growth satisfies the same boundary condition,

$$\lim_{X\to\infty} U - y^L = 0$$

for all solutions y^L with linear growth. Thus, we have established that any solutions to the optimality equation, U and V, satisfy $\lim_{X\to\infty} U(X) - V(X) = 0$ and $\lim_{X\to\infty} U'(X) = \lim_{X\to\infty} V'(X) = k$ where k is the unique fixed point of ψ_∞.[15]

Establishing the boundary conditions for a bounded state space and an unbounded state space with a bounded flow payoff follow a similar pattern. Lastly, similar to Faingold and Sannikov (2011), given two feasible solutions U and V, if there exists an X such that $U(X) - V(X) > 0$, the structure of the optimality equation prevents these solutions from satisfying the same boundary conditions on at least one side of the state space.

3.4 Discussion of Results

3.4.1 Effective Intertemporal Incentives

Theorems 1-3 illustrate that it is possible to structure effective intertemporal incentives in stochastic games; this contrasts with the equilibrium degeneracy in continuous-time repeated games with a long-run and short-run player (Faingold and Sannikov 2011; Fudenberg and Levine 2007; Sannikov and Skrzypacz 2010). As discussed following Lemma 2, games with Brownian information must have a linear and non-tangential incentive structure.[16] In a repeated game, any non-trivial (i.e. non-zero) non-tangential linear incentive structure will escape the boundary of the payoff set with positive probability, a contradiction. However, in a stochastic game, this linear relationship between the Brownian information and the volatility of the continuation value can depend on the state variable. This important feature introduces the possibility of a non-trivial linear incentive structure.

Theorem 1 characterizes the volatility of the continuation value, which depends on the state variable and is proportional to $\beta = (0, U'(X)/r)$. When the state variable is at the value(s) that yields the maximum continuation value across all states, the continuation value

[15] Section 4 illustrates that ψ_∞ is straightforward to calculate, given functional forms for g and μ_x.

[16] There is a linear relationship between the volatility of the continuation value and the Brownian information.

must be independent of the Brownian information to prevent it from escaping the boundary of the payoff set (as is the case in repeated games). In these periods, $\beta = 0$ and the long-run player acts myopically. However, at other values of the state variable, the continuation value depends on Brownian information ($\beta \neq 0$) and the long-run player does not play a myopic best response. This is not possible in a repeated game – trivially, a repeated game is a stochastic game with a single state, and so β must always be zero.

With an unbounded flow payoff, it is possible to provide effective incentives as the state variable becomes large; the same holds true for a bounded flow payoff if the static Nash payoff perpetually oscillates. In both cases, $U'(X)$ doesn't converge to zero and non-trivial incentives can be sustained for large $|X|$. On the other hand, when the flow payoff is bounded and the static Nash payoff is monotonic, $U'(X)$ converges to zero and incentives dissipate as $|X|$ grows large.

Similarly, when the state space is bounded, it is possible to sustain non-trivial incentives as the state variable approaches its endpoints if these endpoints are not absorbing states. On the other hand, if the endpoints points are absorbing states, the stochastic game is eventually reduced to a standard repeated game and incentives dissipate near the boundary. The ability to sustain non-trivial intertemporal incentives across the entire state space is an important and novel insight of this paper.

3.4.2 Temporary or Permanent Incentives

Intertemporal incentives are temporary if the long-run player's continuation value converges to the static Nash payoff; otherwise, incentives are permanent. Whether incentives are temporary or permanent depends on the structure of the state variable and the equilibrium drift and volatility of the state variable near its absorbing states. If the state variable converges to an absorbing point with probability one, then the intertemporal incentives created by the stochastic game will be *temporary.* Once this absorbing state is reached, the stochastic game is reduced to a standard repeated game. On the other hand, if there are no absorbing states, 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.* Similarly, when the state space is unbounded, whether the state variable diverges will depend on the behavior of the state variable as it approaches infinity.

The permanence of incentives has an intuitive relationship to reputation models of incomplete information. Intertemporal incentives generated by incomplete information are temporary. In these models, the state variable is the belief about the long-run player's type and $\widetilde{X} \in \{0, 1\}$ are absorbing states. When the long-run player is a normal type, the transition function governing beliefs has negative drift, and beliefs converge to $\widetilde{X} = 0$.[17] At $X = 0$, short-run players believe the long-run player is a normal type with probability one, and it is not possible for the long-run player to earn payoffs above the static Nash payoff of the complete information game, $g^*(0,0)$. This captures the temporary reputation phenomenon associated with incomplete information models (Cripps et al. 2004; Faingold and Sannikov 2011). Once short-run players learn the long-run player's type, it is not possible to return to a state $X > 0$. Although $X = 1$ is also an absorbing state, the state variable almost surely doesn't converge to $X = 1$ when the long-run player is a normal type.

3.4.3 Empirical Content

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 acti2ons. 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.

4 Return to Examples

I return to the examples introduced in Section 2 to apply the results of Theorems 1, 2 and 3.

4.1 Persistent Effort as a Source of Reputation

Recall that the first example is interested in determining whether persistent effort can provide a firm with endogenous incentives to build its quality. This 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. Let $U(X)$ represent the continuation value of the firm in a Markov

[17] Provided it is not incentive compatible for the strategic player to perfectly mimic a behavioral type.

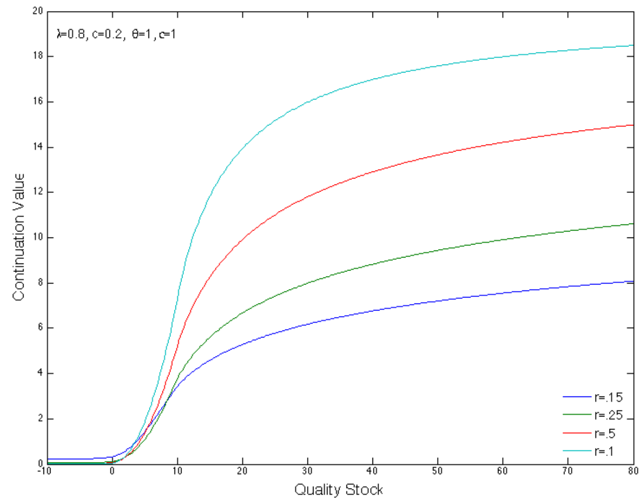

FIGURE 1. Equilibrium Payoffs

perfect equilibrium when $X_t = X$. Then, given equilibrium action profile $\left(a(X), \bar{b}(X)\right)$, the continuation value can be expressed as an ordinary differential equation,

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

where the first term, $\bar{b} - ca^2$, is the payoff that the firm earns today, and the second term, $\theta\left(a - X\right)U'(X) + \frac{1}{2}\sigma^2 U''(X)$, 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, $\theta\left(a - X\right)$.

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 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 non-monotonicity 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

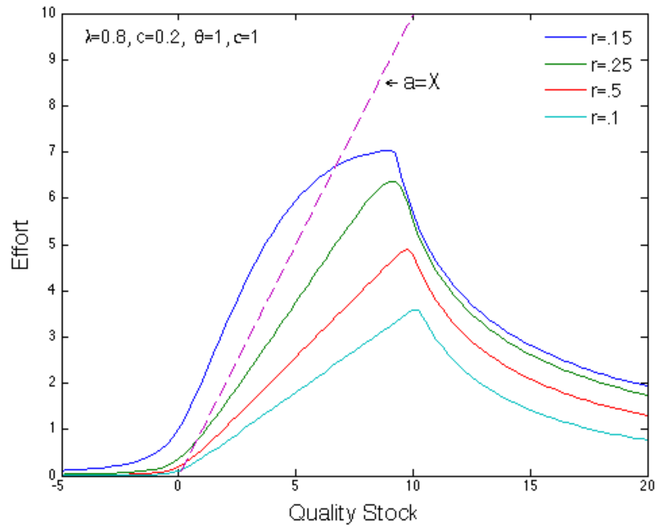

FIGURE 2. Firm Equilibrium Behavior

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.

The slope of the continuation value captures the impact of investment on future payoffs for the firm. In a Markov equilibrium, this 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.

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

When the continuation value is more sensitive to changes in stock quality (captured by a steeper slope), the firm chooses higher effort. Figure 2 graphs equilibrium actions for the firm. The firm's effort level peaks in a region of negative drift in the figure. When the firm receives 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. 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.

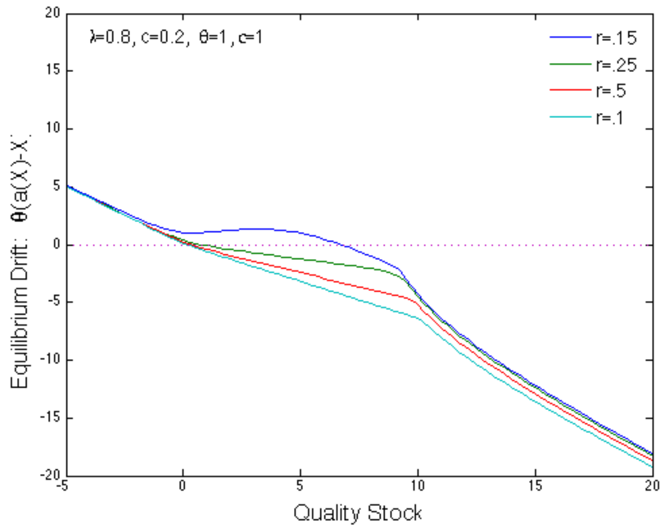

Figure 3. Equilibrium Drift of Stock Quality

These reputation effects are present in the long-run.[18] Product quality is cyclical, and periods of lower quality are followed by periods of higher quality.[19] This contrasts with models in which reputation is derived from behavioral types: as Cripps et al. (2004) 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.

Figure 3 illustrates the equilibrium dynamics of the stock quality. When the equilibrium drift is positive, stock quality increases in expectation, whereas the opposite holds when 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 and the long run player's flow payoff is bounded, 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

[18]Mathematically, 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.

[19]One 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.

depends on the consumers' purchase levels, which is determined by the value of the stock:

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

This payoff is monotonically increasing in X.

Next I compare the firm's payoffs to a game in which actions do not have a persistent effect on payoffs (this corresponds to $\theta = 0$). When effort is transitory, the unique equilibrium of the stochastic game is to choose zero effort each period. 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 be positive. 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. Therefore, the persistent effect of actions create nontrivial intertemporal incentives for the firm to exert effort.

4.1.1 Unbounded Demand

The expressions for equilibrium actions and continuation values in this variation mirrors those characterized above, with the exception of the consumer's action. However, the marginal return to the quality stock does not dissipate as X becomes large. In the unique Markov equilibrium, the slope of the continuation value converges to

$$\lim_{X\to\infty} U'(X) = \lambda r/(r+\theta),$$

as does the average return on stock quality, $U(X)/X$. Therefore, positive investment is sustained for any positive stock quality, and limit investment converges to:

$$\lim_{X\to\infty} a(X) = \frac{\theta\lambda}{2c(r+\theta)}.$$

On the other hand, if stock quality becomes very negative, the slope of the continuation value and equilibrium investment both converge to zero.

When the flow payoff is unbounded, a sufficient condition for uniqueness is that the equilibrium flow payoff and drift are separable in (X, U'), and that the slope of $\psi(X, U')$ is

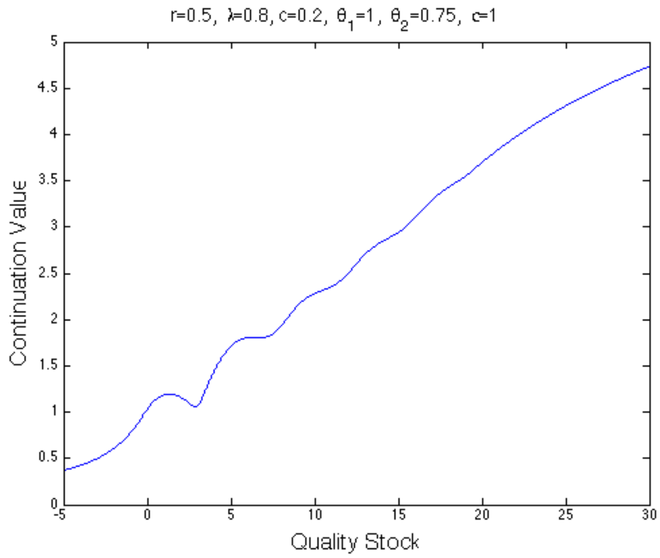

FIGURE 4. Equilibrium Continuation Value

monotonic in X for large X. Both conditions are satisfied in this example:

$$\begin{aligned} g^*(X,U') &= \frac{(1-\lambda)\theta U'}{2cr} + \lambda X - \frac{(\theta U')^2}{4cr^2} \\ \mu^*(X,U') &= \frac{\theta^2 U'}{4cr^2} - \theta X \\ \psi'(X,U') &= \lambda X - \theta U'/r \end{aligned}$$

for large X.

4.1.2 Building a Specialization with Oscillating Demand

The expressions for equilibrium actions and continuation values in this variation also mirror those in the initial quality example, with the exception of the consumer's action. However, the equilibrium dynamics are quite different. Investment oscillates as a function of quality. When the long-run player is in a region where spending is increasing with quality, it continues to build its quality. Once it crosses over to a region of decreasing spending, 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 4 and 5 illustrate equilibrium behavior. As in the initial quality example, the firm's payoff is bounded and $g^*(X,0)$ is eventually monotonic, guaranteeing uniqueness.

