Starting Small and Renegotiation in Discrete-Time Relationships with a Continuum of Types

Xiameng Hua and Joel Watson*

January 2021

Abstract

We add to the literature on long-term relationships with variable stakes and incomplete information by analyzing renegotiation-proofness in a discrete-time partnership game between a principal and agent, with a continuum of types. In each period the principal selects the level of a project and the agent then decides whether to cooperate or betray; payoffs in the period scale with the level. The agent's benefit of betraying is privately known. The discrete-time framework allows renegotiation to be put in terms of an internal consistency condition that compares actual equilibria in the continuation of the game from any period, which improves on the prior literature. Our condition assumes the principal has full power to alter the equilibrium selection. Our main result shows that the resulting perfect Bayesian equilibria converge as the period length shrinks to zero, and we provide a closed-form solution. In equilibrium, the relationship starts small and the level gradually rises until it reaches its maximum; cooperation is viable regardless of the type distribution.

*UC San Diego. The authors thank local colleagues, especially Snehal Banerjee, Renee Bowen, Songzi Du, Simone Galperti, Mark Machina, and Joel Sobel, for their helpful comments. The authors also acknowledge very useful communication with David Kreps about the analysis of discrete-time games like the one studied here, as well as helpful input from Matthias Fahn and anonymous commentators.

1 Introduction

Long-term relationships in business and greater society often begin with a great deal of asymmetric information, where the parties are unsure of each others' incentives and they have choices regarding how to build their relationships. For instance, a manager may not know to what extent a new employee will have the incentive to shirk on his assignments, and the manager can decide what kinds of responsibilities to give this worker over time. Should the manager assign the worker to important projects, where effort would generate substantial profit for the firm but where shirking would translate into great losses?

Conventional wisdom suggests that it is better to start a relationship cautiously with small-stakes projects and then, conditional on good performance, increase the stakes as time goes on. In this way, a manager may be able to induce "bad" types of workers (those who inevitably will shirk at some point) to reveal themselves by shirking when the stakes are low. But if the manager would increase the stakes quickly over time, then a bad type worker would prefer to delay shirking until the stakes are high. Thus, the manager faces a trade-off between the rate at which she increases the worker's responsibilities (conditional on good performance) and when the worker's type will be revealed. Complicating matters, the manager may wish to adjust her plan mid-stream, based on what she learns about the worker.

We explore these dynamics by developing a new game-theoretic model of the interaction between a principal and an agent with private information and a continuum of types. The parties interact in discrete periods. In each period the principal selects the level of a project and the agent then chooses whether to cooperate or betray. We characterize the model's perfect Bayesian equilibria and we propose a renegotiation condition, which we call *alteration-proofness*, that narrows the set of equilibrium outcomes. Alteration-proofness is a notion of internal consistency that assumes the principal has full power to coordinate the players on an altered equilibrium.[1] By way of motivation, in line with the large literature on renegotiation in contractual settings, we think it is natural to assume that the players can revisit and change their equilibrium continuation. More generally, it is useful to work with models that generate narrow equilibrium predictions.

1. Internal consistency is the weakest version of Pareto-perfection that underlies definitions of renegotiation-proof equilibrium in the repeated-game literature, specifically those of Rubinstein (1980), Bernheim and Ray (1989), and Farrell and Maskin (1989).

Although there are multiple alteration-proof equilibria, our main result establishes that these equilibria converge as the period length shrinks to zero, meaning that the model has a unique prediction in the limit. The limit outcome is characterized by a differential initial-value problem. We provide an example for which the solution is easily found in closed form, along with analysis of comparative statics.

Our modeling exercise is closely related to Watson's (1999, 2002) analysis of relationships in continuous time with variable stakes and with two types of players: a "good" type, for whom cooperation would be possible in a setting of complete information, and a single bad type. In Watson's model, an exogenously provided level function gives the stakes of the relationship at every instant of time. The level function is interpreted as jointly determined by the players, and thus the game is not fully noncooperative. These articles show that, by *starting small*, long-term cooperation is always viable between good types of players, regardless of the initial type probabilities. Further, Watson (1999) puts forth renegotiation-proofness conditions that uniquely select a level function and outcome of the game.

We contribute to the literature in three ways. First, because our model is fully noncooperative and in discrete time, the alteration-proofness condition compares actual equilibria in the continuation of the game from any period. This setting provides a better foundation for renegotiation than was possible in Watson (1999). Second, we allow for a continuum of bad types and we obtain a novel characterization of the renegotiation-proof perfect Bayesian equilibria, along with comparative statics. Contrary to the result in Watson (1999), we find a multiplicity of alteration-proof equilibria in the discrete-time setting. This leads to our third contribution, which is to devise a method of bounding the set of equilibria and to characterize the bounds as the period length shrinks. Our method incorporates a new mathematical result on the limit of solutions to discrete-time models defined by transition functions with lags (Watson 2021).

The related literature includes both experimental evidence of starting-small behavior as well as a few other theoretical projects. Andreoni, Kuhn, and Samuelson (2019) reports an experiment in which subjects are able to choose the stakes in a two-period prisoners' dilemma, finding that players utilize a starting-small strategy to achieve cooperation. Likewise, Ye et al. (2020) studies a multi-period weakest-link game in the laboratory, where treatments differ in the exogenously set sequence of levels, finding that cooperation is associated with gradualism (starting small and gradual increase of the level). Kartal, Müller, and Tremewan (2019) provides experimental results on an infinite-horizon partnership game, where treatments differ in the set of level options. This paper find in settings of severe information asymmetry that subjects are able to build trust when they have the option of starting small and gradually raising the stakes of their relationships, and the subjects act accordingly.

Rauch and Watson (2003) develops a model of relationships in which the players have common information but are uncertain of their prospects as a partnership. The article shows theoretically that it is sometimes optimal to start small, and it also provides empirical evidence.[2] Bowen, Georgiadis, and Lambert (2019) examines starting small in a setting where two heterogeneous agents contribute over time to a joint project and collectively decide its level, finding that, in equilibrium, the effective control over the project scale relates to the realized types of players. Atakan, Koçkesen, and Kubilay (2020) studies repeated cheap talk and demonstrates that when the conflict of interest between the receiver and the sender is large, starting-small to communicate is the unique equilibrium arrangement.[3]

Also related is the model of Malcomson (2016, 2020), in which a principal and agent with persistent private information have an ongoing relationship governed by a relational contract (the principal makes voluntary payments to reward the agent's effort choice). Malcomson (2016) shows that when agent's type is on a continuum, then there does not exist a fully separating equilibrium, and Malcomson (2020) characterizes the finest partition equilibria. Separation requires a sufficiently low effort level at the beginning of the relationship, so in this sense some equilibria exhibit a form of starting small. Renegotiation-proofness in the form of external consistency (looking at the frontier of the set of equilibrium payoffs) is also studied in the latter paper.

This paper is organized as follows. In Section 2 we formally describe the model

2. Horstmann and Markusen (1996) models the choice by a multinational firm seeking to enter a new (foreign) market between direct investment and contracting with a local sales agent. Information gained from the agency contract is useful in the decision of whether to pursue direct investment. Hence, the agency contract is analogous to starting small in a variable-stakes games (though it may be desirable to extend it indefinitely). Horstmann and Markusen (2018) analyzes a similar model but relaxes the commitment assumption and studies both moral hazard and adverse selection.

3. Other articles that model some manner of starting small in relationships and/or trust building include Sobel (1985), Ghosh and Ray (1996), Kranton (1996), Blonski and Probst (2001), Rob and Fishman (2005), and Chassang (2010). Tangentially related are Halac (2012) on relationship dynamics under incomplete information and the effect of bargaining power, and Halac (2013) on long-term relationship-specific investment.

and equilibrium concept, and we analyze the agent's incentive conditions. The renegotiation condition is defined and analyzed in Section 3, where we report our main result on the limit of equilibria. Section 4 presents the case of uniformly distributed bad types, and Section 5 offers concluding comments. The appendices contain details of the analysis and proofs.

2 Model

We examine a model of a partnership between a principal and agent in discrete time and with one-sided incomplete information. We start by describing the complete-information version of the game, followed by the incomplete information version, and we then establish notation and review the equilibrium conditions. At the end of this section, we provide a partial characterization of "cooperative equilibria" and give examples of the multiplicity of equilibria.

2.1 Partnership game with complete information

To describe the model, we start by describing the complete-information version, which is a repeated game that terminates under some conditions. There are two players, called player 1 (the principal) and player 2 (the agent). The time period is denoted by $k = 1, 2, \ldots$. We assume the players have a common discount factor $\delta = e^{-r\Delta}$, where $r \in (0, 1)$ is the discount rate and Δ is the length of each period in real time.

In each period, as long as the game was not terminated earlier, players interact in the stage game shown in Figure 1, where player 1 selects a *trust level* $\alpha \in [0, 1]$ and then player 2 observes α and chooses whether to *betray* or *cooperate*. If player 2 cooperates then both players get the payoff $\alpha\Delta$ in the current period and play continues in the next period. On the contrary, if player 2 betrays then the game ends with terminal payoffs of $-\alpha c$ for player 1 and αx for player 2, where parameter x is non-negative and $c > 0$.

In this game of complete information, cooperation can be sustained only if $x \leq \Delta/(1 - \delta)$. Under this condition there is a subgame-perfect equilibrium in which, in every period, player 1 chooses $\alpha = 1$ and player 2 cooperates. Player 2's continuation value of playing this way from the start of any period is $\Delta/(1 - \delta)$, which exceeds the payoff of betraying. Player 1 can be deterred from deviating by specifying that,

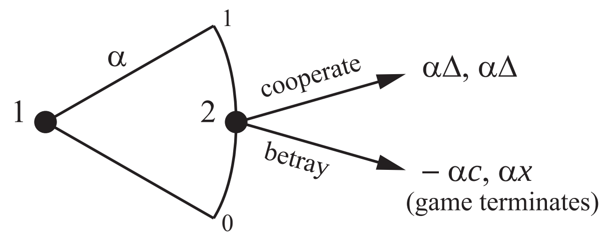


Figure 1: Stage game

following any deviation, the players coordinate on $\alpha = 0$ and betrayal from that point regardless of any further deviations (which is an equilibrium in all future subgames). There are similar equilibria in which, for any $\hat{\alpha} \in [0, 1]$, player 1 chooses $\alpha = \hat{\alpha}$ in each period and so there is cooperation at level $\hat{\alpha}$ over time.

Further, assuming $x \leq \Delta/(1-\delta)$, there are many other equilibria in which the level changes over time. In fact, for any sequence $\{\alpha^k\}$ of feasible levels, there is an equilibrium in which, on the equilibrium path, this sequence of levels is chosen by player 1 and player 2 always cooperates, so long as the following condition holds for each period k: $\alpha^k x \leq \alpha^k \Delta + \delta v_2^{k+1}$, where $v_2^{k+1} = \sum_{\tau=k+1}^{\infty} \delta^{\tau-k+1} \alpha^\tau \Delta$ is player 2's continuation value from the start of period $t+1$.

To summarize, there are a lot of cooperative equilibria if x is not too large. The best equilibrium for both players is clearly that in which player 1 chooses $\alpha = 1$ in every period. On the opposite side, if $x > \Delta/(1-\delta)$ then there is no equilibrium in which cooperation occurs at a positive level in any period, so cooperation is doomed.[4]

2.2 Partnership game with incomplete information

We are interested in the partnership game with incomplete information regarding the payoff parameter x (player 2's gain of betrayal). Specifically, suppose that before the partnership begins, Nature chooses x according to some exogenously given probability distribution F that is common knowledge. Player 2 observes x but player 1 does not observe it. We therefore call *x player 2's type.* Let us call every $x \leq \Delta/(1-\delta)$ a *good type* and every $x > \Delta/(1-\delta)$ a *bad type.*

4. Suppose on the contrary that there exists cooperation at positive levels. In order to induce player 2's cooperation, the levels must satisfy $\alpha^{k+1} \geq (1 - \Delta/x)\alpha^k/\delta$ for all k, so $\alpha^{k+1} \geq [(1 - \Delta/x)/\delta]^k \alpha^1$. Note that when $x > \Delta/(1-\delta)$, i.e. $(1 - \Delta/x)\alpha^k/\delta > 1$, for any $\alpha^1 > 0$, there exists period τ such that $\alpha^k > 1$ for $k > \tau$, which contradicts that α^k must be bounded by 1.

In this game, player 1 may be able to establish perpetual cooperation with a good type, but every bad type must eventually betray. The level of the relationship affects both player 2's betrayal gain and the players' flow payoff of cooperation, so by varying the level over time, player 1 may be able to coax the bad types to betray in periods when the level is small. However, there is a trade off: A bad type of player 2 would be willing to betray in a given period only if this player does not expect that player 1 would choose a much higher level in near future (contingent on player 2 cooperating until then). That is, it may be optimal for a bad type to cooperate for some number of periods and then betray later when α is large. Therefore, player 1 cannot screen out the bad types at a low level and also expect to soon cooperate at a high level with good types. Further, types with higher values of x are essentially less patient than are those with lower values of x, so player 1's choice of levels over time could lead different types of player 2 to betray in different periods.

To simplify our analysis, we assume there is just one good type, for which $x = 0$, and we assume that there is a continuum of bad types given by the interval $[a, b]$ where $b > a > \Delta/(1-\delta)$.[5] Later we will examine sequences of games for Δ converging to 0, so it is useful to note that $\Delta/(1-\delta)$ decreases and converges to $1/r$ as $\Delta \to 0$. We also make the following technical assumption:

Assumption 1. *On the interval of bad types $[a, b]$, cumulative distribution function F has full support, it is continuously differentiable, and its density function f is strictly positive and Lipschitz continuous (there exists a number $\omega > 0$ such that $|f(x') - f(x'')| \leq \omega |x' - x''|$ for all $a \leq x' \leq x'' \leq b$).*

Let $X = \{0\} \cup [a, b]$. From here, "game" and "partnership game" refer to the incomplete-information, discrete-time game with parameters r, Δ, c, a, b, and F just described.

2.3 Strategies and equilibrium conditions

We analyze the game using the Perfect Bayesian Equilibrium (PBE) solution concept. In this subsection, we define and provide notation for histories, strategies, and beliefs.

5. The restriction $a > \Delta/(1-\delta)$ is analogous to the "gap" assumption in the durable-good-monopoly problem of Coase (1972); see Gul, Sonnenschein, and Wilson (1986). The assumption of a single good type with $x = 0$ is innocuous; if we allow for a general set of good types, they all behave the same in the equilbria we identify.

We then describe the equilibrium conditions and, noting the plethora of equilibria, motivate the refinement developed in the next section.

For any $k = 1, 2, \ldots$, a k-period history of level choices is given by $(\alpha^1, \alpha^2, \ldots, \alpha^k)$. This sequence of levels can be interpreted as the public history to the beginning of period $k+1$ (specifying player 1's information set), where player 2 cooperated in periods $1, 2, \ldots, k$. Likewise, for $k > 0$, this same sequence $(\alpha^1, \alpha^2, \ldots, \alpha^k)$ represents the public history to player 2's information set in period k, where the public history to the beginning of period k was $(\alpha^1, \alpha^2, \ldots, \alpha^{k-1})$ and then player 1 selected α^k in period k. Note that player 2's personal history includes both $(\alpha^1, \alpha^2, \ldots, \alpha^k)$ and player 2's type x.

Let $H = \cup_{k=0}^{\infty}[0,1]^k$ be the set of all finite public histories, where $[0,1]^0$ is taken to be the null history at the beginning of period 1. Let $H_+ = \cup_{k=1}^{\infty}[0,1]^k$ be the set of non-null public histories. Also, for any k-period public history h and level α, denote by $h' = h\alpha \in H$ the $(k+1)$-period public history realized when h is followed by level α chosen in period $k+1$.

We focus on pure strategies.[6] Player 1's strategy $s_1 : H \to [0,1]$ specifies the level in each period as a function of the public history to this point. Player 2's strategy specifies whether to cooperate or betray in each period, as a function of history to player 2's information sets, including player 2's type. Thus, player 2's strategy is a function $s_2 : H_+ \times X \to \{1, 0\}$, where $s_2(h', x) = 1$ indicates that player 2 cooperates and $s_2(h', x) = 0$ indicates that player 2 betrays.

We describe player 1's beliefs about player 2's type using an assessment function $Q : H \to \mathcal{P}(X)$, where $\mathcal{P}(X)$ denotes the set of probability distributions over X. That is, for any k-period public history $h \in H$, $Q(h)$ is player 1's belief at the beginning of the following period $k+1$.

Given the strategies s_1 and s_2, any public history h, and player 2's type x, let $v_1(h; s_1, s_2, x)$ and $v_2(h; s_1, s_2, x)$ denote the players' continuation values from the period after history h occurs, assuming that x is player 2's actual type and that play will continue according to s_1 and s_2. Because player 1's assessment is $Q(h)$, player 1's

6. Restricting attention to pure strategies is without loss in our analysis. Because we have assumed a continuum of bad types, indifference conditions will occur for only a subset of measure zero. Further, the refinement developed in the next section would rule out equilibria in which the good type randomizes or player 1 randomizes.

expected continuation value is

$$v_1(h; s_1, s_2, Q(h)) \equiv \mathbb{E}_{Q(h)}[v_1(h; s_1, s_2, x)],$$

where $\mathbb{E}_{Q(h)}$ denotes expectation over $x \sim Q(h)$; this assumes that player 1 continues to believe after history h that player 2's strategy is s_2.

