PENN INSTITUTE for ECONOMIC RESEARCH
UNIVERSITY of PENNSYLVANIA

The Ronald O. Perelman Center for
Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
18-015

Using Persistence to Generate Incentives in a Dynamic Moral Hazard Problem

J. AISLINN BOHREN
University of Pennsylvania

April 1, 2018

Using Persistence to Generate Incentives in a Dynamic Moral Hazard Problem*

J. Aislinn Bohren†
University of Pennsylvania

April 2018

Abstract

I study how the persistence of past choices can be used to create incentives in a continuous time stochastic game in which a large player, such as a firm, interacts with a sequence of short-run players, such as customers. The long-run player faces moral hazard and her past actions are imperfectly observed – they are distorted by a Brownian motion. Persistence refers to the fact that actions impact a payoff-relevant state variable, e.g. the quality of a product depends on both current and past investment choices. I obtain a characterization of actions and payoffs in Markov Perfect Equilibria (MPE), for a fixed discount rate. I show that the perfect public equilibrium (PPE) payoff set is the convex hull of the MPE payoff set. Finally, I derive sufficient conditions for a MPE to be the *unique* PPE. Persistence creates effective intertemporal incentives to overcome moral hazard in settings where traditional channels fail. Several applications illustrate how the structure of persistence impacts the strength of these incentives.

KEYWORDS: Continuous Time Games, Stochastic Games, Moral Hazard
JEL: C73, L1

*I thank Simon Board, 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 Wolitzky and especially Nageeb Ali for useful comments. I also thank numerous seminar participants for helpful feedback.
†Email: abohren@sas.upenn.edu; Web: www.aislinnbohren.com

1 Introduction

This paper studies how the *persistence* of past choices can be used to create incentives in a continuous time stochastic game in which a long-run player interacts with a sequence of short-run players. Persistence refers to the fact that actions noisily impact a payoff-relevant state variable, such as a worker's rating, a firm's product quality or a government's policy variable. This persistence can capture exogenous features of the environment, such as how past investment influences current quality or how past policy choices map into current policy variables. It can also capture endogenous design choices, such as how a rating system aggregates past reviews and rewards a worker based on her rating. The long-run player faces moral hazard and her past actions are not perfectly observed by consumers – they are distorted by a Brownian motion. Incentives can depend on the noisy signal of action choices, as well as on how persistence influences future payoffs through the impact that actions have on the state. The goal of this paper is to determine whether and how persistence strengthens incentives to overcome moral hazard.

The framework captures many economic settings in which past choices shape key features of current and future interactions. For example, a worker's rating on a platform depends on the quality of service she has provided to previous customers. She may be rewarded for earning a good rating and punished for poor performance. This provides an incentive for her to earn and maintain a good rating. Similarly, a firm's ability to make a high quality product is a function not only of its effort today, but also its past investments in developing technology and training its workforce. Quality today is linked to a firm's future quality, in that customers experience similar quality across time due to the persistence of investment. When customers are willing to pay a higher price or buy a larger quantity of a high quality good, this persistence provides an incentive for the firm to invest in developing a high quality product. Finally, a government's success in achieving a policy objective depends on both past and current policy choices. When past policy choices impact the future value of a policy variable, the government may be willing to undertake more costly actions today, since the benefit of such actions continue to accrue in future periods.

I study perfect public equilibria (PPE) in this framework – that is, equilibria in which strategies depend only on public information. I establish that the PPE payoff set is equal to the convex hull of the Markov perfect equilibrium (MPE) payoff set. In a MPE, equilibrium actions and payoffs depend only on the payoff-relevant components of the game – in this case, the observable state. Any paths of signals and states that lead to the same current state prescribe the same continuation play.[1] The intuition for this result stems from the type of incentives that are feasible in this setting. In a stochastic game, incentives can either be *informational* – signals are used to coordinate future equilibrium play, or *structural* – actions impact the structure of future interactions through their impact on the state. This latter channel includes both the state's *direct* impact on future feasible payoffs and its indirect impact through its effect on future *equilibrium* play. In the presence of short-run players, the only feasible type of informational incentives is value-burning, in which players switch to an inefficient action profile after certain signals. But from Sannikov and Skrzypacz (2010), we know that Brownian information is too noisy to create effective incentives via value-burning – punishment is triggered too frequently.[2] Therefore, it is only possible to use structural incentives. This is precisely the channel for incentives in a MPE, as informational channels are precluded by definition.[3]

In establishing this result, I obtain a characterization of the equilibrium payoffs and actions in MPE, for a fixed discount rate. This characterization yields sharp insights – in contrast to a folk theorem, it determines what type of equilibria one expects to emerge and what pattern of behavior will generate a given payoff. It shows the dynamics of behavior depend on observable outcomes, such as the rating of a restaurant or the current level of a policy variable, and how incentives and payoffs depend on the structure of persistence, such as the depreciation rate of investment. The degree to which persistence allows the long-run player to overcome moral hazard depends on the marginal impact of the long-run player's action on the state and the slope of the continuation payoff with respect to the state. If undertaking a costly action moves the state in a direction that yields a higher continuation payoff, this will provide the long-run player with effective incentives. Importantly, this continuation payoff is an equilibrium object that captures how the state impacts both future feasible payoffs and future equilibrium play. For example, when consumers have observed a given level of quality in the recent past, their willingness

[1]In a PPE, this is not the case – equilibria can depend on the path of past signals and states in an arbitrary way.

[2]Value-burning is also the only feasible type of informational incentives in games that have a failure of identifiability. Brownian information is too noisy to create effective informational incentives in these games as well (Abreu, Milgrom, and Pearce 1991; Skrzypacz and Sannikov 2007). A stochastic game with Brownian information, multiple long-run players and a failure of identifiability will have a similar equilibrium characterization to this paper.

[3]In earlier related work, Faingold and Sannikov (2011) establish a similar result when short-run players have incomplete information about the long-run player's type and the state is the belief that the long-run player is committed to choosing a certain action.

to pay for a product today will depend on both the persistence of this past quality in determining today's quality, as well as how they believe a given level of quality influences the firm's current level of investment. The interaction of these *direct* and *equilibrium* channels can significantly strengthen or dampen incentives, depending on the structure of the game.

The second main result determines when a MPE emerges as the *unique* PPE. This result relies on determining when there is a unique MPE in the class of Markov equilibria. When there is a unique MPE, the result described above establishes that this will also be the unique PPE. Uniqueness depends on incentives as the state approaches the boundary of the state space. If boundary incentives are unique – e.g. it is possible to sustain a unique equilibrium action profile and payoff at the boundary – then from the MPE characterization, incentives must also be unique on the interior of the state space. I present sufficient conditions for three cases: (i) an unbounded state space and flow payoff, (ii) a bounded flow payoff and unbounded state space, and (iii) a bounded state space and flow payoff. In case (i), these conditions rule out complementarities between the direct and equilibrium channels for incentives near the boundary. This rules out multiple optimal action profiles due to coordination motives. In case (ii) and (iii), these conditions ensure that incentives collapse as the state approaches the boundary. This rules out the possibility of sustaining multiple equilibrium action profiles at the boundary.

Several applications illustrate how persistence can be used to create effective incentives to overcome moral hazard. In Section 2, I introduce a version of the canonical product choice setting in which a firm's effort has a persistent effect on the quality of its product. I show that this persistence provides effective incentives for the firm to invest in building a high quality product. These incentives are present in the long-run, in that the firm continues to choose a positive level of investment as the time period grows large. In Section 6, I consider a variation of the product choice game in which the marginal return to quality is non-monotonic and show that this can lead to firms specializing in low or high quality. I also consider a setting in which constituents elect a board to implement a policy target, and the policy level depends on both current and past decisions by the board. For example, the Federal Reserve targets interest rates or a board of directors sets growth and return targets for its company. The board's incentive to undertake a costly intervention is strongest when the current policy level is at an intermediate distance from the target. When it is far from the target, the benefit of intervention is too delayed, while when it is close to the target, the benefit of further intervention is small. Finally, I consider a setting in which a government and innovators invest in intellectual capital, and there is a strategic complementarity between their investments. The complementarity leads to multiple Markov equilibria, including one in which neither party invests and several that sustain a positive level of investment. These equilibrium characterizations can be used to address important design questions, such as how to design the optimal reward structure or determining the optimal durability for a production technology.

1.1 Literature

Recent results on repeated games between a long-run and short-run player (Faingold and Sannikov 2011; Fudenberg and Levine 2007) show that the intersection of noise in monitoring and instantaneous adjustment of actions create a genuine challenge in providing intertemporal incentives.[4] In the analogue of this paper with no persistence, the long-run player cannot earn an equilibrium payoff above the best static Nash payoff.[5] In contrast, the equilibrium characterization in this paper demonstrates that persistence creates effective intertemporal incentives and enables the long-run player to overcome moral hazard. This leads to a higher payoff both from the incentive to invest in building the state and from dynamic strategic interaction with the short-run players.

The literature on reputation with behavioral types is another important and well-understood mechanism to overcome moral hazard in similar settings (Faingold and Sannikov 2011; Fudenberg and Levine 1989, 1992). If consumers believe that there is a chance that the firm is committed to choosing high effort, then the firm will be able to charge a higher price for its product. Incomplete information about the firm's type creates a form of persistence, as consumers' beliefs depend on past effort choices. However, fixing a strategic firm's patience, such reputation effects vanish in the ex-ante probability of behavioral types, and so the effectiveness of this persistence via incomplete information requires a non-trivial fraction of behavioral types.[6]

The connection with the reputational literature motivates several key insights. First,

[4] Abreu et al. (1991) first examined incentives in repeated games with imperfect monitoring and frequent actions. They established that shortening the period between actions has a crucial impact on the ability to structure effective incentives.

[5] Skrzypacz and Sannikov (2007) show that this is also the case in games between multiple long-run players in which deviations between individual players are indistinguishable.

[6] Kreps, Milgrom, Roberts, and Wilson (1982); Kreps and Wilson (1982) and Milgrom and Roberts (1982) first demonstrated that reputation, in the form of incomplete information about a player's type, has a dramatic effect on equilibrium behavior. Mailath and Samuelson (2001) show that reputational incentives can also come from a firm's desire to *separate* itself from an incompetent type.

when the firm is known to be strategic, this paper shows that other forms of persistence can also overcome moral hazard. Second, in contrast to the temporary incentives in reputation models (Cripps, Mailath, and Samuelson 2004; Faingold and Sannikov 2011), the incentives in a stochastic game persist in the long-run.[7] Finally, at a theoretical level, this paper explores the general properties of a stochastic game that has powerful intertemporal incentives. The reputational game can be viewed as a specific type of stochastic game. For instance, if instead of influencing the uncertainty about whether it is a behavioral type, a strategic firm makes a costly initial investment in a new production technology that benefits customers today and in the future, we would see similar intertemporal incentives in the resulting stochastic game.

This final point merits a closer comparison with Faingold and Sannikov (2011), who characterize the unique MPE in the stochastic game that corresponds to a continuous time reputation model. In their paper, payoffs and the evolution of the state take a specific form due to Bayesian updating. My characterization builds on the techniques in their paper to understand more generally what properties of stochastic games are needed for uniqueness of MPE and non-degenerate intertemporal incentives. I analyze a general class of stochastic games that places few restrictions on the process governing the evolution of the state and the structure of payoffs. The key technical advancement, relative to their paper, is for the case of an unbounded state space and payoff for the long-run player, as it requires significantly different techniques to complete the analysis.

Beyond reputation models with behavioral types, a rich literature analyzes dynamic games with a state variable, in which effort is directly linked to future payoffs via the state. Ericson and Pakes (1995) were the first to analyze hidden investment and stochastic capital accumulation (the state) in a model that is similar in spirit to the quality example presented in Section 2. They study firm and industry dynamics and establish equilibrium existence. Recent work by Doraszelski and Satterthwaite (2010) modify Ericson and Pakes (1995) to guarantee the existence of a pure strategy MPE, which is computationally tractable. Neither paper establishes uniqueness, but instead focus on the dynamics associated with a particular MPE. More broadly, MPE is the workhorse solution concept across industrial organization and political economy. A comprehensive review of this literature is beyond the scope of this paper.

This paper also relates to a literature on stochastic games with an unobservable state.

[7] Long-run reputation effects are also possible in models with behavioral types when consumers cannot observe all past signals (Ekmekci 2011) or the type of the firm is replaced over time.

In these games, incentives stem from the long-run player's ability to manipulate the public belief about the state through her effort choice. Cisternas (2016) characterizes necessary conditions for the existence of Markov equilibria in a continuous time stochastic game with an unobservable state, and sufficient conditions in two more restrictive classes of games. Hidden states significantly complicate the model, and it is not possible to establish uniqueness results or a full equilibrium characterization. Board and Meyer-ter vehn (2013) study a setting in which a firm's hidden quality depends on past effort and consumers learn about this quality from noisy signals. This paper differs in focus in that there is no adverse selection and I allow for strategic interaction between the long-run and short-run players and a richer class of stage game payoffs.

Several folk theorems exist for discrete time stochastic games with observable states, 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). My setting differs in that there is a single long-run player and information follows a diffusion process. It is already known that these two changes significantly alter incentives in standard repeated games (compare the folk theorem in Fudenberg, Levine, and Maskin (1994) with the equilibrium degeneracy in Faingold and Sannikov (2011); Fudenberg and Levine (2007)). The intuition is similar for the stochastic game folk theorems and the MPE uniqueness result in this paper.[8]

The organization of the paper proceeds as follows. Section 2 presents a product choice example to motivate the model. Section 3 sets up the model and characterizes the structure of PPE. Section 4 presents the three main results: existence of a Markov equilibrium, characterization of the PPE payoff set and uniqueness of a Markov equilibrium in the class of all PPE. Section 5 presents structural results on the shape of equilibrium payoffs. Several additional applications are presented in Section 6. All proofs are in the Appendix.

2 Example: Product Choice with Persistent Quality

Consider a variation of the canonical product choice setting, in which a monopolist firm provides a product to a continuum of short-run consumers and the firm's effort has a persistent effect on the quality of the product. At each instant t, the firm chooses an

[8] The paper also relates to an older literature on stochastic games and existence of Markov equilibria, beginning with Shapley (1953) and Sobel (1973), who examine equilibrium existence in continuous-time stochastic games with perfect monitoring.

unobservable effort level $a_t \in [0, \overline{a}]$, where $\overline{a} > 0$. The quality of the firm's product depends on both current and past effort, $q(a, X) = (1-\lambda)a + \lambda X$, where past effort influences quality through an observable state variable

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

which we refer to as the stock quality.[9] The parameter $\theta > 0$ determines the decay rate of past effort, $\lambda \in [0, 1]$ captures the relative importance of past effort in determining current quality, $\sigma > 0$ is the volatility and $(Z_t)_{t \geq 0}$ is a standard Brownian motion. Effort increases quality both today and in the future. Consumers purchase a single unit of the product. When a consumer believes that the firm will choose effort level $\tilde{a}$, his expected value for the product is $q(\tilde{a}, X)$. Each consumer is willing to pay his expected value for the product if quality is positive, $\overline{b} = q(\tilde{a}, X)$ if $q(\tilde{a}, X) > 0$, and otherwise is willing to pay zero, $\overline{b} = 0$. The average discounted profit of the firm is the difference between revenue and the cost of effort,

$$r \int_0^\infty e^{-rt}(\overline{b}_t - a_t^2/2)dt,$$

where $r > 0$ is the discount rate.

In the unique perfect public equilibrium (PPE) with no persistence, $\lambda = 0$, the firm exerts zero effort and consumers' willingness to pay is zero (Theorem 3 in Faingold and Sannikov (2011)). Intertemporal incentives break down, despite the fact that the firm would earn higher profits if it could commit to higher effort.[10]

In this paper, I show that persistent quality incentivizes the firm to choose a positive level of effort and earn positive profits. Theorems 1-3 establish that there is a *unique* perfect public equilibrium (PPE), which is Markov in the stock quality X_t. The effort level and profit in this unique equilibrium are characterized as a function of the impact of past effort on currently quality λ, the depreciation rate θ of the stock quality and the discount rate r. For any $\lambda > 0$, the firm chooses a positive level of effort and earns positive profits when the quality stock is positive, and at some (possibly all) negative levels of quality stock. For large X, equilibrium effort is approximately $\lambda/(r + \theta)$. This

[9] In a slight abuse of notation, the Lebesgue integral and the stochastic integral are placed under the same integral sign.

[10] In contrast to Abreu et al. (1991) and Skrzypacz and Sannikov (2007), this breakdown of incentives takes place despite there being no failure of identifiability.

effort is larger when past effort plays a larger role in determining current quality, effort depreciates at a slower rate, or the firm is more patient. Further, in this equilibrium, the firm has a long-run incentive to choose high effort. This contrasts with models in which the incentive to produce high quality is derived from consumers' uncertainty over the firm's payoffs and long-run effort converges to zero (Cripps et al. 2004; Faingold and Sannikov 2011).

Persistence increases the firm's payoffs through two complimentary channels. First, the firm's effort increases the stock quality, which increases future revenue through the impact of the stock quality on prices. This is the *direct* effect of persistence on future feasible payoffs, as discussed in the introduction. Second, the link with future payoffs allows the firm to credibly choose a positive level of effort in the current period. If the consumer expects higher effort today, she is willing to pay a higher price today. This arises from the strategic interaction between the firm and consumers – it is the *equilibrium* effect discussed in the introduction. At large values of the quality stock, the continuation value is close to linear and it is possible to approximately quantify the profit arising from each of these channels. The present value of the direct effect on feasible payoffs, minus the cost of effort, is approximately $\lambda^2/2(r+\theta)^2$. This return is higher when past effort plays a larger role in determining current quality (higher λ), quality depreciates at a lower rate (lower θ) or the firm is more patient (lower r). The present value of the equilibrium effect is approximately $(1-\lambda)\lambda/(r+\theta)$. It is largest for intermediate values of λ. The incentive to choose higher effort today is increasing in λ, while the impact of higher effort today on the current price is increasing in $1-\lambda$. Therefore, the equilibrium effect is largest when $\lambda \approx 1/2$.

The equilibrium characterization yields insight into the shape of the equilibrium payoff. It is convex in the quality stock at low levels of stock quality and becomes linear as the quality stock grows large. The shape determines how volatility in quality affects payoffs. When the quality stock is low and consumers are not purchasing, the firm receives substantial gains if quality rises, and small losses if quality falls – revenue is already low, so negative quality shocks can only reduce revenue by a small amount, whereas positive quality shocks can significantly raise revenue. Therefore, higher volatility leads to higher profits. For example, if quality is a measure of innovation, then volatility is beneficial at low levels of quality before a breakthrough occurs. Once a breakthrough occurs and quality becomes is high, the return to quality is approximately linear. The distribution of positive and negative shocks is symmetric, so when the return to quality is linear, volatility has no impact on expected profits.

This example will be used throughout the paper to demonstrate the results. The product choice framework lends itself to other variations, several of which are discussed in Section 6.1.

3 Model

3.1 Model Set-up

States and Actions. A long-run player and a continuum $I = [0,1]$ of identical short-run players, indexed by i, play a continuous time stochastic game with imperfect monitoring. At each instant of time $t \in [0,\infty)$, a publicly observable state variable X_t in nonempty closed interval $\mathcal{X} \subset \mathbb{R}$ determines the action set and feasible flow payoffs. If $\mathcal{X}$ is bounded, denote the upper and lower boundary states by $\overline{X}$ and $\underline{X}$, respectively. Long-run and short-run players simultaneously choose actions a_t from A and b_t^i from $B(X_t)$, respectively, where A is a nonempty compact subset of a Euclidean space and $B(X)$ is a nonempty compact subset of a closed Euclidean space B. Denote the set of feasible short-run player actions and states as $E = \{(b,X) \in B \times \mathcal{X} | b \in B(X)\}$. Assume that the boundary of the feasible set of actions for short-run players grows at most linearly with the state – that is, there exists a $K_b, c_b > 0$ such that for all $(b,X) \in E$, $|b| \leq K_b|X| + c_b$.[11] Individual actions are privately observed. Players observe the aggregate distribution of short-run players' actions, $\bar{b}_t \in \Delta B(X_t)$, and do not observe the long-run player's action.