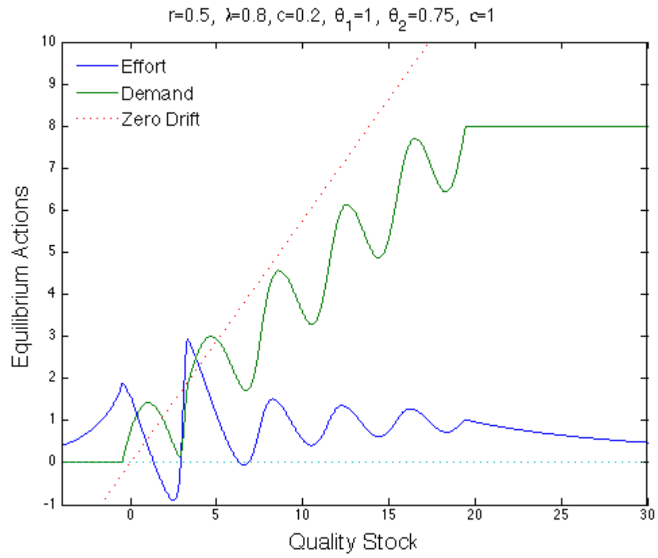

FIGURE 5. Equilibrium Actions

4.2 Policy Targeting

This example seeks to determine when persistent actions create incentives for a governing board to implement its' constituents policy target. From Section 3, the continuation value can be represented as

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

which is plotted in Figure 6.

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 6 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 $X = 1$, $U(1) < v(1) = 1$, which corresponds to the payoff that the governing board would earn if it could somehow keep the policy at the target permanently.

The board's incentives are driven by the slope of the continuation value and the impact

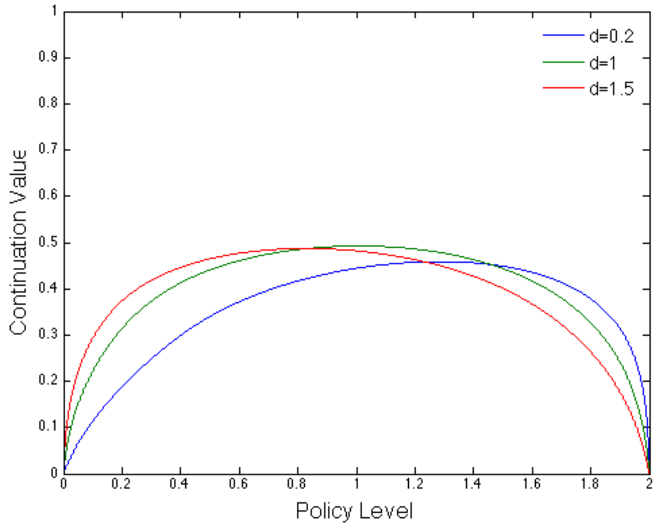

FIGURE 6. Equilibrium Payoffs

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 7 plots the equilibrium effort choice of the board.

5 Properties of Equilibrium Payoffs

The optimality equation provides a rich characterization of the relationship between equilibrium payoffs and the structure of the game. Its form yields insights into the shape of equilibrium payoffs, which states yield the highest or lowest equilibrium payoffs, and how these properties are tied to the underlying stage game.

Recall that $g^*(X, 0)$ represents the static Nash equilibrium payoff at state X. The shape of the static Nash payoff with respect to the state is a key determinant of the shape of the PPE payoff with respect to the state. I call a local maximum or minimum of the static Nash or PPE payoff correspondence an extrema. The number of PPE payoff extrema is bounded

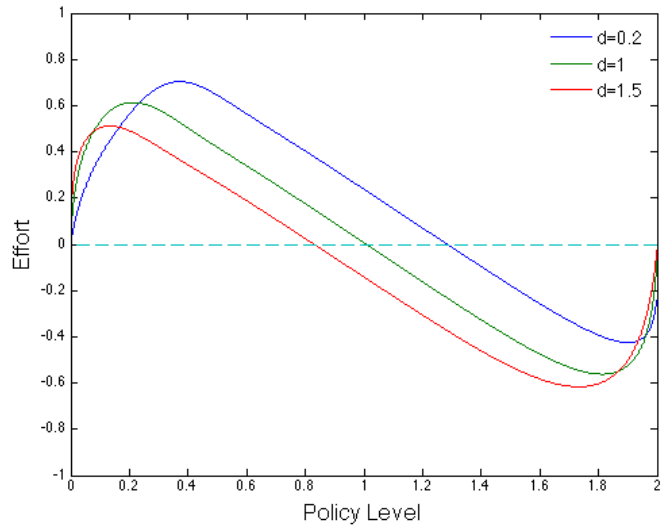

FIGURE 7. Government Equilibrium Behavior

by the number of static Nash extrema. Whether an extrema of the PPE payoff can occur on a given interval of the state space depends on the shape of the static Nash payoff on that interval. On any interval of the state space where the static Nash payoff is monotonic, the PPE payoff can have at most two extrema. Proposition 1 formally establishes these results.

Proposition 1. *Assume 1, 2, 3, 4 and 5. Let k_v and k_U be the number of strict interior extrema for $g^*(X,0)$ and U, with k_v^I and k_U^I representing the restriction to interval $I \subset \Xi$.*

1. *$k_U \leq k_v < \infty$.*
2. *If $g^*(X,0)$ is not constant on I, then $U(X)$ is not constant on I.*
3. *Suppose $g^*(X,0)$ is constant on I. If $\underline{X} \in I$ or $\overline{X} \in I$ then $k_U^I = 0$; otherwise, $k_U^I \leq 1$.*
4. *Suppose $g^*(X,0)$ is increasing (decreasing) on I.*
   - (a) *If the interval includes a boundary point, then $k_U^I \leq 1$.*
   - (b) *If the interval does not include a boundary point, then $k_U^I \leq 2$.*
   - (c) *If $k_U^I = 1$ and $\underline{X} \in I$, then the strict interior extremum is a maximum (minimum) and if $k_U^I = 1$ and $\overline{X} \in I$ then the strict interior extremum is a minimum (maximum).*
   - (d) *When $k_U^I = 2$, then there is a minimum (maximum) followed by a maximum (minimum).*

Proof. For notational simplicity, suppose the state space is bounded. At an interior extremum X, $U'(X) = 0$. From the optimality equation, if X is a minimum, $U(X) \geq g^*(X, 0)$, and if X is a maximum, $U(X) \leq g^*(X, 0)$. For the following proofs, let ξ_I represent the set of states that correspond to strict interior extrema of U on interval I, ordered so that $X_1 < X_2 < ... < X_{k_U^I}$.

Part 1 follows directly from Assumption 5, Lemma 10 and the Lipschitz continuity of g^*. For part 2, I prove the contrapositive. Suppose U is constant on some interval I: there exists a constant c such that $U(X) = c$ for all $X \in I$. Then $U'(X) = 0$ and $U''(X) = 0$ for all $X \in I$. From the optimality equation, $U(X) = g^*(X, 0)$ for all $X \in I$. Therefore, $g^*(X, 0) = c$ for all $X \in I$ and $g^*(X, 0)$ is constant on I.

For part 3, suppose $g^*(X, 0)$ is constant on an interval I and suppose $k_U^I \geq 2$. If X_1 is a minimum and X_2 is a maximum, then $g^*(X_1, 0) \leq U(X_1) < U(X_2) \leq g^*(X_2, 0)$. This is a contradiction, because $g^*(X, 0)$ is constant. The same logic holds if X_1 is a maximum and X_2 is a minimum. Therefore, $k_U^I \leq 1$. If $\underline{X} \in I$ and X_1 is a minimum, then $g^*(X_1, 0) \leq U(X_1) < g^*(\underline{X}, 0) = U(\underline{X})$. This is a contradiction, because $g^*(X, 0)$ is constant – it's not possible to have a minimum. Similarly, it's not possible to have a maximum. Therefore, if I contains a boundary point, $k_U^I = 0$.

For part 4, assume $g^*(X, 0)$ is increasing on an interval I and $k_U^I \geq 2$. If X_i is a strict maximum and X_{i+1} is a strict minimum, then $g^*(X_{i+1}, 0) \leq U(X_{i+1}) < U(X_i) \leq g^*(X_i, 0)$. This is a contradiction, because v is increasing on I – it's not possible to have a maximum followed by a minimum. Therefore, $k_U^I = 1$ if X_1 is a maximum, and $k_U^I \leq 2$ if X_1 is a minimum. Now suppose that $\underline{X} \in I$. By the boundary conditions, $g^*(\underline{X}, 0) = U(\underline{X})$. If X_1 is a strict minimum then $g^*(X_1, 0) \leq U(X_1) < g^*(\underline{X}, 0) = U(\underline{X})$. This is a contradiction, because $g^*(X, 0)$ is increasing – it's not possible to have a minimum. Therefore, $k_U^I = 0$ or $k_U^I = 1$ and X_1 is a maximum. Similarly, if $\overline{X} \in I$, it's not possible to have a maximum. Either $k_U^I = 0$ or $k_U^I = 1$ and X_1 is a minimum. The proof for when $g^*(X, 0)$ is decreasing is analogous.

To extend the proofs to an unbounded state space, replace $g^*(\underline{X}, 0)$ with $\lim_{X \to p} g^*(X, 0)$ for $p \in \{\underline{X}, \overline{X}\}$. □

A direct application of Proposition 1 leads to Proposition 2. If the static Nash payoff increases or decreases with respect to the state variable, then the PPE payoff also increases or decreases, respectively, with respect to the state variable. Similarly, if the static Nash payoff is single-peaked, and takes on the same value at the upper and lower bound of the state space, then the PPE payoff is also single-peaked.

Proposition 2. *Suppose Assumptions 1, 2, 3, 4 and 5 hold.*

1. *If $g^*(X,0)$ (strictly) monotonically increases (decreases) in X, then U (strictly) monotonically increases (decreases) in X.*

2. *If $g^*(X,0)$ is single-peaked with a maximum (minimum) and $g^*(\underline{X},0) = g^*\left(\overline{X},0\right)$, then U is single-peaked with a maximum (minimum).*

Proof. Let k_v and k_U be the number of strict interior extrema for $g^*(X,0)$ and U. Suppose $g^*(X,0)$ is monotonically increasing on Ξ, with $dg^*(X,0)/dx > 0$ for some $X \in \Xi$, but $U(X)$ is not monotonically increasing. Then $U'(X) < 0$ for some $X \in \Xi$. By Proposition **1,** $k_U = 0$. Therefore, it must be that $U'(X) \leq 0$ for all $X \in \Xi$. By the boundary conditions,

$$U(\underline{X}) = g^*(\underline{X},0) < U(\overline{X}) = g^*(\overline{X},0)$$

so there must be an X such that $U'(X) > 0$, a contradiction. Therefore, $U(X)$ is monotonically increasing in X. If $g^*(X,0)$ is strictly monotonically increasing, then U is also strictly monotonically increasing, by part (2) of Proposition **1.** The proof for $g^*(X,0)$ decreasing is analogous.

Suppose $k_v = 1$ and $g^*(X,0)$ has a unique interior maximum at X^*. Then $dg^*(X,0)/dx \geq 0$ for $X < X^*$ and $dg^*(X,0)/dx \leq 0$ for $X > X^*$. By Proposition 1 part (1), $k_U \leq k_v = 1$, and by part (4), if U has a strict interior extrema on $(\underline{X}, X^*]$, then it has a maximum, and if U has a strict interior extrema on $[X^*, \overline{X})$, then it has a maximum. Therefore, if $k_U = 1$, the extremum is a maximum. Suppose $k_U = 0$ and U has no interior extremum. Then U is either increasing, decreasing, or constant on Ξ. Since $g^*(\underline{X},0) = g^*\left(\overline{X},0\right)$, by the boundary conditions, $U(\underline{X}) = U\left(\overline{X}\right)$. Therefore, $k_U = 0$ implies that U is constant on Ξ. This is a contradiction: by Proposition 1 part (2), if $g^*(X,0)$ is not constant on Ξ, then U cannot be constant on Ξ. Therefore, it must be the case that $k_U = 1$ and U is single-peaked with a maximum. □

This Section concludes with a weak bound on the PPE payoff set when the flow payoff is bounded. If the flow payoff is bounded, then the PPE payoff set is bounded and $\overline{W} := \sup_{\Xi} U(X)$ and $\underline{W} := \inf_{\Xi} U(X)$ exist. When $W_t = \overline{W}$, W_t must have zero volatility ($\beta_t = 0$) and a weakly negative drift, so as not to exceed this bound. From Lemma 2, if the drift is weakly negative, then the current flow payoff is weakly greater than the continuation value. But when $\beta_t = 0$, the current flow payoff is the static Nash payoff at the current state. Therefore, the PPE payoff set is bounded above by the highest static Nash payoff across all states. Similarly, PPE payoff set is bounded below by the lowest static Nash payoff across all states.[20] Proposition 3 states this formally.