We extend player 2's strategy $s_2(h', x)$ to the space of type distributions by taking the expectation, so that for any $h \in H_+$ and any type distribution $\hat{F}$,

$$s_2(h, \hat{F}) \equiv \mathbb{E}_{\hat{F}}[s_2(h, x)].$$

Note that if player 1 chooses level α in the period following public history h, then player 1 expects player 2 to cooperate with probability $s_2(h\alpha, Q(h))$.

With this notation, we can describe the equilibrium conditions, starting with the notion of sequential rationality, stated here in terms of single deviations; the one-deviation principle applies.

Definition 1. *Given Q and s_2, player 1's strategy s_1 is called* **sequentially rational** *if for every public history $h \in H$, $s_1(h)$ maximizes*

$$s_2(h\alpha, Q(h))(\alpha\Delta + \delta v_1(h\alpha; s_1, s_2, Q(h\alpha))) + (1 - s_2(h\alpha, Q(h)))(-\alpha)$$

by choice of $\alpha \in [0, 1]$. Given s_1, player 2's strategy s_2 is called **sequentially rational** *if for every $h \in H_+$ and $x \in X$, $s_2(h, x) = 1$ only if*

$$\alpha\Delta + \delta v_2(h; s_1, s_2, x) \geq \alpha x$$

and $s_2(h, x) = 0$ only if the reverse weak inequality holds.

Definition 2. *For the partnership game, (s_1, s_2, Q) constitutes a pure strategy weak* **Perfect Bayesian Equilibrium (PBE)** *if the strategies are sequentially rational and Q obeys Bayes' Rule on the equilibrium path (for every $h \in H$ reached with positive probability given F, s_1, and s_2).*

Note that we have described the weak version of PBE, which requires player 1's beliefs to satisfy Bayes' rule only for information sets reached with positive probability in equilibrium. We could impose stronger consistency conditions (such as Watson

(2017) defines) but it would not matter for the partnership game that we study. This is because we can modify any weak PBE to satisfy strong consistency conditions off the equilibrium path. In fact, there is a simple way to specify equilibrium beliefs and behavior for off-path continuations that trivially achieves the lowest continuation value for all player-types. After any "public deviation," where either player 1 deviated or player 2 cooperated in a contingency in which all types were supposed to betray, specify that all types of player 2 betray and player 1 selects $\alpha = 0$ in every period thereafter (regardless of the history since the public deviation). These continuation strategies are sequentially rational regardless of player 1's beliefs about player 2's type.[7]

Because information sets following public deviations can always be dealt with as just described, incentive conditions on the equilibrium path are sufficient. To describe them, consider a PBE and let $\{\alpha^k\}_{k=1}^K$ be the sequence of levels chosen by player 1 on the equilibrium path, where K can be finite (if all of player 2's types betray on or before this date) or infinity. First, it must be that, for every $k = 1, 2, \ldots, K$, player 1's continuation value from period k is at least zero.[8] Second, each type of player 2 must chose to betray at the optimal time, or to never betray if this is better.

To express player 2's incentive condition mathematically, define:

$$\beta(x) \equiv \operatorname*{arg\,max}_{k \in \{1,2,\ldots,K\}} \sum_{\tau=1}^{k-1} \delta^{\tau-1} \alpha^{\tau} \Delta + \delta^{k-1} x \alpha^k.$$

This is the period at which type x optimally betrays. In the case of $K = \infty$, $\beta(x) = \infty$ means that type x cooperates forever. For each $k = 1, 2, \ldots, K$, let $h^k = (\alpha^1, \alpha^2, \ldots, \alpha^k)$ denote the equilibrium-path public history to player 2's information set in period k. Player 2's type x behaves optimally if $s_2(h^k, x) = 1$ for all $k < \beta(x)$ and $s_2(h^{\beta(x)}, x) = 0$.

7. Also, it is straightforward to describe player 1's beliefs as "appraisals" (Watson, 2017) but this generality is not needed because the partnership game has the simple structure of multiple stages with observed actions.

8. Player 1 can guarantee a payoff of zero by choosing $\alpha = 0$ forever. Further, player 1's continuation value following a deviation can be made zero as described in the previous paragraph.

2.4 Cooperative PBE

We focus on PBE in which, on the equilibrium path, the level is strictly positive in at least one period and the good type cooperates forever, so that $\beta(0) = \infty$. Therefore $K = \infty$ in the expressions above. We call such an equilibrium a *cooperative PBE.* In any other equilibrium, the payoffs of player 1 and the good type of player 2 are strictly lower than in a cooperative PBE, and such an equilibrium would fail the alteration-proofness condition that we develop in the next section.

The first issue to address is whether a cooperative PBE exists.

Theorem 1. *Assuming $F(0) > 0$, a cooperative PBE exists regardless of the other parameters.*

This result extends what was found by previous papers in the literature, in particular Watson (1999, 2002), so it is not surprising. It is worth noting what this means in economic terms. First, an ongoing cooperative relationship between player 1 and the good type of player 2 is viable, and value is created regardless of the type distribution. Second, this conclusion relies on the ability of the players to start small in their relationship. That is, if player 1 had only the choice of, say, $\alpha = 0$ or $\alpha = 1$ then there would be no cooperative PBE for $F(0)$ sufficiently small. Theorem 1 is an implication of the existence result in the next section, so we do not provide a proof here.

We next characterize the cooperative PBE in terms of the relation between the strategy of player 2 (specifically, identifying dates when the various bad types betray) and the sequence of levels on the equilibrium path.

Lemma 1. *For any cooperative PBE of the partnership game, β is weakly decreasing and $\beta(a)$ is finite. That is, there exists a weakly decreasing sequence $\{x^k\}_{k=0}^{\infty}$ such that (i) for all $x \in [a, b]$, in equilibrium type x of player 2 betrays in a period k that satisfies $x \in [x^k, x^{k-1}]$, and (ii) $x^k \leq a$ for large enough values of k.*

Proof of Lemma 1. Consider any cooperative PBE and let $\{\alpha^k\}_{k=1}^{\infty}$ be the sequence of levels chosen on the equilibrium path. Let us define $\omega(k, x)$ as the objective function for the definition of β:

$$\omega(k, x) \equiv \sum_{\tau=1}^{k-1} \delta^{\tau-1} \alpha^{\tau} \Delta + \delta^{k-1} x \alpha^k.$$

Suppose that there exist types x' and x'' such that $x' > x''$ and yet $\beta(x') > \beta(x'')$, and we will find a contradiction. Imagine that player 2 compares betraying in period $\beta(x')$ with betraying in period $\beta(x'')$, ignoring other periods. By the definition of β, type x' prefers betraying in period $\beta(x')$ whereas type x'' prefers betraying in period $\beta(x'')$ only if

$$\omega(\beta(x'),x') - \omega(\beta(x''),x') \geq 0 \text{ and } \omega(\beta(x''),x'') - \omega(\beta(x'),x'') \geq 0,$$

and the preference is strict if the relevant inequality hold strictly. Using the definition of ω and simplifying terms, we get

$$\begin{aligned}\left(\delta^{\beta(x')-\beta(x'')}\alpha^{\beta(x')} - \alpha^{\beta(x'')}\right) x'' \leq -\delta^{-\beta(x'')} \sum_{\tau=\beta(x'')}^{\beta(x')-1} \delta^{\tau}\alpha^{\tau}\Delta& \\ \leq \left(\delta^{\beta(x')-\beta(x'')}\alpha^{\beta(x')} - \alpha^{\beta(x'')}\right) x'&\end{aligned}$$

Because the level is strictly positive in at least one period on the equilibrium path, player 2's incentive condition implies that $\alpha^{\beta(x')}, \alpha^{\beta(x'')} > 0$, which further implies that the middle term in the above expression is strictly negative. It therefore must be the case that $\delta^{\beta(x')-\beta(x'')}\alpha^{\beta(x')} - \alpha^{\beta(x'')} < 0$, which contradicts that $x' > x'' \geq 0$. Therefore, we have proved the monotonicity of β.

Next we show $\beta(a)$ is finite. Define $\overline{\alpha} = \sup\{\alpha^1, \alpha^2, \ldots\}$. Then for any $\epsilon > 0$, there exists a period κ such that $\alpha^{\kappa} \geq \overline{\alpha} - \epsilon$. If player 2 of type a betrays in period κ, then the game ends and he gets terminal payoff $a\alpha^{\kappa}$, which weakly exceeds $a(\overline{\alpha} - \epsilon)$. If this type never betrays, then his continuation value from period κ is $\sum_{k=\kappa}^{\infty} \delta^{k-1}\alpha^k\Delta$, which is bounded above by $\sum_{k=\kappa}^{\infty} \delta^{k-1}\overline{\alpha}\Delta = \overline{\alpha}\Delta/(1-\delta)$. Because $a > \Delta/(1-\delta)$, it is the case that $a(\overline{\alpha}-\epsilon) > \overline{\alpha}\Delta/(1-\delta)$ for sufficiently small values of ϵ, which contradicts that it is rational for type a to cooperate forever. □

Note that point (ii) does not pin down the period of betrayal for the finite number of types of player 2 that may be indifferent between betraying in one period or the next. Because this is a set of measure zero, equilibria that differ in this regard are essentially equivalent. So let us break the ties in favor of betraying later and assume that type x of player 2 betrays in the period k that satisfies $x \in (x^k, x^{k-1}]$. On the equilibrium path, $[a, x^{k-1}]$ is the set of bad types who have cooperated through period $k-1$. Bad types in the subinterval $(x^k, x^{k-1}]$ will betray in period k, at level α^k, and the remaining bad types (those in the interval $[a, x^k]$) cooperate through at least period k.

Furthermore, values of x^k that are less than a (meaning all bad types betray at or before the current period) or that exceed b (meaning that no bad types have yet betrayed) are arbitrary in the sense any such values have the same meaning. So let us add convenient structure to the sequence by assuming that $x^0 = b$ and the only value below a that the sequence takes is $\Delta/(1-\delta)$. That is, the sequence starts at $x^0 = b$ and no bad types betray until the first period k at which $x^k < b$. The last period in which bad types betray is denoted L; it is the first period in which the sequence falls below a and we have $x^k = \Delta/(1-\delta)$ for every $k \geq L$. We will refer to L throughout the analysis in the next section.

To summarize the analysis so far, every cooperative PBE is fully characterized by its sequence of levels $\{\alpha^k\}_{k=1}^{\infty}$ and its sequence of cutoff types $\{x^k\}_{k=1}^{\infty}$. It will also be helpful to keep track of continuation values, so given any cooperative PBE and any period k, we let v_1^k denote the expected continuation value for player 1 from the start of period k on the equilibrium path. Likewise, we let $v_2^k(x)$ denote the continuation value of player 2 of type x from the start of period k conditional on player 2 having always cooperated in the past and player 1 not having deviated from the equilibrium level sequence.

2.5 Intuition and Illustrations

To get a flavor of the relation between the level sequence and player 2's optimal choices, let us take a moment to examine the trade off that player 2 faces locally in time. Because in equilibrium type x^k weakly prefers to cooperate in period k, we have

$$\alpha^k x^k \leq \Delta\alpha^k + \delta v_2^{k+1}(x^k), \tag{1}$$

Note that x^k is the highest type that remains after period k, so this is the type in $[a, x^{k-1}]$ that has the greatest incentive to betray in period k. In some equilibria, type x^k is indifferent between cooperating and betraying in period k, so that

$$\alpha^k x^k = \Delta\alpha^k + \delta v_2^{k+1}(x^k). \tag{2}$$

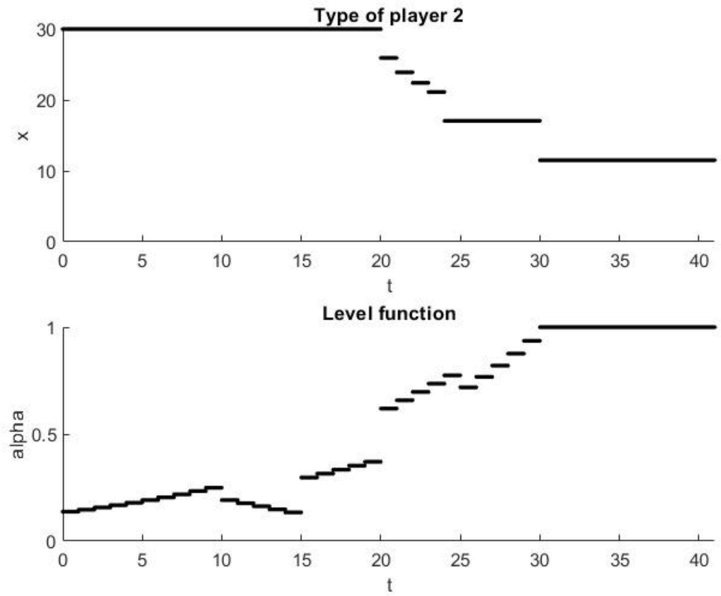


Figure 2: First equilibrium illustration

Then either type x^k betrays in period $k+1$ or we have $x^{k+1} = x^k$. In the event that the indifference condition holds until this type actually betrays, an implication is that $v_2^{k+1}(x^k) = \alpha^{k+1} x^k$. Using this expression to substitute for $v_2^{k+1}(x^k)$ in Equation (2), we obtain:

$$x^k = \frac{\Delta \alpha^k}{\alpha^k - \delta \alpha^{k+1}}. \tag{3}$$

It will turn out that the refinement developed in the next section implies that Equation (3) holds in every period k for which $x^k \geq a$; that is, this indifference condition holds until all bad types have betrayed.

Before proceeding to the equilibrium refinement in the next section, we illustrate the multiplicity of cooperative equilibria, which differ in terms of when bad types betray, how the level changes over time, and players 1's payoff. In Figures 2–4, we construct three equilibria for the same specification of parameters: $\Delta = 1$, $r = 0.1$ (so that $\delta = e^{-r\Delta} = 0.9048$), $a = 11.5083$, and $b = 30$. The distribution F of player 2's type puts probability 0.3836 on the good type and specifies uniform distribution of the bad types. The value of c matters only for player 1's incentives, and the equilibria pictured exist as long as c is not too large. Note that in the constructed equilibria, the last period in which bad types betray is $L = 31$, after which we have $x^k = \Delta/(1-\delta) = 10.5083$.

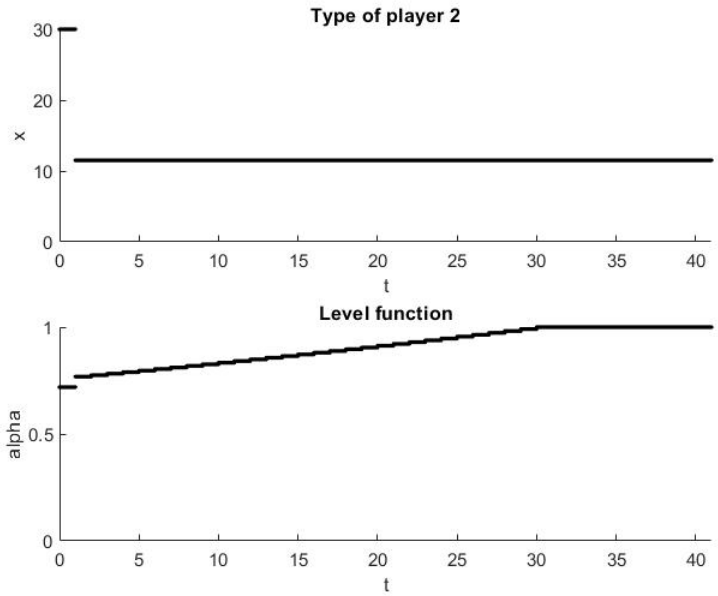


Figure 3: Second equilibrium illustration

For the equilibrium in Figure 2, the indifference Condition (2) does not always hold; player 2 strictly prefers to cooperate in early periods when the level is low, looking forward to betraying in later periods when the level is high. Figures 3 and 4 illustrate equilibria for which Equation (2) holds for all periods; in Figure 3, all bad types betray at the beginning of the game (period $k = 1$), while in Figure 4, no bad type betrays until the level reaches 1 (period $k = 30$). We can show that if $c > b$ then player 1's favorite equilibrium is as pictured in Figure 3, where all bad types of player 2 betray in the first period, whereas if $c < a$ then player 1's favorite equilibrium is as pictured in Figure 4, where all bad types of player 2 wait until period $L-1$ to betray. These findings match with what Watson (2002) demonstrates in the continuous-time model with a single bad type. We mention this in passing because our main objective is to study the implications of renegotiation in our discrete-time setting and we can use the equilibria illustrated here to motivate our renegotiation condition.