Given initial state X_0, the state evolves stochastically according to

$$dX_t = \mu(a_t, \bar{b}_t, X_t)dt + \sigma\left(\bar{b}_t, X_t\right)dZ_t \tag{1}$$

where $(Z_t)_{t\geq 0}$ is a one-dimensional Brownian motion, $\mu : A \times E \to \mathbb{R}$ is the drift and $\sigma : E \to \mathbb{R}$ is the volatility. Assume μ and σ are Lipschitz continuous. The drift depends on the long-run player's action, the aggregate action of the short-run players and the state. Volatility is independent of the long-run player's action to maintain the assumption that the long-run player's action is not perfectly observed. Each function can be linearly extended to $A \times \Delta E$ or ΔE, respectively, where (in a slight abuse of notation), $\Delta E = \{(\bar{b}, X) \in \Delta B \times \mathcal{X} | \operatorname{supp} \bar{b} \subset B(X)\}$. Assume that the volatility of the state variable is positive at all interior states.

[11] I use $|\cdot|$ to denote the Euclidean norm for vectors.

Assumption 1 (Positive Volatility)**.** *For any compact proper subset $I \subset \mathcal{X}$,*

$$\inf_{\{(b,X)\in B\times I \ s.t. \ b\in B(X)\}} \sigma(b,X) > 0.$$

This ensures that the future path of the state variable is stochastic, except possibly at the boundary states.[12] Define a state X as an *absorbing state* if the drift and volatility are both zero, $\mu(a,b,X) = 0$ and $\sigma(b,X) = 0$ for all $(a,b) \in A \times B(X)$. Assumption 1 rules out interior absorbing states.[13]

The path of the state dX_t provides a public signal of the long-run player's action a_t. We assume that there are no additional public signals. This is without loss of generality, as additional signals have no effect on the equilibrium characterization (see discussion in Section 3.3). Let $(F_t)_{t\geq 0}$ represent the filtration generated by public information $(X_t)_{t\geq 0}$. Short-run players receive no information about the long-run player's action beyond what is contained in $(F_t)_{t\geq 0}$.

Payoffs. The payoff of the long-run player depends on her action, the distribution of short-run players' actions and the state. She seeks to maximize the expected value of her discounted payoff,

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

where $r > 0$ is the discount rate and $g : A \times E \to \mathbb{R}$ is a Lipschitz continuous function representing the flow payoff, which is linearly extended to $A \times \Delta E$. Short-run players have identical preferences. The payoff of player i in period t depends on her action, the distribution of short-run players' actions, the action of the long-run player and the state, $h(a_t, \bar{b}_t, b_t^i, X_t)$, where $h : A \times B \times E \to \mathbb{R}$ is a continuous function, which is linearly extended to $A \times \Delta B \times E$. The dependence of payoffs on the state variable creates a form of action persistence, since the state variable depends on prior actions.

To ensure that the expected discounted payoff of the long-run player is well-behaved requires a restriction on either the rate at which the state variable can grow or the flow payoff of the long-run player. Either the drift of the state grows at a linear rate less than the discount rate or the flow payoff is bounded with respect to X. This assumption is

[12] If the state space is bounded, then it must be that for all $(a,b) \in A \times B(\overline{X})$, $\sigma(b,\overline{X}) = 0$ and $\mu(a,b,\overline{X}) \leq 0$, and for all $(a,b) \in A \times B(\underline{X})$, $\sigma(b,\underline{X}) = 0$ and $\mu(a,b,\overline{X}) \geq 0$.

[13] If there is an interior absorbing state, the game can be analyzed as two separate games, since no states above the absorbing state can be reached from states below the absorbing state, and vice versa.

trivially satisfied when the state space is bounded.

Assumption 2 (Bounded Payoff or Drift)**.** *At least one of the following hold.*

1. *The flow payoff g is bounded.*

2. *The drift μ has linear growth at a rate less than r: there exists a $K_\mu \in [0, r)$ and $c_\mu > 0$ such that for all $(a, b, X) \in A \times E$, if $X \geq 0$ then $\mu(a, b, X) \leq K_\mu X + c_\mu$ and if $X \leq 0$ then $\mu(a, b, X) \geq K_\mu X - c_\mu$.*

No lower bound is necessary on the slope of μ for $X > 0$, since a negatively sloped drift pulls the state variable towards zero. Similarly, no upper bound is necessary for $X < 0$.

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 $b_t^i \in B(X_t)$ 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 $(\bar{b}_t)_{t\geq 0}$. The long-run player's expected discounted payoff at time t under strategy $S = (a_t, \bar{b}_t)_{t\geq 0}$ is given by

$$V_t(S) \equiv E_t \left[r \int_0^\infty e^{-rs} g(a_s, \bar{b}_s, X_s) ds \right] \tag{2}$$

I restrict attention to pure strategy perfect public equilibria (PPE), as defined in Sannikov (2007).

Definition 1 (PPE)**.** *A public strategy profile $S = (a_t, \bar{b}_t)_{t\geq 0}$ is a* perfect public equilibrium *if, after all public histories,*

$$V_t(S) \geq V_t(S') \ a.s.$$

for all public strategies $S' = (a'_t, \bar{b}'_t)_{t\geq 0}$ with $(\bar{b}'_t)_{t\geq 0} = (\bar{b}_t)_{t\geq 0}$ almost everywhere, and

$$b \in \arg\max_{b' \in B(X_t)} h(a_t, b', \bar{b}_t, X_t) \qquad \forall b \in \operatorname{supp} \bar{b}_t.$$

Timing. At each instant t, players observe the current state X_t and choose actions. Then nature stochastically determines payoffs and the next state, given the current state and the chosen action profile.

3.2 PPE Structure

This section extends a recursive characterization of PPE to a stochastic game. Given strategy profile $S = (a_t, \bar{b}_t)_{t\geq 0}$, define the long-run player's continuation value as the expected value of the future discounted payoff at time t,

$$W_t(S) \equiv E_t \left[r \int_t^{\infty} e^{-r(s-t)} g(a_s, \bar{b}_s, X_s) ds \right]. \tag{3}$$

The expected average discounted payoff at time t can be represented as

$$V_t(S) = r \int_0^t e^{-rs} g(a_s, \bar{b}_s, X_s) ds + e^{-rt} W_t(S). \tag{4}$$

Lemma 1 characterizes the evolution of the continuation value and the long-run player's incentive constraint in a PPE. It is the analogue of Theorem 2 in Faingold and Sannikov (2011), allowing for an unbounded state space and flow payoff. A challenge in extending the PPE characterization to an unbounded flow payoff is establishing that $E|V_t(S)| < \infty$ for all $t \geq 0$ and characterizing the growth rate of $(W_t(S))_{t\geq 0}$ – steps that are not required for models with a bounded flow payoff. This requires Assumption 2.2, which ensures that the state grows at a slow enough rate relative to the discount rate.[14]

Lemma 1 (PPE Characterization)**.** *Assume Assumptions 1 and 2. A public strategy profile $S = (a_t, \bar{b}_t)_{t\geq 0}$ 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. *Continuation value: $(W_t)_{t\geq 0}$ satisfies*

$$dW_t \;=\; r\left(W_t - g(a_t, \bar{b}_t, X_t)\right) dt + r\beta_t (dX_t - \mu(a_t, \bar{b}_t, X_t) dt) \tag{5}$$

   and there exists a $K, M \geq 0$ such that $|W_t| \leq M + K|X_t|$ for all $t \geq 0$, with $K = 0$ if g is bounded.

2. *Sequential rationality: $(a_t, \bar{b}_t)_{t\geq 0} \in S^*(X_t, r\beta_t)$ for almost all $t \geq 0$, where*

$$S^*(X, z) \equiv \left\{ (a, \bar{b}) : \begin{array}{l} a \in \arg\max_{a' \in A} g(a', \bar{b}, X) + \frac{z}{r}\mu(a', \bar{b}, X) \\ b \in \arg\max_{b' \in B(X)} h(a, b', \bar{b}, X) \ \forall b \in \operatorname{supp} \bar{b} \end{array} \right\} \tag{6}$$

[14] This is similar in spirit to Lemma 1 in Strulovici and Szydlowski (2015), which establishes that the value function of an optimal control problem is finite and satisfies a linear growth condition with respect to the state.

for $(X, z) \in \mathcal{X} \times \mathbb{R}$.

From Lemma 1, the continuation value of the long-run player is a stochastic process that is measurable with respect to public information, $(F_t)_{t\geq 0}$. The drift of the continuation value, $W - g(a, b, X)$, captures the expected change in the continuation value. It is the difference between the current continuation value and the flow payoff. The volatility of the continuation value, β, determines the sensitivity of the continuation value to information. Future payoffs are more sensitive when the volatility of the continuation value is larger. Sequential rationality for the long-run player depends on the trade-off between an action's impact on flow payoffs today and its expected impact on future payoffs through the drift of the state, weighted by β. This condition is analogous to the one-shot deviation principle in discrete time. From the Martingale Representation Theorem, the continuation value and incentive constraint are linear with respect to $(\beta_t)_{t\geq 0}$, which lends significant tractability to the model.

Multiple PPE may arise for several reasons. First, at a state X and volatility β that are on the equilibrium path, there may be multiple sequentially rational action profiles $(a, \bar{b}) \in S^*(X, r\beta)$. In this case, it is clear that there will be multiple PPE. For example, suppose $\beta_t = \beta$ and there are two action profiles in $S^*(X, r\beta)$, $(a_1, \bar{b}_1)$ and $(a_2, \bar{b}_2)$. Then each action profile corresponds to a PPE with $(a_t, \bar{b}_t) = (a_1, \bar{b}_1)$ or $(a_t, \bar{b}_t) = (a_2, \bar{b}_2)$ and the change in the continuation value dW_t determined by (5) evaluated at $(a_1, \bar{b}_1)$ or $(a_2, \bar{b}_2)$, respectively. Second, even if each state and incentive weight $(X, r\beta)$ prescribe a unique sequentially rational action profile, there may be multiple equilibrium paths of incentive weights $(\beta_t)_{t\geq 0}$ that satisfy Lemma 1, and hence, multiple PPE. This paper focuses on the latter class of games, in which there is a unique sequentially rational action profile at each $(X, r\beta)$, but there may be multiple equilibrium paths of incentive weights.

Assumption 3 (Unique Sequentially Rational Action Profile)**.** *For all* $(X, z) \in \mathcal{X} \times \mathbb{R}$*,* S^* *is non-empty, single-valued and returns* $\bar{b} = \delta_b$ *for some* $b \in B(X)$*, where* δ_b *is the Dirac measure on action* b*, and* S^* *is Lipschitz continuous on every bounded subset of* $\mathcal{X} \times \mathbb{R}$.

This assumption is straightforward to verify from the primitives of the model (g, h and μ).[15]

[15]When Assumption 3 fails and S^* is not single-valued, the correspondence may not be lower hemicontinuous. Different techniques are necessary to characterize Markov equilibrium payoffs. Similar to (Faingold and Sannikov 2011), differential inclusions can be used to characterize the "greatest" and "least"

Importantly, Assumption 3 does not preclude the existence of multiple PPE. As illustrated in Section 6.3, there may be multiple equilibria even when it is satisfied. Similarly, in discrete time, many games that satisfy an analogous assumption have multiple non-trivial equilibria.[16] A unique PPE will result from showing that, in addition to Assumption 3, there is a unique path of incentive weights $(\beta_t)_{t\geq 0}$ that satisfies the conditions in Lemma 1.

One implication of Assumption 3 is that the stage game at any state must have a unique static Nash equilibrium, as the static Nash equilibrium profile corresponds to $S^*(X,0)$. This rules out coordination games and some games with strategic complementarities. Assumption 3 still allows for a broad class of games, including games in which actions are strategic substitutes, strategic complements with a unique fixed point, or one-sided complementarities between actions.

Under Assumption 3, we can let $(a(X,z),b(X,z)) \equiv S^*(X,z)$ denote the unique sequentially rational action profile at state X and volatililty of the continuation value z/r. Let $g^*(X,z) \equiv g(a(X,z),b(X,z),X)$, $\mu^*(X,z) \equiv \mu(a(X,z),b(X,z),X)$ and $\sigma^*(X,z) \equiv \sigma(b(X,z),X)$ be the value of the flow payoff, drift and volatility of the state, respectively, at this action profile and state X. By Assumption 3, g^*, μ^* and σ^* are also Lipschitz continuous in (X,z).

Example. Return to the example introduced in Section 2 to demonstrate Assumptions 1-3. The boundary of the feasible action set for short-run players is linear in X $B(X) = [0, \overline{a}+X]$. The volatility of the state is constant, $\sigma(\overline{b},X) = 1$ (Assumption 1). The drift is $\mu(a,\overline{b},X) = a - \theta X$, which is negative when X is high and positive when X is low (Assumption 2.2). The long-run player's payoff is $g(a,\overline{b},X) = \overline{b} - a^2/2$. Given (X,z), sequential rationality for the firm requires

$$a \in \arg\max_{a\in[0,\overline{a}]} \overline{b} - a^2/2 + \frac{z}{r}(a-\theta X),$$

Markov equilibrium payoffs, as a function of the state, and show that the PPE payoff set is bounded by these payoffs.

[16]The analogous assumption in discrete time is more complex, as the incentive weights are functions rather than scalars. The continuation value can change according to any function $z : \mathcal{X} \to W$, where W is the set of feasible payoffs. The simple scalar representation is possible in continuous time because the continuation value changes linearly with respect to Brownian information, a property that does not hold in discrete time.

which yields $a(X, z) = z/r$ for $z \in [0, r\bar{a}]$, $a(X, z) = 0$ for $z < 0$, and $a(X, z) = \bar{a}$ for $z > r\bar{a}$. In equilibrium, consumers' beliefs about the effort choice of the firm are correct. Therefore, consumers are willing to pay $\bar{b}(X, z) = \max\{0, q(a(X, z), X)\}$. This sequentially rational action profile is unique and Lipschitz continuous in (X, z) (Assumption 3). Given $a(X, z)$ and $b(X, z)$, $g^*(X, z) = b(X, z) - a(X, z)^2/2$, $\mu^*(X, z) = a(X, z) - \theta X$ and $\sigma^*(X, z) = 1$.

3.3 Discussion of Model

Equilibrium Actions. In many applications, including rational expectations equilibrium models and learning models (where the state is a belief), the transition of the state depends on both the realized and *equilibrium* action of the long-run player. Similarly, the flow payoff of the long-run player may depend on both her realized and equilibrium action. The framework in this paper indirectly allows for this dependence, since the best response of the short-run player depends on the expected action of the long-run player – which, in equilibrium, is correct. One could also model this dependence directly by defining the drift and volatility to depend on $\tilde{a}$, i.e. $\mu'(a, \tilde{a}, b, X)$ and $\sigma'(\tilde{a}, b, X)$.[17] The analysis is unchanged, provided μ' and σ' satisfy Assumptions 1 - 3.[18]

For example, suppose the drift is $\mu(a, b, X) = \theta_1 b + \theta_2 a$ and the short-run player's payoff is $ab - b^2/2$. Let $\tilde{a}$ denote the equilibrium action. When the short-run player believes that the long-run player will choose $\tilde{a}$, her best response is $b = \tilde{a}$. This is isomorphic to a model in which the drift directly depends on the equilibrium action, $\mu'(a, \tilde{a}b, X) = \theta_1 \tilde{a} + \theta_2 a$.

The framework presented here does rule out some classes of stochastic games. In particular, consider Bayesian learning games with a binary outcome space, and let the state $X \in [0, 1]$ denote the belief that the outcome is high. Assumption 1 rules out games in which there exists an action profile that shuts-down learning at interior beliefs, i.e. there exists an $(\tilde{a}, b) \in A \times B$ such that $\sigma'(\tilde{a}, b, X) = 0$ at some $X \in (0, 1)$. This contrasts with Faingold and Sannikov (2011), in which it is feasible for the normal player to shut-down learning by perfectly mimicking the behavioral type. Therefore, their set-up does not satisfy Assumption 1 and their model requires an alternative approach to establish that the volatility of the state is bounded away from zero in equilibrium.

[17] In contrast to allowing the volatility to depend on the realized action, imperfect monitoring is maintained when the volatility depends on the equilibrium action. This is because the belief about the equilibrium action does not reveal the actual action chosen by the long-run player.

[18] This highlights the distinction from a single-agent decision problem, as the present framework is a fixed-point problem.

Additional Public Signals. In this framework, the path of the state both serves as a public signal of the long-run player's action and directly impacts payoffs. The results and analysis are unchanged if there are additional payoff-irrelevant Brownian public signals. Markov equilibria ignore such signals, so the characterization of Markov equilibria remains the same. Further, it is not possible to effectively use such signals to coordinate additional equilibria, due to reasoning similar to Faingold and Sannikov (2011). An older working paper version of the current paper allows for an arbitrary finite number of public signals (Bohren 2016).

Incentives in Continuous Time Stochastic Games. As discussed in the introduction, in a stochastic game, incentives can either be *informational* – past signals are used to coordinate future equilibrium play, or *structural* – past actions impact the structure of future interactions through their impact on the state. This latter channel includes both the state's *direct* impact on future feasible payoffs and its indirect impact through its effect on future *equilibrium* play. The process $(\beta_t)_{t\geq 0}$ characterized in Lemma 1 captures all of these channels for intertemporal incentives.

The linear structure of the continuation value with respect this process $(\beta_t)_{t\geq 0}$ (as characterized in (5)) plays a key role in determining incentives. In a *repeated* game with a short-run player, this linearity precludes effective intertemporal incentives (Faingold and Sannikov 2011). On the boundary of the equilibrium payoff set, it is not possible to tangentially transfer continuation payoffs between the long-run player and short-run player, since the short-run player is myopic. Further, non-tangential transfers must be constructed linearly with respect to the Brownian information. This results in the continuation value exceeding its boundary for any non-trivial incentive weight $\beta_t > 0$. Thus, $\beta_t = 0$ for all t. In contrast, in a *stochastic* game, β_t can depend on the state. Therefore, it may be possible to use non-zero linear transfers at some states, to create effective incentives without the continuation value escaping its boundary. The remainder of the paper explores whether and when this is possible.[19]

[19]Non-linear incentive structures, such as value-burning, are ineffective in both repeated and stochastic games, because the expected losses from false punishment exceed the expected gains from cooperating (Fudenberg and Levine 2007; Skrzypacz and Sannikov 2007).

4 Equilibrium Analysis

This section presents the main results of the paper. I establish the existence of Markov equilibria, characterize the correspondence of PPE payoffs of the long-run player, and derive conditions under which there is a *unique* PPE, which is Markov.

4.1 Existence of Markov Equilibria

First I characterize the set of Markov equilibria. This establishes existence and characterizes equilibrium behavior and payoffs in any Markov equilibrium. In a Markov equilibrium, the continuation value and equilibrium actions depend solely on the current state variable – they are independent of the past path of the state.[20] In Theorem 1, I show that the continuation values in all Markov equilibria are specified as the solution(s) $U : \mathcal{X} \to \mathbb{R}$ to an ordinary differential equation that maps each state to a payoff. Given a solution $U(X)$, the corresponding Markov equilibrium action profile at state X is the sequentially rational action profile at state X and volatility of the continuation value $U'(X)/r$.

Theorem 1. *Suppose Assumptions 1, 2 and 3 hold. Given initial state X_0, iff $U : \mathcal{X} \to \mathbb{R}$ has linear growth and is a (bounded if g is bounded) solution to the optimality equation*

$$rU(X) = rg^*(X, U'(X)) + U'(X)\mu^*(X, U'(X)) + \frac{1}{2}U''(X)\sigma^*(X, U'(X))^2, \qquad (7)$$

then $U(X)$ characterizes a Markov equilibrium with:

1. *Equilibrium payoff $U(X_0)$;*
2. *Continuation values $(W_t)_{t\geq 0} = (U(X_t))_{t\geq 0}$;*
3. *Equilibrium actions $(a_t, \bar{b}_t) = (a(X_t, U'(X_t)), b(X_t, U'(X_t)))$.*

The optimality equation has at least one continuous solution that has linear growth (is bounded) and lies in the range of feasible payoffs for the long-run player. Thus, there exists at least one Markov equilibrium.

[20] Recall that the path of the state provides a signal of the long-run player's action at each instant. Using these signals to punish or reward the long-run player could lead to PPE in which different paths of the state specify different continuation payoffs and equilibrium actions, even when these paths map to the same current state. In Markov equilibria, this is not allowed.

If there are multiple solutions to the optimality equation, then each solution will characterize a Markov equilibrium. In Section 6.3, I illustrate a setting with multiple Markov equilibria.

