

Stochastic Games in Continuous Time: Persistent Actions in Long-Run Relationships* (Job Market Paper)

J. Aislinn Bohren[†]
University of California, San Diego

November 20, 2011

Abstract

In many instances, a long-run player is influenced not only by considerations of its future but also by actions it has taken in the past. For example, the quality of a firm's product may depend on past as well as current investments. This paper studies a new class of stochastic games in which the actions of a long-run player have a persistent effect on payoffs. The setting is a continuous time game of imperfect monitoring between a long-run and a representative myopic player. The main result of this paper is to establish general conditions for the existence of Markovian equilibria and conditions for the uniqueness of a Markovian equilibrium in the class of all Perfect Public Equilibria. The existence proof is constructive and characterizes, for any discount rate, the explicit form of equilibrium payoffs, continuation values, and actions in Markovian equilibria as a solution to a second order ODE. Action persistence creates a channel to provide intertemporal incentives, and offers a new and different framework for thinking about the reputations of firms, governments, and other long-run agents.

*I thank David Miller, Paul Niehaus, Joel Sobel, Jeroen Swinkels, Joel Watson and especially Nageeb Ali for useful comments. I also thank participants of the UCSD theory seminar and graduate student research seminar for helpful feedback.

[†]Email: abohren@ucsd.edu; Web: <http://econ.ucsd.edu/~abohren>; Address: 9500 Gilman Drive, Department of Economics, La Jolla, CA 92093-0508.

1 Introduction

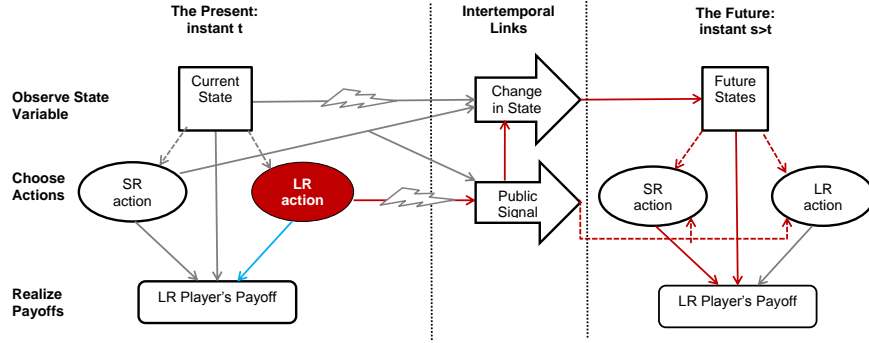
A rich and growing literature on repeated games and reputation has studied how the shadow of the future affects the present. Yet, in many instances, a long-run player is influenced not only by considerations of its future but also by decisions it has made in the past. For example, a firm's ability to make high quality products is a function of not only its effort today but also its past investments in developing technology and training its workforce. A government's ability to offer efficient and effective public services to its citizens depends on its past investments in improving its public services. A university cannot educate students through instantaneous effort alone, but needs to have made past costly investments in hiring faculty and building research infrastructure. In all of these settings, and many others, a long-run player is directly influenced by choices it has made in the past: past actions influence key characteristics of the long-run player's environment, such as the quality of a firm's product or the level of a policy instrument; in turn, these characteristics play a central role in determining current and future profitability. This paper studies a new class of stochastic games in which the actions of a long-run player have a persistent effect on payoffs, and studies how its incentives are shaped by its past and future.

In analyzing this class of stochastic games, the paper develops a new understanding of reputational dynamics. Since [Kreps, Milgrom, Roberts, and Wilson \(1982\)](#), the canonical framework has modeled a long-run player's reputation as the belief that others have that the firm is a behavioral type that takes a fixed action in each period. This framework has been very influential and led to a number of insights across the gamut of economics. Nevertheless, it is unclear across many settings that reputational incentives are driven exclusively by the possibility that players may be non-strategic and are absent when there is common knowledge that the long-run player is rationally motivated by standard incentives. In contrast, my environment returns to a different notion of reputation as an asset ([Klein and Leffler 1981](#)) in which a firm's reputation is shaped by its present and past actions. Persistent actions not only capture an intuitive notion of reputation as a type of capital, but also connect reputation to the aspects of a firm's choices that are empirically identifiable. This environment provides insights on important questions about the dynamics of reputation formation, including: when does a firm build its reputation and when does it allow it to decay; when do reputation effects persist in the long-run, and when are they temporary; how does behavior relate to underlying parameters such as the cost and depreciation rate of investment or the volatility of quality.

I study a continuous-time model with persistent actions and imperfect monitoring between a single long-run player and a continuum of small anonymous players. At each instant, each player chooses an action, which is not observed by others; instead, the long-run player's action generates a noisy *public signal*, and the long-run player observes the aggregate behavior of the short-run players. The stage game varies across time through its dependence on a state variable, whose evolution depends on the long run player's action through the public signal. This state variable determines the payoff structure of the associated stage game, and captures the current value of past investments by the long-run player.

When the long-run player chooses an action, it considers both the impact that this action has on its current payoff and its continuation value via the evolution of the state

Persistent Actions: A Link between the Present and the Future



Two Channels for Intertemporal Incentives:
 (1) Indirect: the state and public signal impact equilibrium play (dotted arrows)
 (2) Direct: actions and the state directly impact payoffs (solid arrows)

variable. For example, a firm may bear the cost of investment today, while reaping the rewards through higher sales and prices tomorrow. This link between current play and future outcomes creates intertemporal incentives for the long-run player. From [Faingold and Sannikov \(2011\)](#) and [Fudenberg and Levine \(2007\)](#), we know that in the absence of this state variable, intertemporal incentives fail: the long-run player cannot attain payoffs beyond those of its best stage game equilibrium. I am interested in determining when action persistence leads the firm to choose an action apart from that which maximizes its instantaneous payoff, and thus investigate whether persistent actions can be used to provide non-trivial intertemporal incentives in settings where those from standard repeated games fail. Figure 1 illustrates the link between current and future periods created by action persistence. The blue arrow indicates the effect of the long-run player’s action on payoffs today, while the red arrows illustrate the channels through which today’s action may impact future payoffs.

The key contributions of this paper are along three dimensions. The theoretical contribution is to establish general conditions for the existence of Markovian equilibria, and conditions for the uniqueness of a Markovian equilibrium in the class of all Perfect Public Equilibria. An explicit characterization of the form of equilibrium payoffs, continuation values and actions, for any discount rate, yields insights into the relationship between the structure of persistence and the decisions of the long-run player. An application of the existence and uniqueness results to a stochastic game without action persistence shows that the long-run player acts myopically in the unique Perfect Public Equilibria of this setting. Lastly, I use these results to describe several interesting properties relating equilibrium payoffs of the stochastic game to the structure of the underlying stage game.

The conceptual contribution of this paper is to illustrate that action persistence creates a channel for effective intertemporal incentive provision in a setting where this is not possible in the absence of such persistence. A stochastic game has two potential channels through which intertemporal incentives can be used to guide behavior. First, the stage game varies across time in response to players’ actions, and thus a long run player’s actions directly impact payoffs in future SR periods as well as the current period. Second, as in repeated games,

players can be rewarded or punished based on the public signal: actions today can affect, in equilibrium, how others behave in the future. In a Markovian equilibrium, intertemporal incentives can only flow through this first channel, as the public signal is ignored. When the unique Perfect Public Equilibria is Markovian, it precludes the existence of any equilibria that use the public signal to generate intertemporal incentives via punishments and rewards. As such, the ability to generate effective intertemporal incentives in a stochastic game of imperfect monitoring stems entirely from the impact that action persistence has on future feasible payoffs.

Lastly, the results of this paper have practical implications for equilibrium analysis in a wide range of applied settings known to exhibit persistence and rigidities, ranging from industrial organization to political economy to macroeconomics. Markovian equilibria are a popular concept in applied work. Advantages of Markovian equilibria include their simplicity and their dependence on payoff relevant variables to specify incentives. Establishing that non-Markovian equilibria do not exist offers a strong justification for focusing on this more tractable class of equilibria. Additionally, this paper derives a tractable expression to construct Markovian equilibria, which can be used to formulate empirically testable predictions about equilibrium behavior. Equilibrium continuation values are specified by the solution to a *nonstochastic* differential equation defined over the state space, while the long-run player's action is determined by the sensitivity of its future payoffs to changes in the state variable (the first derivative of this solution). This result provides a tool that can be utilized for equilibrium analysis in applications. Once functional forms are specified for the underlying game, it is straightforward to derive the relevant differential equation, calibrate it with realistic parameters, and use numerical methods to estimate its solution. This solution is used to explicitly calculate equilibrium payoffs and actions, as a function of the state variable. Note these numerical methods are used for estimation in an equilibrium that has been characterized analytically, and not to simulate an approximate equilibrium; an important distinction.

It may be helpful to describe the contributions of this paper and its relationship to the existing literature using one application that can be studied with the tools developed in this paper: the canonical product choice game. Consider a long-run firm interacting with a sequence of short-run consumers. The firm has a dominant strategy to invest low effort, but would have greater payoffs if it could somehow commit to high quality (its “Stackelberg payoff”). Repeated interaction in discrete time with imperfect monitoring generates a folk theorem (Fudenberg and Levine 1994), but the striking implication from Faingold and Sannikov (2011) and Fudenberg and Levine (2007) is that such intertemporal incentives disappear as the period length becomes small. Since Fudenberg and Levine (1992), we know that if the firm could build a reputation for being a *commitment type* that produces only high quality products, a patient normal firm can approach these payoffs in every equilibrium. Faingold and Sannikov (2011) shows that this logic remains in continuous-time games, but that as in discrete-time, these reputation effects are temporary: eventually, consumers learn the firm's type, and reputation effects disappear in the long-run (Cripps, Mailath, and Samuelson 2007).¹ Departing from standard repeated and reputational games, consider

¹Mailath and Samuelson (2001) show that reputational incentives can also come from a firm's desire to

a simple and realistic modification in which the firm’s current product quality is a noisy function of past investment. Recent investment has a larger impact on current quality than investment further in the past, which is captured by a parameter θ that can be viewed as the rate at which past investment decays. Product quality, X_t , is modeled as a stochastic process:

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

given some initial value X_0 , investment path $(a_s)_{s \leq t}$ and Brownian motion $(Z_s)_{s \leq t}$. The evolution of product quality can be derived using this stochastic process, and takes the form:

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

When investment exceeds the current product quality, the firm is in a reputation building phase, and the product quality drifts upward (perturbed by a Brownian motion). When product quality is high and the firm chooses a lower investment level, it enters a period of reputational decay characterized by declining product quality. Applying the results of this paper to this product choice setting shows that there is a unique Perfect Public Equilibrium, which is Markovian in product quality. The firm’s reputation for quality will follow a cyclical pattern, characterized by phases of reputation building and decay. Importantly, this cyclical pattern does not dissipate with time and reputation effects are permanent; this contrasts with the temporary reputation effects observed in behavioral types models. The product choice game is one of the many settings that can utilize the tools and techniques of this paper to shed light on the relationship between persistent actions and equilibrium behavior.

In terms of techniques, this paper makes an important modeling choice by employing a continuous-time framework. Recent work has shown that the continuous time framework often allows for an explicit characterization of equilibrium payoffs, for any discount rate (Sannikov 2007); this contrasts with the folk-theorem results that typify the discrete time repeated games literature, and characterize the equilibrium payoff set as agents become arbitrarily patient (Fudenberg and Levine 1994). Additionally, continuous time allows for an explicit characterization of equilibrium behavior; an important feature if one wishes to use the model to generate empirical predictions and relate the model to observable behavior. In discrete time, results in similar settings have generally been limited to identifying equilibrium payoffs.

Related Literature: This paper uses tools developed by Faingold and Sannikov (2011), and so I comment on how I generalize their insights. Their setting can be viewed as a stochastic game in which the state variable is the belief that the firm is a commitment type, and the transition function follows Bayes rule. Faingold and Sannikov (2011) characterize the unique Markov equilibrium of this incomplete information game using an ordinary differential equation. This paper extends these tools to establish conditions for the existence and uniqueness of Markov equilibria in a setting with an arbitrary transition function between

separate itself from an incompetent type. Yet, these reputation effects are also temporary unless the type of the firm is replaced over time.

states, which can have a stochastic component independent of the public signal, where the long run player’s payoffs may also depend on the state variable and the state space may be unbounded or have endpoints that are not absorbing states.

My work is also conceptually related to [Board and Meyer-ter-Vehn \(2011\)](#). They model a setting in which product quality takes on high or low values and is replaced via a Poisson arrival process; when a replacement occurs, the firm’s current effort determines the new quality value. Consumers learn about product quality through noisy signals, and reputation is defined as the consumers’ belief that the current product quality is high. Realized product quality in their setting is therefore discontinuous (jumping between low and high), and this discontinuity plays an important role in determining intertemporal incentives. In the product choice application of my setting, the quality of a firm’s product is a smooth function of past investments and its investment today, and thus, the analysis is very different.

The role of persistence in intertemporal incentives can also be contrasted with our understanding of other continuous-time repeated games. [Sannikov and Skrzypacz \(2010\)](#) show that burning value through punishments that affect *all players* is not effective for incentives in settings with imperfect monitoring and Brownian signals, and that in these cases, it is more effective to punish by transferring value from some players to others. But in many settings, including those between long-run and myopic players, it would be impossible to avoid burning value and so intertemporal incentives collapse.² [Fudenberg and Levine \(2007\)](#) examine a product choice game between a long-run and short-run player and demonstrate that it is not possible to earn equilibrium payoffs above the payoffs corresponding to repeated play of the static Nash equilibrium when the volatility of the Brownian component is independent of the long-run player’s action. Thus, the intertemporal incentives that persistent actions induce could not emerge with standard continuous-time repeated games.

The organization of this paper proceeds as follows. Section 2 explores two simple examples to illustrate the main results of the model. Section 3 sets up the model. Section 4.4 analyzes equilibrium behavior and payoffs, while the final section concludes. All proofs are in the Appendix.

2 Examples

2.1 Persistent Investment as a Source of Reputation

Suppose a single long-run firm seeks to provide a continuum of small, anonymous consumers with a service. At each instant t , the firm chooses an unobservable investment level $a_t \in [0, \bar{a}]$. Consumers observe a noisy public signal of the firm’s investment each instant, which can be represented as a stochastic process with a drift term that depends on the firm’s action and a volatility term that depends on Brownian noise

$$dY_t = \theta a_t dt + \sigma dZ_t.$$

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

Investment is costly for the firm, but increases the likelihood of producing a high quality product. The stock quality of a product at time t , represented as X_t , captures the link between past investment levels and current product quality. This stock evolves according to a mean-reverting stochastic process where the change in stock quality at time t is

$$\begin{aligned} dX_t &= \theta dY_t - \theta X_t dt \\ &= \theta (a_t - X_t) dt + \sigma dZ_t. \end{aligned}$$

Stock quality is publicly observed. The expected change in quality is increasing when investment exceeds the current quality level, and decreasing when investment is below the current quality level. The parameter θ captures the persistence of investment: recent investment has a larger impact on current quality than investment further in the past. Thus, θ embodies the rate at which past investment decays: as it increases, more recent investments play a larger role in determining current product quality relative to investments further in the past. This stochastic process, known as the Ornstein-Uhlenbeck process, has a closed form that gives an insightful illustration of how past investments of the firm determine the current product quality. Given a history of investment choices $(a_s)_{s \leq t}$, the current value of product quality is

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

given some initial value of product quality X_0 . As shown in this expression, the impact of past investments decays at a rate proportional to the persistence parameter θ and the time that has elapsed since the investment was made.