It turns out that none of the equilibria pictured are renegotiation-proof. In the first, there are periods in which the level can be increased without affecting player 2's incentives and this increases player 1's payoff.[9] In the second equilibrium, after ob-

9. This feature of the first illustration is familiar, for many signaling models have separating equilibria with nonbinding incentive constraints. Consider, for instance, the standard labor-market signaling game and a separating equilibrium in which the high-ability type chooses an education

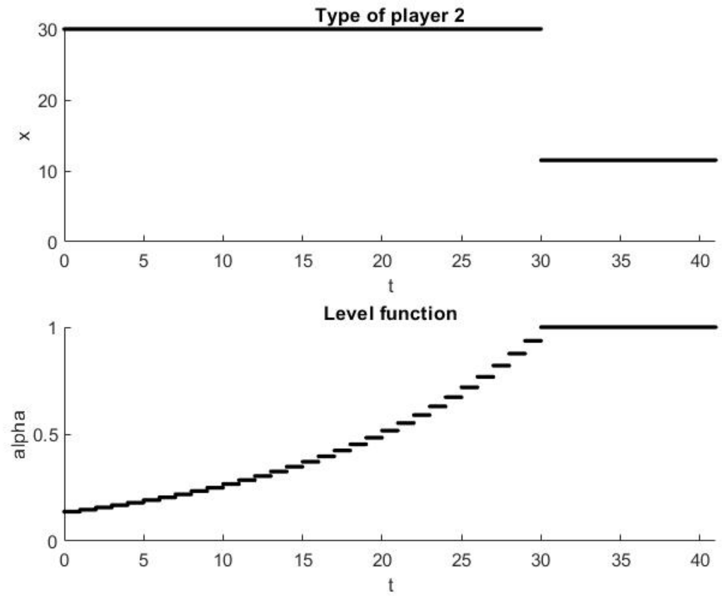


Figure 4: Third equilibrium illustration

serving cooperation in the first period, player 1 would be sure that player 2 is the good type that will never betray. Therefore, in the second period player 1 has the incentive to "jump ahead" to the continuation of the equilibrium from period $L = 30$ where the level is maximal. In the third equilibrium, player 1's payoff decreases as period $L = 30$ approaches and so, in any period before L, player 1 would have the incentive to "stall" as though restarting from the previous period.

3 Alteration Proofness

In this section, we define and analyze a renegotiation-proofness condition that we call *alteration proofness*, where player 1 has the power to dictate an alteration of current equilibrium in the continuation of the game from any period.[10] The concept imposes a form of internal consistency on the equilibrium path: In a given period τ the equilibrium continuation may be altered in any way, so long as in period $\tau + 1$ it returns to a path consistent with the current equilibrium. The new path from

level that is higher than needed for separation. In our model, renegotiation-proofness forces some constraints to bind.

10. That is, we assume that player 1 is the organizational leader; in terms of mechanics, we could imagine that there is pre-play communication at the beginning of each period, players use these messages to coordinate on a continuation path, and only player 1 can speak.

period $\tau + 1$ can pick up the current equilibrium as though in any other period $\tau' \in \{\tau, \tau+1, \tau+2, \ldots\}$.

For instance, if $\tau' = \tau$ then the players are stalling, essentially postponing the equilibrium path by one period. Any $\tau' > \tau + 1$ amounts to "jumping ahead" by $\tau' - \tau - 1$ periods, and $\tau' = \tau + 1$ means that the alteration affects only the current period τ. Restrictions are inherent in how an equilibrium can be altered in this way. In particular, to pick up on the current equilibrium as though in period τ', player 1's belief must be exactly as it would be at the start of period τ', so the alteration must specify for the current period τ behavior that would lead to such a belief at the end of this period.

We motivate the alteration-proofness concept as a minimal notion of renegotiation, where only internally consistent alterations are considered. These alterations may not be the only ones viewed as viable by the players. One can imagine, for instance, that player 1 could demand to switch to some other PBE path, one that does not coincide with any continuation of the current equilibrium. This would strengthen the renegotiation-proofness concept to include a form of external consistency. Because alteration-proofness, based on the small family of equilibrium comparisons described above, will be shown to uniquely select a limit PBE, we assert that the selection is robust to a broader family of alterations.[11]

Our alteration-proofness condition is along the lines of the condition developed in Watson (1999) but has two significant advantages. First, Watson (1999) imposes two separate conditions for a stall and a jump, and these are local in nature; our definition here is a single global condition. Second, because Watson (1999) studies a continuous-time model with a jointly selected level, the conditions there are described as limit conditions that go outside the game being analyzed. In the discrete-time framework here, every feasible alteration is an equilibrium in the continuation game.

3.1 The alteration-proofness condition

Consider any cooperative PBE, characterized by $\{\alpha^k\}_1^\infty$ and $\{x^k\}_0^\infty$, and suppose that any period τ is reached on the equilibrium path. We imagine that player 1 may

11. We could allow $\tau' < \tau$ but it would not change the implications of our theory. In this case, Lemma 1 implies $x^{\tau'} \geq x^{\tau-1}$. If this inequality is strict then the alteration is not feasible; otherwise, player 1 would strictly prefer the alteration only if the alteration with $\tau'\tau$ is strictly preferred. Note as well that the alteration-proofness condition refers only to equilibrium-path contingencies. Imposing similar conditions to contingencies following public deviations would not affect our analysis.

dictate that the equilibrium is to be altered in the continuation of the game, in such a way as to have the path of play from period $\tau + 1$ be as though in the original equilibrium from period $\tau + 1 + m$, where $m \in \{-1, 0, 1, \ldots\}$ denotes by how many periods the altered equilibrium skips ahead in relation to the original equilibrium from period $\tau + 1$. In the altered equilibrium, play from period τ will be described by sequences $\{\tilde{\alpha}^k\}_{k=\tau}^{\infty}$ and $\{\tilde{x}^k\}_{k=\tau}^{\infty}$ where $\tilde{\alpha}^k = \alpha^{k+m}$ for all $k > \tau$ and $\tilde{x}^k = x^{k+m}$ for all $k \geq \tau$. Note that the $\tilde{\alpha}^\tau$ is not nailed down here and so it and m define the alteration.

The level $\tilde{\alpha}^\tau$ is constrained by the requirement that the original and altered equilibrium continuations fit together in terms of player 2's incentives in period τ. To understand the constraint, observe that the alteration is feasible only if, at the beginning of period $\tau + 1$, player 1's belief is exactly what it would have been at the beginning of period $\tau + m + 1$ in the original equilibrium (on the equilibrium path). That is, the continuation game in the altered equilibrium from period $\tau + 1$ must be identical to the continuation game in the original equilibrium from period $\tau+m+1$. In the latter continuation, player 1's belief about player 2's type is exactly the updated version of F conditional on $x \leq x^{\tau+m}$. That is, bad types in the interval $[a, x^{\tau+m}]$ are the only ones that remain in the game at this point.

Therefore $\tilde{\alpha}^\tau$ must be set so that bad types less than or equal to $x^{\tau+m}$ prefer to cooperate in period τ given the altered level sequence, and bad types in the interval $(x^{\tau+m}, x^{\tau-1}]$ prefer to betray in period τ. The first condition is

$$\tilde{\alpha}^\tau x \leq \Delta\tilde{\alpha}^\tau + \delta v_2^{\tau+m+1}(x) \text{ for all } x \in [a, x^{\tau+m}], \tag{4}$$

and the second is

$$\tilde{\alpha}^\tau x \geq \Delta\tilde{\alpha}^\tau + \delta v_2^{\tau+m+1}(x) \text{ for all } x \in (x^{\tau+m}, x^{\tau-1}] \cap [a, b], \tag{5}$$

where $v_2^{\tau+m+1}$ refers to player 2's continuation value in the original equilibrium.[12] The restriction to $x \in [a, b]$ simply ensures application to the actual type space.

Note that the constraints can be vacuous depending on how values $x^{\tau-1}$ and $x^{\tau+m}$ relate to a and b. For instance, if $x^{\tau-1} < a$ then all bad types betray before period τ

12. That is, as defined earlier, $v_2^k(x)$ is the continuation value (here in the original equilibrium) of player 2 of type x from the start of period k conditional on player 2 having always cooperated in the past and player 1 not deviating from the equilibrium level sequence.

in equilibrium, and the second constraint is trivially satisfied. This condition is also vacuous if $x^{\tau+m} = x^{\tau-1}$, which is the case if no bad types are scheduled to betray between periods τ and $\tau + m$ (inclusive) in the original equilibrium.

Definition 3. *Take as given a cooperative PBE, with player 2's equilibrium continuation values denoted by* $\{v_2^k(\cdot)\}_{k=1}^{\infty}$. *For any period* τ*, integer* $m \in \{-1, 0, 1, \ldots\}$*, and level* $\tilde{\alpha}^\tau$*, call the triple* $(\tau, m, \tilde{\alpha}^\tau)$ *an* **alteration** *of the equilibrium. Call* $(\tau, m, \tilde{\alpha}^\tau)$ *a* **feasible alteration** *if Inequalities* (4) *and* (5) *are satisfied.*

Two beneficial alterations were illustrated at the end of Section 2.5. In Figure 3, because all bad types betray in the first period in equilibrium, cooperation in this period would lead player 1 in period 2 to desire the alteration $(2, 28, \tilde{\alpha}^2)$; that is, player 1 would jump ahead from period 2 to period 30. In Figure 4, because all bad types betray in period 30 in equilibrium, when period 30 is reached, player 1 would desire an alteration $(30, -1, \tilde{\alpha}^{29})$, effectively going back to period 29 where all types cooperate.

Recall that, for a given cooperative PBE and any period k, v_1^k denotes the expected continuation value for player 1 from the start of period k on the equilibrium path. Note that on the equilibrium path at the beginning of period τ, player 1 believes that player 2's type is weakly below $x^{\tau-1}$. Further, after selecting the level α^τ that the equilibrium prescribes for period τ, player 1 expects player 2 to cooperate with probability $F(x^\tau)/F(x^{\tau-1})$ and to betray with complementary probability. We can thus express player 1's expected continuation value recursively as

$$v_1^\tau = \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)(-c\alpha^\tau) + \frac{F(x^\tau)}{F(x^{\tau-1})}\left(\alpha^\tau \Delta + \delta v_1^{\tau+1}\right). \tag{6}$$

If at period τ player 1 demands that the players coordinate on a feasible alteration $(\tau, m, \tilde{\alpha}^\tau)$, then player 1's continuation value would instead be

$$\left(1 - \frac{F(x^{\tau+m})}{F(x^{\tau-1})}\right)(-c\tilde{\alpha}^\tau) + \frac{F(x^{\tau+m})}{F(x^{\tau-1})}\left(\tilde{\alpha}^\tau \Delta + \delta v_1^{\tau+m+1}\right).$$

Remember that in the alteration, bad types in the interval $(x^{\tau+m}, x^{\tau-1}]$ betray in period τ, and in period $\tau + 1$ play continues as though from period $\tau + m + 1$ in the original equilibrium.

Definition 4. *Call a cooperative PBE* **alteration proof** *if no feasible alteration im-*

proves player 1's continuation payoff. That is, for every feasible alteration $(\tau, m, \tilde{\alpha}^\tau)$*,*

$$v_1^\tau \geq \left(1 - \frac{F(x^{\tau+m})}{F(x^{\tau-1})}\right)(-c\tilde{\alpha}^\tau) + \frac{F(x^{\tau+m})}{F(x^{\tau-1})}\left(\tilde{\alpha}^\tau\Delta + \delta v_1^{\tau+m+1}\right). \tag{7}$$

We next characterize alteration-proof cooperative PBE. We refer to these equilibria simply as "alteration-proof equilibria." The main implications can be derived by considering local alterations, whereby $m \in \{-1, 0, 1\}$. Recall that L denotes the last period in which bad types betray on the equilibrium path, so that $x^{L-1} \geq a > x^L = \Delta/(1-\delta)$.

Lemma 2. *In every alteration-proof equilibrium, the level is maximal after all bad types have betrayed. That is,* $\alpha^k = 1$ *for every* $k > L$ *(where* $x^{k-1} < a$*).*

Proof. This is an immediate implication of alteration-proofness for $m = 0$ and τ satisfying $x^{\tau-1} < a$. The alteration $(\tau, 0, 1)$ is trivially feasible and strictly increases player 1's continuation payoff from period τ if $\alpha^\tau < 1$ in the original equilibrium. □

Lemma 3. *In every alteration-proof equilibrium, Equation* (3) *holds, that is* $x^k = \Delta\alpha^k/(\alpha^k - \delta\alpha^{k+1})$*, for each period* $k < L$*.*

Recall that this relation between the level sequence and cutoff types was discussed in the previous section and amounts to type x^k, the cutoff type in period k, being indifferent between betraying in period k and betraying in period $k+1$. Rearranging a bit gives an expression for the rate of increase in the level over time, relative to the cutoff type:

$$\frac{\alpha^{k+1}}{\alpha^k} = \frac{x^k - \Delta}{x^k\delta}. \tag{8}$$

The right side strictly exceeds 1, implying that the equilibrium level sequence is strictly increasing. The rate of increase from period to period is itself increasing in the cutoff type and therefore decreasing in k.

Proof of Lemma 3. For convenience in this proof, let us extend v_2^{L+1} to be defined for $x = \Delta/(1-\delta)$ by specifying $v_2^{L+1}(\Delta/(1-\delta)) = \Delta/(1-\delta)$, which would be the continuation value of type $\Delta/(1-\delta)$ in the continuation from period $L+1$ given that the level is 1 thereafter. Of course, there is no type $\Delta/(1-\delta)$ in the model. The extension gives us the starting point for an induction argument.

We begin by proving that, in any alteration-proof equilibrium,

$$v_2^{k+1}(x^k) = \alpha^{k+1} x^k \tag{9}$$

for all $k \leq L$. Note first that this equation holds for $k = L$ because $x^L = \Delta/(1-\delta)$ and $\alpha^{L+1} = 1$. We proceed with an inductive argument.

Suppose that, for a given period $k > 1$, Equation (9) holds. We shall demonstrate that $v_2^k(x^{k-1}) = \alpha^k x^{k-1}$. If $x^{k-1} > x^k$, meaning that type x^{k-1} betrays in period k, we immediately obtain $v_2^k(x^{k-1}) = \alpha^k x^{k-1}$. So let us assume that $x^{k-1} = x^k$, whereby in equilibrium no types betray in period k. Because type x^{k-1} betrays in a future period, it must be that

$$\alpha^k x^{k-1} \leq \Delta\alpha^k + \delta v_2^{k+1}(x^{k-1}) = \Delta\alpha^k + \delta\alpha^{k+1} x^{k-1}. \tag{10}$$

The equality holds because of $x^{k-1} = x^k$ and Equation (9).

Suppose that Inequality (10) is strict. We can find a level $\tilde{\alpha}^k \in (\alpha^k, 1)$ for which, uniquely,

$$\tilde{\alpha}^k x^{k-1} = \Delta\tilde{\alpha} + \delta v_2^{k+1}(x^{k-1}) = \Delta\tilde{\alpha}^k + \delta\alpha^{k+1} x^{k-1}. \tag{11}$$

The existence of this level is implied by the fact that $x^{k-1} > \Delta/(1-\delta) > \Delta$. In fact $(k, 0, \tilde{\alpha}^k)$ is a feasible alteration. Demonstrating feasibility just requires checking that types below x^{k-1} strictly prefer to cooperate in period k in the alteration, which is straightforward.[13] Because no types betray in period k in the original equilibrium and in the alteration, and because the level in period k is higher in the altered equilibrium, player 1's continuation payoff strictly increases.

Therefore it must be that Inequality (10) holds as an equality. That is, in period k type x^{k-1} is indifferent between betraying and cooperating. This implies that $v_2^k(x^{k-1}) = \alpha^k x^{k-1}$, completing the inductive argument.

Next, using Identity (9) we prove the claim of the lemma. Consider any $k < L$ and let us look at two cases. First, if $x^{k-1} = x^k$ then, by the above argument, weak Inequality (10) binds and we have $\alpha^k x^{k-1} = \Delta\alpha^k + \delta\alpha^{k+1} x^{k-1}$. Replacing x^{k-1} with x^k and rearranging terms yields $x^k = \Delta\alpha^k/(\alpha^k - \delta\alpha^{k+1})$. In the second case, we have $x^{k-1} > x^k$. Then types in the nonempty interval $(x^k, x^{k-1}]$ prefer to betray in

13. For $x < x^{k-1}$, $\tilde{\alpha}^k x < \Delta\tilde{\alpha}^k + \delta\alpha^{k+1} x < \Delta\tilde{\alpha} + \delta v_2^{k+1}(x)$ because $(\tilde{\alpha}^k - \delta\alpha^{k+1})x < (\tilde{\alpha}^k - \delta\alpha^{k+1})x^{k-1} = \Delta\tilde{\alpha}^k$.

period k, whereas types in the nonempty interval $[a, x^k]$ prefer to cooperate. Type x^k must be indifferent between cooperation and betrayal in period k, because player 2's continuation value is continuous in player 2's type for any given level sequence.[14] Hence, $\alpha^k x^k = \Delta\alpha^k + \delta v_2^{k+1}(x^k)$. Using Equation (9) to substitute for $v_2^{k+1}(x^k)$ and rearranging terms once again yields $x^k = \Delta\alpha^k/(\alpha^k - \delta\alpha^{k+1})$. $\square$

Together Lemmas 2 and 3 imply that, in an alteration-proof equilibrium, the equilibrium level increases gradually from the first period until it reaches 1, and then remains at 1 thereafter. Further, the level increases in relation to the rate at which the bad types betray, so that in a given period the cutoff bad type is indifferent between betraying in the current period and betraying in the next period.

Note that Equation (8) and monontonicity of the sequences $\{\alpha^k\}_{k=0}^{\infty}$ and $\{x^k\}_{k=0}^{\infty}$ are necessary but not sufficient conditions for alteration-proofness. The set of equilibria that satisfy these properties is quite large and varied. For example, Equation (8) holds in the equilibria illustrated in Figures 3 and 4, which fail to be alteration-proof. Further, there is a large range of equilibria with monotone sequences that satisfy Equation (8) and are between those illustrated in Figures 3 and 4.