The optimality equation lends insight into the shape of payoffs in Markov equilibria. In any Markov equilibrium $U(X)$, the continuation value is equal to the sum of the equilibrium flow payoff rg^* and the equilibrium expected change in the continuation value $U'(X)\mu^*(X,U') + \frac{1}{2}U''(X)\sigma^*(X,U')^2$. This expected change has two components: (i) the interaction between the drift of the state and the slope of the continuation value and (ii) the interaction between the volatility of the state and the concavity of the continuation value. If the continuation value is increasing in the state, $U'(X) > 0$, then higher drift increases the expected change in the continuation value, while if the continuation value is decreasing in the state, the opposite holds. With Brownian information, positive and negative shocks to the state are equally likely. If the continuation value is concave, $U''(X) < 0$, it decreases more following a negative shock, relative to the increase following a positive shock, and volatility in the state decreases the expected change in the continuation value. The opposite holds when the continuation value is convex.

Theorem 1 also characterizes the equilibrium volatility of the continuation value, which determines incentives. The equilibrium path of the process $(\beta_t)_{t\geq 0}$ is proportional to the slope of the continuation value, $r\beta_t = U'(X_t)$. When the state is at the value(s) that yields the maximum equilibrium payoff across all states, $U'(X) = 0$. Therefore, the volatility of the continuation value is zero, which ensures that the continuation value does not escape the payoff set. In these periods, the long-run player acts myopically. At other states, the slope of the continuation value can be nonzero, and the continuation value will be sensitive to the changes in the state. Any state with $U'(X) \neq 0$ creates non-degenerate incentives for the long-run player.

Example, cont. Given $a(X,z)$ and $b(X,z)$ characterized in Section 3.2, any solution $U(X)$ to

$$rU(X) = rb(X,U'(X)) - ra(X,U'(X))^2/2 + U'(X)(a(X,U'(X)) - \theta X) + \frac{1}{2}U''(X) \quad (8)$$

with linear growth as $X \to \infty$ and bounded as $X \to -\infty$ characterizes a Markov equilibrium with equilibrium actions $a(X,U'(X))$ and $b(X,U'(X))$.

Outline of Proof. The first step 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 (7). 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,

$$dU(X_t) = U'(X_t)\mu(a_t^*, \overline{b}_t^*, X_t)dt + U''(X_t)\sigma(\overline{b}_t^*, X_t)^2/2dt + U'(X_t)\sigma(\overline{b}_t^*, X_t)dZ_t.$$

By Lemma 1, the continuation value must also follow the law of motion in Equation (5). Matching the drifts of these two laws of motion yields the optimality equation, while matching the volatilities yields the equilibrium volatility of the continuation value, $r\beta_t = U'(X_t)$. The second step is to show that the optimality equation has at least one solution that lies in the range of feasible payoffs for the long-run player.

The innovative part of the proof lies in establishing existence of a solution to the optimality equation when the state space is unbounded, particularly when g is also unbounded. I show by construction that there exist lower and upper solutions to the optimality equation, $\alpha : \mathcal{X} \to \mathbb{R}$ and $\beta : \mathcal{X} \to \mathbb{R}$, that have linear growth. This is only possible when the maximum drift of the state has linear growth at rate less than r (Assumption 2). The lower and upper solutions characterize bounds on the solution to the optimality equation, $\alpha(X) \leq U(X) \leq \beta(X)$ for all X. Next I show that the bound on the optimality equation grows linearly with respect to $U'(X)$, and therefore the optimality equation does not grow too quickly (technically speaking, it satisfies a growth condition on any compact subset of the state space). These conditions establish that the optimality equation has a twice continuously differentiable solution with linear growth. When g is bounded, the lower and upper solutions are constant, which establishes existence of a bounded solution.

The final step is to show that the continuation value and actions characterized above constitute a Markov equilibrium. Given a solution $U(X)$ and an action profile uniquely specified at state X_t by $(a_t^*, \overline{b}_t^*) = S^*(X_t, U'(X_t))$ (where uniqueness follows from Assumption 3), the state variable evolves uniquely according to (1), the continuation value $(U(X_t))_{t\geq 0}$ satisfies the law of motion (5) and the action profile satisfies the conditions for sequential rationality, (6). Therefore, $(a_t^*, \overline{b}_t^*, U(X_t))$ constitute a PPE.

For a given solution $U(X)$, the state evolves uniquely and actions are uniquely specified as a function of the state. Therefore, *each* solution to the optimality equation characterizes a unique Markov equilibrium. If there are multiple solutions, then there will be multiple

Markov equilibria. There are no Markov equilibria other than those characterized by these solutions.

4.2 The PPE Payoff Set

Next, I show that in any PPE, the long-run player cannot achieve a payoff above the highest or below the lowest Markov equilibrium payoff. Let $\xi : \mathcal{X} \rightrightarrows \mathbb{R}$ denote the correspondence that maps each state onto the corresponding set of PPE payoffs for the long-run player, and let $\Upsilon : \mathcal{X} \rightrightarrows \mathbb{R}$ denote the analogous correspondence for Markov equilibrium payoffs.

Theorem 2. *Assume Assumptions 1, 2 and 3. Then for any state $X \in \mathcal{X}$, the set of PPE payoffs of the long-run player at state X is equal to the convex hull of the set of Markov equilibrium payoffs at state X, $\xi(X) = \text{co}(\Upsilon(X))$.*

The impossibility of the long-run player achieving a PPE payoff above the highest Markov payoff yields insight into the role that persistence plays in generating intertemporal incentives. In a Markov equilibrium, payoffs and actions depend only on the current value of the state. When this Markov equilibrium yields the highest equilibrium payoff, it precludes the existence of equilibria that achieve higher payoffs using *informational* incentives i.e. using past signals to coordinate future equilibrium play by punishing or rewarding the long-run player. Value-burning is the only feasible type of informational incentives when there is a single long-run player, but Brownian information is too noisy to create effective incentives via value-burning (Sannikov and Skrzypacz 2010).Thus, the ability to generate effective intertemporal incentives in stochastic games solely stems from *structural* incentives i.e. the way in which past actions impact the state, which, in turn, impacts future feasible payoffs and equilibrium play.

Outline of Proof. The key argument in the proof shows that any PPE with an initial payoff above the highest Markov equilibrium payoff will eventually yield a continuation value that lies outside of the set of feasible payoffs for the long-run player, which is a contradiction. This escape argument is similar to other papers in the literature, including Faingold and Sannikov (2011).

Given initial state X_0, suppose a PPE with continuation values $(W_t)_{t\geq 0}$ yields a payoff higher than the maximum Markov equilibrium payoff U, i.e. $W_0 > U(X_0)$. Let $D_t = W_t - U(X_t)$ be the difference between these two payoffs. The innovative parts of the proof are to establish that the volatility of D_t is bounded away from zero on an unbounded state space and to show that the above argument can be applied to unbounded flow payoffs when the state does not grow too quickly. When the volatility of D_t is positive, I show that D_t will grow arbitrarily large with positive probability, independent of X_t. By Lemma 2, $|W_t(S)|$ is bounded with respect to X_t. Thus, D_t can only grow arbitrarily large when X_t grows arbitrarily large, leading to a contradiction. Faingold and Sannikov (2011) rely on the compactness of the state space to show that the volatility of D_t is bounded away from zero and rely on the boundedness of the flow payoff to reach a contradiction when D_t grows arbitrarily large. Therefore, their proofs do not trivially extend to an unbounded state space or flow payoff.

Equilibrium Degeneracy without Persistent Actions. If the state evolves independently of the long-run player's action, then there is no link between the current action and the continuation value. It is not possible to generate effective intertemporal incentives and the long-run player acts myopically. In the unique PPE, both players play the static Nash equilibrium action profile $S^*(X, 0)$ at all states X.

Corollary 1. *Assume Assumptions 1, 2 and 3 and suppose μ is independent of a for all X. Then in the unique PPE, $(a_t, \bar{b}_t) = S^*(X_t, 0)$ for all $t \geq 0$ and the continuation value is characterized by the unique solution to the optimality equation (7).*

This is the stochastic game analogue of the equilibrium degeneracy result for repeated games in Fudenberg and Levine (2007) and Faingold and Sannikov (2011).

4.3 Equilibrium Uniqueness

This section establishes sufficient conditions for there to be a unique PPE, which is Markov. The main step is to determine when the optimality equation has a unique feasible solution. When this is the case, by Theorem 1, there is a unique Markov equilibrium and by Theorem 2, PPE payoffs are uniquely specified as the payoffs in this unique Markov equilibrium. The limit behavior of a solution to the optimality equation as the state approaches its boundary plays a key role. From the optimality equation, any two feasible solutions that satisfy the same boundary conditions cannot differ on the interior of the state space – these solutions must be equivalent. Therefore, establishing that all feasible solutions satisfy the same boundary conditions is necessary and sufficient to establish a unique solution.

The next two sections outline three sets of sufficient conditions to guarantee unique boundary conditions, for the case of (i) an unbounded flow payoff and state space, (i) a bounded flow payoff and unbounded state space, and (iii) a bounded flow payoff and state space. The key component of these assumptions is to establish that the volatility of the continuation value is unique at the boundary – and therefore, the long-run player plays a unique action at the boundary. In Section 6.3, I present an application with multiple Markov equilibria, and illustrate how these conditions fail.

4.3.1 Unbounded Flow Payoff.

Assumption 4 outlines a set of sufficient conditions for a unique Markov equilibrium when the flow payoff is unbounded. The first condition requires the long-run player's flow payoff and the drift to be additively separable in the impact of the state X and incentive weight z. This rules out complementarities between the direct and equilibrium channels for incentives near the boundary, preventing multiple optimal action profiles due to coordination motives. The second condition requires the change in the incentive constraint to be monotone with respect to the state, for any incentive weight z. This rules out games in which $U'(X)$ oscillates infinitely often (this follows directly from the optimality equation). From Theorem 1, in any Markov equilibrium, $U(X)$ has linear growth. Taken together, this ensures that the slope of the continuation value $U'(X)$ has a well-defined limit as the state grows large. Combined with additive separability, this establishes that the slope of the continuation value converges to the same limit in all Markov equilibria. The final condition, Lipschitz continuity of $\sigma^*(X,z)^2$, is a technical condition that guarantees two distinct Markov equilibria $U_1(X)$ and $U_2(X)$ cannot have the same limit slope.

Assumption 4. *Suppose g is unbounded and the following conditions hold:*

1. *Additive separability: there exist Lipschitz continuous functions $g_1, \mu_1 : \mathcal{X} \to \mathbb{R}$ and $g_2, \mu_2 : \mathbb{R} \to \mathbb{R}$ such that $g^*(X,z) = g_1(X) + g_2(z)$ and $\mu^*(X,z) = \mu_1(X) + \mu_2(z)$.*

2. *Monotonicity: there exists a $\delta > 0$ such that for $|X| > \delta$, $g_1'(X) + z\mu_1'(X)/r$ is monotone for all $z \in \mathbb{R}$.*

3. *Volatility: $\sigma^*(X,z)^2$ is Lipschitz continuous.*

Combining these insights, Theorem 3 establishes uniqueness and characterizes the limit of the continuation value and its slope in the unique Markov equilibrium.

Theorem 3. *Assume Assumptions 1, 2, 3 and 4. Then for each initial state $X_0 \in \mathcal{X}$, there exists a unique PPE, which is Markov and characterized by the unique solution $U(X)$ of (7) with linear growth. For $p \in \{-\infty, \infty\}$, this solution satisfies*

$$\lim_{X \to p} U(X) - y(X) = g_2(z_p) + z_p \mu_2(z_p)/r \tag{9}$$

$$\lim_{X \to p} U'(X) = z_p, \tag{10}$$

where

$$z_p \equiv \lim_{X \to p} \frac{r g_1(X)}{rX - \mu_1(X)} \tag{11}$$

is the asymptotic slope of the continuation value,

$$y(X) \equiv -\phi(X) \int \frac{r g_1(X)}{\phi(X)\mu_1(X)} dX \tag{12}$$

is the continuation payoff that the long-run player would earn from repeated play of the static Nash equilibrium profile, and $\phi(X) \equiv \exp(\int r/\mu_1(X)dX)$.

From (9), as the state becomes large, the continuation value approaches the sum of a constant, $g_2(z_\infty) + z_\infty \mu_2(z_\infty)/r$, and the payoff that the long-run player would earn in the static Nash equilibrium profile. This constant has an important interpretation. The first term, $g_2(z_\infty)$, is the *equilibrium value* of the strategic interaction between the long-run and short-run players at large states. It is the portion of the equilibrium flow payoff that arises from strategic interaction and captures the effect of the long-run player's action a on the short-run players' aggregate action, net of the cost of a. The second term, $z_\infty \mu_2(z_\infty)/r$, captures the *return on effort*, measured by how the state changes with respect to the long-run player's action and how the continuation value changes with respect to the state. Taken together, if this constant is positive, then as the state becomes large, the long-run player's payoff is strictly higher than the payoff from playing the static Nash profile at each state. Similar intuition holds as the state becomes small.

From (10), as the state becomes large, the slope of the continuation value $U'(X)$ converges to a unique limit slope (11). This limit depends on the ratio of the growth rate of the flow payoff to the growth rate of the drift with respect to the state. When the limit is positive, it is possible to sustain non-trivial intertemporal incentives. This is an important and novel insight of the paper. If it is possible to sustain non-trivial incentives at the boundary of the state space, then incentives are permanent in the sense that they don't dissipate with time, regardless of the asymptotic behavior of the state with respect to time.

In applications, it is straightforward to verify Assumption 4 and derive the boundary conditions in Theorem 3, as illustrated in Section 6. Section 6.3 presents an application in which Assumption 4 fails, and shows that there are multiple Markov equilibria.

Outline of Proof. The innovative part of this proof is to establish the boundary conditions for an unbounded flow payoff and state space. Let

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

be the sum of the long-run player's flow payoff and return on effort at the sequentially rational action profile $(a(X, z), b(X, z))$. Let $U(X)$ be the continuation value in a Markov equilibrium. Suppose that the slope of $U(X)$ doesn't converge as $X \to \infty$. Then for any slope z such that the continuation value has slope z infinitely often at large X, $U(X)$ will alternate between being convex and concave at slope z. From the optimality equation, $\psi(X, z)$ will lie above $U(X)$ when it is concave at slope z, and will lie below $U(X)$ when it is convex at slope z. Therefore, the oscillation of ψ' is at least as large as the oscillation of U'. This violates the monotonicity part of Assumption 4. Therefore, U' is monotonic for large X. By Theorem 1, $U(X)$ has linear growth. Therefore, $\lim_{X\to\infty} U'(X)$ exists and is equal to $\lim_{X\to\infty} U(X)/X$. Denote these limits by $z_\infty \in \mathbb{R}$.

Given the monotonicity assumption and the Lipschitz continuity of μ^* and g^*, the limits of $\psi(X, z)/X$ and $\psi'(X, z)$ exist and are equal as $X \to \infty$. Denote these limits by $\psi_\infty(z)$. The linear growth of $U(X)$ and the Lipschitz continuity of $\sigma^*(X, z)^2$ guarantee that the second derivative term in the optimality equation converges to zero,

$$\lim_{X\to\infty} \sigma^*(X, U'(X))^2 U''(X) = 0.$$

Therefore, from the optimality equation,

$$\lim_{X\to\infty} U(X)/X - \psi(X, U'(X))/X = 0. \tag{13}$$

But $\lim_{X\to\infty} U(X)/X = z_\infty$ and $\lim_{X\to\infty} \psi(X, U'(X))/X = \psi_\infty(z_\infty)$. Therefore, it must be that $z_\infty = \psi_\infty(z_\infty)$ and the limit slope z_∞ is a fixed point of $\psi_\infty(z)$. The additively separable assumption on g^* and μ^* is sufficient to ensure that $\psi_\infty(z)$ has a unique fixed point. This guarantees that all Markov equilibria have the same limit slope as the state grows large. This slope is the unique fixed point of $\psi_\infty(z)$ and can be calculated from (11).

From the optimality equation and the uniqueness of the limit slope, there exists a constant $c \in \mathbb{R}$ such that any Markov equilibrium continuation value $U(X)$ satisfies

$$\lim_{X\to\infty} U(X) - U'(X)\mu_1(X)/r - g_1(X) = c.$$

Consider the linear first order differential equation

$$y(x) - y'(x)\mu_1(x)/r - g_1(x) - c = 0. \tag{14}$$

When the growth rate of μ_1 is in $[0, r)$, then there is a unique linear growth solution y of (14), and when the growth rate of μ_1 is less than zero, then any two linear growth solutions y_1 and y_2 satisfy $\lim_{x\to\infty} y_1(x) - y_2(x) = 0$. Therefore, any Markov equilibrium $U(X)$ satisfies $\lim_{X\to\infty} U(X) - y(X) = 0$ for all linear growth solutions y to (14).

Let U and V be two Markov equilibrium continuation values. By the above reasoning, both continuation values satisfy the same boundary conditions. Therefore, for $p \in \{-\infty, \infty\}$, $\lim_{X\to p} U(X) - V(X) = 0$ and $\lim_{X\to p} U'(X) = \lim_{X\to p} V'(X) = z_p$. Similar to Faingold and Sannikov (2011), if there exists an X such that $U(X) - V(X) > 0$, the structure of the optimality equation prevents these continuation values from satisfying the same boundary conditions for at least one boundary.

4.3.2 Bounded Flow Payoff.

Assumption 4′ outlines two sufficient conditions for a unique Markov equilibrium when the flow payoff is bounded, for an unbounded and bounded state space, respectively. Both conditions guarantee that in any equilibrium, incentives collapse as the state approaches the boundary. This rules out the possibility of sustaining multiple equilibrium action profiles at the boundary – the long-run player's equilibrium action converges to the static Nash action. From here, it is possible to show that all Markov equilibria have the same payoff at the boundary, and it is not possible for two such Markov equilibria to be distinct on the interior of the state space. Therefore, there is a unique Markov equilibrium.

When the state space is unbounded, incentives collapse when the slope of the continuation value converges to zero. From the optimality equation, the oscillation of the continuation value is bounded by the oscillation of the static Nash payoff $g^*(X,0)$. Therefore, if $g^*(X,0)$ becomes monotone as the state grows large, then the continuation value cannot oscillate as the state becomes large. Further, the flow payoff is bounded, and therefore, the continuation value must also be bounded. Taken together, this establishes that the continuation value converges to a well-defined limit, and its slope must converge to zero. When the state space is bounded, absorbing states at the boundary states serve a similar role. At an absorbing state, the drift is equal to zero and the long-run player's action has no impact on the evolution of the state. Therefore, even if the slope of the continuation value is non-zero, the incentive constraint of the long-run player collapses.

Assumption 4′. *Either of the following conditions hold:*

(a) The flow payoff $g(a,b,X)$ is bounded, the state space is unbounded, $\mathcal{X}=\mathbb{R}$, and there exists a $X_1>0$ such that for $|X|>X_1$, the static Nash payoff $g^*(X,0)$ is monotone.

(b) The state space $\mathcal{X}$ is compact and the boundary states $\underline{X}$ and $\overline{X}$ are absorbing states.

Theorem 4 establishes uniqueness and characterizes the boundary conditions when incentives collapse at the boundary. Equation (15) states that the continuation value approaches the static Nash payoff as the state approaches the boundary, and (16) states that intertemporal incentives dissipate at the boundary states.[21]

Theorem 4. *Assume Assumptions 1, 2, 3 and 4′. Then for each initial state $X_0 \in \mathcal{X}$, there exists a unique PPE, which is Markov and characterized by the unique bounded solution $U(X)$ of (7). For $p \in \{\underline{X}, \overline{X}\}$, this solution satisfies*

$$\lim_{X\to p} U(X) - g^*(X,0) = 0 \tag{15}$$

$$\lim_{X\to p} \mu^*(X,U'(X))U'(X) = 0. \tag{16}$$

[21] When Assumption 4′ holds, Theorems 1, 2 and 4 also hold for an alternative version of Assumption 3. Specifically, assume that the restriction of S^* to $\mathcal{X}\times[0,\infty)$ is non-empty, single-valued and returns $\bar{b}=\delta_b$ for some $b\in B(X)$, where δ_b is the Dirac measure on action b, S^* is Lipschitz continuous on every bounded subset of $\mathcal{X}\times[0,\infty)$, and the static Nash payoff $g^*(X,0)$ is increasing in X. Change the definition of S^* to set $S^*(X,z)=S^*(X,0)$ for $z<0$. By Proposition 2, the solution $U(X)$ is increasing, and the values for $z<0$ are irrelevant. An analogous restriction to $(-\infty,0]$ is possible when $g^*(X,0)$ is decreasing in X.

Even when incentives collapse at the boundary, the state does not necessarily converge to a boundary state as $t \to \infty$. Therefore, Theorem 4 does not preclude the existence of long-run incentives, as the state may never reach the boundary.