Consumers simultaneously choose a purchase level $b_t^i \in [0, 10]$. The aggregate action of consumers, \bar{b}_t , is publicly observable, while individual purchase decisions are not. The firm's payoffs are increasing in the aggregate level of purchases by consumers, and decreasing in the level of investment. Average payoffs are represented as

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

where r is the common discount rate and $c < 1$ captures the cost of investment.

Consumers' payoffs depend on the stock quality, the firm's current investment level and their individual purchase decisions. As is standard in games of imperfect monitoring, payoffs can only depend on the firm's unobserved action through the public signal. Consumers are anonymous and their purchase decisions have a negligible impact on the aggregate purchase level. In equilibrium, they choose a purchase level that myopically optimizes their expected flow payoffs of the current stage game, represented as

$$E \left[\min \{b^i, 10\}^{1/2} [(1 - \lambda)dY + \lambda X] - b^i \right]$$

Marginal utility from an additional unit of product is decreasing in the current purchase level, with a saturation point at 10, and is increasing in current investment and stock quality. The parameter λ captures the importance of current investment relative to stock investment.

This product choice game can be viewed as a stochastic game with current product quality X as the state variable and the change in product quality dX as the transition function, which depends on the investment of the firm. I am interested in characterizing equilibrium payoffs and actions in a Markov perfect public equilibrium.

The firm is subject to binding moral hazard in that it would like to commit to a higher level of investment in order to entice consumers to choose a higher purchase level. However, in the absence of such a commitment device, the firm is tempted to deviate to lower investment. This example seeks to characterize when intertemporal incentives, particularly incentives created by the dependence of future feasible payoffs on current investment through persistent quality, can provide the firm with endogenous incentives to choose a positive level of investment. Note that in the absence of intertemporal incentives, the firm always chooses an investment level of $a = 0$.

In a Markov perfect equilibrium, the continuation value can be expressed as an ordinary differential equation that depends on the stock quality. Let $U(X)$ represent the continuation value of the firm when $X_t = X$. Then, given equilibrium action profile (a, \bar{b})

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

describes the relationship between U and its first and second derivatives. The continuation value can be expressed as the sum of the payoff that the firm earns today, $\bar{b} - ca^2$, and the expected change in the continuation value, weighted by the discount rate. The expected change in the continuation value has two components. First, the drift of quality determines whether quality is increasing or decreasing in expectation. Given that the firm's payoffs are increasing in quality ($U' > 0$), positive quality drift increases the expected change in the continuation value, while negative quality drift decreases this expected change. Second, the volatility of quality determines how the concavity of the continuation value relates to its expected change. If the value of quality is concave ($U'' < 0$), then volatility of quality hurts the firm. The firm is more sensitive to negative quality shocks than positive quality shocks, and has a higher continuation value at the expected quality relative to the expected continuation value of quality; in simple terms, the firm is "risk averse" in quality. Positive and negative shocks are equally likely with Brownian noise; thus, volatility has a net negative impact on the continuation value. If the value of quality is convex ($U'' > 0$), then volatility of quality helps the firm: the firm benefits more from positive quality shocks than it is hurt by negative quality shocks. The continuation value is graphed in Figure 1.

The firm faces a trade-off when choosing its investment level: the cost of investment is borne in the current period, but yields a benefit in future periods through higher expected purchase levels by consumers. The impact of investment on future payoffs is captured by the slope of the continuation value, $U'(X)$, which measures the sensitivity of the continuation value to changes in stock quality. In equilibrium, investment is chosen to equate the marginal cost of investment with its future expected benefit:

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

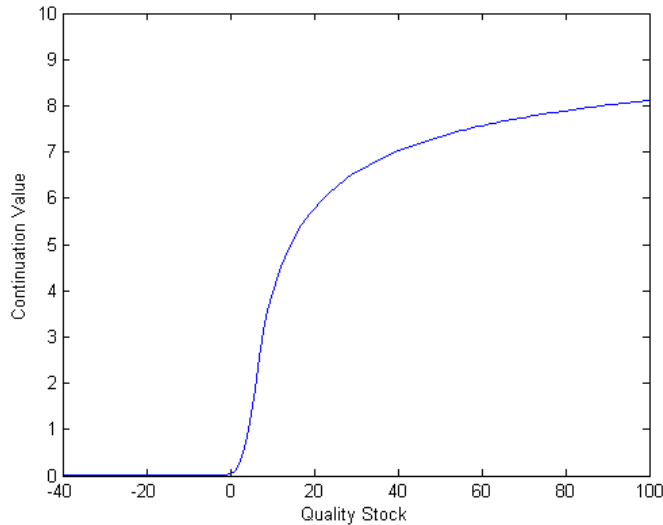


Figure 1: Equilibrium Payoffs

Marginal cost is captured by $2c$, while the marginal future benefit depends on the ratio of persistence to the discount rate. When θ is high, current investment will have a larger immediate impact on future quality, and the firm is willing to choose higher investment. Likewise, when the firm becomes more patient, it cares more about the impact investment today continues to have in future periods, and is willing to choose higher investment. It is interesting to note the trade-off between persistence and the discount rate. When investment decays at the same rate as the firm discounts future payoffs, these two parameters cancel. Thus, only the ratio of persistence to the discount rate is relevant for determining investment; as such, doubling θ has the same impact as halving the discount rate. Investment also depends on the sensitivity of the continuation value to changes in quality; when the continuation value is more sensitive to changes (captured by a steeper slope), the firm chooses a higher level of investment. As θ approaches 0, stock quality is almost entirely determined by its initial level and the intertemporal link between investment and payoffs is very small.

The boundary conditions that characterize the solution to $U(X)$ dictate that the slope of the continuation value converges to 0 as the stock quality approaches positive and negative infinity. Thus, the firm has the strongest incentive to invest at intermediate quality levels - a “reputation building” phase. When quality is very high, the firm’s continuation value is less sensitive to changes in quality and the firm has a weaker incentive to invest. In effect, the firm is “riding” its good reputation for quality. The incentive to invest is also weak when quality is very low, and a firm may wait out a very bad reputation shock before beginning to rebuild its reputation - “reputation recovery”. For interior values of X , the slope of the continuation value is positive, and thus the intertemporal incentives created by persistent actions allows the firm to choose a positive level of investment level.

In equilibrium, consumers myopically optimize flow payoffs by choosing a purchase level

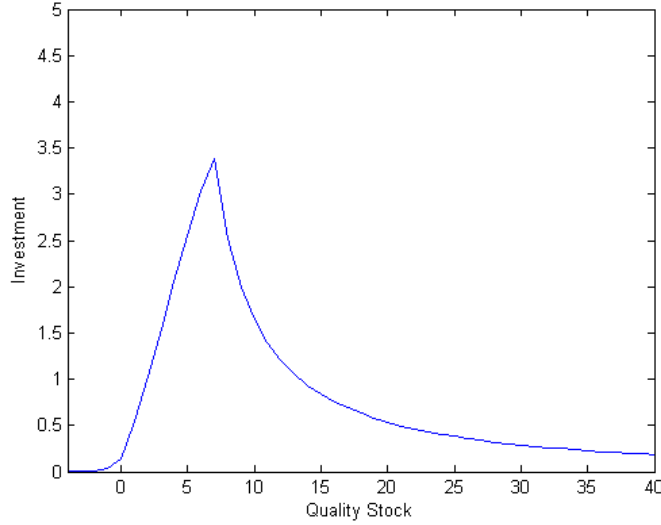


Figure 2: Firm Equilibrium Behavior

such that the marginal utility of an additional unit of product is zero:

$$b^i(a(X), X) = \left\{ \begin{array}{ll} 0 & \text{if } (1 - \lambda)a(X) + \lambda X \leq 0 \\ \frac{1}{4} [(1 - \lambda)a(X) + \lambda X]^2 & \text{if } (1 - \lambda)a(X) + \lambda X \in [0, 2\sqrt{10}] \\ 10 & \text{if } (1 - \lambda)a(X) + \lambda X > 2\sqrt{10} \end{array} \right\}$$

I show that there is a unique Perfect Public Equilibrium, which is Markovian in X_t ; as such, $a(X_t)$ and $b^i(X_t)$, are uniquely determined by X_t , and are also continuous.

Note that $(a(X_t), b^i(X_t))$ is uniquely specified by and continuous in (X_t) . Figures 2 and 3 graph equilibrium actions for the firm and consumers, respectively.

In this model, reputation effects are present in the long-run. Product quality is cyclical, with periods of high quality characterized by lower investment and negative drift, and periods of intermediate quality, where the firm chooses high investment and builds up its product quality. Very negative shocks can lead to periods where the firm chooses low investment and waits for its product quality to recover. Figure 4 illustrates the cycles of product quality across time. This contrasts with models in which reputations come from behavioral types: as [Cripps et al. \(2007\)](#) and [Faingold and Sannikov \(2011\)](#) show, reputation effects are temporary insofar as consumers eventually learn the firm's type, and so asymptotically, a firm's incentives to build reputation disappear. Additionally, conditional on the firm being strategic, reputation in these types models has negative drift.

Lastly, I compare the firm's payoffs in the stochastic game with action persistence to the benchmark without action persistence. The static Nash payoff depends on the value of stock quality. Let

$$v(X) = \min \left\{ 10, \frac{1}{4} \lambda^2 \max \{0, X\}^2 \right\}$$

represent the static Nash payoff of the firm when the stock quality is at level X . This payoff

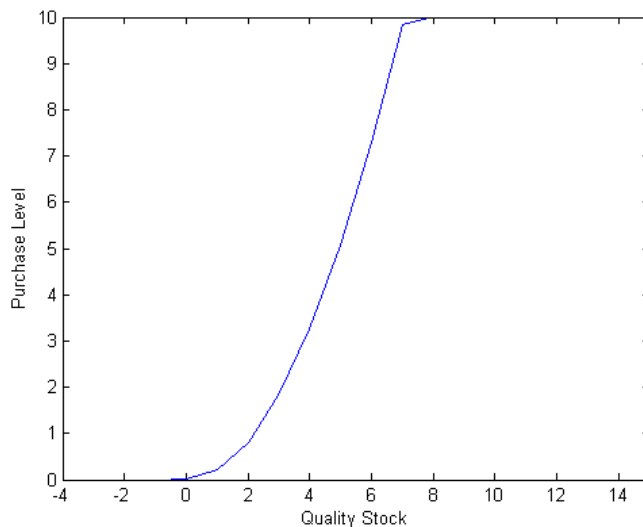


Figure 3: Consumer Equilibrium Behavior

is increasing in the stock quality. In the absence of investment persistence (this corresponds to $\theta = 0$), the unique equilibrium of the stochastic game is to play the static Nash equilibrium each period, which yields an expected continuation value at time t of

$$V(X_t) = r \int_t^\infty e^{-rs} E_t[v(X_s)] ds$$

Note that this expected continuation value may be above or below the static Nash equilibrium payoff of the current stage game, $v(X_t)$, depending on whether X_t is increasing or decreasing in expectation.

The firm achieves higher equilibrium payoffs when its actions are persistent, i.e. $U(X_t) \geq V(X_t)$ for all X_t . There are two complementary channels by which action persistence enhances the firm's payoffs. First, the firm chooses an investment level that equates the marginal cost of investment today with the marginal future benefit. Thus, in order for the firm to be willing to choose a positive level of investment, the future benefit of doing so must exceed the future benefit of choosing zero investment and must also exceed the current cost of this level of investment. Second, the link with future payoffs allows the firm to commit to a positive level of investment in the current period, which increases the equilibrium purchase level of consumers in the current period.

2.2 Policy Targeting

Elected officials and governing bodies often play a role in formulating and implementing policy targets. For example, the Federal Reserve targets interest rates, a board of directors sets growth and return targets for its company, and the housing authority targets home ownership rates. Achieving such targets requires costly effort on behalf of officials, and

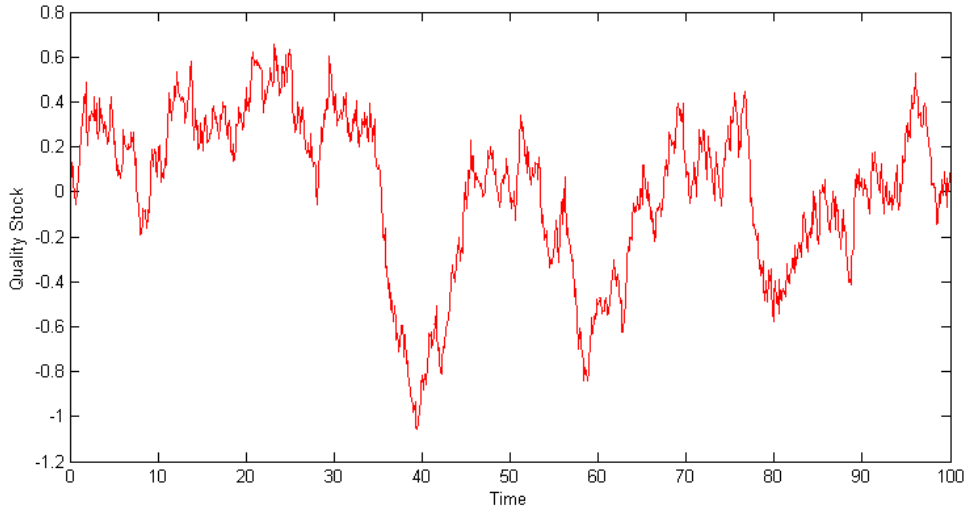


Figure 4: Product Quality Cycles

moral hazard issues arise because the preferences of the officials are not aligned with the population they serve. This example explores when a governing body can be provided with incentives to undertake a costly action in order to implement a target policy when the current level of the policy depends on the history of actions undertaken by the governing body.

Consider a setting where constituents elect a governing body to implement a policy target. The current policy takes on value $X_t \in [0, 2]$, and a policy target of $X_t = 1$ is optimal for constituents. In the absence of intervention, the policy drifts towards its natural level d . Each instant, the governing body chooses an action $a_t \in [-1, 1]$, where a negative action decreases the policy variable and a positive action increases the policy variable, in expectation. The policy evolves over time according to the stochastic process

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

Constituents also choose an action b_t^i each period, which represent their campaign contributions or support for the governing body. Constituents pledge higher support to the governing body when the policy is closer to their optimal target and when the governing body is exerting higher effort to achieve this target. I model the reduced form of the aggregate best response of constituents as

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

in which λ captures the value that constituents place on the governing body's effort to achieve the policy target.

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

$$g(a, \bar{b}_t, X_t) = \bar{b}_t - ca_t^2$$

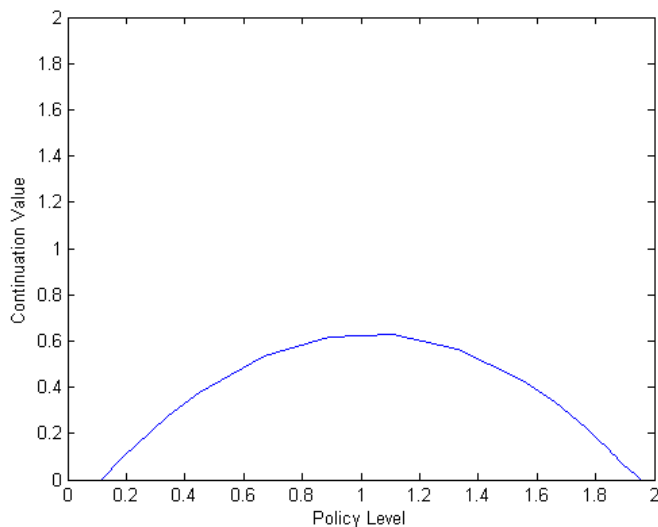


Figure 5: Equilibrium Payoffs

The unique Nash equilibrium of the static game is for the governing body to set $a = 0$ i.e. not intervene in reaching the desired policy, and for the constituents to support the governing body based on the difference between the desired and current policy level, $b = 1 - (1 - X)^2$. Given the current policy level is X , this yields a stage game Nash equilibrium payoff of

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

for the governing body. This payoff is concave in the state variable, and therefore the highest PPE payoff in the stochastic game occurs at the value of the state variable that maximizes the stage game Nash equilibrium payoff, $X = 1$, which yields a stage game payoff of $v(1) = 1$. The highest PPE payoff is strictly less than the highest static game Nash equilibrium payoff. Figure 4 plots the PPE payoff of the governing body, as a function of the current policy level. This payoff is increasing in the policy level for levels below the optimal target, and decreasing in the policy level for levels above the optimal target.