Although these lemmas do not pin down L or the rates at which the level increases and the bad types betray, they provide a useful partial characterization of alteration-proofness in terms of gradualism. Also note that these lemmas do not indicate exactly the equilibrium level in period L. It is not difficult to show that α^L must be in the interval $[a\delta/(a - \Delta), 1]$. The lower endpoint of this interval would make type a indifferent between betraying at L and waiting to do so at $L + 1$, whereas the upper endpoint would make the artificial type $\Delta/(1 - \delta)$ indifferent. In any case, the level at L will not matter as we shrink the period length.

As a step toward our main result, we next utilize Lemma 3 to derive bounds on equilibrium continuation values for every period $\tau < L$. These bounds are constructed by considering alterations in which $m = -1$ or $m = 1$. Our first observation is that Lemma 3 implies that, in a given period $\tau < L$ and for $m \in \{-1, 1\}$, the only feasible alteration $(\tau, m, \tilde{\alpha}^\tau)$ that player 1 could possibly find attractive is that for which $\tilde{\alpha}^\tau$ satisfies

$$\tilde{\alpha}^\tau x^{\tau+m} = \Delta\tilde{\alpha}^\tau + \delta x^{\tau+m}\alpha^{\tau+m+1}.$$

That is, the alteration is supposed to make type $x^{\tau+m}$ indifferent, so that any higher

14. It is easy to verify that $\omega(k, x)$ is continuous in x for fixed k.

types betray in period τ and bad types in $[a, x^{\tau+m}]$ remain into the next period. In the case of $m = -1$, Equation (7) simplifies to

$$v_1^\tau \geq \frac{\Delta}{1-\delta}\alpha^{\tau-1}. \tag{12}$$

In the case of $m = 1$, Equation (7) becomes

$$v_1^\tau \leq \frac{\Delta}{1-\delta}\alpha^\tau \left(\frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})} \cdot \frac{c}{x^\tau} + \frac{F(x^\tau)}{F(x^{\tau-1})} \right). \tag{13}$$

These two conditions provide upper and lower bounds on player 1's continuation value, which constrain how fast the level increases and bad types betray. For these two conditions to hold, the equilibrium must satisfy

$$\frac{F(x^{\tau-1}) - F(x^\tau)}{\Delta F(x^{\tau-1})} \leq \frac{\alpha^\tau - \alpha^{\tau-1}}{(c+\Delta)\alpha^\tau - \delta c \alpha^{\tau+1}}. \tag{14}$$

The derivation of Inequalities (12), (13) and (14) is shown in Appendix A. We turn to the questions of existence of alteration-proof PBE.

Theorem 2. *For any $\Delta > 0$, there exists an alteration-proof equilibrium.*

A constructive proof is provided in Appendix B. We first construct an equilibrium such that player 1 is indifferent between continuing the equilibrium play and altering with $m = -1$ for each period. Next we verify that this equilibrium is alteration-proof.

3.2 Characterization theorem

Our main theorem characterizes alteration-proof equilibria as the period length Δ shrinks to zero. Fix a, b, r, and F. We shall consider any sequence of games indexed by positive integer j, where the period length of game j is denoted $\Delta(j)$ (and the games all share the same parameters a, b, r, and F), such that $\Delta(j)$ converges to zero as $j \to \infty$. For each game j, we consider an arbitrary alteration-proof equilibrium, described by a level sequence $\{\alpha^k(j)\}_{k=1}^\infty$, a sequence of type cutoffs $\{x^k(j)\}_{k=1}^\infty$, and a sequence of player 1's continuation values $\{v^k(j)\}_{k=1}^\infty$. By Theorem 2, we know that there exists an alteration-proof cooperative PBE for each $\Delta(j)$.

To describe what happens as $j \to \infty$, for every j we need to translate the discrete-time equilibrium sequences (levels, type cutoffs, and continuation values) into functions of continuous time. Letting t denote time on the continuum, define $M(t,j) = \min\{k \mid k\Delta(j) \geq t\}$ to be the period in discrete-time game j that contains time t. Then define step functions $\hat{\alpha}(\cdot;j) : [0,\infty] \to [0,1]$, $\hat{x}(\cdot;j) : [0,\infty] \to [0,b]$, and $\hat{v}_1(\cdot;j) : [0,\infty] \to [0,\infty)$ by

$$\hat{\alpha}(t;j) = \alpha^{M(t,j)}(j), \quad \hat{x}(t;j) = x^{M(t,j)}(j), \quad \hat{v}_1(t;j) = v_1^{M(t,j)}(j).$$

Our main theorem shows that each of these functions converges to a continuous-time limit that is independent of the exact sequence of period lengths and the selection of alteration-proof equilibria for each j. That is, alteration-proofness uniquely pins down the equilibrium when the period length is small. We will denote the limit functions by $\alpha(\cdot)$, $x(\cdot)$, and $v_1(\cdot)$, which alters notation in a way that will hopefully not be confusing. Our theorem also shows that these functions are characterized by the solution to a system of differential equations.

Theorem 3. *Fix a, b, r, and F. There exist functions $\alpha : [0,\infty) \to [0,1]$, $x : [0,\infty) \to [0,b]$, and $v_1 : [0,\infty) \to [0,\infty)$, and a positive number T such that the following hold. For any sequence of games given by $\{\Delta(j)\}_{j=1}^{\infty}$, such that $\lim_{j\to\infty} \Delta(j) = 0$, and for any sequence of alteration-proof PBE given by $\{\hat{\alpha}(\cdot;j), \hat{x}(\cdot;j), \hat{v}_1(\cdot;j)\}_{j=1}^{\infty}$, it is the case that $(\hat{\alpha}(\cdot;j), \hat{x}(\cdot;j), \hat{v}_1(\cdot;j))$ converges uniformly to $(\alpha(\cdot), x(\cdot), v_1(\cdot))$. The limit functions and T are uniquely characterized by:*

(a) Level function $\alpha(\cdot)$ is strictly increasing and differentiable on $(0,T)$, $\alpha(0) > 0$, $\lim\limits_{t\to T^-} \alpha(t) = 1$, and $\alpha(t) = 1$ for every $t \geq T$;

(b) The cutoff-type function $x(\cdot)$ is strictly decreasing and differentiable on $(0,T)$, $x(0) = b$, $\lim\limits_{t\to T^-} x(t) = a$, and $x(t) = 1/r$ for every $t \geq T$; and

(c) On interval $(0,T)$, $\alpha(\cdot)$ and $x(\cdot)$ solve the following system of differential equations:

$$\frac{\alpha'}{\alpha} = -(rc+1)\frac{f(x)}{F(x)}x', \tag{15}$$

$$\frac{\alpha'}{\alpha} = r - \frac{1}{x} \ . \tag{16}$$

Further, for every $t \geq 0$, player 1's continuation value satisfies $v_1(t) = \alpha(t)/r$.

The existence of multiple alteration-proof equilibria in our discrete-time setting presents us with a substantial challenge for the limit characterization. Our proof of Theore 3, provided in Appendix C, uses a novel technique in three steps. First, Lemma 6 constructs two sequences of levels, type cutoffs, and continuation values; one is called the "fast sequence" and the the other is called the "slow sequence." These are constructed by setting player 1's continuation value in each period equal to lower and upper bounds that we derive from the equilibrium conditions. Then Lemma 7 shows that all alteration-proof cooperative PBE are bounded by these two sequences for small values of Δ.

Second, Lemmas 8–10 show that the fast and slow sequences uniformly converge to the same system of differential equations, Equation (15) and (16). The proof of uniform convergence utilizes the technical result proved in Watson (2021). Finally, Lemma 11 establishes the uniqueness of the limit functions and shows that T is finite. The second and third step together show that the continuous time limit is characterized by the the unique solution to the system of differential equations. The proof that T is finite depends on the assumption $a > \Delta/(1-\delta)$, which becomes $a > 1/r$ when $\Delta \to 0$.

Here is a heuristic argument. The implications of alteration-proofness expressed above as Inequalities (12) and (13) provide upper and lower bounds on player 1's continuation values on the equilibrium path. Assuming well-behaved convergence, we have

$$\lim_{j\to\infty}\left[\frac{F(\hat{x}(t;j)) - F(\hat{x}(t-\Delta(j);j))}{F(\hat{x}(t-\Delta(j);j))}\cdot\frac{c}{\hat{x}(t;j)} + \frac{F(\hat{x}(t;j))}{F(\hat{x}(t-\Delta(j);j))}\right] = 1, \tag{17}$$

and α^{τ} and $\alpha^{\tau-1}$ converge, so the lower and upper bounds of player 1's continuation value have the same continuous-time limit,

$$\lim_{j\to\infty}\frac{\Delta(j)}{1-e^{-r\Delta(j)}}\hat{\alpha}(t-\Delta(j);j) = \frac{\alpha(t)}{r}.$$

This expression uniquely determines the player 1's continuation value in equilibrium, in relation to the level function. Furthermore, the continuous-time limits of Inequality (14) and player 2's indifference Condition (8) lead to Equations (15) and (16), respectively.

The uniqueness of the system of equations shown in Theorem 3 is guaranteed by the Picard Theorem (Coddington and Levinson, 1984) . One can solve the system in closed form by completing four steps. First, we use Equation (16) to substitute for α'/α in Equation (15) to get the following univariate initial-value problem:

$$\frac{dx}{dt} = \frac{(1-rx)F(x)}{x(rc+1)f(x)}, \quad x(0) = b. \tag{18}$$

Denote

$$I_x(x) = \int \frac{x(rc+1)f(x)}{(1-rx)F(x)}dx. \tag{19}$$

Then we solve Equation (18) to obtain $x(t) = I_x^{-1}(I_x(b)+t)$. Second, to calculate T, we use the terminal condition $a = I_x^{-1}(I_x(b)+T)$. Third, we substitute the solution $x(t)$ into Equation (16) to obtain the following univariate initial-value problem:

$$\frac{d\alpha}{\alpha} = \left(r - \frac{1}{x}\right)dt, \quad \alpha(T) = 1, \tag{20}$$

which yields $\alpha(t) = \exp(I_\alpha(t) - I_\alpha(T))$, where

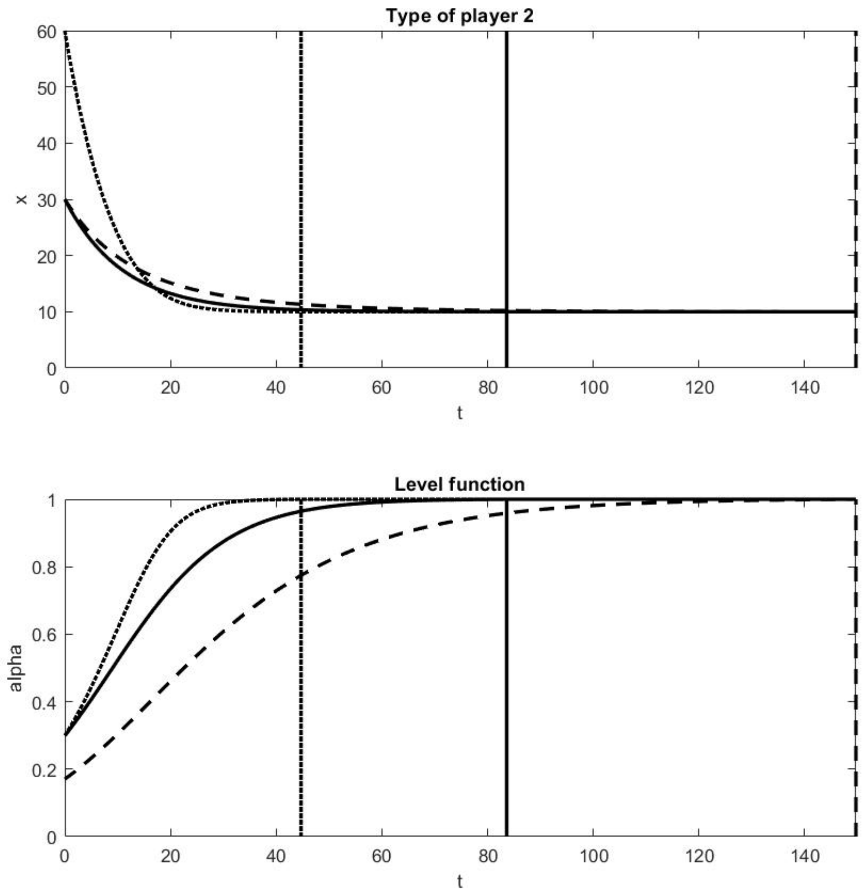
$$I_\alpha(t) = \int r - \frac{1}{I_x^{-1}(I_x(b)+t)}dt.$$

Last, we evaluate the level function at time 0 to obtain $\alpha(0) = \exp(I_\alpha(0) - I_\alpha(T))$.

Note that for any distribution F satisfying Assumption 1, the system of differential equations in Theorem 3 can be easily solved numerically. Furthermore, if the distribution function F is in polynomial form on the interval $[a,b]$, so that $F(x) = a_1x + a_2x^2 + \ldots + a_nx^n$ for real numbers $a_1, a_2, ..., a_3$, a closed-form analytical solution can be derived from a potentially complicated integration.

4 Uniformly Distributed Bad Types

In this section, we provide an example of the limit of alteration-proof equilibria when the bad types of player 2 are uniformly distributed. Fix $c = 1$ and denote the probability of the good type as $q \equiv F(0)$. For any bad type $x \in [a,b]$, we have $F(x) = q + (1-q)(x-a)/(b-a)$ and the density function is $f(x) = (1-q)/(b-a)$. In this special case, we can use the algorithm to solve the equilibrium analytically.



Continuous-time level-curve limit for alteration-proof equilibria with parameters $r = 0.1$, $a = 10.0105$, and $c = 1$. For the comparative statics, we use the following parameters: $q = 0.3337$ and $b = 30$ for solid curves, $q = 0.2002$ and $b = 30$ for dashed curves, and $q = 0.3337$ and $b = 60$ for dotted curves. Vertical lines highlight the time T when the level first hits 1 in each equilibrium.

Figure 5: The limit of alteration-proof equilibria.

In this example, Initial-Value Problems (18) and (20) become

$$\frac{dx}{dt} = \frac{1 - rx}{r+1} \cdot \frac{(b-a)q + (x-a)(1-q)}{(1-q)x}, \quad x(0) = b \tag{21}$$

$$\frac{d\alpha}{\alpha} = \left(r - \frac{1}{x}\right) dt, \quad \alpha(T) = 1. \tag{22}$$

Solving the equations, we have the explicit characterization of the equilibrium:

$$t = \frac{(1-q)(1+r)}{1-q+r(bq-a)}\left(-\frac{1}{r}\ln\frac{1-rx}{1-rb} + \frac{a-bq}{1-q}\ln\frac{(1-q)x+bq-a}{b-a}\right), \quad (23)$$

$$\alpha(t) = \left(\frac{q(b-a)}{(1-q)x(t)+bq-a}\right)^{r+1}, \quad (24)$$

$$T = \frac{(1-q)(1+r)}{1-q+r(bq-a)}\left(-\frac{1}{r}\ln\frac{1-ra}{1-rb} + \frac{a-bq}{1-q}\ln q\right), \quad \alpha(0) = q^{r+1}. \quad (25)$$

Figure 5 graphs the limit level and cutoff-type functions for three cases of parameter values where r, a, and c are fixed. The figure illustrates comparative statics with respect to q and b: Starting with the solid curve, the dashed curve shows the effect of decreasing q while the dotted curve shows the effect of increasing b. Straightforward calculations in Appendix D produce the following comparative statics conclusions. Regarding the last two statements, for any type $\chi \in [a, b]$ we define $\Gamma(\chi)$ to be the time at which the cutoff type is χ; that is, $x(\Gamma(\chi)) = \chi$.

Proposition 1. *Consider the case of uniformly distributed bad types and let q denote the probability of the good type. The equilibrium limit has the following properties, where T denotes the time when the level first reaches 1.*

- $\partial T/\partial q < 0$, $\partial T/\partial b < 0$, $\partial\alpha(0)/\partial q > 0$, *and* $\partial\alpha(0)/\partial b = 0$.
- *For a fixed type $\chi \in [a, b]$, the slope of $x(\cdot)$ at time $\Gamma(\chi)$ is decreasing in q and b, and the same is true for the slope of $\ln\alpha(t)$ at time $\Gamma(\chi)$.*

Player 1's equilibrium payoff is increasing in the quality of player 2's type distribution. Lowering the probability of the good type causes player 1 needs to start the relationship at a smaller level and gradualism slows, so it takes longer to build trust. Increasing b, the worst possible type of player 2, has the same implications.

This result may have empirical implications. For example, consider the interaction between a venture capitalist (player 1) and an entrepreneur (player 2). The venture capitalist controls the investment in a project in successive periods, which is like selecting α in our model. The entrepreneur chooses how to allocate the funds, either to productive use (copperate) or skewed to private benefit (betray). Based on the model, we would expect that the venture capitalist starts small and gradually increases funding before taking the concern public. If the project is in an industry with higher informational barriers, due for instance to sophisticated technologies or geographic distance, then we predict a low initial investment and long period before a public offering. These implications of Proposition 1 are consistent with empirical studies of venture-capital staged financing by Gompers (1995) and Tian (2011).

5 Conclusion

We have extended the starting-small literature by characterizing alteration-proof equilibria in a discrete-time principal-agent setting with a continuum of bad types. Our renegotiation concept has a stronger foundation than the version defined by Watson (1999) in that it is based on comparing actual equilibria of the continuation from any period and is a single condition (allowing alteration with general $m \geq -1$). We hope that our closed-form characterization of equilibrium is appealing and that the model presented here will be a useful step stone for further analysis of the dynamics of relationships under asymmetric information, in particular in more applied settings where multidimensional realistic ingredients are modeled (such as production technology and monitoring).