If the state space is unbounded and $g^*(X, 0)$ oscillates infinitely often, the continuation value can oscillate and effective intertemporal incentives are possible at large or small states. If the state space is bounded and the boundary states are not absorbing, then effective intertemporal incentives are possible at the boundary when the continuation value is pulled away from the boundary quickly enough. In either case, uniqueness will still obtain in some settings. A rigorous characterization of these settings is beyond the scope of this paper.

4.3.3 Example, cont.

The flow payoff is bounded from below as $X \to -\infty$ and unbounded from above as $X \to \infty$. Assumption 4 is the relevant assumption as $X \to \infty$. Effort $a(X, z)$ is independent of X, and therefore, $g^*(X, z)$ and $\mu^*(X, z)$ are linearly separable in (X, z). For large X, $g_1(X) = \lambda X$, $g_2(z) = (1 - \lambda)a(X, z) - a(X, z)^2/2$, $\mu_1(X) = -\theta X$ and $\mu_2(z) = a(X, z)$. Note that $g_1' + z\mu_1'/r = \lambda - z\theta/r$ is monotone in z. Therefore, Assumption 4 is satisfied for large X. Theorem 3 yields the boundary conditions as $X \to \infty$. From (11), the limit slope as X grows large is

$$z_\infty = \lim_{X\to\infty} \frac{r\lambda X}{rX + \theta X} = \frac{r\lambda}{r + \theta}.$$

Equilibrium effort approaches $a(X, z_\infty) = \lambda/(r + \theta)$, which is strictly positive. Plugging in $\mu_1 = -\theta X$ and $\phi(X) = \exp(-\int(r/\theta X)dX) = X^{-r/\theta}$, from (12),

$$y(X) = -X^{-r/\theta} \int \frac{r\lambda X}{-X^{-r/\theta}\theta X} dX = \frac{r\lambda}{r + \theta} X. \tag{17}$$

The right hand side of (9) yields

$$g_2(z_\infty) + z_\infty \mu_2(z_\infty)/r = \frac{(1 - \lambda)\lambda}{r + \theta} + \frac{\lambda^2}{2(r + \theta)^2}. \tag{18}$$

Plugging (17) and (18) into (9) yields the limit continuation value,

$$\lim_{X\to\infty} U(X) - \frac{r\lambda}{r + \theta} X = \frac{(1 - \lambda)\lambda}{r + \theta} + \frac{\lambda^2}{2(r + \theta)^2}.$$

Assumption 4′ is the relevant assumption as $X \to -\infty$. For very negative X, $g^*(X, z) = -a(X, z)^2/2$. Therefore, $g^*(X, 0) = 0$ is monotone (Assumption 4′). Theorem 4 yields the boundary conditions as $X \to -\infty$. From (15), the continuation value converges to zero, $\lim_{X\to-\infty} U(X) = 0$, and from (16), the limit slope converges to zero, $\lim_{X\to-\infty} U'(X) = 0$. Given these unique boundary conditions, by Theorems 3 and 4, there is a unique PPE.

5 Properties of Equilibrium Payoffs

The optimality equation yields rich insights into how the correspondence of PPE payoffs of the long-run player is tied to the underlying structure of the game. The static Nash equilibrium payoff, $g^*(X, 0)$, is a key determinant of the shape of $U(X)$. Proposition 1 shows that the shape of $g^*(X, 0)$ partially determines the number and type of PPE payoff extrema. Note that $g^*(X, 0)$ is straightforward to derive from the primitives of the game.

Proposition 1. *Assume Assumptions 1, 2 and 3. Let $I \subset \mathcal{X}$ denote a closed interval of states.*

1. *If $g^*(X, 0)$ is constant on I, then $U(X)$ has at most one extremum on I.*
2. *If $g^*(X, 0)$ is strictly monotone on I, then $U(X)$ has at most two extrema on I and is not constant on I. If $g^*(X, 0)$ is strictly increasing (decreasing) on I and $U(X)$ has two extrema, a minimum at state X_1 and a maximum at state X_2, then $X_1 < X_2$ ($X_1 > X_2$).*
3. *If $g^*(X, 0)$ has N extrema on I, then $U(X)$ has at most $N + 2$ extrema on I.*

The intuition for Proposition 1 stems from the behavior of $U(X)$ at interior extrema. If there is an extremum at state X, then $U'(X) = 0$ and the optimality equation simplifies to

$$U(X) = g^*(X, 0) + U''(X)\sigma^*(X, 0)^2/2r.$$

This depends on the static Nash payoff $g^*(X, 0)$ and whether the extremum is a maximum or minimum. At a minimum, $U''(X) \geq 0$, and therefore, $U(X) \geq g^*(X, 0)$. Similarly, at a maximum, $U''(X) \leq 0$, and therefore, $U(X) \leq g^*(X, 0)$. Hence, the oscillation of the equilibrium payoff $U(X)$ is bounded by the oscillation of the static Nash payoff $g^*(X, 0)$.

When incentives collapse at the boundary states, then it is possible to characterize additional results. Proposition 2 relates the monotonicity or single-peakedness of $U(X)$ to the monotonicity or single-peakedness of $g^*(X, 0)$.

Proposition 2. *Assume Assumptions 1, 2, 3 and 4′.*

1. $g^*(X,0)$ *constant on* $\mathcal{X} \Leftrightarrow U(X)$ *constant on* $\mathcal{X}$.
2. *If* $g^*(X,0)$ *is (strictly) monotonically increasing (decreasing) on* $\mathcal{X}$*, then* $U(X)$ *is (strictly) monotonically increasing (decreasing) on* $\mathcal{X}$.
3. *If* $g^*(X,0)$ *is single-peaked with a maximum (minimum) and* $g^*(\underline{X},0) = g^*(\overline{X},0)$*, then* $U(X)$ *is single-peaked with a maximum (minimum).*
4. *If* $g^*(X,0)$ *has* N *extrema on* $\mathcal{X}$*, then* $U(X)$ *has at most* N *extrema on* $\mathcal{X}$.

When there are multiple Markov equilibria, an analogous result to Proposition 2 holds in any Markov equilibrium $U(X)$ in which incentives collapse at the boundary states, i.e. $\lim_{X\to p} \mu^*(X, U'(X))U'(X) = 0$ for $p \in \{\underline{X}, \overline{X}\}$.

Applying Propositions 1 and 2 to specific settings will yield structural empirical predictions about how equilibrium behavior and actions change with the state.

Example, cont. The static Nash payoff is $g^*(X,0) = \max\{0, \lambda X\}$, which is increasing in X for $\lambda > 0$. By Proposition 2, $U(X)$ is increasing in X. Combined with Proposition 1.2 establishes that $U(X)$ is strictly increasing for $X \in [0,\infty)$.

If the flow payoff of the long-run player is bounded, then the correspondence of PPE payoffs of the long-run player is also bounded. Let $\overline{W} \equiv \sup_{X\in\mathcal{X}} U(X)$ and $\underline{W} \equiv \inf_{X\in\mathcal{X}} U(X)$ be the highest and lowest PPE payoffs of the long-run player across all states, and let X_H and X_L denote the (possibly infinite) states that yield these payoffs. Proposition 3 establishes a bound on PPE payoffs across all states.

Proposition 3. *Assume Assumptions 1, 2 and 3 and* $g(a,b,X)$ *is bounded. Then the highest (lowest) PPE payoff of the long-run player across all states is bounded above (below) by the static Nash payoff of the long-run player at the corresponding state,*

$$g^*(X_L,0) \le \underline{W} \le \overline{W} \le g^*(X_H,0),$$

where $g^*(X_H,0) = \limsup_{X\to X_H} g^*(X,0)$ *if* $X_H \in \{-\infty,\infty\}$*, and analogously for* $X_L \in \{-\infty,\infty\}$.

This bound follows directly from the optimality equation. Suppose there is an interior state X_H such that $\overline{W} = U(X_H)$. When $X_t = X_H$, W_t must have a weakly negative drift and zero volatility so as not to exceed $\overline{W}$. From Lemma 1, this implies $g(a_t, \bar{b}_t, X_H) \geq \overline{W}$. From Theorem 1, $U'(X_H) = 0$, and therefore $g(a_t, \bar{b}_t, X_H) = g^*(X_H, 0)$. Combining these conditions yields $\overline{W} \leq g^*(X_H, 0)$. If the continuation value is sufficiently flat around X_H (i.e. $U''(X_H) = 0$) or X_H is an absorbing state, then $\overline{W} = g^*(X_H, 0)$. Otherwise, $\overline{W} < g^*(X_H, 0)$, as either the continuation value or the state changes too quickly at X_H to maintain $g^*(X_H, 0)$.

In general, it may be difficult to characterize X_H from the primitives of the game, as X_H does not necessarily correspond to the state that maximizes $g^*(X, 0)$. A weaker bound that follows immediately and can be easily characterized is that $\overline{W}$ is bounded above by the highest static Nash payoff across all states, $\overline{W} \leq \sup_{X \in \mathcal{X}} g^*(X, 0)$, and similarly, $\underline{W} \geq \inf_{X \in \mathcal{X}} g^*(X, 0)$.

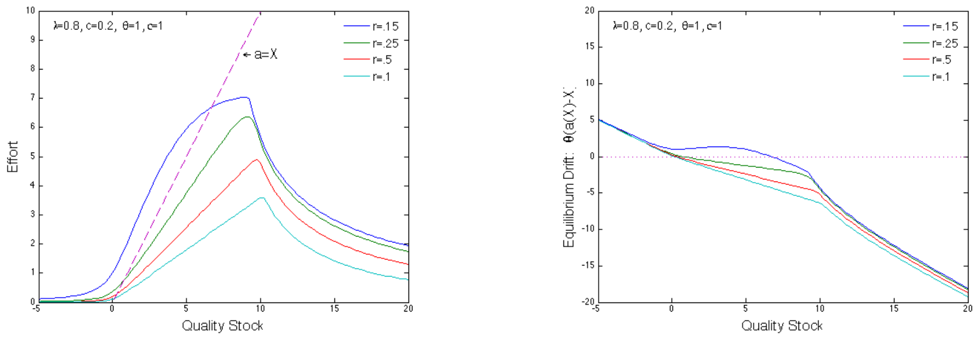
6 Applications

This section develops several applications to illustrate the breadth of the model. Section 6.1 presents two variations on the product choice setting introduced in Section 2: (i) the consumer has a budget constraint and (ii) the marginal value of quality varies. In Section 6.2, a government seeks to implement a policy target, and the policy variable is a persistent function of past effort. In Section 6.3, a government and a sequence of innovators build intellectual capital. Their investments are strategic complements, which leads to multiple equilibria.

6.1 Variations of Persistent Quality

Consumer Budget Constraint. In the example introduced in Section 2, consumers' willingness to pay linearly increases with quality $q(a, X)$ and becomes arbitrarily large as the stock quality grows large. We saw in this example that the firm chooses positive effort and continues to do so for large X. More realistically, suppose that marginal value of quality is decreasing. In particular, assume that each consumer has a budget constraint and is willing to pay up to B for the product, $\bar{b} \in [0, B]$. The best response is the same as in Section 2, except now $\bar{b} = B$ if $q(\tilde{a}, X) \geq B$. Payoffs and the drift of the stock quality are as defined in Section 2.

When $\bar{b}$ is bounded, the flow payoff is also bounded and Assumption 4′ is the relevant



(a) Equilibrium Effort

(b) Equilibrium Drift of Stock Quality

FIGURE 1

condition for uniqueness. The static Nash equilibrium payoff is the willingness to pay of the consumer under zero effort, $g^*(X, 0) = \max\{0, \min\{\lambda X, B\}\}$, which is monotone in X. Therefore, there is a unique Markov perfect equilibrium. Figure 1(a) plots equilibrium effort, as derived from the equilibrium characterization in Theorem 1. The firm has the strongest incentive to invest at intermediate quality stock levels, as revenue rapidly increases with quality. This *reputation building* phase is characterized by high effort and rising quality. When the firm has high stock quality, the firm *rides* its good reputation by enjoying high payoffs today at the expense of allowing the quality to drift down. Very negative shocks lead to periods of *reputation recovery* where the firm chooses low effort and allows quality to recover before beginning to rebuild. Figure 1(b) plots the equilibrium drift of the stock quality, which illustrates these three phases. 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 the stock quality is stable.

Figure 2 illustrates the equilibrium continuation value $U(X)$ for several discount rates. It is convex at low levels of the stock quality and concave at high levels. When the stock quality is high, consumers are purchasing near their maximum level. The firm is risk averse in quality, in that negative quality shocks reduce revenue more than positive quality shocks increase revenue. On the other hand, when the stock quality is low, the firm faces the potential for substantial gains if quality rises, but the risk of loss from a negative quality shock is small. The continuation payoff has an interesting non-monotonicity with respect to the interaction between the discount rate and current stock quality. This 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

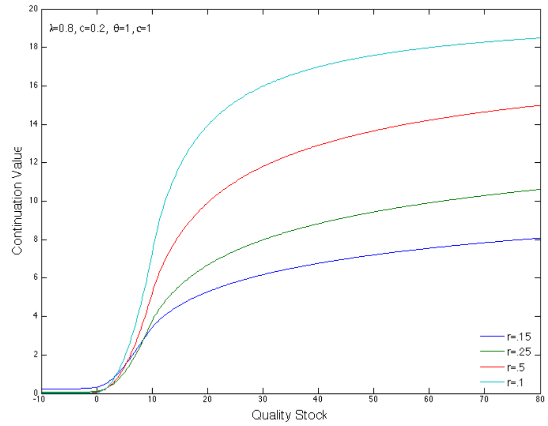


FIGURE 2. Continuation Value

quality. On the other hand, a low discount rate means that transitory positive shocks to quality have a lower value relative to the long-run expected quality. When stock quality is low, the first effect dominates and low discount rates yield higher payoffs; this relationship flips when the stock quality becomes large.

Quality Specialization. Consider a variation of the example from Section 2 in which the marginal value of quality is non-monotonic in the current quality. For example, the marginal value of an upgrade to a new software version is larger for some versions and smaller for others, while the cost of investing in developing this upgrade is constant with respect to the current version (i.e. constant with respect to X). In particular, assume that each consumer's expected value for the product is $q(\tilde{a}, X) + \sin q(\tilde{a}, X)$, where $q(\tilde{a}, X) = (1 - \lambda)\tilde{a} + \lambda X$ as before. Note that the expected value is increasing in X, so higher stock quality is always more valuable, but the marginal value of an increase in the stock quality oscillates between having a large and small impact on the expected value. The consumer chooses $\bar{b} \in [0, B]$ and is willing to pay his expected value for the product. The remainder of the model is as defined in Section 2.

When the *marginal* return to quality is non-monotonic, the equilibrium dynamics are quite different. There are multiple regions in which the marginal return to investing in quality is high versus low, which means that the firm's equilibrium effort will oscillate with respect to the current stock quality. When quality is such that the marginal return is high, the firm chooses a high level of effort to build quality. Once the firm enters a region where the marginal return is flat, it slacks off and lets quality decay. With positive probability, quality drifts back down to a level at which the firm has an incentive to invest again. But also with positive probability, the firm receives a quality shock that pushes the stock quality to a higher level with a high marginal return, and the firm begins investing to maintain this new, higher level of quality. This leads firms to specialize at different levels of quality. A low quality firm may be better off remaining a low quality firm, rather than trying and failing to move up the market. But if a firm gets a breakthrough and reaches a high quality level, it will then have the incentive to invest in maintaining this higher quality. The firm's payoff is bounded and the static Nash equilibrium payoff is eventually monotonic, guaranteeing uniqueness of this Markov perfect equilibrium by Assumption 4′.a.

6.2 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 takes on values in $\mathcal{X} = [0, 2]$. Constituents want to target a policy level $X = 1$, but in the absence of intervention, the policy drifts towards its natural level $d \in [0, 2]$. The board can undertake costly intervention $a \in [-1, 1]$ to alter the level of the policy variable. The policy has drift $\mu(a, b, X) = X(2-X)(a+\theta(d-X))$, where $\theta > 0$ captures the persistence of past interventions by the board. A negative intervention decreases X, while a positive intervention increases X. The policy intervention has the largest impact on X at intermediate values. The policy has volatility $\sigma(b, X) = X(2 - X)$ – it is most volatile at intermediate levels.

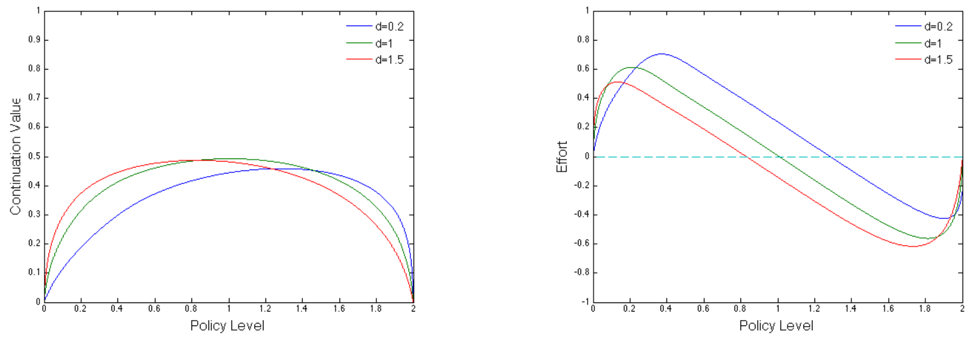
Constituents choose an action each period, which represents their campaign contributions or support for the board. When the policy is at level X and constituents believe the board chooses intervention $\tilde{a}$, they are willing to contribute $\bar{b} = \lambda\tilde{a}^2 + 1 - (1-X)^2$, where $\lambda > 0$. This best response function is a reduced form representation for the constituents' preferences. They pledge higher support when the policy is closer to their preferred target of $X = 1$ and when the board undertakes a stronger intervention. The parameter λ captures constituents' marginal value of an intervention. 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 the magnitude of its intervention, $g(a, \bar{b}, X) = \bar{b} - ca^2$.

This model satisfies the assumptions in Section 3. Volatility is positive, except at the boundary states (Assumption 1). The state space is bounded. Therefore, the board's flow payoff is also bounded (Assumption 2.1). From Lemma 1, given current policy level X and incentive weight $r\beta$, the board chooses intervention $a(X, r\beta) = \max\{-1, \min\{\frac{\beta X(2-X)}{2c}, 1\}\}$. The board will choose an intervention that increases the policy level when the equilibrium incentive weight is positive, and otherwise chooses an intervention that decreases the policy level. The sequentially rational action profile $(a(X, r\beta), \bar{b}(X, r\beta)$ is single-valued and Lipschitz continuous (Assumption 3), where $\bar{b}(X, r\beta) = \lambda a(X, r\beta)^2 + 1 - (1 - X)^2$.

We can use the results from Section 4 to characterize the unique PPE. From Theorem 1, there exists at least one Markov equilibrium. Any solution $U(X)$ to the optimality equation with equilibrium actions $a(X, U'(X)) = \max\{-1, \min\{\frac{U'(X)X(2-X)}{2cr}, 1\}\}$ and $\bar{b}(X, U'(X)) = \lambda a(X, U'(X))^2 + 1 - (1 - X)^2$ characterizes a Markov equilibrium. The state space is compact and the boundary states are absorbing states (Assumption 4′.b). By Theorem 4, there is a unique solution to (7), which characterizes the unique PPE.

Whether the persistence of the policy level incentivizes the governing board to intervene and move the policy toward the constituents' target depends on the shape of the continuation value. From Proposition 2.3, the continuation value is single-peaked with a maximum, and is not constant on any interval of policy levels. Recall that the board's equilibrium intervention is proportional to the slope of the continuation value. Therefore, the board chooses to intervene at almost all policy levels. It intervenes to increase the level at low policy levels, and to decrease the level at high policy levels. The point at which the board switches from a positive to a negative intervention depends on the constituents' target $X = 1$ and the natural drift of the policy d. The continuation value is skewed in the direction of the natural drift. If the natural drift lies above the target, $d > 1$, then at low policy levels, the policy will naturally move toward the target, which benefits the board. At very low policy levels, the board chooses a positive intervention to increase the rate at which the policy moves towards the target. It switches to a negative intervention when the policy is slightly below the target, in order to prevent the policy from overshooting the target. The opposite is the case when $d < 1$.