[20]Note that when the continuation value is not at its upper or lower bound, the PPE payoff may lie above

Proposition 3. *Suppose Assumptions 1, 2, 3 and 4 hold, and the PPE payoff set is bounded from above (below) across all states. Then the highest (lowest) PPE payoff is bounded above (below) by the static Nash payoff at the corresponding state:*

1. *If $\exists X$ such that $\overline{W} = U(X)$, then $\overline{W} \leq g^*(X,0)$. If $\overline{W} = \limsup_{X \to p} U(X)$, then $\overline{W} \leq \limsup_{X \to p} g^*(X,0)$ for $p \in \{-\infty, \infty\}$.*

2. *If $\exists X$ such that $\underline{W} = U(X)$, then $\underline{W} \geq g^*(X,0)$. If $\underline{W} = \liminf_{X \to p} U(X)$, then $\underline{W} \geq \liminf_{X \to p} g^*(X,0)$ for $p \in \{-\infty, \infty\}$.*

Proof. Suppose $U(X)$ is bounded from above. Then $U(X)$ is continuous and bounded on Ξ, a closed set. Therefore, $U(X)$ either attains its maximum on Ξ, in which case $\overline{W} = U(X)$ for some X, or $\overline{W} = \limsup_{X \to p} U(X)$ for $p \in \{-\infty, \infty\}$. Suppose $\overline{W} = U(X)$ occurs at an interior non-absorbing state X. Then $U'(X) = 0$ and $U''(X) \leq 0$. From the optimality equation,

$$U''(X) = \frac{2r\left(\overline{W} - g^*(X,0)\right)}{\left|\sigma^*\left(X,0\right)\right|^2} \leq 0.$$

and therefore $\overline{W} \leq g^*(X,0)$. If $\overline{W} = U(X)$ occurs at an absorbing state $\widetilde{X}$, then $\overline{W} = g^*(\widetilde{X},0)$. Suppose the state space is bounded from above and $\overline{W} = U(\overline{X})$. By the definition of $\overline{X}$, $\mu^*(\overline{X}, U'(\overline{X})) \leq 0$ and $\sigma^*\left(\overline{X}, U'(\overline{X})\right) = 0$. Also, $U'(\overline{X}) \geq 0$. From the optimality equation,

$$\overline{W} - g^*(\overline{X},0) = \frac{1}{r} U'(\overline{X}) \mu^*(\overline{X}, U'(\overline{X})) \leq 0$$

The proof for $\overline{W} = U(\underline{X})$ is analogous. Otherwise, by Lemma 10, the local minima and maxima of U are bounded by the local minima and maxima of $g^*(X,0)$. Therefore, for any sequence local maxima $\{U_k\}$ converging to $\overline{W}$, there exists a sequence of states $\{X_k\}$ such that $U_k = U(X_k)$ and $\overline{W} \leq \lim_{k \to \infty} g^*(X_k,0) = \limsup_{X \to \infty} g^*(X,0)$. The proofs for the lower bound $\underline{W}$ are analogous. □

If the continuation value is sufficiently flat around a state X that yields $\overline{W}$ (i.e. if $U''(X) = 0$) or X is an absorbing state, then the highest PPE payoff is $\overline{W} = g^*(X,0)$. Otherwise, the highest PPE payoff will be strictly below $g^*(X,0)$, as either the continuation value or the state changes too quickly at X.

6 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

or below the static Nash payoff at the corresponding state.

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 equilibrium, which is Markov. These results develop a tractable method to characterize equilibrium behavior and payoffs, which can be applied to models in industrial organization, political economy and macroeconomics. The equilibrium dynamics can be directly related to observable features of a firm or government or other long-run player, and used to generate empirically testable predictions. Once more structure is placed on payoff functions and the state variable, it will be possible to address questions about efficiency, comparative statics, steady-state dynamics, the impact of shocks and market design issues such as the trade-off between different types of persistence.

This paper leaves open several interesting avenues for future research, including whether the model is robust in the sense that nearby discrete time games exhibit similar equilibrium properties and whether it is possible to say anything about a game with multiple long-run players or states. Analyzing a setting with multiple state variables is technically challenging; whether it is possible to reduce such games to a simpler setting that yields a tractable characterization of equilibrium dynamics remains an open question.

References

Abreu, D., P. Milgrom, and D. Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59, 1713–1733.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

Back, K. (1992): "Insider Trading in Continuous Time," *The REview of Financial Studies*, 5, 387–409.

Board, S. and M. Meyer-ter vehn (2013): "Reputation for Quality," *Econometrica*.

Cisternas, G. (2014): "Two-sided Learning and Moral Hazard," .

Cripps, M., G. Mailath, and L. Samuelson (2004): "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72, 407–432.

Daley, B. and B. Green (2012): "Waiting for News in the Market for Lemons," *Econometrica*, 80, 1433–1504.

Dutta, P. K. (1995): "A Folk Theorem for Stochastic Games," *Journal of Economic Theory*, 66, 1–32.

Dutta, P. K. and R. Sundaram (1992): "Markovian equilibrium in a class of stochastic games: existence theorems for discounted and undiscounted models," *Economic Theory*, 2, 197–214.

Ericson, R. and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62, 53–82.

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

Flesch, J., F. Thuijsman, and O. J. Vrieze (2003): "Stochastic games with non-observable actions," *Mathematical Methods of Operations Research*, 58, 459–475.

Fudenberg, D. and D. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759–778.

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

——— (2009): "Repeated games with frequent signals," *Quarterly Journal of Economics*, 233–265.

Fudenberg, D. and D. K. Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103–135.

——— (2007): "Continuous time limits of repeated games with imperfect public monitoring," *Review of Economic Dynamics*, 10, 173–192.

Fudenberg, D. and Y. Yamamoto (2011): "The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring," *Journal of Economic Theory*, 146, 1664–1683.

Hörner, J., T. Sugaya, S. Takahashi, and N. Vieille (2011): "Recursive Methods in Discounted Stochastic Games: An Algorithm and a Folk Theorem," *Econometrica*, 79, 1277–1318.

Karatzas, I. and S. Shreve (1991): *Brownian Motion and Stochastic Calculus*, New York: Springer-Verlag.

Klein, B. and K. Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 615–641.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Dilemma Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245–252.

Kreps, D. and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

Mailath, G. J. and L. Samuelson (2001): "Who Wants a Good Reputation?" *Review of Economic Studies*, 68, 415–41.

——— (2006): *Repeated Games and Reputations: Long-Run Relationships*, no. 9780195300796 in OUP Catalogue, Oxford University Press.

Milgrom, P. and J. Roberts (1982): "Predation, reputation, and entry deterrence," *Journal of Economic Theory*, 27, 280–312.

Murto, P. (2004): "Exit in Duopoly under Uncertainty," *The Rand Journal of Economics*, 35, 111–127.

Papageorgiou, T. (2013): "Learning Your Comparative Advantages," *The Review of Economic Studies*, 81, 1263–1295.

Sannikov, Y. (2007): "Games With Impercetly Observable Actions," *Econometrica*, 75, 1285–1329.

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.

Schmitt, K. (1969): "Bounded Solutions of Nonlinear Second Order DIfferential Equations," *Duke Mathematical Journal*, 36, 237–243.

Sobel, M. (1973): "Continuous Stochastic Games," *Journal of Applied Probability*, 10, 597–604.

Stigler, G. (1961): "The Economics of Information," *The Journal of Political Economy*, 69, 213–225.

Strulovici, B. and M. Szydlowski (2014): "On the Smoothness of Value Functions and the Existence of Optimal Strategies," .

7 Appendix

7.1 Proof of Lemma 1

By Assumption 3, there exists a $k \in [0, r)$ and $c > 0$ such that for all $X \geq 0$, $\mu_x(a, b, X) \leq kX + c$ for all $a \in A, b \in B$, and for all $X \leq 0$, $\mu_x(a, b, X) \geq kX - c$ for all $a \in A, b \in B$. By Lipshitz continuity, there exists a $K_g, K_\sigma > 0$ such that for all $X \in \Xi$, $|g(a, b, X)| \leq K_g(\frac{c}{k} + |X|)$ and $\sigma(b, X) \leq K_\sigma(1 + |X|)$ for all $a \in A, b \in B$.

Step 1: Derive a bound on $E_\tau\left[g(a_t, \overline{b}_t, X_t)\right]$, the expected flow payoff at time t conditional on available information at time $\tau \leq t$. Note this bound will be independent of the strategy profile. Define

$$f(X) \quad : \quad = \begin{cases} K_g(\frac{c}{k} - X) & \text{if } X \leq -1 \\ -\frac{1}{8}K_g X^4 + \frac{3}{4}K_g X^2 + \frac{3}{8}K_g + K_g \frac{c}{k} & \text{if } X \in (-1, 1) \\ K_g(\frac{c}{k} + X) & \text{if } X \geq 1 \end{cases}$$

Note $f \in \mathcal{C}^2$, $f(X) \geq 0$, $|f'(X)| \leq K_g$ and $f''(X) \leq \frac{3}{2}K_g$. Ito's Lemma holds for any $\mathcal{C}^2$ function, so for any strategy profile $S = \left(a_t, \overline{b}_t\right)_{t\geq 0}$, any initial value $X_\tau < \infty$ and path of the state variable $(X_t)_{t\geq\tau}$ that evolves according to 1, given S, and any $t \geq \tau$,

$$\begin{aligned} f(X_t) &= f(X_\tau) + \int_\tau^t \left(f'(X_s)\mu_x(a_s, \overline{b}_s, X_s) + \frac{1}{2}f''(X_s)\left|\sigma_x\left(\overline{b}_s, X_s\right)\right|^2 \right) ds + \int_\tau^t f'(X_s)\sigma_x\left(\overline{b}_s, X_s\right) dZ_s \\ &\leq f(X_\tau) + \int_\tau^t \left(K_g\left(k\left|X_s\right| + c\right) + 3K_g K_\sigma^2 \right) ds + K_g K_\sigma \int_\tau^t \left(1 + X_s\right) dZ_s \\ &\leq f(X_\tau) + k\int_\tau^t f(X_s) ds + 3K_g K_\sigma^2\left(t - \tau\right) + K_g K_\sigma \int_\tau^t \left(1 + X_s\right) dZ_s \end{aligned}$$

where $f'X_s)\mu_x(a_s, \overline{b}_s, X_s) \leq K_g\left(k\left|X_s\right| + c\right)$ and $\frac{1}{2}f''(X_s)\left|\sigma_x\left(\overline{b}_s, X_s\right)\right|^2 \leq 3K_g K_\sigma^2$ for all $X \in \Xi$ and $\left(a, \overline{b}\right) \in A \times B$. Taking expectations, and noting that $(1 + X_s)$ is square-integrable on $[\tau, t]$, so the expectation of the stochastic integral is zero,

$$\begin{aligned} E_\tau[f(X_t)] &\leq f(X_\tau) + 3K_g K_\sigma^2 (t-\tau) + k\int_\tau^t E_\tau[f(X_s)]\, ds \\ &\leq \left(f(X_\tau) + 3K_g K_\sigma^2 (t-\tau)\right) e^{k(t-\tau)} \end{aligned}$$

where the last line follows from Gronwall's inequality. Note that for all $X \in \Xi$, $|g(a,b,X)| \leq f(X)$ for all $a \in A$, $b \in B$. Therefore,

$$e^{-r(t-\tau)} E_\tau \left|g(a_t, \bar{b}_t, X_t)\right| \leq e^{-r(t-\tau)} E_\tau[f(X_t)] \leq \left(f(X_\tau) + 3K_g K_\sigma^2 (t-\tau)\right) e^{-(r-k)(t-\tau)}$$

Step 2: Show that if $X_t < \infty$, then $W_t(S) < \infty$.

$$\begin{aligned} |W_t(S)| &= \left| E_t \left[r \int_t^\infty e^{-r(s-t)} g(a_s, \bar{b}_s, X_s) ds \right] \right| \\ &\leq rE_t \left[\int_t^\infty e^{-r(s-t)} \left| g(a_s, \bar{b}_s, X_s) \right| ds \right] \\ &\leq r \int_t^\infty e^{-r(s-t)} E_t \left| g(a_s, \bar{b}_s, X_s) \right| ds \\ &\leq r \int_t^\infty \left(f(X_t) + 3K_g K_\sigma^2 (s-t) \right) e^{-(r-k)(s-t)} ds \\ &= r \left(\frac{f(X_t)(r-k) + 3K_g K_\sigma^2}{(r-k)^2} \right) \end{aligned}$$

which is finite for any $X_t < \infty$ and $k < r$. Also, given that $f(X_t)$ has linear growth, there exists a $K > 0$ such that:

$$|W_t(S)| \leq K(1 + |X_t|).$$

Step 3: Show $E\,|V_t(S)| < \infty$ for any $X_0 < \infty$. By similar reasoning to Step 2,

$$E\,|V_t(S)| = E \left| E_t \left[r \int_0^\infty e^{-rs} g(a_s, \bar{b}_s, X_s) ds \right] \right| \leq E \left[r \int_0^\infty e^{-rs} \left| g(a_s, \bar{b}_s, X_s) \right| ds \right]$$

which is finite for any $X_0 < \infty$ and $k < r$.

Step 4: Show $E_t[V_{t+k}(S)] = V_t(S)$

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

By steps 3 and 4, $V_t(S)$ is a martingale.

Note that if $g(a,b,X)$ is bounded with respect to X, then, trivially, $V_t(S)$ and $W_t(S)$ are bounded for all $t \geq 0$ and $X_0 \in \Xi$. Also note that $g(a,b,X)$ is bounded whenever the state space is bounded.

7.2 Proof of Lemma 2

Evolution of the continuation value: From Lemma 1, $V_t(S)$ is a martingale. 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 $dV_t(S) = re^{-rt}\beta_t^{\top}\sigma\left(\bar{b}_t,X_t\right)dZ_t$. Combining these two expressions for $dV_t(S)$ yields the law of motion for the continuation value:

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

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. Lemma 1 establishes that any continuation value is bounded when g is bounded, and otherwise bounded with respect to the state variable.

Sequential Rationality: Consider strategy profile $(a_t,\bar{b}_t)_{t\geq 0}$ played from period τ onwards and alternative strategy $(\widetilde{a}_t,\bar{b}_t)_{t\geq 0}$ 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

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

whereas action $\widetilde{a}_\tau$ will induce

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

Let $\widetilde{V}_\tau$ be the expected average payoff conditional on information at time τ when the long-run player follows $\widetilde{a}$ 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.