We conclude with a few technical notes on the analysis. First, we wish to emphasize that multiple alteration-proof equilibria exist in our model, as suggested by the gap between the equilibrium bounds related to Equations (12) and (13). Only in the limit does this gap shrink. We have not found a straightforward way of identifying extremal equilibria because, as shown in the Appendix, one of the bounds is not an equilibrium outcome in itself.

Second, our alteration-proofness condition may be usefully applied to other settings with incomplete information where a notion of internal consistency is desired (further expanding beyond the standard application of repeated games). A key to its applicability is that posterior beliefs have a threshold form and are monotone over time in equilibrium. In dynamic games with this property, it may be possible to describe an altered equilibrium path in terms of an adjustment in one period and a continuation that essentially jumps ahead or stalls relative to the original equilibrium.

Third, since internal consistency leads to a unique solution in the limit, there is no need for us to invoke a consideration of *external consistency*, whereby equilibria surviving internal consistency are compared. In this regard, there may be something special about games with incomplete information. For example, analysis of renegotiation of perfect public equilibrium in repeated-game settings typically involves both internal and external consistency (see Miller and Watson (2013) and Miller, Olsen, and Watson (2020), for instance).

Fourth, the assumption that player 1 can dictate the selection of an equilibrium alteration makes alteration-proofness a tight condition. Some "stall alterations" (where $m = -1$) that appeal to player 1 would not appeal to any type of player 2, and so if we required an agreement between players to coordinate on an alteration then the resulting renegotiation-proofness condition would be weaker. However, any alteration with $m \geq 0$ such as a "jump alteration" (where $m > 0$) that is desired by player 1 would also be desired by all types of player 2. If we were to assume that player 2 dictates the terms of alterations, then in an alteration-proof equilibrium, the level would start higher and rise at a rate that holds player 1's continuation value to 0 until all bad types have betrayed.

Finally, if we allow a to be arbitrarily close to $\Delta/(1-\delta)$, then in the limit, $a = 1/r$. In this case, all of our analysis still goes through, but in the limit, T becomes infinite, as Integration (19) becomes an improper integral.[15]

15. The case of $a = 1/r$ is similar to the "no gap" case in the durable-good-monopoly setting, where sales are distributed over infinite time; see Ausubel and Deneckere (1989).

Appendix A Derivation of (12), (13) and (14)

In the original equilibrium, player 1's continuation value from period τ satisfies

$$v_1^\tau = \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)(-c\alpha^\tau) + \frac{F(x^\tau)}{F(x^{\tau-1})}\left(\alpha^\tau\Delta + \delta v_1^{\tau+1}\right).$$

In the main text, we argue that if player 1 demands an alteration $m \in \{-1.1\}$ in period τ, then it suffices to consider alteration $(\tau, m, \alpha^{\tau+m})$.

First consider the case where $m = -1$. If at period τ player 1 demands that the players coordinate on a feasible alteration $(\tau, -1, \alpha^{\tau-1})$, then player 1's continuation value would instead be

$$\left(1 - \frac{F(x^{\tau-1})}{F(x^{\tau-1})}\right)\left(-c\alpha^{\tau-1}\right) + \frac{F(x^{\tau-1})}{F(x^{\tau-1})}\left(\alpha^{\tau-1}\Delta + \delta v_1^\tau\right) = \alpha^{\tau-1}\Delta + \delta v_1^\tau.$$

Therefore, alteration-proofness Condition (7) in this case simplifies to

$$v_1^\tau \geq \alpha^{\tau-1}\Delta + \delta v_1^\tau,$$

which yields Inequality (12).

Second consider the case of $m = 1$. If at period τ player 1 demands that the players coordinate on a feasible alteration $(\tau, 1, \alpha^{\tau+1})$, then player 2 with type $x \in (x^\tau, x^{\tau-1})$ betrays at level $\alpha^{\tau+1}$, while player 1 and player 2 with type $x \leq x^\tau$ continue in period τ as if they were in period $\tau + 1$ of the original equilibrium. Therefore, player 1's continuation value is

$$v_1^{\tau+1}\frac{F(x^\tau)}{F(x^{\tau-1})} - \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)c\alpha^{\tau+1}, \tag{26}$$

and we need this continuation value to be no greater than the continuation value without alteration, v_1^τ.

To apply the Condition (7), we need to express $x_1^{\tau+1}$ in terms v_1^τ. From Equation (6), we can solve for $v_1^{\tau+1}$ and get

$$v_1^{\tau+1} = \frac{1}{\delta}\left(v_1^\tau + \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)c\alpha^\tau - \frac{F(x^\tau)}{F(x^{\tau-1})}\Delta\alpha^\tau\right)\frac{F(x^{\tau-1})}{F(x^\tau)}.$$

Hence, plug $v_1^{\tau+1}$ into (26) and compare it with v_1^τ to get the following alteration-proofness condition:

$$v_1^\tau \geq \frac{1}{\delta}\left(v_1^\tau + \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)c\alpha^\tau - \frac{F(x^\tau)}{F(x^{\tau-1})}\Delta\alpha^\tau\right) - \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)c\alpha^{\tau+1}.$$

Then we multiply both sides by $-\delta$ and add v_1^τ to get

$$-(1-\delta)v_1^\tau \geq \left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)c\alpha^\tau - \frac{F(x^\tau)}{F(x^{\tau-1})}\Delta\alpha^\tau - \delta\left(1 - \frac{F(x^\tau)}{F(x^{\tau-1})}\right)c\alpha^{\tau+1},$$

which simplifies to

$$v_1^\tau \leq \frac{1}{1-\delta}\left[-(\alpha^\tau - \delta\alpha^{\tau+1})c + \frac{F(x^\tau)}{F(x^{\tau-1})}\left(\alpha^\tau - \delta\alpha^{\tau+1}\right)c + \frac{F(x^\tau)}{F(x^{\tau-1})}\alpha^\tau\Delta\right].$$

Further notice that Equation (8) can be written $\alpha^\tau - \delta\alpha^{\tau+1} = \Delta\alpha^\tau/x^\tau$. This substitution yields Inequality (13).

Last, we combine Inequalities (12) and (13). Note that player 1's continuation value satisfying both conditions only if

$$\frac{\Delta}{1-\delta}\alpha^{\tau-1} \leq \frac{\Delta}{1-\delta}\alpha^\tau\left(\frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})}\frac{c}{x^\tau} + \frac{F(x^\tau)}{F(x^{\tau-1})}\right).$$

Using player 2's indifference Condition (8) and rearranging terms, we obtain Inequality (14).

Appendix B Proof of Theorem 2

We start with a useful lemma.

Lemma 4. *If a cooperative equilibrium is "locally alteration-proof," defined as applying the conditions for $m \in \{-1, 0, 1\}$ only, then it is alteration-proof.*

Proof. We will show that the upper bound of player 1's continuation value is increasing in m, and hence the local alteration-proofness conditions are tighter than the other conditions required for alteration-proofness.

Player 1's continuation value in the original equilibrium is

$$v_1^k = \sum_{n=0}^{m}\left((c+\Delta)\frac{F(x^{k+n})}{F(x^{k-1})} - c\frac{F(x^{k+n-1})}{F(x^{k-1})}\right)\delta^n\alpha^{k+n} + \delta^{m+1}\frac{F(x^{k+m})}{F(x^{k-1})}v_1^{k+m+1}, m \geq 0.$$

Denote player 1's continuation value in a $(\tau, m, \alpha^{\tau+m})$ alteration as $\tilde{v}_1^{\tau}(m)$:

$$\tilde{v}_1^k(m) = \left((c+\Delta)\frac{F(x^{k+m})}{F(x^{k-1})} - c\right)\alpha^{k+m} + \delta\frac{F(x^{k+m})}{F(x^{k-1})}v_1^{k+m+1}.$$

For player 1 to prefer $m=0$ to $m \geq 1$, the incentive condition is $v_1^k \geq \tilde{v}_1^k(m)$, which simplifies to

$$\begin{aligned}(1-\delta^m)\delta\frac{F(x^{k+m})}{F(x^{k-1})}v_1^{k+m+1} \leq \sum_{n=0}^{m}\left((c+\Delta)\frac{F(x^{k+n})}{F(x^{k-1})} - c\frac{F(x^{k+n-1})}{F(x^{k-1})}\right)\delta^n\alpha^{k+n} \\ -\left((c+\Delta)\frac{F(x^{k+m})}{F(x^{k-1})} - c\right)\alpha^{k+m}.\end{aligned}$$

Plugging into the continuation value of original equilibrium, we get

$$\begin{aligned}v_1^k &\leq \frac{1}{1-\delta^m}\sum_{n=0}^{m-1}\left((c+\Delta)\frac{F(x^{k+n})}{F(x^{k-1})} - c\frac{F(x^{k+n-1})}{F(x^{k-1})}\right)\delta^n\alpha^{k+n} \\ &\qquad + \frac{\delta^m}{1-\delta^m}\left(1 - \frac{F(x^{k+m-1})}{F(x^{k-1})}\right)c\alpha^{k+m} \\ &= \frac{1}{1-\delta^m}\sum_{n=1}^{m}\delta^{n-1}\left[\frac{F(x^{k+n-1})}{F(x^{k-1})}[(c+\Delta)\alpha^{k+n-1} - \delta c\alpha^{k+n}] + c(\delta\alpha^{k+n} - \alpha^{k+n-1})\right] \\ &= \frac{1}{1-\delta^m}\sum_{n=1}^{m}\zeta(n),\end{aligned}$$

where

$$\zeta(n) \equiv \delta^{n-1}\left[\frac{F(x^{k+n-1})}{F(x^{k-1})}[(c+\Delta)\alpha^{k+n-1} - \delta c\alpha^{k+n}] + c(\delta\alpha^{k+n} - \alpha^{k+n-1})\right].$$

Note that this upper bound increases in m if $\delta\zeta(n) \leq \zeta(n+1)$; that is

$$\begin{aligned}0 &\geq \frac{F(x^{k-1})}{F(x^{k+n-1})}\delta^{-n}(\delta\zeta(n) - \zeta(n+1)) \\ &= \frac{F(x^{k-1}) - F(x^{k+n-1})}{F(x^{k+n-1})}c[\delta(\alpha^{k+n} - \alpha^{k+n+1}) - (\alpha^{k+n-1} - \alpha^{k+n})] \\ &\quad + \frac{F(x^{k+n-1}) - F(x^{k+n})}{F(x^{k+n-1})}[(c+\Delta)\alpha^{k+n} - \delta c\alpha^{k+n+1}] - \Delta(\alpha^{k+n} - \alpha^{k+n-1}).\end{aligned}$$

By Inequality (14), the last line is non-positive, so it suffices to require

$$\alpha^{k+n} - \alpha^{k+n-1} \geq \delta(\alpha^{k+n+1} - \alpha^{k+n}),$$

which is easily verified by plugging in Equation (8). $\square$

To prove the existence theorem, we first construct an equilibrium in which, in each period, player 1 is indifferent between continuing on the equilibrium path and altering it with $m = -1$. By construction, there is no gain to an alteration with $m = 0$. Then we show that in this equilibrium, player 1 cannot ever gain by an alteration with $m = 1$. Finally, we use Lemma 4 to argue that the constructed equilibrium is alteration-proof.

We first construct an equilibrium which, in each period, player 1 is indifferent between continuing on the equilibrium path and altering it with $m = -1$. We call this the *fast equilibrium* and denote its cutoff type sequence $\{\bar{x}^\tau(\Delta)\}_{\tau=1}^{\bar{L}(\Delta)}$, its level sequence $\{\bar{\alpha}^\tau(\Delta)\}_{\tau=1}^{\bar{L}(\Delta)}$, and its sequence of player 1's continuation values $\{\bar{v}_1^{\tau+1}(\Delta)\}_{\tau=1}^{\bar{L}(\Delta)}$. We define these sequences to period $\bar{\Delta}$. Recall that, in all later periods, the level is 1 (from Lemma 2), the cutoff type is constant at $\Delta/(1-\delta)$, and player 1's continuation value is also $\Delta/(1-\delta)$. Note that the sequences, including the period $\bar{L}$, depend on Δ; we will suppress the dependence on Δ in most of our analysis for notational simplicity.

In the fast equilibrium, Condition (12) binds. Using the equilibrium Identity (6), we get that $\{\bar{x}^\tau\}$ and $\{\bar{\alpha}^\tau\}$ must satisfy

$$\frac{F(\bar{x}^{\tau-1}) - F(\bar{x}^\tau)}{\Delta F(\bar{x}^{\tau-1})} = \frac{\bar{\alpha}^\tau - \bar{\alpha}^{\tau-1}}{(\Delta + c(1-\delta))\bar{\alpha}^\tau}, \text{ for } \tau \in \{2, 3, ..., \bar{L} - 1\} \tag{27}$$

Furthermore, facing this sequence of levels, player 2 of type $\bar{x}^\tau$ satisfying Equation (8) cooperates until period τ and betrays in period $\tau + 1$.

The fast equilibrium is constructed recursively as follows, and pictured in Figure 6. Whatever $\bar{L}$ turns out to be, the terminal condition requires $\bar{x}^{\bar{L}-1} = a$, $\bar{\alpha}^{\bar{L}} = a\delta/(a - \Delta)$ and $\bar{v}_1^{\bar{L}} = \Delta/(1-\delta)$. We also set $\bar{x}^{\bar{L}} = \Delta/(1-\delta)$ to match with Lemma 2 and 3. For each $\tau \leq \bar{L} - 1$ and with equilibrium values x^τ and α^τ given, the equilibrium values for the previous period, $\bar{x}^{\tau-1}$ and $\bar{\alpha}^{\tau-1}$, are set to solve the system of Equations (8) and (27). We continue in an iterative fashion until the value of the cutoff type weakly exceeds b, which identifies the first period and thus the number

$\bar{L}$. With the sequences $\{\bar{x}^\tau\}_{\tau=1}^{\bar{L}}$ and $\{\bar{\alpha}^\tau\}_{\tau=1}^{\bar{L}}$, the sequence of player 1's continuation values is given by $\bar{v}_1^\tau = \bar{\alpha}^\tau \Delta/(1-\delta)$.

It is easy to verify that in the constructed equilibrium, player 1 weakly prefers staying on the equilibrium path rather than altering it with $m = 1$, because

$$\frac{F(\bar{x}^{\tau-1}) - F(\bar{x}^\tau)}{\Delta F(\bar{x}^{\tau-1})} = \frac{\bar{\alpha}^\tau - \bar{\alpha}^{\tau-1}}{(\Delta + c(1-\delta))\bar{\alpha}^\tau} \leq \frac{\bar{\alpha}^\tau - \bar{\alpha}^{\tau-1}}{(c+\Delta)\bar{\alpha}^\tau - \delta c \bar{\alpha}^{\tau+1}}$$

is the case when (8) is satisfied. Using Lemma 4, we have the result.

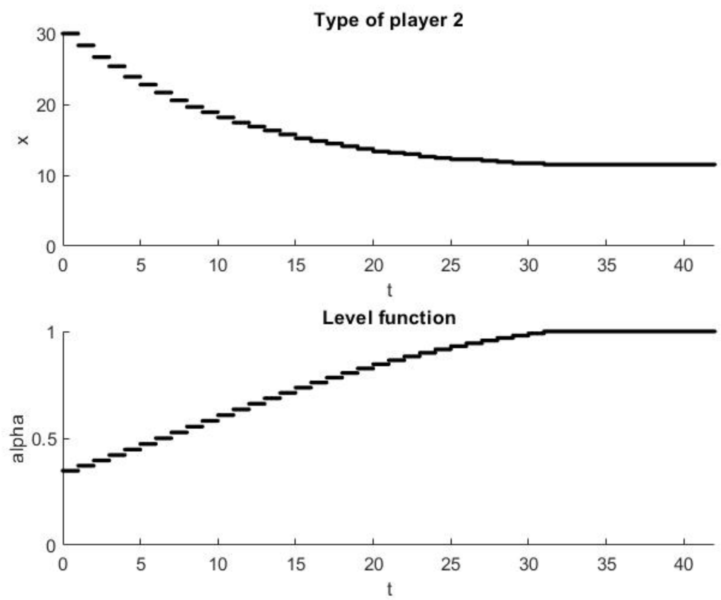


Figure 6: Existence of alteration-proof equilibrium

Appendix C Proof of Theorem 3

Parallel to the fast equilibrium constructed in Appendix B, for any given Δ we define what we call the *slow sequence* denoted $\{\underline{x}^\tau(\Delta), \underline{\alpha}^\tau(\Delta), \underline{v}_1^{\tau+1}(\Delta)\}_{\tau=1}^{\underline{L}(\Delta)}$. This sequence is constructed to impose in every period a condition that relates to player 1 being indifferent to alterations with $m = 1$. The condition will combine values of α^τ, x^τ, and $v_1^{\tau+1}$ that do not necessarily arise in an equilibirum, and therefore the slow sequence is not necessarily an equilibrium, but it will provide a bound on the equilibria that helps us to characterize the continuous-time limit. Note again that these sequences depend on Δ, an argument we will sometimes suppress for notational simplicity.