While the direction of the intervention is determined by the slope of the continuation value, the magnitude is determined by both this slope and the marginal impact of the



(a) Equilibrium Payoff

(b) Board's Equilibrium Behavior

FIGURE 3

intervention on the policy level, $X(2 - X)$. When the current policy level is very far from its target, the board's intervention has a small marginal impact, and the board has a low incentive to undertake a costly intervention. When the policy level is close to its target, the continuation value approaches its maximum and the slope of the continuation value approaches zero. Therefore, the board also has a low incentive to intervene. The strongest incentive to intervene is when the policy variable is an intermediate distance from its target – in this case, the policy level is sensitive to an intervention, and the continuation value is sensitive to a change in the policy. Figure 3(a) plots the equilibrium payoffs and intervention for several levels of d.

6.3 Complementary Investment and Multiple Equilibria

Suppose a government and a sequence of short-run innovators can invest to generate intellectual capital. The state X represents the current level of intellectual capital. The government chooses an investment level $a \in [0, \overline{a}]$, where $\overline{a} > 0$ is the maximum feasible investment for the government, and each innovator chooses investment $b \in [-\gamma|X|, \gamma|X|]$, where $c > 0$ and $\gamma > 0$ are constants. Both government and innovator investment contribute to the growth of intellectual capital, with returns $\theta_1 > 0$ and $\theta_2 > 0$, respectively. Intellectual capital depreciates at rate $\theta_3 > 0$. Therefore, the expected change in intellectual capital is

$$\mu(a, b, X) = \theta_1 b + \theta_2 a - \theta_3 X.$$

Assume the volatility of intellectual capital is constant, $\sigma(b, X) = 1$ (which satisfies Assumption 1) and $\gamma < (r + \theta_3)/\theta_1$, to satisfy Assumption 2.2.[22]

For the innovators, the investment of the government is a strategic complement with their own investment. The current level of intellectual capital is also complementary. For example, when an innovator invests in a new project, her return depends on both the stock of intellectual capital available in the economy, as well as investment from the government to make this intellectual capital accessible. This is captured by payoff

$$h(a, b, X) = abX - cb^2/2,$$

where $cb^2/2$ captures the cost of effort. Note that intellectual capital is only valuable to investors if they invest and the government invests.

The government does not directly value intellectual capital. It receives a return of $\alpha > 0$ on each innovator's investment, and therefore, indirectly values intellectual capital through its impact on future investment by the innovators. For example, the government can tax investment at rate α. Therefore, it may be willing to invest today if this will increase future tax returns. This is captured by payoff

$$g(a, b, X) = \alpha b - a^2/2.$$

It is straightforward to compute the sequentially rational action profile for any (z, X). Given incentive weight z, the government chooses an investment level to solve

$$\max_{a \in [0, \overline{a}]} -a^2/2 + \frac{z}{r}\theta_2 a.$$

This results in sequentially rational investment level

$$a(X, z) = \begin{cases} \theta_2 \frac{z}{r} & \text{if } z/r \in [0, \overline{a}/\theta_2] \\ \overline{a} & \text{if } z/r > \overline{a}/\theta_2 \\ 0 & \text{if } z < 0 \end{cases}$$

for the government. Investment is increasing in the impact that it has on the growth of

[22]Note that innovator investment can be unboundedly negative. The advantage of this specification is that it yields a closed-form solution for the continuation value, which makes it straightforward to illustrate the existence of multiple equilibria. In the more realistic case that the lower bound on investor effort is zero, the equilibrium characterization is qualitatively similar.

intellectual capital θ_2 and the incentive weight z. When an innovator believes that the government will choose investment level $\tilde{a}$ and the current stock of intellectual capital is X, the innovator's best response is to select investment $\tilde{a}X/c$ if $\tilde{a}/c \leq \gamma$, and otherwise to choose the maximum possible investment, γ X. To reduce the number of cases, assume that an interior solution is always feasible for the innovator, $\overline{a} \leq \gamma c$. This results in sequentially rational investment level

$$b(X,z) = \begin{cases} \frac{\theta_2}{c}\frac{z}{r}X & \text{if } z/r \in [0, \overline{a}/\theta_2] \\ \frac{\overline{a}}{c}X & \text{if } z/r > \overline{a}/\theta_2 \\ 0 & \text{if } z < 0 \end{cases}$$

for each innovator. The innovator's investment is increasing in the investment of the government and the current stock of intellectual capital, reflecting the complementarity of these two inputs. Note that $a(X,z)$ and $b(X,z)$ are unique for each (X,z) and Lipschitz continuous (Assumption 3).

By Theorem 1, there exists at least one Markov equilibrium, and by Theorem 2, the PPE payoff set is bounded by the highest and lowest Markov equilibrium payoffs. The state space is unbounded, so Assumption 4 is the relevant sufficient condition for uniqueness. We first search for equilibria in which the government's optimal investment is an interior solution. If $z/r \in [0, \overline{a}/\theta_2]$, then the government chooses an interior level of investment, yielding

$$g^*(X,z) = \frac{\alpha\theta_2}{c}\frac{z}{r}X - \frac{1}{2}\theta_2^2\left(\frac{z}{r}\right)^2 \tag{19}$$

and

$$\mu^*(X,z) = \frac{\theta_1\theta_2}{c}\frac{z}{r}X + \theta_2^2\frac{z}{r} - \theta_3 X. \tag{20}$$

Neither g^* nor μ^* are additively separable in (X,z). Therefore, Assumption 4 does not hold.

Conjecture that there exists an equilibrium in which the continuation value is linear in the current level of intellectual capital. This means that $U'(X)$ is constant with respect to X and $U''(X) = 0$. Taking the derivative of the optimality equation, (7), such an equilibrium must satisfy

$$rU'(X) = r\frac{dg^*}{dX} + U'(X)\frac{d\mu^*}{dX}. \tag{21}$$

Rearranging terms and plugging in the derivatives of (19) and (20), any solution to

$$z = \frac{\frac{\alpha\theta_2}{c} z}{r - \frac{\theta_1\theta_2}{c}\frac{z}{r} + \theta_3} \tag{22}$$

with $z/r \in [0, \overline{a}/\theta_2]$ is a candidate equilibrium slope. It is straightforward to verify that $z^* = 0$ is a solution. In an equilibrium with slope $z^* = 0$, neither the government nor the innovators invest, $a(X) = b(X) = 0$ for all X, and the government's equilibrium payoff is $U(X) = 0$. Due to the strategic complementarity, if the government doesn't invest, then neither will the innovators, yielding a payoff of zero for all players. If $\alpha = 0$ or $\theta_2 = 0$, $z^* = 0$ is also the unique solution, and therefore, the unique equilibrium. Intuitively, if the government doesn't receive a return on the innovators' investment or its own investment does not contribute to building intellectual capital, then it has no incentive to undertake costly investment.

There are also non-trivial equilibria that sustain positive investment. The unique non-zero solution to (22) is

$$\frac{z^*}{r} = \frac{rc - \alpha\theta_2 + c\theta_3}{\theta_1\theta_2}. \tag{23}$$

In order for this to be a valid solution, it must satisfy $z^*/r \in [0, \overline{a}/\theta_2]$. Recall that by assumption, $\overline{a} \leq \gamma c$ and $\gamma < (r + \theta_3)/\theta_1$. Allowing $\overline{a}$ and γ to be as large as possible, subject to these constraints, yields $\overline{a}/\theta_2 \approx (rc + c\theta_3)/\theta_1\theta_2$ as the upper bound for z^*/r. It is clear from (23) that $z^*/r < (rc + c\theta_3)/\theta_1\theta_2$ for all $\alpha > 0$ and $\theta_2 > 0$. For the lower bound, $z^*/r > 0$ for all r when $c\theta_3 > \alpha\theta_2$. Therefore, there exists an equilibrium that sustains positive investment and has slope (23) when $\gamma \approx (r + \theta_3)/\theta_1$, $\overline{a} = \gamma c$ and $c\theta_3 > \alpha\theta_2$. This equilibrium has non-zero equilibrium investment levels,

$$\begin{aligned} a(X) &= \frac{rc - \alpha\theta_2 + c\theta_3}{\theta_1} \\ b(X) &= \left(\frac{rc - \alpha\theta_2 + c\theta_3}{c\theta_1}\right) X, \end{aligned}$$

and continuation value

$$U(X) = r\left(\frac{rc - \alpha\theta_2 + c\theta_3}{\theta_1\theta_2}\right) X + \frac{(rc - \alpha\theta_2 + c\theta_3)^2}{2\theta_1^2}.$$

The slope captures the government's net present value of the current stock of intellectual capital, while the constant term captures the equilibrium effect stemming from the value of future strategic interaction between the government and the innovators. This latter effect is positive, given $c\theta_3 > \alpha\theta_2$. As the government becomes arbitrarily patient, $r \to 0$, the net present value of the current stock of intellectual capital approaches zero and $U(X) \to \frac{(c\theta_3-\alpha\theta_2)^2}{2\theta_1^2}$. Intuitively, the government cares more about the long-run return from the strategic interaction, rather than the short-run return from the current stock of intellectual capital. This long-run return has a natural interpretation: it is the equilibrium flow payoff for the patient government when the stock of intellectual capital is at its long-run average, which depends on equilibrium investment and the rate of depreciation.[23]

To complete the characterization of Markov equilibria, it remains to check whether there are equilibria with slopes $z < 0$ or $z/r > \overline{a}/\theta_2$. If there is an equilibrium with slope $z < 0$, then from sequential rationality, $a(X, z) = b(X, z) = 0$, which leads to $U(X) = 0$. But then $U'(X) = 0$, which contradicts $z < 0$. Therefore, there are no Markov equilibria with slope $z < 0$. There may be a Markov equilibrium with slope $z/r > \overline{a}/\theta_2$. In such an equilibrium, from sequential rationality, $a(X, z) = \overline{a}$ and $b(X, z) = \overline{a}X/c$. Computing $g^*(X, z)$ and $\mu^*(X, z)$ for this case and plugging the derivatives into (21), the equilibrium slope must satisfy

$$\frac{z^*}{r} = \frac{\alpha\overline{a}}{cr - \theta_1\overline{a} + c\theta_3}. \tag{24}$$

and $z^*/r > \overline{a}/\theta_2$. These conditions are simultaneously satisfied when $\alpha > (cr - \theta_1\overline{a} + c\theta_3)/\theta_2$ – therefore, there is an equilibrium with slope (24) and equilibrium investment levels $a(X) = \overline{a}$ and $b(X) = \overline{a}X/c$.

This example illustrates that both trivial and non-trivial Markov equilibria can exist in a stochastic game when there are complementarities between the players' actions. Even when uniqueness does not hold, Theorem 1 can be used to characterize these Markov equilibria, and Theorem 2 can be used to characterize the PPE payoff set.

[23] Given process $dX_t = \theta(M - X_t)dt + \sigma dZ_t$, the ergodic distribution of X has mean M. From $\mu(a, b, X) = \theta_1 b(X) + \theta_2 a(X) - \theta_3 X$, $\theta = -\theta_1 b(X)/X + \theta_3 = \alpha\theta_2/c - r$. and

$$M = \frac{\theta_2 a(X)}{\theta} = \frac{c\theta_2(rc - \alpha\theta_2 + c\theta_3)}{\theta_1(\alpha\theta_2 - rc)} = \frac{c^2\theta_2\theta_3}{\theta_1(\alpha\theta_2 - rc)} - \frac{\theta_2 c}{\theta_1}.$$

Therefore, the equilibrium ergodic distribution of intellectual capital has mean $\frac{c^2\theta_2\theta_3}{\theta_1(\alpha\theta_2-rc)} - \frac{\theta_2 c}{\theta_1}$. As $r \to 0$, $M \to \frac{c^2\theta_3}{\alpha\theta_1} - \frac{\theta_2 c}{\theta_1} = \frac{c}{\alpha}a(X)$. The expected flow payoff is

$$E[g(a(X), b(X), X)] = E[\alpha b(X) - a(X)^2/2] = \alpha a(X)E[X]/c - a(X)^2/2.$$

As $r \to 0$, $E[g(a(X), b(X), X)] \to \frac{(c\theta_3-\alpha\theta_2)^2}{2\theta_1^2}$, which is equal to $\lim_{r\to 0} U(X)$ derived above.

7 Conclusion

This paper shows that persistence provides an important channel for intertemporal incentives and develops a tractable method to characterize Markov equilibrium behavior and payoffs. The tools developed in this paper will yield insights into equilibrium behavior in a broad range of settings, from industrial organization to political economy to macroeconomics. Once functional forms are specified for payoffs and the evolution of the state, it is straightforward to use Theorem 1 to construct Markov equilibria. This in turn can be used to derive empirically testable comparative statics and predictions about the dynamics of equilibrium behavior based on observable features of the environment. Future research can use this framework to address design questions in specific applications, such as determining the optimal structure of persistence in a rating mechanism.

Furthermore, the equilibrium characterization can be used for structural estimation. Markov equilibria have an intuitive appeal in empirical work, due to their simplicity and dependence on payoff relevant variables to structure incentives. Players do not need to condition on past behavior in a complex way, as actions and payoffs are fully determined by the current value of the state. Establishing that a Markov equilibrium exists and is unique provides a strong justification for focusing on this equilibrium concept, while the equilibrium characterization yields expressions for payoffs and actions that can calibrated and estimated.

A Appendix

A.1 Proofs from Section 3

The following lemma establishes that $(V_t(S))_{t\geq 0}$ is a martingale and $(W_t(S))_{t\geq 0}$ is bounded with respect to $(X_t)_{t\geq 0}$. It is used in the proof of Lemma 1.

Lemma 2. *Assume Assumption 2. For any public strategy profile $S = (a_t, \bar{b}_t)_{t\geq 0}$, initial state X_0 and path of the state variable $(X_t)_{t\geq 0}$ that evolves according to (1) given S,*

1. *$V_t(S)$ is a martingale.*
2. *There exists a $K_W > 0$ such that $|W_t(S)| \leq K_W(1 + |X_t|)$ for all $t \geq 0$.*

Proof of Lemma 2. Suppose g is unbounded. By Assumption 2, there exists a $k \in [0, r)$ and $c > 0$ such that for all $(a, b, X) \in A \times E$, if $X \geq 0$ then $\mu(a, b, X) \leq kX + c$ and if $X \leq 0$ then $\mu(a, b, X) \geq kX - c$. Lipschitz continuous functions have linear growth. Therefore, by Lipschitz continuity of g and σ, the compactness of A and the assumption that $|b| \leq K_b|X| + c_b$ for all $(b, X) \in E$, there exists a $K_g, K_\sigma, c > 0$ such that for all $(a, b, X) \in A \times E$, $|g(a, b, X)| \leq K_g(\frac{c}{k} + |X|)$ and $|\sigma(b, X)| \leq K_\sigma(1 + |X|)$.

Step 1: Derive a bound on $E_\tau|g(a_t, \bar{b}_t, X_t)|$, the expected flow payoff at time t conditional on available information at time $\tau \leq t$. This bound will be independent of the strategy profile. Define $f : \mathcal{X} \to \mathbb{R}$ as

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

Note $f \in \mathcal{C}^2$, $f \geq 0$, $|f'| \leq K_g$ and

$$f''(X) = \begin{cases} 0 & \text{if } |X| \geq 1 \\ \frac{3}{2}K_g\,(1 - X^2) & \text{if } |X| < 1 \end{cases}$$

Ito's Lemma holds for any $\mathcal{C}^2$ function. Given a strategy profile $S = \left(a_t, \bar{b}_t\right)_{t\geq 0}$, initial state $X_\tau < \infty$ and path of the state variable $(X_t)_{t\geq\tau}$ that evolves according to (1),

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

for all $t \geq \tau$, where the first inequality follows from $f'(X)\mu(a, b, X) \leq K_g(k|X| + c)$, $\frac{1}{2} f''(X)\sigma(b, X)^2 \leq 3K_g K_\sigma^2$ and $f'(X)\sigma(b, X)z \leq K_g K_\sigma (1 + |X|)z$ for all $z \in \mathbb{R}$ and for all $(a, b, X) \in A \times E$, and the second inequality follows from the definition of f. The addition of the absolute value sign in $f'(X)\mu(a, b, X) \leq K_g(k|X| + c)$ follows from the sign of f' and the bound on $\frac{1}{2} f''(X)\sigma(b, X)^2$ follows from $f''(X)\sigma(b, X)^2 = 0$ if $|X| \geq 1$ and

$$\begin{aligned} f''(X)\sigma(b, X)^2 &= \frac{3}{2} K_g (1 - X^2)\sigma(b, X)^2 \\ &\leq \frac{3}{2} K_g (1 - X^2) K_\sigma^2 (1 + |X|)^2 \\ &\leq 6K_g K_\sigma^2 \end{aligned}$$

if $|X| < 1$. 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 $|g(a, b, X)| \leq f(X)$ for all $(a, b, X) \in A \times E$. Therefore,

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

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

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

which is finite for any $X_t < \infty$ and $k < r$. Also, given that f 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,

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

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.

If g is bounded with respect to X, then, trivially, $V_t(S)$ and $W_t(S)$ are bounded for all $t \geq 0$ and $X_0 \in \mathcal{X}$, and showing $V_t(S)$ is a martingale follows from step 4. Also note that if $\mathcal{X}$ is bounded, then g is bounded. $\square$

Proof of Lemma 1.

Evolution of the continuation value. From Lemma 2, $V_t(S)$ 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 $dV_t(S) = re^{-rt}\beta_t\sigma\left(\bar{b}_t, X_t\right)dZ_t$. Combining these two expressions for $dV_t(S)$ yields the law of motion for the continuation value:

$$\begin{aligned} dW_t(S) &= r\left(W_t(S) - g(a_t, \bar{b}_t, X_t)\right)dt + r\beta_t\sigma(\bar{b}_t, X_t)dZ_t \\ &= r\left(W_t(S) - g(a_t, \bar{b}_t, X_t)\right)dt + r\beta_t\left(dX_t - \mu(a_t, \bar{b}_t, X_t)dt\right), \end{aligned}$$

where β_t captures the sensitivity of the continuation value to the state variable. Lemma 2 establishes that any continuation value has linear growth with respect to X_t and is bounded when g is bounded.

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

$$dX_\tau = \mu(a_\tau, \bar{b}_\tau, X_\tau)dt + \sigma(\bar{b}_\tau, X_\tau)dZ_\tau$$

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

$$dX_\tau = \mu(\widetilde{a}_\tau, \bar{b}_\tau, X_\tau)dt + \sigma(\bar{b}_\tau, X_\tau)dZ_\tau.$$

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 $(\widetilde{a}_t, \overline{b}_t)$ 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,

$$\begin{aligned} d\widetilde{V}_\tau &= re^{-r\tau}(g(\widetilde{a}_\tau, \overline{b}_\tau, X_\tau) - W_\tau)d\tau + e^{-r\tau}dW_\tau \\ &= re^{-r\tau}(g(\widetilde{a}_\tau, \overline{b}_\tau, X_\tau) - g(a_\tau, \overline{b}_\tau, X_\tau))d\tau + re^{-r\tau}\beta_\tau(dX_\tau - \mu(a_\tau, \overline{b}_\tau, X_\tau)d\tau) \\ &= re^{-r\tau}(g(\widetilde{a}_\tau, \overline{b}_\tau, X_\tau) - g(a_\tau, \overline{b}_\tau, X_\tau) + \beta_\tau(\mu(\widetilde{a}_\tau, \overline{b}_\tau, X_\tau) - \mu(a_\tau, \overline{b}_\tau, X_\tau)))d\tau + \beta_\tau\sigma(\overline{b}_\tau, X_\tau)dZ_\tau. \end{aligned}$$

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

$$\begin{aligned} \widetilde{W}_0 &= E_0\left[\widetilde{V}_\infty\right] = E_0\left[\widetilde{V}_0 + \int_0^\infty d\widetilde{V}_t\right] \\ &= W_0 + E_0\left[r\int_0^\infty e^{-rt}(g(\widetilde{a}_t, \overline{b}_t, X_t) + \beta_t\mu(\widetilde{a}_t, \overline{b}_t, X_t) - g(a_t, \overline{b}_t, X_t) - \beta_t\mu(a_t, \overline{b}_t, X_t))dt\right] \end{aligned}$$

If

$$g(a_t, \overline{b}_t, X_t) + \beta_t\mu(a_t, \overline{b}_t, X_t) \geq g(\widetilde{a}_t, \overline{b}_t, X_t) + \beta_t\mu(\widetilde{a}_t, \overline{b}_t, X_t)$$

holds for all $t \geq 0$, then $W_0 \geq \widetilde{W}_0$ and deviating to $S = (\widetilde{a}_t, \overline{b}_t)$ 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:

$$a_t \in \arg\max_a g(a, \overline{b}_t, X_t) + \beta_t\mu(a, \overline{b}_t, X_t).$$

A.2 Proof of Theorem 1

Let $\psi(X, z) \equiv g^*(X, z) + \frac{z}{r}\mu^*(X, z)$ be the value of the long-run player's incentive constraint at the sequentially rational action profile for incentive weight z.