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

Consider changing τ so that long-run player plays strategy $(\widetilde{a}_t, \bar{b}_t)$ for another instant: $d\widetilde{V}_\tau$ 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. When long-run player switches strategies at time τ,

$$\begin{aligned}
d\widetilde{V}_\tau &= re^{-r\tau}\left[g(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) - W_\tau\right] d\tau + e^{-r\tau} dW_\tau \\
&= re^{-r\tau}\left[g(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) - g(a_\tau, \bar{b}_\tau, X_\tau)\right] d\tau + re^{-r\tau}\beta_{y\tau}\left[dY_\tau - \mu_y(a_\tau, \bar{b}_\tau, X_\tau) d\tau\right] \\
&\quad + re^{-r\tau}\beta_{xt}\left[dX_t - \mu_x(a_t, \bar{b}_t, X_t) dt\right] \\
&= re^{-r\tau}\left[\left[g(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) - g(a_\tau, \bar{b}_\tau, X_\tau)\right] d\tau + \beta_{y\tau}\left[\mu_y(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) - \mu_y(a_\tau, \bar{b}_\tau, X_\tau)\right] d\tau\right] \\
&\quad + re^{-r\tau}\left[\beta_{x\tau}\left[\mu_x(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) - \mu_x(a_\tau, \bar{b}_\tau, X_\tau)\right] d\tau + \beta_\tau^\top \sigma\left(\bar{b}_t, X_t\right) dZ_t\right]
\end{aligned}$$

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 future state X_t, which impact the continuation value. The profile $(\widetilde{a}_t, \bar{b}_t)_{t\geq 0}$ yields the long-run player a payoff of:

$$\begin{aligned}
\widetilde{W}_0 &= E_0\left[\widetilde{V}_\infty\right] = E_0\left[\widetilde{V}_0 + \int_0^\infty d\widetilde{V}_t\right] \\
&= W_0 + E_0\left[r\int_0^\infty e^{-rt}\left\{\begin{array}{c} g(\widetilde{a}_t, \bar{b}_t, X_t) + \beta_{y\tau}\mu_y(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) + \beta_{x\tau}\mu_x(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau) \\ -g(a_t, \bar{b}_t, X_t) - \beta_{y\tau}\mu_y(a_\tau, \bar{b}_\tau, X_\tau) - \beta_{x\tau}\mu_x(a_\tau, \bar{b}_\tau, X_\tau) \end{array}\right\} dt\right]
\end{aligned}$$

If

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

holds for all $t \geq 0$, then $W_0 \geq \widetilde{W}_0$ and deviating to $S = (\widetilde{a}_t, \bar{b}_t)$ is not a profitable deviation. 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:

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

7.3 Notation

Given state X and $\beta_y = 0, \beta_x = z$, let

$$(a(X,z), b(X,z)) := S^*(X, 0, z/r)$$

be the optimal actions and

$$\begin{aligned} g^*(X,z) &:= g(a(X,z), b(X,z), X) \\ \mu^*(X,z) &:= \mu_x(a(X,z), b(X,z), X) \\ \sigma^*(X,z) &:= \sigma_x(b(X,z), X) \end{aligned}$$

be the value of the flow payoff, drift and volatility, respectively, as the optimal actions. All of these functions are Lipschitz continuous, either by assumption or because they are the composite of Lipschitz continuous functions. Also, for any fixed X, all of these functions are bounded with respect to z, since $a(X,z)$ is bounded by the compactness of A and the remaining functions depend on z through $a(X,z)$. Let

$$\psi(X,z) := g^*(X,z) + \frac{z}{r}\mu^*(X,z)$$

be the value of the long-run player's incentive constraint at the optimal action. For a fixed z, $\psi(X,z)$ is Lipschitz continuous in X, but not necessarily Lipschitz continuous in (X,z). Also, ψ is not bounded in z.

7.4 Proof of Theorem 1

Form of Optimality Equation: 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, if a Markov equilibrium exists, the continuation value will evolve according to:

$$\begin{aligned} dU(X_t) &= U'(X_t)dX_t + \frac{1}{2}U''(X_t)\left|\sigma_x\left(\overline{b}_t^*, X_t\right)\right|^2 dt \\ &= U'(X_t)\mu_x(a_t^*, \overline{b}_t^*, X_t)dt + \frac{1}{2}U''(X_t)\left|\sigma_x\left(\overline{b}_t^*, X_t\right)\right|^2 dt + U'(X_t)\sigma_x\left(\overline{b}_t^*, X_t\right)dZ_t \end{aligned}$$

Matching the drift of this expression with the drift of the continuation value characterized in Lemma 2 yields the optimality equation for strategy profile $(a^*, \overline{b}^*)$,

$$U''(X) = \frac{2r\left(U(X) - g(a^*, \overline{b}^*, X)\right)}{\left|\sigma_x\left(\overline{b}_t^*, X_t\right)\right|^2} - \frac{2\mu_x(a^*, \overline{b}^*, X)}{\left|\sigma_x\left(\overline{b}_t^*, X_t\right)\right|^2}U'(X)$$

which is a second order non-homogenous differential equation.

Matching the volatility characterizes the process governing incentives, $\beta_y = 0$ and $\beta_x = U'(X)/r$. 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 $\beta_y = 0$. Plugging these into the condition for sequential rationality,

$$S^*(X, 0, U'(X)/r) = \left\{(a^*, \overline{b}^*) \text{ s.t. } a^* = \arg\max_a \ g(a, \overline{b}^*, X) + \frac{1}{r}U'(X)\mu_x(a, \overline{b}^*, X)\ , \overline{b}^* = \overline{\mathcal{B}}(a^*, X)\right\}$$

which is unique by Assumption 4.

Existence of solution to optimality equation: Let $f : \Xi \times \mathbb{R}^2 \to \mathbb{R}$ denote the right hand side of the optimality equation:

$$f(X, U, U') = \frac{2r}{|\sigma^*(X, U')|^2}\left(U - \psi(X, U')\right)$$

which is continuous on $int(\Xi) \setminus \tilde{\Xi}$. I establish that the second order differential equation

$$U'' = f(X, U, U')$$

has at least one $\mathcal{C}^2$ solution that takes on values in the interval of feasible payoffs for the long-run player.

Case 1: Unbounded State Space. I use the following theorem from Schmitt (1969), which gives sufficient conditions for the existence of a solution to a second order differential equation defined on $\mathbb{R}^3$. The Theorem is reproduced below.

Theorem 4 (Schmitt (1969))**.** *Let $\alpha, \beta \in C^2(\mathbb{R})$ be functions such that $\alpha \leq \beta$, $E = \{(t, u, v) \in \mathbb{R}^3 | \alpha(t) \leq u(t) \leq \beta(t)\}$ and $f : E \to \mathbb{R}$ be continuous. Assume that α and β*

are such that for all $t \in \mathbb{R}$*,*

$$f(t,\alpha(t),\alpha'(t)) \leq \alpha''(t) \text{ and } \beta''(t) \leq f(t,\beta(t),\beta'(t)).$$

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

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

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

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

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

Lemma 3. *Suppose* $\Xi = \mathbb{R}$ *and there are no interior absorbing points. 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.*

Suppose $\Xi = \mathbb{R}$ and there are no interior absorbing points. The optimality equation is continuous on $\mathbb{R}^3$. Consider the functions

$$\alpha(X) \;:=\; \begin{cases} \alpha_1 X - c_a & \text{if } X \leq -1 \\ \frac{1}{8}\alpha_1 X^4 - \frac{3}{4}\alpha_1 X^2 - \frac{3}{8}\alpha_1 - c_a & \text{if } X \in (-1,1) \\ -\alpha_1 X - c_a & \text{if } X \geq 1 \end{cases}$$

$$\beta(X) \;:=\; \begin{cases} -\beta_1 X + c_b & \text{if } X \leq -1 \\ -\frac{1}{8}\beta_1 X^4 + \frac{3}{4}\beta_1 X^2 + \frac{3}{8}\beta_1 + c_b & \text{if } X \in (-1,1) \\ \beta_1 X + c_b & \text{if } X \geq 1 \end{cases}$$

for some $\alpha_1, \beta_1, c_a, c_b \geq 0$. Note that $\alpha, \beta \in C^2(\mathbb{R})$ and $\alpha \leq \beta$. Then α, β are lower and upper solutions to the optimality function if there exist $\alpha_1, \beta_1, c_a, c_b \geq 0$ such that:

$$\frac{2r}{\sigma^2(X,\alpha')}\left(\alpha - \psi(X,\alpha')\right) \leq \alpha'' \text{ and } \beta'' \leq \frac{2r}{\sigma^2(X,\beta')}\left(\beta - \psi(X,\beta')\right)$$

for all $X \in \mathbb{R}$.

Step 1: Assume that there exists a $k \in [0,r)$ such that for some $c \geq 0$, $\mu^*(X,z) \leq kX + c$ for all $X \geq 0$ and $\mu^*(X,z) \geq kX - c$ for all $X \leq 0$. Show that there exist $\alpha_1, \beta_1, c_a, c_b \geq 0$

such that α, β are lower and upper solutions to the optimality function. Note Step 1 does not require g to be bounded.

Step 1a: Find a bound on $\psi(X, z)$. By Lipschitz continuity and the fact that g^* and μ^* are bounded in z, $\exists k_g, k_m \geq 0$ such that

$$\begin{aligned} |g^*(X,z) - g^*(0,z)| &\leq k_g |X| \\ |\mu^*(X,z) - \mu^*(0,z)| &\leq k_m |X| \end{aligned}$$

for all (X, z). Therefore, $\exists \underline{g}_1, \underline{g}_2, \overline{g}_1, \overline{g}_2 \geq 0$, $\underline{\mu}_1, \overline{\mu}_2 \in [0, r)$, $\underline{\mu}_2, \overline{\mu}_1 > 0$ and $\overline{\gamma}, \underline{\gamma}, \overline{m}, \underline{m} \in \mathbb{R}$ such that:

$$\begin{cases} \underline{g}_1 X + \underline{\gamma} \\ -\underline{g}_2 X + \underline{\gamma} \end{cases} \leq g^*(X,z) \leq \begin{cases} -\overline{g}_1 X + \overline{\gamma} & \text{if } X < 0 \\ \overline{g}_2 X + \overline{\gamma} & \text{if } X \geq 0 \end{cases}$$

$$\begin{cases} \underline{\mu}_1 X + \underline{m} \\ -\underline{\mu}_2 X + \underline{m} \end{cases} \leq \mu^*(X,z) \leq \begin{cases} -\overline{\mu}_1 X + \overline{m} & \text{if } X < 0 \\ \overline{\mu}_2 X + \overline{m} & \text{if } X \geq 0 \end{cases}$$

and

$$\begin{cases} \left(\underline{g}_1 - \frac{\overline{\mu}_1}{r} z\right) X + \underline{\gamma} + \frac{\overline{m}}{r} z \\ \left(-\underline{g}_2 + \frac{\overline{\mu}_2}{r} z\right) X + \underline{\gamma} + \frac{\overline{m}}{r} z \\ \left(\underline{g}_1 + \frac{\underline{\mu}_1}{r} z\right) X + \underline{\gamma} + \frac{\underline{m}}{r} z \\ -\left(\underline{g}_2 + \frac{\underline{\mu}_2}{r} z\right) X + \underline{\gamma} + \frac{\underline{m}}{r} z \end{cases} \leq \psi(X,z) \leq \begin{cases} \left(-\overline{g}_1 + \frac{\underline{\mu}_1}{r} z\right) X + \overline{\gamma} + \frac{\underline{m}}{r} z & \text{if } X < 0, z \leq 0 \\ \left(\overline{g}_2 - \frac{\underline{\mu}_2}{r} z\right) X + \overline{\gamma} + \frac{\underline{m}}{r} z & \text{if } X \geq 0, z \leq 0 \\ -\left(\overline{g}_1 + \frac{\overline{\mu}_1}{r} z\right) X + \overline{\gamma} + \frac{\overline{m}}{r} z & \text{if } X < 0, z \geq 0 \\ \left(\overline{g}_2 + \frac{\overline{\mu}_2}{r} z\right) X + \overline{\gamma} + \frac{\overline{m}}{r} z & \text{if } X \geq 0, z \geq 0 \end{cases}$$

Step 1b: Find conditions on $(\alpha_1, \beta_1, c_a, c_b)$ such that α, β are lower and upper solutions to the optimality function when $X \leq -1$. Note $\alpha''(X) = \beta''(X) = 0$, so this corresponds to showing $\psi(X, \alpha_1) \geq \alpha_1 X - c_a$ and $\psi(X, -\beta_1) \leq -\beta_1 X + c_b$. From the bound on ψ :

$$\begin{aligned} \psi(X, \alpha_1) &\geq \left(\underline{g}_1 + \frac{\underline{\mu}_1}{r}\alpha_1\right) X + \underline{\gamma} + \frac{\underline{m}}{r}\alpha_1 \\ \psi(X, -\beta_1) &\leq -\left(\overline{g}_1 + \frac{\underline{\mu}_1}{r}\beta_1\right) X + \overline{\gamma} - \frac{\underline{m}}{r}\beta_1 \end{aligned}$$

Therefore, this requires:

$$\begin{aligned}\alpha_1 &\geq \frac{r\underline{g}_1}{r-\underline{\mu}_1} \text{ and } c_a \geq -\underline{\gamma} - \frac{\underline{m}}{r}\alpha_1 := c_a^1 \\ \beta_1 &\geq \frac{r\overline{g}_1}{r-\underline{\mu}_1} \text{ and } c_b \geq \overline{\gamma} - \frac{\underline{m}}{r}\beta_1 := c_b^1\end{aligned}$$

Step 1c: Find conditions on $(\alpha_1, \beta_1, c_a, c_b)$ such that α, β are lower and upper solutions to the optimality function when $X \geq 1$. This corresponds to showing $\psi(X, -\alpha_1) \geq -\alpha_1 X - c_a$ and $\psi(X, \beta_1) \leq \beta_1 X + c_b$. From the bound on ψ :

$$\begin{aligned}\psi(X, -\alpha_1) &\geq -\left(\underline{g}_2 + \frac{\overline{\mu}_2}{r}\alpha_1\right) X + \underline{\gamma} - \frac{\overline{m}}{r}\alpha_1 \\ \psi(X, \beta_1) &\leq \left(\overline{g}_2 + \frac{\overline{\mu}_2}{r}\beta_1\right) X + \overline{\gamma} + \frac{\overline{m}}{r}\beta_1\end{aligned}$$

