Title

Essays on Reputation Dynamics in Agency Relationships, Venture Capital Markets, and Lending Relationships

Author
Hua, Xiameng

Publication Date
2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Reputation Dynamics in Agency Relationships, Venture Capital Markets, and Lending Relationships

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Xiameng Hua

Committee in charge:

Professor Joel Watson, Chair
Professor Snehal Banerjee
Professor Renee Bowen
Professor Songzi Du
Professor Richard Townsend

2024

The Dissertation of Xiameng Hua is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

This dissertation is dedicated to the many people who have supported and inspired me throughout this journey.

To my parents, for their unwavering love, support, and belief in me. Your guidance has been my foundation and strength.

To my husband, Aaron, for your endless patience, encouragement, and sacrifices. Your love has been my guiding light through every step of this journey.

To my advisor, Prof. Joel Watson, whose mentorship and guidance have been invaluable. Your wisdom and support have profoundly shaped my academic career.

To my mentors, colleagues, and friends from University of California San Diego, Workshop on Entrepreneurial Finance and Innovation Fellows Program, Lehigh University, and Southwestern University of Finance and Economics. Your camaraderie, encouragement, and insights have made this journey both possible and enjoyable.

Thank you all for your unwavering support and belief in me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Joel Watson for his support as the chair of my committee. Through multiple drafts and many long nights, his guidance has proved to be invaluable.

Chapter 1 is coauthored with Joel Watson and has been published in Journal of Economic Theory 204 (September 2022). The dissertation author was a coauthor of this chapter.

Chapter 2, in part is currently being prepared for submission for publication of the material. The dissertation author was the author of this material.

VITA

2014	Bachelor of Science, Southwestern University of Finance and Economics
2016	Master of Science, Lehigh University
2024	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

"Starting Small in Project Choice: a Discrete-Time Setting with a Continuum of Types" with Joel Watson, Journal of Economic Theory, Volume 204, September 2022, 105490.

FIELDS OF STUDY

Major Field: Economic Theory

Studies in Relational Contract
Professor Joel Watson

Studies in Finance
Professors Snehal Banerjee and Richard Townsend

ABSTRACT OF THE DISSERTATION

Essays on Reputation Dynamics in Agency Relationships, Venture Capital Markets, and Lending Relationships

by

Xiameng Hua

Doctor of Philosophy in Economics

University of California San Diego, 2024

Professor Joel Watson, Chair

Chapter 1 contributes to the understanding of long-term relationships characterized by variable stakes and incomplete information through an analysis of a discrete-time trust game between a principal and agent with a continuum of types. The principal selects the project level in each period, and the agent decides whether to cooperate or betray, with payoffs scaling with the project level. The agent's benefit of betraying is privately known. A novel internal consistency condition for renegotiation is introduced, allowing for a comprehensive examination of equilibrium selection, with the principal assumed to have full power to alter it. The main result demonstrates convergence of perfect Bayesian equilibria as the period length approaches zero,

with a closed-form solution provided. Cooperation remains viable across all type distributions, with the relationship starting small and gradually reaching maximum level.

Chapter 2 investigates the reputation-spillover phenomenon of venture capitalists (VCs) creating value for startups by attracting high-quality labor. I analyze a dynamic matching model between long-lived VCs with persistent but unknown abilities and short-lived workers with varying productivity, and characterize stable, positive assortative matching within each period, with a worker's wage potentially decreasing with productivity. Production technology determines the steady-state distribution by influencing the speed at which VCs build reputations.

Chapter 3 presents analysis of the borrowing-lending game where the borrower has private information about riskiness, and the bank starts small to learn about it. Empirical examination of model predictions using loan-level datasets from the U.S. Small Business Administration corroborates several findings. Repeated borrowers exhibit lower default rates, and default rates increase with the distance between borrower and lender. Repeated borrowers face lower interest rates initially but higher rates when refinancing. Economic downturns stall existing loans at a low level, and first-time borrowers secure larger loan amounts than repeat borrowers.

Chapter 1

Starting Small in Project Choice

1.1 Introduction

Long-term relationships in business and greater society often begin with 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. Also, it is useful to work with models that generate narrow equilibrium predictions.

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.[2] 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 most 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.

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).

2. The convergence result pertains to nontrivial equilibria, in which the principal sets a positive level in at least one period. For some parameter values, there may also exist a trivial no-trust equilibrium, but we argue that it would be ruled out by a weak form of external consistency.

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 alteration-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 (Watson 2021).

The related literature on starting small in relationships includes both seminal theoretical contributions and experimental evidence. Theoretical origins reside in Sobel (1985), Ghosh and Ray (1996), and Watson (1999, 2002). Sobel (1985) focuses on a "sender-receiver" model in which the level is determined by an exogenous random draw in each period; this paper also describes a "loan model" in which one player chooses the level and the other, who could be a friend or enemy, chooses whether to invest or default. The equilibrium of the loan model entails gradual increase in the loan level over time. Our model has the same form of stage game, but the payoffs are different for the player we call the good type of agent, and in our model incentives relate to an infinite horizon. Ghosh and Ray (1996) examines a setting in which players in a community randomly match to form long-term relationships. Players can exit their relationships at any time and then rematch, and newly matched players receive no information about the past behavior of their partners. A fraction of the population is myopic. Players are motivated to weed out myopic types by reducing the level of cooperation in the first period of new relationships, and this serves as a punishment for non-myopic players who might otherwise cheat and rematch without consequences.

On the empirical side, 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 finds 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 and empirically that it is sometimes optimal to start small.[3] 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.[4]

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 if 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

3. 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.

4. Other articles that model some manner of starting small in relationships and/or trust building include Kranton (1996), Blonski and Probst (2001), Rob and Fishman (2005), and Chassang (2010).

requires a 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 1.2 we formally describe the model and equilibrium concept, and we analyze the agent's incentive conditions. The renegotiation condition is defined and analyzed in Section 1.3, where we report our main result on the limit of equilibria. Section 1.4 presents the case of uniformly distributed bad types. Section 1.5 provides additional technical notes, including on off-equilibrium-path alterations and how our results extend to the "no-gap case" of agent types assumed away earlier, and discusses additional connections with literature. Section 1.6 offers concluding comments. The appendices contain details of the analysis and proofs.

1.2 Model

We examine a model of a relationship between a principal and agent in discrete time and with one-sided incomplete information. In this section we first describe the complete-information version of the game, followed by the incomplete-information version. We then establish notation, review the equilibrium conditions, and provide a partial characterization of "trusting equilibria" along with examples.

1.2.1 Trust game with complete information

The complete-information version of the model 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 $k \in \{1, 2, \ldots\}$. We assume the players have a common discount factor $\delta \equiv e^{-r\Delta}$, where $r \in (0,1)$ is the discount rate and $\Delta > 0$ 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 1.1, 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 $c > 0$. Players seek to maximize the discounted sum of their period payoffs.

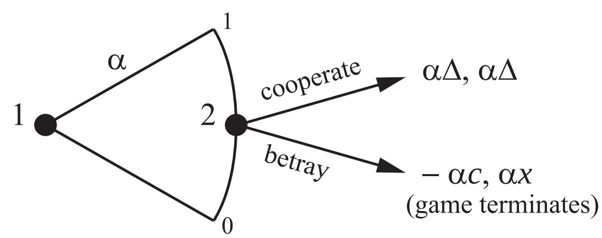


Figure 1.1. Stage game

It is easy to verify that cooperation can be sustained if and 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. Furthermore, if $0 < x < \Delta/(1-\delta)$, then there are many other equilibria. 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 $k+1$.[5]

To summarize, there are a lot of equilibria featuring trust and cooperation 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 contrary, if $x > \Delta/(1-\delta)$ then there is no equilibrium in which cooperation occurs at a positive level in any period.

5. Player 1 can be deterred from deviating by specifying that, 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).

1.2.2 Trust game with incomplete information

We are interested in the trust game with incomplete information regarding the payoff parameter x. Specifically, suppose that before the relationship begins, Nature chooses x according to a given probability distribution F that is common knowledge, with support denoted by $X \subset \mathbb{R}$. Player 2 privately observes x, which we therefore refer to as *player 2's type*. Let us label every $x \leq \Delta/(1-\delta)$ a *good type* and every $x > \Delta/(1-\delta)$ a *bad type*. We generally express F as a cumulative probability function, so that $F(x')$ denotes the probability that $x \leq x'$.

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.

We assume that the types are bounded and there is a gap between the sets of good types and bad types.[6] We will later examine sequences of games for Δ converging to zero, and we want the type labels to hold for every Δ close to zero. Because $\Delta/(1-\delta)$ decreases and converges to $1/r$ as $\Delta \to 0^+$, we therefore assume that the good types are below $1/r$ and the lowest bad type is strictly above $1/r$. Additional technical assumptions are included in the following assumption.

Assumption 1. *Distribution function F is continuous. There are numbers $\hat{\Delta}$, a, and b satisfying*

6. This is analogous to the "gap" case of the durable-good-monopoly problem (Coase (1972), Gul, Sonnenschein, and Wilson (1986)).

$0 < \hat{\Delta} < 1/r < a < b$ such that $F(\hat{\Delta}) = 0$; $F(a) = F(1/r) > 0$; $F(b) = 1$; and restricted to subdomain $[a,b]$, F is twice continuously differentiable with a strictly positive density function f. Finally, $\Delta \leq \min\{\hat{\Delta}, \overline{\Delta}\}$, where

$$\overline{\Delta} \equiv a \left(\frac{1}{a \min_{x \in [a,b]} f(x)} + 1 \right)^{-1}.$$

Note that the set of bad types is the interval $[a,b]$. We define the derivative of F at endpoint a as its right derivative, and at endpoint b as its left derivative, so that f is well-defined and can have the assumed properties when restricted to $[a,b]$. We define the derivative of f similarly. Because f is strictly positive and continuous on subdomain $[a,b]$, it reaches a minimum that is strictly positive, and so $\overline{\Delta}$ is well-defined. The assumption that the set of good types is bounded away from 0 ensures the existence of a class of simple equilibria but is not needed for existence or used in our characterization theorem. From here, "game" and "trust game" refer to our incomplete-information, discrete-time game with parameters r, Δ, X, and F just now described.

1.2.3 Strategies and equilibrium conditions

We analyze the game using the weak Perfect Bayesian Equilibrium (PBE) solution concept. In this subsection, we define and provide notation for histories, strategies, and beliefs. We then describe the equilibrium conditions and, noting the plethora of equilibria, motivate the refinement developed in the next section.

For any $k \in \{1,2,\ldots\}$, 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, 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.[7] 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 $Q : H \to \mathscr{P}(X)$, where $\mathscr{P}(X)$ 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 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

7. Restricting attention to pure strategies for player 2 is without loss in our analysis. Because we have a continuum of types, indifference conditions would occur for only a subset of measure zero. Accounting for randomization by player 1 would complicate the statement of the alteration-proofness conditions, we think with no effect on the main results, as discussed in Section 1.5.

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 $s_2(h\alpha, Q(h))$.

We next review the notion of sequential rationality, stated here in terms of single deviations, and the equilibrium definition. 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*

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

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. *A pure-strategy weak* **Perfect Bayesian Equilibrium (PBE)** *is a strategy profile (s_1, s_2) and beliefs Q such that s_1 and s_2 are sequentially rational and Q obeys Bayes' Rule for all histories reached with positive probability given F, s_1, and s_2.*

Note that in periods in which a positive mass of types is supposed to cooperate, player 1 cannot detect a deviation by a type that was meant to betray, and so standard Bayes updating applies. Weak PBE does not constrain belief updating following any "public deviation," where either player 1 deviated or player 2 cooperated in a contingency in which all types were supposed to betray, because the conditional probability formula does not apply in such a contingency. We could impose stronger consistency conditions, such as Watson (2017) defines, but it would be of no consequence because in the game studied here, we can modify any weak PBE to satisfy strong consistency conditions off the equilibrium path where needed.

For any PBE, let $\{\alpha^k\}_{k=1}^K$ denote the sequence of levels chosen by player 1 *on the equilibrium path*, where there is no public deviation. In this expression, K denotes the last period that occurs in equilibrium; K is finite if all types of player 2 betray in bounded time, and $K = \infty$ if for every period k, a positive mass of types cooperate through period k on the equilibrium path. We will show shortly that $K = \infty$ for any PBE, but for now we must allow for the possibility of K finite.[8]

Let us characterize player 2's incentive conditions on the equilibrium path. Type x optimally betrays in some period in the set

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

Note that β is defined relative to a given PBE and it constrains player 2 to betray at or before the equilibrium K. In the case of $K = \infty$, $\infty \in \beta(x)$ is allowed and means that type x optimally cooperates forever.[9] For each $k \in \{1,2,\ldots,K\}$, 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 equilibrium strategy must have the property that, for $x \in X$ and for the lowest k for which $s_2(h^k, x) = 0$, it is the case that $k \in \beta(x)$.

Regarding player 1's incentives, observe that the following specification of beliefs and behavior for off-path continuations achieves a continuation value of zero for both players. After any deviation by player 1, all types of player 2 would immediately betray, ending the game. If

8. Equilibrium strategies must specify behavior after all histories, including ones in which a period $k > K$ is reached following a public deviation. We could describe, for instance, the infinite sequence of levels that would result from player 1 never deviating and player 2 always cooperating; this sequence would be on the equilibrium path through period K and off the equilibrium path thereafter. Such a sequence will not be needed for our analysis.

9. The set $\beta(x)$ is nonempty even if $K = \infty$ due to discounting. Also, if $\beta(x)$ contains an infinite number of periods then it must also contain ∞. If K is finite, then it would be feasible for player 2 to cooperate through period K and perhaps betray later, and we would need to check such a deviation to determine whether player 2 best responds.

player 2 instead cooperates, which constitutes a further public deviation, then player 1's updated belief would assign probability 1 to a bad type. Then in every period thereafter, regardless of the interim history, player 1 is supposed select $\alpha = 0$ and all types of player 2 are supposed to betray. These continuation strategies are sequentially rational.[10] Because player 1 can guarantee a payoff of zero by choosing $\alpha = 0$ forever, player 1's incentive conditions on the equilibrium path amount to having a nonnegative continuation value.

1.2.4 Trusting PBE

We are particularly interested in PBE in which, on the equilibrium path, the level is strictly positive in at least one period.

Definition 3. *A perfect Bayesian equilibrium in the trust game is called a* **trusting PBE** *if* $\alpha^k > 0$ *for some* $k \in \{1, 2, \ldots, K\}$.

A trusting equilibrium exhibits some degree of cooperation at positive levels of trust, for otherwise player 1 would strictly prefer to set the level to zero in every period. We first characterize trusting PBE in terms of the relation between the strategy of player 2 and the sequence of levels on the equilibrium path.

Lemma 1. *Every trusting PBE has the following properties:* $K = \infty$. *There is an integer* L *and a weakly decreasing sequence* $\{x^k\}_{k=0}^{\infty}$ *such that (i) for every* $x \in X$, *if player 2 of type* x *betrays on the equilibrium path then this betrayal occurs in a period* k *that satisfies* $x \in [x^k, x^{k-1}]$, *and (ii) type* $x = a$ *betrays in period* L.

Proof of Lemma 1. Consider any trusting PBE and let $\{\alpha^k\}_{k=1}^{K}$ 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:

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

10. In fact, they are sequentially rational regardless of player 1's beliefs about player 2's type, but the particular belief specified here will be helpful later for an extension of the model.

We first show that for any two types x' and x'' such that $x' > x''$, on the equilibrium path type x' betrays in a period weakly earlier than does type x''.

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

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

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

$$\left(\alpha^{k''} - \delta^{k'-k''}\alpha^{k'}\right)x'' \geq \delta^{-k''}\sum_{\tau=k''}^{k'-1}\delta^{\tau}\alpha^{\tau}\Delta \geq \left(\alpha^{k''} - \delta^{k'-k''}\alpha^{k'}\right)x'.$$

Because the level is strictly positive in at least one period on the equilibrium path, player 2's incentive condition implies that $\alpha^{k''} > 0$, which further implies that the middle term in the above expression is strictly positive. Using the left inequality and $x'' > 0$, we obtain $\alpha^{k''} - \delta^{k'-k''} > 0$. Combining the inequalities and dividing by $\alpha^{k''} - \delta^{k'-k''}$, we get $x'' \geq x'$, contradicting our presumption that $x' > x''$.

Next, we show that $\beta(a)$ is bounded above. Define $\overline{\alpha} = \sup\{\alpha^1, \alpha^2, \ldots\}$. Then for any $\varepsilon > 0$, there exists a period κ such that $\alpha^{\kappa} \geq \overline{\alpha} - \varepsilon$. 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 $a(\overline{\alpha} - \varepsilon)$. If $\beta(a)$ were unbounded then $K = \infty$ and $\infty \in \beta(a)$. By cooperating forever, this type's 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 $a > \Delta/(1-\delta)$, we know that $a(\overline{\alpha} - \varepsilon) > \overline{\alpha}\Delta/(1-\delta)$ for sufficiently small values of ε, which contradicts that it is rational for type a to cooperate forever. We conclude that the lowest bad type betrays in some period L on the equilibrium path.

Finally, we show that $K = \infty$. Assume otherwise, meaning that on the equilibrium path all types of player 2 betray at or before period K and some types wait until K to do so. It must be that $\alpha^K > 0$, for otherwise the types that are supposed to betray in period K would strictly prefer to betray in an earlier period where the level is strictly positive (a time which must exist in a trusting equilibrium). But then in period K player 1's continuation value must be strictly negative because he expects player 2 to betray with probability one. This contradicts player 1's rationality because he would strictly gain by selecting $\alpha^k = 0$ for all $k \geq K$. □

Lemma 1 does not pin down the periods of betrayal for the countable number of types of player 2 that may be indifferent between betraying in one period and the next. Because this is a set of measure zero, equilibria that differ in this regard are essentially equivalent.

To summarize the analysis so far, every trusting PBE has an infinite equilibrium path and is partially characterized by its sequence of levels $\{\alpha^k\}_{k=1}^{\infty}$ and its sequence of cutoff types $\{x^k\}_{k=0}^{\infty}$. On the equilibrium path, for any integer k, all types below x^{k-1} cooperate through period $k-1$ and then types in the subinterval (x^k, x^{k-1}), and possibly one or both endpoints, will betray in period k at level α^k. The monotonicity of betrayal dates established by Lemma 1 applies to all types, good types included. All bad types betray in or before period L. Note that the lemma does not indicate whether any good types betray in equilibrium.

As the analysis continues, we will need to keep track of continuation values. Given any trusting 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.

We conclude this subsection with an existence result for trusting PBE, which is a corollary of our main existence result in the next section.

Theorem 1. *Under Assumption 1, a trusting PBE exists in the trust game.*

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 good types 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 trusting PBE for a sufficiently small mass of good types.

1.2.5 Intuition and Illustrations

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

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

In some equilibria, type x^k is indifferent between cooperating and betraying in period k, so that the above inequality holds as an equation. 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 $v_2^{k+1}(x^k)$, we obtain:

$$\alpha^k x^k = \Delta\alpha^k + \delta\alpha^{k+1}x^k. \tag{1.1}$$

The refinement developed in the next section will be shown to imply that Equation (1.1) holds in every period k for which $x^k > a$; 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 trusting equilibria, which differ in terms of when bad types betray, how the level changes over time, and players 1's payoff. Figures 1.2–1.4 depict three equilibria that we constructed 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 has a mass of 0.3836 of good types and specifies a uniform distribution of 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. In each of these equilibria, on the equilibrium path all good types cooperate in every period.

For the equilibrium in Figure 1.2, Equation (1.1) does not hold in some periods. Every type of 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 1.3 and 1.4 illustrate equilibria for which Equation (1.1) holds for all periods. In the equilibrium shown in Figure 1.3, all bad types betray in period 1 at the beginning of the game, so $L = 1$. In the equilibrium shown in Figure 1.4, no bad type betrays until the level reaches 1 in period $L = 30$.[11]

It turns out that none of the equilibria pictured satisfy the renegotiation-proofness condition developed in the next section. In the first equilibrium, there are periods in which the level can be increased without affecting player 2's incentives, and this increases player 1's payoff.[12] In the second equilibrium, after observing cooperation in the first period, player 1 would be sure that player 2 is a 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.

11. We can show that if $c > b$ then player 1's favorite equilibrium is as pictured in Figure 1.3, 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 1.4, where all bad types of player 2 wait until period $L - 1$ to betray. These findings match with what Watson (2002) demonstrates in a continuous-time model with a single bad type.

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

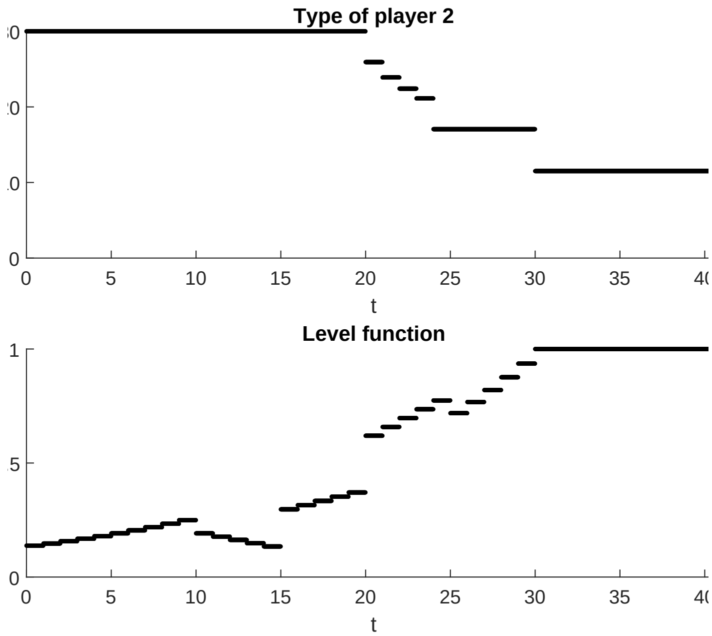


Figure 1.2. First equilibrium illustration

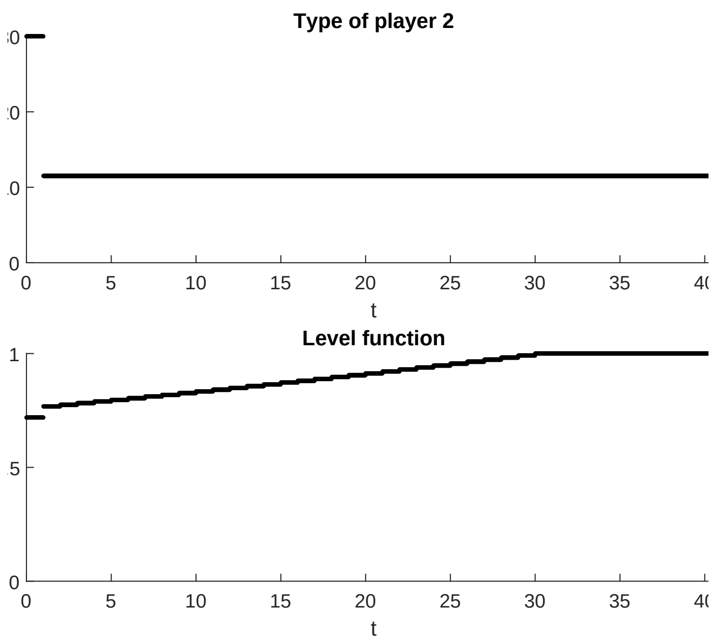


Figure 1.3. Second equilibrium illustration

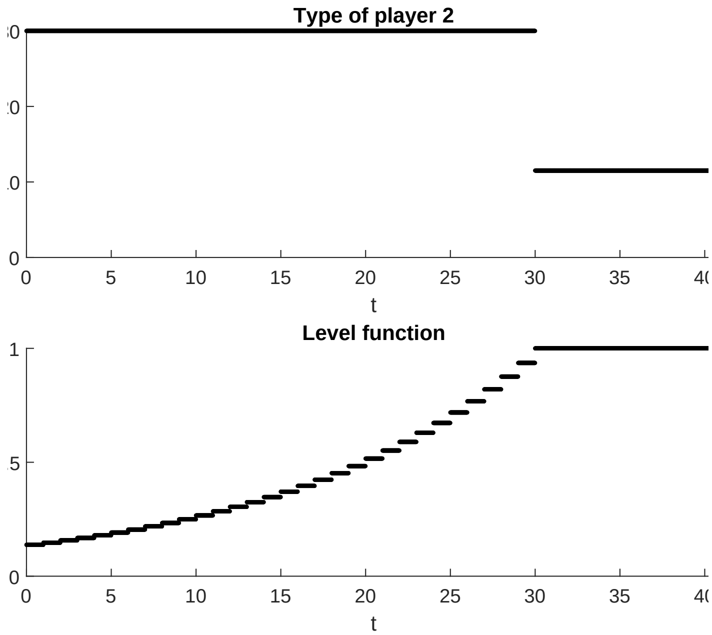


Figure 1.4. Third equilibrium illustration

1.3 Alteration Proofness

In this section, we define and analyze a minimal notion of renegotiation 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.[13] The concept imposes a form of internal consistency: In a given period k the equilibrium continuation may be altered in any way, so long as in period $k+1$ it returns to a path consistent with the current equilibrium. The new path from period $k+1$ can pick up the current equilibrium as though in any other period $k' \in \{k, k+1, k+2, \ldots, K\}$.

For instance, if $k' = k$ then the players are stalling, essentially postponing the equilibrium path by one period. Any $k' > k+1$ amounts to "jumping ahead" by $k' - k - 1$ periods, and $k' = k+1$ means that the alteration affects only the current period k. Restrictions are inherent in

13. 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.

how an equilibrium can be altered in this way. In particular, to pick up on the current equilibrium as though in period k', player 1's belief must be exactly as it would be at the start of period k', so the alteration must specify for the current period k behavior that would lead to such a belief at the end of this period.[14]

To focus on what drives our main characterization result and to avoid complicated notation, we shall define alteration-proofness in reference only to continuations of the game *on the equilibrium path*. It is appropriate to also apply alteration-proofness to continuations of the game following public deviations, so that the conditions are imposed both on and off the equilibrium path. In fact, our results extend to this wider application of alteration-proofness, as explained in Section 1.5 and in the Appendix. The wider imposition of alteration-proofness turns out to not further constrain equilibrium outcomes.

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.

1.3.1 The alteration-proofness condition

Consider any trusting PBE, partly characterized by $\{\alpha^k\}_{k=1}^{\infty}$ and $\{x^k\}_{k=0}^{\infty}$, and suppose that period k is reached on the equilibrium path. We imagine that player 1 may 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 $k+1$ be as though in the original equilibrium from period $k+1+m$, where $m \in \{-1, 0, 1, \ldots, K-k-1\}$ denotes by how many periods the altered equilibrium skips

14. We could allow alterations with $k' < k$, but it would not change the implications of our theory. In this case, Lemma 1 implies $x^{k'} \geq x^{k-1}$. If this inequality is strict then the alteration is not feasible; otherwise, player 1 would strictly prefer the alteration only if an alteration with $k' \geq k$ is strictly preferred.

ahead in relation to the original equilibrium from period $k+1$. In the altered equilibrium, play from period k will be described by level and cutoff sequences $\{\tilde{\alpha}^\tau\}_{\tau=k}^\infty$ and $\{\tilde{x}^\tau\}_{\tau=k-1}^\infty$ where $\tilde{\alpha}^\tau = \alpha^{\tau+m}$ for all $\tau > k$ and $\tilde{x}^\tau = x^{\tau+m}$ for all $\tau \geq k$. Note that the $\tilde{\alpha}^k$ is not nailed down here and so it and m define the alteration.

The level $\tilde{\alpha}^k$ 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 $k+1$, player 1's belief is exactly what it would have been at the beginning of period $k+1+m$ in the original equilibrium (on the equilibrium path). That is, the continuation game in the altered equilibrium from period $k+1$ must be identical to the continuation game in the original equilibrium from period $k+1+m$. In the latter continuation, player 1's belief about player 2's type is exactly the updated version of F conditional on $x \leq x^{k+m}$ because only these types remain in the game at this point.

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

$$\tilde{\alpha}^k x \leq \Delta\tilde{\alpha}^k + \delta v_2^{k+1+m}(x) \quad \text{for all } x \in X \text{ such that } x \leq x^{k+m}, \tag{1.2}$$

and the second is

$$\tilde{\alpha}^k x \geq \Delta\tilde{\alpha}^k + \delta v_2^{k+1+m}(x) \quad \text{for all } x \in X \text{ such that } x \in (x^{k+m}, x^{k-1}], \tag{1.3}$$

where v_2^{k+1+m} refers to player 2's continuation value in the original equilibrium.

Note that the constraints can be vacuous depending on how values x^{k-1} and x^{k+m} relate to X. For instance, if no types are scheduled to betray between periods k and $k+m$, which would be the case if $x^{k+m} = x^{k-1}$ or if these values are both between $1/r$ and a, then the second constraint is trivially satisfied.

Definition 4. *Take as given a trusting PBE, with player 2's equilibrium continuation values denoted by* $\{v_2^k(\cdot)\}_{k=1}^{\infty}$*. For any period* k*, integer* $m \in \{-1, 0, 1, \ldots\}$*, and level* $\tilde{\alpha}^k$*, call the triple* $(k, m, \tilde{\alpha}^k)$ *an* **alteration** *of the equilibrium. Call* $(k, m, \tilde{\alpha}^k)$ *a* **feasible alteration** *if Inequalities* (1.2) *and* (1.3) *are satisfied.*

Two beneficial alterations were illustrated at the end of Section 1.2.5. In Figure 1.3, 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 $(2, 28, 1)$; that is, player 1 would jump ahead from period 2 to period 30. In Figure 1.4, because all bad types betray in period 30 in equilibrium, when period 30 is reached, player 1 would desire an alteration $(30, -1, \alpha^{29})$, effectively going back to period 29 where all types cooperate.