The characterization of a unique Markovian equilibrium can be used to determine the equilibrium effort level of the governing body. Let $U(X)$ represent the continuation value as a function of the policy level in such an equilibrium, which is plotted in figure 5

The optimal effort choice of the governing body depends on the slope of the continuation value, the sensitivity of the change in the policy level to the effort level, and the cost of effort.

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

When the current policy level is very far from its optimal target, the effort of the governing body has a smaller impact on the policy level, and the governing body has a lower incentive to undertake costly effort. When the policy level is close to the optimal target, the continuation value approaches its maximum, and the slope of the continuation value approaches zero.

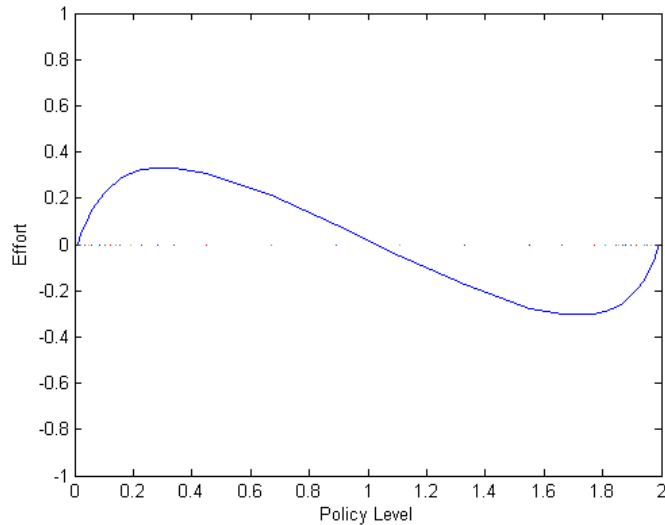


Figure 6: Government Equilibrium Behavior

Thus, the governing body also has a lower incentive to undertake costly effort when the policy is close to its target. Figure 6 plots the equilibrium effort choice of the governing body as a function of the policy level. As illustrated in the figure, the governing body exerts the highest effort when the policy variable is an intermediate distance from the optimal target. Figure 7 shows the equilibrium constituent support, which is highest when the policy level is closest to its optimal target.

3 Model

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

The Stage Game: At each instant t , the agency and collective members simultaneously choose actions a_t from A and b_t^i from B , respectively, where A and B are compact sets of a Euclidean space. Individual actions privately observed. Rather, the aggregate distribution of the collective's action, $\bar{b}_t \in \Delta B$ and a public signal of the agency's action, dY_t , are publicly observed. The public signal evolves according to the stochastic differential equation

$$dY_t = \mu_Y(a_t, \bar{b}_t)dt + \sigma_Y dZ_t^Y$$

where $(Z_t^Y)_{t \geq 0}$ is a Brownian motion, $\mu_Y : A \times B \rightarrow R$ is the drift and $\sigma_Y \in R$ is the volatility. Assume μ_Y is a Lipschitz continuous function. The drift term provides a signal of the agency's action and can also depend on the aggregate action of the collective, but is

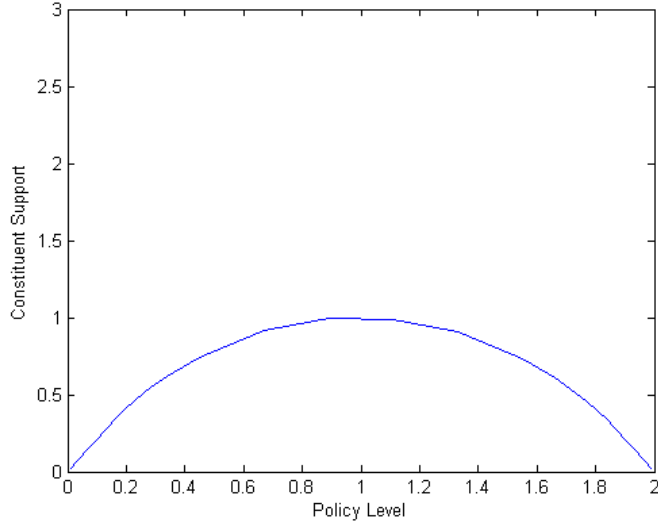


Figure 7: Constituent Equilibrium Behavior

independent of the individual actions of the collective to preserve anonymity. The volatility is independent of players' actions.

The Stochastic Game: The stage game varies across time through its dependence on a state variable $(X_t)_{t \geq 0}$, which takes on values in the state space $\Xi \subset R$ and evolves stochastically as a function of the current state and players' actions. The path of the state variable is publicly observable. As the state variable is not intended to provide any additional signal of players' actions, its evolution depends on actions solely through the available public information. The transition of the state variable is governed by the stochastic differential equation:

$$dX_t = f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t)dt + f_2(\bar{b}_t, X_t)dt + f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t)dZ_t^X$$

where $f_1 : B \times \Xi \rightarrow R$, $f_2 : B \times \Xi \rightarrow R$ and $\sigma_X^2 : \Xi \rightarrow R$ are Lipschitz continuous functions, and $(Z_t^X)_{t \geq 0}$ is a Brownian motion which is assumed to be orthogonal to $(Z_t^Y)_{t \geq 0}$. The drift of the state variable has two components: the first component, $f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t)$, specifies how the agency's action influences the transition of the state, while the second component, $f_2(\bar{b}_t, X_t)$, is independent of the firm's action and allows the model to capture other channels that influence the transition of the state variable. The volatility of the state variable depends on the volatility of the public signal, $f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y$, as well as a volatility term that is independent of the public signal, $\sigma_X(X_t)dZ_t^X$. Note that the same function multiplies the drift and volatility of the public signal; this ensures that no additional information about the agency's action is revealed by the evolution of the state variable. Let $\{F_t\}_{t \geq 0}$ represent the filtration generated by the public information, $(Y_t, X_t)_{t \geq 0}$.³

³The state space may or may not be bounded. It is bounded if (i) there exists an upper bound \bar{X} at

I assume that the volatility of the state variable is positive at all interior points of the state space. This ensures that the future path of the state variable is always stochastic. Brownian noise can take on any value in R , and as such, this assumption means that any future path of the state variable, $(X_s)_{s>t}$ can be reached from the current state $X_t \in \Xi$. This assumption is analogous to a strong form of irreducibility, since any state $X_s \in \Xi$ can be reached from the current state X_t at all times $s > t$.

Assumption 1. *For any compact proper subset $I \subset \Xi$, there exists a c such that*

$$\sigma_I = \inf_{\bar{b} \in B, X \in I} [f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)] > c$$

Note that this assumption does not preclude the possibility that the state variable evolves independently of the public signal, which corresponds to $f_1 = 0$.

Define a state X as an *absorbing state* if the drift and volatility of the transition function are both zero. The following definition formalizes the conditions that characterize an absorbing state.

Definition 1. *$X \in \Xi$ is an absorbing state if there exists an action profile $\bar{b} \in B$ such that $f_1(\bar{b}, X) = 0$, $f_2(\bar{b}, X) = 0$ and $\sigma_X(X) = 0$.*

Remark 1. *The assumption that the volatility of the state variable is positive at all interior points of the state space precludes the existence of interior absorbing points. Given that Brownian motion is continuous, this is without loss of generality. To see why, suppose that $\Xi = [\underline{X}, \bar{X}]$ and there is an interior absorbing point X^* , and the initial state is $X_0 < X^*$. Then states $X > X^*$ are never reached under any strategy profile, and the game can be redefined on the state space $\Xi = [\underline{X}, X^*]$.*

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

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

where r is the discount rate. Assume g is Lipschitz continuous and bounded for all $a \in A$, $\bar{b} \in \Delta B$ and $X \in \Xi$. The dependence of payoffs on the state variable creates a form of action persistence for the firm, since the state variable is a function of prior actions.

Collective members' have identical preferences, and each member seeks to maximize its expected flow payoff at time t ,

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

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

which the volatility is zero and the drift is weakly negative, i.e. $f_1(\bar{b}, \bar{X}) = 0$; $\sigma_X(\bar{X}) = 0$, and $f_2(\bar{b}, \bar{X}) \leq 0$ for all $\bar{b} \in B$; and (ii) there exists a lower bound $\underline{X} < \bar{X}$ such that the volatility is zero and the drift is weakly positive, i.e. $f_1(\bar{b}, \underline{X}) = 0$; $\sigma_X(\underline{X}) = 0$ and $f_2(\bar{b}, \underline{X}) \geq 0$ for all $\bar{b} \in B$.

Thus, in the stochastic game, at each instant t , given the current state X_t , players choose actions, and then nature stochastically determines payoffs, the public signal and next state as a function of the current state and action profile. The game defined here includes several subclasses of games, including a game where the state variable evolves independently of the agency's action ($f_1(\bar{b}, X) = 0$), the state variable evolves deterministically given the public signal ($\sigma_X(X) = 0$), or the agency's payoffs only depend on the state indirectly through the actions of the collective ($g(a, \bar{b}, X) = g(a, \bar{b})$).

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

3.1 Equilibrium Structure

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

In any PPE, collective members choose b_t^i to myopically optimize expected flow payoffs each instant.⁴ Let $\mathcal{B} : A \times \Delta B \times \Xi \rightrightarrows B$ represent the best response correspondence that maps an action profile and a state to the set of collective member actions that maximize payoffs in the current stage game, and $\bar{\mathcal{B}} : A \times \Xi \rightrightarrows \Delta B$ represent the aggregate best response function. In many applications, it will be sufficient to specify the aggregate best response function as a reduced form for the collective's behavior.

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

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

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

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

⁴The individual actions of a collective member, b_t^i , has a negligible impact on the aggregate action \bar{b}_t (and therefore X_t) and is not observable by the agency. Therefore, the model could also allow for long-run small, anonymous players.

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

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

$$\begin{aligned} dW_t(S) &= r(W_t(S) - g(a_t, \bar{b}_t, X_t)) dt \\ &\quad + r\beta_{1t} [dY_t - \mu_Y(a_t, \bar{b}_t)dt] \\ &\quad + r\beta_{2t}\sigma_X(X_t)dZ_t^X \end{aligned}$$

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

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

$$\begin{aligned} a_t &\in \arg \max g(a', \bar{b}_t, X_t) + \beta_{1t}\mu_Y(a', \bar{b}_t) \\ b_t^i &\in \mathcal{B}(a_t, X_t) \end{aligned}$$

The continuation value of the agency is a stochastic process that is measurable with respect to public information, $\{F_t\}_{t \geq 0}$. Two components govern the motion of the continuation value, a drift term that captures the difference between the current continuation value and the current flow payoff. This is the expected change in the continuation value. A volatility term β_{1t} determines the sensitivity of the continuation value to the public signal: the agency’s future payoffs are more sensitive to good or bad signal realizations when the volatility of the continuation value is larger. A second volatility term β_{2t} captures the sensitivity of the continuation value to the stochastic element of the state variable that is independent of the public signal.

The condition for sequential rationality depends on the process $(\beta_t)_{t \geq 0}$, which specifies how the continuation value changes with respect to the public information. Today’s action impacts future payoffs through the drift of the public signal, $\mu_Y(a, \bar{b})$, and the sensitivity of the continuation value to the public signal, β_1 , while it impacts current payoffs through the flow payoff of the agency, $g(a, \bar{b}, X)$. A strategy for the agency is sequentially rational if it maximizes the sum of flow payoffs today and the expected impact of today’s action on future payoffs. This condition is analogous to the one-shot deviation principle in discrete time.

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

Remark 2. *The key aspect of this model that allows for this tractable characterization of the agency’s incentive constraint is the assumption that the volatility of the state variable is always positive (except at the boundary of the state space), which ensures that any future path of states can be reached from the current state. This assumption, coupled with the linear*

incentive structure of Brownian information, ensures the condition for sequential rationality takes the form in Lemma 1. To see this, consider a deviation from a_t to \tilde{a}_t at time t . This deviation impacts future payoffs by inducing a different probability measure over the future path of the state variable, $(X_s)_{s>t}$, but doesn't affect the set of feasible sample paths. Given that all paths of the state variable are feasible under a_t and \tilde{a}_t , the continuation value under both strategies is a non-degenerate expectation with respect to the future path of the state variable. Thus, the change in the continuation value when the agency deviates from a_t to \tilde{a}_t depends solely on the different measures a_t and \tilde{a}_t induce over future sample paths, and, given the requirement that Brownian information is used linearly, this change is linear with respect to the difference in the drift of the public signal, $\mu_Y(\tilde{a}_t, \bar{b}_t) - \mu_Y(a_t, \bar{b}_t)$.

Remark 3. It is of interest to note that it is precisely this linear structure with respect to the Brownian information, coupled with the inability to transfer continuation payoffs between players, that precludes the effective provision of intertemporal incentives in a standard repeated game between a long-run and short-run player. The short-run player acts myopically, so it is not possible to tangentially transfer continuation values between players. Using Brownian information linearly, but non-tangentially, results in the continuation value escaping the boundary of the payoff set with positive probability, and Brownian information cannot be used effectively in a non-linear manner. This paper will illustrate that a stochastic game permits the provision of intertemporal incentives by introducing the possibility of linearly using Brownian information for some values of the state variable.

The sequential rationality condition can be used to specify an auxiliary stage game parameterized by the state variable and the process linking current play to the continuation value. Let $S^*(X, \beta_1) = \{(a, \bar{b})\}$ represent the correspondence of static Nash equilibrium action profiles in this auxiliary game, defined as:

Definition 2. Define $S^*(X, \beta_1) = \Xi \times R \rightrightarrows A \times \Delta B$ as the correspondence that describes the Nash equilibrium of the static game parameterized by $(X, \beta_1) \in \Xi \times R$:

$$S^*(X, \beta) = \left\{ \begin{array}{l} a \in \arg \max_{a'} g(a', \bar{b}, X) + \beta_1 \mu_Y(a', \bar{b}) \\ \bar{b} \in \bar{B}(a, X) \end{array} \right\}$$

In any PPE strategy profile $(a_t, \bar{b}_t)_{t \geq 0}$ of the stochastic game, given some processes $(X_t)_{t > 0}$ and $(\beta_{1t})_{t > 0}$, the action profile at each instant must be a static Nash equilibrium of the auxiliary game i.e. $(a_t, \bar{b}_t) \in S^*(X_t, \beta_{1t})$ for all t . I assume that this auxiliary stage game has a unique static Nash equilibrium with an atomic distribution over small players' actions. While this assumption is somewhat restrictive, it still allows for a broad class of games, including those discussed in the previous examples.

Assumption 2. Assume $S^*(X, \beta)$ is non-empty and single-valued for all $(X, \beta) \in \Xi \times R$, Lipschitz continuous on any subset of $\Xi \times R$, and the small players choose identical actions $b^i = \bar{b}$.

Note that $S^*(X, 0)$ corresponds to the Nash equilibrium of the stage game in the current model when the state variable is equal to X .