Therefore, this requires:

$$\begin{aligned}\alpha_1 &\geq \frac{r\underline{g}_2}{r-\overline{\mu}_2} \text{ and } c_a \geq -\underline{\gamma} + \frac{\overline{m}}{r}\alpha_1 := c_a^2 \\ \beta_1 &\geq \frac{r\overline{g}_2}{r-\overline{\mu}_2} \text{ and } c_b \geq \overline{\gamma} + \frac{\overline{m}}{r}\beta_1 := c_b^1\end{aligned}$$

Step 1d: Find conditions on $(\alpha_1, \beta_1, c_a, c_b)$ such that α, β are lower and upper solutions to the optimality function when $X \in (-1, 1)$. Note $\alpha''(X) = -\frac{3}{2}\alpha_1 (1 - X^2) \geq -\frac{3}{2}\alpha_1$ and $\alpha(X) \leq -\frac{3}{8}\alpha_1 - c_a$ and $\beta''(X) = \frac{3}{2}\beta_1 (1 - X^2) \leq \frac{3}{2}\beta_1$ and $\beta(X) \geq \frac{3}{8}\beta_1 + c_b$, so this is equivalent to showing:

$$\begin{aligned}c_a &\geq \frac{3}{4}\left(\frac{|\sigma^*(X, \alpha')|^2}{r} - \frac{1}{2}\right)\beta_1 - \psi(X, \alpha') \\ c_b &\geq \frac{3}{4}\left(\frac{|\sigma^*(X, \alpha')|^2}{r} - \frac{1}{2}\right)\beta_1 + \psi(X, \beta')\end{aligned}$$

for $X \in (-1, 1)$. Let $\overline{\sigma} = \sup_{X \in [0,1], z} |\sigma^*(X, z)|$, which exists since σ^* is Lipschitz continuous in X and bounded in z.

First consider $X \in (-1, 0]$, which means that $\beta' = \frac{1}{2}\beta_1 X (3 - X^2) \in (-\beta_1, 0]$ and $a' =$

$-\frac{1}{2}\alpha_1 X\left(3-X^2\right)\in[0,\alpha_1)$. From the bound on ψ :

$$\begin{aligned}\psi\left(X,\alpha'\right) &\geq \left(\underline{g}_1+\frac{\underline{\mu}_1}{r}\alpha'\right)X+\underline{\gamma}+\frac{m}{r}\alpha'\geq -\underline{g}_1+\underline{\gamma}-\frac{\underline{\mu}_1}{r}\alpha_1+\frac{\alpha_1}{r}\min\left\{\underline{m},0\right\}\\ \psi\left(X,\beta'\right) &\leq \left(-\overline{g}_1+\frac{\underline{\mu}_1}{r}\beta'\right)X+\overline{\gamma}+\frac{m}{r}\beta'\leq \overline{g}_1+\overline{\gamma}+\frac{\underline{\mu}_1}{r}\beta_1-\frac{\beta_1}{r}\min\left\{\underline{m},0\right\}\end{aligned}$$

Therefore, this requires:

$$\begin{aligned}c_a &\geq \frac{3}{4}\left(\frac{\overline{\sigma}^2}{r}-\frac{1}{2}\right)\alpha_1+\underline{g}_1-\underline{\gamma}+\frac{\underline{\mu}_1}{r}\alpha_1-\frac{\alpha_1}{r}\min\left\{\underline{m},0\right\}:=c_a^3\\ c_b &\geq \frac{3}{4}\left(\frac{\overline{\sigma}^2}{r}-\frac{1}{2}\right)\beta_1+\overline{g}_1+\overline{\gamma}+\frac{\underline{\mu}_1}{r}\beta_1-\frac{\beta_1}{r}\min\left\{\underline{m},0\right\}:=c_b^3\end{aligned}$$

Next consider $X\in[0,1)$, which means that $\beta'=\frac{1}{2}\beta_1 X\left(3-X^2\right)\in[0,\beta_1)$ and $a'=-\frac{1}{2}\alpha_1 X\left(3-X^2\right)\in(-\alpha_1,0]$. From the bound on ψ :

$$\begin{aligned}\psi\left(X,\alpha'\right) &\geq \left(-\underline{g}_2+\frac{\overline{\mu}_2}{r}\alpha'\right)X+\underline{\gamma}+\frac{\overline{m}}{r}\alpha'\geq -\underline{g}_2+\underline{\gamma}-\frac{\overline{\mu}_2}{r}\alpha_1-\frac{\alpha_1}{r}\max\left\{\overline{m},0\right\}\\ \psi\left(X,\beta'\right) &\leq \left(\overline{g}_2+\frac{\overline{\mu}_2}{r}\beta'\right)X+\overline{\gamma}+\frac{\overline{m}}{r}\beta'\leq \overline{g}_2+\overline{\gamma}+\frac{\overline{\mu}_2}{r}\beta_1+\frac{\beta_1}{r}\max\left\{\overline{m},0\right\}\end{aligned}$$

This requires:

$$\begin{aligned}c_a &\geq \frac{3}{4}\left(\frac{\overline{\sigma}^2}{r}-\frac{1}{2}\right)\alpha_1+\underline{g}_2-\underline{\gamma}+\frac{\overline{\mu}_2}{r}\alpha_1+\frac{\alpha_1}{r}\max\left\{\overline{m},0\right\}:=c_a^4\\ c_b &\geq \frac{3}{4}\left(\frac{\overline{\sigma}^2}{r}-\frac{1}{2}\right)\beta_1+\overline{g}_2+\overline{\gamma}+\frac{\overline{\mu}_2}{r}\beta_1+\frac{\beta_1}{r}\max\left\{\overline{m},0\right\}:=c_b^4\end{aligned}$$

Step 1e: Compiling these conditions and choosing

$$\begin{aligned}\alpha_1 &= \max\left\{\frac{r\underline{g}_1}{r-\underline{\mu}_1},\frac{r\underline{g}_2}{r-\overline{\mu}_2}\right\}\\ \beta_1 &= \max\left\{\frac{r\overline{g}_1}{r-\underline{\mu}_1},\frac{r\overline{g}_2}{r-\overline{\mu}_2}\right\}\end{aligned}$$

yields $\alpha_1,\beta_1\geq 0$ that satisfy the slope conditions in steps 1b-1d, and choosing

$$c_a=\max\left\{0,c_a^1,c_a^2,c_a^3,c_a^4\right\}\text{ and }c_b=\max\left\{0,c_b^1,c_b^2,c_b^3,c_b^4\right\}$$

yields c_a,c_b that satisfy the intercept conditions in steps 1b-1d. Conclude that α and β are lower and upper solutions to the optimality equation.

Step 2: Assume g is bounded. Show that there exist $\alpha_1, \beta_1, c_a, c_b \geq 0$ such that α, β are lower and upper solutions to the optimality function. Note step 2 places no restrictions on the relationship between the growth rate of μ_x and r.

Define $\overline{g} := \sup_{a,X} g(a, \overline{\mathcal{B}}(a, X), X)$ and $\underline{g} := \inf_{a,X} g(a, \overline{\mathcal{B}}(a, X), X)$, which exist since g is bounded. Set $\alpha_1 = 0$ and $c_a = -\underline{g}$. Then $\psi(X, \alpha') = g^*(X, 0)$, so $\alpha - \psi(X, \alpha') = \underline{g} - g^*(X, 0) \leq 0$ and $\alpha(X) = \underline{g}$ is a lower solution. Similarly, set $\beta_1 = 0$ and $c_b = \overline{g}$. Then $\psi(X, \beta') = g^*(X, 0)$, so $\beta - \psi(X, \beta') = \overline{g} - g^*(X, 0) \geq 0$ and $\beta(X) = \overline{g}$ is an upper solution.

Step 3: Show that the Nagumo condition is satisfied. Given a compact proper subset $I \subset \Xi$, there exists a $K_I > 0$ such that

$$|f(X, U, U')| = \left| \frac{2r}{|\sigma^*(X, U')|^2} \left(U - g^*(X, U') - \frac{U'}{r} \mu^*(X, U') \right) \right| \leq K_I (1 + |U'|)$$

for all $(X, U, U') \in \{I \times \mathbb{R}^2 \text{ s.t. } \alpha(X) \leq U \leq \beta(X)\}$. This follows directly from the fact that $X \in I$, $\alpha(X)$ and $\beta(X)$ are bounded on I, $\alpha(X) \leq U \leq \beta(X)$, g^*, μ^* are bounded on $(X, U') \in I \times \mathbb{R}$ and the lower bound on σ_x. Define $\phi_I(z) = K_I (1 + z)$. Therefore, $\int_0^\infty \phi_I(s)^{-1} s ds = \infty$.

Conclude that $f(X, U(X), U'(X))$ has at least one C^2 solution U such that for all $X \in \mathbb{R}$, $\alpha(X) \leq U \leq \beta(X)$. If α and β are bounded, then U is bounded.

Case 2: Bounded State Space. I use standard existence results from Coster and Habets (2006) and an extension in Faingold and Sannikov (2011), which is necessary because the optimality equation is undefined at the upper and lower bound of the state space. The result applied to the current setting is reproduced below.

Lemma 4 (Faingold Sannikov (2011))**.** *Let $E = \left\{(t, u, v) \in (\underline{t}, \overline{t}) \times \mathbb{R}^2\right\}$ and $f : E \to \mathbb{R}$ be continuous. Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha \leq \beta$ and $f(t, \alpha, 0) \leq 0 \leq f(t, \beta, 0)$ for all $t \in \mathbb{R}$. Assume also that for any closed interval $I \subset (\underline{t}, \overline{t})$, there exists a $K_I > 0$ such that $|f(t, u, v)| \leq K_I \left(1 + |v|^2\right)$ for all $(t, u, v) \in I \times [\alpha, \beta] \times \mathbb{R}$. Then the differential equation $U'' = f(t, U(t), U'(t))$ has at least one $\mathcal{C}^2$ solution U on $(\underline{t}, \overline{t})$ such that $\alpha \leq U(t) \leq \beta$.*

When Ξ is compact, the feasible payoff set for the long-run player is also bounded, since g is Lipschitz continuous and A is compact. Define $\underline{g} := \inf_{A,\Xi} g(a, \overline{\mathcal{B}}(a, X), X)$ and $\overline{g} := \sup_{A,\Xi} g(a, \overline{\mathcal{B}}(a, X), X)$.

Lemma 5. *Suppose Ξ is compact and there are no interior absorbing points. The optimality equation has at least one $\mathcal{C}^2$ solution U that lies in the range of feasible payoffs for the long-run player i.e. $\underline{g} \leq U(X) \leq \overline{g}$.*

Suppose Ξ is compact and there are no interior absorbing points. Then the optimality equation is continuous on the set $E = \left\{(X,U,U') \in (\underline{X},\overline{X}) \times \mathbb{R}^2\right\}$. For any closed interval $I \subset (\underline{X},\overline{X})$, there exists a $K_I > 0$ such that

$$|f(X,U,U')| = \left|\frac{2r}{|\sigma^*(X,U')|^2}\left(U - g^*(X,U') - \frac{U'}{r}\mu^*(X,U')\right)\right| \leq K_I\,(1+|U'|)$$

for all $(X,U,U') \in I \times \left[\underline{g},\overline{g}\right] \times \mathbb{R}$. This follows directly from the fact that $X \in I$, $U \in \left[\underline{g},\overline{g}\right]$, g^*, μ^* are bounded on $(X,U') \in I \times \mathbb{R}$, and the lower bound on σ_x. Therefore, $|f(X,U,U')| \leq 2K_I\left(1+|U'|^2\right)$. Also note that

$$f(X,\underline{g},0) = \frac{2r}{|\sigma^*(X,0)|^2}\left(\underline{g} - g^*(X,0)\right) \leq 0 \leq f(X,\overline{g},0) = \frac{2r}{|\sigma^*(X,0)|^2}\left(\overline{g} - g^*(X,0)\right)$$

for all $X \in \Xi$. By Lemma 4, $f(X,U(X),U'(X))$ has at least one C^2 solution U such that for all $X \in \Xi$, $\underline{g} \leq U(X) \leq \overline{g}$.

Case 3: Absorbing States. By Assumption 1, the state space can be represented as the union of a finite number of subsets, $\Xi = \cup_{i=1}^n \Xi_i$ such that for any compact proper subset $I \subset \Xi_i$, there exists a c_I such that $\sigma_I = \inf_{b\in B, X\in I} |\sigma_x(b,X)|^2 > c_I$. By Lemmas 3 and 5, there exists a solution on each subset Ξ_i.[21] Also, at each absorbing point $\widetilde{X}$, $U(\widetilde{X}) = v(\widetilde{X})$, by Lemma 24. Pasting the solutions on each subset together yields a continuous solution on Ξ, which is $\mathcal{C}^2$ on $\Xi \setminus \tilde{\Xi}$.