Recall that, for a given trusting 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, and at the beginning of this period, player 1 believes that player 2's type is weakly below x^{k-1}. Further, after selecting the level α^k that the equilibrium prescribes for period k, player 1 expects player 2 to cooperate with probability $F(x^k)/F(x^{k-1})$ and to betray with complementary probability. We can thus express player 1's expected continuation value recursively as

$$v_1^k = \left(1 - \frac{F(x^k)}{F(x^{k-1})}\right)\left(-c\alpha^k\right) + \frac{F(x^k)}{F(x^{k-1})}\left(\alpha^k \Delta + \delta v_1^{k+1}\right). \tag{1.4}$$

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

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

Definition 5. *Call a PBE* **alteration proof** *if it is trusting and no feasible alteration improves* *player 1's continuation value. That is, for every feasible alteration $(k, m, \tilde{\alpha}^k)$,*

$$v_1^k \geq \left(1 - \frac{F(x^{k+m})}{F(x^{k-1})}\right)\left(-c\tilde{\alpha}^k\right) + \frac{F(x^{k+m})}{F(x^{k-1})}\left(\tilde{\alpha}^k\Delta + \delta v_1^{k+1+m}\right). \tag{1.5}$$

We refer to these as alteration-proof equilibria.

1.3.2 Partial characterization

We next partially characterize alteration-proof equilibria. Recall that L denotes the last period in which bad types betray on the equilibrium path, so that $x^{L-1} \geq a \geq x^L$.

Lemma 2. *In every alteration-proof equilibrium, good types never betray and the level is maximal after all bad types have betrayed. That is, $\alpha^k = 1$ for every $k > L$, and it can be assumed that $x^k = a$ for every $k \geq L$.*

Proof. Consider any alteration-proof PBE and let $\eta \equiv \sup\{\alpha^k \mid k > L\}$. We first prove that $\eta = 1$ by assuming otherwise and finding a contradiction. Presuming $\eta < 1$, let $\varepsilon > 0$ be small enough to satisfy $\eta + \varepsilon \leq 1$ and

$$\frac{1}{r} < \frac{\Delta}{1-\delta} + \varepsilon \cdot \frac{\Delta - (1+\delta)(1/r)}{\eta(1-\delta)}. \tag{1.6}$$

There is such a value of ε because $1/r < \Delta/(1-\delta)$. Let $\tau > L$ be a period at which $\alpha^\tau > \eta - \varepsilon$, and note that player 1's belief at the beginning of period τ puts positive probability on only good types, which are weakly below $1/r$. By definition of η, we know that $v_1^\tau \leq \eta\Delta/(1-\delta)$.

Then consider alteration $(\tau, -1, \eta + \varepsilon)$. Observe that $(\eta+\varepsilon)x < (\eta+\varepsilon)\Delta + \delta(\eta-\varepsilon)x$ follows from Inequality (1.6) and $x < 1/r$. Further, the left side is type x's value of betraying immediately in the alteration, whereas the right side is weakly less than the value of waiting until period $\tau+1$ to betray. This implies that all types weakly below $1/r$ strictly prefer to cooperate in period τ given the altered sequence of levels, and so the alteration is feasible. The alteration gives player 1 the continuation value $(\eta+\varepsilon)\Delta + \delta v_1^\tau$, which strictly exceeds v_1^τ, contradicting alteration-proofness.

Having established that $\eta = 1$, we can use a similar argument to show that there is a period $\ell > L$ at which $\alpha^{\ell} = 1$. If there were no such period, then for any $\varepsilon > 0$ we could find a period $\tau > L$ such that $\alpha^{\tau} \in (1-\varepsilon, 1)$ where it would have to be the case that $v_1^{\tau} < \Delta/(1-\delta)$. For sufficiently small ε, the alteration given by $(\tau, -1, 1)$ is feasible and yields player 1 a strictly higher continuation value than v_1^{τ}. Thus there is a period $\ell > L$ where $\alpha^{\ell} = 1$. The same logic implies also that $v_1^{\ell} = \Delta/(1-\delta)$, and so $\alpha^k = 1$ for all $k \geq \ell$.

The penultimate step is to realize that, in the case of $\ell > L+1$, it must also be true that $\alpha^{\ell-1} = 1$ and $v^{\ell-1} = \Delta/(1-\delta)$. This follows from the fact that good types strictly prefer to cooperate in period $\ell - 1$ regardless of the level, given that their continuation value is $\Delta/(1-\delta)$ from period ℓ. If $\alpha^{\ell-1} < 1$ then alteration $(\ell-1, 0, 1)$ is trivially feasible and strictly increases player 1's continuation payoff from period $\ell - 1$. It follows by induction that $\alpha^k = 1$ for all $k > L$. Finally, note that, because $\alpha^L \leq 1$, all good types strictly prefer to cooperate forever rather than betray in period L or any later period. No good type betrays prior to period L in equilibrium, and therefore the good types never betray. □

Lemma 2 establishes that the cutoff sequence $\{x^k\}_{k=0}^{\infty}$ never falls below $1/r$, and without loss of generality we can assume that $x^0 = b$ and $x^k = \Delta/(1-\delta)$ for every $k \geq L$. 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 L, where the value of the sequence drops to $\Delta/(1-\delta)$, which is below a, and is then constant.

Lemma 3. *In every alteration-proof equilibrium,* $\alpha^k x^k = \Delta\alpha^k + \delta\alpha^{k+1}x^k$ *for all* $k < L$.

Recall that this relation between the level sequence and cutoff types was discussed and appears as Equation (1.1) in the previous section. It means type x^k is 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:

$$\frac{\alpha^{k+1}}{\alpha^k} = \frac{x^k - \Delta}{x^k \delta}. \tag{1.7}$$

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,

$$v_2^{k+1}(x^k) = \alpha^{k+1} x^k \tag{1.8}$$

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 (1.8) 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

$$\alpha^k x^{k-1} \leq \Delta \alpha^k + \delta v_2^{k+1}(x^{k-1}) = \Delta \alpha^k + \delta \alpha^{k+1} x^{k-1}. \tag{1.9}$$

The equality holds because of $x^{k-1} = x^k$ and Equation (1.8).

Suppose that Inequality (1.9) is strict. We can find a level $\tilde{\alpha}^k \in (\alpha^k, 1)$ for which, uniquely,

$$\tilde{\alpha}^k x^{k-1} = \Delta\tilde{\alpha} + \delta v_2^{k+1}(x^{k-1}) = \Delta\tilde{\alpha}^k + \delta\alpha^{k+1}x^{k-1}. \tag{1.10}$$

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.[15] 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 (1.9) 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 (1.8) 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 (1.9) 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 yields $\alpha^k x^k = \Delta\alpha^k + \delta\alpha^{k+1}x^k$. 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 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.[16] Hence, $\alpha^k x^k = \Delta\alpha^k + \delta v_2^{k+1}(x^k)$. Using Equation (1.8) to substitute for $v_2^{k+1}(x^k)$ once again yields $\alpha^k x^k = \Delta\alpha^k + \delta\alpha^{k+1}x^k$. □

Together Lemmas 2 and 3 imply that, in an alteration-proof equilibrium, the level increases gradually until it reaches 1, and then remains at 1 thereafter. 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. Good types cooperate forever.

Note that Equation (1.7) and monontonicity of the sequences $\{\alpha^k\}_{k=1}^{\infty}$ and $\{x^k\}_{k=0}^{\infty}$ are

15. 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 $(\tilde{\alpha}^k - \delta\alpha^{k+1})x < (\tilde{\alpha}^k - \delta\alpha^{k+1})x^{k-1} = \Delta\tilde{\alpha}^k$.
16. It is easy to verify that $\omega(k,x)$ is continuous in x for fixed k.

necessary but not sufficient conditions for alteration-proofness. The set of equilibria that satisfy these properties is quite large and varied. For example, Equation (1.7) holds in the equilibria illustrated in Figures 1.3 and 1.4, but these equilibria fail to be alteration-proof.[17]

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

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

That is, the alteration is supposed to make type x^{k+m} indifferent, so that any higher types betray in period k and bad types in $[a, x^{k+m}]$ remain into the next period.

We'll use this equation, Identity (1.4), Equation (1.7), and Inequality (1.5). In the case of $m = -1$, we obtain

$$v_1^k \geq \frac{\Delta}{1-\delta}\alpha^{k-1}. \tag{1.11}$$

In the case of $m = 1$, we get

$$v_1^k \leq \frac{\Delta}{1-\delta}\alpha^k \left(\frac{F(x^k) - F(x^{k-1})}{F(x^{k-1})} \cdot \frac{c}{x^k} + \frac{F(x^k)}{F(x^{k-1})} \right). \tag{1.12}$$

We thus have 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

$$\frac{F(x^{k-1}) - F(x^k)}{\Delta F(x^{k-1})} \leq \frac{\alpha^k - \alpha^{k-1}}{(c+\Delta)\alpha^k - \delta c \alpha^{k+1}}. \tag{1.13}$$

The derivation of Inequalities (1.11), (1.12), and (1.13) is shown in Appendix 1.7.1.

---

17. Incidentally, Lemmas 2 and 3 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 an artificial type $\Delta/(1-\delta)$ indifferent.

1.3.3 Existence and multiplicity

We pointed out earlier that the set of trusting PBE is large. There are even multiple alteration-proof equilibria. The following definition describes a class of equilibria in which v_1^k/α^{k-1} is constant over time:

Definition 6. *Call an alteration-proof equilibrium a* **constant-proportion equilibrium** *(CPE) if there exists a number $\gamma \geq 1$ such that*

$$v_1^k = \gamma \cdot \frac{\Delta}{1-\delta}\alpha^{k-1} \tag{1.14}$$

for all $k \in \{2, 3, \ldots, L\}$.

Note that $\gamma = 1$ is the case in which Inequality (1.11) binds in each period. The next theorem, proved by construction in Appendix 1.7.2, establishes existence of CPE.

Theorem 2. *Under Assumption 1, there is a number $\bar{\gamma} > 1$ such that for all $\gamma \in [1, \bar{\gamma}]$, the trust game has a constant-proportion equilibrium with parameter γ.*

Let make three remarks on Theorem 2. First, $\bar{\gamma}$ depends on the parameters of the game including Δ. Fixing the other parameters, as Δ decreases toward zero, the value of $\bar{\gamma}$ derived in the proof decreases. Second, a constant-proportion equilibrium with $\gamma = 1$ always exists, regardless of Δ and the other parameters of the game. Substituting $v_1^k/\alpha^{k-1} = v_1^{k+1}/\alpha^k = \Delta/(1-\delta)$ into the equilibrium Identity (1.4) and rearranging terms yields

$$\frac{\alpha^k - \alpha^{k-1}}{\alpha^k} = -\left(c\frac{1-\delta}{\Delta} + 1\right)\frac{F(x^k) - F(x^{k-1})}{F(x^{k-1})}. \tag{1.15}$$

This equation, together with indifference Condition (1.7), characterizes the $\gamma = 1$ constant-proportion equilibrium. Figure 1.5 illustrates this equilibrium for the same set of parameters used in Section 1.2.5.

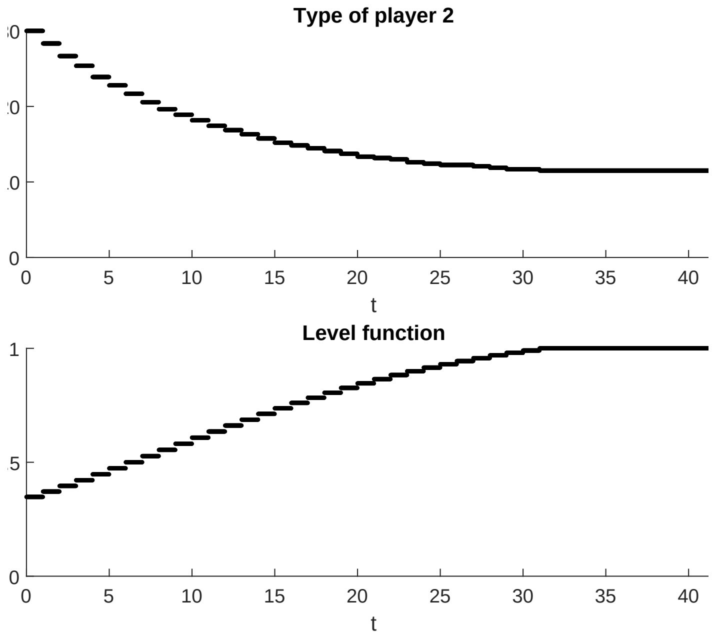


Figure 1.5. Constant-proportion equilibrium with $\gamma = 1$

Third, Theorem 2 says nothing about alteration-proof equilibria that are outside the constant-proportion class. We have been able to calculate some other equilibria, but we have been unable to show that they are bounded in some way by constant-proportion equilibria. Thus, although the constant-proportion class provides a useful illustration of alteration-proof equilibria and demonstrates multiplicity, we cannot restrict attention to this class in the next stage of our analysis.

1.3.4 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 by $\Delta(j)$, 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 $\{x^k(j)\}_{k=0}^{\infty}$, 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

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

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

The theorem establishes 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. 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 uniquely solve a specific initial-value problem (differential equation) that depends on the parameters.

Theorem 3. *Fix the parameters of a trust game, with the exception of Δ, satisfying Assumption 1. 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 equilibria 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 $(0,T)$, $\alpha(0) > 0$, $\lim_{t\to T^-} \alpha(t) = 1$, and $\alpha(t) = 1$ for every $t \geq T$;

(b) The cutoff-type function $x(\cdot)$ is strictly decreasing and differentiable on $(0,T)$, $x(0) = b$, $\lim_{t\to T^-} x(t) = a$, 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:

$$\frac{\alpha'}{\alpha} = -(rc+1)\frac{f(x)}{F(x)}x', \tag{1.16}$$

$$\frac{\alpha'}{\alpha} = r - \frac{1}{x}\ . \tag{1.17}$$

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

The existence of multiple and varied alteration-proof equilibria in the discrete-time setting presents a substantial challenge for the limit characterization. Our proof of Theorem 3, provided in Appendix 1.7.3, uses novel techniques in two steps. First, we use equilibrium identities and alteration-proofness conditions to find bounds on $\{x^k\}_{k=1}^{L}$, working backward from period L and using a calculation that maximizes and minimizes the cutoff type in a given period by making adjustments to the cutoff type in the next period, along with other variables in these periods. We discover a monotone relation on these bounds for successive periods, which allows us to use an inductive argument to construct bounds. Second, we apply the new convergence result proved by Watson (2021) to show that these two sequences of bounds converge uniformly to the same continuous time limit initial-value problem. The corresponding bounds of sequences $\{\alpha^k, v_1^{k+1}\}_{k=1}^{L}$ are constructed similarly. 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 (1.11) and (1.12) provide upper and lower bounds on player 1's continuation values on the equilibrium path. Assuming well-behaved convergence, we have

$$\lim_{j\to\infty}\left[\frac{F(\hat{x}(t;j)) - F(\hat{x}(t-\Delta(j);j))}{F(\hat{x}(t-\Delta(j);j))}\cdot\frac{c}{\hat{x}(t;j)} + \frac{F(\hat{x}(t;j))}{F(\hat{x}(t-\Delta(j);j))}\right] = 1, \tag{1.18}$$

and α^k and α^{k-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 Equation (1.15) and player 2's indifference Condition (1.7) lead to Equations (1.16) and (1.17), respectively.

The system of equations shown in Theorem 3 can be solved as follows. First, use Equation (1.17) to substitute for α'/α in Equation (1.16) to get the following univariate initial-value problem:

$$\frac{dx}{dt}=\frac{(1-rx)F(x)}{x(rc+1)f(x)},\quad x(0)=b. \tag{1.19}$$

Denote

$$I_x(x)=\int\frac{x(rc+1)f(x)}{(1-rx)F(x)}dx. \tag{1.20}$$

Then we solve Equation (1.19) 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 (1.17) to obtain the following univariate initial-value problem:

$$\frac{d\alpha}{\alpha}=\left(r-\frac{1}{x(t)}\right)dt,\quad \alpha(T)=1, \tag{1.21}$$

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

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

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.

1.4 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 types as $q \equiv F(1/r)$. 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.

In this example, Initial-Value Problems (1.19) and (1.21) become

$$\frac{dx}{dt} = \frac{1-rx}{r+1} \cdot \frac{(b-a)q + (x-a)(1-q)}{(1-q)x}, \quad x(0) = b \tag{1.22}$$

$$\frac{d\alpha}{\alpha} = \left(r - \frac{1}{x}\right) dt, \quad \alpha(T) = 1. \tag{1.23}$$

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

$$t = \frac{(1-q)(1+r)}{1-q+r(bq-a)} \left(-\frac{1}{r} \ln \frac{1-rx}{1-rb} + \frac{a-bq}{1-q} \ln \frac{(1-q)x+bq-a}{b-a} \right), \tag{1.24}$$

$$\alpha(t) = \left(\frac{q(b-a)}{(1-q)x(t)+bq-a} \right)^{r+1}, \tag{1.25}$$

$$T = \frac{(1-q)(1+r)}{1-q+r(bq-a)} \left(-\frac{1}{r} \ln \frac{1-ra}{1-rb} + \frac{a-bq}{1-q} \ln q \right), \ \alpha(0) = q^{r+1}. \tag{1.26}$$

Figure 1.6 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 1.7.4 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, $x(\Gamma(\chi)) = \chi$.

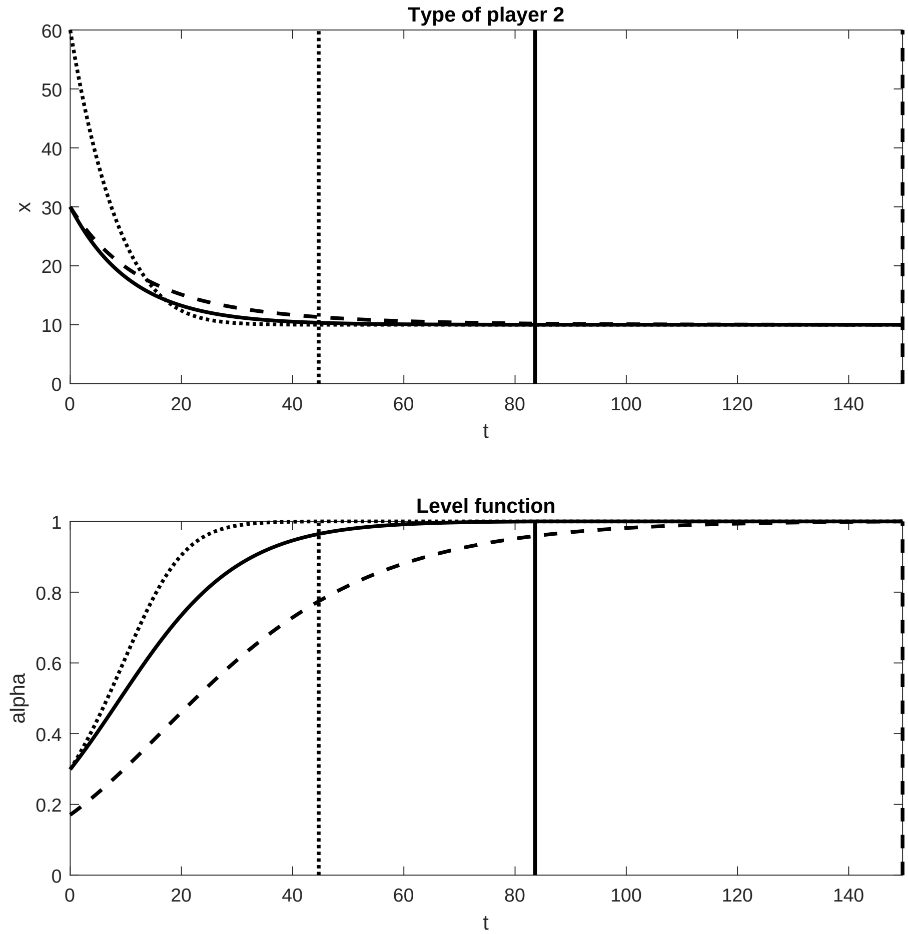


Figure 1.6. The limit of alteration-proof equilibria. Continuous-time level-curve limit for alteration-proof equilibria with parameters $r = 0.1$, $a = 10.0105$, and $c = 1$. For the comparative statics, we use the following parameters: $q = 0.3337$ and $b = 30$ for solid curves, $q = 0.2002$ and $b = 30$ 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.

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 (cooperate) 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 Paul A. Gompers (1995) and Tian (2011).

1.5 Technical Notes and Extensions

In this section, we elaborate on the foundations of the model, we discuss extensions of the modeling exercise related to both the alteration-proofness concept and the parameters, and we comment on connections with literature.

1.5.1 Further selection and foundations

In this subsection, we elaborate on the scope and foundation of alteration-proofness. First, recall that our definitions in Section 1.3 impose alteration-proofness only on the equilibrium path. To the extent that renegotiation is also plausible in off-equilibrium contingencies, we should further require alteration-proofness after histories in which player 1 deviated (the only histories entailing a public deviation in a trusting PBE).

In fact, the equilibria that we construct to prove Theorem 2 satisfy the extended version of alteration-proofness. In a period in which player 1 deviates, every type of player 2 is supposed to betray, ending the game. If player 2 deviates by cooperating, then player 1 subsequently believes with certainty that player 2 is a bad type, and the strategies specify that player 1 set $\alpha = 0$ and player 2 betray in all future periods regardless of the interim history. From every period following player 1's deviation, there are no trusting equilibria in the continuation game due to player 1's posterior belief, and therefore no feasible alteration can give player 1 a positive payoff.

Second, recall that we defined alteration-proofness for trusting equilibria, and this is the class of equilibria to which our convergence result applies. There are also non-trusting equilibria in which player 1 always chooses level 0. Further, using the logic noted in the previous paragraph, we can find a non-trusting equilibrium that satisfies the alteration-proofness conditions. This leads to the question of whether there is an argument along the lines of alteration-proofness that would rule out non-trusting equilibria, so that our convergence implies a unique selection.

Non-trusting equilibria would not survive a reasonable notion of *external consistency* imposed on top of our alteration-proofness concept, because player 1 and every type of player 2 strictly prefer the alteration-proof PBE that we construct for Theorem 2 to every non-trusting equilibrium (which gives a payoff of zero to both players), both at the beginning of the relationship and in later periods. For instance, suppose in addition to the alterations studied in Section 1.3, the players view as viable a suggestion to switch to another alteration-proof equilibrium that would improve the payoff of every player-type by at least ε for some fixed $\varepsilon > 0$. Then for sufficiently small Δ, none of the alteration-proof equilibria could dominate others in this way (due to Theorem 3) but they all dominate every non-trusting equilibrium.[18]

Continuing with the topic of non-trusting equilibria, it is worth noting that the existence of such equilibria depends on good types having positive betrayal benefits ($x > 0$), so that they have an incentive to betray at positive levels if they expect that the level would be zero in future

18. This comparison is similar to the "Pareto external agreement consistency" condition of Miller and Watson (2013). We utilize the bound ε because we have not determined whether and how alteration-proof equilibria are ranked for fixed Δ.

periods. We conjecture that inclusion of good types with $x \leq 0$ would narrow the set of PBE in interesting ways without even imposing alteration-proofness. This may be a good topic for future work. We have been able to show that, in this setting, there is no non-trusting equilibrium, and all good types cooperate forever in every PBE.[19] This is also true for continuations following a deviation of player 1. Therefore, in any equilibrium construction, we cannot specify reversion to a non-trusting path to punish player 1, complicating the argument for existence. As a bridge to further research, in the Supplementary Appendix, we provide a new existence result that does not rely on non-trusting equilibria following deviations.

Finally, it would not be difficult to provide non-cooperative foundations for the alteration-proofness condition, along the lines of Watson (2013) and Miller and Watson (2013). We could model renegotiation at the beginning of each period as a simple dictator game, whereby player 1 has the option to declare an alteration, which the players then coordinate on if feasible. In fact, since, in the construction behind Theorem 4, deviations in the level chosen for the current period are associated with feasible alterations, one could imagine that player 1 triggers an alteration by simply deviating from the level that the equilibrium specifies for the current period.

1.5.2 No-gap case

Recall that we assumed distribution F is constant in an open interval containing $\Delta/(1-\delta)$; that is, there is a gap at the value where player 2 would be indifferent between betraying and cooperating if the level were constant. Our analysis can be extended to the "no-gap" case, which we can describe by setting $a = 1/r$ and where bad types are those above $\Delta/(1-\delta)$. Without going into details, here is a summary of our findings in this case.

The PBE characterization developed in Section 1.2 carries over to the no-gap case except that Lemma 1 must be modified to allow $L = \infty$. The intuition is that bad types with x close to $\Delta/(1-\delta)$ are willing to cooperate in periods where α rises slowly, so a strictly increasing

19. From a non-trusting strategy profile, if player 1 deviates to a positive level, good types with $x \leq 0$ optimally must cooperate regardless of the anticipated future play. If player 1 deviates in this way and player 2 then cooperates, the probability that player 1 puts on the good types must weakly increase.

sequence $\{\alpha^k\}$ that converges to a number in the interval $(0,1]$ is associated with a strictly decreasing sequence of cutoff types $\{x^k\}$ that converges to $\Delta/(1-\delta)$. Thus, every bad type betrays at some point, but there are PBE in which, in every period, some bad types have yet to betray.[20]

Likewise, the analysis of alteration-proofness extends to the no-gap case. Lemmas 2 and 3 hold, allowing for $L=\infty$. Calculations underlying our convergence theorem are valid, but in the limit T becomes infinite and Expression (1.20) becomes an improper integral, with terminal condition $x^L = a = \Delta/(1-\delta)$ at $L=\infty$. The convergence result now identifies a class of solutions that is unique up to a constant $\alpha^* \in (0,1]$. Specifically, in the limit as $\Delta \to 0$, alteration-proof equilibria all share the same path of cutoff types. For every $\alpha^* \in (0,1]$ there is a sequence of alteration-proof equilibria whose level sequences converge to a function that approaches α^* as $t \to \infty$.

This conclusion is in contrast to alteration-proofness in the gap case, where the level converges to 1 (meaning $\alpha^* = 1$ for every alteration-proof equilibrium) and does so in finite time. But the equilibria are Pareto-ranked in α^*, so an appeal to external consistency would justify selection of the equilibria that entail $\alpha^* = 1$ as a unique limit prediction. Another way to restore our unique prediction is to expand the definition of "feasible alteration" to include player 1 setting $\alpha = 1$ in every period of the continuation, with good types cooperating perpetually and bad types betraying immediately.[21] In an equilibrium with $\alpha^* < 1$, player 1 would strictly prefer such an alteration once only a small fraction of bad types remain. Further, in line with the various notes made here about including good types with $x \leq 0$, we conjecture that in settings with such types, there would be no substantive difference between the analysis of the gap case and the no-gap case.

20. The no gap case here is similar to the no-gap case of the durable-good-monopoly model; see Ausubel and Deneckere (1989).

21. Such an alteration is not included in the definition of "feasible" given in Section 1.3 because the continuation from the next period does not coincide with any continuation on the path of the current equilibrium.

1.5.3 More on related literature

Let us expand a bit on our discussion of related literature in the Introduction. As noted, the most closely related paper is Watson (1999). While our modeling exercise shares heuristic logic with Watson's modeling exercise, they have distinct structures and very different analytical approaches.

Watson's model has an exogenously provided level function and therefore is not a noncooperative game, whereas we develop a fully noncooperative model of a principal-agent setting in which the principal selects the level in each period. We define alteration-proofness with respect to actual alternative equilibria in the continuation of the game from any period, and we are able also to study alteration-proofness at contingencies off the equilibrium path. Further, we study a discrete-time setting with a continuum of types, whereas Watson (1999) looked at a continuous-time setting with just two types. The analytical methods developed herein are novel and unique to our setting, with particular challenges owing to discrete time. The analysis of convergence (as well as the existence of multiple alteration-proof equilibria) has no counterpart in Watson (1999). Our work in this regard involves a new method of characterizing bounds on the set of equilibria and the first application of a result on the limit of solutions to discrete-time models.

A number of past game theoretic modeling exercises have delivered notable results on the convergence of equilibrium outcomes as the frequency of interaction increases. For instance, Gul, Sonnenschein, and Wilson (1986) substantiate the Coase conjecture on dynamic pricing by a durable-good monopolist (similarly bargaining under incomplete information with one-sided offers by the uninformed-party). The seller in a given period evaluates the rate at which sales increase as the price is lowered (attracting additional lower-valuation buyers) and this interacts with the relative patience of different buyer types. As the period length shrinks, the balance tilts in favor of the seller's interest in lowering the price to increase present-period sales, and in the limit the seller sets a low price from the beginning.

There is a similar trade-off in our model. By raising the level in an equilibrium alteration, the principal can induce more bad types to betray, hastening the time when the principal enjoys cooperation with the good types. One might expect, in line with the Coase conjecture, that as the period length shrinks, the principal would be resigned to start with a high level and suffer the consequences in the event of a bad type of player 2. This is not the case, however, because the relative intertemporal trade-offs for the principal and agent in the trust game are different than for the seller and buyer in the dynamic pricing game. The principal's choice of project level scales the payoffs of both players rather than splitting the the surplus for the seller and buyer, and the agent's action is a choice of whether to divert gains at the principal's expense. Importantly, betrayal by the bad types imposes a loss on the principal, whereas selling to a high-valuation buyer at a low price still generates value to a monopolist. Further, in periods when the level is low, the principal still earns a cooperative flow payoff from the good types, whereas for the durable-good monopolist, pricing high earns no flow benefit from low-valuation buyers. For these reasons, the principal's incentive to increase the level is tempered as the period length shrinks, and the result is gradualism.