In the slow sequence, (13) binds in each period. Together with Equation (6), we get that $\{\underline{x}^\tau\}$ and $\{\underline{\alpha}^\tau\}$ must satisfy

$$\frac{F(\underline{x}^\tau) - F(\underline{x}^{\tau-1})}{F(\underline{x}^\tau)} = \frac{F(\underline{x}^\tau) - F(\underline{x}^{\tau+1})}{F(\underline{x}^\tau)} \frac{-\delta\underline{\alpha}^{\tau+2} + \underline{\alpha}^{\tau+1}(1+\Delta/c)}{\underline{\alpha}^{\tau+1} - \underline{\alpha}^\tau} - \frac{\Delta}{c}, \tag{28}$$

for $\tau \in \{2, 3, ..., \underline{L} - 2\}$. Similarly, facing this sequence of levels, player 2 of type $\underline{x}^\tau$ satisfying Equation (8) cooperates until period τ and betrays in period $\tau + 1$.

We construct the slow sequence recursively as follows. Whatever is $\underline{L}$, the terminal condition requires $\underline{x}^{\underline{L}-1} = a$ and $\underline{\alpha}^{\underline{L}} = \underline{\alpha}^{\underline{L}-1} = 1$. We also set $\underline{x}^{\underline{L}} = \Delta/(1-\delta)$ to match with Lemma 2 and 3. For each $\tau \leq \underline{L}$ and values x^τ, α^τ, and $\alpha^{\tau+1}$ given, values $x^{\tau-1}$ and $\alpha^{\tau-1}$ for the previous period are set to solve the system of Equations (8) and (28). We continue in an iterative fashion until the value of the cutoff type weakly exceeds b, which identifies the first period and hence the number $\underline{L}$. With the sequences $\{\underline{x}^\tau, \underline{\alpha}^\tau\}_{\tau=1}^{\underline{L}}$, player 1's continuation value sequence is derived from

$$\underline{v}_1^\tau = \frac{\Delta}{1-\delta}\underline{\alpha}^\tau \left(\frac{F(\underline{x}^\tau) - F(\underline{x}^{\tau-1})}{F(\underline{x}^{\tau-1})} \cdot \frac{c}{\underline{x}^\tau} + \frac{F(\underline{x}^\tau)}{F(\underline{x}^{\tau-1})} \right).$$

To prove Theorem 3, we first show that all alteration-proof equilibria are bounded by the fast equilibrium and the slow sequence (Lemma 6–7), and then show that the $\{\alpha^\tau\}$ and $\{x^\tau\}$ in both the fast equilibrium and slow sequence converge to the solution to differential Equations (15) and (16) (Lemma 8–10). Finally, we apply the Picard Theorem (Coddington and Levinson, 1984) to show that the solution to the system of differential equations is unique.

Lemma 5. *When Δ is small, for any alteration-proof PBE, $x^{\tau+1} - x^\tau$ and $\alpha^{\tau+1} - \alpha^\tau$ are both on the order of Δ.*

Proof. From player 2's indifference condition, we can write x^τ as

$$x^\tau = \frac{\Delta\alpha^\tau}{\alpha^\tau - \delta\alpha^{\tau+1}} = \frac{\Delta}{1 - \delta\alpha^{\tau+1}/\alpha^\tau}.$$

Because x^τ is in $[a, b]$ for all $\Delta > 0$, so is the right most term, which simplifies to

$$\frac{\Delta}{\delta}\left(\frac{1-\delta}{\Delta} - \frac{1}{a}\right) \leq \frac{\alpha^{\tau+1} - \alpha^\tau}{\alpha^\tau} \leq \frac{\Delta}{\delta}\left(\frac{1-\delta}{\Delta} - \frac{1}{b}\right). \tag{29}$$

Note the left and right most terms are on the order of $\alpha^\tau \Delta$, and when we have α^τ strictly positive, then $\alpha^{\tau+1} - \alpha^\tau$ must be on the order of Δ. Second, rewrite Inequality (14) to get

$$0 \geq \frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})} \geq \frac{c(\alpha^\tau - \delta\alpha^{\tau+1}) + \Delta\alpha^{\tau-1}}{(c+\Delta)\alpha^\tau - \delta c\alpha^{\tau+1}} - 1 = \frac{\alpha^{\tau-1} - \alpha^\tau}{\alpha^\tau(c/x^\tau + 1)}, \tag{30}$$

where the left-most term is 0 and right-most term is on the order of Δ, so $F(x^\tau) - F(x^{\tau-1})$ is also on the order of Δ. By Assumption 1 (in particular that F' is bounded away from 0) and because $F(x^{\tau-1})$ is bounded away from zero owing to the probability of the good type, we conclude that $x^\tau - x^{\tau-1}$ is on the order of Δ. □

Corollary 1. *When Δ is small, for any alteration-proof PBE and period τ, $(v_1^\tau/\alpha^{\tau-1}) - \Delta/(1-\delta)$ is on the order of Δ.*

Proof. From Inequalities (12) and (13), we have

$$\frac{\Delta}{1-\delta}\alpha^{\tau-1} \leq v_1^\tau \leq \frac{\Delta}{1-\delta}\alpha^\tau \left(\frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})} \cdot \frac{c}{x^\tau} + \frac{F(x^\tau)}{F(x^{\tau-1})} \right).$$

Dividing each term by $\alpha^{\tau-1}$ and subtracting $\Delta/(1-\delta)$, we get

$$0 \leq \frac{v_1^\tau}{\alpha^{\tau-1}} - \frac{\Delta}{1-\delta} \leq \frac{\Delta}{1-\delta} \left[\frac{\alpha^\tau}{\alpha^{\tau-1}} \left(\frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})} \cdot \frac{c}{x^\tau} + \frac{F(x^\tau)}{F(x^{\tau-1})} \right) - 1 \right]$$

where the expression in the square brackets simplifies to

$$\frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})} \cdot \frac{\alpha^\tau}{\alpha^{\tau-1}} \left(\frac{c}{x^\tau} + 1 \right) + \frac{\alpha^\tau - \alpha^{\tau-1}}{\alpha^{\tau-1}}.$$

Apply Inequalities (29) and (30), this term is bounded above by a function of order Δ, that is

$$\frac{F(x^\tau) - F(x^{\tau-1})}{F(x^{\tau-1})} \cdot \frac{\alpha^\tau}{\alpha^{\tau-1}} \left(\frac{c}{x^\tau} + 1 \right) + \frac{\alpha^\tau - \alpha^{\tau-1}}{\alpha^{\tau-1}} \leq \frac{\Delta}{\delta} \left(\frac{1-\delta}{\Delta} - \frac{1}{b} \right).$$

The claim follows from the fact that $v_1^\tau/\alpha^{\tau-1} - \Delta/(1-\delta)$ is positive and bounded above by a function of order Δ. □

Lemma 6. *Fix a, b, r, and F. There exists $\Delta_0 > 0$, such that for all $\Delta < \Delta_0$, the following holds. Consider any alteration-proof PBE given by cutoff type se-*

quence $\{x^\tau\}_{\tau=1}^{L-1}$*, level sequence* $\{\alpha^\tau\}_{\tau=1}^{L-1}$*, and player 1's continuation value sequence* $\{v_1^{\tau+1}\}_{\tau=1}^{L-1}$*, in period* $\tau \in \{2, 3, ..., L-1\}$*. It is true that* $x^{\tau-1} \in [\underline{\tilde{x}}^{\tau-1}, \tilde{\tilde{x}}^{\tau-1}]$*, where*

$$\left(\frac{\Delta}{1-\delta}\frac{\delta\tilde{\tilde{x}}^{\tau-1}}{\tilde{\tilde{x}}^{\tau-1}-\Delta}+c\right)F(\tilde{\tilde{x}}^{\tau-1}) = F(x^\tau)\left(\delta\frac{v_1^{\tau+1}}{\alpha^\tau}+c+\Delta\right), \tag{31}$$

$$F(\underline{\tilde{x}}^{\tau-1}) = F(x^\tau)+\delta\frac{\frac{v_1^{\tau+1}}{\alpha^\tau}-\frac{\Delta}{1-\delta}}{c(1-\frac{\Delta}{1-\delta}\frac{1}{x^\tau})}F(x^\tau). \tag{32}$$

Note that $\underline{\tilde{x}}^{\tau-1}$ and $\tilde{\tilde{x}}^{\tau-1}$ are not necessarily equilibrium values. Mathematically, for any given values of x^τ, α^τ and $v_1^{\tau+1}$, they are defined as the values of $x^{\tau-1}$ that make Inequalities (12) and (13) bind.

Proof. Fix continuation value $v_1^{\tau+1}$ and consider an arbitrary $x^{\tau-1}$. Equilibrium Identity (6) and alteration-proofness Condition (12) imply

$$\left(1-\frac{F(x^\tau)}{F(x^{\tau-1})}\right)(-c\alpha^\tau)+\frac{F(x^\tau)}{F(x^{\tau-1})}\left(\alpha^\tau\Delta+\delta v_1^{\tau+1}\right) \geq \frac{\Delta}{1-\delta}\alpha^{\tau-1}.$$

Rearrange terms and we get

$$\left(\frac{\Delta}{1-\delta}\frac{\delta x^{\tau-1}}{x^{\tau-1}-\Delta}+c\right)F(x^{\tau-1}) \leq \left(\delta\frac{v_1^{\tau+1}}{\alpha^\tau}+c+\Delta\right)F(x^\tau),$$

where the left side increases in $x^{\tau-1}$ when Δ is small, because Assumption 1 requires F' to be bounded away from 0. So this condition translates into $F(x^{\tau-1}) \leq F(\tilde{\tilde{x}}^{\tau-1})$, where $\tilde{\tilde{x}}^{\tau-1}$ satisfies Equation (31).

Combining Equations (6) and (13), we get

$$\left(1-\frac{F(x^\tau)}{F(x^{\tau-1})}\right)(-c\alpha^\tau)+\frac{F(x^\tau)}{F(x^{\tau-1})}\left(\alpha^\tau\Delta+\delta v_1^{\tau+1}\right) \leq \frac{\Delta}{1-\delta}\alpha^\tau\left(\frac{F(x^\tau)-F(x^{\tau-1})}{F(x^{\tau-1})}\cdot\frac{c}{x^\tau}+\frac{F(x^\tau)}{F(x^{\tau-1})}\right),$$

which simplifies to

$$\frac{c+\Delta+\delta\frac{v_1^{\tau+1}}{\alpha^\tau}-\frac{\Delta}{1-\delta}(\frac{c}{x^\tau}+1)}{c(x^\tau-\frac{\Delta}{1-\delta})\frac{1}{x^\tau}} \leq \frac{F(x^{\tau-1})}{F(x^\tau)}.$$

The right side increases in $x^{\tau-1}$ for fixed x^{τ}. So this condition translates into $F(x^{\tau-1}) \geq F(\underline{\tilde{x}}^{\tau-1})$, where $\underline{\tilde{x}}^{\tau-1}$ satisfies Equation (32). $\square$

From Lemma 6, player 2's cutoff type in period $\tau - 1$ (in any alteration-proof equilibrium) is between $\underline{\tilde{x}}^{\tau-1}$ and $\bar{\tilde{x}}^{\tau-1}$. Lemma 7 shows if $\alpha^{\tau}, x^{\tau}, v_1^{\tau+1}$ turn out to be from the same equilibrium, then $\underline{\tilde{x}}^{\tau-1}$ and $\bar{\tilde{x}}^{\tau-1}$ are the cutoff-type values in the slow and fast sequences, provided that Δ is small. Recall that the fast equilibrium and slow sequence are denoted $\{\bar{x}^{\tau}(\Delta), \bar{\alpha}^{\tau}(\Delta), \bar{v}_1^{\tau+1}(\Delta)\}_{\tau=1}^{\bar{L}(\Delta)}$ and $\{\underline{x}^{\tau}(\Delta), \underline{\alpha}^{\tau}(\Delta), \underline{v}_1^{\tau+1}(\Delta)\}_{\tau=1}^{\underline{L}(\Delta)}$, and these depend on Δ (and hence on j, suppressed here).

Lemma 7. *Fix a, b, r, and F. There exists $\Delta_1 > 0$ such that for all $\Delta < \Delta_1$ and any alteration-proof PBE given by cutoff-type sequence $\{x^{\tau}\}_{\tau=1}^{L}$, level sequence $\{\alpha^{\tau}\}_{\tau=1}^{L}$, and player 1's continuation value sequence $\{v_1^{\tau+1}\}_{\tau=1}^{L}$, this equilibrium has the property that for all $\pi = 0, 1, 2, \ldots, \min\{\underline{L}(\Delta), L(\Delta), \bar{L}(\Delta)\}$,*

$$x^{L-\pi} \in [\underline{x}^{\underline{L}-\pi}(\Delta), \bar{x}^{\bar{L}-\pi}(\Delta)], \quad \alpha^{L-\pi} \in [\bar{\alpha}^{\bar{L}-\pi}(\Delta), \underline{\alpha}^{\underline{L}-\pi}(\Delta)],$$
$$\textit{and } v_1^{L-\pi+1} \in [\bar{v}_1^{\bar{L}-\pi+1}(\Delta), \underline{v}_1^{\underline{L}-\pi+1}(\Delta)]. \quad (33)$$

Proof. When $\pi = 0$, by construction, $\underline{x}^{\underline{L}} = \bar{x}^{\bar{L}} = \Delta/(1-\delta)$, $\underline{\alpha}^{\underline{L}} = 1$ and $\bar{\alpha}^{\bar{L}} = a\delta/(1-\Delta)$. By Lemma 2 and 3, we know that for any alteration-proof PBE, $x^L = \Delta/(1-\delta)$ and $\alpha^L \in [a\Delta/(1-\delta), 1]$. When $\pi = 1$, then by construction $\underline{x}^{\underline{L}-1} = \bar{x}^{\bar{L}-1} = a$. When Δ is close to 0, then for period L to be the last period betrayal happens, it must also be that $x^{L-1} = a$. Therefore (33) trivially holds for $\pi = 0$ and $\pi = 1$. It suffices to consider $\pi > 1$.

First, we want to show $v_1^{L-\pi+1} \in [\bar{v}_1^{\bar{L}-\pi+1}(\Delta), \underline{v}_1^{\underline{L}-\pi+1}(\Delta)]$. Note that $\bar{v}_1^{\bar{L}-\pi+1}$ is defined such that Inequality (12) binds, and $\underline{v}_1^{\underline{L}-\pi+1}$ is defined such that Inequality (13) binds. For the equilibrium to be alteration-proof, player 1's continuation value $v_1^{L-\pi+1}$ must satisfy Inequalities (12) and (13). So $v_1^{L-\pi+1} \in [\bar{v}_1^{\bar{L}-\pi+1}(\Delta), \underline{v}_1^{\underline{L}-\pi+1}(\Delta)]$ is trivially satisfied.

Second, if we knew $x^{L-\pi} \in [\underline{x}^{\underline{L}-\pi}(\Delta), \bar{x}^{\bar{L}-\pi}(\Delta)]$, then $\alpha^{L-\pi} \in [\bar{\alpha}^{\bar{L}-\pi}(\Delta), \underline{\alpha}^{\underline{L}-\pi}(\Delta)]$ trivially holds. This result follows directly from player 2's indifference Condition (8).

The rest of the proof shows $x^{L-\pi} \in [\underline{x}^{\underline{L}-\pi}(\Delta), \bar{x}^{\bar{L}-\pi}(\Delta)]$ by induction. Assume $x^{L-\pi} \in [\underline{x}^{\underline{L}-\pi}(\Delta), \bar{x}^{\bar{L}-\pi}(\Delta)]$ holds, then we want to show it is also the case that $x^{L-\pi-1} \in [\underline{x}^{\underline{L}-\pi-1}(\Delta), \bar{x}^{\bar{L}-\pi-1}(\Delta)]$. From Lemma 6, we know that upper and lower bounds of $x^{L-\pi}$, that is $\bar{\tilde{x}}^{L-\pi-1}$ and $\underline{\tilde{x}}^{L-\pi-1}$, are characterized by Equations (31) and

(32). Hence, it suffices to show that $\tilde{\underline{x}}^{L-\pi-1} = \underline{x}^{\underline{L}-\pi-1}$ and $\tilde{\bar{x}}^{L-\pi-1} = \bar{x}^{\bar{L}-\pi-1}$.

For fixed $x^{L-\pi} \in [\underline{x}^{\underline{L}-\pi}(\Delta), \bar{x}^{\bar{L}-\pi}(\Delta)]$ and corresponding $\alpha^{L-\pi}$ and $v_1^{L-\pi}$ satisfying Equations (6) and (8), if the right side of Equations (31) and (32) are monotonically increasing in $x^{L-\pi}$, then the maximum $\tilde{\bar{x}}^{L-\pi-1}$ and minimum $\tilde{\underline{x}}^{\underline{L}-\pi-1}$ are achieved when $x^{L-\pi}$ is at its upper and lower bound, respectively. Note that Equation (31) is defined with $v_1^{L-\pi-1}$ at its lower bound; when we further impose $x^{L-\pi} = \bar{x}^{\bar{L}-\pi}$ making $v_1^{L-\pi}$ at its lower bound, Equation (31) coincide with Equation (27) and hence $\tilde{\bar{x}}^{L-\pi-1} = \bar{x}^{\bar{L}-\pi-1}$. Similarly, Equation (32) is defined with $v_1^{L-\pi-1}$ at its upper bound; when we further impose $x^{L-\pi} = \underline{x}^{\underline{L}-\pi}$ making $v_1^{L-\pi}$ at its upper bound, Equation (32) coincide with Equation (28) and hence $\tilde{\underline{x}}^{L-\pi-1} = \underline{x}^{\underline{L}-\pi-1}$.