Construct a Markov equilibrium: Suppose the state variable initially starts at X_0 and U is a solution to the optimality equation. The action profile

$$(a^*,\overline{b}^*) = \begin{cases} S^*(X,0,U'(X)/r) & \text{if } X \in \Xi \setminus \widetilde{\Xi} \\ S(X,0,0) & \text{if } X \in \widetilde{\Xi} \end{cases}$$

is unique and Lipschitz continuous in X and U'. Given X_0, U and $(a_t^*,\overline{b}_t^*)_{t\geq 0}$, the state variable evolves according to the unique strong solution $(X_t)_{t\geq 0}$ to the stochastic differential equation

$$dX_t = \mu^*(X_t,U'(X_t))dt + \sigma^*\left(X_t,U'(X_t)\right)dZ_t$$

which exists since μ^* and σ^* are Lipschitz continuous. The continuation value evolves ac-

[21] It is a straightforward extension to show that a solution exists on an interval $\Xi_i = [-\infty, X_1)$ or $\Xi_i = [X_2,\infty)$.

cording to:

$$\begin{aligned} dU(X_t) &= U'(X_t)\mu^*(X_t, U'(X_t))dt + \frac{1}{2}U''(X_t)\left|\sigma^*\left(X_t, U'(X_t)\right)\right|^2 dt + U'(X_t)\sigma^*\left(X_t, U'(X_t)\right)dZ_t \\ &= r\left(U(X_t) - g^*\left(X_t, U'(X_t)\right)\right)dt + U'(X_t)\sigma^*\left(X_t, U'(X_t)\right)dZ_t \end{aligned}$$

on $\Xi \backslash \widetilde{\widetilde{\Xi}}$ and $dU(X_t) = 0$ on $X \in \widetilde{\widetilde{\Xi}}$. This process satisfies the conditions for the continuation value in a PPE characterized in Lemma 2. Additionally, $(a_t^*, \bar{b}_t^*)_{t\geq 0}$ satisfies the condition for sequential rationality given process $(\beta_t)_{t\geq 0}$ with $\beta_t = (0, U'(X_t))$ if $X \in \Xi \backslash \widetilde{\widetilde{\Xi}}$ and $\beta_t = (0,0)$ if $X \in \widetilde{\widetilde{\Xi}}$. Thus, the strategy profile $(a_t^*, \bar{b}_t^*)_{t\geq 0}$ is a PPE yielding equilibrium payoff $U(X_0)$.

7.5 Proof of Theorem 2

Let X_0 be the initial state, and let $V(X)$ be the upper envelope of the set of solutions to the optimality equation. Suppose there is a PPE $(a_t, \bar{b}_t)_{t\geq 0}$ that yields an equilibrium payoff $W_0 > V(X_0)$. The continuation value in this equilibrium must evolve according to:

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

for some process $(\beta_t)_{t\geq 0}$. By sequential rationality, $(a_t, \bar{b}_t) = S^*(X_t, \beta_{yt}, \beta_{xt})$ for all t, and by Assumption 4, these actions are unique for each (X, β_y, β_x). Define

$$\begin{aligned} \widehat{g}(X, \beta_y, \beta_x) &:= g(S^*(X, \beta_y, \beta_x), X) \\ \widehat{\mu}(X, \beta_y, \beta_x) &:= \mu(S^*(X, \beta_y, \beta_x), X) \\ \widehat{\sigma}(X, \beta_y, \beta_x) &:= \sigma(S^*(X, \beta_y, \beta_x), X) \end{aligned}$$

which are Lipschitz continuous, given g, μ, σ and S^* are Lipschitz. The process $V(X_t)$ evolves according to:

$$dV(X_t) = V'(X_t)\widehat{\mu}(X_t, \beta_{yt}, \beta_{xt})dt + \frac{1}{2}V''(X_t)\left|\widehat{\sigma}_x\left(X_t, \beta_{yt}, \beta_{xt}\right)\right|^2 dt + V'(X_t)\widehat{\sigma}_x\left(X_t, \beta_{yt}, \beta_{xt}\right)dZ_t$$

on $\Xi \backslash \widetilde{\widetilde{\Xi}}$ and $dV(X_t) = 0$ on $X \in \widetilde{\widetilde{\Xi}}$. Define a process $D_t = W_t - V(X_t)$ with initial condition $D_0 = W_0 - V(X_0) > 0$. Then D_t evolves with drift $rD_t + d(X_t, \beta_{yt}, \beta_{xt})$, where

$$\begin{aligned} d(X, \beta_y, \beta_x) &= r\left(V(X) - \widehat{g}(X, \beta_y, \beta_x)\right) - V'(X)\mu_x(X, \beta_y, \beta_x) - \frac{1}{2}V''(X)\left|\widehat{\sigma}_x\left(X, \beta_y, \beta_x\right)\right|^2 \\ &= r\left(\widehat{g}(X, 0, V'/r) - \widehat{g}(X, \beta_y, \beta_x)\right) + V'(X)\left(\mu_x(X, 0, V'/r) - \mu_x(X, \beta_y, \beta_x)\right) \end{aligned}$$

if $X \in \Xi \setminus \widetilde{\Xi}$, where the second line follows from substituting the optimality equation for $V(X)$, and $d(X, \beta_y, \beta_x) = r(\widehat{g}(X, 0, 0) - \widehat{g}(X, \beta_y, \beta_x))$ if $X \in \widetilde{\Xi}$, and volatility

$$f(X, \beta_y, \beta_x) = \begin{cases} \begin{bmatrix} r\beta_y^{\intercal}\widehat{\sigma}_y\left(X, \beta_y, \beta_x\right) \\ \left(r\beta_x - V'(X)\right)\widehat{\sigma}_x\left(X, \beta_y, \beta_x\right) \end{bmatrix} & \text{if } X \in \Xi \setminus \widetilde{\Xi} \\ 0 & \text{if } X \in \widetilde{\Xi} \end{cases}$$

Lemma 6. *If $|f(X, \beta_y, \beta_x)| = 0$, then $d(X, \beta_y, \beta_x) = 0$.*

Suppose $|f(X, \beta_y, \beta_x)| = 0$ for some (X, β_y, β_x). Then $\beta_y = 0$ and

$$r\beta_x = \begin{cases} V'(X) & \text{if } X \in \Xi \setminus \widetilde{\Xi} \\ 0 & \text{if } X \in \widetilde{\Xi} \end{cases}$$

The action profile associated with $(X, 0, V'(X)/r)$ corresponds to the actions played in a Markov equilibrium at state X. Therefore, $d(X, \beta_y, \beta_x) = 0$.

Lemma 7. *For every $\varepsilon > 0$, there exists a $\eta > 0$ such that either $d(X, \beta_y, \beta_x) > -\varepsilon$ or $|f(X, \beta_y, \beta_x)| > \eta$.*

Suppose the state space is unbounded, $\Xi = \mathbb{R}$, and there are no interior absorbing states. Fix $\varepsilon > 0$ and suppose $d(X, \beta_y, \beta_x) \leq -\varepsilon$. Show that there exists a $\eta > 0$ such that $|f(X, \beta_y, \beta_x)| > \eta$ for all $(X, \beta) \in \Xi \times \mathbb{R}$.

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

$$(X, \beta) \in \Omega_a = \{\Xi \times \mathbb{R} : |\beta| > M\ \}$$

$V'(X)$ 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 $y \in \mathbb{R}^d$, bounding σ away from 0, so there exists an $M > 0$ and $\eta_1 > 0$ such that $|f(X, \beta_y, \beta_x)| > \eta_1$ for all $|\beta| > M$, regardless of d.

Step (b): Show that there exists an $X^* > 0$ such that this is true for:

$$(X, \beta) \in \Omega_b = \{\Xi \times \mathbb{R} : |\beta| \leq M \text{ and } |X| > X^*\}$$

Consider the set $\Phi_b \subset \Omega_b$ with $d(X, \beta_y, \beta_x) \leq -\varepsilon$. It must be that (β_x, β_y) is bounded away from $(V'(X)/r, 0)$ on Φ_b. Suppose not. Then either (i) there exists some $(X, \beta) \in \Phi_b$ with $\beta_x = V'(X)/r$ and $\beta_y = 0$, which implies $f(X, \beta_y, \beta_x) = 0$ and therefore $d(X, \beta_y, \beta_x) = 0$, a contradiction, or (ii) as X becomes large, the boundary of the set Φ_b approaches $(\beta_x, \beta_y) = (V'/r, 0)$, which implies that for any $\delta_1 > 0$, there exists an $(X, \beta) \in \Phi_b$ with

$\max\{r\beta_x - V'(X), \beta_y\} < \delta_1$. Choose δ_1 so that $|\widehat{g}(X, 0, V'(X)/r) - \widehat{g}(X, \beta_y, \beta_x)| < \varepsilon/4r$ and $|\widehat{\mu}(X, 0, V'(X)/r) - \widehat{\mu}(X, \beta_y, \beta_x)| < \varepsilon/4k$, where $|V'(X)| \leq k$ is the bound on V'. Then

$$|d(X, \beta_y, \beta_x)| < \varepsilon/4 + \varepsilon/4 = \varepsilon/2$$

which is a contradiction. Therefore, there exists a η_2 such that $|f(X, \beta_y, \beta_x)| > \eta_2$ on Φ_b. Then on the set Ω_b, if $d(X, \beta_y, \beta_x) \leq -\varepsilon$ then $|f(X, \beta_y, \beta_x)| > \eta_2$.

Step (c): Show this is true for:

$$(X, \beta) \in \Omega_c = \{\Xi \times \mathbb{R} : |\beta| \leq M \text{ and } |X| \leq X^*\}$$

Consider the set $\Phi_c \subset \Omega_c$ where $d(X, \beta_y, \beta_x) \leq -\varepsilon$. The function d is continuous and Ω_c is compact, so Φ_c is compact. The function $|f|$ is also continuous, and therefore achieves a minimum η_3 on Φ_c. If $\eta_3 = 0$, then $d = 0$ by Lemma 6, a contradiction. Therefore, $\eta_3 > 0$ and $|f(X, \beta_y, \beta_x)| > \eta_3$ for all $(X, \beta) \in \Phi_c$.

Take $\eta = \min\{\eta_1, \eta_2, \eta_3\}$. Then when $d(X, \beta_y, \beta_x) \leq -\varepsilon$, $|f(X, \beta_y, \beta_x)| > \eta$. The proof for a bounded state space is analogous, omitting step 2b. To allow for interior absorbing points, simply apply the proof to each Ξ_i that doesn't contain an absorbing point. Q.E.D.

Lemma 8. *Any PPE payoff W_0 is such that $V_1(X_0) \leq W_0 \leq V_2(X_0)$ where V_1 and V_2 are the upper and lower envelope of the set of solutions to the optimality equation.*

Lemma 7 implies that whenever the drift of D_t is less than $rD_t - \varepsilon$, the volatility is greater than η. 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 η such that $|f(X, \beta_y, \beta_x)| > \eta$. Thus, whenever $D_t \geq D_0/2 > 0$, it has either positive drift or positive volatility, and grows arbitrarily large with positive probability, irrespective of X_t. This is a contradiction, since D_t is the difference of two processes that are bounded with respect to X_t. Thus, it cannot be that $D_0 > 0$ and it must be the case that $W_0 \leq V(X_0)$. Similarly, if $\underline{V}$ is the lower envelope of the set of solutions to the optimality equation, it is not possible to have $D_0 < 0$ and therefore it must be the case that $W_0 \geq \underline{V}(X_0)$. Q.E.D.

The second part of Theorem 2 follows directly from Lemma 8, and the fact that it is possible the achieve any equilibrium payoff in the convex hull of the highest and lowest Markov equilibrium payoff with randomization. Q.E.D.

7.6 Proof of Theorem 3

The proof proceeds in three steps.

1. Any solution to the optimality equation has the same boundary conditions.

2. If all solutions have the same boundary conditions, then there is a unique solution.

3. When there is a unique solution, then there is a unique PPE.

Step 1a: Boundary Conditions for Unbounded Support. Suppose $\Xi = \mathbb{R}$, g is bounded and there are no absorbing points.

Lemma 9. *The limits* $g_p = \lim_{X \to p} g^*(X, 0)$ *exist for* $p \in \{-\infty, \infty\}$.

The fact that g is bounded, $g^*(X, 0)$ is Lipschitz continuous and monotone for large enough $|X|$ guarantees the existence of $\lim_{X \to p} g^*(X, 0)$. Q.E.D.

Lemma 10. *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,* $U_p = \lim_{X \to p} U(X)$ *exist for* $p \in \{-\infty, \infty\}$.

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 static Nash equilibria is played when $U'(X) = 0$, yielding flow payoff $g^*(X, 0)$. From the optimality equation, this implies $g^*(X_n, 0) \geq U(X_n)$ at the maximum and $g^*(X_m, 0) \leq U(X_m)$ at the minimum. Thus, the oscillation of $g^*(X, 0)$ is at least as large as the oscillation of $U(X)$. This is a contradiction, as there exists a δ such that for $|X| > \delta$, $g^*(X, 0)$ is monotonic.The fact that U is bounded and monotone for large enough X guarantees the existence of $\lim_{X \to p} U(X)$. Q.E.D.

Lemma 11. *Suppose a function* $f(X)$ *is Lipschitz continuous on* $\mathbb{R}$. *Then any bounded solution* U *of the optimality equation satisfies:*

$$\begin{aligned} \lim_{X \to p} \inf |f(X)| \, U'(X) &\leq 0 \leq \lim_{X \to p} \sup |f(X)| \, U'(X) \\ \lim_{X \to p} \inf f(X)^2 U''(X) &\leq 0 \leq \lim_{X \to p} \sup f(X)^2 U''(X) \end{aligned}$$

for $p \in \{-\infty, \infty\}$.