Also relevant to the theme of frequent play is the topic of "repeated games with incomplete information" (though technically not repeated games), such as analyzed by Hart (1985) and Shalev (1994). One strand of this literature looks at how incomplete information about stage-game payoffs leads to a "reputation-based refinement" of equilibrium predictions relative to a complete-information benchmark in settings with sufficiently patient players.[22] In the typical setting, a long-run player with private information is, in the limit, infinitely more patient than the other players (a special case being a sequence of short-run players). Any type of long-run player can patiently pretend to be any other type, and the short-run players must either best respond to the mimicked type's strategy in the short run or "learn" that the long-run player is this type, which implies bounds on the long-run player's payoffs. Our setting is quite different because the

22. Prominent entries include Fudenberg and Levine (1989), Cripps, Schmidt, and Thomas (1996), and Cripps and Thomas (2003). Watson (1993) and Battigalli and Watson (1997) examine implications for rationalizable beliefs. Pei (2021) explores reputations based on randomization that straddles multiple types.

flow payoff of cooperation is scaled by the period length, whereas the terminal payoff of betrayal is not scaled. The intertemporal incentives are fixed (the discount rate is held constant) while we shrink the period length. This smooths the equilibrium behavior, but neither player becomes infinitely more patient than the other, and so there is no extreme reputation effect.[23]

1.5.4 Other extensions

Our analysis took place under the assumption that $c > 0$, but it easily extends to the case of $c = 0$, and curiously it still has interesting things to say. With no cost of betrayal, player 1's evaluation of alterations trades off discounting with the probability of cooperation. Player 1 prefers that bad types not betray immediately and so starts small. Alteration-proofness implies a unique equilibrium in discrete time, where player 1 is always indifferent between continuing on the equilibrium path and altering with $m = -1$ and $m = 1$. The limit differential equations can be derived from the expressions shown in Theorem 3 by setting $c = 0$.

The assumption that player 1 can dictate the selection of an equilibrium alteration makes alteration-proofness a tight condition. A renegotiation-proofness condition requiring agreement between the players would be weaker, because there are stall alterations that appeal to player 1 but would not appeal to any type of player 2. Any alteration with $m > 0$ 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.

That the game ends after betrayal is an assumption made for convenience. We think that our results would not be different in a model in which play continues following betrayal. For instance, to extend the trusting equilibria we have studied, we need to deal with histories in which betrayal has occurred in the past. If the betrayal occurred on the equilibrium path, player 1's updated belief would be that player 2 is a bad type and player 1 would choose $\alpha = 0$ thereafter.

23. Another line in the literature on incomplete information examines the effect of "behavioral types," but this is farther afield from our project.

For a public-deviation betrayal, such as in a period in which all types were supposed to cooperate at the equilibrium level, then we could specify that player 1's posterior belief is concentrated on bad types and player 1 selects $\alpha = 0$ thereafter. Further, if player 1 were to deviate, we could prescribe that all types of player 2 betray in the current period, so player 1's belief is unchanged after this betrayal, and continuation play from the next period is exactly what the players were supposed to do from the current period (a stall of sorts). These provisions would not complicate alteration-proofness.

1.6 Conclusion

We have added to the literature on relationship building by characterizing alteration-proof equilibria in a discrete-time principal-agent setting with a continuum of bad types. We hope that our closed-form characterization of equilibrium will motivate 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 think the 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.

1.7 Appendix

1.7.1 Derivation of (1.11), (1.12) and (1.13)

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

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

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

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

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

Therefore, alteration-proofness Condition (1.5) in this case simplifies to

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

which yields Inequality (1.11).

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

$$v_1^{k+1} \frac{F(x^k)}{F(x^{k-1})} - \left(1 - \frac{F(x^k)}{F(x^{k-1})}\right) c\alpha^{k+1}, \tag{1.27}$$

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

To apply the Condition (1.5), we need to express x_1^{k+1} in terms v_1^{k+1}. From Equation (1.4), we can solve for v_1^{k+1} and get

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

Using this to substitute for v_1^{k+1} in Expression (1.27) and comparing it with v_1^k, we obtain the following alteration-proofness condition:

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

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

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

which simplifies to

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

Notice that Equation (1.7) can be written $\alpha^k - \delta\alpha^{k+1} = \Delta\alpha^k/x^k$. This substitution yields Inequality (1.12).

Last, we combine Inequalities (1.11) and (1.12). Note that player 1's continuation value satisfying both conditions only if

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

Using player 2's indifference Condition (1.7) and rearranging terms, we obtain Inequality (1.13).

1.7.2 Proof of Theorem 2

Local and global alteration-proofness

We start with a useful lemma on the sufficiency of local alteration-proofness, which refers to values $m \in \{-1,0,1\}$.

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

Proof. For any integer $m \geq 0$, we first write player 1's continuation value v_1^k in the original equilibrium as

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

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

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

We will show that if player 1 has no incentive to alter the equilibrium from period k by jumping to period $k+m-1$ and also no incentive to alter from period $k+m-1$ by jumping to period $k+m$, then player 1 also has no incentive to alter from period k by jumping to period $k+m$. We prove this by contradiction: assuming $v_1^k \geq \tilde{v}_1^k(m-1)$ and $v_1^k < \tilde{v}_1^k(m)$, we will derive $v_1^{k+m-1} < \tilde{v}_1^{k+m-1}(1)$ meaning that, in the original equilibrium, a local alteration-proofness condition in period $k+m-1$ is violated.

Alteration-proofness condition $v_1^k \geq \tilde{v}_1^k(m-1)$ can be simplified to

$$(1-\delta^{m-1}) \delta \frac{F(x^{k+m-1})}{F(x^{k-1})} v_1^{k+m} \leq \sum_{n=0}^{m-1} \left((c+\Delta) \frac{F(x^{k+n})}{F(x^{k-1})} - c \frac{F(x^{k+n-1})}{F(x^{k-1})} \right) \delta^n \alpha^{k+n} - \left((c+\Delta) \frac{F(x^{k+m-1})}{F(x^{k-1})} - c \right) \alpha^{k+m-1}. \quad (1.28)$$

Likewise, alteration-proofness condition $v_1^k < \tilde{v}_1^k(m)$ can be simplified to

$$(1-\delta^m)\delta\frac{F(x^{k+m})}{F(x^{k-1})}v_1^{k+m+1} > \sum_{n=0}^{m}\left((c+\Delta)\frac{F(x^{k+n})}{F(x^{k-1})} - c\frac{F(x^{k+n-1})}{F(x^{k-1})}\right)\delta^n\alpha^{k+n} - \left((c+\Delta)\frac{F(x^{k+m})}{F(x^{k-1})} - c\right)\alpha^{k+m}. \quad (1.29)$$

Next we rewrite the equilibrium Identity (1.4) for period $k+m$,

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

and then multiply both sides by $(1-\delta^m)F(x^{k+m-1})/F(x^{k-1})$, allowing us to substitute the lower bound of $(1-\delta^m)\delta v_1^{k+m+1}F(x^{k+m})/F(x^{k-1})$ using Inequality (1.29). After collecting terms, we get a lower bound of v_1^{k+m}:

$$v_1^{k+m}(1-\delta^m)\frac{F(x^{k+m-1})}{F(x^{k-1})} > \left(c - c\frac{F(x^{k+m-1})}{F(x^{k-1})}\right)\alpha^{k+m} + \sum_{n=0}^{m-1}\left((c+\Delta)\frac{F(x^{k+n})}{F(x^{k-1})} - c\frac{F(x^{k+n-1})}{F(x^{k-1})}\right)\delta^n\alpha^{k+n}. \quad (1.30)$$

Note that the summation term on the right here also appears in Inequality (1.28). Using Inequality (1.28) to substitute for the summation term in Inequality (1.30) yields the following:

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

Applying Condition (1.7) to substitute $\alpha^{k+m-1}(x^{k+m-1}-\Delta)/(x^{k+m-1}\delta)$ for α^{k+m}, this inequality simplifies to

$$v_1^{k+m} > \alpha^{k+m-1}\left(\frac{\Delta}{1-\delta} + \frac{F(x^{k-1}) - F(x^{k+m-1})}{F(x^{k+m-1})}\cdot\frac{x^{k+m-1}-\Delta/(1-\delta)}{x^{k+m-1}\delta}c\right). \quad (1.31)$$

On the other hand, by Inequality (1.12), local alteration-proofness in period $k+m-1$ requires

$$v_1^{k+m-1} \leq \frac{\Delta}{1-\delta}\alpha^{k+m-1}\left(\frac{F(x^{k+m-1})-F(x^{k+m-2})}{F(x^{k+m-2})}\cdot\frac{c}{x^{k+m-1}}+\frac{F(x^{k+m-1})}{F(x^{k+m-2})}\right),$$

which translates into an upper bound on v_1^{k+m} using the period $k+m-1$ equilibrium Identity

$$v_1^{k+m-1} = \left((c+\Delta)\frac{F(x^{k+m-1})}{F(x^{k+m})}-c\right)\alpha^{k+m-1}+\frac{F(x^{k+m-1})}{F(x^{k+m-2})}\delta v_1^{k+m}.$$

Therefore, local alteration-proofness requires

$$v_1^{k+m} \leq \alpha^{k+m-1}\left(\frac{\Delta}{1-\delta}+\frac{F(x^{k+m-2})-F(x^{k+m-1})}{F(x^{k+m-1})}\cdot\frac{x^{k+m-1}-\Delta/(1-\delta)}{\delta x^{k+m-1}}c\right). \tag{1.32}$$

Recall Inequalities (1.28) and (1.29), together with equilibrium Conditions (1.4) and (1.7), imply that v_1^{k+m} is bounded below by the right side of Inequality (1.31). We also showed that the local alteration-proofness Condition (1.12) and equilibrium Identity (1.4) imply that v_1^{k+m} is bounded above by the right side of Inequality (1.32).

However, for all $m \geq 1$, we have $x^{k+m-2} \leq x^{k-1}$, so the upper bound of v_1^{k+m} (the right side of Inequality (1.32)) is weakly lower than the lower bound of v_1^{k+m} (the right side of Inequality (1.31)), which is a contradiction. Therefore, we conclude that if player 1 has no incentive to jump from period k to period $k+m-1$, and player 1's local alteration-proof conditions are satisfied, then player 1 must have no incentive to jump from period k to period $k+m$. Applying this argument recursively for $m \in \{1,2,...\}$, we have proved that if player 1's alteration-proof conditions for $m \in \{-1,0,1\}$ are satisfied, then the global conditions for $m > 1$ are also satisfied. □

PBE construction: equilibrium sequence

Fix any trust game with period length $\Delta < \overline{\Delta}$. Define

$$\tilde{\gamma} \equiv \frac{a-\Delta}{a\delta}, \tag{1.33}$$

and consider any $\gamma \in [1, \tilde{\gamma}]$. We will construct a constant-proportion (alteration-proof) equilibrium with constant of proportion γ.

We begin by defining a sequence $\{x^k, \alpha^k, v_1^{k+1}\}_{k=0}^{L}$, where L will be determined in the construction. For $k > 0$ the meaning of x^k, α^k, and v_1^k is the same as in the text; α^k is the level chosen in period k on the equilibrium path, x^k is the type cutoff, and v_1^k is player 1's continuation value. The sequence gives these values from period 1 up to the last period L in which bad types betray in equilibrium. The initial cutoff type x^0 will equal b as with every PBE, whereas α^0 will be an artificial value that aids in the equilibrium construction. The sequence is defined by induction, starting with $k = L$ and working backward in time. The period offset for v_1 helps to organize the variables in the recursive step.

The inductive procedure uses Equations (1.4), (1.7) and (1.14), which we restate here:

$$v_1^k = \left(1 - \frac{F(x^k)}{F(x^{k-1})}\right)\left(-c\alpha^k\right) + \frac{F(x^k)}{F(x^{k-1})}\left(\alpha^k\Delta + \delta v_1^{k+1}\right), \tag{1.34}$$

$$\frac{\alpha^k}{\alpha^{k-1}} = \frac{x^{k-1}-\Delta}{x^{k-1}\delta}, \tag{1.35}$$

$$v_1^k = \gamma \cdot \frac{\Delta}{1-\delta}\alpha^{k-1}, \tag{1.36}$$

for $k \in \{1, 2, ..., L\}$. Recall that the first equation is the identity relating player 1's continuation values in adjacent periods, the second equation is the indifference condition for the cutoff type (here stated for periods $k-1$ and k), and the third is the constant-proportion condition that defines the equilibrium we are working to form.

We next describe the induction procedure to construct the sequence. Letting L be an arbitrary positive integer, set $x^L = a$, $\alpha^L = 1$, and $v_1^{L+1} = \Delta/(1-\delta)$. Then, for any k for which $(x^k, \alpha^k, v_1^{k+1})$ has been set and if the procedure has not yet terminated, derive $(x^{k-1}, \alpha^{k-1}, v_1^k)$ as follows. If for the given $(x^k, \alpha^k, v_1^{k+1})$, there is a vector $(x^{k-1}, \alpha^{k-1}, v_1^k)$ that solves the system of Equations (1.34)-(1.36), then $(x^{k-1}, \alpha^{k-1}, v_1^k)$ is taken to be this vector; further, if $x^{k-1} < b$ then the procedure continues by lowering the value of k by one unit and restarting the calculations. Otherwise (if there is no solution or if the solution has $x^{k-1} = b$), set $x^{k-1} = b$, set $\alpha^{k-1} = \delta\alpha^k b/(b-\Delta)$, redefine L so that the current value of k is 1, and terminate the procedure.[24]

The construction can be put in terms of the type cutoffs only, where a transition function relates x^k to x^{k-1}, without reference to the other variables α and v_1. Here are the calculations that yield the transition function:

For $k = L$, we plug in the terminal values $\alpha^L = 1$, $x^L = a$, and $v_1^{L+1} = \Delta/(1-\delta)$ into Equations (1.34)-(1.36), use Equation (1.35) to substitute for α^{k-1} in Equation (1.36), and then use the resulting equation to substitute for v_1^k in Equation (1.34). This yields the following equation that implicitly identifies x^{L-1}:

$$\Delta \frac{\frac{1-\gamma\delta}{1-\delta} - \frac{\gamma\delta}{x^{L-1}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\frac{\Delta}{1-\delta}} = \frac{F(x^{L-1}) - F(a)}{F(x^{L-1})}. \tag{1.37}$$

For $k \in \{1, 2, ..., L-1\}$, we can use Equation (1.36) to substitute for both v_1^k and v_1^{k+1} in Equation (1.34), which yields

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

Dividing both sides by α^k, using Equation (1.35) to substitute for α^{k-1}/α^k, and rearranging

24. By definition, type b would be indifferent between betraying in a given period at level α^0 and waiting until the next period to betray at level α^1.

terms yields

$$\Delta\frac{1-\frac{\gamma\delta}{x^{k-1}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}}=\frac{F(x^{k-1})-F(x^{k})}{F(x^{k-1})}. \tag{1.38}$$

Thus, the sequence $\{x^k\}_{k=0}^{L}$ is formed by first letting L be an arbitrary number to be defined later, setting $x^L = a$, and finding x^{L-1} to solve Equation (1.37). For $k < L$, x^{k-1} is defined inductively by Equation (1.38). At the point where there is no solution to the transition function or where the solution is exactly b, the procedure terminates, L is set to the number of rounds that occurred, and x^0 is set to b. Once $\{x^k\}_{k=0}^{L}$ has been determined, the corresponding values of α^k and v_1^{k+1} are easily calculated using Equations (1.35) and (1.36) and the initial value $\alpha^L = 1$.

We next show that the procedure to construct $\{x^k\}_{k=0}^{L}$ is well defined in that the solution to the system of equations, when it exists, is unique. For values of Δ, γ, and x^k for which Equation (1.38) has an interior unique solution, denote this solution by $x^{k-1} = \mu(x^k;\gamma,\Delta)$.

Lemma 5. *Under Assumption 1, for any $\gamma \in [1,\tilde{\gamma}]$ and $x^k \in [a,b]$, Equation* (1.38) *has at most one solution. At such a point where $\mu(x^k;\gamma,\Delta) \in (a,b)$, the solution is uniquely defined on a neighborhood of γ and x^k, the conditions of the implicit function theorem hold, and $d\mu(x^k;\gamma,\Delta)/dx^k \geq 0$ and $d\mu(x^k;\gamma,\Delta)/d\gamma \leq 0$. Equations* (1.37) *has the same properties.*

Proof. We first show that, under Assumption 1, the solution to Equation (1.38) is unique if it exists. Denote

$$\mathscr{F}(x^{k-1};x^k,\gamma) \equiv \Delta\frac{1-\frac{\gamma\delta}{x^{k-1}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}}-\frac{F(x^{k-1})-F(x^{k})}{F(x^{k-1})},$$

so that $\mathscr{F}(x^{k-1};x^k,\gamma) = 0$ at a solution point x^{k-1}. Observe that for all $x^k \in (a,b)$ and $\gamma \in [1,\tilde{\gamma}]$,

$$\mathscr{F}(x^k;x^k,\gamma)=\Delta\frac{1-\frac{\gamma\delta}{x^{k}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}}>\Delta\frac{1-\frac{\tilde{\gamma}\delta}{x^{k}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\tilde{\gamma}\frac{\Delta}{1-\delta}}=\Delta\frac{1-\frac{a-\Delta}{x^{k}-\Delta}\frac{\Delta/(1-\delta)}{a}}{c+\Delta+(a-\Delta)\frac{\Delta/(1-\delta)}{a}}>0.$$

The derivative of $\mathscr{F}$ with respect to x^{k-1} exists for every $x^{k-1} \in (a,b)$ and is equal to

$$\frac{\partial\mathscr{F}}{\partial x^{k-1}}=\frac{\Delta\frac{\gamma\delta}{(x^{k-1}-\Delta)^2}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}}-\frac{F(x^k)f(x^{k-1})}{(F(x^{k-1}))^2}=\frac{\Delta\frac{\gamma\delta}{(x^{k-1}-\Delta)^2}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}}-\frac{f(x^{k-1})}{F(x^{k-1})}\frac{c+\gamma\delta\frac{x^{k-1}}{x^{k-1}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}},$$

which is strictly negative under the assumption $\Delta \le \overline{\Delta}$. This implies that at most one value of x^{k-1} solves Equation (1.38).

Function $\mathscr{F}$ is continuously differentiable on the set of $x^k, x^{k-1} \in (a,b)$ and $\gamma \in (0,\tilde{\gamma})$, and the derivative with respect to x^{k-1} is nonzero. Applying the implicit function theorem to the identity $\mathscr{F} = 0$ yields

$$\frac{dx^{k-1}}{dx^k} = \frac{f(x^k)}{F(x^{k-1})} \cdot \left(\frac{F(x^k) f(x^{k-1})}{(F(x^{k-1}))^2} - \Delta \frac{\frac{\gamma\delta}{(x^{k-1}-\Delta)^2} \frac{\Delta}{1-\delta}}{c + \Delta + \delta\gamma\frac{\Delta}{1-\delta}} \right)^{-1}$$

and

$$\frac{dx^{k-1}}{d\gamma} = \left(\frac{\Delta}{x^{k-1} - \Delta} + \frac{F(x^{k-1}) - F(x^k)}{F(x^{k-1})} \right) \cdot \left(\frac{\Delta\gamma}{(x^{k-1}-\Delta)^2} - \frac{F(x^k) f(x^{k-1})}{F^2(x^{k-1})} \frac{c + \Delta + \delta\gamma\frac{\Delta}{1-\delta}}{\delta\frac{\Delta}{1-\delta}} \right)^{-1}.$$

Because

$$\frac{F(x^k) f(x^{k-1})}{(F(x^{k-1}))^2} > \Delta \frac{\frac{\gamma\delta}{(x^{k-1}-\Delta)^2} \frac{\Delta}{1-\delta}}{c + \Delta + \delta\gamma\frac{\Delta}{1-\delta}}, \text{ and } \frac{\Delta\gamma}{(x^{k-1}-\Delta)^2} \le \frac{F(x^k) f(x^{k-1})}{F^2(x^{k-1})} \frac{c + \Delta + \delta\gamma\frac{\Delta}{1-\delta}}{\delta\frac{\Delta}{1-\delta}}$$

for $\Delta < \overline{\Delta}$, we have $d\mu(x^k;\gamma,\Delta)/dx^k \ge 0$ and $d\mu(x^k;\gamma,\Delta)/d\gamma \le 0$.

The calculations for Equation (1.37) are similar to those above for Equation (1.38). $\square$

We conclude the constructive step by showing that L is finite. Because $\gamma \le \tilde{\gamma} \le (x^{k-1} - \Delta)/(x^{k-1}\delta)$, the left side of Equation (1.38) is positive and bounded away from zero. Assumption 1 guarantees that the slope of F is bounded away from zero. This implies that there is a number $\varepsilon > 0$ such that $x^{k-1} - x^k > \varepsilon$ for each k, proving that the inductive step terminates in a finite number of rounds, and this number of rounds is defined to be L so that the resulting sequence runs from $k = 1$ to $k = L$.

PBE construction: strategies

We next specify the strategy profile and verify that it is a PBE. On the equilibrium path through period L, the sequence of levels chosen by player 1 and the cutoff types for player 2 will be $\{\alpha^k, x^k\}_{k=1}^{L}$ as constructed above, and then $\alpha^k = 1$ and $x^k = a$ for every $k > L$. For any history to period k on this path, player 1's strategy prescribes level α^k in period k. Likewise, on this path through the middle of any period k, every type $x \geq x^k$ is supposed to betray and types below x^k cooperate. Player 1's updated belief at the beginning of each period $k+1$ is then given by F conditional on $x < x^k$.

For every history of play in which player 1 had at some point deviated from the prescribed sequence of levels, all types of player 2 are prescribed to immediately betray. For every history of play in which player 1 had at some point deviated from the prescribed sequence of levels and yet player 2 continued to cooperate (a further public deviation), player 1's updated belief assigns probability 1 to the bad type $x = a$ and player 1 is supposed to select $\alpha = 0$.

The specifications just described cover all histories. Beliefs accord to the conditional probability formula on the equilibrium path. It is easy to see that the strategies are sequentially rational. By the construction of $\{\alpha^k, x^k\}_{k=1}^{\infty}$, on the equilibrium path every type of player 2 optimally behaves as prescribed, with bad types betraying at the appointed periods and good types cooperating forever, and player 1 cannot gain by deviating (player 1's continuation value is strictly positive, whereas deviating would lead to a continuation value of zero). Clearly the prescribed behavior is rational off the equilibrium path.

Alteration-proofness of the constructed PBE

We next show that, for γ close to 1, the PBE constructed above is alteration-proof. We do this by proving that it is locally alteration-proof and then applying Lemma 4.

Lemma 6. *For any $\Delta > 0$, there exists $\bar{\gamma} > 1$, such that for all $\gamma \in [1, \bar{\gamma}]$, the constructed sequences $\{\alpha^k, x^k, v_1^{k+1}\}_{k=0}^{L}$ constitute a locally alteration-proof PBE.*

Proof. To verify the local alteration-proof Conditions (1.11) and (1.12), we first show the constructed $\{x^k, \alpha^k, v_1^{k+1}\}_{k=0}^{L}$ satisfies

$$\frac{\Delta}{1-\delta}\alpha^{k-1} \leq v_1^k \leq \frac{\Delta}{1-\delta}\alpha^k \left(\frac{F(x^k)-F(x^{k-1})}{F(x^{k-1})}\cdot\frac{c}{x^k}+\frac{F(x^k)}{F(x^{k-1})}\right), \tag{1.39}$$

for $k \in \{1, 2, ..., L+1\}$.

First, Inequality (1.39) holds for $k = L+1$, given $v_1^{L+1} = \Delta/(1-\delta)$, $\alpha^L = \alpha^{L+1} = 1$, and $x^L = x^{L+1} = a$. For $k = L$, substituting in $\alpha^L = 1$, $x^L = a$, $v_1^L = \gamma\alpha^{L-1}\Delta/(1-\delta)$, and $\alpha^{L-1} = \delta x^{L-1}/(x^{L-1}-\Delta)$, Inequality (1.39) becomes

$$\frac{\Delta}{1-\delta}\frac{\delta x^{L-1}}{x^{L-1}-\Delta} \leq \gamma\frac{\Delta}{1-\delta}\frac{\delta x^{L-1}}{x^{L-1}-\Delta} \leq \frac{\Delta}{1-\delta}\left(\frac{F(a)-F(x^{L-1})}{F(x^{L-1})}\cdot\frac{c}{a}+\frac{F(a)}{F(x^{L-1})}\right).$$

The left side of this condition is satisfied with $\gamma \geq 1$, and the right side of this condition can be simplified to

$$\frac{a}{c+a}\left(1-\gamma\frac{\delta x^{L-1}}{x^{L-1}-\Delta}\right) \geq \frac{F(x^{L-1})-F(a)}{F(x^{L-1})}$$

Note that the right side of this inequality is the same as the right side of Equation (1.37), so a sufficient condition is

$$\frac{a}{c+a}\left(1-\gamma\frac{\delta x^{L-1}}{x^{L-1}-\Delta}\right) \geq \Delta\frac{\frac{1-\gamma\delta}{1-\delta}-\frac{\gamma\delta}{x^{L-1}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\frac{\Delta}{1-\delta}}.$$

For $\gamma \leq \tilde{\gamma}$, this condition can be simplified to

$$\frac{a}{c+a} \geq \frac{\frac{\Delta}{1-\delta}}{c+\frac{\Delta}{1-\delta}},$$

which is satisfied for all $a > \Delta/(1-\delta)$. Hence, we conclude that Inequality (1.39) holds for $k = L$.

Next, we verify Inequality (1.39) for $k \in \{2, 3, ..., L-1\}$. The left inequality of this condition is trivially satisfied with $\gamma \geq 1$. Simplifying the right inequality of this condition using Equation (1.35), we have

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

We use the right side of Equation (1.38) to substitute for the right side of this inequality to get

$$\frac{x^k}{c+x^k}\left(1-\frac{x^{k-1}\gamma\delta}{x^{k-1}-\Delta}\right) \geq \Delta\frac{1-\frac{\gamma\delta}{x^{k-1}-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}},$$

which simplifies to

$$0 \geq (\delta\gamma)^2-\delta\gamma\left(\delta+\frac{\Delta}{x^{k-1}}\frac{c}{x^k}-\frac{1-\delta}{\Delta}c\right)-c\frac{1-\delta}{\Delta}\left(1-\frac{\Delta}{x^k}\right)\left(1-\frac{\Delta}{x^{k-1}}\right). \tag{1.40}$$

Denote the right side of this inequality as $\mathscr{G}(\gamma,x^{k-1};x^k,\Delta)$. Recall that x^{k-1} relates to γ and x^k according to $\mu(x^k;\gamma,\Delta)$, so we can write $\mathscr{G}(\gamma,\mu(x^k;\gamma,\Delta);x^k,\Delta)$ to substitute for x^{k-1}. We will show that $\mathscr{G}(\gamma,\mu(x^k;\gamma,\Delta);x^k,\Delta) \leq 0$ for all $x^k \in [a,b]$ and for all $\gamma \in [1,\min\{\tilde{\gamma},\hat{\gamma}\}]$, where

$$\hat{\gamma} \equiv \frac{1}{2\delta}\left(\left(\delta+\Delta\frac{c}{a^2}-\frac{1-\delta}{\Delta}c\right)+\sqrt{\left(\delta+\Delta\frac{c}{a^2}-\frac{1-\delta}{\Delta}c\right)^2+4c\frac{1-\delta}{\Delta}\left(1-\frac{\Delta}{a}\right)^2}\right).$$

When $\gamma=1$, $\mathscr{G}(1,\mu(x^k;1,\Delta);x^k,\Delta)$ is strictly negative. For $\gamma \in [1,\tilde{\gamma}]$, we apply the chain rule to obtain

$$\frac{d\mathscr{G}}{d\gamma}=\delta\left(2\delta\gamma-\delta+\frac{1-\delta}{\Delta}c-\frac{\Delta}{x^{k-1}}\frac{c}{x^k}\right)-\Delta\left(\frac{1-\delta}{\Delta}-\frac{1-\delta}{x^k}-\frac{\delta\gamma}{x^k}\right)\frac{c}{(x^{k-1})^2}\frac{dx^{k-1}}{d\gamma},$$

which is positive when $\Delta < \overline{\Delta}$. Note $\mathscr{G}(\gamma,\mu(x^k;\gamma,\Delta);x^k,\Delta)$ is continuous in γ, so there exists $\hat{\gamma}(x^k)$, such that $\mathscr{G}(\gamma,\mu(x^k;\gamma,\Delta);x^k,\Delta) \leq 0$ for all $\gamma \in [1,\hat{\gamma}(x^k)]$.

In fact, it suffices to set $\hat{\gamma}(x^k)=\hat{\gamma}$, so that this upper bound on γ implies $\mathscr{G}(\gamma,\mu(x^k;\gamma,\Delta);x^k,\Delta) \leq 0$ for all x^k. To see this, let us apply the chain rule to calculate the derivative of $\mathscr{G}$ with respect to x^k:

$$\frac{d\mathscr{G}}{dx^k} = (1-\delta)\frac{c}{(x^k)^2}\left(\frac{1-\delta+\delta\gamma}{1-\delta}\frac{\Delta}{x^{k-1}}-1\right)+(1-\delta)\frac{c}{(x^{k-1})^2}\left(\frac{1-\delta+\delta\gamma}{1-\delta}\frac{\Delta}{x^k}-1\right)\frac{dx^{k-1}}{dx^k}.$$

This value is negative when $\Delta < \overline{\Delta}$. The same conclusion holds for x^{k-1} because it enters the expression in a way that is symmetric to x^k. Thus $\mathscr{G}$ is decreasing in both x^k and x^{k-1}, and so plugging in a for these variables yields the highest value of $\mathscr{G}$ for a given γ. Because $\hat{\gamma}$ is the root of $\mathscr{G}(\gamma, a; a, \Delta) = 0$, we conclude that Inequality (1.40) is satisfied for all $\gamma \in [1, \min\{\hat{\gamma}, \tilde{\gamma}\}]$.