We complete the proof by showing that the right side of Equations (31) and (32) are monotonically increasing in $x^{L-\pi}$ if Δ is close to 0. From the indifference Condition (8), fixing $x^{L-\pi+1}$, $\alpha^{L-\pi}$ decreases in $x^{L-\pi}$. From the equilibrium Identity (6), we know that

$$\begin{aligned}
\frac{v_1^{L-\pi+1}}{\alpha^{L-\pi}} F(x^{L-\pi}) &= \Big(F(x^{L-\pi}) - F(x^{L-\pi+1})\Big)\left(-c\frac{\alpha^{L-\pi+1}}{\alpha^{L-\pi}}\right) \\
&\quad + F(x^{L-\pi+1})\left(\frac{\alpha^{L-\pi+1}}{\alpha^{L-\pi}}\Delta + \delta\frac{v_1^{L-\pi+2}}{\alpha^{L-\pi}}\right) \\
&= \Delta\left(-c\frac{F(x^{L-\pi}) - F(x^{L-\pi+1})}{\Delta} + F(x^{L-\pi+1})\right)\frac{\alpha^{L-\pi+1}}{\alpha^{L-\pi}} \\
&\quad + F(x^{L-\pi+1})\delta\frac{v_1^{L-\pi+2}}{\alpha^{L-\pi}}.
\end{aligned}$$

Roughly speaking, the first term could be decreasing in $x^{L-\pi}$ but is on the order of Δ, while the second term must be increasing in $x^{L-\pi}$ and is on the order of 1. Therefore, for Δ sufficiently small, the monontonicity of the left side is determined by the second term. Hence, the right side of Equation (31) is increasing in $x^{L-\pi}$ when Δ is small. Similarly, consider the right side of Equation (32). The first term must be increasing in $x^{L-\pi}$ and is on the order of 1, while the second term could be decreasing in $x^{L-\pi}$ but is on the order of Δ. By the same argument, for Δ sufficiently small, the monontonicity is determined by the first term, so the right side of Equation (32) is increasing in $x^{L-\pi}$ when Δ is small. $\square$

We have shown that all discrete-time alteration-proof PBE are bounded by fast equilibrium and the slow sequence, and the rest of the proof will show that the fast and slow sequences uniformly converge to the same set of differential equations. Lemmas 8–10 apply the general result of the relationship between discrete time and continuous time equilibrium in Watson (2021) to derive the differential Equations (15) and (16) that characterize the continuous time alteration-proof PBE.

Take as given a sequence $\{\Delta(j)\}_{j=1}^{\infty}$ that converges to 0. Recall that the slow and fast sequences depend on the value of Δ, which depends on j. As with the earlier calculations, we sometimes suppress the argument j as well as $\Delta(j)$ to ease notation.

We first define step functions using the fast and slow sequences, and then derive the limit of the step functions as $j \to \infty$. Recall $M(t,j) = \min\{k \mid k\Delta(j) \geq t\}$ is the period in discrete-time game j that contains time t. Define step functions

$$\hat{\bar{\alpha}}(\cdot;j) : [0,\infty] \to [0,1],\ \ \hat{\bar{x}}(\cdot;j) : [0,\infty] \to [0,b],\ \ \hat{\bar{v}}_1(\cdot;j) : [0,\infty] \to [0,\infty),$$

$$\hat{\underline{\alpha}}(\cdot;j) : [0,\infty] \to [0,1],\ \ \hat{\underline{x}}(\cdot;j) : [0,\infty] \to [0,b],\ \ \hat{\underline{v}}_1(\cdot;j) : [0,\infty] \to [0,\infty)$$

by

$$\hat{\bar{\alpha}}(t;j) = \bar{\alpha}^{M(t,j)}(\Delta(j)),\ \ \hat{\bar{x}}(t;j) = \bar{x}^{M(t,j)}(\Delta(j)),\ \ \hat{\bar{v}}_1(t;j) = \bar{v}_1^{M(t,j)}(\Delta(j)),$$

$$\hat{\underline{\alpha}}(t;j) = \underline{\alpha}^{M(t,j)}(\Delta(j)),\ \ \hat{\underline{x}}(t;j) = \underline{x}^{M(t,j)}(\Delta(j)),\ \ \hat{\underline{v}}_1(t;j) = \underline{v}_1^{M(t,j)}(\Delta(j)).$$

We introduce the following notation to simplify computation:

$$A(x^\tau,\Delta) \equiv \frac{x^\tau(1-\delta)-\Delta}{x^\tau-\Delta},\ \ B(\Delta) \equiv \frac{\Delta}{c(1-\delta)+\Delta},\ \ C(x^{\tau+1}) \equiv \frac{1}{x^{\tau+1}}+\frac{1}{c},$$
$$D(x^{\tau+1},x^\tau) \equiv F(x^{\tau+1}) - F(x^\tau),\ \ E(x^\tau,x^{\tau-1},\Delta) \equiv F(x^{\tau-1}) - \frac{\Delta+c}{c}F(x^\tau).$$

Lemma 8. *As $j \to \infty$, $\hat{\bar{x}}(\cdot,j)$ uniformly converges to function $x : [0,\infty) \to [0,b]$ solving* (18).

Proof. Recall that $\hat{\bar{x}}(\cdot,j)$ is defined by player 2's cutoff type sequence $\{\bar{x}^{M(t,j)}(\Delta(j))\}$ in the fast equilibrium that satisfies (8) and (27). We first rewrite Equation (27) as

$$F(\bar{x}^{M(t,j)+1}) = F(\bar{x}^{M(t,j)}) - F(\bar{x}^{M(t,j)})A(\bar{x}^{M(t,j)},\Delta)B(\Delta), \tag{34}$$

which implicitly defines the recursive formulation of $\bar{x}^{M(t,j)}(\Delta(j))$. In the rest of this proof, we use $\tau = M(t,j)$ and suppress the dependence on j and Δ to simplify notation. We define the function $\bar{\sigma} : [a,b] \times [0,\infty] \to [a,b]$ by $\bar{x}^{\tau+1} \equiv \bar{\sigma}(\bar{x}^{\tau}, \Delta)$ for $\Delta > 0$, and $\bar{x}^{\tau} = \bar{\sigma}(\bar{x}^{\tau}, 0) \equiv \lim_{\Delta\to 0} \bar{\sigma}(\bar{x}^{\tau}, \Delta)$, such that $\bar{x}^{\tau}$, $\bar{x}^{\tau+1}$ and Δ satisfy Equation (34).

By Watson (2021), the sequence $\{\bar{x}^{\tau}\}$ with $\tau = M(t;j)$, defined by recursive formulation $\bar{x}^{\tau+1} = \bar{\sigma}(\bar{x}^{\tau}, \Delta)$ and initial condition $\bar{x}^1 = \bar{x}_1$, converges to function $\bar{x} : [0,\infty) \to [a,b]$ uniformly as $\Delta(j) \to 0$, and function $\bar{x}$ satisfies

$$\frac{d\bar{x}}{dt} = \lim_{\Delta\to 0} \frac{\partial \bar{\sigma}}{\partial \Delta}(\bar{x}^{\tau}, \Delta).$$

To apply this result, we take the partial derivative of $\bar{x}^{\tau+1} = \bar{\sigma}(\bar{x}^{\tau}, \Delta)$ with respect to Δ. From (34), we use Implicit Function Theorem to differentiate with respect to Δ and get

$$-\frac{f(\bar{x}^{\tau+1})}{F(\bar{x}^{\tau})}\frac{\partial \bar{x}^{\tau+1}}{\partial \Delta} = \frac{\partial A(\bar{x}^{\tau}, \Delta)}{\partial \Delta} B(\Delta) + \frac{\partial B(\Delta)}{\partial \Delta} A(\bar{x}^{\tau}, \Delta), \tag{35}$$

where

$$\frac{\partial A(\bar{x}^{\tau}, \Delta)}{\partial \Delta} = \frac{\bar{x}^{\tau}\delta r - 1}{\bar{x}^{\tau} - \Delta} + \frac{\bar{x}^{\tau}(1-\delta) - \Delta}{(\bar{x}^{\tau} - \Delta)^2}, \quad \frac{\partial B(\Delta)}{\partial \Delta} = \frac{1}{c(1-\delta)+\Delta} - \frac{\Delta(cr\delta + 1))}{(c(1-\delta)+\Delta)^2}.$$

Note that the above two equations are continuous and bounded. With initial condition $\bar{x} = a$ and $\bar{x}(t) = \lim_{j\to\infty} \hat{\bar{x}}(t;j)$, we have

$$\lim_{\Delta\to 0} \frac{A(\bar{x}^{\tau}, \Delta)}{\Delta} = \frac{\bar{x}(t)r - 1}{\bar{x}(t)}, \quad \lim_{\Delta\to 0} \Delta \frac{\partial B(\Delta)}{\partial \Delta} = \frac{1}{cr+1} - \frac{cr+1}{(cr+1)^2} = 0,$$

$$\lim_{\Delta\to 0} \frac{\partial A(\bar{x}^{\tau}, \Delta)}{\partial \Delta} = \frac{\bar{x}(t)r - 1}{\bar{x}(t)}, \quad \lim_{\Delta\to 0} B(\Delta) = \frac{1}{cr+1}.$$

Take limit of Equation (35) as $\Delta \to 0$ and plug in the above limits, we get

$$\frac{d\bar{x}}{dt} = \lim_{\Delta\to 0} \frac{\partial \bar{x}^{\tau+1}}{\partial \Delta} = -\frac{F(\bar{x})}{f(\bar{x})} \cdot \frac{\bar{x}r - 1}{\bar{x}(cr+1)},$$

which is (18). □

Lemma 9. *As $j \to \infty$, $\hat{\underline{x}}(\cdot, j)$ uniformly converges to function $x : [0,\infty) \to [0,b]$ solving* (18).

Proof. Recall that, $\hat{\underline{x}}(\cdot, j)$ is defined by player 2's cutoff type sequence $\{\underline{x}^{M(t,j)}(\Delta(j))\}$ in the slow sequence that satisfies (8) and (28), To simplify Equation (28), we first multiply both sides by $F(\underline{x}^{\tau})(\underline{\alpha}^{\tau}/\underline{\alpha}^{\tau+1}-1)$

$$\begin{aligned}\left(1-\frac{\underline{\alpha}^{\tau}}{\underline{\alpha}^{\tau+1}}\right)&(F(\underline{x}^{\tau-1})-F(\underline{x}^{\tau}))\\&=\left(F(\underline{x}^{\tau+1})-F(\underline{x}^{\tau})\right)\left(-\delta\frac{\underline{\alpha}^{\tau+2}}{\underline{\alpha}^{\tau+1}}+(1+\Delta/c)\right)+\left(1-\frac{\underline{\alpha}^{\tau}}{\underline{\alpha}^{\tau+1}}\right)\frac{\Delta}{c}F(\underline{x}^{\tau}),\end{aligned}$$

then substitute for $\underline{\alpha}^{\tau}/\underline{\alpha}^{\tau+1}$ and $\underline{\alpha}^{\tau+2}/\underline{\alpha}^{\tau+1}$ using player 2's indifference Condition (8), and finally collect terms to get

$$0=A(\underline{x}^{M(t,j)},\Delta)E(\underline{x}^{M(t,j)},\underline{x}^{M(t,j)-1},\Delta)-\Delta C(\underline{x}^{M(t,j)+1})D(\underline{x}^{M(t,j)+1},\underline{x}^{M(t,j)}).\quad(36)$$

which implicitly defines the recursive formulation of $\underline{x}^{M(t,j)}(\Delta(j))$. In the rest of this proof, we use $\tau=M(t,j)$ and suppress the dependence on j and Δ to simplify notation. We define the function $\underline{\sigma}:[a,b]\times[a,b]\times[0,\infty]\to[a,b]$ by $\underline{x}^{\tau+1}\equiv\underline{\sigma}(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)$ for $\Delta>0$, and $\underline{x}^{\tau}=\underline{\sigma}(\underline{x}^{\tau},\underline{x}^{\tau},0)\equiv\lim_{\Delta\to0}\underline{\sigma}(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)$, such that $\underline{x}^{\tau-1}$, $\underline{x}^{\tau}$, $\underline{x}^{\tau+1}$ and Δ satisfy Equation (36).

By Watson (2021), the sequence $\{\underline{x}^{\tau}\}$ with $\tau=M(t;j)$, defined by initial conditions $\underline{x}^0=\underline{x}_0$, $\underline{x}^1=\underline{x}_1$ and recursive formulation $\underline{x}^{\tau+1}=\underline{\sigma}(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)$, converges to function $\underline{x}:[0,\infty)\to[a,b]$ uniformly as $\Delta(j)\to0$, and function $\underline{x}$ satisfies

$$\frac{d\underline{x}}{dt}=\lim_{\Delta\to0}\frac{\frac{\partial\underline{\sigma}}{\partial\Delta}(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)}{1+\frac{\partial\underline{\sigma}}{\partial\underline{x}^{\tau-1}}(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)}.\quad(37)$$

To apply this result, we next find partial derivatives of $\underline{x}^{\tau+1}=\underline{\sigma}(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)$ with respect to $\underline{x}^{\tau-1}$ and Δ. We do so by using Implicit Function Theorem and differentiate Equation (36).

First, we find $\partial\underline{x}^{\tau+1}/\partial\Delta$, we differentiate Equation (36) with respect to Δ and get

$$\begin{aligned}0=\frac{\partial A(\underline{x}^{\tau},\Delta)}{\partial\Delta}&E(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)+A(\underline{x}^{\tau},\Delta)\frac{\partial E(\underline{x}^{\tau},\underline{x}^{\tau-1},\Delta)}{\partial\Delta}-C(\underline{x}^{\tau+1})D(\underline{x}^{\tau+1},\underline{x}^{\tau})\\&-\Delta\frac{\partial C(\underline{x}^{\tau+1})}{\partial\underline{x}^{\tau+1}}D(\underline{x}^{\tau+1},\underline{x}^{\tau})\frac{\partial\underline{x}^{\tau+1}}{\partial\Delta}-\Delta C(\underline{x}^{\tau+1})\frac{\partial D(\underline{x}^{\tau+1},\underline{x}^{\tau})}{\partial\underline{x}^{\tau+1}}\cdot\frac{\partial\underline{x}^{\tau+1}}{\partial\Delta}.\end{aligned}$$

Use (36) to substitute $E(\underline{x}^\tau, \underline{x}^{\tau-1}, \Delta)$ and then solve for $\partial \underline{x}^{\tau+1}/\partial\Delta$ and get

$$\frac{\partial \underline{x}^{\tau+1}}{\partial \Delta} = \frac{C(\underline{x}^{\tau+1})\frac{D(\underline{x}^{\tau+1},\underline{x}^\tau)}{\Delta}\left(\frac{\partial A(\underline{x}^\tau,\Delta)}{\partial\Delta}\cdot\frac{\Delta}{A(\underline{x}^\tau,\Delta)} - 1\right) + \frac{A(\underline{x}^\tau,\Delta)}{\Delta}\cdot\frac{\partial E(\underline{x}^\tau,\underline{x}^{\tau-1},\Delta)}{\partial\Delta}}{\frac{\partial C(\underline{x}^{\tau+1})}{\partial \underline{x}^{\tau+1}}D(\underline{x}^{\tau+1},\underline{x}^\tau) + C(\underline{x}^{\tau+1})\frac{\partial D(\underline{x}^{\tau+1},\underline{x}^\tau)}{\partial\underline{x}^{\tau+1}}}.$$

Second, to find $\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}$, we take derivatives of Equation (36) with respect to $\underline{x}^{\tau-1}$ to obtain:

$$\begin{aligned} 0 = A(\underline{x}^\tau, \Delta)&\frac{\partial E(\underline{x}^\tau, \underline{x}^{\tau-1}, \Delta)}{\partial \underline{x}^{\tau-1}} \\ &- \Delta\left(\frac{\partial C(\underline{x}^{\tau+1})}{\partial \underline{x}^{\tau+1}}D(\underline{x}^{\tau+1},\underline{x}^\tau) + C(\underline{x}^{\tau+1})\frac{\partial D(\underline{x}^{\tau+1},\underline{x}^\tau,\Delta)}{\partial\underline{x}^{\tau+1}}\right)\frac{\partial \underline{x}^{\tau+1}}{\partial \underline{x}^{\tau-1}}, \end{aligned}$$

and solve for $\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}$ to get

$$\frac{\partial \underline{x}^{\tau+1}}{\partial \underline{x}^{\tau-1}} = \frac{\frac{A(\underline{x}^\tau,\Delta)}{\Delta}\cdot\frac{\partial E(\underline{x}^\tau,\underline{x}^{\tau-1},\Delta)}{\partial \underline{x}^{\tau-1}}}{\frac{\partial C(\underline{x}^{\tau+1})}{\partial \underline{x}^{\tau+1}}D(\underline{x}^{\tau+1},\underline{x}^\tau) + C(\underline{x}^{\tau+1})\frac{\partial D(\underline{x}^{\tau+1},\underline{x}^\tau,\Delta)}{\partial\underline{x}^{\tau+1}}}.$$

Note that $\partial \underline{x}^{\tau+1}/\partial\Delta$ and $\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}$ are continuous and bounded. The uniform convergence result requires the absolute value of $\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}$ to be bounded below by 1. In our model, this condition translates into c being small. However, when c is large, that is, when the absolute value of $\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}$ is greater than 1, we can still have the convergence result by treating $\underline{x}^{\tau-1}$ as an implicit function of $\underline{x}^\tau$ and $\underline{x}^{\tau+1}$. In this backward method, the convergence requires the absolute value of $\partial \underline{x}^{\tau-1}/\partial \underline{x}^{\tau+1}$ to be bounded below by 1. In fact, because our model has well-defined both initial condition ($\underline{x}^1 = a$) and terminal condition ($\underline{x}^L = b$), it is easy to verify that the backward and forward methods lead to the same convergence result.