Suppose $f(X)$ is Lipschitz continuous and $\lim_{X \to p} \inf |f(X)| \, U'(X) > 0$. By Lemma 31, 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 $\lim_{X \to p} \inf |f(X)| \, U'(X) > 0$, there exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $|X| >$

δ_2, $|f(X)|\, U'(X) > \varepsilon$. Take $\delta = \max\{\delta_1, \delta_2\}$. Then for $|X| > \delta$, $|U'(X)| > \frac{\varepsilon}{|f(X)|} \geq \frac{\varepsilon}{M|X|}$. Then the antiderivative of $\frac{\varepsilon}{M|X|}$ is $\frac{\varepsilon}{M} \ln |X|$ which converges to ∞ as $|X| \to \infty$. This violates the boundedness of U. Therefore $\lim_{X\to p} \inf |f(X)|\, U'(X) \leq 0$. The proof is analogous for the other case. Suppose $f(X)$ is Lipschitz continuous and $\lim_{X\to\infty} \inf f(X)^2 U''(X) > 0$. By Lemma 31, f is $O(X)$, so there exists an $M \in \mathbb{R}$ and a $\delta_1 \in \mathbb{R}$ such that when $|X| > \delta_1, |f(X)| \leq M\,|X|$ and therefore, $f(X)^2 \leq M^2 X^2$. There also exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $|X| > \delta_2$, $f(X)^2 U''(X) > \varepsilon$. Take $\delta = \max\{\delta_1, \delta_2\}$. Then for $|X| > \delta$, $|U''(X)| > \frac{\varepsilon}{f(X)^2} > \frac{\varepsilon}{M^2 X^2}$. Then the antiderivative of $\frac{\varepsilon}{M^2 X^2}$ is $\frac{-\varepsilon}{M^2} \ln |X|$ which converges to $-\infty$ as $|X| \to \infty$. This violates the boundedness of U. Therefore $\lim_{X\to p} \inf f(X)^2 U''(X) \leq 0$. The proof is analogous for the other case. Q.E.D.

Lemma 12. *Suppose $f(X)$ is Lipschitz continuous on $\mathbb{R}$. Then any bounded solution U of the optimality equation satisfies $\lim_{X\to p} f(X)U'(X) = 0$ for $p \in \{-\infty, \infty\}$.*

Suppose that $\lim_{X\to p} \sup |X|\, U'(X) > 0$. By Lemma 10, there exists a $\delta > 0$ such that $U(X)$ is monotonic for $|X| > \delta$. Then for $|X| > \delta$, $XU'(X)$ doesn't change sign. Therefore, if $\lim_{X\to p} \sup |X|\, U'(X) > 0$, then $\lim_{X\to p} \inf |X|\, U'(X) > 0$. This is a contradiction. Thus, $\lim_{X\to p} \sup |X|\, U'(X) = 0$. By similar reasoning, $\lim_{X\to p} \inf |X|\, U'(X) = 0$, and therefore $\lim_{X\to p} |X|\, U'(X) = 0$. 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 $|X| > \delta_1$, $|f(X)U'(X)| \leq M\,|XU'(X)| \to 0$. Note this result also implies that $\lim_{X\to p} U'(X) = 0$. Q.E.D.

Lemma 13. *Let U be a bounded solution of the optimality equation. Then the limit of U converges to the limit of the static Nash payoff, $U_p = g_p$ for $p \in \{-\infty, \infty\}$.*

By Lemmas 9 and 10, g_p and U_p exist. Suppose $U_p < g_p$. The equilibrium drift $\mu^*(X, U')$ is Lipschitz continuous, so by Lemma 12, $\lim_{X\to p} U'\mu^*(X, U') = 0$. By the Lipschitz continuity of g^*,

$$\lim_{X\to p} g^*(X, U') = \lim_{X\to p} g^*(X, 0) = g_p.$$

Plugging the above conditions in to the optimality equation,

$$\begin{aligned} \limsup_{X\to p} |\sigma^*(X, U')|^2\, U'' &= \limsup_{X\to p} 2r\,(U - g^*(X, U')) - 2\mu^*(X, U')U' \\ &= 2r\,(U_p - g_p) < 0 \end{aligned}$$

This violates Lemma 11 since σ^* is Lipschitz continuous. A similar contradiction holds for $U_p > g_p$. Thus, $U_p = g_p$. Q.E.D.

Lemma 14. *Any bounded solution U of the optimality equation satisfies*

$$\lim_{X\to p} \left|\sigma^*(X,U')\right|^2 U''(X) = 0$$

for $p \in \{-\infty, \infty\}$. Note this also implies $U''(X) \to 0$, since $\lim_{X\to p} \left|\sigma^*(X,U')\right|^2 > 0$.

By Lemmas 12 and 13 and the squeeze theorem,

$$\lim_{X\to p} \left| \left|\sigma^*(X,U')\right|^2 U'' \right| = \lim_{X\to p} \left|2r\left(U - g^*(X,U')\right) - 2\mu^*(X,U')U'\right| = 0$$

Step 1b: Boundary Conditions for Unbounded Support and Unbounded g. Suppose $\Xi = \mathbb{R}$, g is unbounded and there are no absorbing points. Let $\overline{\psi}(X,k) = \psi(X,k)/X$, and $\overline{U}(X) = U(X)/X$, and ψ' and $\overline{\psi}'$ refer to the partial derivative with respect to X.

Lemma 15. *For $p \in \{-\infty, \infty\}$, there exists a function $\psi_p(k)$ such that*

$$\lim_{X\to p} \overline{\psi}(X,k) = \lim_{X\to p} \psi'(X,k) := \psi_p(k).$$

Given that g^* and μ^* are Lipschitz continuous, there exists $M, X_0 \in R$ such that $\left|\overline{\psi}(X,k)\right| \leq M(1+k)$ for all $|X| > X_0$ and $k \in R$. Therefore, $\overline{\psi}(X,k)$ is bounded in X for large enough X. By Assumption 5, $\psi'(X,k)$ is monotone in X for large enough X. By Lemma 30, $\overline{\psi}(X,k)$ is monotone in X for large enough X. By Lemma 28 and the fact that $\overline{\psi}$ continuous (except possibly at 0), $\lim_{X\to p} \overline{\psi}(X,k)$ exists and , $\lim_{X\to p} X\overline{\psi}'(X,k) = 0$. By Lemma 29, $\lim_{X\to p} \psi'(X,k) = \lim_{X\to p} \overline{\psi}(X,k)$. Q.E.D.

Lemma 16. *If U is a solution of the optimality equation with linear growth, then there exists a δ such that for $|X| > \delta$, U' and $\overline{U}$ are monotonic. Additionally, for $p \in \{-\infty, \infty\}$,*

$$\lim_{X\to p} \overline{U}(X) = \lim_{X\to p} U'(X) \ := U_p$$

exist and are equal.

Suppose U' is not eventually monotonic. Then for any δ, there exists a k and a $|X_n|, |X_m| > \delta$ such that $U'(X_n) = k$ and $U''(X_n) \leq 0$ and $U'(X_m) = k$ and $U''(X_m) \geq 0$, by the continuity of U'. From the optimality equation, this implies $U(X_n) \leq \psi(X_n,k)$ and $\psi(X_m,k) \leq U(X_m)$. Thus, the oscillation of $\psi(X,k)$ is at least as large as the oscillation of U. Given that $\psi'(X,k)$ is eventually monotonic, it must be that U' is also eventually monotonic, a contradiction. Therefore, it must be that U' is monotonic. By Lemma 30, if U' is eventually monotonic, then $\overline{U}$ is also eventually monotonic. Note that $\overline{U}$ is bounded, given that U has linear growth. Therefore, since $\overline{U}$ is eventually monotonic, bounded and differentiable, by Lemma 28, $\lim_{X\to p}\overline{U}(X)$ exists and $\lim_{X\to p}X\overline{U}'(X)=0$ for $p\in\{-\infty,\infty\}$. Therefore, by Lemma 29, $\lim_{X\to p}U'(X)=\lim_{X\to p}\overline{U}(X)$, which establishes existence and equality. Q.E.D.

Lemma 17. *Suppose a function $f(X)$ is Lipschitz continuous on $\mathbb{R}$. Then any solution U of the optimality equation with linear growth satisfies:*

$$\lim_{X\to p}\inf|f(X)|\,U''(X)\leq 0\leq\lim_{X\to p}\sup|f(X)|\,U''(X)$$

for $p\in\{-\infty,\infty\}$.

Suppose $f(X)$ is Lipschitz continuous and $\lim_{X\to p}\inf|f(X)|\,U''(X)>0$. Then f has linear growth, 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 $\lim_{X\to p}\inf|f(X)|\,U''(X)>0$, 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 $X>\delta$, $U''(X)>\frac{\varepsilon}{|f(X)|}\geq\frac{\varepsilon}{M|X|}$. Then the antiderivative of $\frac{\varepsilon}{M|X|}$ is $\frac{\varepsilon}{M}\ln|X|$ which converges to ∞ as $X\to p$. Therefore, U' must grow unboundedly large as $X\to p$, which violates the result that U has linear growth. Therefore $\lim_{X\to p}\inf|f(X)|\,U''(X)\leq 0$. The proof is analogous for the other case. Q.E.D.

Lemma 18. *Suppose a function $f(X)$ is Lipschitz continuous on $\mathbb{R}$.. Then any solution U of the optimality equation with linear growth satisfies*

$$\lim_{X\to p}f(X)U''(X)=0$$

for $p\in\{-\infty,\infty\}$.

Suppose that $\lim_{X\to p}\sup|X|\,U''(X)>0$. By Lemma 29, there exists a $\delta>0$ such that $U'(X)$ is monotonic for $|X|>\delta$. Then for $|X|>\delta$, $|X|\,U''(X)$ doesn't change sign. Therefore, if $\lim_{X\to p}\sup|X|\,U''(X)>0$, then $\lim_{X\to p}\inf|X|\,U''(X)>0$. This is a contradiction. Thus, $\lim_{X\to p}\sup|X|\,U''(X)=0$. By similar reasoning, $\lim_{X\to p}\inf|X|\,U''(X)=0$, and therefore $\lim_{X\to p}|X|\,U''(X)=0$. Suppose $f(X)$ has linear growth. 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 $|X|>\delta_1$, $|f(X)U''(X)|\leq M\,|XU''(X)|\to 0$. Note this result also implies that $\lim_{X\to p}U''(X)=0$. Q.E.D.

Lemma 19. *Suppose U is a solution of the optimality equation with linear growth. Then*

$$\lim_{X\to p}U'(X)=\lim_{X\to p}\overline{U}(X)=k_p$$

for $p \in \{-\infty, \infty\}$*, where* k_p *is the unique fixed point of* $\psi_p(k)$*.*

Consider $p \in \{-\infty, \infty\}$ and let U be a solution of the optimality equation. Note that $|\sigma^*(X, U')|^2 / X$ has linear growth, given that σ^* is Lipschitz continuous. Therefore, by Lemma 18,

$$\lim_{X \to p} \frac{|\sigma^*(X, U')|^2}{X} U''(X) = 0$$

which, from the optimality equation, implies

$$\lim_{X \to p} \overline{U}(X) - \overline{\psi}(X, U') = 0$$

From Lemma 16 that $\lim_{X \to p} U'(X) = \lim_{X \to p} \overline{U}(X) = U_p$ exists, and given that $\overline{\psi}$ is continuous and has a limit at p, $\lim_{X \to p} \overline{\psi}(X, U') = \lim_{X \to p} \overline{\psi}(X, U_p) = \psi_p(U_p)$. Therefore, $U_p - \psi_p(U_p) = 0$, and U_p must be a fixed point of $\psi_p(k)$.

By Assumption 5, g^* and μ^* are additively separable in (X, k). Therefore, $\psi_p(k)$ has a unique fixed point:

$$\begin{aligned} k_p &= \lim_{X \to p} \frac{g_1(X)}{X} + \frac{k_p}{r} \frac{\mu_1(X)}{X} \\ \Rightarrow k_p &= \lim_{X \to p} \frac{r g_1(X)/X}{r - \mu_1(X)/X} \end{aligned}$$

and $U_p = k_p$. Q.E.D.

Lemma 20. *Suppose* U *is a solution of the optimality equation with linear growth. Then*

$$\lim_{X \to p} U(X) - y_p^L(X) = 0$$

for $p \in \{-\infty, \infty\}$*, where*

$$y_p^L(X) \quad : \quad = -r \exp\left(\int \frac{r}{\mu_1(X)} dX \right) \int \exp\left(- \int \frac{r}{\mu_1(X)} dX \right) \frac{g_1(X)}{\mu_1(X)} dX + \phi_p$$

and $\phi_p = g_2(k_p) + \frac{k_p}{r} \mu_2(k_p)$*.*

Suppose U is a solution to the optimality equation with linear growth. From the optimality equation, the separability of g^* and μ^*, and the Lipschitz continuity of σ^2 :

$$\begin{aligned}
U &= \frac{\sigma^2(X,U')}{2r}U'' + g_1(X) + g_2(U') + \frac{U'}{r}\left(\mu_1(X) + \mu_2(U')\right) \\
\Rightarrow \quad \lim_{X\to p} U - \frac{U'}{r}\mu_1(X) - \frac{\sigma^2(X,U')}{2r}U'' - g_1(X) &= \lim_{X\to p} g_2(U') + \frac{U'}{r}\mu_2(U') \\
\Rightarrow \quad \lim_{X\to p} U - \frac{U'}{r}\mu_1(X) - g_1(X) &= \phi_p
\end{aligned}$$

Consider the linear, non-homogenous ODE

$$y - \frac{\mu_1(X)}{r}y' = g_1(X) + \phi_p \tag{2}$$

The general solution to Equation 2 is

$$y = \left(-r\int\left(\frac{1}{f(x)}\right)\frac{g_1(x)}{\mu_1(x)}dx + c\right)f(x) + \phi_p$$

where

$$f(x) = \exp\left(\int \frac{r}{\mu_1(x)}dx\right)$$

and c is a constant. By Assumption 3, μ_1 is $O(X)$ with a growth rate slower than r. Let m be the asymptotic growth rate of μ_1.