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 $\overline{b}_t^* = \overline{b}(X_t)$. By Ito's formula, if a Markov equilibrium exists, the continuation value will evolve according to:

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

Matching the drift of this expression with the drift of the continuation value characterized in Lemma 1 yields the optimality equation for strategy profile $(a^*, \overline{b}^*)$,

$$U''(X) = \frac{2r(U(X) - g(a^*, \overline{b}^*, X))}{\sigma(\overline{b}^*, X)^2} - \frac{2\mu(a^*, \overline{b}^*, X)}{\sigma(\overline{b}^*, X)^2}U'(X) \tag{25}$$

which is a second order non-homogenous differential equation.

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

$$S^*(X, U'(X)) = \left\{ (a^*, \overline{b}^*) : \begin{array}{l} a^* = \arg\max_{a \in A} g(a, \overline{b}^*, X) + U'(X)\mu(a, \overline{b}^*, X)/r \\ \overline{b}^* = \arg\max_{b \in B(X)} h(a^*, b, \overline{b}^*, X) \end{array} \right\}.$$

which is unique by Assumption 3.

Existence of solution to optimality equation. Define $f : \mathcal{X} \times \mathbb{R}^2 \to \mathbb{R}$ as:

$$f(X, U, U') \equiv \frac{2r}{\sigma^*(X, U')^2}(U - \psi(X, U'))$$

which is continuous on $\text{int}(\mathcal{X})$. I establish that the second order differential equation $U'' = f(X, U, U')$ has at least one solution $U \in \mathcal{C}^2$ that takes on values in the interval of feasible payoffs for the long-run player, and therefore, is a solution to (25).

Case 1: Unbounded State Space. Theorem 5.6 from Coster and Habets (2006), which is based on Schmitt (1969), 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 5 (Coster Habets (2006))**.** *Let $\alpha, \beta \in C^2(\mathbb{R})$ be functions such that $\alpha \leq \beta$, $D = \{(t, u, v) \in \mathbb{R}^3 | \alpha(t) \leq u \leq \beta(t)\}$ and $f : D \to \mathbb{R}$ be a continuous function. 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 interval I, there exists a positive continuous function H_I :

$\mathbb{R}^+ \to \mathbb{R}$ *that satisfies the Nagumo condition,*[24]

$$\int_0^\infty \frac{sds}{H_I(s)} = \infty, \tag{26}$$

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

Lemma 3. *If* $\mathcal{X} = \mathbb{R}$, *then* (7) *has at least one solution* $U \in C^2(\mathbb{R})$ *that lies in the range of feasible payoffs for the long-run player.*

Proof. Suppose $\mathcal{X} = \mathbb{R}$. Then (7) is continuous on $\mathbb{R}^3$. Define $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ as

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

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

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 and β are lower and upper solutions to (7) if there exist $\alpha_1, \beta_1, c_a, c_b \geq 0$ such that for all $X \in \mathbb{R}$,

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

and

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

Step 1: By Assumption 2, $\exists k \in [0, r)$ and $c \geq 0$ such that $\mu^*(X, z) \leq kX + c$ 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 (7). Note this step does not require g to be bounded.

[24]The Nagumo condition is a growth condition on the second order differential equation $f(X, U, U')$ and plays an important role in demonstrating the existence of solutions of the boundary value problem.

Step 1a: Find a bound on ψ. 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 (7) 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, showing α and β are lower and upper solutions requires

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

Step 1c: Find conditions on $(\alpha_1, \beta_1, c_a, c_b)$ such that α, β are lower and upper solutions to (7) 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

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

Step 1d: Find conditions on $(\alpha_1, \beta_1, c_a, c_b)$ such that α, β are lower and upper solutions to (7) 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 $\beta' = \frac{1}{2}\beta_1 X\left(3 - X^2\right) \in (-\beta_1, 0]$ and $a' = -\frac{1}{2}\alpha_1 X\left(3 - X^2\right) \in [0,\alpha_1)$. From the bound on ψ,

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

Therefore, this requires

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

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

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

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 (7).

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 (7). Note this step places no restrictions on the relationship between the growth rate of μ and r. Define $\overline{g} \equiv \sup_{(a,b,X)\in A\times E} g(a,b,X)$ and $\underline{g} \equiv \inf_{(a,b,X)\in A\times E} g(a,b,X)$, which exist since g is bounded. Let $\alpha_1 = 0$ and $c_a = -\underline{g}$. Then $\psi(X, \alpha'(X)) = g^*(X,0)$, so $\alpha - \psi(X,\alpha') = \underline{g} - g^*(X,0) \leq 0$ and $\alpha(X) = \underline{g}$ is a lower solution. Similarly, let $\beta_1 = 0$ and $c_b = \overline{g}$. Then $\psi(X,\beta') = g^*(X,0)$, so $\beta - \psi(X,\beta'(X)) = \overline{g} - g^*(X,0) \geq 0$ and $\beta(X) = \overline{g}$ is an upper solution.

Step 3: Show that the Nagumo condition (26), which is a growth condition on $f(X,U,U')$ that Theorem 5 uses to establish existence of a solution to the boundary value problem, is satisfied. Given a compact proper subset $I \subset \mathcal{X}$, 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 $(X,U,U') \in \{I \times \mathbb{R}^2 \text{ s.t. } \alpha(X) \leq U \leq \beta(X)\}$. 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. Define $H_I(z) = K_I(1+z)$. Therefore, $\int_0^\infty \frac{sds}{H_I(s)} = \infty$.

Conclude that $f(X,U(X),U'(X))$ has at least one $\mathcal{C}^2$ solution U such that for all $X \in \mathbb{R}$, $\alpha(X) \leq U(X) \leq \beta(X)$. 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 (7) is undefined at $\{\underline{X}, \overline{X}\}$. The result applied to the current setting is reproduced below.

Lemma 4 (Faingold Sannikov (2011))**.** *Let $E = \{(t,u,v) \in (\underline{t},\overline{t}) \times \mathbb{R}^2\}$ 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*

$|f(t,u,v)| \leq K_I(1+|v|))$ *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* $\alpha \leq U(t) \leq \beta$.

Proof. When $\mathcal{X}$ is compact, the feasible payoff set for the long-run player is bounded, since g is Lipschitz continuous and A is compact. Define $\underline{g} \equiv \inf_{(a,b,X)\in A\times E} g(a,b,X)$ and $\overline{g} \equiv \sup_{(a,b,X)\in A\times E} g(a,b,X)$. □

Lemma 5. *Suppose* $\mathcal{X}$ *is compact. Then* (7) *has at least one* $\mathcal{C}^2$ *solution* U *that lies in the range of feasible payoffs for the long-run player,* $\underline{g} \leq U(X) \leq \overline{g}$.

Proof. Suppose $\mathcal{X}$ is compact. Then (7) is continuous on the set $D = \left\{(X,U,U') \in (\underline{X},\overline{X}) \times \mathbb{R}^2\right\}$. For any closed interval $I \subset (\underline{X},\overline{X})$, there exists a $K_I > 0$ such that

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

for all $(X,U,U') \in I \times [\underline{g},\overline{g}] \times \mathbb{R}$. This follows directly from the fact that $X \in I$, $U \in [\underline{g},\overline{g}]$, g^*,μ^* are bounded on $\mathcal{X}\times\mathbb{R}$, and σ is bounded away from zero on I. Also note that

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

for all $X \in \mathcal{X}$. By Lemma 4, (7) has at least one C^2 solution U such that $\underline{g} \leq U(X) \leq \overline{g}$ for all $X \in \mathcal{X}$. □

Construct a Markov equilibrium. Suppose the state variable initially starts at X_0 and U is a solution to (7). The action profile $(a^*, \overline{b}^*) = S^*(X, U'(X)/r)$ is unique and Lipschitz continuous. 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 according to:

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

This process satisfies (5). Additionally, $(a_t^*, \overline{b}_t^*)_{t\geq 0}$ satisfies (6) given process $(\beta_t)_{t\geq 0}$ with $\beta_t = U'(X_t)/r$. Thus, the strategy profile $(a_t^*, \overline{b}_t^*)_{t\geq 0}$ is a PPE yielding equilibrium payoff

$U(X_0)$.

A.3 Proof of Theorem 2

Let X_0 be the initial state, and let $\overline{U}$ be the upper envelope of the set of solutions to the optimality equation (7). Suppose there is a PPE $S = (a_t, \bar{b}_t)_{t\geq 0}$ that yields an equilibrium payoff $W_0 > \overline{U}(X_0)$. The continuation value in this equilibrium must evolve according to

$$dW_t(S) = r\left(W_t(S) - g(a_t, \bar{b}_t, X_t)\right)dt + r\beta_t\left[dX_t - \mu(a_t, \bar{b}_t, X_t)dt\right] \tag{27}$$

for some process $(\beta_t)_{t\geq 0}$. By sequential rationality, $(a_t, \bar{b}_t) = S^*(X_t, \beta_t)$ for all t, and by Assumption 3, these actions are unique for each (X, β). Define

$$\begin{aligned}
\widehat{g}(X,\beta) &\equiv g(S^*(X,\beta), X) \\
\widehat{\mu}(X,\beta) &\equiv \mu(S^*(X,\beta), X) \\
\widehat{\sigma}(X,\beta) &\equiv \sigma(S^*(X,\beta), X)
\end{aligned}$$

which are Lipschitz continuous, given g, μ, σ and S^* are Lipschitz. The state $(X_t)_{t\geq 0}$ evolves according to (1), given PPE action profile $S = (a_t, \bar{b}_t)_{t\geq 0}$. By Ito's formula, the process $(\overline{U}(X_t))_{t\geq 0}$ evolves according to

$$d\overline{U}(X_t) = \overline{U}'(X_t)\widehat{\mu}(X_t,\beta_t)dt + \frac{1}{2}\overline{U}''(X_t)\widehat{\sigma}(X_t,\beta_t)^2dt + \overline{U}'(X_t)\widehat{\sigma}(X_t,\beta_t)dZ_t. \tag{28}$$

Define a process $D_t \equiv W_t(S) - \overline{U}(X_t)$ with initial condition $D_0 = W_0(S) - \overline{U}(X_0) > 0$. Then D_t evolves according to $dD_t = dW_t(S) - d\overline{U}(X_t)$. Plugging in (27) and (28), the process has drift $rD_t + d(X_t, \beta_t)$, where

$$\begin{aligned}
d(X,\beta) &\equiv r(\overline{U}(X) - \widehat{g}(X,\beta)) - \overline{U}'(X)\widehat{\mu}(X,\beta) - \overline{U}''(X)\widehat{\sigma}(X,\beta)^2/2 \\
&= r(\widehat{g}(X,\overline{U}'(X)/r) - \widehat{g}(X,\beta)) + \overline{U}'(X)(\widehat{\mu}(X,\overline{U}'(X)/r) - \widehat{\mu}(X,\beta)) \\
&\quad +\overline{U}''(X)(\widehat{\sigma}(X,U'(X)/r)^2 - \widehat{\sigma}(X,\beta)^2)/2,
\end{aligned}$$

where the second line follows from substituting (7) for $\overline{U}$, and volatility

$$f(X,\beta) \equiv (r\beta - \overline{U}'(X))\widehat{\sigma}(X,\beta).$$

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

Proof. Suppose $|f(X,\beta)| = 0$ for some (X,β). Then $r\beta = \overline{U}'(X)$. The action profile associated with $(X, \overline{U}'(X)/r)$ corresponds to the actions played in a Markov equilibrium at state X. Therefore, $d(X,\beta) = 0$. □

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

Proof. Suppose the state space is unbounded, $\mathcal{X} = \mathbb{R}$. Fix $\varepsilon > 0$ and suppose $d(X,\beta) \leq -\varepsilon$. Show that there exists a $\eta > 0$ such that $|f(X,\beta)| > \eta$ for all $(X,\beta) \in \mathcal{X} \times \mathbb{R}$.

Step 1. Show $\exists M > 0$ such that this is true for $(X,\beta) \in \Omega_a \equiv \{\mathcal{X} \times \mathbb{R} : |\beta| > M\}$.
$\overline{U}'$ is bounded, by Assumption 1, σ is bounded away from 0. Therefore, there exists an $M > 0$ and $\eta_1 > 0$ such that $|f(X,\beta)| > \eta_1$ for all $|\beta| > M$, regardless of d.

Step 2. Show $\exists X^* > 0$ such that this is true for $(X,\beta) \in \Omega_b \equiv \{\mathcal{X} \times \mathbb{R} : |\beta| \leq M, |X| > X^*\}$.
Consider the set $\Phi_b \subset \Omega_b$ with $d(X,\beta) \leq -\varepsilon$. It must be that β is bounded away from $\overline{U}'(X)/r$ on Φ_b. Suppose not. Then either (i) there exists some $(X,\beta) \in \Phi_b$ with $\beta = \overline{U}'(X)/r$, which implies $|f(X,\beta)| = 0$ and therefore $d(X,\beta) = 0$, a contradiction, or (ii) as X becomes large, the boundary of the set Φ_b approaches $\beta = \overline{U}'(X)/r$, which implies that for any $\delta_1 > 0$, there exists an $(X,\beta) \in \Phi_b$ with $r\beta - \overline{U}'(X) < \delta_1$. Choose δ_1 so that $|\widehat{g}(X, \overline{U}'(X)/r) - \widehat{g}(X,\beta)| < \varepsilon/4r$ and $|\widehat{\mu}(X, \overline{U}'(X)/r) - \widehat{\mu}(X,\beta)| < \varepsilon/4k$, where $|\overline{U}'(X)| \leq k$ is the bound on $\overline{U}'$. Then $|d(X,\beta)| < \varepsilon/4 + \varepsilon/4 = \varepsilon/2$ which is a contradiction. Therefore, there exists a η_2 such that $|f(X,\beta)| > \eta_2$ on Φ_b. Then on the set Ω_b, if $d(X,\beta) \leq -\varepsilon$ then $|f(X,\beta)| > \eta_2$.

Step 3. Show this is true for $(X,\beta) \in \Omega_c \equiv \{\mathcal{X} \times \mathbb{R} : |\beta| \leq M \text{ and } |X| \leq X^*\}$.
Consider the set $\Phi_c \subset \Omega_c$ where $d(X,\beta) \leq -\varepsilon$. 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 $|f(X,\beta)| > \eta_3$ for all $(X,\beta) \in \Phi_c$.

Take $\eta = \min\{\eta_1, \eta_2, \eta_3\}$. Then when $d(X,\beta) \leq -\varepsilon$, $|f(X,\beta)| > \eta$. The proof for a bounded state space is analogous, omitting step 2b. □

Lemma 8. *Any PPE payoff W_0 is such that $\underline{U}(X_0) \leq W_0 \leq \overline{U}(X_0)$ where $\overline{U}$ and $\underline{U}$ are the upper and lower envelope of the set of solutions to (7).*

Proof. 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 $|f(X, \beta)| > \eta$. 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 by Lemma 2, 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 $W_0 \leq \overline{U}(X_0)$. Similarly, if $\underline{U}$ is the lower envelope of the set of solutions to (7), it is not possible to have $D_0 < 0$ and therefore it must be the case that $W_0 \geq \underline{U}(X_0)$. $\square$

The proof of Theorem 2 follows directly from Lemma 8, and the fact that at any state $X \in \mathcal{X}$, it is possible for the long-run player to achieve any payoff in the convex hull of the set of Markov equilibrium payoffs at state X by using randomization.

Proof of Corollary 1. Existence of a Markov equilibrium follows from Theorem 1. When μ is independent of a, the sequential rationality condition (6) in a Markov equilibrium collapses to maximizing the static flow payoff, and the long-run player plays the unique static Nash action profile $S^*(X, 0)$ in each state. Therefore, the measure over the state is independent of the solution to (7). Any solution to (7) must satisfy

$$U(X_t) = E_t \left[r \int_t^\infty e^{-rs} g^*(X_s, 0) dt \right]. \tag{29}$$

Given the the RHS of (7) is independent of U, (7) must have a unique solution and there is a unique Markov equilibrium. By Theorem 2, this is also the unique PPE. The solution to (7) evaluated at the current state X_t analytically characterizes the RHS of (29).

A.4 Proof of Theorems 3 and 4

I prove Theorems 3 and 4 simultaneously. 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 linear growth (bounded) solution.
3. When there is a unique solution, then there is a unique PPE.

As before, let $\psi(X,z) \equiv g^*(X,z) + z\mu^*(X,z)/r$. All intermediate theorems and lemmas assume Assumptions 1-3 and Assumption 4 or 4′.

Step 1: Boundary Conditions. Theorems 6, 7 and 8 characterize the boundary conditions for (i) unbounded $\mathcal{X}$ and g, (ii) unbounded $\mathcal{X}$ and bounded g and (iii) bounded $\mathcal{X}$, respectively, to establish step 1.

Step 1a: Boundary Conditions for Unbounded $\mathcal{X}$ and g (Theorem 3).

Theorem 6. *Suppose $\mathcal{X} = \mathbb{R}$ and g is unbounded and assume Assumptions 1-4. Then any solution U of (7) with linear growth satisfies*

$$\begin{aligned}
&\lim_{X\to p} U(X) - y^L(X) = g_2(z_p) + z_p\mu_2(z_p)/r \\
&\lim_{X\to p} U'(X) = z_p \\
&\lim_{X\to p} XU''(X) = 0
\end{aligned}$$

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

$$\begin{aligned}
y^L(X) &= -f(X)\int \frac{rg_1(X)}{f(X)\mu_1(X)}dX \\
f(X) &= \exp\left(\int \frac{r}{\mu_1(X)}dX\right) \\
z_p &= \lim_{X\to p} \frac{rg_1(X)/X}{r - \mu_1(X)/X}.
\end{aligned}$$

The proof proceeds by a series of lemmas. Define $\overline{\psi}(X,z) \equiv \psi(X,z)/X$, $\overline{U}(X) \equiv U(X)/X$, and ψ' and $\overline{\psi}'$ refer to the partial derivative with respect to X where, given Assumption 4, $\psi'(X,z) = g_1'(X) + z\mu_1'(X)/r$.

Lemma 9. *Given $p \in \{-\infty, \infty\}$, there exists a $\psi_p : \mathbb{R} \to \mathbb{R}$ such that for all $z \in \mathbb{R}$, $\psi_p(z) \equiv \lim_{X\to p} \overline{\psi}(X,z) = \lim_{X\to p} \psi'(X,z)$ and $\overline{\psi}(\cdot,z)$ is monotone for large $|X|$.*

Proof. Let $p \in \{-\infty, \infty\}$ and fix $z \in \mathbb{R}$. Given that g^* and μ^* are Lipschitz continuous, there exists $M, \delta > 0$ such that $|\overline{\psi}(X,z)| \le M(1 + |z| + z^2)$ for all $|X| > \delta$. Therefore, for fixed z, $\overline{\psi}(\cdot,z)$ is bounded in $|X|$ for $|X| > \delta$. By Assumption 4, $\psi'(\cdot,z)$ is monotone for large X, and therefore, $\overline{\psi}(\cdot,z)$ is monotone for large X (Lemma 35). Therefore, by Lemma 33, $\psi_p(z) \equiv \lim_{X\to p} \overline{\psi}(X,z)$ exists and $\lim_{X\to p} X\overline{\psi}'(X,z) = 0$. By Lemma 34, $\lim_{X\to p} \psi'(X,z) = \psi_p(z)$. ∎

Lemma 10. *Suppose U is a solution of (7) with linear growth. Then $\exists \delta > 0$ such that for $|X| > \delta$, U' and $\overline{U}$ are monotone and either both increasing or both decreasing. For $p \in \{-\infty, \infty\}$, there exists a $U'_p \in \mathbb{R}$ such that $\lim_{X\to p} \overline{U}(X) = \lim_{X\to p} U'(X) = U'_p$.*

Proof. Let $p \in \{-\infty, \infty\}$. Suppose U' is not monotone for large $|X|$. Then by the continuity of U', for any $\delta > 0$, there exists a z and a $|X_n|, |X_m| > \delta$ such that $U'(X_n) = z$ and $U''(X_n) \leq 0$ and $U'(X_m) = z$ and $U''(X_m) \geq 0$. From (7), this implies $U(X_n) \leq \psi(X_n, z)$ and $\psi(X_m, z) \leq U(X_m)$. Thus, the oscillation of $\psi'(\cdot, z)$ is at least as large as the oscillation of U'. By Assumption 4, $\psi'(\cdot, z)$ is monotone for large X, so U' cannot be non-monotone for large X. This is a contradiction; therefore, U' is monotone for large X.