Static Equilibria Payoffs: The feasible payoffs of the current stage game depend on the state variable, as do stage game Nash equilibrium payoffs. The presence of myopic players imposes restrictions on the payoffs that can be achieved by the long-run player, given that the myopic players must play a static best response.

Define $v : \Xi \rightarrow R$ as the payoff to the agency in the Nash equilibrium of the stage game, parameterized by the state variable, where $v(X) := g(S^*(X, 0), X)$. The assumption that the Nash equilibrium correspondence of the stage game is Lipschitz continuous, non-empty and single-valued guarantees $v(X)$ is a Lipschitz continuous function. When the state space is bounded, $\Xi = [\underline{X}, \overline{X}]$, $v(X)$ has a well-defined limit as it approaches the highest and lowest state. If the state space is unbounded, an additional assumption is necessary to guarantee that $v(X)$ has well-defined limits.

Assumption 3. *If the state space is unbounded, $\Xi = R$, then there exists a δ such that for $|X| > \delta$, $v(X)$ is monotonic in X .*

This assumption ensures that $v(X)$ doesn't oscillate as it approaches infinity, a technical assumption that is necessary for the equilibrium uniqueness result. Represent the highest and lowest stage Nash equilibrium payoffs across all states as:

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

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

This subsection illustrates the model and definitions introduced above.

Example 1. Pricing Quality: *Consider a setting where a firm invests in developing a product, and consumers choose the price they are willing to pay to purchase this product. Each instant, a firm chooses an investment level $a_t \in [0, 1]$ and consumers choose a price $b^i \in [0, \overline{B}]$. A public signal provides information about the firm's instantaneous investment through the process*

$$dY_t = a_t dt + dZ_t^Y$$

The state variable is product quality, which is a function of past investments and takes on values in the bounded support $\Xi = [0, \overline{X}]$. The change in product quality is governed by the process

$$dX_t = X_t(\overline{X} - X_t) (a_t dt + dZ_t^Y) - X_t dt$$

which is increasing in the firm's investment, and decreasing in the current product quality. Note this corresponds to $\mu_Y = a_t$, $\sigma_Y = 1$, $f_1 = X_t(\overline{X} - X_t)$, $f_2 = -X_t$ and $\sigma_X^2 = 0$. At the upper bound of product quality, investment no longer impacts product quality and the process has negative drift. At the lower bound, investment also no longer impacts product quality, and the process has zero drift. As such, $X = 0$ is an absorbing state but $X = \overline{X}$ is not.

The firm earns the price the consumers are willing to pay for the product, and pays a cost of $c(a)$ for an investment level of a , with $c(0) = 0, c' > 0$ and $c'' > 0$. Its payoff function is

$$g(a, \bar{b}, X) = \bar{b} - c(a)$$

which is independent of product quality.

Consumers receive an instantaneous value of $X + a$ from purchasing a unit of the product. Their flow payoff is the difference between the purchase price and the value of the product,

$$h(a, b^i, \bar{b}, X) = - (b^i - X - a)^2$$

In equilibrium, consumers myopically optimize flow payoffs, and thus pay a price equal to their expected utility from purchasing the product, which is increasing in the stock quality and investment of the firm.⁵ The aggregate consumer best response function takes the form

$$\bar{b}(a, X) = X + a$$

In the static game, given a product quality of X , the unique Nash equilibrium is for the firm to choose an investment level of $a^* = 0$ and the consumers to pay a price of $\bar{b}^*(0, X) = X$ for the good. This yields a stage game Nash equilibrium payoff of $v(X) = X$ for the firm, a maximum stage NE payoff of $\bar{v}^* = \bar{X}$ at $X = \bar{X}$ and a minimum stage NE payoff of $\underline{v}^* = 0$ at $X = 0$.

In the stochastic game, the firm also considers the impact that current investment has on future product quality. Using the condition for sequential rationality specified in Lemma 1, the firm chooses an investment level to maximize

$$a \in \arg \max_{a'} X + a - c(a') + \beta a'$$

which yields an equilibrium action

$$a^*(X, \beta) = (c')^{-1}(\beta)$$

Equilibrium investment is strictly positive in the stochastic game when $\beta > 0$. Thus, persistent investment allows the firm to overcome the binding moral hazard present in the static game and earn a higher price for its product.

Note that

$$S^*(X, \beta) = \left((c')^{-1}(\beta), X + (c')^{-1}(\beta) \right)$$

which is non-empty, single-valued, unique and Lipschitz continuous for each (X, β) .

⁵While it may seem unusual that the consumer receives negative utility when they pay a price lower than the value of the product, this setting can be interpreted as the reduced form for a monopolistic market in which the firm captures all of the surplus from the product quality. Such a setting would yield the same aggregate best response function, which is the only relevant aspect of consumer behavior for equilibrium analysis.

4 Equilibrium Analysis

This section presents the main results of the paper, and proceeds as follows. First, I construct a Markovian equilibrium in the state variable, which simultaneously establishes the existence of at least one Markovian equilibria and characterizes equilibrium behavior and payoffs in such an equilibrium. Next, I establish conditions for a Markovian equilibrium to be the unique equilibrium in the class of all Perfect Public Equilibria. Following is a brief discussion on the role action persistence plays in using Brownian information to create effective intertemporal incentives. An application of the existence and uniqueness results to a stochastic game *without* action persistence shows that the agency acts myopically in the unique Perfect Public Equilibria of this setting. Finally, I use the equilibrium characterization to describe several interesting properties relating the agency's equilibrium payoffs to the structure of the underlying stage game.

4.1 Existence of Markov Perfect Equilibria

The first main result of the paper establishes the existence of a Markovian equilibrium in the state variable. The existence proof is constructive, and as such, characterizes the explicit form of equilibrium continuation values and actions in Markovian equilibria. This result applies to a general setting in which:

- The state space may be bounded or unbounded.
- The transition function governing the law of motion of the state variable is stochastic and depends on the agency's action through a public signal, as well as the aggregate action of the collective and the current value of the state.
- There may or may not be absorbing states at the endpoints of the state space.

Theorem 1. *Suppose Assumptions 1 and 2 hold. Then given an initial state X_0 and action profile $(a, \bar{b}) = S^*(X, U'(X)f_1(\bar{b}, X))$, any bounded solution $U(X)$ to the second order differential equation:*

$$U''(X) = \frac{2r [U(X) - g(a, \bar{b}, X)]}{f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)} - \frac{2 [f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]}{f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)} U'(X)$$

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

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

$$S^*(X, U'(X)f_1(\bar{b}, X)) = \left\{ \begin{array}{l} a = \arg \max_{a'} rg(a', \bar{b}, X) + U'(X)f_1(\bar{b}, X)\mu(a', \bar{b}) \\ \bar{b} = \bar{B}(a, X) \end{array} \right\}$$

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

Theorem 1 shows that the stochastic game has at least one Markovian equilibrium. Continuation values in this equilibrium are represented by a second order ordinary differential equation. Rearranging the optimality equation as:

$$U(X) = g(a, \bar{b}, X) + \frac{1}{r} [f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] U'(X) + \frac{1}{2r} U''(X) [f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)]$$

lends insight into the relationship between the continuation value and the transition of state variable. The continuation value is equal to the sum of the flow payoff today, $g(a, \bar{b}, X)$, and the expected change in the continuation value, weighted by the discount rate. The second term captures how the continuation value changes with respect to the drift of the state variable. For example, if the state variable has positive drift ($f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X) > 0$), and the continuation value is increasing in the state variable ($U' > 0$), then this increases the expected change in the continuation value. The third term captures how the continuation value changes with respect to the volatility of the state variable. If U is concave ($U'' < 0$), it is more sensitive to negative shocks than positive shocks. Positive and negative shocks are equally likely, and therefore, the continuation value is decreasing in the volatility of the state variable. If U is linear ($U'' = 0$), then the continuation value is equally sensitive to positive and negative shocks, and the volatility of the state variable does not impact the continuation value.

Now consider a value of the state variable that yields a local maximum $U(X^*)$ (note this implies $U' = 0$). Since the continuation value is at a local maximum, it must be decreasing as X moves away from X^* in either direction. This is captured by the fact that $U''(X) < 0$. Larger volatility of the state variable or a more concave function lead to a larger expected decrease in the continuation value.

I now outline the intuition behind the proof of Theorem 1. The first step in proving this existence is to show that if a Markovian equilibrium exists, then continuation values must be characterized by the solution to the optimality equation. In a Markovian equilibrium, continuation values take the form $W_t = U(X_t)$ 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 Markovian equilibrium $dW_t = dU(X_t)$, as a function of the law of motion for the state variable:

$$\begin{aligned} dU(X_t) &= U'(X_t) [f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t) + f_2(\bar{b}_t, X_t)] dt \\ &\quad + \frac{1}{2} U''(X_t) [f_1(\bar{b}_t, X_t)^2 \sigma_Y^2 + \sigma_X^2(X_t)] dt \\ &\quad + U'(X_t) [f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t) dZ_t^X] \end{aligned}$$

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

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

Matching the drifts of these two laws of motion yields the optimality equation, a second order ordinary differential equation that specifies continuation payoffs as a function of the state variable.

The next step in the existence proof is to show that this ODE has at least one solution that lies in the range of feasible payoffs for the agency. The technical condition to guarantee the existence of a solution is that the second derivative of U is bounded with respect the first derivative of U on any bounded interval of the state space. The denominator of the optimality equation depends on the volatility of the state variable. Thus, the assumption that the volatility of the state variable is positive on any open interval of the state space (Assumption 1) is crucial to ensure this condition is satisfied. The numerator of the optimality equation depends on the drift of the state variable, and the agency's flow payoff. Lipschitz continuity of these functions ensures that they are bounded on any bounded interval of the state space. These conditions are sufficient to guarantee the optimality equation has at least one bounded solution that lies in the range of feasible payoffs for the agency.

The final step of the existence proof is to construct a Markovian equilibrium that satisfies the conditions of a PPE established in Lemma 1. The incentive constraint for the agency is constructed by matching the volatility of the laws of motion for the continuation value established in Lemma 1 with the volatility of the law of motion for the continuation value as a function of the state variable, $dU(X_t)$. Lemma 1 established that the volatility of the continuation value must be

$$r\beta_{1t}\sigma_Y dZ_t^Y + r\beta_{2t}\sigma_X(X_t)dZ_t^X$$

in any PPE. Thus, in a Markovian equilibrium

$$\begin{aligned} r\beta_{1t}\sigma_Y &= U'(X_t)f_1(\bar{b}_t, X_t)\sigma_Y \\ r\beta_{2t}\sigma_X(X_t) &= U'(X_t)\sigma_X(X_t) \end{aligned}$$

This characterizes the process $(\beta_t)_{t \geq 0}$ governing incentives, and as such, the incentive constraint for the agency. This incentive constraint takes an intuitive form. The impact of the current action on future payoffs is captured by the impact the current action has on the state variable, $f_1(\bar{b}, X)\mu(a', \bar{b})$, as well as the slope of the continuation value, $U'(X_t)$, which captures how the continuation value changes with respect to the state variable.

Theorem 1 also establishes that *each* solution U to the optimality equation characterizes a single Markovian equilibrium. This is a direct consequence of the assumption that the Nash equilibrium correspondence of the auxiliary stage game $S^*(X, \beta)$ is single-valued, Assumption 2, which guarantees that U uniquely determines equilibrium actions. Note that if there are multiple solutions to the optimality equation, then each solution characterizes a single Markovian equilibrium. The formal proof of Theorem 1 is presented in the Appendix.

Markovian equilibria have an intuitive appeal in stochastic games. Advantages of Markovian equilibria include their simplicity and their dependence on payoff relevant variables to specify incentives. Theorem 1 yields a tractable expression that can be used to construct equilibrium behavior and payoffs in a Markovian equilibrium. The continuation value of

the agency is specified by the solution to a second order differential equation defined over the state space. The agency's incentives are governed by the slope of this solution, which determines how the continuation value changes with the state variable. As such, this result provides a tool to analyze equilibrium behavior in a broad range of applied settings. Once functional forms are specified for the agency's payoffs and the transition function of the state variable, it is straightforward to use Theorem 1 to characterize the optimality equation and incentive constraint for the agency, as a function of the state variable. This constructs a Markovian equilibrium. Numerical methods for ordinary differential equations can then be used to estimate a solution to the optimality equation and explicitly calculate equilibrium payoffs and actions. These calculations yield empirically testable predictions about equilibrium behavior. Note that numerical methods are used for estimation in an equilibrium that has been characterized analytically, and not to simulate an approximate equilibrium. This is an important distinction.

4.1.1 Example to illustrate Theorem 1

The following example illustrates how to use Theorem 1 to construct equilibrium behavior.

Example 2. Consider the persistent investment model presented in Section 2.1. The state variable evolves according to:

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

which corresponds to $f_1 = 1$, $\mu_Y = \theta a$, $f_2 = -\theta X$, $\sigma_Y = \sigma$ and $\sigma_X = 0$, and the firm's flow payoff is:

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

Using Theorem 1 to characterize the optimality equation yields

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

The sequential rationality condition for the firm is

$$a = \arg \max_{a'} \bar{b} - ca^2 + U'(X)\theta a'$$

In equilibrium, the firm chooses action

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

This constructs equilibrium behavior and payoffs as a function of the current product quality X and the solution to the optimality equation, U . Numerical methods can now be used to estimate a solution U to the optimality equation. Calibrating the model with a set of parameters will then fully determine equilibrium actions and payoffs as a function of the current product quality. As discussed in 2.1, the empirical predictions of this application are:

1. The firm's continuation value is increasing in the current product quality

2. *The firm's incentives to invest in quality are highest when the current product quality is at an intermediate level. As such, the firm goes through phases of "reputation building", during which the firm chooses high investment levels and product quality increases in expectation, and "reputation riding", during which the firm chooses low investment levels and reaps the benefits of having a high product quality.*

The firm's equilibrium payoff captures the future value of owning a product of a given quality level, and as such, can be interpreted as the asset value of the firm.

4.2 Uniqueness of Markovian Equilibrium

The second main result of the paper establishes conditions under which there is a unique Markovian equilibrium, which is also the *unique equilibrium* in the class of *all Perfect Public Equilibria*. The first step of this result is to establish when the optimality equation has a unique bounded solution. Recall that each solution to the optimality equation characterized in Theorem 1 characterizes a single Markovian equilibrium. Thus, when the optimality equation has a unique solution, there is a unique Markovian equilibrium. The second step of the result is to prove that there are no non-Markovian PPE, and as such, this unique Markovian equilibrium is the unique PPE.

The optimality equation will have a unique solution when its solution satisfies certain boundary conditions as the state variable approaches its upper and lower bound (in the case of an unbounded state space, as the state variable converges to positive or negative infinity). The boundary conditions for the optimality equation depend on the rate at which the drift and volatility of the state variable converge as the state variable approaches its upper and lower bound. As such, the key condition that ensures a unique solution to the optimality equation is an assumption on the limiting behavior of the drift and volatility of the state variable.