Finally, to complete the proof, we verify Inequality (1.39) for $k = 1$. We define α^0 as the artificial level such that type $x^0 = b$ is indifferent between betraying in period 0 with level α^0 and in period 1 with level α^1, that is $\alpha^0 \equiv \alpha^1\delta b/(b-\Delta)$. Together with $x^0 = b$ and Equation (1.34), Inequality (1.39) evaluated at $k = 1$ becomes

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

which rearranges to

$$\frac{\frac{(1-\delta)b-\Delta}{b-\Delta}+\delta(\gamma-1)}{c\frac{1-\delta}{\Delta}+1+\delta(\gamma-1)} \geq 1-F(x^1) \geq \frac{\delta(\gamma-1)}{c\left(\frac{1-\delta}{\Delta}-\frac{1}{x^1}\right)+\delta(\gamma-1)}. \tag{1.41}$$

Recall that the sequence is constructed backward in k and terminates at $x^0 = b$ if the remaining mass of bad types is too small for Equation (1.38) to hold—that is

$$\Delta\frac{1-\frac{\gamma\delta}{b-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}} \geq 1-F(x^1). \tag{1.42}$$

The left inequality of Expression (1.41) is implied by Inequality (1.42), as the following is easily verified:

$$\frac{\frac{(1-\delta)b-\Delta}{b-\Delta}+\delta(\gamma-1)}{c\frac{1-\delta}{\Delta}+1+\delta(\gamma-1)} \geq \Delta\frac{1-\frac{\gamma\delta}{b-\Delta}\frac{\Delta}{1-\delta}}{c+\Delta+\delta\gamma\frac{\Delta}{1-\delta}}.$$

The right inequality of Expression (1.41) is satisfied if

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

The right side of this inequality is always greater than 1 because $a \leq x^1 < x^0 = b$ by construction. For $\gamma = 1$, this condition is satisfied as a strict inequality. Because μ is continuous in γ and x^1 results from a finite number of applications of μ, x^1 is continuous in γ except possibly at discrete values where L changes in the sequence construction. Therefore, there is a bound $\check{\gamma}$ such that the right inequality of Expression (1.41) is satisfied for all $\gamma \in [1, \min\{\check{\gamma}, \tilde{\gamma}\}]$. So we can define

$$\overline{\gamma} \equiv \min\{\tilde{\gamma}, \hat{\gamma}, \check{\gamma}\}.$$

and then for all $\gamma \in [1, \overline{\gamma}]$, the constructed sequence $\{x^k, \alpha^k, v_1^{k+1}\}_{k=0}^{L}$ satisfies the local alteration-proof Condition (1.39). □

1.7.3 Proof of Theorem 3

Two lemmas

We start with a couple of lemmas that will be used at a few points in the main analysis.

Lemma 7. *For any alteration-proof PBE and for all $k \in \{2, ..., L\}$,*

$$\frac{v_1^k}{\alpha^{k-1}} \in \left[\frac{\Delta}{1-\delta}, \frac{\Delta}{1-\delta} \cdot \frac{x^{k-1}-\Delta}{x^{k-1}\delta} \right].$$

Further, $v_1^k/\alpha^{k-1} - \Delta/(1-\delta)$ is on the order of Δ.

Proof. From Inequality (1.11), we have $v_1^k/\alpha^{k-1} \geq \Delta/(1-\delta)$. From Inequality (1.12) and Condition (1.35), we have

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

which proves the uniform upper bound of v_1^k/α^{k-1}. On the second claim, note that the difference between the upper and lower bounds on v_1^k/α^{k-1} can be rewritten by rearranging terms as follows:

$$\frac{\Delta}{1-\delta}\frac{x^{k-1}-\Delta}{x^{k-1}\delta}-\frac{\Delta}{1-\delta}=\frac{\Delta}{\delta}\cdot\frac{\Delta}{1-\delta}\cdot\left(\frac{1-\delta}{\Delta}-\frac{1}{x^{k-1}}\right).$$

The term in parentheses is bounded away from zero for $x^{k-1}\in[a,b]$ and $\Delta/(1-\delta)$ converges to $1/r$ as Δ approaches zero, and so the entire expression is on the order of Δ. □

Lemma 8. *For any alteration-proof PBE, $x^{k+1}-x^k$ and $\alpha^{k+1}-\alpha^k$ are both on the order of Δ.*

Proof. From player 2's indifference Condition (1.7), rearranging terms gives us

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

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

$$\frac{\Delta}{\delta}\left(\frac{1-\delta}{\Delta}-\frac{1}{a}\right)\leq\frac{\alpha^{k+1}-\alpha^k}{\alpha^k}\leq\frac{\Delta}{\delta}\left(\frac{1-\delta}{\Delta}-\frac{1}{b}\right). \tag{1.43}$$

Note the left and right most terms are on the order of Δ, and when we have α^k strictly positive, then $\alpha^{k+1}-\alpha^k$ must be on the order of Δ.

Second, rewrite Inequality (1.13) to get

$$0\geq\frac{F(x^k)-F(x^{k-1})}{F(x^{k-1})}\geq\frac{c(\alpha^k-\delta\alpha^{k+1})+\Delta\alpha^{k-1}}{(c+\Delta)\alpha^k-\delta c\alpha^{k+1}}-1=\frac{\alpha^{k-1}-\alpha^k}{\alpha^k(c/x^k+1)}, \tag{1.44}$$

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

Change of variables and adjacent periods

We next move on to the main analysis for the proof of Theorem 3. Our calculations are simplified by introducing a new variable $w^k \equiv v_1^k/\alpha^{k-1}$, for all $k \in \{2,...,L+1\}$, which reduces the number of dimensions by alleviating v_1^k and α^{k-1} as separate variables. Note that, with the new notation, any alteration-proof PBE sequence $\{x^k,\alpha^k,v_1^{k+1}\}_{k=1}^L$ has a corresponding sequence $\{x^k,w^{k+1}\}_{k=1}^L$.

In fact, the PBE and alteration-proofness Conditions (1.4), (1.7), (1.11) and (1.12) can be expressed in terms of only the sequence $\{x^k,w^{k+1}\}_{k=1}^L$ as follows:

$$w^k = \frac{x^{k-1}-\Delta}{\delta x^{k-1}}\left(-c+\frac{F(x^k)}{F(x^{k-1})}(\Delta+c+\delta w^{k+1})\right) \tag{1.45}$$

$$\frac{\Delta}{1-\delta} \leq w^k \leq \frac{\Delta}{1-\delta}\frac{x^{k-1}-\Delta}{\delta x^{k-1}}\left(\frac{F(x^k)-F(x^{k-1})}{F(x^{k-1})}\cdot\frac{c}{x^k}+\frac{F(x^k)}{F(x^{k-1})}\right). \tag{1.46}$$

Equation (1.45) is derived by starting from Equation (1.4), substituting for v_1^k and v_1^{k+1} using $w^k = v_1^k/\alpha^{k-1}$ and $w^{k+1} = v_1^{k+1}/\alpha^k$, dividing by α^k, and using Equation (1.7) to substitute for α^{k-1}/α^k. The inequalities in Expression (1.46) are Inequalities (1.11) and (1.12) after substituting in w^k.

We define function $g\colon [a,b]\times[a,b]\times[0,\infty)\to[0,\infty)$ to give w^k as a function of x^k, x^{k-1}, and w^{k+1} according to Equation (1.45). That is,

$$w^k = g(x^k,x^{k-1},w^{k+1}) \equiv \frac{x^{k-1}-\Delta}{\delta x^{k-1}}\left(-c+\frac{F(x^k)}{F(x^{k-1})}(\Delta+c+\delta w^{k+1})\right).$$

By next studying the properties of g, we will be able to construct bounds on equilibrium sequences. In particular, we need to look at how the values of w^{k+1} and x^{k-1} relate over the two adjacent periods k and $k+1$, while fixing w^k, w^{k+1}, and x^{k+1}, and allowing x^k to vary.

Lemma 9. *Fix numbers w^k, $x^{k+1} \in [a,b]$ and w^{k+2}. Under Assumption 1, suppose that the* *following equations hold for some values* $\hat{x}^k, \hat{x}^{k-1} \in (a,b)$ *and* $\hat{w}^{k+1} \in \mathbb{R}$*:*

$$w^k = g(x^k, x^{k-1}, g(x^{k+1}, x^k, w^{k+2})) \text{ and } w^{k+1} = g(x^{k+1}, x^k, w^{k+2}). \tag{1.47}$$

These equations implicitly define x^{k-1} *and* w^{k+1} *as functions of* x^k *on a neighborhood of* $(\hat{x}^k, \hat{x}^{k-1}, \hat{w}^{k+1})$*. The implicit function theorem applies and* $dw^{k+1}/dx^k < 0$ *and* $dx^{k-1}/dx^k > 0$*.*

Proof. The first part of the lemma follows directly from equilibrium Identity (1.45) and Assumption 1. To show that w^{k+1} decreases in x^k, we differentiate

$$g(x^{k+1}, x^k, w^{k+2}) = \frac{x^k - \Delta}{\delta x^k}\left(-c + \frac{F(x^{k+1})}{F(x^k)}(\Delta + c + \delta w^{k+2})\right)$$

with respect to x^k to get

$$\frac{dw^{k+1}}{dx^k} = -c\frac{\Delta}{\delta (x^k)^2} + \frac{F(x^{k+1})}{F(x^k)}\left(\frac{\Delta}{x^k} - \frac{f(x^k)}{F(x^k)}(x^k - \Delta)\right)\frac{\Delta + c + \delta w^{k+2}}{\delta x^k}.$$

Under Assumption 1 and $\Delta \leq \overline{\Delta}$, it is easy to verify that $dw^{k+1}/dx^k < 0$.

Next, we show that x^{k-1} increases in x^k. Denote by g_j the partial derivative of function g with respect to its jth argument. Applying the implicit function theorem to condition $w^k = g(x^k, x^{k-1}, g(x^{k+1}, x^k, w^{k+2}))$ and rearranging terms, we get

$$\frac{dx^{k-1}}{dx^k} = -\frac{g_1(x^k, x^{k-1}, w^{k+1}) + g_3(x^k, x^{k-1}, w^{k+1}) g_2(x^{k+1}, x^k, w^{k+2})}{g_2(x^k, x^{k-1}, w^{k+1})}, \tag{1.48}$$

where

$$
\begin{aligned}
g_1(x^k, x^{k-1}, w^{k+1}) &= \frac{x^{k-1}-\Delta}{\delta x^{k-1}} \frac{f(x^k)}{F(x^{k-1})}(\Delta + c + \delta w^{k+1}), \\
g_2(x^k, x^{k-1}, w^{k+1}) &= \frac{\Delta}{\delta (x^{k-1})^2} \left(-c + \frac{F(x^k)}{F(x^{k-1})}(\Delta + c + \delta w^{k+1}) \right) \\
&\qquad - \frac{x^{k-1}-\Delta}{\delta x^{k-1}} \frac{F(x^k) f(x^{k-1})}{F^2(x^{k-1})}(\Delta + c + \delta w^{k+1}), \\
g_3(x^k, x^{k-1}, w^{k+1}) &= \frac{x^{k-1}-\Delta}{x^{k-1}} \frac{F(x^k)}{F(x^{k-1})}, \text{ and} \\
g_2(x^{k+1}, x^k, w^{k+2}) &= \frac{\Delta}{\delta (x^k)^2} \left(-c + \frac{F(x^{k+1})}{F(x^k)}(\Delta + c + \delta w^{k+2}) \right) \\
&\qquad - \frac{x^k-\Delta}{\delta x^k} \frac{F(x^{k+1}) f(x^k)}{F^2(x^k)}(\Delta + c + \delta w^{k+2}).
\end{aligned}
$$

Note that Identity (1.45) can be rearranged to form

$$
c + w^k \frac{\delta x^{k-1}}{x^{k-1}-\Delta} = \frac{F(x^k)}{F(x^{k-1})}(\Delta + c + \delta w^{k+1}),
$$

which we use to simplify g_2 and get

$$
g_2(x^k, x^{k-1}, w^{k+1}) = \frac{\Delta}{x^{k-1}} \frac{w^k}{x^{k-1}-\Delta} - \frac{f(x^{k-1})}{F(x^{k-1})} \left(c \frac{x^{k-1}-\Delta}{\delta x^{k-1}} + w^k \right)
$$

and

$$
g_2(x^{k+1}, x^k, w^{k+2}) = \frac{\Delta}{x^k} \frac{w^{k+1}}{x^k-\Delta} - \frac{f(x^k)}{F(x^k)} \left(c \frac{x^k-\Delta}{\delta x^k} + w^{k+1} \right).
$$

Finally, we substitute the partial derivatives into Equation (1.48) and collect terms to obtain

$$
\frac{dx^{k-1}}{dx^k} = - \frac{\frac{x^{k-1}-\Delta}{\delta x^{k-1}} \frac{\Delta}{x^k} \frac{1}{F(x^{k-1})} \left(f(x^k)(x^k + c) + \frac{\delta w^{k+1}}{x^k-\Delta} F(x^k) \right)}{\frac{\Delta}{x^{k-1}} \frac{w^k}{x^{k-1}-\Delta} - \frac{f(x^{k-1})}{F(x^{k-1})} \left(c \frac{x^{k-1}-\Delta}{\delta x^{k-1}} + w^k \right)}.
$$

The numerator is always positive and the denominator is negative when $\Delta \leq \overline{\Delta}$, so we have found that x^{k-1} increases in x^k. □

Lemma 10. *Under Assumption 1 and assuming that the values of function g's first and second arguments satisfy $x^k \in (a,b)$ and $x^{k-1} \in (a,b)$, g is strictly increasing in its first and third arguments and strictly decreasing in its second argument.*

Proof. Clearly g is increasing in its third argument, w^{k+1}, given that $x^{k-1} > a > \Delta$. In the proof of Lemma 9 we calculated g_1 and g_2. The first of these is strictly positive and the second, under Assumption 1, is strictly negative. □

Construction of bounds

We will construct two sequences, $\{\bar{x}^\ell\}_{\ell=0}^\infty$ and $\{\underline{x}^\ell\}_{\ell=0}^\infty$, that bound the type cutoffs of all alteration-proof equilibria. These sequences will be indexed in *reverse time* by integer ℓ, which counts the number of periods before period $L-1$ in any alteration-proof equilibrium.[25] The construction incorporates bounds on the w^k values. By construction, for any given alteration-proof equilibrium and its associated sequence $\{x^k, w^{k+1}\}_{k=1}^L$, it will be the case that $x^k \in [\underline{x}^{L-1-k}, \bar{x}^{L-1-k}]$ for each $k \in \{1,2,...,L-1\}$. There are two parts of the construction. The first constructs $\{\bar{x}^\ell\}_{\ell=0}^\infty$ and the second in similar fashion constructs $\{\underline{x}^\ell\}_{\ell=0}^\infty$.

Each of the two bounding sequences will be derived starting from an arbitrary alteration-proof equilibrium sequence $\{x^\tau, w^{\tau+1}\}_{\tau=1}^L$ under Assumption 1. We will adjust the sequence in recursive steps and then reverse the time index. So as not to overly complicate the presentation, we continue to call the adjusted sequence $\{x^\tau, w^{\tau+1}\}_{\tau=1}^L$ after every modification (that is, we redefine the sequence as the result of each adjustment), rather than create new notation to track the adjustments. It will turn out that the resulting bounds do not depend on the equilibrium that we started with.

The adjustments needed to construct $\{\bar{x}^\ell\}_{\ell=0}^\infty$ will utilize the following two operations, indexed by a given period k:

Operation 1: For a given sequence $\{x^\tau, w^{\tau+1}\}_{\tau=1}^L$ and integer $k \le L$ satisfying $w^k \le$

25. From Lemma 2 we know that every alteration-proof equilibrium has $x^L = a$, so the bounding sequences will be for the periods $L-1$ and earlier.

$g(x^k,x^{k-1},w^{k+1})$, hold fixed all values in the sequence except x^{k-1}. Raise the value of x^{k-1} to the point x' at which $w^k = g(x^k,x',w^{k+1})$, and then redefine $x^{k-1} \equiv x'$. If no such $x' \in [a,b]$ exists, then stop the procedure and set $x^{k-1} \equiv b$. Note that Lemma 10 ensures that x' is uniquely determined and weakly exceeds the starting value x^{k-1}.

Operation 2: For a given sequence $\{x^\tau, w^{\tau+1}\}_{\tau=1}^L$ and integer $k \leq L-1$ satisfying $w^\tau = g(x^\tau,x^{\tau-1},w^{\tau+1})$ for every $\tau \in \{k,k+1,\dots,L\}$, hold fixed all values in the sequence except x^{k-1}, x^k, and w^{k+1}. Raise the value of x^k to x' and simultaneously raise x^{k-1} to x'' and lower w^{k+1} to w', such that the system of Equations (1.47) is maintained, to the point at which $w' = \Delta/(1-\delta)$. Then redefine $x^k \equiv x'$, $x^{k-1} \equiv x''$, and $w^{k+1} \equiv w'$. If no such point exists (because x'' goes above b), then stop the procedure and set $x^{k-1} \equiv b$ and set x^k and w^{k+1} to the values that satisfy (1.47). Note that Lemma 9 ensures (x',x'',w') is uniquely determined, with x' and x'' weakly exceeding their starting values of x^k and x^{k-1}.

Here are the steps to construct $\{\bar{x}^\ell\}_{\ell=0}^\infty$. Take any alteration-proof equilibrium and let $\{x^\tau, w^{\tau+1}\}_{\tau=1}^L$ be its sequence of levels and w values. From Equation (1.45), we know that $w^\tau = g(x^\tau,x^{\tau-1},w^{\tau+1})$ for every $\tau \leq L$. From Lemmas 2 and 7, and because $(x-\Delta)/\delta x$ is increasing in $x > \Delta$, we know that $x^L = a$, $w^{L+1} \leq (\Delta/(1-\delta)) \cdot ((a-\Delta)/\delta a)$, and $w^\tau \geq \Delta/(1-\delta)$ for every τ. Our adjustments will eventually push w^τ down to exactly this lower bound, for each $\tau \leq L$.

The first step in the construction is to reset w^{L+1} to equal $(\Delta/(1-\delta)) \cdot ((a-\Delta)/\delta a)$, which causes $g(x^L,x^{L-1},w^{L+1})$ to rise. Then perform Operation 1 for $k = L$, which restores $w^L = g(x^L,x^{L-1},w^{L+1})$ but in raising x^{L-1} causes $g(x^{L-1},x^{L-2},w^L)$ to increase. Perform Operation 1 for $k = L-1$, restoring $w^{L-1} = g(x^{L-1},x^{L-2},w^L)$.

The second step is to apply Operations 1 and 2 recursively as follows, starting with $k' = L-1$. For any integer $k' \leq L-1$ satisfying $w^\tau = g(x^\tau,x^{\tau-1},w^{\tau+1})$ for every $\tau \in \{k',k'+1,\dots,L\}$, perform Operation 2 for $k = k'$, which results in $w^{k'+1} = \Delta/(1-\delta)$ and $w^{k'-1} < g(x^{k'-1},x^{k'-2},w^{k'})$, and then perform Operation 1 for $k = k'-1$, which restores $w^{k'-1} =$

$g(x^{k'-1}, x^{k'-2}, w^{k'})$. Decrease k by one and repeat this function until either operation triggers the process to stop. Let N denote the period at which the procedure stops, where $x^M = b$.

Note that Operations 1 and 2 adjust the cutoff-type sequence only by raising values of x^τ, so we have constructed upper bounds on the type cutoffs in any alteration-proof equilibrium. The procedure also results in $w^\tau = g(x^\tau, x^{\tau-1}, w^{\tau+1})$ and $w^\tau = \Delta/(1-\delta)$ for each $\tau \in \{N, N+1, \ldots, L\}$. Letting $\ell = L-1-\tau$, so we count backward in time, we thus have found a bounding sequence $\{\bar{x}^\ell\}_{\ell=0}^\infty$ defined recursively by $\bar{x}^0 = x^{L-1}$,

$$\frac{\Delta}{1-\delta} = \frac{\bar{x}^{\ell+1} - \Delta}{\delta \bar{x}^{\ell+1}} \left(-c + \frac{F(\bar{x}^\ell)}{F(\bar{x}^{\ell+1})} \left(\Delta + c + \delta \frac{\Delta}{1-\delta} \right) \right) \tag{1.49}$$

for each $\ell \in \{0, 1, \ldots, L-N-2\}$, and $\bar{x}^\ell = b$ for $\ell \geq L-N-1$. Note that Equation (1.49) comes from plugging $w^\tau = w^{\tau+1} = \Delta/(1-\delta)$ into $w^\tau = g(x^\tau, x^{\tau-1}, w^{\tau+1})$ and replacing τ with ℓ and replacing $\tau - 1$ with $\ell + 1$ to reverse the index direction. This equation gives $\bar{x}^{\ell+1}$ implicitly as a function of $\bar{x}^\ell$.

We construct $\{\underline{x}^\ell\}_{\ell=0}^\infty$ using the same steps but working in the opposite direction for the adjustments. Note that from Lemmas 2 and 7, we know that $x^L = a$ and $\Delta/(1-\delta) \leq w^\tau < (\Delta/(1-\delta)) \cdot ((x^{\tau-1} - \Delta)/\delta x^{\tau-1})$ for every τ. Let us define Operations 1R and 2R just as we did Operations 1 and 2 except with adjustments in the opposite direction. That is, Operation 1R begins with a sequence satisfying $w^k \geq g(x^k, x^{k-1}, w^{k+1})$ and lowers x^{k-1} to the point x' at which $w^k = g(x^k, x', w^{k+1})$. Operation 2R lowers x^k to x' and simultaneously lowers x^{k-1} to x'' and raises w^{k+1} to w', such that the system of Equations (1.47) is maintained, to the point at which $w' = (\Delta/(1-\delta)) \cdot ((x^{k-1} - \Delta)/\delta x^{k-1})$. It is not difficult to verify that both operations are well defined, yielding cutoff-type values satisfying $x^{k-1} \geq x^k$.

Starting with an arbitrary alteration-proof equilibrium sequence $\{x^\tau, w^{\tau+1}\}_{\tau=1}^L$, we first reset w^{L+1} down to equal $(\Delta/(1-\delta))$, which causes $g(x^L, x^{L-1}, w^{L+1})$ to fall. Then we perform Operation 1R for $k = L$ to restore $w^L = g(x^L, x^{L-1}, w^{L+1})$, and again for $k = L-1$ to restore $w^{L-1} = g(x^{L-1}, x^{L-2}, w^L)$. We proceed to the recursive step, applying Operations 1R and 2R, starting with $k' = L-1$ and ending when the boundary b is reached for $x^{\tau-1}$, at a period denoted by N. All operations adjust the cutoff-type sequence only by lowering values of x^{τ}, so we have constructed lower bounds on the type cutoffs in any alteration-proof equilibrium. The procedure also results in $w^{\tau} = g(x^{\tau}, x^{\tau-1}, w^{\tau+1})$ and $w^{\tau} = (\Delta/(1-\delta)) \cdot ((x^{\tau-1}-\Delta)/\delta x^{\tau-1})$ for each $\tau \in \{N, N+1, \ldots, L\}$. Letting $\ell = L-1-\tau$, we thus have found a bounding sequence $\{\underline{x}^{\ell}\}_{\ell=0}^{\infty}$ defined recursively by $\underline{x}^0 = x^{L-1}$,

$$\frac{\Delta}{1-\delta} = -c + \frac{F(\underline{x}^{\ell})}{F(\underline{x}^{\ell+1})}\left(\Delta + c + \delta\frac{\Delta}{1-\delta}\cdot\frac{\underline{x}^{\ell}-\Delta}{\delta\overline{x}^{\ell}}\right) \tag{1.50}$$

for each $\ell \in \{0, 1, \ldots, L-N-2\}$, and $\overline{x}^{\ell} = b$ for $\ell \geq L-N-1$. Equation (1.50) comes from plugging $w^{\tau} = (\Delta/(1-\delta)) \cdot ((x^{\tau-1}-\Delta)/\delta x^{\tau-1})$ and $w^{\tau+1} = \Delta/(1-\delta) \cdot ((x^{\tau}-\Delta)/\delta x^{\tau})$ into $w^{\tau} = g(x^{\tau}, x^{\tau-1}, w^{\tau+1})$ and replacing τ with ℓ and replacing $\tau-1$ with $\ell+1$. This equation gives $\underline{x}^{\ell+1}$ as a function of $\underline{x}^{\ell}$.

Convergence of bounding sequences

In summary, we have constructed sequences $\{\overline{x}^{\ell}\}_{\ell=0}^{\infty}$ and $\{\underline{x}^{\ell}\}_{\ell=0}^{\infty}$ that bound all alteration-proof equilibrium type-cutoff sequences. The final step of the proof is to apply a convergence result of Watson (2021) to show that as $\Delta \to 0^{+}$, the upper and lower bounds converge uniformly to the same continuous-time function defined by a differential equation. To be precise, let us make explicit the dependence of the bounding sequences on Δ by writing $\{\overline{x}^{\ell}(\Delta)\}_{\ell=0}^{\infty}$ and $\{\underline{x}^{\ell}(\Delta)\}_{\ell=0}^{\infty}$, and define step functions $\hat{\overline{x}}\colon [0,\infty) \times (0,\infty) \to [a,b]$ and $\hat{\underline{x}}\colon [0,\infty) \times (0,\infty) \to [a,b]$ by $\hat{\overline{x}}(t;\Delta) = \overline{x}^{[t/\Delta]}$ and $\hat{\underline{x}}(t;\Delta) = \underline{x}^{[t/\Delta]}$, where $[t/\Delta]$ denotes the largest integer that is weakly below t/Δ. As shown below, $\hat{\overline{x}}(\cdot;\Delta)$ and $\hat{\underline{x}}(\cdot;\Delta)$ converge, which implies the convergence of equilibrium type cutoffs stated in Theorem 3.

Lemma 11. *As $\Delta \to 0$, step functions $\hat{\overline{x}}(\cdot;\Delta)$ and $\hat{\underline{x}}(\cdot;\Delta)$ uniformly converge to the same function*

$z : [0,\infty) \to [0,b]$ *that solves this initial-value problem:*

$$\frac{dz}{dt} = \frac{F(z)}{f(z)} \frac{rz-1}{z(1+cr)}, \;\; z(0) = a. \tag{1.51}$$

Proof. We simplify Equations (1.49) and (1.50) by rearranging terms to obtain, respectively,

$$\frac{\Delta}{1-\delta} \frac{\delta \overline{x}^{\ell+1}}{\overline{x}^{\ell+1} - \Delta} + c = \frac{F(\overline{x}^{\ell})}{F(\overline{x}^{\ell+1})} \left(c + \frac{\Delta}{1-\delta} \right), \tag{1.52}$$

and

$$F(\underline{x}^{\ell+1}) = F(\underline{x}^{\ell}) \frac{\Delta + c + \frac{\Delta}{1-\delta} \cdot \frac{\underline{x}^{\ell} - \Delta}{\underline{x}^{\ell}}}{\frac{\Delta}{1-\delta} + c}. \tag{1.53}$$

Define the transition function $\overline{\sigma} \colon \; [a,b] \times (0,\overline{\Delta}) \to [a,b]$ so that, for every $x^{\ell} \in [a,b]$ and $\Delta \in (0,\overline{\Delta})$, $\overline{\sigma}(x^{\ell},\Delta)$ is the value of $x^{\ell+1}$ that solves (1.52). Likewise, define $\underline{\sigma} \colon \; [a,b] \times (0,\overline{\Delta})$ so that $\underline{x}^{\ell+1} = \underline{\sigma}(\underline{x}^{\ell},\Delta)$ solves (1.53).

We next extend the domain of functions $\overline{\sigma}$ and $\underline{\sigma}$ to $\mathbb{R} \times \mathbb{R}$. Regarding Δ, Expressions (1.52) and (1.53) are already well-defined for $\Delta \in \mathbb{R} \setminus \{0\}$ and the limits as $\Delta \to 0$ exist because $\lim_{\Delta \to 0} \Delta/(1-\delta) = 1/r$. So for $\Delta = 0$ we simply replace $\Delta/(1-\delta)$ with $1/r$ in (1.52) and (1.53), which extends $\overline{\sigma}$ and $\underline{\sigma}$ to $\Delta \in \mathbb{R}$. The extension to $x^{\ell} \in \mathbb{R}$ can be done arbitrarily.

Under Assumption 1, the extended functions $\overline{\sigma}$ and $\underline{\sigma}$ are twice continuously differentiable on $(a,b) \times \mathbb{R}$. Clearly, the initial states $\overline{x}^0(\Delta)$ and $\underline{x}^0(\Delta)$ converge to a. Note as well that $\overline{\sigma}(x,0) = x$ and $\underline{\sigma}(x,0) = x$ for all $x \in [a,b]$. Finally, the implicit function theorem applies to calculate $d\overline{\sigma}/d\Delta$ and $d\underline{\sigma}/d\Delta$, and these derivatives are bounded on a neighborhood of $\Delta = 0$ and for all $x^{\ell} \in [a,b]$. The properties just stated allow us to apply Theorem 2 of Watson (2021), which establishes that $\hat{\overline{x}}(\cdot;\Delta)$ and $\hat{\underline{x}}(\cdot;\Delta)$ uniformly converge to, respectively, functions $\overline{z} : [0,\infty) \to [a,b]$ and $\underline{z} : [0,\infty) \to [a,b]$ that solve initial-value problems given by

$$\frac{d\overline{z}}{dt} = \frac{d\overline{\sigma}}{d\Delta}(\overline{z},0), \quad \frac{d\underline{x}}{dt} = \frac{d\underline{\sigma}}{d\Delta}(\underline{z},0), \quad \text{and } \overline{z}(0) = \underline{z}(0) = a.$$

To complete the proof, we evaluate $d\overline{\sigma}(\bar{z},0)/d\Delta$ and $d\underline{\sigma}(\underline{z},0)/d\Delta$.