By Lemma 35, if U' is monotone for large $|X|$, then $\overline{U}$ is monotone for large $|X|$. Given that U has linear growth, $\overline{U}$ is bounded. Therefore, $\lim_{X\to p} \overline{U}(X)$ exists and by Lemma 33, $\lim_{X\to p} X\overline{U}'(X) = 0$. By Lemma 34, $\lim_{X\to p} U'(X) = \lim_{X\to p} \overline{U}(X)$. Let U'_p denote this limit. □

Lemma 11. *Suppose U is a solution of (7) with linear growth. Then $\lim_{X\to p} \overline{\psi}(X, U'(X)) = \psi_p(U'_p)$ for $p \in \{-\infty, \infty\}$, where $U'_p \equiv \lim_{X\to p} U'(X)$.*

Proof. Let $p \in \{-\infty, \infty\}$ and U be a solution of (7) with linear growth. Given μ^* and g^* are Lipschitz continuous, there exists a $\delta, M_1, M_2, M_3, c > 0$ such that for $|X| > \delta$,

$$|\psi(X, z_1) - \psi(X, z_2)| \leq M_1|z_1 - z_2| + M_2|z_1||z_1 - z_2| + M_3|z_1 - z_2|(|X| + |z_2|)$$

From Lemma 10, there exists a $U'_p \in \mathbb{R}$ such that $\lim_{X\to p} U'(X) = U'_p$. Therefore, for $|X| > \delta$,

$$\begin{aligned} & \lim_{X\to p} |\overline{\psi}(X, U'(X)) - \overline{\psi}(X, U'_p)| \\ = \quad & \lim_{X\to p} \frac{|\psi(X, U'(X)) - \psi(X, U'_p)|}{|X|} \\ \leq \quad & \lim_{X\to p} \frac{M_1|U'(X) - U'_p| + M_2|U'(X)||U'(X) - U'_p| + M_3|U'(X) - U'_p|(|X| + |U'_p|)}{|X|} \\ = \quad & 0 \end{aligned}$$

Therefore, $\lim_{X\to p} \overline{\psi}(X, U'(X)) = \lim_{X\to p} \overline{\psi}(X, U'_p)$. From Lemma 9, $\lim_{X\to p} \overline{\psi}(X, U'_p) = \psi_p(U'_p)$. Therefore, $\lim_{X\to p} \overline{\psi}(X, U'(X)) = \psi_p(U'_p)$. □

Lemma 12. *Suppose $f : \mathbb{R} \to \mathbb{R}$ has linear growth. Then any solution U of (7) 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 $p \in \{-\infty, \infty\}$.

Proof. Let $p \in \{-\infty, \infty\}$. Suppose f has linear growth and $\lim_{X\to p}\inf |f(X)|U''(X) > 0$. There exists an $\delta_1, M > 0$ 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 $\delta_2, \varepsilon > 0$ 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 linear growth of U. Therefore $\lim_{X\to p}\inf |f(X)|U''(X) \leq 0$. The proof is analogous for $\lim_{X\to p}\sup |f(X)|U''(X) \geq 0$. ☐

Lemma 13. *Suppose* $f : \mathbb{R} \to \mathbb{R}$ *has linear growth. Then any solution* U *of* (7) *with linear growth satisfies* $\lim_{X\to p} f(X)U''(X) = 0$ *for* $p \in \{-\infty, \infty\}$.

Proof. Let $p \in \{-\infty, \infty\}$. Suppose that $\lim_{X\to p}\sup |X|U''(X) > 0$. By Lemma 10, there exists a $\delta > 0$ such that for $|X| > \delta$, U' is monotone. 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, given Lemma 12. 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 has linear growth. Then there exists an $\delta_1, M > 0$ 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 also implies that $\lim_{X\to p} U''(X) = 0$. ☐

Lemma 14. *Suppose* U *is a solution of* (7) *with linear growth. Then* U'_p *is a fixed point of* ψ_p, $U'_p = \psi_p(U'_p)$ *for* $p \in \{-\infty, \infty\}$.

Proof. Let $p \in \{-\infty, \infty\}$ and U be a solution of (7) with linear growth. Given that σ^* is Lipschitz continuous, there exists an $\delta, M > 0$ such that for $|X| > \delta$,

$$\sigma^*(X,z)^2/|X| \leq M|(X,z)|^2/|X| = M(|X| + z^2/|X|).$$

Then

$$\lim_{X\to p} \frac{\sigma^*(X,U'(X))^2}{|X|}|U''(X)| \leq \lim_{X\to p} M|XU''(X)| + M|U'(X)^2|U''(X)/X| = 0.$$

where the equality follows from Lemmas 10 and 13. Plugging this into (7) yields $\lim_{X\to p}\overline{U}(X) - \overline{\psi}(X, U'(X)) = 0$ and therefore, by Lemma 10, $\lim_{X\to p}\overline{\psi}(X,U'(X)) = U'_p$. From Lemma 11, $\lim_{X\to p}\overline{\psi}(X,U'(X)) = \psi_p(U'_p)$. Combining these equations, $U'_p = \psi_p(U'_p)$, and U'_p must be a fixed point of ψ_p. ☐

Lemma 15. *Suppose U is a solution of* (7) *with linear growth. Then for $p \in \{-\infty, \infty\}$, the uni que fixed point of ψ_p is $r\overline{g}_p/(r-\overline{\mu}_p)$, where $\overline{g}_p \equiv \lim_{X\to p} g^*(X,z)/X$ and $\overline{\mu}_p \equiv \lim_{X\to p} \mu^*(X,z)/X$ for all $z \in \mathbb{R}$, and $\overline{\mu}_p < r$. Therefore, $U'_p = r\overline{g}_p/(r-\overline{\mu}_p)$.*

Proof. Let $p \in \{-\infty, \infty\}$. Given $\overline{\psi}(\cdot, z)$ is monotone for large $|X|$ and μ^* and g^* are additively separable in (X,z), $g^*(X,z)/X$ and $\mu^*(X,z)/X$ are monotone for large $|X|$. Since g^* and μ^* are Lipschitz continuous, $g^*(X,z)/X$ and $\mu^*(X,z)/X$ are bounded. Therefore, the limits $\overline{g}_p \equiv \lim_{X\to p} g^*(X,z)/X$ and $\overline{\mu}_p \equiv \lim_{X\to p} \mu^*(X,z)/X$ exist. By additive separability, these limits are independent of z. Therefore,

$$\psi_p(z) = \lim_{X\to p} \frac{g^*(X,z) + z\mu^*(X,z)/r}{X} = \overline{g}_p + z\overline{\mu}_p/r.$$

At a fixed point z^*, $\psi_p(z^*) = z^*$. The unique fixed point of ψ_p is $z^* = \frac{r\overline{g}_p}{r-\overline{\mu}_p}$.

By Lemma 14, for any linear growth solution U of (7) and $p \in \{-\infty, \infty\}$, U'_p is a fixed point of ψ_p. Thus, the final statement follows. □

Lemma 16. *Suppose y is a solution to the ODE*

$$y'(x) - (r/\mu_1(x))y(x) = 0 \tag{30}$$

with linear growth. Then for $p \in \{-\infty, \infty\}$, $\lim_{x\to p} y(x) = 0$.

Proof. The general solution to (30) is

$$y(x) = c\exp\left(\int \frac{r}{\mu_1(x)}dx\right) \tag{31}$$

where $c \in \mathbb{R}$ is a constant. Trivially, there always exists a solution with linear growth because $y(x) = 0$ is a solution. Consider $p = \infty$. By Assumption 2, μ_1 has linear growth with rate slower than r. By Assumption 4, μ_1 is monotone for large x. Therefore, there exists a $\delta > 0$ such that for $x > \delta$, either (i) there exists a $k \in (0,r)$ such that $\mu_1(x) \in (0, kx]$ or (ii) there exists a $k > 0$ such that $\mu_1(x) \in [-kx, 0)$.

Case (i): Suppose there exists a $k \in (0,r)$ and $\delta > 0$ such that for $x > \delta$, $\mu_1(x) \in (0, kx]$. Then $\frac{1}{\mu_1(x)} \geq \frac{1}{kx}$. But $\exp\left(\int \frac{r}{kx}dx\right) = \exp\left(\frac{r}{k}\ln x\right) = x^{r/k}$ is not in $O(x)$ since $r/k > 1$. Therefore, $\exp\left(\int \frac{r}{\mu_1(x)}dx\right)$ is not in $O(x)$. Therefore, any solution to (31) that has linear growth must have $c = 0$. The unique solution with linear growth is $y(x) = 0$, which trivially satisfies $\lim_{x\to\infty} y(x) = 0$.

Case (ii): Suppose there exists a $k, \delta > 0$ such that for $x > \delta$, $\mu_1(x) \in [-kx, 0)$. Then $\frac{1}{\mu_1(x)} \leq -\frac{1}{kx}$. But $\exp\left(\int -\frac{r}{kx} dx\right) = \exp\left(-\frac{r}{k}\ln x\right) = x^{-r/k}$ and $\lim_{x\to\infty} x^{-r/k} \to 0$. Therefore, $\lim_{x\to\infty} \exp\left(\int \frac{r}{\mu_1(x)} dx\right) = 0$. Therefore, for all c, $\lim_{x\to\infty} y(x) = 0$ and any solution to (31) satisfies this property.

The case for $p = -\infty$ is analogous. □

Lemma 17. *Suppose U and V are solutions of (7) with linear growth. Then for $p \in \{-\infty, \infty\}$, $\lim_{X\to p} U(X) - V(X) = 0$.*

Proof. Let $p \in \{-\infty, \infty\}$, U and V be solutions of (7) with linear growth. Then

$$\begin{aligned}
\Rightarrow \lim_{X\to p} \quad & U(X) - g_1(X) - U'(X)\mu_1(X)/r - g_2(U'(X)) - U'(X)\mu_2(U'(X))/r \\
& -U''(X)|\sigma^*(X, U'(X))|^2/2r = 0 \\
\Rightarrow \lim_{X\to p} \quad & U(X) - g_1(X) - U'(X)\mu_1(X)/r = g_2(z_p) + z_p\mu_2(z_p)/r \qquad (32)
\end{aligned}$$

where the first line follows from (7) and the additive separability of g^* and μ^*, and the second line follows from the Lipschitz continuity of $(\sigma^*)^2$, $\lim_{X\to p} U'(X) = z_p$ and the Lipschitz continuity of g_2 and μ_2. By Lemma 15, $\lim_{X\to p} U'(X) = \lim_{X\to p} V'(X) = z_p$.

Define $D = U - V$. Then $D' = U' - V'$, D has linear growth since U and V have linear growth and

$$\begin{aligned}
&\lim_{X\to p} D(X) - \mu_1(X)D'(X)/r = 0 \\
&\lim_{X\to p} D'(X) = 0
\end{aligned}$$

where the first line follows from (32). Therefore, there exists a solution y to (30) with linear growth such that $\lim_{X\to p} D(X) - y(X) = 0$. By Lemma 16, $\lim_{X\to p} y(X) = 0$ for any solution y with linear growth. Therefore, $\lim_{X\to p} D(X) = 0$ and any two solutions U and V with linear growth have the same boundary conditions, $\lim_{X\to p} U(X) - V(X) = 0$. □

Lemma 18. *Suppose y is a solution to the ODE*

$$y(x) - g_1(x) - \mu_1(x)y'(x)/r = 0 \qquad (33)$$

with linear growth. Then for $p \in \{-\infty, \infty\}$, $\lim_{x \to p} y(x) - y^L(x) = 0$, *where*

$$y^L(x) \equiv -\phi(x) \int \left(\frac{1}{\phi(x)} \right) \frac{r g_1(x)}{\mu_1(x)} dx \tag{34}$$

is a solution with linear growth and $\phi(x) \equiv \exp\left(\int r/\mu_1(x) dx \right)$.

Proof. The general solution to (33) is

$$y(x) = -\phi(x) \int \left(\frac{1}{\phi(x)} \right) \frac{r g_1(x)}{\mu_1(x)} dx - \phi(x) c \tag{35}$$

where ϕ is as defined above and $c \in \mathbb{R}$ is a constant. Consider $p = \infty$. By Assumption 2, μ_1 has linear growth with rate slower than r. By Assumption 4, μ_1 is monotone for large x. Therefore, there exists a $\delta > 0$ such that for $x > \delta$, either (i) there exists a $k \in (0, r)$ such that $\mu_1(x) \in (0, kx]$ or (ii) there exists a $k > 0$ such that $\mu_1(x) \in [-kx, 0)$.

Case (i): Suppose there exists a $k \in (0, r)$ and $\delta > 0$ such that for $x > \delta$, $\mu_1(x) \in (0, kx]$. Then $\frac{1}{\mu_1(x)} \geq \frac{1}{kx}$. But $\exp\left(\int \frac{r}{kx} dx \right) = \exp\left(\frac{r}{k} \ln x \right) = x^{r/k}$ is not in $O(x)$ since $r/k > 1$. Therefore, $\exp\left(\int \frac{r}{\mu_1(x)} dx \right)$ is not in $O(x)$ and ϕ doesn't have linear growth. Therefore, any solution to (35) that has linear growth must have $c = 0$. The unique solution with linear growth is (34), which trivially satisfies $\lim_{x \to \infty} y(x) - y^L(x) = 0$.

Case (ii): Suppose there exists a $k, \delta > 0$ such that for $x > \delta$, $\mu_1(x) \in [-kx, 0)$. Then $\frac{1}{\mu_1(x)} \leq -\frac{1}{kx}$. But $\exp\left(\int -\frac{r}{kx} dx \right) = \exp\left(-\frac{r}{k} \ln x \right) = x^{-r/k}$ and $\lim_{x \to \infty} x^{-r/k} = 0$. Therefore, $\lim_{x \to \infty} \phi(x) = \lim_{x \to \infty} \exp\left(\int \frac{r}{\mu_1(x)} dx \right) = 0$ and c does not affect the limit properties of a solution y. Therefore, for all c, $\lim_{x \to \infty} y(x) - y^L(x) = 0$ and any solution to (35) satisfies this property.

The case for $p = -\infty$ is analogous. □

Lemma 19. *Suppose* U *is a solution of* (7) *with linear growth. Then for* $p \in \{-\infty, \infty\}$, $\lim_{X \to p} U(X) - y^L(X) = g_2(z_p) + z_p \mu_2(z_p)/r$, *where* y_L *is defined by* (34).

Proof. Let $p \in \{-\infty, \infty\}$ and U be a solution of (7) with linear growth. Then

$$\begin{aligned} \lim_{X \to p} \quad & U(X) - g_1(X) - U'(X)\mu_1(X)/r - U''(X)|\sigma^*(X, U'(X))|^2/2r \\ & -g_2(U'(X)) - U'(X)\mu_2(U'(X))/r = 0 \\ \Rightarrow \lim_{X \to p} \quad & U(X) - g_1(X) - U'(X)\mu_1(X)/r = g_2(z_p) + z_p \mu_2(z_p)/r \end{aligned}$$

where the first line follows from (7) and the additive separability of g^* and μ^*, and the second line follows from the Lipschitz continuity of $(\sigma^*)^2$, $\lim_{X\to p} U'(X) = z_p$ and the Lipschitz continuity of g_2 and μ_2. Therefore, there exists a solution y to (33) with linear growth such that $\lim_{X\to p} U(X) - y(X) = g_2(z_p) + z_p\mu_2(z_p)/r$. By Lemma 18, $\lim_{X\to p} y(X) - y^L(X) = 0$. Therefore, $\lim_{X\to p} U(X) - y^L(X) = g_2(z_p) + z_p\mu_2(z_p)/r$ which establishes the boundary condition for U. □

Step 1b: Boundary Conditions for Unbounded $\mathcal{X}$ and Bounded g (Theorem 4).

Theorem 7. *Suppose $\mathcal{X} = \mathbb{R}$ and g is bounded and assume Assumptions 1-3 and Assumption 4′.a. Then any bounded solution U of (7) satisfies*

$$\lim_{X\to p} U(X) = g_p$$
$$\lim_{X\to p} XU'(X) = 0$$
$$\lim_{X\to p} X^2U''(X) = 0$$

for $p \in \{-\infty, \infty\}$, where $g_p \equiv \lim_{X\to p} g^*(X, 0)$.

The proof proceeds by a series of lemmas. Note g_p exists given g bounded and $g^*(\cdot, 0)$ monotone for large $|X|$.

Lemma 20. *If U is a bounded solution of (7), then there exists a $\delta > 0$ such that for $|X| > \delta$, U is monotone and for $p \in \{-\infty, \infty\}$, there exists a $U_p \in \mathbb{R}$ such that $\lim_{X\to p} U(X) = U_p$.*

Proof. Let $p \in \{-\infty, \infty\}$. Suppose U is not monotone for large X. Then for all $\delta > 0$, there exists a $|X_n| > \delta$ that corresponds to a local maximum 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 minimum 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 equilibrium is played at any X such that $U'(X) = 0$, yielding flow payoff $g^*(X, 0)$. From (7), 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 $g^*(\cdot, 0)$ is at least as large as the oscillation of U. This is a contradiction, as $g^*(\cdot, 0)$ is monotone for large $|X|$. Thus, there exists a δ such that for $|X| > \delta$, U is monotone. The existence of $\lim_{X\to p} U(X)$ follows from U is bounded and monotone for large $|X|$. □

Lemma 21. *Suppose $f : \mathbb{R} \to \mathbb{R}$ has linear growth. Then any bounded solution U of* (7) *satisfies*

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

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

Proof. Let $p \in \{-\infty,\infty\}$. Suppose f has linear growth and $\lim_{X\to p}\inf|f(X)|U'(X) > 0$. Then there exists an $\delta_1, M > 0$ 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 $\delta_2, \varepsilon > 0$ 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 $|X| \to \infty$. This violates the boundedness of U. Therefore $\lim_{X\to p}\inf|f(X)|U'(X) \leq 0$. The proof is analogous for $\lim_{X\to p}\sup|f(X)|U'(X) \geq 0$.

Suppose f has linear growth and $\lim_{X\to\infty}\inf f(X)^2U''(X) > 0$. Then there exists a $M, \delta_1 > 0$ such that when $|X| > \delta_1, |f(X)| \leq M|X|$ and therefore, $f(X)^2 \leq M^2X^2$. There also exists a $\delta_2, \varepsilon > 0$ such that when $|X| > \delta_2$, $f(X)^2U''(X) > \varepsilon$. Take $\delta = \max\{\delta_1, \delta_2\}$. Then for $|X| > \delta$, $|U''(X)| > \frac{\varepsilon}{f(X)^2} > \frac{\varepsilon}{M^2X^2}$. The antiderivative of $\frac{\varepsilon}{M^2X^2}$ is $\frac{-\varepsilon}{M^2}\ln|X|$ which converges to $-\infty$ as $|X| \to \infty$. This violates the boundedness of U. Therefore $\lim_{X\to p}\inf f(X)^2U''(X) \leq 0$. The proof is analogous for $\lim_{X\to p}\sup f(X)^2U''(X) \geq 0$. □

Lemma 22. *Suppose $f : \mathbb{R} \to \mathbb{R}$ has linear growth. Then for $p \in \{-\infty,\infty\}$, any bounded solution U of* (7) *satisfies* $\lim_{X\to p} f(X)U'(X) = 0$.

Proof. Let $p \in \{-\infty,\infty\}$. Suppose that $\lim_{X\to p}\sup|X|U'(X) > 0$. By Lemma 20, there exists a $\delta > 0$ such that U is monotone 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 has linear growth. Then there exists an $M, \delta_1 > 0$ such that when $|X| > \delta_1, |f(X)| \leq M|X|$. Thus, for $|X| > \delta_1$, $|f(X)U'(X)| \leq M|XU'(X)| \to 0$. Therefore, $\lim_{X\to p} f(X)U'(X) = 0$. This also implies that $\lim_{X\to p} U'(X) = 0$. □

Lemma 23. *Let U be a bounded solution of* (7)*. Then for $p \in \{-\infty,\infty\}$,* $\lim_{X\to p} U(X) = g_p$.