Case (i): If $m \in [0, r)$, then $f(x)$ is not $O(X)$, and therefore, any solution y^L that is $O(X)$ must have $c = 0$.

$$y^L = -rf(x)\int\left(\frac{1}{f(x)}\right)\frac{g_1(x)}{\mu_1(x)}dx + \phi_p$$

Conclude that there is a unique solution y^L that has linear growth.

Case (ii) If $m < 0$, then $\lim_{X\to\infty} f(x) = 0$, and c does not affect the limit properties of y. Therefore, $\lim_{X\to\infty} y_L - y = 0$ for all c, and it is without loss of generality to consider the limit properties of y_L.

Let U be a solution to the optimality equation with linear growth. Then, given that there is a unique solution y^L with linear growth to Equation 2, it must be that:

$$\lim_{X\to p} U - y^L = 0$$

Note this implies that any two solutions with linear growth U and V have the same boundary conditions, $\lim_{X\to p} U - V = 0$. Q.E.D.

Step 1c: Boundary Conditions for Bounded Support. Suppose $\Xi = \left[\underline{X}, \overline{X}\right]$ a closed and bounded interval, and there are no absorbing points.

Lemma 21. *Any bounded solution U of the optimality equation has bounded variation.*

Suppose U has unbounded variation. Then there exists a sequence $(X_n)_{n\in N}$ 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 static Nash equilibria is played when $U'(X) = 0$, yielding flow payoff $g^*(X, 0)$. From the optimality equation, this implies $g^*(X_n, 0) \geq U(X_n)$. Likewise, there exists a sequence $(X_m)_{m\in N}$ that correspond to local minima of U, so $U'(X_m) = 0$ and $U''(X_m) \geq 0$. This implies $g^*(X_m, 0) \leq U(X_m)$. Thus, g^* also has unbounded variation. This is a contradiction, since $g^*(X, 0)$ is Lipschitz continuous. Q.E.D.

Lemma 22. *Suppose a function $f(X)$ is Lipschitz continuous on $\left[\underline{X}, \overline{X}\right]$ with $f(\overline{X}) = f(\underline{X}) = 0$. Then any bounded solution U of the optimality equation satisfies*

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

for $p \in \left\{\underline{X}, \overline{X}\right\}$.

Let $f(X)$ be a Lipschitz continuous function on $\left[\underline{X}, \overline{X}\right]$ with $f(p) = 0$ and suppose $\lim_{X\to p}\inf |f(X)| U'(X) > 0$. Given $f(X)$ is $O(p - X)$, there exists an $M \in \mathbb{R}$ and a $\delta_1 > 0$ such that when $|p - X| < \delta_1, |f(X)| \leq M|p - X|$. Given $\lim_{X\to p}\inf |f(X)| U'(X) > 0$, there exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $|p - X| < \delta_2$, $|f(X)| U'(X) > \varepsilon$. Take $\delta = \min\{\delta_1, \delta_2\}$. Then for $|p - X| < \delta$, $U'(X) > \frac{\varepsilon}{|f(X)|} \geq \frac{\varepsilon}{M|p-X|}$. Then the antiderivative of $\frac{\varepsilon}{M|p-X|}$ is $\frac{\varepsilon}{M}\ln|p - X|$ which diverges to $-\infty$ as $X \to p$. This violates the boundedness of U. Therefore $\lim_{X\to p}\inf |f(X)| U'(X) \leq 0$.

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

Lemma 23. *Suppose $f(X)$ is Lipschitz continuous on $\left[\underline{X},\overline{X}\right]$ with $f(\overline{X}) = f(\underline{X}) = 0$. Then any bounded solution U of the optimality equation satisfies $\lim_{X\to p} f(X)U'(X) = 0$ for $p \in \left\{\underline{X},\overline{X}\right\}$.*

By Lemma 22, $\lim_{X\to p}\inf |p-X|\, U'(X) \leq 0 \leq \lim_{X\to p}\sup |p-X|\, U'(X)$. Suppose, without loss of generality, that $\lim_{X\to p}\sup |p-X|\, U'(X) > 0$. Then there exist constants k and K such that $|p-X|\, U'$ crosses the interval (k,K) infinitely many times as X approaches p. Additionally, there exists an $L > 0$ such that

$$\begin{aligned} |U''(X)| &= \left|\frac{2r\left[U(X) - g^*(X,U')\right] - 2\mu^*(X,U')U'(X)}{\left|\sigma^*(X,U')\right|^2}\right| \leq \left|\frac{L_1 - L_2\,|p-X|\, U'(X)}{(p-X)^2}\right| \\ &\leq \left|\frac{L_1 - L_2 k}{(p-X)^2}\right| = \frac{L}{(p-X)^2} \end{aligned}$$

This implies that

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

where the first line follows from differentiating $(p-X)\,U'(X)$ 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 $(p-X)U'(X) \in (k,K)$. Then

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

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

Lemma 24. *Let U be a bounded solution of the optimality equation. Then $\lim_{X\to p} U(X) = g^*(p,0)$ for $p \in \left\{\underline{X},\overline{X}\right\}$.*

Suppose not. By Lemma 21 and the continuity of U, $\lim_{X\to p} U(X) = U(p)$ exists. Suppose $U(p) < g^*(p,0)$. The function $\mu^*(X,U')$ is Lipschitz continuous. By Lemma

23 and the assumption that $\mu_x(a, b, p) = 0$ for all $a, b \in A, B$, $\lim_{X\to p} \mu^*(X, U')U'(X) = 0$. Additionally, by the Lipschitz continuity of g^*, $\lim_{X\to p} g^*(X, U') = g^*(p, 0)$. From the optimality equation,

$$\begin{aligned}\limsup_{X\to p} \left|\sigma^*(X, U')\right|^2 U'' &= \limsup_{X\to p} \left[2r\left(U - g^*(X, U')\right) - 2\mu^*(X, U')U'\right] \\ &= 2r\left(U(p) - g^*(p, 0)\right) < 0\end{aligned}$$

which violates Lemma 11 since σ^* is Lipschitz continuous and $\sigma(b, p) = 0$ for all $b \in B$. Thus, $U(p) = g^*(p, 0)$. Q.E.D.

Lemma 25. *Any bounded solution U of the optimality equation satisfies*

$$\lim_{X\to p} \left|\left|\sigma^*(X, U')\right|^2 U''\right| = 0$$

Applying Lemmas 23 and 24 and the squeeze theorem,

$$\lim_{X\to p} \left|\left|\sigma^*(X, U')\right|^2 U''\right| = \lim_{X\to p} \left|2r\left(U - g^*(X, U')\right) - 2\mu^*(X, U')U'\right| = 0.$$

The above analysis directly carries over to a state space that is bounded from one side, or contains a finite number of interior absorbing points.

Step 2: Uniqueness of Solution to Optimality Equation. The first lemma follows directly from Lemma C.7 in Faingold and Sannikov (2011).

Lemma 26. *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 $U(X) < V(X)$ and $U'(X) < V'(X)$ 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 $X_1 = \inf\left\{X \in [X_0, \overline{X}) : U'(X) \geq V'(X)\right\}$.

Note it's possible for two solutions to be parallel / cross once, and then converge to the same limit slope despite the slope of one solution always remaining strictly higher than the other.

Lemma 27. *There exists a unique solution U to the optimality equation.*

Suppose U and V are both solutions to the optimality equation. Suppose $V(X) > U(X)$ for some $X \in \Xi$. Let X^* be the point where $V(X) - U(X)$ is maximized, which is well-defined given U and V are continuous functions and $\lim_{X\to p} U(X)-V(X)=0$ for $p\in\{\underline{X},\overline{X}\}$. Then $U'(X^*)=V'(X^*)$ and $V(X^*)>U(X^*)$. By Lemma 26, $V'(X)>U'(X)$ for all $X>X^*$, and $V(X)-U(X)$ is strictly increasing, a contradiction since X^* maximizes $U(X)-V(X)$. The proof for $\underline{X}$ is analogous. Q.E.D.

Step 3: Uniqueness of PPE. When there is a unique solution to the optimality equation, it is obvious to see that there is a unique Markov equilibrium, by Theorem 1. It remains to show that there are no other PPE. When there is a unique Markov equilibrium, Theorem 2 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 are equal and actions are uniquely specified by $S^*(X,0,U'(X)/r)$. Q.E.D.

7.7 Intermediate Results

Lemma 28. *Suppose a function f is bounded and there exists a δ such that for $|X|>\delta$, f is monotonic. Then for $p\in\{-\infty,\infty\}$, $\lim_{X\to p} f(X)$ exists. If f is also differentiable, then $\lim_{X\to p} Xf'(X)=0$.*

The first result is a standard result in analysis. For the second, suppose that $\lim_{X\to p} Xf'(X)\neq 0$. Suppose that $\lim_{X\to\infty}\inf|X|\,f'(X)>0$. Given $\lim_{X\to\infty}\inf|X|\,f'(X)>0$, there exists a $\delta\in\mathbb{R}$ and an $\varepsilon>0$ such that when $|X|>\delta$, $|X|\,f'(X)>\varepsilon$. Then for $|X|>\delta$, $f'(X)>\frac{\varepsilon}{|X|}$. Then the antiderivative of $\frac{\varepsilon}{|X|}$ is $\varepsilon\ln|X|$ which converges to ∞ as $|X|\to p$. This violates the boundedness of f. Therefore $\lim_{X\to p}\inf|X|\,f'(X)\leq 0$. Similarly, $\lim_{X\to p}\sup|X|\,f'(X)\geq 0$. Suppose that $\lim_{X\to p}\sup|X|\,f'(X)>0$. There exists a $\delta>0$ such that $f(X)$ is monotonic for $|X|>\delta$. Then for $|X|>\delta$, $|X|\,f'(X)$ doesn't change sign. Therefore, if $\lim_{X\to p}\sup|X|\,f'(X)>0$, then $\lim_{X\to p}\inf|X|\,f'(X)>0$. This is a contradiction. Thus, $\lim_{X\to p}\sup|X|\,f'(X)=0$. By similar reasoning, $\lim_{X\to p}\inf|X|\,f'(X)=0$, and therefore $\lim_{X\to p}|X|\,f'(X)=0$. Note this result also implies that $\lim_{X\to p} f'(X)=0$. Q.E.D.

Lemma 29. *Suppose a function f is differentiable. Define $\overline{f}(X)=f(X)/X$. For $p\in\{-\infty,\infty\}$, if $\lim_{X\to p}\overline{f}(X)=M$ and $\lim_{X\to p} X\overline{f}'(X)=0$, then $\lim_{X\to p} f'(X)=M$. In other words,*

$$\lim_{X\to p} f'(X)=\lim_{X\to p}\overline{f}(X)$$

Suppose $\overline{f}(X)=f(X)/X$, $\lim_{X\to p}\overline{f}(X)=M$ and $\lim_{X\to p} X\overline{f}'(X)=0$. Then

$$\begin{aligned} &\overline{f}'=(1/X)(f'-\overline{f})\\ \Rightarrow\quad &\lim_{X\to p} f'=\lim_{X\to p} X\overline{f}'+\overline{f}=M \end{aligned}$$

Q.E.D.

Lemma 30. *If f is a twice continuously differentiable function and there exists a δ such that for $|X| > \delta$, f' is monotonic, then there exists a δ_2 such that for $|X| > \delta_2$, f/X is monotonic.*

Let $\overline{f}(X) = f/X$. Note that

$$\begin{aligned} \overline{f}' &= (1/X)(f' - \overline{f}) \\ \overline{f}'' &= (1/X)(f'' - 2\overline{f}') \end{aligned}$$

and $\overline{f}' = 0$. when $f' = \overline{f}$, with a maximum if $f'' \leq 0$ and a minimum if $f'' \geq 0$. Let f' be monotonically increasing for $X > \delta$ i.e. $f'' \geq 0$ for all $X > \delta$. From $\overline{f}'' = (1/X)(f'' - 2\overline{f}')$, if $\overline{f}' < 0$ and $f'' \geq 0$, then $\overline{f}'' > 0$ and $\overline{f}'$ is increasing. Suppose there exists a $\delta_2 > \delta$ such that $\overline{f}'(\delta_2) \geq 0$. Then, by continuity of $\overline{f}'$ and the fact that $\overline{f}' < 0$ and $f'' \geq 0 \Rightarrow \overline{f}'' > 0$, it is not possible to have $\overline{f}' < 0$ for $X > \delta_2$. Therefore, $\overline{f}\,' \geq 0$ for all $X > \delta_2$ and $\overline{f}$ is monotonically increasing for all $X > \delta_2$. Otherwise, $\overline{f}' < 0$ for all $X > \delta$, and therefore $\overline{f}$ is monotonically decreasing for all $X > \delta$. The proof is similar when f' is monotonically decreasing. Q.E.D.

Lemma 31. *Convergence rate of Lipschitz functions.*

1. *Suppose a function $f(x)$ is Lipschitz continuous on $\mathbb{R}$. Then there exists a $K \in \mathbb{R}$ and a $\delta \in \mathbb{R}$ such that when $|x| > \delta, |f(x)| \leq K|x|$.*
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 $x \to x^*$.*

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

$$|f(x)| \leq |f(x) - f(0)| + |f(0)| \leq K_1 |x| + K_2$$

Therefore, there exists a K and an x^* such that for $|x| > x^*$, $|f(x)| \leq K|x|$.

Suppose $f(x)$ is Lipschitz continuous on $[\underline{x}, \overline{x}]$ with Lipschitz constant $K \in \mathbb{R}$. Then $|f(x) - f(\underline{x})| \leq K|x - \underline{x}|$. Note that $f(\underline{x})$ is a constant. Therefore, $f(x) - f(\underline{x})$ is $O(x - \underline{x})$ as $x \to \underline{x}$. As a special case, this implies that if $f(\underline{x}) = 0$, then $f(x)$ is $O(x - \underline{x})$. Q.E.D.