To derive $d\overline{\sigma}/d\Delta$, we apply implicit function theorem by differentiating both sides of Equation (1.52) with respect to Δ, which yields

$$\frac{1-\frac{\Delta}{1-\delta}r\delta}{1-\delta}\frac{\delta\bar{x}^{\ell+1}}{\bar{x}^{\ell+1}-\Delta}+\frac{\Delta}{1-\delta}\frac{(-r\delta\bar{x}^{\ell+1}+\delta\frac{d\overline{\sigma}}{d\Delta})(\bar{x}^{\ell+1}-\Delta)-\delta\bar{x}^{\ell+1}(\frac{d\overline{\sigma}}{d\Delta}-1)}{(\bar{x}^{\ell+1}-\Delta)^2}$$
$$=-\frac{F(\bar{x}^{\ell})f(\bar{x}^{\ell+1})}{F^2(\bar{x}^{\ell+1})}\left(c+\frac{\Delta}{1-\delta}\right)\frac{d\overline{\sigma}}{d\Delta}+\frac{F(\bar{x}^{\ell})}{F(\bar{x}^{\ell+1})}\frac{1-\frac{\Delta}{1-\delta}r\delta}{1-\delta}.$$

We solve for $d\overline{\sigma}/d\Delta$ and evaluate it at $\Delta=0$, replacing $\Delta/(1-\delta)$ with $1/r$ as required by the extension, and setting $\bar{x}^{\ell}=\bar{z}$ and $\bar{x}^{\ell+1}=\overline{\sigma}(\bar{z},0)=\bar{z}$. This yields

$$\frac{d\bar{z}}{dt}=\frac{d\overline{\sigma}}{d\Delta}(\bar{z},0)=\frac{F(\bar{z})}{f(\bar{z})}\frac{r\bar{z}-1}{\bar{z}(cr+1)}.$$

Similarly, we differentiate both sides of Equation (1.53) with respect to Δ,

$$\frac{d\underline{\sigma}}{d\Delta}=\frac{F(\underline{x}^{\ell})}{f(\underline{x}^{\ell+1})}\frac{(\frac{\Delta}{1-\delta}+c)\left(1+\frac{1-\frac{\Delta}{1-\delta}r\delta}{1-\delta}\frac{\underline{x}^{\ell+1}-\Delta}{\underline{x}^{\ell+1}}-\frac{\Delta}{1-\delta}\frac{1}{\underline{x}^{\ell+1}}\right)-\left(\Delta+c+\frac{\Delta}{1-\delta}\cdot\frac{\underline{x}^{\ell}-\Delta}{\underline{x}^{\ell}}\right)\frac{1-\frac{\Delta}{1-\delta}r\delta}{1-\delta}}{(\frac{\Delta}{1-\delta}+c)^2},$$

and evaluate it at $\Delta=0$, $\underline{x}^{\ell}=\underline{z}$, and $\underline{x}^{\ell+1}=\underline{\sigma}(\underline{z},0)=\underline{z}$. This yields

$$\frac{d\underline{z}}{dt}=\frac{d\underline{\sigma}}{d\Delta}(\underline{z},0)=\frac{F(\underline{z})}{f(\underline{z})}\frac{r\underline{z}-1}{\underline{z}(1+cr)}.$$

Clearly functions $\bar{z}$ and $\underline{z}$ are the same, identical to the function z described in the statement of the lemma. □

Note that the initial-value problem described in Lemma 11 is identical to that described for x in Theorem 3,

$$\frac{dx}{dt}=-\frac{F(x)}{f(x)}\frac{rx-1}{x(1+cr)},$$

except with the direction of time reversed, so we have proved the result with respect to the cutoff-type sequences. As shown in the text, the value T is derived by integrating the differential equation and solving for the time at which $x = b$.

To show that the equilibrium level sequences and continuation values for player 1 are also characterized as the theorem states, it is enough to observe that we can trivially rewrite the transition functions that define sequences $\{\bar{x}^{\ell}\}_{\ell=0}^{\infty}$ and $\{\underline{x}^{\ell}\}_{\ell=0}^{\infty}$ as vector-valued functions that include the level and player 1's continuation value. In any alteration-proof equilibrium, the transitions of the level and player 1's continuation value obey

$$\alpha^{k+1} = \frac{x^k - \Delta}{\delta x^k}\alpha^k \text{ and } v_1^{k+1} = w^{k+1}\alpha^k,$$

$\alpha^{L+1} = 1$, and $v_1^{L+1} = 1/r$. Corresponding to the lower-bound sequence $\{\underline{x}^{\ell}\}_{\ell=0}^{\infty}$ is a lower-bound sequence for α and an upper-bound sequence for w; likewise, the upper-bound sequence $\{\bar{x}^{\ell}\}_{\ell=0}^{\infty}$ corresponds to an upper-bound sequence for α and a lower-bound sequence for w. The convergence theorem of Watson (2021) applies to vector sequences. Thus, the characterization of the limit of level sequences and player 1's continuation values then follows from the characterization of the limit of type-cutoff sequences derived above.

1.7.4 Proof of Proposition 1

We first derive comparative statics of T. From Equation (1.26), we use the fact $\ln(z) \leq z - 1$ and get

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

and

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

Second, from Equation (1.26), comparative statics of $\alpha(0)$ is

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

Third, for the comparative statics of the slope of x for fixed $\chi\in[a,b]$ at time $\Gamma(\chi)$, we use Equation (1.22)

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

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 (1.25), 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,

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

1.8 Supplementary Appendix

In this appendix, we provide an additional existence result: an alteration-proof equilibrium that exhibits gradualism, trust, and cooperation by good types in every continuation (including after a deviation by player 1).

Theorem 4. *Under Assumption 1, the trust game has a constant-proportion equilibrium with parameter $\gamma = 1$ that specifies a trusting equilibrium in the continuation game following any history. In fact, after a deviation by player 1, within two periods equilibrium play coincides with a continuation on the original equilibrium path.*

Incidentally, in this equilibrium, when considering an alteration the players anticipate no further alterations in the future.

Proof of Theorem 4:

In reference to the constant-proportion equilibrium definition, consider the case of $\gamma = 1$. We use the same on equilibrium path sequence as in Appendix 1.7.2, but different off-equilibrium-path specifications.

In the equilibrium we now construct, for every history of play to the beginning of any period, player 1's updated belief about player 2's type will be given by the posterior of F conditioned on $x \leq \bar{x}$ for some number $\bar{x} \geq \Delta/(1-\delta)$. In other words, every continuation game from the start of any period will be defined by an upper-truncated type space. For a given number $\bar{x}$, let us call this continuation game the *$\bar{x}$-truncation continuation game*. Thus, we can fully describe player 1's equilibrium strategy by stating the level player 1 is prescribed to choose in the first period of the $\bar{x}$-truncation game, for every $\bar{x} \geq \Delta/(1-\delta)$. Likewise, player 2's equilibrium strategy will be fully described by stating the set of types that betray in the first period of the $\bar{x}$-truncation game after player 1's choice α in this period, for every $\bar{x} \geq \Delta/(1-\delta)$ and for every $\alpha \in [0,1]$.

Before describing the strategies, let us make a few notes. Recall that indifference Condition (1.35) means that type x^{k-1} is indifferent between betraying at level α^{k-1} in one period and waiting to betray at level α^k in the next period. Rearranging this equation yields

$$\alpha^{k-1} = \alpha^k \cdot \frac{x^{k-1}\delta}{x^{k-1} - \Delta}.$$

Now think about the level $\breve{\alpha}^k$ such that type x^k of player 2 would be indifferent between betraying at level $\breve{\alpha}^k$ in one period and waiting to betray at level α^k in the next period. This level is given by

$$\breve{\alpha}^k = \alpha^k \cdot \frac{x^k \delta}{x^k - \Delta},$$

and clearly $\breve{\alpha}^k \in (\alpha^{k-1}, \alpha^k)$. Observe that $\breve{\alpha}^k$ is increasing in k.

Here is the specification of strategies. Consider any $\bar{x}$-truncation continuation game and let ℓ be such that $\bar{x} \in (x^\ell, x^{\ell-1}]$. If $\bar{x} < x^{\ell-1}$ then player 1 is prescribed to choose $\alpha = \alpha^\ell$ in the current period. If $\bar{x} = x^{\ell-1}$ then player 1's specified behavior depends on whether player 1 deviated in the previous period. If player 1 did not deviate in the previous period then player 1 is supposed to choose $\alpha = \alpha^\ell$. If player 1 deviated in the previous period then player 1 is supposed to randomize between $\alpha = \alpha^\ell$ and $\alpha = \alpha^{\ell-1}$, with the probabilities described below.

For whatever level α' that is actually chosen by player 1, player 2's prescribed behavior is determined as follows. If $\alpha' \leq \alpha^{\ell-1}$ then all types above $\Delta\alpha^{\ell-1}/(\alpha^{\ell-1} - \delta\alpha^\ell)$ cooperate and types below betray. If $\alpha' > \alpha^{\ell-1}$ then find the integer ℓ' such that $\alpha' \in [\breve{\alpha}^{\ell'-1}, \breve{\alpha}^{\ell'})$ and $\alpha' \geq \alpha^{\ell-1}$. Player 2's action is then specified as follows:

If $\alpha' \in [\breve{\alpha}^{\ell'-1}, \alpha^{\ell'-1})$ then all types strictly greater than $x^{\ell'-1}$ betray and all types weakly below $x^{\ell'-1}$ cooperate. This is rational because player 1 in the following period will randomize between $\alpha^{\ell'}$ and $\alpha^{\ell'-1}$ with exactly the probabilities that make type $x^{\ell'-1}$ indifferent between betraying at level α' in the current period and waiting to betray in the next period. Note that in this case the continuation game from the next period is a truncation with cutoff $\bar{x}' \equiv x^{\ell'-1}$.

Let us calculate the probability p that player 1 must put on level $\alpha^{\ell'-1}$ in the next period to make type $x^{\ell'-1}$ indifferent. That $\alpha' \in [\breve{\alpha}^{\ell'-1}, \alpha^{\ell'-1})$ ensures that such a probability exists because, given the definition of $\breve{\alpha}^{\ell'-1}$, type $x^{\ell'-1}$ would strictly prefer to betray immediately if player 1 would choose $\alpha^{\ell'-1}$ in the next period, and would strictly prefer to wait if player 1 would choose $\alpha^{\ell'}$ in the next period. Type $x^{\ell'-1}$'s indifference condition is:

$$x^{\ell'-1}\alpha' = \Delta\alpha' + \delta x^{\ell'-1}(p\alpha^{\ell'-1} + (1-p)\alpha^{\ell'}),$$

which yields

$$p = \frac{\frac{x^{\ell'-1}-\Delta}{\delta x^{\ell'-1}}\alpha' - \alpha^{\ell'}}{\alpha^{\ell'-1} - \alpha^{\ell'}}. \tag{1.54}$$

If $\alpha' \in [\alpha^{\ell'-1}, \breve{\alpha}^{\ell'})$ then all types strictly greater than $\bar{x}' \equiv \alpha'\Delta/(\alpha' - \delta\alpha^{\ell'})$ betray and all types weakly below $\bar{x}'$ cooperate. This is rational because player 1 in the following period will choose $\alpha^{\ell'}$ for sure, making type $\bar{x}'$ indifferent between betraying at level α' in the current period and waiting to betray at level $\alpha^{\ell'}$ in the next period. Note that in this case the continuation game from the next period is a truncation with cutoff $\bar{x}' \in (x^{\ell}, x^{\ell-1}]$.

Denote $\hat{v}_1(\bar{x};\alpha')$ as player 1's continuation value in $\bar{x}-$truncation game, when player 1 chooses α', assuming players follow prescribed strategies after player 1 deviates. Therefore, player 1's continuation value is

$$\begin{aligned}\hat{v}_1(\bar{x};\alpha') = &\left(1 - \frac{F(\bar{x}')}{F(\bar{x})}\right)(-c\alpha') + \frac{F(\bar{x}')}{F(\bar{x})}\Delta\alpha' \\ &+ \delta\frac{F(\bar{x}')}{F(\bar{x})}\left(\left(1 - \frac{F(x^{\ell'})}{F(\bar{x}')}\right)(-c\alpha^{\ell'}) + \frac{F(x^{\ell'})}{F(\bar{x}')}\left(\Delta\alpha^{\ell'} + \delta v_1^{\ell'+1}\right)\right),\end{aligned} \tag{1.55}$$

for $\alpha' \in [\alpha^{\ell'-1}, \breve{\alpha}^{\ell'})$, and

$$\hat{v}_1(\bar{x};\alpha') = \left(1 - \frac{F(x^{\ell'-1})}{F(\bar{x})}\right)(-c\alpha') + \frac{F(x^{\ell'-1})}{F(\bar{x})}\left(\Delta\alpha' + \delta\frac{\Delta}{1-\delta}\alpha^{\ell'-1}\right), \tag{1.56}$$

for $\alpha' \in [\breve{\alpha}^{\ell'-1}, \alpha^{\ell'-1})$.

Lemma 12. *In an $\bar{x}$-truncation continuation game and given the strategy, player 1's continuation value from any deviation α' will be weakly lower than some alteration.*

Proof. We define ℓ such that $\bar{x} \in (x^{\ell}, x^{\ell-1}]$. Suppose player 1 deviates to $\alpha' \geq \alpha^{\ell-1}$, we first find the $\ell' \geq \ell$ such that $\alpha' \in [\breve{\alpha}^{\ell'-1}, \breve{\alpha}^{\ell'})$. Next according to the strategy, we discuss the following two cases: $\alpha' \geq \alpha^{\ell'-1}$ and $\alpha' < \alpha^{\ell'-1}$.

In the case of $\alpha' \in [\alpha^{\ell'-1}, \breve{\alpha}^{\ell'})$, all types strictly greater than $\bar{x}' = \alpha'\Delta/(\alpha' - \delta\alpha^{\ell'})$ betray and all types weakly below $\bar{x}'$ cooperate. By choosing α' in current period and $\alpha^{\ell'}$ in the following period, player 1's continuation value becomes Equation (1.55). We substitute $v_1^{\ell'+1} = \alpha^{\ell'}\Delta/(1-\delta)$, and $\alpha'/\alpha^{\ell'} = \delta\bar{x}'/(\bar{x}'-\Delta)$ into Equation (1.55) rearrange terms and get

$$\hat{v}_1(\bar{x};\alpha') = \delta\alpha^{\ell'}\left(c + \frac{\Delta}{1-\delta}\right)\frac{F(x^{\ell'})}{F(\bar{x})} + \left(\left(-c + (c+\Delta)\frac{F(\bar{x}')}{F(\bar{x})}\right)\frac{\delta\bar{x}'}{\bar{x}'-\Delta} - c\delta\frac{F(\bar{x}')}{F(\bar{x})}\right)\alpha^{\ell'}.$$

To find player 1's best deviation in this case, we differentiate $\hat{v}_1(\bar{x};\alpha')$ with respect to $\bar{x}'$ and get

$$\alpha^{\ell'}\frac{\delta\Delta}{(\bar{x}'-\Delta)^2}\left(c\frac{F(\bar{x}) - F(\bar{x}')}{F(\bar{x})} + \frac{f(\bar{x}')}{F(\bar{x})}(\bar{x}'+c)(\bar{x}'-\Delta) - \frac{F(\bar{x}')}{F(\bar{x})}\Delta\right),$$

which is positive for $\Delta \leq \overline{\Delta}$. Hence, we conclude that $\hat{v}_1(\bar{x};\alpha')$ increases in $\bar{x}'$. Further, because of $\bar{x}' = \alpha'\Delta/(\alpha' - \delta\alpha^{\ell'})$, this implies that $\hat{v}_1(\bar{x};\alpha')$ decreases in α' and the optimal deviation for $\alpha' \in [\alpha^{\ell'-1}, \breve{\alpha}^{\ell'})$ is $\alpha' = \alpha^{\ell'-1}$.

In the case of $\alpha' \in [\breve{\alpha}^{\ell'-1}, \alpha^{\ell'-1})$, all types weakly below $x^{\ell'-1}$ cooperate, and player 1 randomizes in the following period by putting probability $1-p$ on $\alpha^{\ell'}$ and probability p on $\alpha^{\ell'-1}$, with p given by Equation (1.54). One can verify that $\alpha' \in [\breve{\alpha}^{\ell'-1}, \alpha^{\ell'-1})$ implies $p \in (0,1]$.

Player 1 's continuation value from the α' deviation, Equation (1.56), simplifies to

$$\hat{v}_1(\bar{x};\alpha') = \alpha'\left(-c + \frac{F(x^{\ell'-1})}{F(\bar{x})}(c+\Delta)\right) + \delta\frac{F(x^{\ell'-1})}{F(\bar{x})}\frac{\Delta}{1-\delta}\alpha^{\ell'-1}.$$

Note that $\hat{v}_1(\bar{x};\alpha')$ is linear in α', allowing us to conclude that the best way to deviate within the interval $[\breve{\alpha}^{\ell'-1},\alpha^{\ell'-1}]$ is to set α' equal to one of the boundaries. Since the interval is open at the upper boundary, we are using the fact that player 1's continuation value is continuous there. To see this, let us look at the lower boundary and the interval below it. At $\alpha' = \breve{\alpha}^{\ell'-1}$, the continuation value is

$$\hat{v}_1(\bar{x};\alpha^{\ell'-1}) = \alpha^{\ell'-1}\frac{x^{\ell'-1}\delta}{x^{\ell'-1}-\Delta}\left(-c + \frac{F(x^{\ell'-1})}{F(\bar{x})}\left(c+\Delta+\frac{\Delta}{1-\delta}\frac{x^{\ell'-1}-\Delta}{x^{\ell'-1}}\right)\right),$$

which is the same as

$$\lim_{\alpha'\to\breve{\alpha}^{\ell'-1}} \hat{v}_1(\bar{x};\alpha' \in [\alpha^{\ell'-2},\breve{\alpha}^{\ell'-1})).$$

Recall in the case of $\alpha' \in [\alpha^{\ell'-2},\breve{\alpha}^{\ell'-1})$, we have proved that $\hat{v}_1(\bar{x};\alpha')$ monotonically decreases in α', so we conclude that comparing to $\alpha' = \breve{\alpha}^{\ell'-1}$, player 1 is able to obtain a higher payoff by choosing $\alpha' = \alpha^{\ell'-2}$. Therefore, combing the two cases, we have that player 1's optimal deviation in the interval $[\alpha^{\ell'-1},\alpha^{\ell'-2}]$ is one of the endpoints $\{\alpha^{\ell'-1}$ and $\alpha^{\ell'-2}\}$. Finally, because deviations with $\alpha' = \alpha^{\ell'-1}$ and $\alpha^{\ell'-2}$ are equivalent to alterations with $(\ell,\ell'-\ell-1,\alpha^{\ell'-1})$ and $(\ell,\ell'-\ell-2,\alpha^{\ell'-2})$ respectively, we conclude that player 1 has no incentive to deviate in an alteration-proof PBE. □

Lemma 12 implies that if player 1 has no incentive to alter the game, then she also has no incentive to deviate. It remains to show that the PBE with the specified strategy is alteration-proof. However, as the on equilibrium path outcome for the this equilibrium is a special case of the alteration-proof PBE in Appendix 1.7.2, we apply Lemma 4 and 6 with $\gamma = 1$ to conclude that the prescribed strategy constitutes an alteration-proof PBE.

Acknowledgements

Chapter 1 is coauthored with Joel Watson and has been published in Journal of Economic Theory 204 (September 2022). The dissertation author was a coauthor of this chapter.

Chapter 2

Reputation Spillover of Venture Capitalists

2.1 Introduction

Bernstein, Mehta, Townsend, and Xu (2022) show that job seekers show more interest in reputable venture capitalist (VC)-backed startups, particularly in early-stage startups, after controlling for the level of financing raised by the startup. This suggests passive value creation attributed to VCs through their association with these firms. However, the mechanisms and implications of this reputation spillover remain unclear, primarily due to the limitations of traditional economic models that overlook the interplay between capital and labor. The current paper considers a fundamental question: What drives the mechanism behind VC reputation spillover, and what are its multifaceted implications? Moreover, how does VC reputation depend on the startup's production technology and labor input?

To explore this question, this paper employs a dynamic matching model with transfers between VCs and workers. VCs are endowed with heterogeneous abilities, which are persistent and unknown, and their reputations are modeled as public beliefs of their ability. Over time, VCs invest in startups, and the outcomes of these startups—whether they succeed or fail—generate noisy signals about the VC's ability, which, in turn, affect the VC's reputation. Startups hire short-lived workers from a pool of workers with heterogeneous levels of skill. Workers derive payoffs from their immediate wages, which are negotiated during their employment, as well as from the potential for future career benefits, which is assumed to increase with the VC's reputation. In the long term, the distribution of VC reputations converges to a steady state. Further, the learning speed, steady-state reputation distribution, and wage contracts are all contingent on the degree of labor-VC complementarity in startup success probability.

My research makes several contributions to the entrepreneurial finance literature. Firstly, it offers a formal framework for understanding VC reputation, addressing how VC reputation accumulates over time and is distributed among various VCs in the market. While the literature has widely acknowledged the positive impact of reputable VCs on startups, there has been a lack of clarity regarding what precisely constitutes VC reputation, especially in distinguishing between their true underlying ability and reputation based on experience. By conceptualizing a VC's reputation as the public beliefs about its hidden abilities, this model aligns with the empirical findings that reputable VC-backed startups have better short and long-run performance (Nahata (2008), and Krishnan et al. (2011)). Moreover, it accommodates the idea of imperfect information about VC abilities, as suggested by Hochberg, Ljungqvist, and Vissing-Jørgensen (2014).

Secondly, the theoretical analysis offers fresh insights into empirical observations. Due to the trade-off between a worker's current wage and the future career benefits brought about by working experience at reputable VC-backed startups, workers are often willing to accept lower wages. On one hand, this sheds light on the previously observed phenomenon of reputable VCs' substantial "bargaining power" (Gompers and Lerner (1999), and Hsu (2004)) and their tendency to request less downside protection in contracts with startups (Bengtsson and Sensoy (2011)): The reputation spillover mechanism enables VCs to create and retain more surplus, and the lower cost of labor decreases VC's losses when a startup fails. On the other hand, it unveils a novel driver of VC's return persistence (Kaplan and Schoar (2005)): Past high investment returns lead to enhanced reputation, subsequently resulting in improved labor inputs and reduced labor costs.

Finally, the model elucidates the link between production technology and the VC market. In the early stages, there is a strong complementarity between labor and VC inputs in determining a startup's probability of success, leading to faster learning when VC reputation is high. The model predicts a positively skewed steady-state VC distribution, with a heavy concentration of VCs having low reputations. Since fewer VCs can accumulate a high reputation, the reputation spillover effect becomes more prevalent. Consequently, the model yields novel predictions that the VC market, specializing in startups with a strong labor-VC complementarity, exhibits a positively skewed distribution with a high concentration towards the lower end of the reputation space.

The paper also contributes to the economic theory literature by exploring the interplay between learning and matching. The model captures a realistic feature that signal structure is contingent on the current matching outcome, and belief updates following each additional signal influence the next matching, creating a feedback loop that poses challenges in an infinite horizon matching model. To tackle this challenge, the analysis utilizes the Kullback-Leibler divergence as a metric for assessing the informativeness of each additional signal, thereby formalizing the speed of learning when signals depend on matching outcomes.

The structure of this paper is as follows: Section 2.2 provides an overview of the model, and Section 2.3 offers preliminary analysis and an intuitive example of the model. Section 2.4 delves into the analysis of when learning interacts with matching. In Section 2.5, we present numerical simulations and comparative static analyses. Section 2.6 discusses the model's extensions and its additional links to existing literature, and Section 2.7 concludes.

2.2 Model

Consider a dynamic matching model among three sets of players in a continuum: venture capitalists (VCs), workers, and startups. In each period $t \in \{1, 2, 3, ...\}$, a VC and a worker are matched and run a startup, which generates an uncertain payoff. VCs are long-run players and have persistent types, which are unknown to all the players. Workers and startups are short-run players, but observe VCs' historical performance.

2.2.1 Players

Venture capitalists

The set of VCs $\mathscr{I}$ has a measure of 1, and each VC can invest in at most one startup. Each VC $i \in \mathscr{I}$ is endowed with an unknown binary type x^i, which measures the VC's managerial ability and non-monetary resources that add value to startups. The type of each VC is independently drawn from a common prior distribution, such that a VC is of high type $(x^i = 1)$ with probability $\theta_0 \in (0,1)$ and low type $(x^i = 0)$ with probability $1-\theta_0$. Define VC i's reputation at the end of period $t \in \{1,2,...\}$ as the public belief of the probability that $x^i = 1$, denoted by $\theta_t^i \in \Theta \equiv [0,1]$. Further, denote $\mathscr{Q}_t$ as the probability measure of VCs' reputation Θ at the end of period t; denote the set of all VCs' reputation at the end of period t by $\Theta_t \equiv \{\theta_t^i\}_{i\in\mathscr{I}}$. We make the following technical assumption on the reputation measure:

Assumption 2. *Set* Θ_0 *is atomic and measure* $\mathscr{Q}_0$ *is discrete.*

Each VC $i \in \mathscr{I}$ lives for multiple periods until it exists exogenously, after which it is replaced with a new VC whose type is drawn from the unconditional distribution. Define VC turnover rate $q \in [0,1]$ as the probability that VC i is replaced by a new entrant in period t, for all $i \in \mathscr{I}$ and $t \in \{1,2,...\}$. Denote the age of VC i in period t by τ_t^i. If VC i is replaced in period t, then $\tau_t^i = 0$, and otherwise $\tau_t^i = \tau_{t-1}^i + 1$. Denote $\mathscr{T}_t : \mathbb{N} \to [0,1]$ as the cross-sectional measure of VC ages in period t. The interpretation of VCs in this model aligns with individual general partners rather than venture capital firms. As documented in Gompers et al. (2010) and Ewens and Rhodes-Kropf (2015), individual partners have a greater impact on target startups than VC firms do. Turning over on the individual level evolves with idiosyncrasies, so each partner's entry and exit become exogenous shocks to VC firm's true ability in aggregation.[1]

1. Different from empirical studies that use VC's age and number of past target startups as a proxy of reputation, such as Gompers and Lerner (1999) and Hochberg, Ljungqvist, and Lu (2007), the current model assumes VC's age to be an exogenous process. This simplifying assumption allows us to identify the feedback effect of reputation spillover and characterize of steady state distribution of VC reputation. As in Section 2.4, even without endogenous exits, VCs' reputations are priced in, which causes VC's return persistence.

In each period, VC i invests 1 unit of capital in a target startup and collects investment returns. Further, because each period of the model corresponds to one investment cycle from deal sorting to existing, a VC's age corresponds to the number of past investments. Hence, the unit of VC's age is not in years, but as a multiple of around 10 years. The baseline model assumes VC to be myopic in the sense that VC maximizes expected payoff and does not trade off current period return for higher continuation value. However, this assumption is not unrealistic, not only because the period length is sufficiently high in the model, but also because VCs are obligated in their relation with limited partners to maximize return from each investment.

Workers

In each period, a measure of 1 of workers intend to work at startups, and each worker can work for at most one startup. Assume all workers exit the startup labor market after one period of working experience, after which an identical measure of workers enter the startup labor market such that the labor market is stable over time. Each worker is endowed with an observable productivity $\ell \in [0,1]$; denote $\mathscr{P}$ as the probability measure of worker's productivity on the set of workers $\mathscr{L} \equiv [0,1]$. The model assumes that workers are heterogeneous only in ℓ and that workers do not have liquidity constraints or cost of effort. We make the following technical assumption on the worker space:

Assumption 3. *The set $\mathscr{L}$ is non-atomic and measure $\mathscr{P}$ is absolutely continuous with respect to the Lebesgue measure. The density of measure $\mathscr{P}$, denoted by function p, is bounded away from 0; that is, there exists $\varepsilon > 0$, such that $p(\ell) \geq \varepsilon$ for all $\ell \in \mathscr{L}$.*

By working in a startup, a worker's payoff includes two parts: salaries, determined in an equilibrium wage contract, and future lifetime earnings, assumed to increase the reputation of the startup's investor. This assumption captures startup job seekers' belief that working at a reputable VC-backed startup provides a better resume value and improves their future careers. Define function $v : \mathscr{L} \times \Theta \to [0,\infty)$ as the present value of worker's future lifetime earnings, such that worker ℓ gains a quasi-linear utility of $w(\ell,\theta)+v(\ell,\theta)$ from working in startup backed by VC of reputation θ with wage contract $w:\mathscr{L}\times\Theta\to\mathbb{R}$.

The function v can be interpreted as the workers discounted continuation payoff. The model assumes that workers leave the startup labor market in one period, after which workers could still benefit from VC's reputation spillover while seeking their next job in an established firm, entering into a management role, exercising equity options, or starting their own business. Despite that workers' future career decision is not in the current model, the startup job experience has a non-negligible externality on the worker's payoffs, and v is the reduced-form model of this externality. Explicit modeling of startup workers' career path includes Coles and Mortensen (2016) and Dinlersoz, Hyatt, and Janicki (2019).[2]

Startups

The set of startups has a measure strictly greater than 1, and each startup needs one unit of capital and one labor input to produce the numéraire. Startups only live for one period and are homogeneous. The direct implication of this assumption is that startups only facilitate matching between VCs and workers without sharing any surplus. Therefore, the model focuses on the two-sided matching between VCs and workers and assumes that the surplus created from matching is shared between VCs and workers and that startups do not retain any economic profit in equilibrium.

If a startup is invested by a VC of type $x\in\{0,1\}$ and hires a worker of productivity $\ell\in[0,1]$, then the startup will produce a positive output $R(\ell)$ (succeed) with probability $y_x(\ell)$ and a zero output (fail) with complementary probability, where $R:\mathscr{L}\to(0,\infty)$ and $y_x:\mathscr{L}\to(0,1)$ for $x\in\{0,1\}$ are deterministic functions of ℓ.

2. The assumption of worker's future career benefit is motivated by the experiment in Bernstein et al. (2022), but the career benefit has more foundation in the labor economics literature, such as Terviö (2009). In the current model, the career benefit as a worker's continuation value not only includes the externalities that VC's reputation creates for a worker's career, but also could contain the discounted future payoffs from the startup's stock options. Because the model simplified the term sheet and compensation contracts, we do not distinguish different sources of worker's continuation value. However, one could experimentally test VC's reputation impact on workers' sensitivity to wages and stock options to further identify the pure resume value caused by reputation spillover.