Proof. Let $p \in \{-\infty,\infty\}$. Suppose $U_p < g_p$. Given μ^* is Lipschitz continuous, there exists a $\delta, M > 0$ such that for $|X| > \delta$,

$$|\mu^*(x,z)| \leq M|(x,z)| \leq M(|x| + |z|).$$

Therefore,

$$\lim_{X\to\infty} |U'(X)\mu^*(X, U'(X))| \leq \lim_{x\to\infty} |U'(X)|M(|X| + |U'(X)|) = 0$$

where the equality follows from Lemma 22. Similarly,

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

since g^* is Lipschitz continuous. Plugging these limits into (7),

$$\begin{aligned}\limsup_{X\to p} \frac{1}{2}\sigma^*(X, U'(X))^2 U''(X) &= \limsup_{X\to p} (rU(X) - rg^*(X, U'(X)) - \mu^*(X, U'(X))U'(X)) \\ &= r(U_p - g_p) < 0.\end{aligned}$$

But by Lemma 21, $\limsup_{X\to p} \sigma^*(X, U'(X))^2 U''(X) > 0$ since σ^* is Lipschitz continuous, a contradiction. Thus, $U_p \geq g_p$. A similar contradiction holds for $U_p > g_p$. Therefore, $U_p = g_p$. □

Lemma 24. *Any bounded solution U of (7) satisfies $\lim_{X\to p} \sigma^*(X, U'(X))^2 U''(X) = 0$ for $p \in \{-\infty, \infty\}$.*

Proof. Let $p \in \{-\infty, \infty\}$. By Lemmas 22 and 23 and the squeeze theorem,

$$\lim_{X\to p} \frac{1}{2}|\sigma^*(X, U'(X))^2 U''(X)| = \lim_{X\to p} |rU(X) - rg^*(X, U'(X)) - \mu^*(X, U'(X))U'(X)| = 0.$$

If $\lim_{X\to p} \sigma^*(X, U'(X))^2 > 0$, this implies that $\lim_{X\to p} U''(X) = 0$. □

Step 1c: Boundary Conditions for Bounded $\mathcal{X}$ (Theorem 4).

Theorem 8. *Suppose $\mathcal{X}$ is compact and assume Assumptions 1-3 and Assumption 4'. Then* *any bounded solution U of (7) satisfies*

$$\lim_{X \to p} U(X) = g^*(p, 0)$$
$$\lim_{X \to p} (X - p)U'(X) = 0$$
$$\lim_{X \to p} (X - p)^2 U''(X) = 0.$$

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

The proof proceeds by a series of lemmas.

Lemma 25. *Any bounded solution U of (7) has bounded variation.*

Proof. 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 (6), a static Nash equilibrium is played at any X such that $U'(X) = 0$, yielding flow payoff $g^*(X, 0)$. From (7), 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, $g^*(\cdot, 0)$ has unbounded variation. This is a contradiction, since $g^*(\cdot, 0)$ is Lipschitz continuous. □

Lemma 26. *Suppose $f : \mathcal{X} \to \mathbb{R}$ is Lipschitz continuous with $f(\overline{X}) = f(\underline{X}) = 0$. Then any bounded solution U of (7) satisfies*

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

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

Proof. Let $p \in \{-\underline{X}, \overline{X}\}$ and $f : \mathcal{X} \to \mathbb{R}$ be Lipschitz continuous with $f(p) = 0$. Then f is $O\left(p - X\right)$, so there exists an $M, \delta_1 > 0$ such that when $|p - X| < \delta_1, |f(X)| \leq M|p - X|$.

Suppose $\lim_{X \to p} \inf |f(X)|U'(X) > 0$. Then there exists a $\delta_2, \varepsilon > 0$ 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 $\lim_{X \to p} \inf |f(X)|U'(X) \leq 0$. The proof is analogous for $\lim_{X \to p} \sup |f(X)|U'(X) \geq 0$.

Suppose $\lim_{X \to p} \inf f(X)^2 U''(X) > 0$. There exists a $\delta_3, \varepsilon_2 > 0$ such that when $|p - X| < \delta_3$, $f(X)^2 U''(X) > \varepsilon_2$. Take $\delta = \min\{\delta_1, \delta_3\}$. Then for $|p - X| < \delta$, $U''(X) > \frac{\varepsilon_2}{f(X)^2} >$

$\frac{\varepsilon_2}{M^2(p-X)^2}$. The second antiderivative of $\frac{\varepsilon_2}{M^2(p-X)^2}$ is $\frac{-\varepsilon_2}{M^2}\ln|p-X|$ which converges to $-\infty$ as $X \to p$. This violates the boundedness of U. Therefore $\lim_{X\to\infty}\inf f(X)^2U''(X) \leq 0$. The proof is analogous for $\lim_{X\to p}\sup f(X)^2U''(X) \geq 0$. □

Lemma 27. *Suppose $f : \mathcal{X} \to \mathbb{R}$ is Lipschitz continuous, with $f(\overline{X}) = f(\underline{X}) = 0$. Then any bounded solution U of (7) satisfies $\lim_{X\to p} f(X)U'(X) = 0$ for $p \in \{\underline{X}, \overline{X}\}$.*

Proof. Let $p \in \{-\underline{X}, \overline{X}\}$. Suppose that $\limsup_{X\to p}|p-X|U'(X) > 0$. By Lemma 26, $\lim_{X\to p}\inf|p-X|U'(X) \leq 0$. Then there exist constants $k, K > 0$ such that $|p-X|U'(X)$ crosses the interval (k, K) infinitely many times as X approaches p. Additionally, there exists an $L > 0$ such that

$$\begin{aligned}
|U''(X)| &= \left|\frac{2r(U(X)-g^*(X,U'(X))) - 2\mu^*(X,U'(X))U'(X)}{\sigma^*(X,U'(X))^2}\right| \leq \left|\frac{L_1 - L_2|p-X|U'(X)}{(p-X)^2}\right| \\
&\leq \left|\frac{L_1 - L_2k}{(p-X)^2}\right| = \frac{L}{(p-X)^2}
\end{aligned}$$

This implies that

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

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

$$U'(X) \geq \frac{|((p-X)U'(X))'|}{(1+\frac{L}{k})}$$

Therefore, the total variation of U is at least $\frac{K-k}{(1+\frac{L}{k})}$ 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, $\lim_{X\to p}\inf(p-X)U'(X) = 0$. Therefore $\lim_{X\to p}(p-X)U'(X) = 0$.

Suppose f is $O(p-X)$. Then there exists an $M, \delta > 0$ such that for $|p-X| < \delta$, $|f(X)| \leq M|(p-X)| \to 0$. Thus for $|p-X| < \delta$, $|f(X)U'(X)| \leq M|(p-X)U'(X)| \to 0$. Therefore $\lim_{X\to p} f(X)U'(X) = 0$. □

Lemma 28. *Let U be a bounded solution of* (7)*. Then for $p \in \{\underline{X}, \overline{X}\}$, $\lim_{X \to p} U(X) = g^*(p, 0)$.*

Proof. Let $p \in \{-\underline{X}, \overline{X}\}$. Given that U is continuous, bounded and has bounded variation, $U_p \equiv \lim_{X \to p} U(X)$ exists. Suppose $U_p < g^*(p, 0)$. By Lemma 27, the Lipschitz continuity of μ^* and the assumption that $\mu(a, b, p) = 0$ for all $(a, b) \in A \times B(p)$,

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

If X is an absorbing state, then $S^*(X, z) = S^*(X, 0)$ for all $z \in \mathbb{R}$. Therefore, $g^*(X, z) = g^*(X, 0)$ for all $z \in \mathbb{R}$. By the Lipschitz continuity of g^*,

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

Plugging these limits into (7),

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

But by Lemma 26, $\limsup_{X \to p} \sigma^*(X, U'(X))^2 U''(X) > 0$ since σ^* is Lipschitz continuous and $\sigma(b, p) = 0$ for all $b \in B(p)$. This is a contradiction. Thus, $U_p \geq g^*(p, 0)$. A similar contradiction holds for $U_p > g^*(p, 0)$. Therefore, $U_p = g^*(p, 0)$. □

Lemma 29. *Any bounded solution U of* (7) *satisfies*

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

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

Proof. Let $p \in \{-\underline{X}, \overline{X}\}$. Applying Lemmas 27 and 28 and the squeeze theorem,

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

□

Step 2: Uniqueness of Solution to Optimality Equation (Theorems 3 and 4).

Lemma 30. *If U and V are two bounded solutions of (7) such that $U(X_0) \leq V(X_0)$ 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 $X < X_0$.*

Proof. This follows directly from Lemma C.7 in Faingold and Sannikov (2011), defining $X_1 = \inf\left\{X \in [X_0, \overline{X}) : U'(X) \geq V'(X)\right\}$. □

Lemma 31. *There exists a unique linear growth (bounded) solution U to (7).*

Proof. Suppose U and V are both solutions to (7). Suppose $V(X) > U(X)$ for some $X \in \mathcal{X}$. 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 $p \in \left\{\underline{X}, \overline{X}\right\}$. Then $U'(X^*) = V'(X^*)$ and $V(X^*) > U(X^*)$. By Lemma 30, $V'(X) > U'(X)$ for all $X > X^*$, and $V(X) - U(X)$ is strictly increasing, a contradiction since X^* maximizes $U(X) - V(X)$. □

Step 3: Uniqueness of PPE (Theorems 3 and 4).

Lemma 32. *There exists a unique PPE.*

Proof. By Lemma 31, there is a unique linear growth (bounded) solution to (7). It is obvious to see that this implies 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 $(W_t)_{t\geq 0}$, $W_t = U(X_t)$ for all t. Therefore, the volatility of the two continuation values are equal, otherwise they both cannot be equal to $U(X_t)$. Given equal volatilities, actions are uniquely specified by $S^*(X, U'(X)/r)$. □

A.5 Proofs from Section 5

Proof of Proposition 1.

Proof. Let $n_U(I)$ denote the number of strict interior extrema of $U(X)$ on a closed interval $I \subset \mathcal{X}$. At a state X corresponding to an interior extremum, $U'(X) = 0$. From (7), if X is a minimum, $U(X) \geq g^*(X, 0)$, and if X is a maximum, $U(X) \leq g^*(X, 0)$.

Part 1: Suppose $g^*(X, 0)$ is constant on I and $n_U(I) > 1$. 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 on I. The same logic holds if X_1 is a maximum and X_2 is a minimum. Therefore, $n_U(I) \leq 1$.

Part 2: Suppose $g^*(X,0)$ is strictly increasing on I and $n_U(I) > 2$. 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 $g^*(X,0)$ is increasing on I. Therefore, it is not possible to have a maximum followed by a minimum. If X_1 is a maximum, $n_U(I) = 1$ and if X_1 is a minimum, $n_U(I) \leq 2$.

Suppose $U(X)$ is constant on I. There exists a constant u such that $U(X) = u$ for all $X \in I$. Then $U'(X) = 0$ and $U''(X) = 0$ for all $X \in I$. From (7), $U(X) = g^*(X,0)$ for all $X \in I$. Therefore, $g^*(X,0) = u$ for all $X \in I$ and $g^*(X,0)$ is constant on I, a contradiction.

The proof for when $g^*(X,0)$ is decreasing is analogous.

Part 3: Follows directly from $U(X) \geq g^*(X,0)$ at a minimum, $U(X) \leq g^*(X,0)$ at a maximum, and the Lipschitz continuity of g^*.

□

Proof of Proposition 2.

Proof. Let n_U denote the number of strict interior extrema of $U(X)$ on and n_g denote the number of strict interior extrema of $g^*(X,0)$ on $\mathcal{X}$. Suppose $\mathcal{X}$ is bounded.

Part 1: Suppose $g^*(X,0)$ is constant on $\mathcal{X}$. Then there exists a $c \in \mathbb{R}$ such that $g^*(X,0) = c$ for all X and $n_g = 0$. By Proposition 2.4, $n_U = 0$. By the boundary conditions, $U(\underline{X}) = c$ and $U(\overline{X}) = c$, which implies $U(\underline{X}) = U(\overline{X})$. Combined with $n_U = 0$, this implies that U is constant on $\mathcal{X}$.

I prove the contrapositive to establish that if $U(X)$ is constant on $\mathcal{X}$, then $g^*(X,0)$ is constant on $\mathcal{X}$. Suppose $g^*(X,0)$ is not constant on $\mathcal{X}$. Then there exists an interval $I \subset \mathcal{X}$ such that $g^*(X,0)$ is strictly increasing or decreasing on I. By Proposition 1.2, $U(X)$ is not constant on I. Therefore, $U(X)$ is not constant on $\mathcal{X}$.

Part 2: Suppose $g^*(X,0)$ is monotonically increasing on $\mathcal{X}$, but $U(X)$ is not monotonically increasing. Then $U'(X) < 0$ for some $X \in \mathcal{X}$. Since $n_g = 0$, by Proposition 2.4, $n_U = 0$. Therefore, it must be that $U'(X) \leq 0$ for all $X \in \mathcal{X}$. This implies that $U(\underline{X}) > U(\overline{X})$. By the boundary conditions, $U(\underline{X}) = g^*(\underline{X},0)$ and $U(\overline{X}) = g^*(\overline{X},0)$, and by monotonicity, $g^*(\underline{X},0) \leq g^*(\overline{X},0)$. This implies $U(\underline{X}) \leq U(\overline{X})$, a contradiction. Therefore, $U(X)$ is monotonically increasing. If $g^*(X,0)$ is strictly increasing on $\mathcal{X}$, then by Proposition 1.2, $U(X)$ is not constant on any $I \subset \mathcal{X}$, so $U(X)$ is also strictly increasing. The proof for $U(X)$ monotonically decreasing is analogous.

Part 3: Suppose $n_g = 1$, $g^*(X,0)$ has a unique interior maximum at X^* and $g^*(\underline{X},0) = g^*(\overline{X},0)$. By Proposition 2.4, $n_U \leq 1$. Consider the interval $I_1 = [\underline{X}, X^*]$. On this interval, $g^*(X,0)$ is monotonically increasing, and strictly so on some subinterval. By Proposition 1.2, $U(X)$ is not constant on I_1. Similarly, $g^*(X,0)$ is monotonically decreasing on $I_2 = [X^*, \overline{X}]$, and strictly so on some subinterval. Therefore, $U(X)$ is not constant on I_2. From the boundary conditions, $U(\underline{X}) = g^*(\underline{X},0)$ and $U(\overline{X}) = g^*(\overline{X},0)$. Therefore, $U(\underline{X}) = U(\overline{X})$. Since $U(X)$ is not constant, it must be that $n_U = 1$.

Suppose the extremum for $U(X)$ occurs at state $\tilde{X}$ and is a minimum. Then it must be that $U(\tilde{X}) \geq g^*(\tilde{X},0)$. But $U(\tilde{X}) < U(\underline{X}) = g^*(\underline{X},0)$ by the definition of a minimum. Therefore, $g^*(\tilde{X},0) < g^*(\underline{X},0) < g^*(X^*,0)$. But then $g^*(X,0)$ must have two interior extrema. This is a contradiction. Therefore, $U(X)$ is single-peaked with a maximum. The proof for $U(X)$ single-peaked with a minimum is analogous.

Part 4: Follows directly from $U(X) \geq g^*(X,0)$ at a minimum, $U(X) \leq g^*(X,0)$ at a maximum, $U(\underline{X}) = g^*(\underline{X},0)$, $U(\overline{X}) = g^*(\overline{X},0)$ and the Lipschitz continuity of g^*.

For $\mathcal{X}$ unbounded, replace $g^*(p,0)$ with $\lim_{X\to p} g^*(X,0)$ and $U(p)$ with $\lim_{X\to p} U(X)$ for $p \in \{\underline{X}, \overline{X}\}$. □

Proof of Proposition 3.

Proof. Suppose g is bounded. Then U is continuous and bounded on a closed set. Therefore, U either attains its maximum on $\mathcal{X}$, in which case $\overline{W} = U(X_H)$ for some X_H, or if $\mathcal{X}$ is unbounded, $\overline{W} = \limsup_{X\to X_H} U(X)$ for $X_H \in \{-\infty, \infty\}$. Suppose U attains a maximum at an interior state X_H. Then $U'(X_H) = 0$ and $U''(X_H) \leq 0$. From (7),

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

and therefore $\overline{W} \leq g^*(X_H,0)$. Suppose $\mathcal{X}$ is unbounded and U doesn't attain a maximum at an interior state. By Lemma 21, the local minima and maxima of U are bounded by the local minima and maxima of $g^*(\cdot,0)$. Therefore, for any sequence of local maxima $\{U_k\}$ converging to $\overline{W}$, there exists a sequence of states $\{X_k\}$ such that $U_k = U(X_k)$ and $\overline{W} \leq \lim_{k\to\infty} g^*(X_k,0) = \limsup_{X\to\infty} g^*(X,0)$. Suppose $\mathcal{X}$ is bounded 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 (7),

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

The proof for $\overline{W} = U(\underline{X})$ and the lower bound $\underline{W}$ are analogous. □

B Intermediate Results

Lemma 33. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is bounded, differentiable and there exists a $\delta > 0$ such that for $|x| > \delta$, f is monotone. Then for $p \in \{-\infty, \infty\}$, $\lim_{x\to p} xf'(x) = 0$.*

Proof. Suppose that $\lim_{x\to\infty} \inf |x|f'(x) > 0$. Then there exists a $\delta_2, \varepsilon > 0$ such that when $|x| > \delta_2$, $|x|f'(x) > \varepsilon$. Then for $|x| > \delta_2$, $f'(x) > \frac{\varepsilon}{|x|}$. 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 $\lim_{x\to p} \sup |x|f'(x) > 0$. For $|x| > \delta$, f is monotone and therefore $|x|f'(x)$ 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 $\lim_{x\to p} f'(x) = 0$. □

Lemma 34. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable and define $\overline{f}(x) = f(x)/x$. For $p \in \{-\infty, \infty\}$, if $\lim_{x\to p} \overline{f}(x) = c$ and $\lim_{x\to p} x\overline{f}'(x) = 0$, then $\lim_{x\to p} f'(x) = c$.*

Proof. Suppose $\lim_{x\to p} \overline{f}(x) = c$ and $\lim_{x\to p} x\overline{f}'(x) = 0$. Given $\overline{f}' = (f' - \overline{f})/x$, $\lim_{x\to p} f' = \lim_{x\to p}(x\overline{f}' + \overline{f}) = c$. □

Lemma 35. *If $f : \mathbb{R} \to \mathbb{R} \in \mathcal{C}^2$ and there exists a $\delta > 0$ such that for $|x| > \delta$, f' is monotone increasing (decreasing), then there exists a δ_2 such that for $|x| > \delta_2$, $\overline{f}(x) \equiv f(x)/x$ is monotone increasing (decreasing).*

Proof. Note that $\overline{f}' = (1/x)(f' - \overline{f})$, $\overline{f}'' = (1/x)(f'' - 2\overline{f}')$ and if $f' = \overline{f}$, then $\overline{f}' = 0$ with a maximum if $f'' \leq 0$ and a minimum if $f'' \geq 0$. Let f' be monotone 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 analogous when f' is monotone decreasing. □

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.

BOARD, S. AND M. MEYER-TER VEHN (2013): "Reputation for Quality," *Econometrica.*

BOHREN, J. A. (2016): "Using Persistence to Generate Incentives in a Dynamic Moral Hazard Problem," .

CISTERNAS, G. (2016): "Two-Sided Learning and Moral Hazard," *Princeton University.*

COSTER, C. AND P. HABETS (2006): *Two-Point Boundary Value Problems: Lower and Upper Solutions*, Elsivier.

CRIPPS, M., G. MAILATH, AND L. SAMUELSON (2004): "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72, 407–432.

DORASZELSKI, U. AND M. SATTERTHWAITE (2010): "Computable Markov-Perfect Industry Dynamics," *RAND Journal of Economics*, 41, 215–243.

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.

EKMEKCI, M. (2011): "Sustainable Reputations with Rating Systems," *Journal of Economic Theory*, 146, 479–503.

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.

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., D. Levine, and E. Maskin (1994): “The Folk Theorem with Imperfect Public Information,” *Econometrica*, 62, 997–1039.

Fudenberg, D. and D. K. Levine (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.

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*, Oxford University Press.

Milgrom, P. and J. Roberts (1982): “Predation, Reputation and Entry Deterrence,” *Journal of Economic Theory*, 27, 280–312.

Sannikov, Y. (2007): “Games With Imperfectly Observable Actions in Continuous Time,” *Econometrica*, 75, 1285–1329.

Sannikov, Y. and A. Skrzypacz (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.

Shapley, L. (1953): "Stochastic Games," *Proceedings of the National Academy of Sciences*, 39, 1095.

Skrzypacz, A. and Y. Sannikov (2007): "Impossibility of Collusion Under Imperfect Monitoring with Flexible Production," *American Economic Review*, 97, 1794–1823.

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

Strulovici, B. and M. Szydlowski (2015): "On the Smoothness of Value Functions and the Existence of Optimal Strategies," *Journal of Economic Theory*, 1016–1055.