Assumption 4. *1. If the state space is bounded, $\Xi = [\underline{X}, \bar{X}]$, then as X approaches its upper and lower bound $\{\underline{X}, \bar{X}\}$, the functions governing the transition of the state variable satisfy the following limiting behavior:*

- (a) *The drift of the state variable converges to zero at a linear rate, or faster: $f_2(\bar{b}, X)$ and $f_1(\bar{b}, X)$ are $O(X^* - X)$ as $X \rightarrow X^* \in \{\underline{X}, \bar{X}\}$.*
- (b) *The volatility of the state variable converges to zero at a linear rate, or faster: $1/f_1(\bar{b}, X)\sigma_Y(\bar{b}) + \sigma_X(X)$ is $O(1/(X^* - X))$ as $X \rightarrow X^* \in \{\underline{X}, \bar{X}\}$.*

2. *If the state space is unbounded, $\Xi = R$, then as X approaches positive and negative infinity, the functions governing the transition of the state variable satisfy the following limiting behavior:*

- (a) *The drift of the state variable grows linearly, or slower: $f_2(\bar{b}, X)$ and $f_1(\bar{b}, X)$ are $O(X)$ as $X \rightarrow \{-\infty, \infty\}$.*

- (b) *The volatility of the state variable is bounded: $f_1(\bar{b}, X)\sigma_Y(\bar{b}) + \sigma_X(X)$ is $O(1)$ as $X \rightarrow \{-\infty, \infty\}$.*

When the support is bounded, this assumption requires that the upper and lower bounds of the state space are absorbing points. The drift and volatility of the state variable must converge to zero at a linear rate, or faster, as the state variable approaches its boundary. When the support is unbounded, these assumptions require that the drift of the state variable grows at a linear rate, or slower, as the magnitude of the state becomes arbitrarily large, and that the volatility of the state variable is uniformly bounded. The role this assumption plays in establishing equilibrium uniqueness is discussed following the presentation of the result.

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

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

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

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

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

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

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

I briefly relate the boundary conditions characterized in Theorem 2 to equilibrium behavior and payoffs, and then outline the intuition behind the uniqueness result. These boundary conditions have several implications for equilibrium play. Recall the incentive constraint for the agency from Theorem 1. The link between the agency's action and future payoffs is proportional to the slope of the continuation value and the drift component of the state variable that depends on the public signal, $U'(X)f_1(\bar{b}, X)$. The assumption on the growth

rate of $f_1(\bar{b}, X)$ ensures that $U'(X)f_1(\bar{b}, X)$ converges to zero at the boundary points (in the unbounded case, as the state variable approaches positive or negative infinity). When this is the case, the agency's incentive constraint is reduced to the myopic optimization of its instantaneous flow payoff at the boundary points. Thus, at the upper and lower bound of the state space (in the limit for an unbounded state space), the agency plays a static Nash equilibrium action. Additionally, continuation payoffs are equal to the Nash equilibrium payoff of the static game at the boundary points.

I next provide a sketch of the proof for the existence of a unique PPE, which is Markovian. The first step in proving this result is establishing that the optimality equation has a unique solution. This is done so in two parts: (i) showing that any solution to the optimality equation must satisfy the same boundary conditions, and (ii) showing that it is not possible for two different solutions to the optimality equation to satisfy the same boundary conditions.

I discuss the boundary conditions for an unbounded state space; the case of a bounded state space is analogous. Suppose U is a bounded solution to the optimality equation. The U , and its first and second derivative, must satisfy the following set of boundary conditions. U will have a well-defined limit at the boundary points when the static Nash equilibrium payoff function has a well-defined limit, which is guaranteed given the Lipschitz continuity of the Nash equilibrium correspondence and the agency's payoff function. The boundedness of U coupled with the assumption that the static Nash equilibrium payoff function is monotonic for large X ensures that the first derivative of U converges to zero, and does so faster than $1/X$.⁶ This establishes the boundary condition on U' presented in Theorem 2,

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

The boundedness of U also ensures that the second derivative, U'' , doesn't converge to a constant.⁷ The optimality equation in Theorem 1 specifies the relationship between U and its first and second derivative. This relationship, coupled with Assumption 4 is used to establish the boundary condition for U . From the optimality equation,

$$(f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)) U''(X) = 2r [U(X) - g(a, \bar{b}, X)] - 2 [f_1(\bar{b}, X) \mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] U'(X)$$

Consider the limit of the optimality equation as the state variable approaches positive infinity. Under the assumption that the drift of the state variable has linear growth (f_1 and f_2), the second term on the right hand side converges to zero, and the flow payoff $g(a, \bar{b}, X)$ converges to v_∞ . (Recall that the agency plays a myopic best response at the boundaries, which yields a flow payoff equal to the static Nash equilibrium payoff v_∞). Then when the volatility of the state variable, $f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)$, is bounded, as is assumed in Assumption 4, U must also converge to v_∞ to prevent the U'' from converging to a constant. This establishes the

⁶The monotonicity assumption on the static Nash equilibrium payoff function, $v(X)$, (Assumption 3) plays a key role in ensuring the limit of the first derivative exists. It is possible for a bounded function to converge to a finite limit, but have a derivative that oscillates. This assumption guarantees that U is monotonic for large X , and prevents U' from oscillating. A similar assumption is not necessary in the bounded state space case, as the Lipschitz continuity of v is sufficient to ensure the limit of U' is well-defined.

⁷The boundedness of U ensures that U'' either converges to zero, or oscillates around zero.

boundary condition on U presented in Theorem 2,

$$\lim_{X \rightarrow \infty} U(X) = v_\infty \text{ and } \lim_{X \rightarrow -\infty} U(X) = v_{-\infty}$$

Given Assumption 4, any solution to the optimality equation must satisfy these boundary conditions. Showing that it is not possible for two different solutions U_1 and U_2 to both satisfy these boundary conditions concludes the proof that the optimality equation has a unique solution. This establishes the existence of a unique Markovian equilibrium.

The second step in proving the existence of a unique PPE is showing that there are no non-Markovian PPE, and as such, this unique Markovian equilibrium is the unique PPE. The intuition behind this result, and its relationship with the continuous time literature, is discussed in depth following an example to illustrate Theorem 2.

4.2.1 Example to illustrate Theorem 2

I next illustrate that the persistent investment example satisfies the assumptions for Theorem 2 and has a unique PPE, which is Markovian.

Example 3. *In the persistent investment model, the state variable evolves according to:*

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

The drift of the state variable is $\theta (a_t - X_t)$, which grows linearly as $|X|$ approaches infinity. The volatility of the state variable is σ , which is bounded uniformly with respect to X , since it is constant. This example satisfies Assumption 4. As characterized in Section 2.1, the unique stage game Nash equilibrium is for the firm to choose zero investment, and consumers to choose a purchase level of $\bar{b} = 3$ when $X > 3/\lambda$. This ensures the monotonicity assumption of the static Nash payoffs, Assumption 3, is satisfied. Section 4.1 established that this example satisfies the other required assumptions for Theorem 2. Therefore, this example has a unique Markovian equilibrium that satisfies the following boundary conditions:

$$\begin{aligned} \lim_{X \rightarrow \infty} U(X) &= 3 \text{ and } \lim_{X \rightarrow -\infty} U(X) = 0 \\ \lim_{X \rightarrow \infty} XU'(X) &= \lim_{X \rightarrow -\infty} XU'(X) = 0 \end{aligned}$$

In this equilibrium, the firm's action converges to the static Nash best response of zero investment as the product quality becomes large, and the firm receives an equilibrium payoff of 3, the highest feasible payoff for the firm.

4.2.2 Intertemporal Incentives in Stochastic Games

The fact that the unique PPE is Markovian yields an important insight on the role action persistence plays in generating intertemporal incentives. In a stochastic game of imperfect monitoring, intertemporal incentives can be generated through two potential channels: (1) conditioning future equilibrium play on the public signal and (2) the effect of the current action on the set of future feasible payoffs via the state variable. Equilibrium play in a

Markovian equilibrium is completely specified by the current value of the state variable, and the public signal is ignored. As such, the sole source of intertemporal incentives in a Markovian equilibrium is from the impact that the current action has on the set of future feasible payoffs. When this equilibrium is unique, it precludes the existence of any equilibria that use the public signal to generate intertemporal incentives via continuation play. *As such, the ability to generate effective intertemporal incentives in a stochastic game of imperfect monitoring stems entirely from the effect of the current action on the set of future feasible payoffs via the state variable.*

This insight relates to equilibrium degeneracy results from the continuous time repeated games literature, which show that it is not possible to provide effective intertemporal incentives in an imperfect monitoring game between a long-run and short-run player. In a standard repeated game, conditioning future equilibrium play on the public signal is the only potential channel for generating intertemporal incentives, and, as is the case in the stochastic game, this is not an effective channel for incentive provision. Thus, the introduction of action persistence creates an essential avenue for intertemporal incentive provision, and the ability to create effective intertemporal incentives in the stochastic game is entirely due to this additional this channel.

I next comment on the features of a stochastic game that allow Brownian information to be used to effectively provide intertemporal incentives. First consider the intuition behind why coordinating future equilibrium play is not an effective channel for incentive provision in a game between a short-run and long-run player. As discussed following lemma 1, Brownian information must be used linearly. Given this restriction, consider what happens when the long-run player's continuation value is at its upper bound. Using Brownian information linearly in a direction that is non-tangential to the boundary of the equilibrium payoff set will result in the continuation value escaping its upper bound with positive probability, a contradiction. Using Brownian information linearly in a tangential manner is precluded by the presence of myopic short-run players. Thus, it is not possible to linearly use Brownian information to structure incentives at the long-run player's highest continuation payoff, and both the long-run player and short-run player will play a myopic best response at this point. But this is precisely the definition of a static Nash equilibrium, and therefore long-run player's highest continuation payoff is bounded above by the highest static Nash equilibrium payoff.

Now consider the introduction of action persistence. The firm's incentive constraint and the evolution of the continuation value is still linear with respect to the Brownian information, as captured by the process $(\beta_t)_{t \geq 0}$ that governs incentives and the volatility of the continuation value. However, it is possible to characterize this process in a manner that depends on the state variable, and ensures the continuation value has zero volatility at states that yield the highest continuation payoff. This prevents the continuation value from escaping its upper bound with positive probability. Note this also implies that the long-run player must be playing a myopic best response at the state that yields the highest continuation payoff. However, for other states, it is possible to structure incentives such that the firm plays a non-myopic action.

Recall the characterization of $(\beta_t)_{t \geq 0}$ in Theorem 1, where β_t is the volatility of the

agency's continuation value and also governs the agency's incentive constraint:

$$\beta_{1t} = U'(X_t)f_1(\bar{b}_t, X_t)$$

Suppose an interior state yields the highest continuation value. Then the slope of the continuation value is zero at this point, $U'(X_t) = 0$, which ensures the volatility of the continuation value is zero at its upper bound. Suppose that a boundary of the state space yields the highest continuation value. Then the boundary conditions characterized in Theorem 2 ensure that β_{1t} converges to zero at this boundary, and therefore the continuation value has zero volatility at its upper bound. However, for states that do not yield the highest or lowest continuation value, it is possible for $|\beta_{1t}| > 0$, which allows intertemporal to be structured in a manner such that the agency plays a nonmyopic action.

4.3 Stochastic Games without Action Persistence

Suppose that the state variable evolves independently of the public signal. This removes the link between the agency's action and evolution of the state variable, creating a stochastic game without action persistence. This section establishes that the unique PPE in this setting is one in which the firm acts myopically and plays the static Nash equilibrium of the current stage game.

Given an initial state X_0 , define the average discounted payoff from playing the static Nash equilibrium action profile in each state as:

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

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

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

where these expectations are taken with respect to the state variable, given that the state evolves according to the measure generated by the static Nash equilibrium action profile $S^*(X, 0) = (a(X, 0), \bar{b}(X, 0))_{X \in \Xi}$. This expression defines the stochastic game payoff that the agency will earn if it myopically optimizes flow payoffs each instant. It is important to note that repeated play of the static Nash action profile is not necessarily an equilibrium strategy profile of the stochastic game; this is a general property of stochastic games.

The following Corollary establishes that in a stochastic game without action persistence, repeated play of the static Nash action profile is an equilibrium strategy profile; in fact, it is the unique equilibrium strategy profile, and yields the firm an equilibrium payoff of $V_{NE}(X_0)$. The corollary also directly characterizes $V_{NE}(X_0)$ as the solution to a *non-stochastic* second order differential equation (recall that $V_{NE}(X_0)$ is an expectation with respect to the state variable).

Corollary 1. *Suppose that the transition function of the state variable is independent of the public signal, $f_1 = 0$, for all X , and suppose Assumptions 1 and 2 hold. Then, given*

an initial state X_0 , there is a unique perfect public equilibrium characterized by the unique bounded $U(X)$ solution to:

$$U''(X) = \frac{2r [U(X) - g(a, \bar{b}, X) - f_2(\bar{b}, X)U'(X)]}{\sigma_X^2(X)}$$

This solution characterizes a Markovian 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 $(a_t, \bar{b}_t)_{t \geq 0}$ uniquely specified by the static Nash equilibrium action profile, for each X :

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

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

This result is a direct application of Theorem 1. The equilibrium characterization in Theorem 1 is used to characterize the optimality equation and incentive constraint for the firm when $f_1 = 0$. In the absence of a link between the agency's action and the state variable, the firm plays a static best response each instant.

It is interesting to note that the uniqueness result in a game without action persistence stems directly from the existence characterization in Theorem 1, and does not require the additional assumptions (Assumptions 3 and 4) necessary for Theorem 2. To see why, note that the incentive constraint is independent of the solution U to the optimality equation. As such, any solution to the optimality equation yields the same equilibrium action profile - namely, the action profile in which the agency plays a static best response each instant. Thus, all Markovian equilibrium action profiles induce the same measure over the path of the state variable. When a static best response is played each period, the continuation value in any Markovian equilibrium evolves according to the expected payoff from playing the static Nash equilibrium action profile in each state:

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

where the expectation is taken with respect to the measure over the state variable. Given that this measure over the state variable is the same in any Markovian equilibrium, any solution to the optimality equation must yield the same continuation values $(U(X_t))_{t \geq 0} = (W_{NE}(X_t))_{t \geq 0}$ for all $X \in \Xi$. Therefore, this solution must be unique. The solution to the optimality

equation in Corollary 1 can be used to explicitly characterize the expectation $V_{NE}(X_0)$ and $W_{NE}(X_t)$.

The existence of a solution to the optimality equation requires that the volatility of the state variable is bounded away from zero at interior points (Assumption 1). Note that when the state variable does not depend on the public signal, this assumption requires $\sigma_X^2(X_t)$ to be bounded away from zero to ensure that the state variable evolves stochastically.

Although Assumption 4 and 3 are not required to establish the existence of a unique PPE in the stochastic game without action persistence, if they do hold, then Theorem 2 can be used to characterize continuation payoffs at the boundary points of the state space. As established in Theorem 2, as the state variable approaches its boundary points (when the state space is unbounded, positive and negative infinity), then continuation payoffs approach the static Nash equilibrium payoff at the boundary points. Thus, the expected payoff from repeated play of the static game Nash equilibrium action profile approaches the payoff of the static Nash equilibrium at the boundary point of the state space.

Corollary 2. *Suppose Assumptions 3 and 4 also hold. Then if the state space is bounded,*

$$\lim_{X \rightarrow \bar{X}} W_{NE}(X) = v(\bar{X}) \text{ and } \lim_{X \rightarrow \underline{X}} W_{NE}(X) = v(\underline{X})$$

and if the state space is unbounded

$$\lim_{X \rightarrow \infty} W_{NE}(X) = v_\infty \text{ and } \lim_{X \rightarrow -\infty} W_{NE}(X) = v_{-\infty}$$

This result is a direct implication of Theorem 2.

The degeneracy result of this section relates to the discussion of intertemporal incentive provision in Section 4.2.2. When the state variable evolves independently of the agency's action, this removes the second channel for intertemporal incentives that links the agency's action to the set of future feasible payoffs. Therefore, the only potential channel for intertemporal incentive provision is the coordination of equilibrium play, and, as discussed in 4.2.2, it is not possible to effectively structure incentives via this channel. The agency plays myopically in the unique equilibrium of this setting.