With initial condition $\underline{x}^0 = \underline{x}^1 = a$ and Equation (37), we then take limit of $\partial \underline{x}^{\tau+1}/\partial\Delta$ and $\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}$ as $\Delta \to 0$. By definition of C, D, and E, we need the following partial derivatives

$$\frac{\partial C(\underline{x}^{\tau+1})}{\partial \underline{x}^{\tau+1}} = -\frac{1}{(\underline{x}^{\tau+1})^2}, \quad \frac{\partial D(\underline{x}^{\tau+1},\underline{x}^\tau)}{\partial \underline{x}^{\tau+1}} = f(\underline{x}^{\tau+1}),$$

$$\frac{\partial E(\underline{x}^\tau, \underline{x}^{\tau-1}, \Delta)}{\partial \Delta} = -\frac{1}{c}F(\underline{x}^\tau), \quad \frac{\partial E(\underline{x}^\tau, \underline{x}^{\tau-1}, \Delta)}{\partial \underline{x}^{\tau-1}} = f(\underline{x}^{\tau-1}).$$

With $\underline{x}(t) = \lim_{j\to\infty}\hat{\underline{x}}(t;j)$ and the partial derivatives of A as derived in proof of

Lemma 8, we get

$$\lim_{\Delta\to 0}\frac{\partial \underline{x}^{\tau+1}}{\partial \Delta} = -\frac{(\underline{x}(t)r-1)F(\underline{x}(t))}{(c+\underline{x}(t))f(\underline{x}(t))}, \quad \lim_{\Delta\to 0}\frac{\partial \underline{x}^{\tau+1}}{\partial \underline{x}^{\tau-1}} = \frac{(\underline{x}(t)r-1)c}{c+\underline{x}(t)}.$$

Therefore, plugging the above limits into (37) and simplifying yields the same equation as (18)

$$\frac{d\underline{x}}{dt} = \lim_{\Delta\to 0}\frac{\partial \underline{x}^{\tau+1}/\partial\Delta}{1+\partial \underline{x}^{\tau+1}/\partial \underline{x}^{\tau-1}} = \frac{-\frac{(\underline{x}r-1)F(\underline{x})}{(c+\underline{x})f(\underline{x})}}{1+\frac{(\underline{x}r-1)c}{c+\underline{x}}} = -\frac{(\underline{x}r-1)F(\underline{x})}{c(1+\underline{x}r)f(\underline{x})}.$$

□

Lemma 10. *As $j\to\infty$, $\hat{\alpha}(\cdot,j)$ uniformly converges to function $\alpha:[0,\infty)\to[0,1]$ solving* (16)*.*

Proof. With Lemma 8 and Lemma 28, we can think x^τ as a predetermined sequence. Recall that $\hat{\alpha}(\cdot,j)$ is defined by level sequence $\{\alpha^{M(t,j)}(\Delta(j))\}$ in the any equilibrium that satisfies (8). Similar to the argument in proof of Lemma 8, consider $\tau = M(t;j)$, we can rewrite Equation (8) as the recursive formulation

$$\alpha^\tau = \alpha^{\tau+1}\frac{\delta x^\tau}{x^\tau-\Delta}.$$

In this case, the recursive formulation is backward. To apply Watson (2021), we need a well-defined terminal condition $\alpha^L=1$ and partial derivative with respect to $(-\Delta)$, which yields

$$\frac{\partial\alpha^\tau}{\partial(-\Delta)} = \alpha^{\tau+1}\frac{x^\tau\delta r(x^\tau-\Delta)-\delta x^\tau}{(x^\tau-\Delta)^2} = \alpha^{\tau+1}\frac{x^\tau\delta}{x^\tau-\Delta}\left(r-\frac{1}{x^\tau-\Delta}\right).$$

With $\alpha(t)=\lim_{j\to\infty}\hat{\alpha}(t;j)$, we have Equation (16)

$$\frac{d\alpha}{dt} = \lim_{\Delta\to 0}\frac{\partial\alpha^\tau}{\partial(-\Delta)} = \alpha\left(r-\frac{1}{x}\right).$$

□

From Lemma 8 and 9, we show that the step functions $\hat{\bar{x}}(\cdot;j)$ and $\underline{\hat{x}}(\cdot;j)$ converge uniformly to the same continuous function $x(\cdot)$ characterized by Equation (18). With the same function $x(\cdot)$, Lemma 10 shows that the step functions $\hat{\bar{\alpha}}(\cdot;j)$ and $\underline{\hat{\alpha}}(\cdot;j)$

converge to the same level function $\alpha(\cdot)$ characterized by Equation (16). The following lemma proves that the solution to this system of differential equations derived from Lemma 8–10 is unique.

Lemma 11. *Conditions (a)–(c) uniquely determine $\alpha(\cdot), x(\cdot)$, $\alpha(0)$, and T. Furthermore, T is finite.*

Proof. The existence of a solution is a direct result of solving the system of differential equations in the main text. Thus, it suffices to show the solution is unique by applying the Picard Theorem.

First, combine Equations (15) and (16), we get the uni-variate differential equation with initial condition (the same as Equation (18))

$$\frac{dx}{dt} = \frac{(1-rx)F(x)}{x(rc+1)f(x)}, \quad x(0) = b.$$

To apply the Picard Theorem, we check that the conditions are satisfied. Denote $G(x)$ as the right side of the differential equation above, and denote the domain of (t, x) as set $R_x = \{(t, x) : 0 \le t \le T, a \le x \le b\}$. It is straightforward to check that function G is continuous and bounded for all $(t, x) \in R_x$, $|G(x)| \le K_1$, and dG/dx is continuous and bounded in R_x, $|dG/dx| \le K_2$, where K_1, K_2 are positive constant real numbers. Denote $\tau_x = \min\{T, (b-a)/K_1\}$. Apply the Picard Theorem, the ODE has unique solution for $t \le \tau_x$. For $t \in [\tau_x, T]$, we can apply the theorem recursively and the uniqueness of the solution to ODE can be proved for the entire interval $[0, T]$. Note that T is finite because in the solution $T = I_x(a) - I_x(b)$, where $I_x(x)$ is well defined integration bounded above by assumption $x \ge a > 1/r$.

In particular, denote $\tau_x^0 = 0$. First notice that if $(b-a)/K_1 < T$, so, by the Picard Theorem, the uniqueness is guaranteed for $t \in [\tau_x^0, \tau_x^1]$, where $\tau_x^1 = (b-a)/K_1$. Second, for $t \in [\tau_x^1, T]$, the initial value is determined by $x(\tau_x^1)$, so the uniqueness is guaranteed by the Picard Theorem for $t \in [\tau_x^1, \tau_x^2]$, where $\tau_x^2 = \min\{T, \tau_x^1 + (x(\tau_x^1) - a)/K_1\}$. Applying the Picard Theorem in each round guarantees the uniqueness of the solution for

$$t \in \cup_{n=1}^{\infty}[\tau_x^{n-1}, \tau_x^n], \text{ where } \tau_x^n = \min\left\{T, \tau_x^{n-1} + \frac{x(\tau_x^{n-1}) - a}{K_1}\right\}.$$

To show the uniqueness of solution for the entire interval $[0, T]$, it suffices to show that $\lim_{n\to\infty} \tau_x^{n-1} + (x(\tau_x^{n-1}) - a)/K_1 \ge T$. Note that the sequence $\{\tau_x^n\}$ is weakly increase and bounded above with limit as $\lim_{n\to\infty} \tau_x^{n-1} = T$. Further because we set

$x(T) = a$, and because dx/dt is negative and $|dx/dt|$ is bounded by a finite number K_1, then when τ_x^{n-1} is close to T, it is easy to verify that this condition holds.

Second, for the differential Equation (16), or Equation (20)

$$\frac{d\alpha}{dt} = \left(r - \frac{1}{x(t)}\right)\alpha \quad \alpha(T) = 1,$$

where $x(t)$ is the unique solution solved from previous analysis. Now we check regularity conditions and apply the Picard Theorem to this uni-variate differential equation and solve for α. Denote $H(t,\alpha)$ as the right side of the differential equation above, and denote the domain of (t,α) as the set $R_\alpha = \{(t,\alpha) : 0 \le t \le T, 0 < \alpha \le 1\}$. It is straightforward to check that function H is continuous and bounded for all $(t,\alpha) \in R_\alpha$, $|H(t,\alpha)| \le K_3$, and $dH/d\alpha$ is continuous and bounded in R_α, $|dH/d\alpha| \le K_4$, where K_3, K_4 are positive constant real numbers. Denote $\tau_\alpha = \min\{T, 1/K_3\}$. Applying the Picard Theorem, the ODE has unique solution for $t \le \tau_\alpha$. The same recursive argument applies and the uniqueness of the solution to ODE can be proved for the entire interval $[0,T]$. $\square$

Appendix D Proof of Proposition 1

We first derive comparative statics of T. From Equation (25), we use the fact $\ln(z) \le z - 1$ and get

$$\begin{aligned}
\frac{dT}{dq} &= \frac{(b-a)(1+r)}{(1-q+r(bq-a))^2}\left(\ln\frac{1-ra}{q(1-rb)} - \frac{1-q+r(bq-a)}{q}\frac{bq-a}{b-a}\right) \\
&\le \frac{(b-a)(1+r)}{(1-q+r(bq-a))^2}\left(\frac{1-ra}{q(1-rb)} - 1 - \frac{1-q+r(bq-a)}{q}\frac{bq-a}{b-a}\right) \\
&= \frac{(b-a)(1+r)}{1-q+r(bq-a)}\left(\frac{1}{1-rb} - \frac{bq-a}{b-a}\right)\frac{1}{q} = \frac{b(1+r)}{q(1-rb)} < 0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{dT}{db} &= -\frac{(1-q)(1+r)}{1-q+r(bq-a)}\left(\frac{rq\left(-\frac{1}{r}\ln\frac{1-ra}{1-rb} + \frac{a-bq}{1-q}\ln q\right)}{1-q+r(bq-a)} + \left(\frac{1}{1-rb} + \frac{q}{1-q}\ln q\right)\right) \\
&= \frac{q(1-q)(1+r)}{(1-q+r(bq-a))^2}\left(\ln\frac{1-ra}{q(1-rb)} - \frac{1-ra}{q(1-rb)} + 1\right) \le 0.
\end{aligned}$$

Second. from Equation (25), comparative statics of $\alpha(0)$ is

$$\frac{d\alpha(0)}{dq} = (r+1)q^r > 0, \quad \frac{d\alpha(0)}{db} = 0.$$

Third, for the comparative statics of the slope of x for fixed $\chi \in [a,b]$ at time $\Gamma(\chi)$, we use Equation (21)

$$\frac{d\Gamma}{d\chi} = \frac{(1+r)\chi}{1-r\chi} \cdot \frac{1-q}{(b-a)q + (\chi-a)(1-q)} \equiv g(b,q;\chi),$$

and take partial derivatives of g, we get

$$\begin{aligned}
\frac{dg(b,q;\chi)}{dq} &= \frac{(1+r)\chi}{r\chi-1} \cdot \frac{b-a}{((b-a)q+(\chi-a)(1-q))^2} > 0, \\
\frac{dg(b,q;\chi)}{db} &= \frac{(1+r)\chi}{r\chi-1} \cdot \frac{q(1-q)}{((b-a)q+(\chi-a)(1-q))^2} > 0.
\end{aligned}$$

Last, we consider the comparative statics of the slope of $\ln\alpha$ for fixed $\chi \in [a,b]$ at time $\Gamma(\chi)$. Similarly, with Equation (24), we have

$$\frac{d\ln\alpha}{d\chi} = -\frac{(1-q)(1+r)}{(b-a)q+(\chi-a)(1-q)} \equiv h(b,q;\chi).$$

Therefore,

$$\begin{aligned}
\frac{dh(b,q;\chi)}{dq} &= \frac{(1+r)(b-a)}{((b-a)q+(\chi-a)(1-q))^2} > 0, \\
\frac{dh(b,q;\chi)}{db} &= \frac{q(1-q)(1+r)}{((b-a)q+(\chi-a)(1-q))^2} > 0.
\end{aligned}$$

References

Andreoni, James, Michael A. Kuhn, and Larry Samuelson. 2019. “Building Rational Cooperation on Their Own: Learning to Start Small.” *Journal of Public Economic Theory* 21 (5): 812–825.

Atakan, Alp, Levent Koçkesen, and Elif Kubilay. 2020. “Starting Small to Communicate.” *Games and Economic Behavior* 121:265–296.

Ausubel, Lawrence M., and Raymond J. Deneckere. 1989. "Reputation in Bargaining and Durable Goods Monopoly." *Econometrica* 57 (3): 511–531.

Bernheim, B Douglas, and Debraj Ray. 1989. "Collective dynamic consistency in repeated games." *Games and Economic Behavior* 1 (4): 295–326.

Blonski, Matthias, and Daniel André Probst. 2001. "The Emergence of Trust."

Bowen, T. Renee, George Georgiadis, and Nicolas S. Lambert. 2019. "Collective Choice in Dynamic Public Good Provision." *American Economic Journal: Microeconomics* 11, no. 1 (February): 243–98.

Chassang, Sylvain. 2010. "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts." *American Economic Review* 100, no. 1 (March): 448–65.

Coase, R. H. 1972. "Durability and Monopoly." *The Journal of Law and Economics* 15 (1): 143–149.

Coddington, Earl A., and Norman Levinson. 1984. *Theory of Ordinary Differential Equations.* Malabar, Fla R.E. Krieger.

Farrell, Joseph, and Eric Maskin. 1989. "Renegotiation in repeated games." *Games and economic behavior* 1 (4): 327–360.

Ghosh, Parikshit, and Debraj Ray. 1996. "Cooperation in Community Interaction Without Information Flows." *The Review of Economic Studies* 63, no. 3 (July): 491–519.

Gompers, Paul A. 1995. "Optimal Investment, Monitoring, and the Staging of Venture Capital." *The Journal of Finance* 50 (5): 1461–1489.

Gul, Faruk, Hugo Sonnenschein, and Robert Wilson. 1986. "Foundations of Dynamic Monopoly and the Coase Conjecture." *Journal of Economic Theory* 39 (1): 155–190.

Halac, Marina. 2012. "Relational Contracts and the Value of Relationships." *American Economic Review* 102, no. 2 (April): 750–79.

Halac, Marina. 2013. “Relationship Building: Conflict and Project Choice over Time.” *The Journal of Law, Economics, and Organization* 30, no. 4 (August): 683–708.

Horstmann, Ignatius, and James Markusen. 1996. “Exploring New Markets: Direct Investment, Contractual Relations and the Multinational Enterprise.” *International Economic Review* 37 (February): 1–19.

———. 2018. “Learning to Sell in New Markets: A Preliminary Analysis of Market Entry by a Multinational Firm.” *Review of International Economics* 26 (5): 1040–1052.

Kartal, Melis, Wieland Müller, and James Tremewan. 2019. “Building Trust: The Costs and Benefits of Gradualism.” *Available at SSRN 3324993.*

Kranton, Rachel E. 1996. “The Formation of Cooperative Relationships.” *The Journal of Law, Economics, and Organization* 12, no. 1 (April): 214–233.

Malcomson, James M. 2016. “Relational Incentive Contracts With Persistent Private Information.” *Econometrica* 84 (1): 317–346.

———. 2020. “Grouping Agents with Persistent Types.” *unpublished draft.*

Miller, David A., Trond E. Olsen, and Joel Watson. 2020. “Relational Contracting, Negotiation, and External Enforcement.” *American Economic Review* 110, no. 7 (July): 2153–97.

Miller, David A., and Joel Watson. 2013. “A Theory of Disagreement in Repeated Games With Bargaining.” *Econometrica* 81 (6): 2303–2350.

Rauch, James, and Joel Watson. 2003. “Starting Small in an Unfamiliar Environment.” *International Journal of Industrial Organization* 21 (7): 1021–1042.

Rob, Rafael, and Arthur Fishman. 2005. “Is Bigger Better? Customer Base Expansion through Word-of-Mouth Reputation.” *Journal of Political Economy* 113 (5): 1146–1162.

Rubinstein, Ariel. 1980. “Strong perfect equilibrium in supergames.” *International Journal of Game Theory* 9 (1): 1–12.

Sobel, Joel. 1985. "A Theory of Credibility." *The Review of Economic Studies* 52, no. 4 (October): 557–573.

Tian, Xuan. 2011. "The causes and consequences of venture capital stage financing." *Journal of Financial Economics* 101 (1): 132–159.

Watson, Joel. 1999. "Starting Small and Renegotiation." *Journal of Economic Theory* 85 (1): 52–90.

———. 2002. "Starting Small and Commitment." *Games and Economic Behavior* 38 (1): 176–199.

———. 2017. "A General, Practicable Definition of Perfect Bayesian Equilibrium." *unpublished draft.*

———. 2021. "A Note on the Relation Between Models in Discrete and Continuous Time." *unpublished draft.*

Ye, Maoliang, Jie Zheng, Plamen Nikolov, and Sam Asher. 2020. "One Step at a Time: Does Gradualism Build Coordination?" *Management Science* 66 (1): 113–129.