For $x \in \{0,1\}$, denote $Y_x : \mathscr{L} \to [0,\infty)$ as output of a startup with worker ℓ and VC x, that is,

$$Y_x(\ell) = \begin{cases} R(\ell), & \text{with probability } y_x(\ell). \\ 0, & \text{with probability } 1 - y_x(\ell). \end{cases}$$

Recall that VC's type x is unknown to everyone, so matching and contracting are based upon expected output. Define two functions $y : \mathscr{L} \times \Theta \to (0,1)$ and $Y : \mathscr{L} \times \Theta \to [0,\infty)$ as the expected probability of positive output and expected output level conditional on VC's reputation. That is, for $\ell \in \mathscr{L}$ and $\theta \in \Theta$,

$$y(\ell,\theta) = \theta y_1(\ell) + (1-\theta) y_0(\ell), \text{ and } Y(\ell,\theta) = R(\ell) y(\ell,\theta).$$

A startup pays wage w to its worker, as a fixed cost of production, and pays rents on capital out of its output Y_x. When a startup fails, VC bears the cost of hiring the worker and loses $w(\ell,\theta)$; when the startup succeeds, VC collects $R(\ell) - w(\ell,\theta)$, all production output net of wage payment. Denote $r : \mathscr{L} \times \Theta \to \mathbb{R}$ as VC's expected return, then

$$r(\ell,\theta) = Y(\ell,\theta) - w(\ell,\theta).$$

Definition 7. *Under Assumptions 2 and 3, define functions $r : \mathscr{L} \times \Theta \to \mathbb{R}$ and $u : \mathscr{L} \times \Theta \to \mathbb{R}$ as VC's and worker's payoff; define function $f : \mathscr{L} \to \Theta \cup \{\emptyset\}$. Matching is (r,u,f) such that for all $\ell \in \mathscr{L}$,*

i *if $f(\ell) \neq \emptyset$, $Y(\ell, f(\ell)) + v(\ell, f(\ell)) = r(\ell, f(\ell)) + u(\ell, f(\ell))$,*

ii *if $f(\ell) = \emptyset$ then $u(\ell,\emptyset) = 0$, and if $f^{-1}(\theta) = \emptyset$ then $r(\emptyset,\theta) = 0$, and*

iii $\mathscr{Q}(\theta) = \mathscr{P}(\{\ell : f(\ell) = \theta\})$

For given functions Y and v, matching (r,u,f) is stable if for all ℓ, $r(\ell, f(\ell)) \geq 0$, $u(\ell, f(\ell)) \geq 0$, and $Y(\ell, f(\ell)) + v(\ell, f(\ell)) \leq r(\ell, f(\ell)) + u(\ell, f(\ell))$.

This concept was first introduced by Shapley and Shubik (1969) and is equivalent to a competitive equilibrium with a worker's equilibrium wage clearing the market. We make the following assumption on the total value created from matching $Z \equiv Y + v$. This assumption makes sure that the expected total value created increases in VC's reputation and worker's productivity, as well as the complementarity between the two inputs.

Assumption 4. *Functions y and v are non-decreasing and differentiable almost everywhere in both arguments, and Z is supermodular. For $x \in \{0,1\}$, there exist uniform bounds $\underline{y}_x$, $\bar{y}_x$, $\underline{y}'_x$ and $\bar{y}'_x$ such that for all $\ell \in [0,1]$, $\underline{y}_x < y_x(\ell) < \bar{y}_x$ and $\underline{y}'_x < dy_x(\ell)/d\ell < \bar{y}'_x$. Further, $\underline{y}_0 > 0$, $\bar{y}_0 < \underline{y}_1$, and $\bar{y}_1 < 1$.*

We make additional assumptions on the tie-breaking rule:

Assumption 5. *When a worker is indifferent between VCs with two different reputations, the worker works for the VC with a higher reputation. When VCs with the same reputation θ are indifferent among workers in set L, then they match randomly; in this case, matching generates values Y and v from inputs θ and the expected worker's productivity $\int_L \ell d\mathscr{P}$.*

This tie-breaking assumption models a scenario where the market can observe matches between VCs with a specific reputation level θ and workers in the set L, but it does not observe the exact match between a particular VC i and a particular worker ℓ. The interpretation of Assumption 5 is that that on average, the workers matched with VCs of the same reputation contribute at the level of the expected worker's type, and the value created from these matches is determined by the average type of the worker who is matched with VCs of the same reputation.

2.2.2 Timeline

Define h_t as the public history of the game in period t, and $H_t = \cup_{\tau=0}^{t}\{h_\tau\}$ as the public history of the game until period t. At the beginning of the game, a mass 1 of VCs is born with types drawn from the prior distribution, and all VCs have reputation θ_0. In each period $t \in \{1,2,...\}$, the game has three stages: turnover, matching, and production.

Stage 1: turnover

At the beginning of each period, workers and startups are born, and a mass q of existing VCs are replaced with new entrants, whose types are independently drawn from the prior distribution. In the current period, if VC i was not replaced with a new entrant, its reputation carries over from the end of the previous period θ^i_{t-1}, and its age increases by 1 to $\tau^i_t = \tau^i_{t-1}+1$. If VC i is a new entrants, then $\theta^i_{t-1} = \theta_0$ and $\tau^i_t = 0$.

Stage 2: matching

Each startup tries to match with one VC and one worker. We select the VC-preferred stable match, and the total value selects a VC-preferred stable matching $Z(\ell,\theta)$ is shared between VCs and workers with a wage transfer $w(\ell,\theta)$. Unmatched players get a null contract. Matched pair of VC i and worker ℓ agree on the contract of sharing expected surplus: VC pays $w(\ell,\theta^i_{t-1})$ to worker regardless of startup output, and VC takes all risks and returns of the startup.

Stage 3: production

Unmatched VCs, workers, and startups each obtain a payoff of zero. Startup matched with worker ℓ and VC i produces with technology $Y_{x^i}(\ell)$. Worker's payoff is $w(\ell,\theta^i_{t-1})+v(\ell,\theta^i_{t-1})$, and VC's payoff is $Y^t_{x^i}(\ell)-w(\ell,\theta^i_{t-1})$. Period t history includes $h_t = (\ell,\theta^i_{t-1},Y^t_{x^i})$, and VC's end of period reputation updates by Bayes' rule $\theta^i_t = \mathbb{E}[x^i|H_{t-1}\cup h_t]$.

If realized period t revenue $Y_{x^i}(\ell) > 0$,

$$\theta^i_t = \frac{\theta^i_{t-1}y_1(\ell)}{\theta^i_{t-1}y_1(\ell)+(1-\theta^i_{t-1})y_0(\ell)},$$

and if $Y_{x^i}(\ell) = 0$,

$$\theta^i_t = \frac{\theta^i_{t-1}(1-y_1(\ell))}{\theta^i_{t-1}(1-y_1(\ell))+(1-\theta^i_{t-1})(1-y_0(\ell))}.$$

2.3 Preliminary Analysis

We start by noticing that within each period, because of Assumption 4, the complementarity between worker's productivity ℓ and VC's reputation θ results in positive assortative matching (Becker (1973)). For example, VCs of the top 10% reputation are matched with the workers of the top 10% productivity. This provides the theory foundation of the VC reputation spillover phenomenon as Bernstein et al. (2022) document. However, it remains unclear how the top 10% VCs are determined in each period.

To understand how VCs build reputation, we simplify the Bayesian belief updating rule and obtain the dynamics of reputation:

$$\frac{\theta_t^i}{1-\theta_t^i} = \mu(\ell) \cdot \frac{\theta_{t-1}^i}{1-\theta_{t-1}^i}, \tag{2.1}$$

for all $i \in \mathscr{I}$ and $t \in \{1, 2, ...\}$, where μ is defined as the likelihood ratio by

$$\mu(\ell) \equiv \begin{cases} y_1(\ell)/y_0(\ell), & \text{if } Y_{x^i}(\ell) > 0, \\ (1-y_1(\ell))/(1-y_0(\ell)), & \text{if } Y_{x^i}(\ell) = 0. \end{cases}$$

Note by Assumption 4, $y_1(\ell) > y_0(\ell)$, so after observing a successful startup $Y_{x^i}(\ell) > 0$, everyone is more optimistic about VC's ability: $\theta_t^i > \theta_{t-1}^i$ with $\mu(\ell) > 1$. Similarly, bad news ($Y_{x^i}(\ell) = 0$) lead to a decrease in VC's reputation with $\mu(\ell) < 1$.

A notable feature of this model is that the belief updating process is contingent on whom the VC is matched with. The dependence of μ on ℓ, as well as positive assortative matching, leads to a novel reputation process. To see this, consider a VC with a reputation of 0.8 under two Scenarios:

1. *If* 10% *VCs have a reputation of* 0.8 *and* 90% *VCs have a reputation of* 0.2*, then the VC will be matched with one of the top* 10% *workers.*

2. *If* 10% *VCs have a reputation of* 0.2 *and* 90% *VCs have a reputation of* 0.8*, then the VC could be matched anyone but the bottom* 10% *workers.*

In this example, the VC with the same prior reputation 0.8 could be matched with workers with significantly different qualities, which also leads to different likelihoods of success $y_x(\ell)$. Intuitively, one VC's own reputation is meaningful only in the context of the cross-sectional distribution of all VCs in the market. Mathematically, the interaction between matching and learning implies that the state variables for VC i in period t include both its current reputation θ_{t-1}^i and the distribution of reputation of all VCs $\mathcal{Q}_t$.

Moreover, the model also speaks to the equilibrium speed of learning. In the example, even when the startup backed by VC with reputation 0.8 succeeds (or fails) in both Scenarios, $\mu(\ell)$ makes the VC's posterior reputation diverge substantially. With the complementarity of x and ℓ in probability $y_x(\ell)$, each signal embeds some information about x, which is sensitive to labor input.

Heuristically, after differentiating μ with respect to ℓ, we could observe that the monotonicity of μ in ℓ depends on how $(dy_1/d\ell)/(dy_0/d\ell)$ is related to the likelihood ratios y_1/y_0 and $(1-y_1)/(1-y_0)$. With strong complementarity (high $dy_1/d\ell - dy_0/d\ell$), bad news is "more informative" in Scenario 1 than in Scenario 2: with the availability of top 10% workers, bad news is more likely due to VC's incapability of complementing value creation. In Section 2.4, we will formalize this intuition with different cases of complementarity, before which we define the informativeness of a signal as follows.

We measure the value of additional information by Kullback-Leibler divergence. Specifically, for period $t-1$ posterior distribution, or period t prior distribution $Prob(x=1|H_{t-1}) = \theta_{t-1}$ and period t posterior distribution $Prob(x=1|H_t) = \theta_t$ defined on type space $x \in \{0,1\}$, the relative entropy from posterior to prior in period t is defined to be

$$D_{KL}(\theta_t||\theta_{t-1}) \equiv \theta_t \log \frac{\theta_t}{\theta_{t-1}} + (1-\theta_t) \log \frac{1-\theta_t}{1-\theta_{t-1}}.$$

Divergence $D_{KL}(\theta_t||\theta_{t-1})$ measures the informativeness of period t signal, and learning is faster when $D_{KL}(\theta_t||\theta_{t-1})$ is higher. Next, we omit the dependence in t by using θ' to represent posterior belief when prior belief is θ. After some algebra in Appendix 2.8.1

$$D_{KL}(\theta'||\theta) = \frac{\theta\mu(\ell)}{1-\theta(1-\mu(\ell))}\log\mu(\ell) - \log(1-\theta(1-\mu(\ell))). \tag{2.2}$$

Another feature of the model is the exogenous turnover of VCs. Because turnover is independent among VCs and is independent of matching and learning, we characterize VC's age distribution $\mathscr{T}$ before the main analysis. Recall that $\mathscr{T}_t(\tau)$ represents the measure of VCs with the age of τ, for all $\tau \in \{0,1,...,t\}$. Appendix 2.8.1 shows its proof.

Lemma 13. *As $t \to \infty$, $\mathscr{T}_t$ converges in distribution to $\mathscr{T}$, and $\mathscr{T}(\tau) = q(1-q)^\tau$.*

2.4 Matching Interacts with Learning

In this section, we first analyze the learning benchmark, where startups generate identically distributed signals each period, and then for a given distribution of VC reputation in each period, we characterize the stable matching and contracts. Finally, we will allow the matching outcome to influence the signals and show how it interferes with learning of VC's ability.

2.4.1 Learning benchmark: independent signal

When $y_x(\ell)$ does not depend on ℓ, learning of VC's hidden type x is time-invariant. In this subsection, we omit the dependence of ℓ by writing y_x as probability of $Y_x = R(\ell)$ for $x \in \{0,1\}$. We begin with the following lemma that explores the properties of speed of learning.

Lemma 14. *Divergence $D_{KL}(\theta'||\theta)$ achieves maximum at*

$$\theta^* = \frac{\mu(\log\mu)/(\mu-1)-1}{\mu-1}. \tag{2.3}$$

$\lim_{\theta\to 0} D_{KL}(\theta'||\theta) = 0$ *and* $\lim_{\theta\to 1} D_{KL}(\theta'||\theta) = 0$. *Further, θ^* decreases in μ and $\theta^* \to 0.5$ as $\mu \to 1$.*

Proofs of all results in this subsection are presented in Appendix 2.8.2. This result echos the intuitive finding that a successful startup is more informative than a failed startup when reputation is low: upon good news, $\mu = y_1/y_0 > 1$, $\theta^* < 0.5$, and upon bad news, $\mu = (1-y_1)/(1-y_0) < 1$, $\theta^* > 0.5$. Further, learning is faster when uncertainty about VC's ability is still high, whereas learning is slower when prior belief already contains most of the information about VC's ability. As the next lemma shows, VCs' reputation reveals their ability with probability 1 for infinitely lived VCs.

Lemma 15. *For infinitely lived VC i with $x^i \in \{0,1\}$, $\lim_{t\to\infty} Prob(|\theta_t^i - x^i| < \varepsilon) = 1$, for all $\varepsilon > 0$.*

Mathematically, when signals are independent and identically distributed in each period, the Bayesian posterior is consistent, which is a standard result in Bayesian learning. Lemma 15 implies that in the benchmark case, the market would learn about VC's abilities perfectly if VC could maintain a history of infinitely many investments. Combined with Lemma 14, this implies that if the turnover rate is 0, then in the long run, VCs with different x are perfectly separated and each new startup investment does not contain any additional information about VC's ability. After allowing VCs to turnover with probability $q \in (0,1)$, the model produces a more realistic reputation process and long-run distribution of VCs, as presented in Lemma 16 and 17.

Lemma 16. *Consider VC i in period t with age τ. If n out of τ investments are successful with $Y_x > 0$ and the remaining are unsuccessful with $Y_x = 0$, then the VC's reputation is $\theta_t^i = \theta_t^i(\tau, n)$, where*

$$\theta_t^i(\tau,n) \equiv \left(1 + \frac{1-\theta_0}{\theta_0}\left(\frac{y_0}{y_1}\right)^n \left(\frac{1-y_0}{1-y_1}\right)^{\tau-n}\right)^{-1}. \tag{2.4}$$

Further, this happens with probability

$$p(\tau,n) \equiv \begin{cases} \binom{\tau}{n} y_x^n (1-y_x)^{\tau-n}, & \text{if } n \le \tau; \\ 0, & \text{otherwise}. \end{cases}$$

Lemma 16 characterizes the process of VC's reputation in closed form. Because y_x is time-invariant, the state of the reputation process is now only the number of past investments and the number of successful investments. Moreover, from a recursive substitution of Equation (2.1), $\{\theta_t^i\}$ process can be transformed into a Binomial process. Note that the right side of Equation (2.4) is not contingent on the values of i or t, which allows a closed-form characterization of the long-run distribution of reputations: a measure of VCs with same reputation $\theta_t^i(\tau,n)$ is the accumulative measure of VCs with same success proportion τ/n among all possible ages $\tau \leq t$.

Lemma 17. *Period t distribution of VC reputation $\mathscr{Q}_t$ is*

$$\mathscr{Q}_t(\theta = \theta_t^i(\tau,n)) = \sum_{m=0}^{n} q(1-q)^{\tilde{\tau}(m)} p(\tilde{\tau}(m),m)$$

for $\tau \in \{0,1,...,t\}$ and $n \in \{0,1,...,\tau\}$, where

$$\tilde{\tau}(m) \equiv \begin{cases} \tau + (n-m)\left(\frac{\log\left(\frac{y_0}{y_1}\right)}{\log\left(\frac{1-y_0}{1-y_1}\right)} - 1\right), & \textit{if } (n-m)\left(\frac{\log\left(\frac{y_0}{y_1}\right)}{\log\left(\frac{1-y_0}{1-y_1}\right)} - 1\right) \textit{ is integer;} \\ 0, & \textit{otherwise.} \end{cases}$$

In period t, there are at most $\sum_{\tau=0}^{t}\sum_{n=0}^{\tau} n = \sum_{\tau=0}^{t} \tau(\tau+1)/2 = t(t+1)(t+2)/6$ different reputation values. As $t \to \infty$, VC's reputation space Θ_t is still atomic and the measure $\mathscr{Q}_t$ is discrete.

2.4.2 Stable matching with transfers

The following proposition characterizes the stable matching and wage contract for a given continuous distribution of workers $\mathscr{P}$ and discrete distribution of VCs $\mathscr{Q}_t$. This result is a special case of Azevedo and Leshno (2016), with the payoffs of two sides of the matching in analytical form. The proof of this proposition and the following two corollaries are in Appendix 2.8.3.

Proposition 2. *Denote $\theta^{(k)}$ as the $k-th$ smallest reputation in set Θ for $k \in \{1,2,...,K\}$. Create a $K-$partition of $\mathscr{L}$ such that $\underline{\ell}^{(1)} = 0$, $\underline{\ell}^{(K+1)} = 1$, $L^{(k)} = [\underline{\ell}^{(k)}, \underline{\ell}^{(k+1)})$, and $\mathscr{Q}(\theta^{(k)}) = \mathscr{P}(L^{(k)})$.*

Denote $\ell^{(k)}$ as the expected value of worker's type in interval $L^{(k)}$. Stable matching exists and is positive assortative. The VC-preferred stable matching is unique and transfer satisfies that $w(\ell^{(1)},\theta^{(1)}) = -v(\ell^{(1)},\theta^{(1)})$, and for $k \in \{2,3,...,K\}$,

$$\begin{aligned} w(\ell^{(k)},\theta^{(k)}) &= w(\ell^{(k-1)},\theta^{(k-1)}) \\ &\quad + Y(\ell^{(k)},\theta^{(k-1)}) - Y(\ell^{(k-1)},\theta^{(k-1)}) - v(\ell^{(k)},\theta^{(k)}) + v(\ell^{(k)},\theta^{(k-1)}). \end{aligned} \tag{2.5}$$

For $k \in \{1,2,...,K\}$ and $\ell \in L^{(k)}$, payoffs in stable matching are

$$r(\ell,\theta^{(k)}) = Y(\ell^{(k)},\theta^{(k)}) - w(\ell^{(k)},\theta^{(k)}), \text{ and } u(\ell,\theta^{(k)}) = w(\ell^{(k)},\theta^{(k)}) + v(\ell^{(k)},\theta^{(k)}).$$

This proposition states that in equilibrium, worker space is partitioned into different tiers, with the mass of each tier matching the mass of VC with that reputation. This result explains the mechanism of reputation spillover: VCs' reputation and workers' productivity have complementarity in total value created from matching, and wage transfers allow two sides of the market to sort efficiently.

From Equation (2.5), we have that workers' wage differentials between two tiers, $w(\ell^{(k)},\theta^{(k)}) - w(\ell^{(k-1)},\theta^{(k-1)})$ increases in the marginal product of labor $Y(\ell^{(k)},\theta^{(k-1)}) - Y(\ell^{(k-1)},\theta^{(k-1)})$ and decreases in future career benefit $v(\ell^{(k)},\theta^{(k)}) - v(\ell^{(k)},\theta^{(k-1)})$. Moreover, because of the quantile matching, the effect of career benefit is more significant if $\mathscr{Q}(\theta^{(k)})$ is low relative to $\theta^{(k)} - \theta^{(k-1)}$.

Generically, fix $\mathscr{P}$ as uniform distribution and rearrange Equation (2.5), then we have

$$\begin{aligned} \frac{w(\ell^{(k)},\theta^{(k)}) - w(\ell^{(k-1)},\theta^{(k-1)})}{\ell^{(k)} - \ell^{(k-1)}} &= \frac{Y(\ell^{(k)},\theta^{(k-1)}) - Y(\ell^{(k-1)},\theta^{(k-1)})}{\ell^{(k)} - \ell^{(k-1)}} \\ &\quad - \frac{v(\ell^{(k)},\theta^{(k)}) - v(\ell^{(k)},\theta^{(k-1)})}{\theta^{(k)} - \theta^{(k-1)}} \frac{\theta^{(k)} - \theta^{(k-1)}}{\ell^{(k)} - \ell^{(k-1)}}. \end{aligned}$$

On the right side, the first term shows additional value created by higher labor type, which increases worker's surplus. The second term shows a sensitivity of worker's future career benefit, adjusted by the relative mass of VC's reputation and worker's types. With $f(\ell) = \theta^{(k-1)}$ for $\ell \in L^{(k-1)}$ and $\mathscr{Q}(\theta^{(k-1)}) = \mathscr{P}(L^{(k-1)})$,if $\ell^{(k)} - \ell^{(k-1)}$ is small relative to $\theta^{(k)} - \theta^{(k-1)}$, then the effect of worker's future career benefit is magnified. Intuitively, worker's current period wage payoff and future career benefits are substitutes, so workers are willing to trade off their current wage for working at a reputable VC-backed startup. Further, when VC's high reputation is a scarce resource, workers' willingness to sacrifice current period payoff increases.

To illustrate the welfare implications in the following two corollaries, we make a simplifying assumption that workers' career benefit increases in VC's hidden ability x and that $v(\ell, \theta)$ is the expected career benefit given VC's ability is $x = 1$ with probability θ. That is: $v(\ell, \theta) = \theta v(\ell, 1) + (1-\theta)v(\ell, 0)$. Under this specification, worker's wages can decrease in their quality in the presence of reputation spillover

Corollary 1. *If stable matching satisfies that for some k,*

$$v(\ell^{(k)}, \theta^{(k)}) - v(\ell^{(k)}, \theta^{(k-1)}) > Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k-1)}, \theta^{(k-1)}), \tag{2.6}$$

the $w(\ell^{(k)}, \theta^{(k)}) < w(\ell^{(k-1)}, \theta^{(k-1)})$*.*

If $(\ell^{(k)} - \ell^{(k-1)})|p'(\ell^{(k)})|$ *and* $(\ell^{(k)} - \ell^{(k-1)})|d^2Y(\ell^{(k)}, \theta^{(k-1)})/d(\ell^{(k)})^2|$ *are close to* 0 *relative to* $p(\ell^{(k)})$ *and* $dY(\ell^{(k)}, \theta^{(k-1)})/d\ell^{(k)}$*, then Inequality* (2.6) *is approximately*

$$\mathscr{Q}(\theta^{(k)}) < \frac{p(\ell^{(k)})}{dY(\ell^{(k)}, \theta^{(k-1)})/d\ell^{(k)}}(\theta^{(k)} - \theta^{(k-1)})(v(\ell^{(k)}, 1) - v(\ell^{(k)}, 0)), \tag{2.7}$$

so wage decreases in worker's types when (1) marginal product of labor is low, (2) marginal career benefit is high, and (3) supply of high reputation VC is low relative to the density in the labor market.

To illustrate this result, we consider the two scenarios in Section 2.2. Suppose $Y(\ell, \theta) = \ell(\theta y_1 + (1-\theta)y_0)$, $y_1 = 2/3$, $y_0 = 1/3$, $v(\ell, \theta) = 0.5\ell\theta$, and ℓ is uniformly distributed. In this case, $p(\ell) = 1$ for all $\ell \in [0,1]$, $dY/d\ell = \theta y_1 + (1-\theta)y_0$, $p'(\ell) = 0$, $d^2Y/d\ell^2 = 0$, and $v(\ell,1) - v(\ell,0) = 0.5\ell$.

In Scenario 1, the 0.9 mass of VCs with reputation 0.2 are matched with workers $\ell \in [0,0.9)$, whose average type is 0.45; the remaining 0.1 mass of VCs with reputation 0.8 are matched with workers $\ell \in [0.9,1]$, whose average type is 0.95. With $v(0.95,0.8) = 0.38$, $v(0.95,0.2) = 0.095$, $Y(0.95,0.2) = 0.38$, and $Y(0.45,0.2) = 0.18$, we have Inequality (2.6) hold ($0.38 - 0.095 > 0.38 - 0.18$), and wage for workers with type $\ell \in (0.9,1]$ is lower than wage for workers with type $\ell \in [0.0.9]$.

In Scenario 2, the 0.1 mass of VCs with reputation 0.2 are matched with workers $\ell \in [0,0.1)$, whose average type is 0.05; the remaining 0.9 mass of VCs with reputation 0.8 are matched with workers $\ell \in [0.1,1]$, whose average type is 0.55. With $v(0.55,0.8) = 0.22$, $v(0.55,0.2) = 0.055$, $Y(0.55,0.2) = 0.22$, and $Y(0.05,0.2) = 0.02$, we have Inequality (2.6) does not hold ($0.22 - 0.055 < 0.22 - 0.02$), and wage for workers with type $\ell \in (0.9,1]$ is higher than wage for workers with type $\ell \in [0.0.9]$.

Inequality (2.7) provides a more intuitive explanation. The left side of the inequality is the market "supply" of VC's reputation $\theta^{(k)}$, and the right side shows workers' "demand" for VC's reputation. Workers' demand decreases in worker's marginal product and increases in sensitivity of career benefit with respect to VC's reputation. In the first Scenario, an insufficient supply of VCs with a high reputation increases workers' willingness to pay, and hence wage for higher-type workers is lower. In the second scenario, the mass of VCs with high reputations is sufficiently high relative to the worker's demand for such working experience. This example shows that if VC's reputation distribution has a thin right tail, then wage decreases for high-productivity workers. However, it is worth noting that VC's reputation distribution is an equilibrium object, and we will discuss the condition under which VC's distribution has a thin right tail in Section 2.4.3.

We complete the analysis of this section by looking at VC's payoff:

Corollary 2. *For all k and let $a \in (0,1)$ such that $\theta^{(k)} = a\theta^{(k-1)} + (1-a)\theta^{(k+1)}$, in stable* *matching, we have* $r(\ell', \theta^{(k)}) < ar(\ell, \theta^{(k-1)}) + (1-a)r(\ell'', \theta^{(k+1)})$*, with* $\ell \in L^{(k-1)}$*,* $\ell' \in L^{(k)}$ *and* $\ell'' \in L^{(k+1)}$*.*

This corollary shows the "convexity" of VC's return in its reputation, and because of such convexity, VC shows risk-loving payoffs. As Paul A Gompers (1996), young VCs take startups to the public earlier and are more underpriced than old VCs. The current model explains this phenomenon: in the presence of reputation spillover, learning of VC's underlying ability is valuable. Even if Bayesian updating is a martingale process (expected posterior belief equals prior belief), the convexity of VC's payoff yields that the expected payoff under posterior distribution is higher than the payoff under prior distribution.

2.4.3 Matching interacts with learning

In Section 2.4.1 and 2.4.2, matching and learning are independent processes, because $y_x(\ell)$ is not contingent on ℓ. However, it is more realistic to assume that the probability of a startup succeeding depends on both the worker's and VC's ability, and this setup allows us to analyze more interesting dynamics: At the beginning of each period, VCs and workers match in a positive assortative way. Regardless of the VC's true ability, a higher reputation leads to better worker input and hence a higher probability of success, which in turn allows the VC to further build a reputation. This feedback loop makes the self-fulfilling prophesy possible, in which VC's initial startup investment returns correlate with future ones and VCs of different abilities pooled with similar reputations.

We formalize this intuition in this section. We will first apply the positive assortative matching results in Proposition 2 for VC distribution in each period, and then analyze the new learning process as a result of positive assortative matching and derive the steady-state distribution of VCs. A crucial assumption of Proposition 2 is that Θ is atomic and $\mathscr{Q}_t$ is discrete. In general, this is not necessarily true when $y_x(\ell)$ can take a continuum of values. To resolve this issue, we make the tie-breaking rule Assumption 5, which makes sure that $y_x(\ell)$ can take at most countable distinct values. This modification preserves Assumption 4 and makes sure Θ atomic.[3]

At the beginning of the game, $\Theta_0 = \{\theta_0\}$, and $\mathscr{Q}_0(\theta_0) = 1$. In period t, given previous period reputation space Θ_{t-1} and distribution $\mathscr{Q}_{t-1}$, we have the sorted reputation $\{\theta_{t-1}^{(k)}\}_{k=1}^{K(t)}$ and corresponding mass $\{\mathscr{Q}_t(\theta_{t-1}^{(k)})\}_{k=1}^{K(t)}$. We next partition $\mathscr{L} = [0,1]$ into $K(t)$ intervals $\{L^{(k)}\}_{k=1}^{K(t)}$ such that $\mathscr{P}(L^{(k)}) = \mathscr{Q}_t(\theta_{t-1}^{(k)})$. Construct the sequence of average labor types $\{\ell^{(k)}\}_{k=1}^{K(t)}$ from $\ell^{(k)} = \int_{L^{(k)}} \ell d\mathscr{P}$. VCs with reputation $\theta_{t-1}^{(k)}$ match randomly with any worker $\ell \in L^{(k)}$, and startup has inputs $\theta^{(k)}$ and $\ell^{(k)}$.