4.4 Properties of Equilibrium Payoffs

This section uses Theorem 1 and 2 to describe several interesting properties relating the agency's equilibrium payoffs to the structure of the underlying stage game. The main results of this section are to provide an upper and lower bound on the PPE payoffs of the stochastic game across all states and characterize how the PPE payoff of the agency varies with the state variable

First, I examine properties of the highest and lowest PPE payoffs across all states, represented by \bar{W} and \underline{W} , respectively. Incentives for agency to choose a non-myopic action in the current period are provided through the link between the agency's action, the transition of the state variable and future feasible payoffs. When the continuation value is at its upper or lower bound, then the continuation value must have zero volatility so as not to escape

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

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

Theorem 3. $\bar{W} \leq \bar{v}^*$ and $\underline{W} \geq \underline{v}^*$

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

This result has an intuitive relation to reputation models of incomplete information. Recall that the state variable in this model is the consumers' beliefs about the firm's type. Therefore, $X = 0$ and $X = 1$ are absorbing states. When $X = 0$, consumers place probability one on the firm being a normal type, and it is not possible for the firm to earn payoffs above the static Nash equilibrium payoff of the game with complete information. Provided the firm's payoffs are increasing in the belief it is a behavioral type, the lowest PPE payoff occurs at $X = 0$ and is equal to the Nash equilibrium payoff of the complete information game, \underline{v}^* . Conditional on the firm being a normal type, the transition function governing beliefs has negative drift, and beliefs converge to the absorbing state $X = 0$. This captures the temporary reputation phenomenon associated with reputation models of incomplete information. Once consumers learn the firm's type, it is not possible to return to a state $X > 0$. Note that although $X = 1$ is also an absorbing state, but conditional on the firm being the normal type, the state variable never converges to $X = 1$.

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

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

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

1. *Suppose $v(X)$ is increasing (decreasing) in X . Then $U(X)$ is also increasing (decreasing) in X . The state that yields the highest static Nash payoff also yields the highest PPE payoff; likewise, the state that yields the lowest static Nash payoff also yields the lowest PPE payoff.*
2. *Suppose $v(X)$ has a unique interior maximum X^* , and v is monotonically increasing (decreasing) for $X < X^*$ ($X > X^*$). Then $U(X)$ has a unique interior maximum at X^* , and U is monotonically increasing (decreasing) for $X < X^*$ ($X > X^*$). The state that yields the highest static Nash payoff also yields the highest PPE payoff, whereas and the state that yields the lowest PPE payoff is a boundary point.*
3. *Suppose $v(X)$ has a unique interior minimum X^* , and v is monotonically decreasing (increasing) for $X < X^*$ ($X > X^*$). Then $U(X)$ has a unique interior minimum at X^* , and U is monotonically decreasing (increasing) for $X < X^*$ ($X > X^*$). The state that yields the lowest static Nash payoff also yields the lowest PPE payoff, whereas and the state that yields the highest PPE payoff is a boundary point.*

If the stage Nash equilibrium payoff of the agency is increasing in the state variable, then the PPE payoff to the agency is also increasing in the the state variable. The state that yields the highest PPE payoff to the agency corresponds to the state that yields the highest stage Nash equilibrium payoff, and the state that yields the lowest PPE payoff to the agency corresponds to the state that yields the lowest PPE payoff. Stochastic games differ from standard repeated games in that it is not necessarily possible to achieve an equilibrium payoff of the stochastic game that is equal to the stage Nash equilibrium payoff. Thus, for intermediate values of the state variable, the PPE payoff may lie above or below the static Nash equilibrium payoff. A symmetric result holds if the stage Nash equilibrium payoff of the agency is decreasing in the state variable.

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

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

5 Conclusion

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

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

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

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

Additionally, examining other classes of stochastic games, such as games between two long-run players whose actions jointly determine a stock variable, or games with different

information structures governing the state transitions, remain unexplored.

References

- BOARD, S. AND M. MEYER-TER-VEHN (2011): “Reputation for Quality,” Tech. rep., UCLA.
- CRIPPS, M. W., G. J. MAILATH, AND L. SAMUELSON (2007): “Disappearing private reputations in long-run relationships,” *Journal of Economic Theory*, 134, 287 – 316.
- FAINGOLD, E. (2008): “Building a Reputation under Frequent Decisions,” Tech. rep.
- FAINGOLD, E. AND Y. SANNIKOV (2011): “Reputation in Continuous-Time Games,” *Econometrica*, 79, 773–876.
- FUDENBERG, D. AND D. LEVINE (2007): “Continuous Time Limits of Repeated Games with Imperfect Public Monitoring,” *Review of Economic Dynamics*, 10, 173–192.
- FUDENBERG, D. AND D. K. LEVINE (1992): “Maintaining a Reputation When Strategies Are Imperfectly Observed,” *Review of Economic Studies*, 59, 561–79.
- (1994): “Efficiency and Observability with Long-Run and Short-Run Players,” *Journal of Economic Theory*, 62, 103–135.
- KLEIN, B. AND K. B. LEFFLER (1981): “The Role of Market Forces in Assuring Contractual Performance,” *Journal of Political Economy*, 89, 615–41.
- KREPS, D. M., P. MILGROM, J. ROBERTS, AND R. WILSON (1982): “Rational cooperation in the finitely repeated prisoners’ dilemma,” *Journal of Economic Theory*, 27, 245–252.
- MAILATH, G. J. AND L. SAMUELSON (2001): “Who Wants a Good Reputation?” *Review of Economic Studies*, 68, 415–41.
- SANNIKOV, Y. (2007): “Games with Imperfectly Observable Actions in Continuous Time,” *Econometrica*, 75, 1285–1329.
- SANNIKOV, Y. AND A. SKRZYPACZ (2007): “Impossibility of Collusion under Imperfect Monitoring with Flexible Production,” *American Economic Review*, 97, 1794–1823.
- (2010): “The Role of Information in Repeated Games With Frequent Actions,” *Econometrica*, 78, 847–882.

6 Appendix

6.1 Proof of Lemma 1

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

6.1.1 Evolution of the continuation value

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

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

Lemma 2. *The average discounted payoff at time t , $V_t(S)$, is a martingale.*

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

Lemma 3. *In any PPE, the continuation value evolves according to the stochastic differential equation*

$$dW_t(S) = r (W_t(S) - g(a_t, \bar{b}_t, X_t)) dt + r\beta_{1t} [dY_t - \mu_Y(a_t, \bar{b}_t)dt] + r\beta_{2t}\sigma_X(X_t)dZ_t^X$$

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:

$$dV_t(S) = re^{-rt} [\beta_{1t}\sigma_Y dZ_t^Y + \beta_{2t}\sigma_X(X_t)dZ_t^X]$$

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

$$\begin{aligned}
& re^{-rt}g(a_t, \bar{b}_t, X_t)dt - re^{-rt}W_t(S)dt + e^{-rt}dW_t(S) = re^{-rt} [\beta_{1t}\sigma_Y dZ_t^Y + \beta_{2t}\sigma_X(X_t)dZ_t^X] \\
\Rightarrow & dW_t(S) = rW_t(S)dt - rg(a_t, \bar{b}_t, X_t)dt + r\beta_{1t}\sigma_Y dZ_t^Y + r\beta_{2t}\sigma_X(X_t)dZ_t^X \\
\Rightarrow & dW_t(S) = r(W_t(S) - g(a_t, \bar{b}_t, X_t))dt + r\beta_{1t} [dY_t - \mu_Y(a_t, \bar{b}_t)dt] + r\beta_{2t}\sigma_X(X_t)dZ_t^X
\end{aligned}$$

Q.E.D.

6.1.2 Sequential Rationality

Lemma 4. *A strategy $(a_t)_{t \geq 0}$ is sequentially rational for the agency if, given $(\beta_t)_{t \geq 0}$, for all t :*

$$a_t \in \arg \max g(a', \bar{b}_t, X_t) + \beta_{1t}\mu_Y(a', \bar{b}_t)$$

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

At time τ , the state variable is equal to X_τ . Action a_τ will induce $dY_\tau = \mu_Y(a_\tau, \bar{b}_\tau)d\tau + \sigma_Y dZ_\tau^Y$ whereas action \tilde{a}_τ will induce $dY_\tau = \mu_Y(\tilde{a}_\tau, \bar{b}_\tau)d\tau + \sigma_Y dZ_\tau^Y$. Let \tilde{V}_τ be the expected average payoff conditional on info at time τ when follows \tilde{a} up to τ and a afterwards, and let W_τ be the continuation value when the firm follows strategy $(a_t)_{t \geq 0}$ starting at time τ .

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

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

$$\begin{aligned}
dW_\tau &= r(W_\tau - g(a_\tau, \bar{b}_\tau, X_\tau))d\tau \\
&\quad + r\beta_{1\tau} [dY_\tau - \mu_Y(a_\tau, \bar{b}_\tau)d\tau] \\
&\quad + r\beta_{2\tau}\sigma_X(\bar{b}_\tau, X_\tau)dZ_\tau^X
\end{aligned}$$

when firm switches strategies at time τ .

$$\begin{aligned}
d\tilde{V}_\tau &= re^{-r\tau} [g(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - W_\tau] d\tau + e^{-r\tau}dW_\tau \\
&= re^{-r\tau} [g(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - g(a_\tau, \bar{b}_\tau, X_\tau)] d\tau \\
&\quad + re^{-r\tau}\beta_{1\tau} [dY_\tau - \mu_Y(a_\tau, \bar{b}_\tau)d\tau] \\
&\quad + re^{-r\tau}\beta_{2\tau}\sigma_X(\bar{b}_\tau, X_\tau)dZ_\tau^X \\
&= re^{-r\tau} \left\{ \begin{array}{l} [g(\tilde{a}_\tau, \bar{b}_\tau, X_\tau) - g(a_\tau, \bar{b}_\tau, X_\tau)] d\tau \\ + re^{-r\tau}r\beta_{1\tau} [\mu_Y(\tilde{a}_\tau, \bar{b}_\tau)d\tau - \mu_Y(a_\tau, \bar{b}_\tau)d\tau + \sigma_Y dZ_\tau^Y] \\ + re^{-r\tau}\beta_{2\tau}\sigma_X(\bar{b}_\tau, X_\tau)dZ_\tau^X \end{array} \right\}
\end{aligned}$$

There are two components to this strategy change: how it affects the immediate flow payoff and how it affects future public signals Y_t , which impacts the continuation value (captured in process β). The profile $(\tilde{a}_t, \bar{b}_t)_{t \geq 0}$ yields the firm a payoff of:

$$\begin{aligned}\tilde{W}_0 &= E_0 [\tilde{V}_\infty] = E_0 \left[\tilde{V}_0 + \int_0^\infty d\tilde{V}_t \right] \\ &= W_0 + E_0 \left[r \int_0^\infty e^{-rt} \left\{ \begin{array}{l} g(\tilde{a}_t, \bar{b}_t, X_t) + \beta_{1t} \mu_Y(\tilde{a}_t, \bar{b}_t) \\ -g(a_t, \bar{b}_t, X_t) - \beta_t \mu_Y(a_t, \bar{b}_t) \end{array} \right\} dt \right]\end{aligned}$$

If

$$g(a_t, \bar{b}_t, X_t) + \beta_{1t} \mu_Y(a_t, \bar{b}_t) \geq g(\tilde{a}_t, \bar{b}_t, X_t) + \beta_{1t} \mu_Y(\tilde{a}_t, \bar{b}_t)$$

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

Q.E.D.

6.2 Proof of Theorem 1: Characterization of Markovian Equilibrium

Theorem 5. *Suppose Assumptions 1 and 2 hold. Then given X_0 , any solution $U(X)$ to the second order differential equation,*

$$U''(X) = \frac{2r [U(X) - g(a, \bar{b}, X)]}{f_1^2(\bar{b}, X) \sigma_Y^2 + \sigma_X^2(X)} - \frac{2 [f_1(\bar{b}, X) \mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] U'(X)}{f_1^2(\bar{b}, X) \sigma_Y^2 + \sigma_X^2(X)}$$

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

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

$$S^*(X, U'(X) f_1(\bar{b}, X)) = \left\{ \begin{array}{l} a = \arg \max_{a'} r g(a', \bar{b}, X) + U'(X) f_1(\bar{b}, X) \mu(a', \bar{b}) \\ \bar{b} = \bar{\mathcal{B}}(a, X) \end{array} \right\}$$

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

6.2.1 Proof of form of Optimality Equation:

Lemma 5. *If a Markovian equilibrium exists, it takes the following form:*

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

$$U''(X) = \frac{2r(U(X) - g(a, \bar{b}, X))}{[f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)]} - \frac{2[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]}{[f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)]}U'(X)$$

2. Given solution $U(X)$, the process governing incentives for the agency is characterized by:

$$\begin{aligned} r\beta_1 &= U'(X)f_1(\bar{b}, X) \\ r\beta_2 &= U'(X) \end{aligned}$$

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

$$\begin{aligned} dW_t &= U'(X_t)dX_t + \frac{1}{2}U''(X_t)[f_1(\bar{b}_t, X_t)^2\sigma_Y^2 + \sigma_X^2(X_t)]dt \\ &= U'(X_t)[f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t) + f_2(\bar{b}_t, X_t)]dt \\ &\quad + \frac{1}{2}U''(X_t)[f_1(\bar{b}_t, X_t)^2\sigma_Y^2 + \sigma_X^2(X_t)]dt \\ &\quad + U'(X_t)[f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t)dZ_t^X] \end{aligned}$$

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

$$dW_t = r(W_t - g(a_t, \bar{b}_t, X_t))dt + r\beta_{1t}\sigma_Y dZ_t^Y + r\beta_{2t}\sigma_X(X_t)dZ_t^X$$

We can match the drift of these two characterizations to obtain the optimality equation:

$$\begin{aligned} r(U(X) - g(a, \bar{b}, X)) &= U'(X)[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] + \frac{1}{2}U''(X)[f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)] \\ \Rightarrow U''(X) &= \frac{2r(U(X) - g(a, \bar{b}, X))}{[f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)]} - \frac{2[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]}{[f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)]}U'(X) \end{aligned}$$

for strategy profile $(a, \bar{b}) = (a(X), \bar{b}(X))$, which is a second order non-homogenous differential equation. Matching the volatility characterizes the process governing incentives:

$$\begin{aligned} r\beta_{1t} &= U'(X_t)f_1(\bar{b}_t, X_t) \\ r\beta_{2t} &= U'(X_t) \end{aligned}$$

Plugging these into the constraints for sequential rationality yields

$$\begin{aligned} S^*(X, U'(X)f_1(\bar{b}_t, X_t)) &= (a, \bar{b}) \text{ s.t.} \\ a &= \arg \max rg(a', \bar{b}, X) + U'(X)f_1(\bar{b}, X)\mu(a, \bar{b}) \\ \bar{b} &= \bar{\mathcal{B}}(a, X) \end{aligned}$$

which are unique by Assumption 2.

Q.E.D.

6.2.2 Prove existence of bounded solution to optimality equation

Linear Growth

Lemma 6. *The optimality equation has linear growth. Suppose Assumption 1 holds. For all $M > 0$ and compact intervals $I \subset \Xi$, there exists a $K_I > 0$ such that for all $X \in I$, $(a, \bar{b}) \in A \times B$, $u \in [-M, M]$ and $u' \in R$,*

$$u'' = \frac{r [u - g(a, \bar{b}, X)] - [f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] u'}{[f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X(X)^2]} \leq K (1 + |u'|)$$

Follows directly from the fact that $u \in [-M, M]$, g, f_1, μ_Y and f_2 are Lipschitz continuous, the bound on $f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X(X)$, and $X \in I$.

Q.E.D.

Existence for Unbounded support

Theorem 6. *The optimality equation has at least one solution $U \in C^2(R)$ that lies in the range of feasible payoffs for the agency i.e. for all $X \in R$*

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

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

Theorem 7. *Let $\alpha, \beta \in R$ be such that $\alpha \leq \beta$, $E = \{(t, u, v) \in R^3 | \alpha \leq u \leq \beta\}$ and $f : E \rightarrow R$ be continuous. Assume that α and β are such that for all $t \in R$*

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

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

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

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

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

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

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

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

$$\begin{aligned} h(X, \underline{g}, 0) &= \frac{2r}{[f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)]} (\underline{g} - g(a, \bar{b}, X)) \leq 0 \\ h(X, \bar{g}, 0) &= \frac{2r}{[f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)]} (\bar{g} - g(a, \bar{b}, X)) \geq 0 \end{aligned}$$

for all X . For any bounded interval I , define

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

where $\sigma_I = \inf_{\bar{b} \in B, X \in I} [f_1(\bar{b}_t, X_t)^2 \sigma_Y^2 + \sigma_X^2(X_t)]$ which is positive by assumption. and $\mu_I = \sup_{\bar{b} \in B, X \in I} f_1(\bar{b}, X) \mu_Y(a, \bar{b}) + f_2(\bar{b}, X)$, which are well-defined given $f_1, f_2, \mu_Y, \sigma_Y$ and σ_X are Lipschitz continuous and B is compact. Note

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

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

$$\begin{aligned} |h(X, u, v)| &= \left| \frac{2r}{f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)} (u - g(a(X), \bar{b}(X), X)) \right. \\ &\quad \left. - \frac{2[f_1(\bar{b}, X) \mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]}{f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)} v \right| \\ &\leq \phi_I(|v|) \end{aligned}$$

Additionally, $h(X_t, U(X_t), U'(X_t))$ is continuous given that $f_1, f_2, \mu_Y, \sigma_Y$ and σ_X are Lipschitz continuous and $g(a(X), \bar{b}(X), X)$ is continuous. Thus, $h(X, U(X), U'(X))$ has at least one solution on $U \in C^2(R)$ such that for all $X \in R$,

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

Q.E.D.

Existence for Bounded support

Theorem 8. *The optimality equation has at least one solution $U \in C^2(R)$ that lies in the range of feasible payoffs for the agency i.e. for all $X \in \Xi$*

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

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

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

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

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

Lemma 7. *Let $E = \{(t, u, v) \in \Xi \times R^2\}$ and $f : E \rightarrow R$ be continuous. Assume that for any interval $I \subset \Xi$, there exists a $K_I > 0$ such that for all $t \in I$, $(u, v) \in R^2$ with $\alpha \leq u \leq \beta$,*

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

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

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

Consider the optimality equation $h(X, U(X), U'(X))$. Let $\bar{g} = \sup g(a, \bar{b}, X)$ and $\underline{g} = \inf g(a, \bar{b}, X)$, which are well defined since g is bounded. By 6, for any bounded interval I and $u \in [\underline{g}, \bar{g}]$, there exists a K_I such that

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

Additionally, $h(X_t, U(X_t), U'(X_t))$ is continuous given that $f_1, f_2, \mu_Y, \sigma_Y$ and σ_X are Lipschitz continuous and $g(a(X), \bar{b}(X), X)$ is continuous. Let $\alpha = \underline{g}$ and $\beta = \bar{g}$. Then $h(X, U(X), U'(X))$ has at least one solution on $U \in C^2(R)$ such that for all $X \in \Xi$,

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

Q.E.D.

6.2.3 Construct a Markovian equilibrium

Suppose the state variable initially starts at X_0 and U is a bounded solution to the optimality equation. The action profile satisfying $(a, \bar{b}) = S^*(X, U'(X)f_1(\bar{b}, X))$ is unique and Lipschitz continuous in X and U . Thus, given X_0, U and $(a_t, \bar{b}_t)_{t \geq 0} = [S^*(X_t, U'(X_t)f_1(\bar{b}_t, X_t))]_{t \geq 0}$, the state variable uniquely evolves according to the stochastic differential equation

$$dX_t = [f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t) + f_2(\bar{b}_t, X_t)] dt + f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t)dZ_t^X$$

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

$$\begin{aligned} dU(X_t) &= U'(X_t) [f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t) + f_2(\bar{b}_t, X_t)] dt \\ &\quad + \frac{1}{2}U''(X_t) [f_1(\bar{b}_t, X_t)^2\sigma_Y^2 + \sigma_X^2(X_t)] dt \\ &\quad + U'(X_t) [f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t)dZ_t^X] \\ &= r(U(X_t) - g(a_t, \bar{b}_t, X_t)) + U'(X_t) [f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t)dZ_t^X] \end{aligned}$$

this process satisfies the conditions for the continuation value in a PPE characterized in Lemma 1. Additionally, $(a_t, \bar{b}_t)_{t \geq 0}$ satisfies the condition for sequential rationality given process $(\beta_t)_{t \geq 0} = (U'(X_t)f_1(\bar{b}_t, X_t), U'(X_t))_{t \geq 0}$. Thus, the strategy profile $(a_t, \bar{b}_t)_{t \geq 0}$ is a PPE yielding equilibrium payoff $U(X_0)$.

Q.E.D.

6.3 Proof of Theorem 2: Characterize Unique Markovian Equilibrium

Theorem 9. *Suppose 4, 2 and 1 hold. Then, for each $X_0 \in \Xi$, there exists a unique perfect public equilibrium with continuation values characterized by the unique bounded solution U of the optimality equation, yielding equilibrium payoff $U(X_0)$.*

1. If $\Xi = [\underline{X}, \bar{X}]$ then the solution satisfies the following boundary conditions:

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

2. If $\Xi = \mathcal{R}$ then the solution satisfies the following boundary conditions:

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

6.3.1 Boundary Conditions

Boundary Conditions for Unbounded Support

Theorem 10. *Any bounded solution U of the optimality equation satisfies the following boundary conditions*

$$\begin{aligned} \lim_{X \rightarrow \infty} U(X) &= v_\infty \text{ and } \lim_{X \rightarrow -\infty} U(X) = v_{-\infty} \\ \lim_{X \rightarrow \infty} XU'(X) &= \lim_{X \rightarrow -\infty} XU'(X) = 0 \\ \lim_{X \rightarrow \infty} (f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)) U''(X) &= 0 \\ \lim_{X \rightarrow -\infty} (f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)) U''(X) &= 0 \end{aligned}$$

The proof proceeds by a series of lemmas.

Lemma 8. *If U is a bounded solution of the optimality equation, then $\lim_{X \rightarrow \infty} U(X)$ and $\lim_{X \rightarrow -\infty} U(X)$ are well-defined.*

Proof: Assume Assumption 3. This guarantees $\lim_{X \rightarrow \infty} v(X)$ exists. U is bounded, and therefore $\lim_{X \rightarrow \infty} \sup U(X)$ and $\lim_{X \rightarrow \infty} \inf U(X)$ are well-defined. Suppose $\lim_{X \rightarrow \infty} \sup U(X) \neq \lim_{X \rightarrow \infty} \inf U(X)$. Then there exists a sequence $(X_n)_{n \in \mathbb{N}}$ that correspond to local maxima of U , so $U'(X_n) = 0$ and $U''(X_n) \leq 0$. Given the incentives for the agency, a stage nash equilibria is played when $U'(X) = 0$, yielding flow payoff $v(X)$. From the optimality equation, this implies $v(X_n) \geq U(X_n)$. Likewise, there exists a sequence $(X_m)_{m \in \mathbb{N}}$ that correspond to local minima of U , so $U'(X_m) = 0$ and $U''(X_m) \geq 0$. This implies $v(X_m) \leq U(X_m)$. Thus, $\lim_{X \rightarrow \infty} \sup v(X) \neq \lim_{X \rightarrow \infty} \inf v(X)$. This is a contradiction, as $\lim_{X \rightarrow \infty} v(X)$ is well-defined. Thus, $\lim_{X \rightarrow \infty} U(X)$ exists. The case of $\lim_{X \rightarrow -\infty} U(X)$ is similar. Q.E.D.

Lemma 9. *If $U(X)$ is a bounded solution of the optimality equation, then there exists a δ such that for $|X| > \delta$, $U(X)$ is monotonic.*

Proof: Assume Assumption 3. Suppose that there does not exist a δ such that for $X > \delta$, U is monotonic. Then for all $\delta > 0$, there exists a $X_n > \delta$ that corresponds to a local maxima of U , so $U'(X_n) = 0$ and $U''(X_n) \leq 0$ and there exists a $X_m > \delta$ that corresponds to a local minima of U , so $U'(X_m) = 0$ and $U''(X_m) \geq 0$, by the continuity of U . Given the incentives for the agency, a stage nash equilibria is played when $U'(X) = 0$, yielding flow payoff $v(X)$. From the optimality equation, this implies $v(X_n) \geq U(X_n)$ at the maximum and $v(X_m) \leq U(X_m)$ at the minimum. Thus, the oscillation of $v(X)$ is at least as large as the oscillation of $U(X)$. This is a contradiction, as there exists a δ such that for $X > \delta$, $v(X)$ is monotonic. The case of $-X > \delta$ is similar. Q.E.D.

Lemma 10. *Given a function $f(X)$ that is $O(X)$ as $X \rightarrow \{-\infty, \infty\}$, then any bounded solution U of the optimality equation satisfies*

1.

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

2.

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

Note this is trivially satisfied if $f(X)$ is $O(1)$.

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

Q.E.D.

Lemma 11. *Given a function $f(X)$ that is $O(X)$ as $X \rightarrow \{-\infty, \infty\}$, then any bounded solution U of the optimality equation satisfies*

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

Assume Assumption 3. By Lemma 10, $\lim_{X \rightarrow \infty} \inf XU'(X) \leq 0 \leq \lim_{X \rightarrow \infty} \sup XU'(X)$. Suppose, without loss of generality, that $\lim_{X \rightarrow \infty} \sup XU'(X) > 0$. By Lemma 9, there exists a $\delta > 0$ such that $U(X)$ is monotonic for $X > \delta$. Then for $X > \delta$, $U'(X)$ doesn't change sign and therefore, $XU'(X)$ doesn't change sign. Therefore, if $\lim_{X \rightarrow \infty} \sup XU'(X) > 0$, then $\lim_{X \rightarrow \infty} \inf XU'(X) > 0$. This is a contradiction. Thus, $\lim_{X \rightarrow \infty} \sup XU'(X) = 0$. By similar reasoning, $\lim_{X \rightarrow \infty} \inf XU'(X) = 0$, and therefore $\lim_{X \rightarrow \infty} XU'(X) = 0$. Suppose $f(X)$ is $O(X)$. Then there exists an $M \in R$ and a $\delta_1 \in R$ such that when $X > \delta_1$, $|f(X)| \leq M|X|$. Thus, for $X > \delta_1$, $|f(X)U'(X)| \leq M|XU'(X)| \rightarrow 0$. The case for $\lim_{X \rightarrow -\infty} f(X)U'(X) = 0$ is analogous. Note this result also implies that

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

Q.E.D.

Lemma 12. *Let U be a bounded solution of the optimality equation. Then the limit of U converges to the limit of the stage Nash equilibrium payoffs as $X \rightarrow \{-\infty, \infty\}$*

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

Proof: Assume Assumption 2, 3, and 4. By 8, $\lim_{X \rightarrow \infty} U(X) = U_\infty$ exists. Suppose $U_\infty < v_\infty$, where v_∞ is the limit of the stage game Nash equilibrium payoff at positive infinity. The function f_1 is $O(X)$, and therefore by Lemma 11, $\lim_{X \rightarrow \infty} U'(X)f_1(\bar{b}, X) = 0$ and $S^*(X, U'(X)f_1(\bar{b}, X)) \rightarrow (a_\infty^N, \bar{b}_\infty^N)$ which is the stage Nash equilibrium as $X \rightarrow \infty$. Thus, $\lim_{X \rightarrow \infty} g(a(X), \bar{b}(X), X) = v_\infty$.

By Lemma 11 and the assumption that $(f_1\mu_Y + f_2)$ is $O(X)$

$$\lim_{X \rightarrow \infty} [f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] U'(X) = 0$$

By the assumption that $f_1\sigma_Y + \sigma_X$ is $O(1)$, there exists an $M > 0$ and a δ such that when $X > \delta_1$, then $f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X) \leq M$.

Plugging the above conditions in to the optimality equation yields

$$\begin{aligned}\limsup_{X \rightarrow \infty} U''(X) &= \limsup_{X \rightarrow \infty} \left[\frac{2r(U(X) - g(a, \bar{b}, X))}{f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)} - \frac{2[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]}{f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)} U'(X) \right] \\ &\leq \frac{2r(U_\infty - v_\infty)}{M} < 0\end{aligned}$$

which violates Lemma 10, and therefore U is unbounded. This is a contradiction. Thus, $U_\infty = v_\infty$. The proof for the other cases is analogous.

Q.E.D.

Lemma 13. *Any bounded solution U of the optimality equation satisfies*

$$\begin{aligned}\lim_{X \rightarrow \infty} |(f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)) U''(X)| &= 0 \\ \lim_{X \rightarrow -\infty} |(f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)) U''(X)| &= 0\end{aligned}$$

Note this also implies $U''(X) \rightarrow 0$.

Applying the squeeze theorem to the optimality equation yields

$$\begin{aligned}&\lim_{X \rightarrow \infty} |(f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)) U''(X)| \\ &= \lim_{X \rightarrow \infty} |2r(U(X) - g(a, \bar{b}, X)) - 2[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] U'(X)| \\ &= 0\end{aligned}$$

by applying Lemmas 11 and 12 and the assumption that $[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]$ is $O(X)$.

Q.E.D.

Boundary Conditions for Bounded Support

Theorem 11. *Any bounded solution U of the optimality equation satisfies the following boundary conditions:*

$$\begin{aligned} \lim_{X \rightarrow \bar{X}} U(X) &= v(\bar{X}) \text{ and } \lim_{X \rightarrow \underline{X}} U(X) = v(\underline{X}) \\ \lim_{X \rightarrow \bar{X}} (\bar{X} - X)U'(X) &= \lim_{X \rightarrow \underline{X}} (\underline{X} - X)U'(X) = 0 \\ \lim_{X \rightarrow \bar{X}} (f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)) U''(X) &= 0 \\ \lim_{X \rightarrow \underline{X}} (f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)) U''(X) &= 0 \end{aligned}$$

The proof proceeds by a series of lemmas.

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

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

Q.E.D.

Lemma 15. *Given a function $f(X)$ that is $O(a - X)$ as $X \rightarrow a \in \{\underline{X}, \bar{X}\}$, then any bounded solution U of the optimality equation satisfies*

1.

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

2.

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

Note this is trivially satisfied if $f(X)$ is $O(1)$.

1. Suppose $f(X)$ is $O(\bar{X} - X)$ as $X \rightarrow \bar{X}$ and $\liminf_{X \rightarrow \bar{X}} |f(X)U'(X)| > 0$. There exists an $M \in \mathbb{R}$ and a $\delta_1 > 0$ such that when $|\bar{X} - X| < \delta_1$, $|f(X)| \leq M|\bar{X} - X|$. Given $\liminf_{X \rightarrow \bar{X}} |f(X)U'(X)| > 0$, there exists a $\delta_2 \in \mathbb{R}$ and an $\varepsilon > 0$ such that when $|\bar{X} - X| < \delta_2$, $|f(X)U'(X)| > \varepsilon$. Take $\delta = \min\{\delta_1, \delta_2\}$. Then for $|\bar{X} - X| < \delta$, $|U'(X)| > \frac{\varepsilon}{|f(X)|} \geq \frac{\varepsilon}{M|\bar{X} - X|}$. Then the antiderivative of $\frac{\varepsilon}{M|\bar{X} - X|}$ is $\frac{\varepsilon}{M} \ln |\bar{X} - X|$ which diverges to $-\infty$ as $X \rightarrow \bar{X}$. This violates the boundedness of U . Therefore $\liminf_{X \rightarrow \bar{X}} f(X)U'(X) \leq 0$. The proof is analogous for the other cases.