The following period reputation space is composed of $\Theta_t = \Theta_0 \cup \Theta_{t-1}^H \cup \Theta_{t-1}^L$, where

$$\Theta_{t-1}^H = \left\{ \theta_t^{H(k)} \equiv \frac{\theta_{t-1}^{(k)} y_1(\ell^{(k)})}{\theta_{t-1}^{(k)} y_1(\ell^{(k)}) + (1-\theta_{t-1}^{(k)}) y_0(\ell^{(k)})} : k = 1,2,...,K(t) \right\}$$

and

$$\Theta_{t-1}^L = \left\{ \theta_t^{L(k)} \equiv \frac{\theta_{t-1}^{(k)} (1 - y_1(\ell^{(k)}))}{\theta_{t-1}^{(k)} (1 - y_1(\ell^{(k)})) + (1-\theta_{t-1}^{(k)})(1 - y_0(\ell^{(k)}))} : k = 1,2,...,K(t) \right\}.$$

Probability mass in period $\mathscr{Q}_t$ is linked to $\mathscr{Q}_{t-1}$ in the following way:

$$\begin{aligned}
&\mathscr{Q}_t(\theta_0) = q, \\
&\mathscr{Q}_t(\theta_t^{H(k)}) = (1-q)(\theta_{t-1}^{(k)} y_1(\ell^{(k)}) + (1-\theta_{t-1}^{(k)}) y_0(\ell^{(k)}))\mathscr{Q}_{t-1}(\theta_{t-1}^{(k)}), \\
&\mathscr{Q}_t(\theta_t^{L(k)}) = (1-q)(\theta_{t-1}^{(k)} (1 - y_1(\ell^{(k)})) + (1-\theta_{t-1}^{(k)})(1 - y_0(\ell^{(k)})))\mathscr{Q}_{t-1}(\theta_{t-1}^{(k)})
\end{aligned}$$

for all $k \in \{1,2,...,K(t)\}$. The following lemma shows that this distribution converges to a steady state distribution, which can be numerically solved.

Lemma 18. *As $t \to \infty$, the cross-sectional distribution of VC reputation converges in law to a*

3. To see this, consider in period 1 when all VCs have identical reputation θ_0, workers with types $\ell \in \mathscr{L} = [0,1]$ randomly match with VCs. Without the tie-breaking rule, after the startup's output is realized, each value of ℓ could correspond to two possible VC's posterior reputation, and at the end of period 1, there are unaccountably many possible reputations of VC.

steady state distribution.

Next, we examine the impact of reputation spillover on the learning speed of VC's ability. Recall that learning is faster when $D_{KL}(\theta'||\theta)$ is higher.

Proposition 3. *For fixed θ, monotonicity of $D_{KL}(\theta'||\theta)$ depends on the complementarity between ℓ and x in function $y_x(\ell)$:*

1. $\frac{dy_1(\ell)/d\ell}{dy_0(\ell)/d\ell} > \frac{y_1(\ell)}{y_0(\ell)} > \frac{1-y_1(\ell)}{1-y_0(\ell)}$, *then $D_{KL}(\theta'||\theta)$ increases in ℓ for both good and bad news.*
2. $\frac{y_1(\ell)}{y_0(\ell)} > \frac{1-y_1(\ell)}{1-y_0(\ell)} > \frac{dy_1(\ell)/d\ell}{dy_0(\ell)/d\ell}$ *, then $D_{KL}(\theta'||\theta)$ decreases in ℓ for both good and bad news.*
3. $\frac{y_1(\ell)}{y_0(\ell)} > \frac{dy_1(\ell)/d\ell}{dy_0(\ell)/d\ell} > \frac{1-y_1(\ell)}{1-y_0(\ell)}$, *then $D_{KL}(\theta'||\theta)$ decreases in ℓ for good news and increases in ℓ for bad news.*

In Case 1, success requires both VC and worker's inputs. This could correspond to the development of the product and service in the local market. VC's managerial expertise and workers' skills complement each other, and the startup needs both to succeed. In Case 2, success requires only resources from one of the VCs and the worker. This could model the resources to make the product could go viral, or access to big clients, financing parties, and investment bankers. If a startup's founder already has a good network and is capable of achieving these milestones (high ℓ), then the marginal effect on startup success rate that a $x = 1$ VC could add is limited. Hence, we could classify early-stage startups to Case 1, and late-state startups to Case 2, with Case 3 being the transition phase. [4]

In Case 1, when reputation is low, each additional signal is less informative (divergence measure is lower) as reputation decreases, so the mass of the distribution is concentrated on the left tail. When reputation is higher, each additional signal is more informative (divergence

4. Note the second case implies that $dy_1/d\ell < dy_0/d\ell$, which does not contradict the standard interpretation that VC's value creation improves the startup's marginal product of labor. Recall that we made the simplifying assumption that the startup's production is separable with two components: the probability of success $y_x(\ell)$ and revenue conditional on success $R(\ell)$. The common assumption that VC's value creation improves the startup's marginal product of labor, which would correspond to the complementarity of ℓ and θ in the expected revenue $Y(\ell,\theta) = R(\ell)y(\ell,\theta)$.

measure is higher) as reputation increases, until θ gets sufficiently close to 1 and the effect of $d\theta$ dominates. More information about VC is revealed along the right tail so the right tail is longer.

Case 2 mirrors the first case: when reputation is high, each additional signal provides diminishing information (lower divergence measure), resulting in a concentration of the distribution's mass in the right tail. In contrast, with a lower reputation, each added signal offers more information (higher divergence measure) as reputation increases, until θ approaches sufficiently close to 0, where the impact of $d\theta$ becomes dominant. This reveals more information about the VC in the left tail, resulting in a longer left tail.

Case 3 does not generate skewed reputation distribution. Good news is informative when reputation is low, and bad news is informative when reputation is high. This leads to better information revelation in both tails of the distribution. The following corollary summarizes this result.

Corollary 3. *If $\theta_0 = 0.5$, then skewness of reputation steady state distribution depends on $y_x(\ell)$.*

1. *If $\frac{dy_1(\ell)/d\ell}{dy_0(\ell)/d\ell} > \frac{y_1(\ell)}{y_0(\ell)} > \frac{1-y_1(\ell)}{1-y_0(\ell)}$, then reputation steady state distribution has positive skewness.*
2. *If $\frac{y_1(\ell)}{y_0(\ell)} > \frac{1-y_1(\ell)}{1-y_0(\ell)} > \frac{dy_1(\ell)/d\ell}{dy_0(\ell)/d\ell}$, then reputation steady state distribution has negative skewness.*

This result links the VC market reputation distribution with the complementarity of worker's productivity and VC's ability to create value, which could be tested empirically. Startups of different industries, stages, and geographical locations usually have different production technologies, and their likelihood of success requires different worker-VC complementarity. VCs specializing in investing in such startups could have different life cycles, market concentrations, and competition.

An example of existing literature supporting this idea is Hsu (2004), which shows that startups in different industries have a different propensity of receiving multiple offers from VCs. According to Hsu (2004), startups in the internet industry are most likely to receive multiple offers, while startups in the healthcare industry are least likely. This is explained by the model that production technology in the healthcare industry is closer to Case 1 and the Internet industry is closer to Case 2. VCs with above-average reputations concentrate more on Case 2, which leads to high competition.

2.5 Numerical Analysis

In this section, we start by simulating the model with functional form assumptions, and we will illustrate the different cases in Proposition 3 with three ranges of parameter values. We assume strong VC-worker complementarity in the baseline analysis, where we compute the steady-state distribution numerically, as well as the matching outcome and players' payoffs in the steady state. Next, we consider comparative statics with respect to VC-worker complementarity and the case without the worker's future career benefit. Finally, we will illustrate the implied learning speed for each possible match between VCs and workers, and we show how the learning speed in steady-state stable matching explains the imperfect information revelation.

2.5.1 Production technology

In this section, we illustrate the main results with a numerical simulation. Start by making the following functional form assumptions:

$$y_x(\ell) = 0.2 \times 2^{-1/\rho} \times \left((1+x)^\rho + (1+\ell)^\rho\right)^{1/\rho}. \tag{2.8}$$

The function $y_x(\ell)$ is scaled by factor 0.2 to match the fact that startups have low success rate, and by factor $2^{-1/\rho}$ such that $y_0(0)$ does not vary with ρ. The functional form assumes constant elasticity of substitution $1/(1-\rho)$, between VC input x and worker input ℓ.

With additional assumptions $R(\ell) = \ell + \ell^2$ and $v(\ell, \theta) = \theta \ell$, we have expected output conditional on true ability x as $Y_x = y_x(\ell)R(\ell)$, and ex ante expected total value created from matching:

$$Z(\ell,\theta) = Y(\ell,\theta) + v(\ell,\theta) = R(\ell)(\theta y_1(\ell) + (1-\theta)y_0(\ell)) + \theta\ell.$$

One advantage of this numerical specification is that we can investigate the impact of the elasticity of substitution on the matching, payoff, and steady-state distribution. We can verify Assumption 4, where different parameter values of ρ correspond to the three cases in Proposition 3.

Lemma 19. *With the assumed functional form above,*

1. *For $\rho < 0$, then $\frac{dy_1}{dy_0} > \frac{y_1}{y_0} > \frac{1-y_1}{1-y_0}$ holds for all $\ell \in [0,1]$.*
2. *There exists $\hat{\rho} \in (\tilde{\rho}, 2)$, such that for $\rho \geq \hat{\rho}$, $\frac{y_1}{y_0} > \frac{1-y_1}{1-y_0} > \frac{dy_1}{dy_0}$ holds for all $\ell \in [0,1]$.*
3. *There exists $\tilde{\rho} \in (0,1)$, such that for $\rho \in [0,\tilde{\rho}]$, $\frac{y_1}{y_0} > \frac{dy_1}{dy_0} > \frac{1-y_1}{1-y_0}$ holds for all $\ell \in [0,1]$*

Figure 2.1 illustrates the functional form assumptions of production technology: y_x, (probability of success conditional on true ability x), Y_x (expected output conditional on true ability x), and increasing difference $Y_1 - Y_0$. Throughout the simulation, I consider three cases of parameter $\rho \in \{-5, 1, 5\}$ that represent three possible labor-VC complementarity cases in the main results. To illustrate the effect of the interaction of matching and learning, I adopt parameters $y_0 = 0.2$ and $y_1 = 0.4$ in the case of independent and identically distributed signals.

2.5.2 Convergence and steady state

In the simulation, we consider $\theta_0 = 0.5$, and the probability of replacement $q = 0.1$. With these parameters, the average age of VC is $(1-q)/q = 9$, and with probability 95%, each VC is replaced before age 28.5. In the baseline analysis, we adopt $\rho = -5$, and in this case, $y_x(\ell)$ follows a typical constant elasticity of substitution production function with the elasticity of substitution equal to $1/6$. In the next subsection, we will investigate the convergence and steady state of different ρ values.

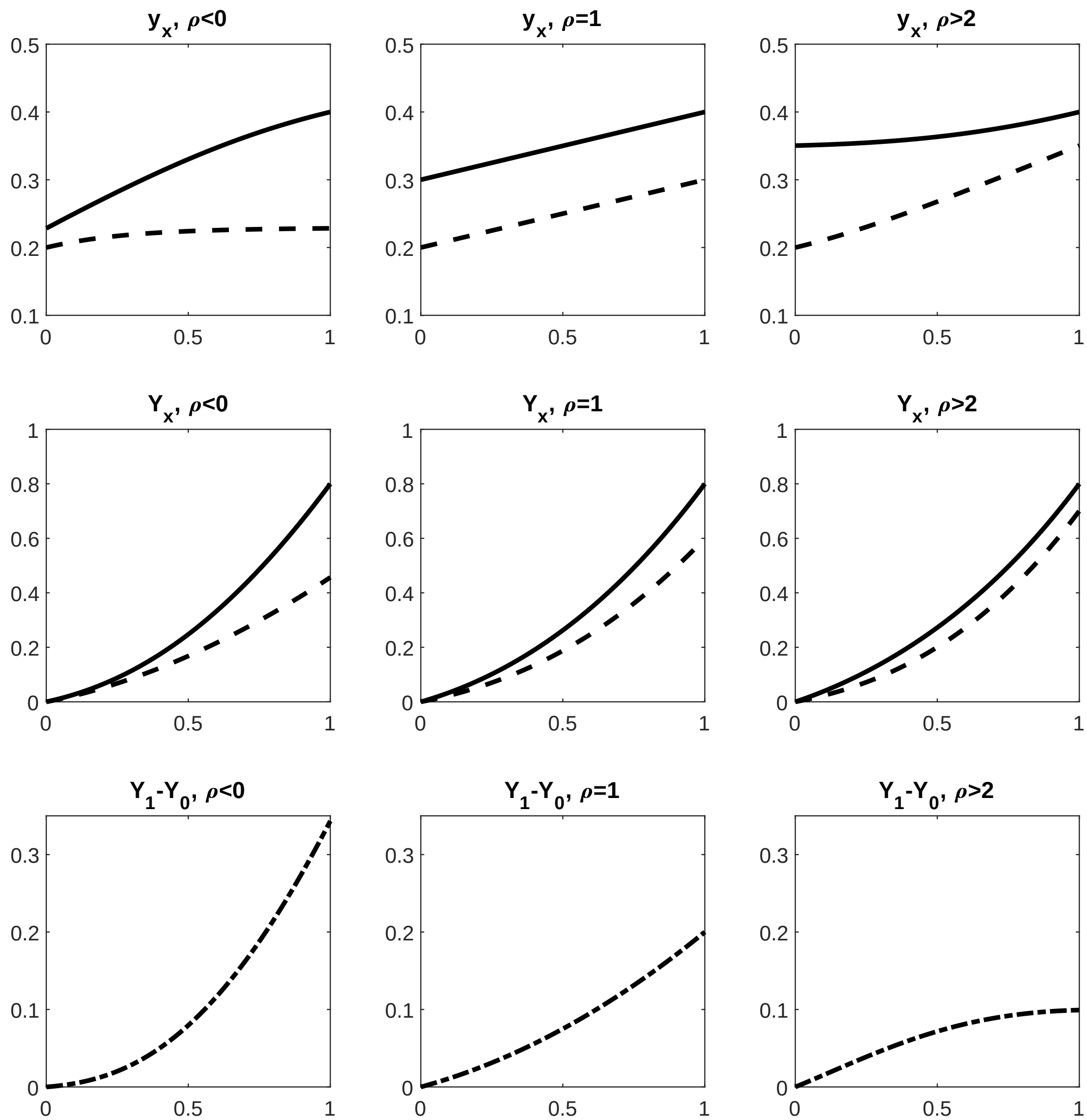


Figure 2.1. Illustration of production technology. Production technology assumption; nine plots above show how productivity varies across horizontal axes labor input ℓ from 0 to 1. Top and middle panels are y_x and $Y_x = Ry_x$ conditional on known $x \in \{0,1\}$, where solid lines represent $x = 1$ and dashed lines represent $x = 0$. The bottom panels show the difference between solid and dashed lines in the middle panels $Y_1 - Y_0$. From left to right, the production technology plots use parameters $\rho = -5$, $\rho = 1$, and $\rho = 5$.

Figure 2.2 shows the convergence of cross-sectional distribution of VC reputation $\mathscr{Q}_t(\theta)$. In a steady state, the distribution of VC reputation has a mean of equal to $\theta_0 = 0.5$, because of the martingale property of Bayesian learning. The steady state distribution has a standard deviation of 14.85% and skewness of 1.1921. The mass of the distribution is concentrated on the left tail (steep CDF for $\theta \in [0.3, 0.5]$) and the right tail is longer (flat CDF for $\theta \in [0.5, 1]$). No VC has a reputation lower than 0.3; incomplete learning is mainly due to VCs' finite life, but because of the interaction of matching and learning with $\rho < 0$, learning is slower when reputation is low (Proposition 2) at the left tail.

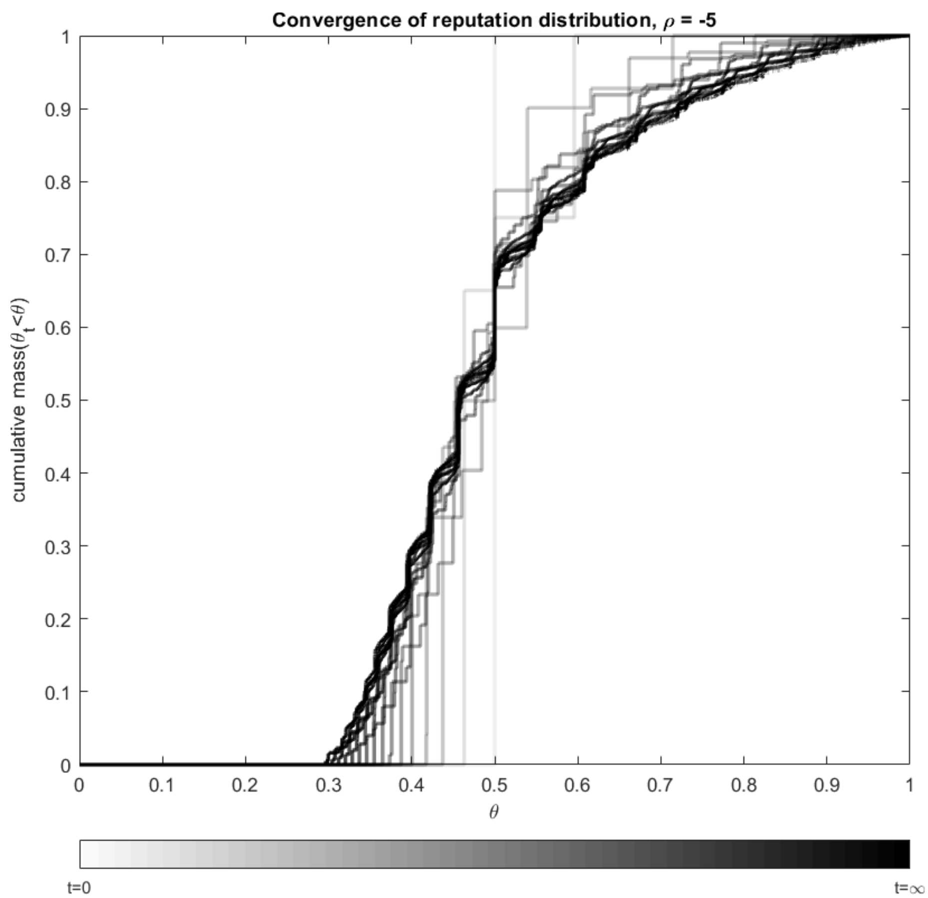
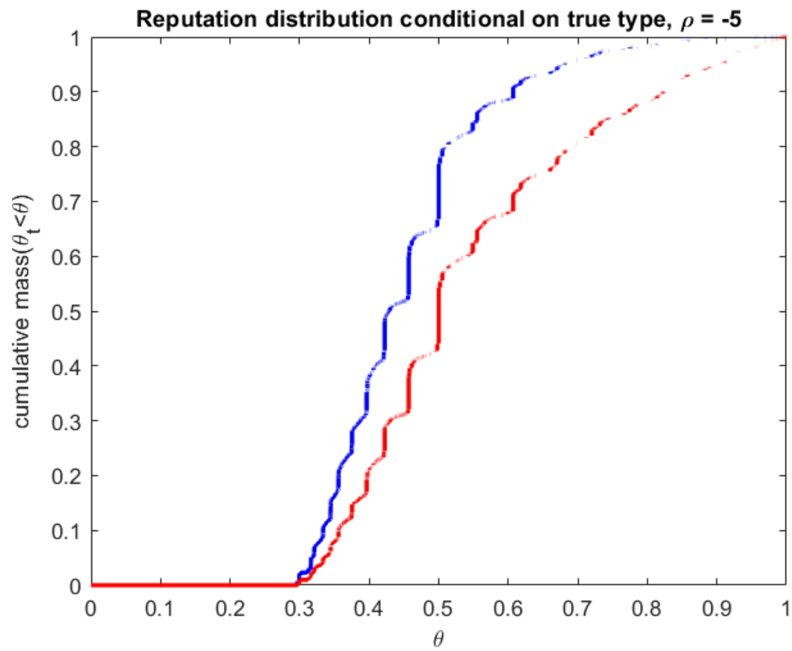


Figure 2.2. Convergence of reputation distribution in the case with high VC-worker complementarity

Figure 2.3 shows the steady state distribution conditional on true ability x. Good VC's reputation has first-order stochastic dominance over bad VC's reputation, as the blue curve ($x = 0$) lies below the red curve ($x = 1$), which means that for any reputation θ, $x = 1$ VC gives at least as high a mass of having at least reputation θ as does $x = 0$ VC. The mean and standard



Blue curve is steady state distribution of $x = 0$ VC, and red curve is steady state distribution of $x = 1$ VC.

Figure 2.3. Steady state distribution by true type

deviation of reputation distribution conditional on $x = 1$ and $x = 0$ are 0.5451, 16.89% and 0.4549, 11.18% respectively. In conclusion, $x = 1$ has a better reputation than $x = 0$ on average, but with positive probability, $x = 1$ VC has a worse reputation than $x = 0$, and pooling of $x = 1$ and $x = 0$ is most significant when reputation is low.

In a steady state, VC and workers match according to Proposition 1. Figure 2.4 shows the matching and contracts in a steady state. Panel (a) shows positive assortative matching between VC with reputation θ (horizontal axis) and worker with type ℓ (vertical axis). It is not a coincidence that this curve is identical to VC's steady state reputation CDF, because of the assumption that workers are uniformly distributed on the interval $[0, 1]$.

Panel (b) shows the wage transfers in steady-state equilibrium. In particular, it illustrates the non-monotonicity of wage contracts as in Corollary 1. Because of the VC-worker complementarity in production technology, the worker can retain part of the marginal product of labor; yet because workers' future career payoff and current wage payoff are substitutes, workers are

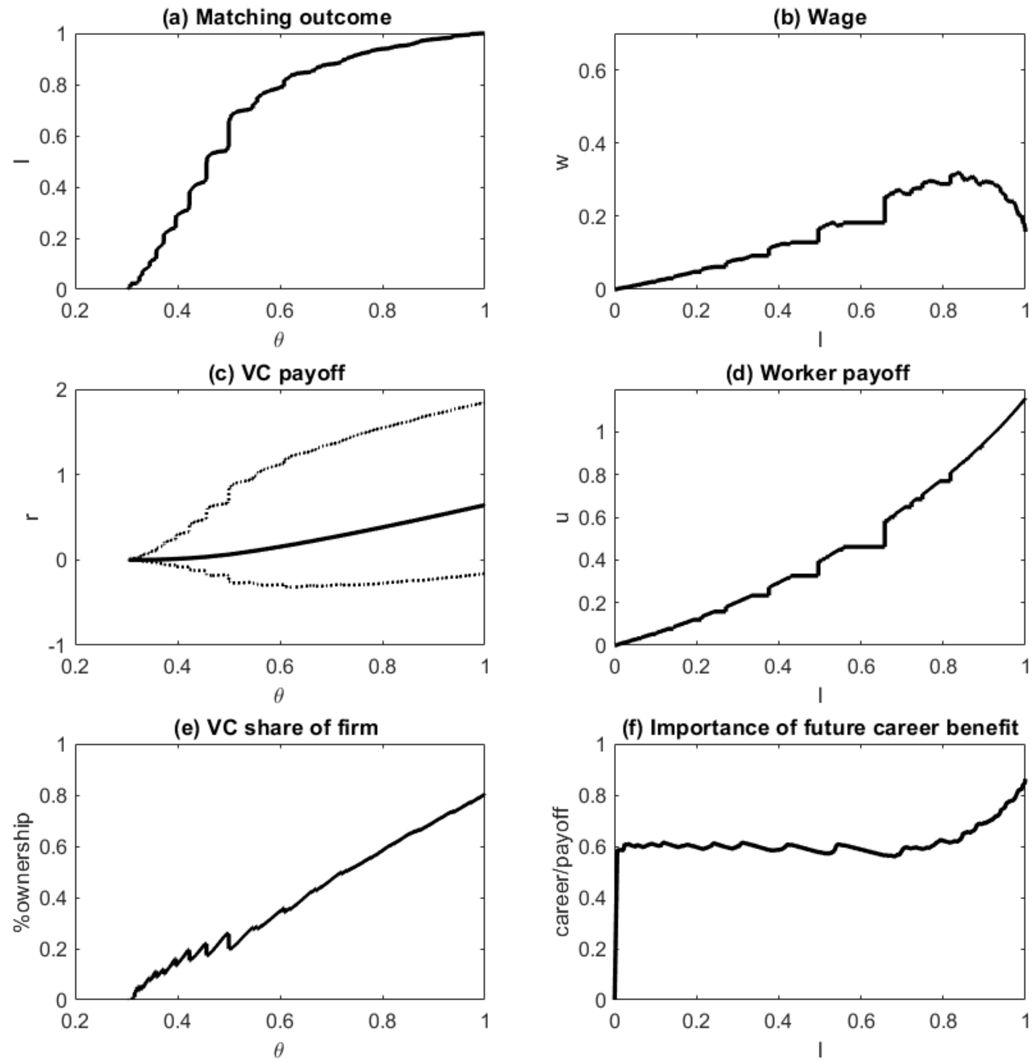


Figure 2.4. Steady state contracts and payoffs

willing to sacrifice some wages to work at a reputable VC-backed startup. When the effect of the marginal product of labor dominates, the wage increases in ℓ, and when the future career payoff effect dominates, the wage decreases in ℓ. Which effect dominates depends on the distribution of θ: in the left tail, when more VC pools around close reputation, workers get higher market power in that 1% increase of worker is more valuable than 1% increase in VC quantile.

Panel (c) shows the expected payoff (solid curve) and realized payoffs (dashed curves) under good and bad outcomes. Upon a bad outcome, VC's losses are worker's wage, and VC's net gain in a good state is total revenue $R(\ell)$ net of worker's wage. As in Corollary 2, VC's expected payoff is convex in θ. Panel (e) shows the VC's share of the firm as a function of the VC's reputation, and this is consistent with empirical evidence that reputable VC acquires larger shares of startups Hsu (2004).

Further, notice that VC's realized payoff distribution increases as VC's reputation increases, but downside risk (losses after a bad outcome) stops increasing, if not decreasing, when VC's reputation is sufficiently high. Hence, despite that reputable VCs claim a higher share of the firm in expectation, the model explains why reputable VCs demand lower downside protection, as documented by Bengtsson and Sensoy (2011).

By assumption, VC undertakes all risks of a startup and pays workers certain wages regardless of the startup outcome. When VC's reputation is high, higher types of workers are willing to work at lower wages, which helps VC to manage its downside risks and provides a cushion for VC's long-term return. In this sense, the model provides an alternative explanation of the VC's return persistence (Kaplan and Schoar (2005)).

Panel (d) shows worker's payoffs. Despite workers' wage transfers are not monotonic in their type, payoffs are still monotonic. Note that the stable matching is unique but the transfer scheme is not unique, and the transfer scheme is unique up to a constant set by a boundary condition. This current equilibrium is selected by setting $u(\ell^{(1)}, \theta^{(1)}) = 0$. Panel (f) shows how much if worker's payoff comes from future career benefits. The impact of future career benefit is most significant for higher type workers, which is consistent with the finding in Bernstein et al. (2022) that being funded by a top-tier investor attracts both high and low-quality workers, but only high quality workers get hired.

2.5.3 Comparative statics

We consider comparative statics with respect to the worker's future career benefit $v(\ell, \theta)$ and production technology ρ.

First, if we shut down the channel of the worker's future career benefit, that is to set $v(\ell, \theta) = 0$ for all ℓ and θ, then convergence, steady-state distribution, and stable matching are

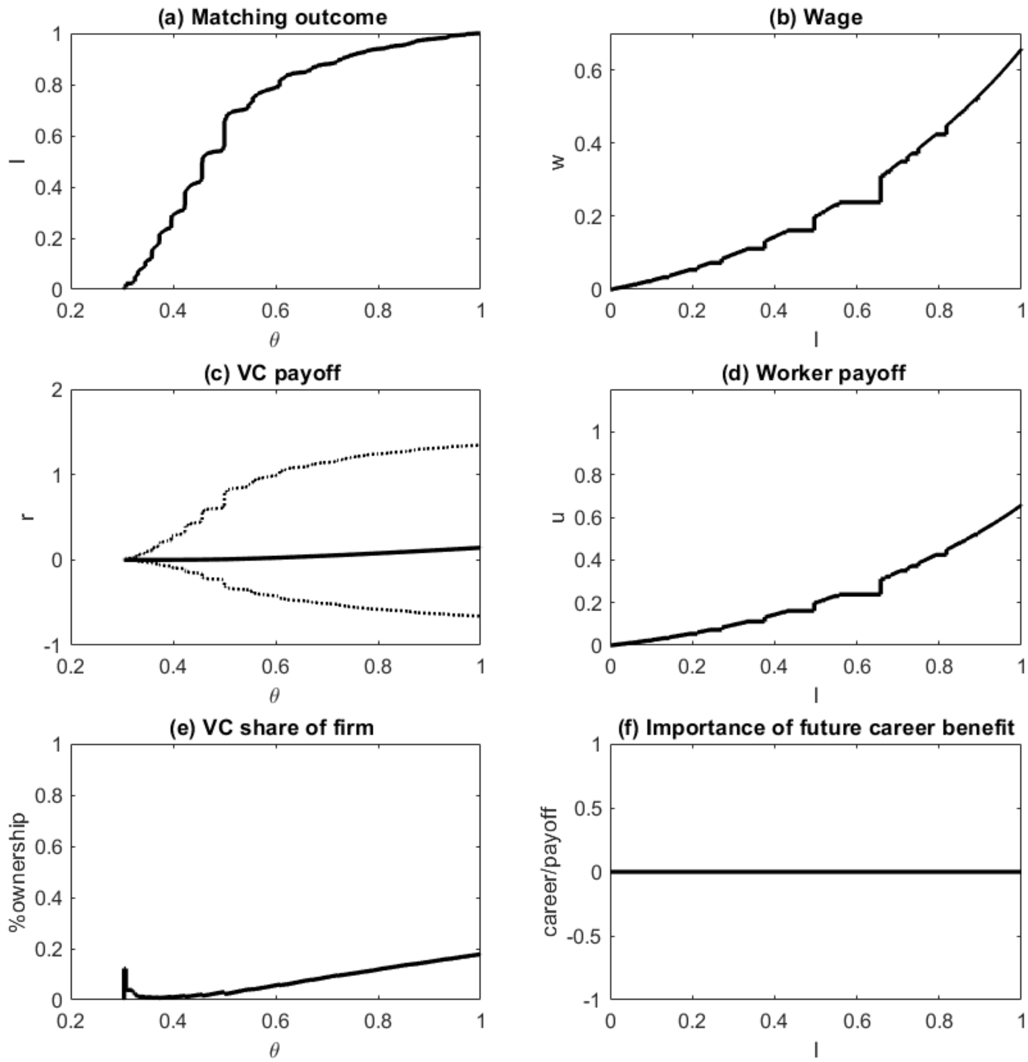


Figure 2.5. Steady state contracts and payoffs, without worker's future career benefit

the same as the baseline simulation, but the transfer and payoffs are different. While learning and matching mainly depend on the signal structure $y_x(\ell)$, career benefit influences value created from matching and how the surplus is shared between VC and workers.