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

Q.E.D.

Lemma 16. *Given a differentiable function $f(X)$ that is $O(X^* - X)$ as $X \rightarrow \{\underline{X}, \bar{X}\}$, then any bounded solution U of the optimality equation satisfies*

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

By Lemma 15, $\lim_{X \rightarrow \bar{X}} \inf(\bar{X} - X)U'(X) \leq 0 \leq \lim_{X \rightarrow \bar{X}} \sup(\bar{X} - X)U'(X)$. Suppose, without loss of generality, that $\lim_{X \rightarrow \bar{X}} \sup(\bar{X} - X)U'(X) > 0$. Then there exist constants k and K such that $(\bar{X} - X)U'$ crosses the interval (k, K) infinitely many times as X approaches \bar{X} . Additionally, there exists an $L > 0$ such that

$$\begin{aligned} |U''(X)| &= \left| \frac{2r [U(X) - g(a, \bar{b}, X)] - 2 [f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] U'(X)}{f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)} \right| \\ &\leq \left| \frac{L_1 - L_2 (\bar{X} - X) U'(X)}{(\bar{X} - X)^2} \right| \\ &\leq \left| \frac{L_1 - L_2 k}{(\bar{X} - X)^2} \right| = \frac{L}{(\bar{X} - X)^2} \end{aligned}$$

Therefore,

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

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

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

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

Q.E.D.

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

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

Suppose not. By 14, $\lim_{X \rightarrow \bar{X}} U(X) = U(\bar{X})$ exists. Suppose $U(\bar{X}) < v(\bar{X})$, where $v(\bar{X})$ is the limit of the stage game Nash equilibrium payoff at \bar{X} . The function f_1 is $O(\bar{X} - X)$, and therefore by Lemma 16, $\lim_{X \rightarrow \bar{X}} U'(X)f_1(\bar{b}, X) = 0$ and $S^*(X, U'(X)f_1(\bar{b}, X)) \rightarrow (a_{\bar{X}}^N, \bar{b}_{\bar{X}}^N)$ which is the stage Nash equilibrium as $X \rightarrow \bar{X}$. Thus, $\lim_{X \rightarrow \bar{X}} g(a(X), \bar{b}(X), X) = v(\bar{X})$.

By Lemma 16 and the assumption that $f_1\mu_Y + f_2$ is $O(\bar{X} - X)$

$$\lim_{X \rightarrow \infty} (f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X))U'(X) = 0$$

By the assumption that $1/(f_1\sigma_Y + \sigma_X)$ is $O(1/(\bar{X} - X))$, there exists an $M > 0$ and a δ such that when $|\bar{X} - X| < \delta_1$, then $1/(f_1\sigma_Y + \sigma_X) \leq M/(\bar{X} - X)$

Plugging the above conditions in to the optimality equation yields

$$\begin{aligned}\limsup_{X \rightarrow \bar{X}} U''(X) &= \limsup_{X \rightarrow \bar{X}} \left[\frac{2r(U(X) - g(a, \bar{b}, X))}{f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)} - \frac{2[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]}{f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)} U'(X) \right] \\ &\leq \frac{2r(U(\bar{X}) - v(\bar{X}))}{M(\bar{X} - X)^2} < 0\end{aligned}$$

which violates Lemma 15, and therefore U is unbounded. This is a contradiction. Thus, $U(\bar{X}) = v(\bar{X})$. The proof for the other cases is analogous.

Q.E.D.

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

$$\begin{aligned}\lim_{X \rightarrow \bar{X}} |(f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X))U''(X)| &= 0 \\ \lim_{X \rightarrow \underline{X}} |(f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X))U''(X)| &= 0\end{aligned}$$

Applying the squeeze theorem to the optimality equation yields

$$\begin{aligned}&\lim_{X \rightarrow \bar{X}} |(f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X))U''(X)| \\ &= \lim_{X \rightarrow \bar{X}} |2r(U(X) - g(a, \bar{b}, X)) - 2[f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)]U'(X)| \\ &= 0\end{aligned}$$

by applying Lemmas 16 and 17.

Q.E.D.

6.3.2 Uniqueness of Solution to Optimality Equation

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

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

Lemma 19. *If two bounded solutions of the optimality equation, U and V , satisfy $U(X_0) \leq V(X_0)$ and $U'(X_0) \leq V'(X_0)$, with at least one strict inequality, then $U(X) \leq V(X)$ and $U'(X) \leq V'(X)$ for all $X > X_0$. Similarly if $U(X_0) \leq V(X_0)$ and $U'(X_0) \geq V'(X_0)$, with at least one strict inequality, then $U(X) < V(X)$ and $U'(X) > V'(X)$ for all $X < X_0$.*

The proof is analogous to the proof in Faingold and Sannikov (2011), defining

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

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

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

for $\Xi = R$.

Q.E.D.

Lemma 20. *There exists a unique solution U to the optimality equation.*

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

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

Take an interval $X \in [X_1, X_2]$, and suppose $U(X) > V(X)$ for some $X \in [X_1, X_2]$. Let X^* be the point where $U - V$ is maximized, which is well-defined given U and V are continuous functions on a compact interval. Suppose the maximum occurs at an interior point $X^* \in (X_1, X_2)$. Then $U'(X^*) = V'(X^*)$. By Lemma 19, $U'(X) \geq V'(X)$ for all $X > X^*$, and this difference is strictly increasing, a contradiction. Suppose the maximum occurs at an endpoint, $X^* = X_2$, and let $U(X_2) - V(X_2) = M > 0$. Then it must be the case that $U'(X_2) \geq V'(X_2)$. By Lemma 19, $U'(X) \geq V'(X)$ for all $X > X_2$, and this difference is strictly increasing for $X > X_2$. But then for $\varepsilon < M$, there does not exist a δ such that $|U(X) - V(X)| < \varepsilon$ when $X > \delta$. This violates the boundary condition. The

argument is analogous if the maximum occurs at $X^* = X_1$. Thus, it is not possible to have $U(X) > V(X)$.

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

Q.E.D.

6.3.3 Uniqueness of Markovian Equilibrium in class of PPE

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

$$\begin{aligned} dW_t(S) &= r(W_t(S) - g(a_t, \bar{b}_t, X_t)) dt \\ &\quad + r\beta_{1t} [dY_t - \mu_Y(a_t, \bar{b}_t)dt] \\ &\quad + r\beta_{2t}\sigma_X(X_t)dZ_t^X \end{aligned}$$

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

$$\begin{aligned} dU(X_t) &= U'(X) [f_1(\bar{b}, X)\mu_Y(a, \bar{b}) + f_2(\bar{b}, X)] dt + \frac{1}{2}U''(X) [f_1(\bar{b}, X)^2\sigma_Y^2 + \sigma_X^2(X)] dt \\ &\quad + U'(X_t) [f_1(\bar{b}_t, X_t)\sigma_Y dZ_t^Y + \sigma_X(X_t)dZ_t^X] \end{aligned}$$

Define a process $D_t = W_t - U(X_t)$ with initial condition $D_0 = W_0 - U(X_0) > 0$. Then dD_t evolves with drift

$$\begin{aligned} &r [W_t - g(a_t, \bar{b}_t, X_t)] - U'(X_t) [f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t) + f_2(\bar{b}_t, X_t)] \\ &\quad - \frac{1}{2}U''(X_t) [f_1(\bar{b}_t, X_t)^2\sigma_Y^2 + \sigma_X^2(X_t)] \\ &= rD_t + r [U(X_t) - g(a_t, \bar{b}_t, X_t)] - U'(X_t) [f_1(\bar{b}_t, X_t)\mu_Y(a_t, \bar{b}_t) + f_2(\bar{b}_t, X_t)] \\ &\quad - \frac{1}{2}U''(X_t) [f_1(\bar{b}_t, X_t)^2\sigma_Y^2 + \sigma_X^2(X_t)] \\ &= rD_t + d(a_t, \bar{b}_t, X_t) \end{aligned}$$

and volatility

$$f(\bar{b}, X, \beta) = \frac{r\beta_{1t}\sigma_Y}{r\beta_{2t}\sigma_X(X_t)} - \frac{U'(X_t)f_1(\bar{b}_t, X_t)\sigma_Y}{U'(X_t)\sigma_X(X_t)}$$

Lemma 21. *For every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all (a, \bar{b}, X, β) satisfying the condition for sequential rationality*

$$\begin{aligned} a &\in \arg \max rg(a', \bar{b}, X) + r\beta_1\mu_Y(a, \bar{b}) \\ \bar{b} &\in \bar{\mathcal{B}}(a, X) \end{aligned}$$

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

Proof:

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

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

Let $|f(\bar{b}, X, \beta_1)| = 0$. Then $r\beta_1 = U'(X)f_1(\bar{b}, X)$ and for each X , there is a unique action (a, \bar{b}) profile that satisfies

$$\begin{aligned} a &\in \arg \max rg(a', \bar{b}, X) + U'(X)f_1(\bar{b}, X)\mu_Y(a, \bar{b}) \\ \bar{b} &\in \bar{\mathcal{B}}(a, X) \end{aligned}$$

This action profile corresponds to the actions played in a Markovian equilibrium, and therefore $d(a_t, \bar{b}_t, X_t) = 0$, by the optimality equation.

Step 2: Fix ε . Show if $d(a_t, \bar{b}_t, X_t) \leq -\varepsilon$, then there exists a $\delta > 0$ such that $|f(\bar{b}, X, \beta_1)| > \delta$ for all (a, \bar{b}, X, β) such that the sequential rationality condition is satisfied.

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

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

$U'(X)f_1(\bar{b}, X)$ and $U'(X)$ are bounded, so there exists an $M > 0$ and $m > 0$ such that $|f(\bar{b}, X, \beta)| > m$ for all $|\beta| > M$.

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

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

Show if $r(v_\infty - g_\infty(a, \bar{b})) \leq -\gamma$ then there exists a $\eta_2 > 0$ such that $|\beta| > \eta_2$. Let $\lim_{X \rightarrow \infty} g(a, \bar{b}, X) := g_\infty(a, \bar{b})$ be the limit flow payoff for actions (a, \bar{b}) , and $\lim_{X \rightarrow \infty} \bar{\mathcal{B}}(a, X) := \bar{\mathcal{B}}_\infty(a)$. These limits *exist*, since g and $\bar{\mathcal{B}}(a)$ are Lipschitz continuous and bounded. Consider the set $\Phi = (a, \bar{b}, \beta)$ that satisfies

$$\begin{aligned} a &\in \arg \max rg_\infty(a', \bar{b}) + r\beta_1\mu_Y(a, \bar{b}) \\ \bar{b} &\in \bar{\mathcal{B}}_\infty(a) \end{aligned}$$

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

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

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

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

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

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

Q.E.D.

This lemma implies that whenever the drift of D_t is less than $rD_t - \varepsilon$, the volatility is greater than δ . Take $\varepsilon = rD_0/4$ and suppose $D_t \geq D_0/2$. Then whenever the drift is less than $rD_t - \varepsilon > rD_0/2 - rD_0/4 = rD_0/4 > 0$, there exists a δ such that $|f(\bar{b}, X, \beta)| > \delta$. Thus, whenever $D_t \geq D_0/2 > 0$, it has either positive drift or positive volatility, and grows arbitrarily large with positive probability. This is a contradiction, since D_t is the difference of two bounded processes. Thus, it cannot be that $D_0 > 0$. Likewise, it is not possible to have $D_0 < 0$. Thus, in any PPE with continuation values $(W_t)_{t \geq 0}$, it must be the case that $W_t = U(X_t)$ for all t . Therefore, it must be that $|f(\bar{b}, X, \beta)| = 0$, and actions are uniquely specified by $S^*(X, U'(X)f_1(\bar{b}_t, X_t))$.

Q.E.D.

6.4 Proofs: Equilibrium Payoffs

Theorem 12. *The highest PPE payoff across all states is bounded above by the highest static Nash equilibrium payoff across states, $\bar{W} \leq \bar{v}^*$ and the lowest PPE payoff across all states is bounded below by the lowest static Nash equilibrium payoff across states $\underline{W} \geq \underline{v}^*$.*

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

$$U''(X) = \frac{2r [\bar{W} - v(X)]}{f_1(\bar{b}, X)^2 \sigma_Y^2 + \sigma_X^2(X)} \leq 0$$

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

Q.E.D.

Theorem 13. *Assume 1, 2, 3 and 4. The following properties show how PPE payoffs change with the state variable:*

1. *Suppose $v(X)$ is increasing (decreasing) in X . Then $U(X)$ is also increasing (decreasing) in X . The state that yields the highest static Nash payoff also yields the highest PPE*

payoff; likewise, the state that yields the lowest static Nash payoff also yields the lowest PPE payoff.

2. Suppose $v(X)$ has a unique interior maximum X^* , and v is monotonically increasing (decreasing) for $X < X^*$ ($X > X^*$). Then $U(X)$ has a unique interior maximum at X^* , and U is monotonically increasing (decreasing) for $X < X^*$ ($X > X^*$). The state that yields the highest static Nash payoff also yields the highest PPE payoff, whereas and the state that yields the lowest PPE payoff is a boundary point.
3. Suppose $v(X)$ has a unique interior minimum X^* , and v is monotonically decreasing (increasing) for $X < X^*$ ($X > X^*$). Then $U(X)$ has a unique interior minimum at X^* , and U is monotonically decreasing (increasing) for $X < X^*$ ($X > X^*$). The state that yields the lowest static Nash payoff also yields the lowest PPE payoff, whereas and the state that yields the highest PPE payoff is a boundary point.
1. Suppose $v(X)$ is increasing in X , but $U(X)$ is not increasing in X . Thus, $U'(X) < 0$ for some $X \in \Xi$. Let $(X_1, X_2) \subset \Xi$ be a maximal subinterval such that $U'(X) < 0$ for all $X \in (X_1, X_2)$. Note $\lim_{X \rightarrow -\infty} U(X) = \underline{v}^* \leq \lim_{X \rightarrow \infty} U(X) = \bar{v}^*$ since v is increasing in X , so $U'(X)$ is not strictly decreasing on Ξ . Without loss of generality, assume $U(X)$ is increasing on $(-\infty, X_1)$. Then X_1 is an interior local maximum with $U'(X_1) = 0$ and $U''(X_1) \leq 0$. Then by the optimality equation,

$$U''(X_1) = \frac{2r}{[f_1(\bar{b}, X_1)^2 \sigma_Y^2 + \sigma_X^2(\bar{b}, X_1)]} (U(X_1) - v(X_1)) \leq 0$$

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

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

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

$$\begin{aligned} U''(X_2) &= \frac{2r}{[f_1(\bar{b}, X_2)^2 \sigma_Y^2 + \sigma_X^2(\bar{b}, X_2)]} (U(X_2) - v(X_2)) \\ &\leq \frac{2r}{[f_1(\bar{b}, X_2)^2 \sigma_Y^2 + \sigma_X^2(\bar{b}, X_2)]} (U(X_2) - v(X_1)) < 0 \end{aligned}$$

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

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

and by the optimality equation, $U(X_1) \leq v(X_1)$. Also, X_2 is a local minimum with $U'(X_2) = 0$ and $U''(X_2) \geq 0$, and by the optimality equation, $U(X_2) \geq v(X_2)$. This implies:

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

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

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

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

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

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

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

3. The proof is analogous to part 2.

Q.E.D.