Similar to Figure 2.4, the steady state contracts and payoffs without the worker's future career benefit are illustrated in Figure 2.5. In this case, the worker's payoff solely comes from wage transfer. Despite that VC's share of the firm still increases in VC's reputation, VC's downside exposure keeps increasing as reputation increases. Reputable VC-backed startups are attractive to workers only to the extent of higher wage offers. Hence, to explain empirical findings, a worker's career benefit is crucial.

Further, without a worker's future career benefit, VC's bargaining power is strongly limited. Despite that in the baseline simulation, workers retain smaller shares of the revenue, workers' total payoff is still higher than in the case without future career benefits. This justifies the reputation premium in Hsu (2004), as reputable VC not only claims higher share surplus but also creates more values for both parties, and VC's reputation premium can be interpreted as a way of internalizing the positive externality.

Second, we examine the cases of different production technology $y_x(\ell)$ and illustrate the results in Corollary 3. Qualitative properties of convergence for each case are similar to Figures 2.2 and 2.3, and comparative statics results are summarized in Table 2.1.

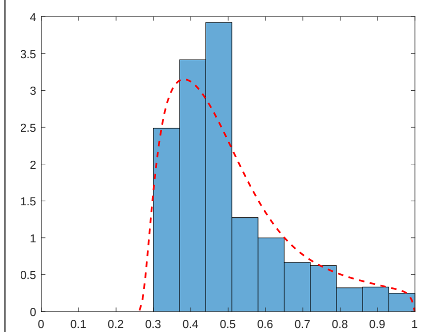
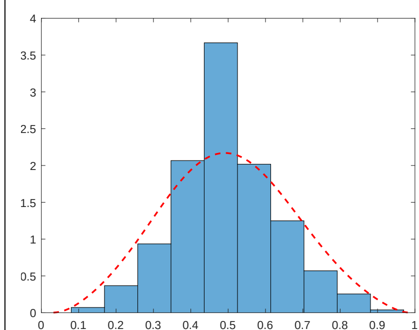
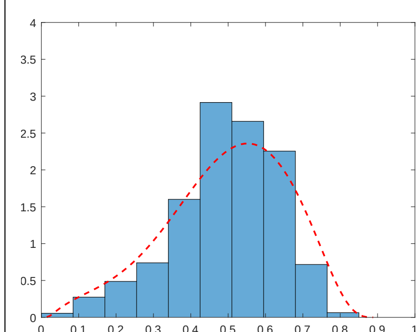
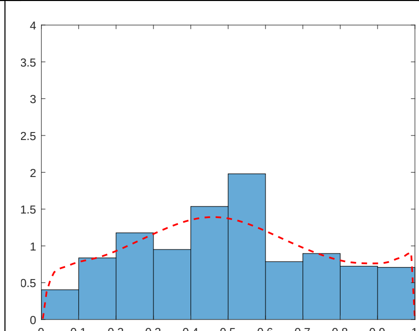
When VC input and worker input are perfect substitutes ($\rho = 1$), the steady-state distribution is close to symmetric; because of the high substitutability of labor and VC inputs, the good news is more informative for low reputation VCs and bad news is more informative for high reputation VCs. In the case of $\rho = 5$, the mass of the distribution is concentrated on the right tail and the left tail is longer (negative skewness). Comparing the cases with and without the interaction between matching and learning, we have that the case with IID signals has the highest standard deviation and lowest kurtosis. This shows that such interactions reduce the degree of information revelation and lead to a more concentrated VC market. The comparative static show that different production technology corresponds to different VC market concentrations.

2.5.4 Speed of information revelation

Figure 2.6 shows the divergence $D_{KL}(\theta'||\theta)$ for all possible combination of inputs θ and ℓ. For each fixed ℓ, the figure illustrates Lemma 14: divergence achieves its maximum at θ^*, which is less (greater) than 0.5 upon a good (bad) news. It also shows Proposition 3 and Lemma 19: when ρ is low (high), D_{KL} increases (decreases) in ℓ for fixed θ.

In steady state equilibrium, Figure 2.7 shows the divergence $D_{KL}(\theta'||\theta)$ for pairs of stable matching ℓ and $\theta = f(\ell)$. This shows the joint effect of θ and ℓ on learning speed. In the

Table 2.1. Summary of comparative statics of steady-state distribution of VC reputation with respect to different cases of substitutability of production inputs (ρ).

Cases	Mean	Standard Deviation	Skewness	Kurtosis	Histogram and Kernel Density
$\rho = -5$	0.5	0.1485	1.1921	3.9801	
$\rho = 1$	0.5	0.1373	0.1459	3.2369	
$\rho = 5$	0.5	0.1378	-0.7229	3.4447	
IID signals	0.5	0.2455	0.1310	2.2854	

case of $\rho < 0$, learning is fastest for θ greater than 0.5 upon both good and bad news. Upon good news, the effect of ℓ dominates, and, together with positive assortative matching, this shifts the reputation with the fastest information revelation from below to above 0.5. In the case of bad news, learning is fastest at $\theta^* > 0.5$, and learning speed further increases in ℓ as θ increases. The two effects reinforce each other, and hence learning is faster for a high reputation. Further, for VCs with low reputation, as θ keeps decreasing, the speed at which D_{KL} decreases in θ also decreases, This explains the tendency of pooling around low reputation in a steady state: when reputation is sufficiently low, the market does not infer much new information from an additional bad news.

2.6 Discussion

In this section, we discuss two extensions of the main model, followed by additional connections to related literature.

2.6.1 Extension with forward-looking VCs

The model implicitly assumes that VCs are myopic and maximizes the expected return without considering intertemporal tradeoffs. While this assumption aligns with the practical constraints that real-world VCs often face, the following lemma outlines the essential criteria under which positive assortative matching can persist, even in scenarios where VCs exhibit forward-looking behavior. It is worth noting that the numerical analysis specifications adhere to these criteria, ensuring the sustainability of positive assortative matching.

Lemma 20. *Under the following two cases, stable matching in Proposition 2 sustains in the presence of forward-looking VCs.*

1. $y_1(\ell) - y_0(\ell)$ *increases in* ℓ *and* $d(Z(\ell,1) - Z(\ell,1)))/d\ell$ *increases in* ℓ*,*

2. $y_1(\ell) - y_0(\ell)$ *decreases in* ℓ *and* $d(Z(\ell,1) - Z(\ell,1)))/d\ell$ *decreases in* ℓ*,*

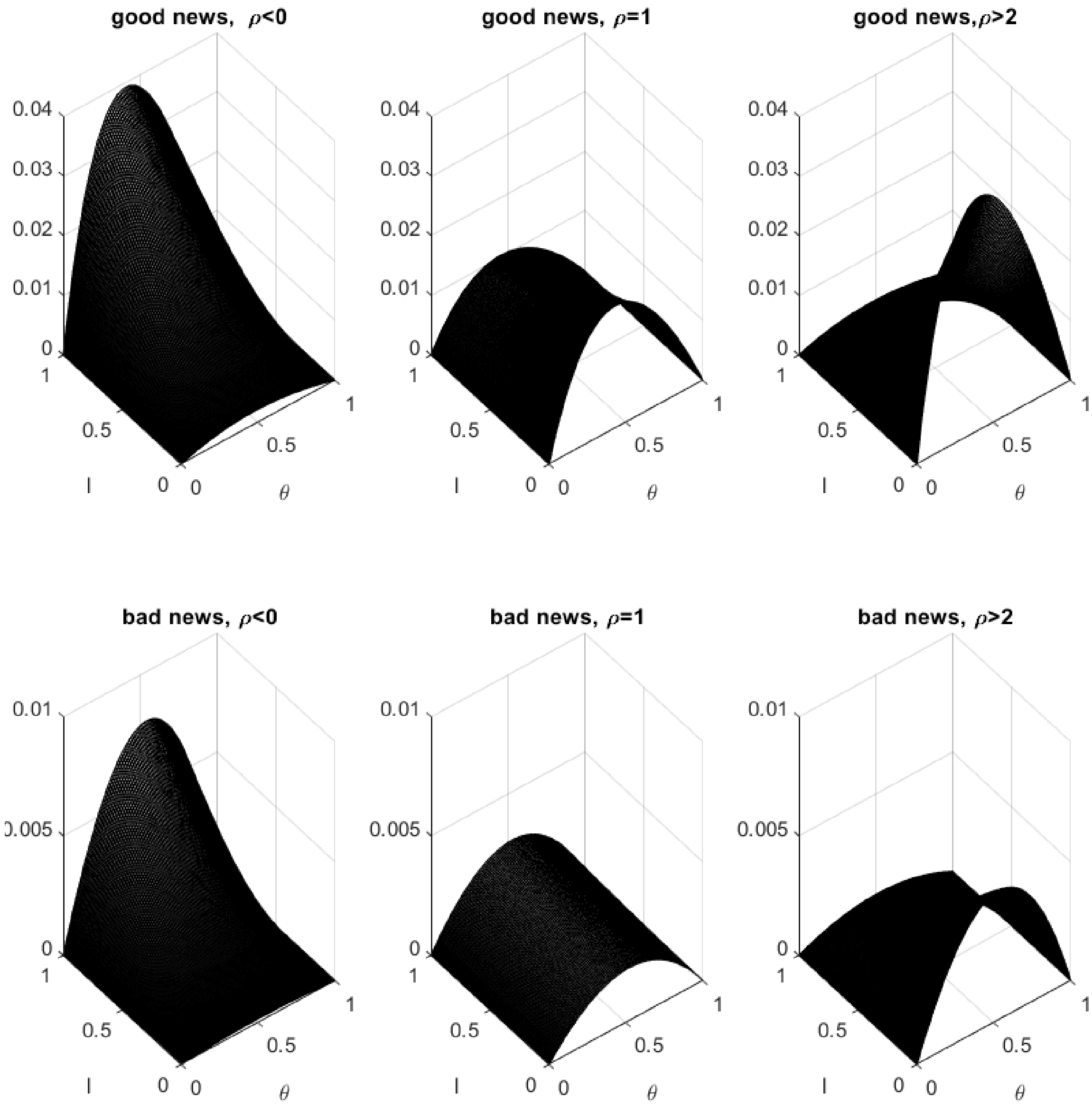


Figure 2.6. Speed of information revelation

2.6.2 Additional connections to the literature

VCs add value to startups in various ways. Nahata (2008) finds that reputable VC-backed startups are more likely to exit successfully. Krishnan et al. (2011) find that reputable VC-backed firms have superior long-run performance. Chemmanur, Krishnan, and Nandy (2011) find that reputable VC-backed startups have high growth of productivity and sales, and low growth of production costs.

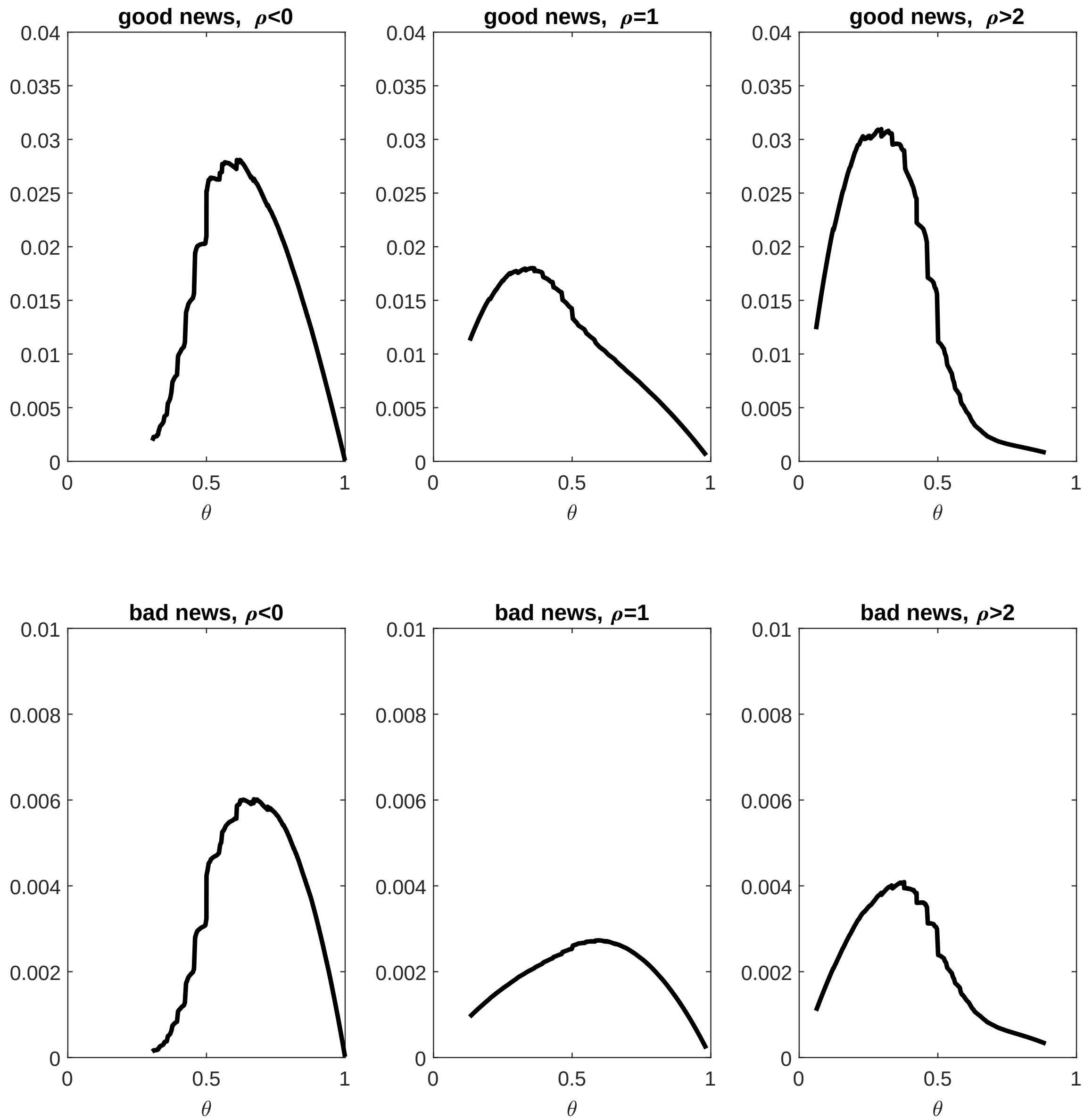


Figure 2.7. Speed of information revelation in steady state

Hochberg, Ljungqvist, and Lu (2007) measure the effect of VC's networks on investment performance. They find that networks not only give VCs a better deal flow but also enable VCs to provide better value-added services to their portfolio companies. Further, they find that VC networks correlate with experiences; the effect of networks is significant after controlling for a variety of dimensions of experience, but the effect of experience is reduced after controlling for networks. They also found a significant network effect after controlling for return persistence.

This paper is connected to the literature on VC term sheets. Hsu (2004) finds that offers made by VCs with a high reputation are three times more likely to be accepted, and high-reputation VCs acquire start-up equity at a 10–14% discount. On the other hand, Bengtsson and Sensoy (2011) find that more experienced VCs obtain economically and statistically significantly weaker downside protections, suggesting that entrepreneurs pay less (in an expected value sense) for affiliation with high-quality VCs than Hsu (2004) suggests. Bengtsson and Sensoy (2011) also find that more experienced VCs are more likely to join their portfolio companies' boards of directors, which is consistent with Bottazzi, Da Rin, and Hellmann (2008).

Ewens, Gorbenko, and Korteweg (2022) study matching model between heterogeneous VCs and heterogeneous entrepreneurs. Empirically they find that VCs "use their bargaining power to receive more investor-friendly terms compared to the contract that maximizes startup value", and that "better VCs still benefit the startup and the entrepreneur due to their positive value creation".

Kaplan and Schoar (2005) document the persistence of private equity returns and that performance increases with fund size and with the VC's experience. They conjectured that such results are driven by VCs' heterogeneous skills. Hochberg, Ljungqvist, and Vissing-Jørgensen (2014) reconcile the puzzle that despite VC's skills being valuable and scarce resources, VC does not raise fees to eliminate excess demand. Hochberg, Ljungqvist, and Vissing-Jørgensen (2014) also empirically estimate that up to two-thirds of VC firms lack skill.

This paper provides an additional channel of return persistence while existing literature has emphasized the deal flow channel. Cong and Xiao (2022) study a delegate investment model where general partners (investors) contract with limited partners (VCs). LPs offer a hierarchy of contracts based on GPs' track record, which creates dynamic incentives for GPs to exert effort. The contracts assigned by LPs are incentive-compatible for GPs to invest in certain types of projects. Consequently, homogeneous fund managers have heterogeneous deal flow and fund performance. The model explains the empirical evidence of performance persistence and mean-reversion in long-term performance.

Sørensen (2007) develops a structural model to distinguish sorting and influence effects of VCs and to estimate the relative importance. The structural model has two parts; the first part consists of the outcome of the investments, and the second part controls for sorting by applying a two-sided matching model (the Gale and Shapley (1962) college admissions model), in a static, complete information, exogenous and homogeneous bargaining power setting. They find that both effects are significant, with sorting almost twice as important as influence for the difference in IPO rates.

2.7 Conclusion

We provide a model of VC financing that captures the dynamics of VC reputation spillover, matching processes, and their welfare implications. Within each period, the stable matching between workers and VCs is positive assortative, which provides a micro foundation of the reputation spillover effect. A notable result of our model is the counterintuitive observation that more productive workers can earn lower wages, which provides new channels of explaining empirical regularities, including VC's return persistence.

The model also emphasizes the interdependence between the reputation process and the cross-sectional distribution. As a result of reputation spillover, VCs with high reputations are matched with productive workers, which improves the chance of VC gaining an even better reputation. Yet the learning speed is sensitive to the degree of complementarity between VCs and workers, which further influences the steady-state distribution of reputation among VCs. Such complementarity characterizes different roles of VCs in different stages, industries, or locations of startups. Moreover, as the contract and matching depend on steady state distribution of VCs, the production technology also leads to different variations in welfare predictions.

For future research, several potential directions emerge from our findings. The model could be extended to allow $y_x(\ell)$ to take continuous values and explore non-atomic two-sided matching models, offering a broader perspective on real-world scenarios. Additionally, introducing heterogeneity in workers' wage lower bounds could lead to more realistic dynamics, where some workers are unable to make the trade-off between current wages and future career benefits. Incorporating endogenous VC exits could enhance the model's alignment with empirical reputation measures, connecting theory with real-world data. Furthermore, explicitly modeling workers' career benefits, especially in cases where benefits are tied to stock options or higher future wages within the startups, would offer insights into different incentive structures.

2.8 Appendix

2.8.1 Proof of useful results in Section 2.3

Derivation of Equation (2.2)

Rearrange Equation (2.1) and we have

$$\theta_t = \frac{\mu(\ell)\theta_{t-1}}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}}.$$

Using this condition to substitute θ_t in the definition of

$$D_{KL}(\theta_t||\theta_{t-1}) = \theta_t \log\frac{\theta_t}{\theta_{t-1}} + (1-\theta_t)\log\frac{1-\theta_t}{1-\theta_{t-1}},$$

collecting terms, and we obtain

$$\begin{aligned} D_{KL}(\theta_t||\theta_{t-1}) = \frac{\mu(\ell)\theta_{t-1}}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}} \log\frac{\mu(\ell)}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}} \\ + \frac{1-\theta_{t-1}}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}} \log\frac{1}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}}. \end{aligned}$$

Next, we use properties of logarithm to further simplify this equation:

$$\begin{aligned} D_{KL}(\theta_t||\theta_{t-1}) = &\frac{\mu(\ell)\theta_{t-1}}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}}\log\mu(\ell) \\ &- \frac{\mu(\ell)\theta_{t-1}}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}}\log\left(1-\theta_{t-1}+\mu(\ell)\theta_{t-1}\right) \\ &- \frac{1-\theta_{t-1}}{1-\theta_{t-1}+\mu(\ell)\theta_{t-1}}\log\left(1-\theta_{t-1}+\mu(\ell)\theta_{t-1}\right). \end{aligned}$$

Collect the last two terms, then we get Equation (2.2) before omitting dependence on t:

$$D_{KL}(\theta_t||\theta_{t-1}) = \frac{\mu(\ell)\theta_{t-1}}{1-\theta_{t-1}(1-\mu(\ell))}\log\mu(\ell) - \log\left(1-\theta_{t-1}(1-\mu(\ell))\right).$$

Proof of Lemma 13

Proof. In period $t>0$, each VC is replaced with probability q, so $\mathscr{T}_t(0)=q$. For $\tau\in\{1,2,...,t\}$, the measure of $\tau_t=\tau$ comes from the VCs with last period age $\tau_{t-1}=\tau-1$ not replaced, that is $\mathscr{T}_t(\tau)=(1-q)\mathscr{T}_{t-1}(\tau-1)$. Similarly, the probability of $\tau_t=\tau$ comes from the VCs born in period $t-\tau$ to survive until period t, $\mathscr{T}_t(\tau)=(1-q)^{\tau}\mathscr{T}_{t-\tau}(0)$. Further, because $\mathscr{T}_{t-\tau}(0)=q$ for all $t\geq\tau$, $\mathscr{T}_t(\tau)=q(1-q)^{\tau}$ does not vary across t. Denote $\mathscr{T}$ as steady state distribution of VC ages, then VC with age τ is of mass $\mathscr{T}(\tau)=q(1-q)^{\tau}$. □

2.8.2 Proof of results in benchmark case

Proof of Lemma 14

Proof. Differentiate $D_{KL}(\theta'||\theta)$ with respect to θ to get

$$\frac{dD_{KL}(\theta'||\theta)}{d\theta} = \left(\frac{\mu\log\mu}{1-\theta(1-\mu)}+(1-\mu)\right)\frac{1}{1-\theta(1-\mu)}. \tag{2.9}$$

First note that $1-\theta(1-\mu)>0$, because when $\mu=y_1/y_0>1$, $1-\mu<0$ and when $\mu=(1-y_1)/(1-y_0)<1$, $1-\mu\in(0,1)$. Hence, the right side of Equation 2.9 is positive if and only if

$$\frac{\mu \log \mu}{1-\theta(1-\mu)} + (1-\mu) > 0.$$

Subtract $1-\mu$ from both sides and then multiply by $1-\theta(1-\mu)$, we get

$$\mu \log \mu > -(1-\mu)(1-\theta(1-\mu))$$

Add $1-\mu$ to both sides and then divide $(1-\mu)^2$, we get

$$\frac{\mu \log \mu + (1-\mu)}{(1-\mu)^2} > \theta.$$

The left side is equal to the θ^* as in Equation (2.3). Further, because of the property of logarithm, $1-1/\mu \leq \log \mu \leq \mu - 1$, the θ^* is achieved in the interval $[0,1]$.

Next, we show that the limit of $D_{KL}(\theta'||\theta)$ is 0 when θ converges to 0 or 1:

$$\lim_{\theta \to 0} D_{KL}(\theta'||\theta) = \lim_{\theta \to 0} \frac{\mu\theta}{1-\theta(1-\mu)} \log \mu - \log(1-\theta(1-\mu)) = \mu \log \mu \lim_{\theta \to 0} \theta = 0,$$

$$\lim_{\theta \to 1} D_{KL}(\theta'||\theta) = \lim_{\theta \to 1} \frac{\mu\theta}{1-\theta(1-\mu)} \log \mu - \log(1-\theta(1-\mu)) = \log \mu - \log \mu = 0.$$

Finally, we look at the property of θ^*. Differentiate θ^* with respect to μ, we have

$$\frac{d\theta^*}{d\mu} = \frac{(1-\mu)^2 \log \mu + 2(1-\mu)(\mu \log \mu + (1-\mu))}{(1-\mu)^4} = \frac{\frac{1+\mu}{1-\mu} \log \mu + 2}{(1-\mu)^2}.$$

The numerator increases in μ for $\mu \in (0,1)$, decreases in μ for $\mu > 1$, and is maximized at $\mu = 1$ with

$$\lim_{\mu \to 1} \frac{1+\mu}{1-\mu} \log \mu + 2 = 2 + 2 \lim_{\mu \to 1} \frac{1/\mu}{-1} = 0.$$

Therefore, $d\theta^*/d\mu \leq 0$ for all μ, and θ^* decreases in μ, and

$$\lim_{\mu\to 1}\theta^* = \lim_{\mu\to 1}\frac{\log\mu}{-2(1-\mu)} = \lim_{\mu\to 1}\frac{1/\mu}{2} = 0.5.$$

□

Proof of Lemma 15

Proof. First, take logarithm of both sides of Equation (2.1), and then omit dependence on i and ℓ, we get

$$\log\frac{\theta_t}{1-\theta_t} = \log\mu + \log\frac{\theta_{t-1}}{1-\theta_{t-1}}. \tag{2.10}$$

Define sequence $\{b_t\}_{t=0}^{\infty}$ by $b_t \equiv \log(\theta_t/(1-\theta_t))$. Given θ_0, we can write Eqnation (2.10) as $b_t = \log\mu + b_{t-1}$. Hence, $b_t = b_0 + \sum_{j=0}^{t}\log\mu(\ell_t)$.

Apply the definition of μ, we have

$$\log\mu = \begin{cases} \log\frac{y_1}{y_0}, & \text{with probability } y_x, \\ \log\frac{1-y_1}{1-y_0}, & \text{with probability } 1-y_x. \end{cases}$$

Denote random variable t_x as the number of successful startups, conditional on VC's ability x. By Law of Large Number, as $t\to\infty$, $t_x/t \to y_x$, so

$$\frac{b_t - b_0}{t} = \frac{1}{t}\sum_{j=0}^{t}\log\mu \to y_x\log\frac{y_1}{y_0} + (1-y_x)\log\frac{1-y_1}{1-y_0}. \tag{2.11}$$

Apply property of logorithm that $1-1/z < \log z < z$, we have the following bounds of Equation (2.11). If $x=1$, right side of Equation (2.11) is bounded below by

$$y_1\frac{y_1-y_0}{y_1} - (1-y_1)\frac{y_1-y_0}{1-y_1} = 0.$$

If $x = 0$, right side of Equation (2.11) is above below by

$$y_0 \frac{y_1 - y_0}{y_0} - (1 - y_0) \frac{y_1 - y_0}{1 - y_0} = 0.$$

Therefore, when $t \to \infty$, $b_t \to \infty$ for $x = 1$ and $b_t \to -\infty$ for $x = 0$. Apply $b_t = \log(\theta_t/(1-\theta_t))$, and this implies that θ_t converges to x in probability. □

Proof of Lemma 16

Proof. If VC with age τ has n successful startups, then start from Equation (2.10), we have

$$\log \frac{\theta(\tau,n)}{1-\theta(\tau,n)} = n \log \frac{y_1}{y_0} + (\tau - n) \log \frac{1-y_1}{1-y_0} + \log \frac{\theta(0,0)}{1-\theta(0,0)}$$

Take exponential on both sides,

$$\frac{\theta(\tau,n)}{1-\theta(\tau,n)} = \left(\frac{y_1}{y_0}\right)^n \cdot \left(\frac{1-y_1}{1-y_0}\right)^{(\tau-n)} \cdot \frac{\theta_0}{1-\theta_0}.$$

Rearrange terms and we have Equation (2.4). Finally, the probability of n success out of τ investments follows from Biomial distribution. □

2.8.3 Proof of results in matching with contracts

Lemma 21. *In a stable matching, for any two matched pairs (ℓ, θ) and (ℓ', θ') with payoffs*

$$r = y - w, \;\; u = w + v,$$

wage contracts $w(\ell, \theta)$ and $w(\ell', \theta')$ satisfy

$$w(\ell,\theta) - w(\ell',\theta') \;\; \leq \;\; Y(\ell,\theta) - Y(\ell',\theta) + v(\ell',\theta') - v(\ell',\theta) \quad (2.12)$$

$$w(\ell,\theta) - w(\ell',\theta') \;\; \geq \;\; Y(\ell,\theta') - Y(\ell',\theta') + v(\ell,\theta') - v(\ell,\theta) \quad (2.13)$$

Proof. In a stable matching, we need four incentive conditions for ℓ, ℓ', θ, and θ':

$$\begin{aligned} u(\ell,\theta) \geq u(\ell,\theta') &\implies w(\ell,\theta)+v(\ell,\theta) \geq w(\ell,\theta')+v(\ell,\theta') \\ u(\ell',\theta') \geq u(\ell',\theta) &\implies w(\ell',\theta')+v(\ell',\theta') \geq w(\ell',\theta)+v(\ell',\theta) \\ r(\ell,\theta) \geq r(\ell',\theta) &\implies Y(\ell,\theta)-w(\ell,\theta) \geq Y(\ell',\theta)-w(\ell',\theta) \\ r(\ell',\theta') \geq r(\ell,\theta') &\implies Y(\ell',\theta')-w(\ell',\theta') \geq Y(\ell,\theta')-w(\ell,\theta') \end{aligned}$$

Adding the second and third conditions, we have

$$w(\ell',\theta')+v(\ell',\theta')+Y(\ell,\theta)-w(\ell,\theta) \geq v(\ell',\theta)+Y(\ell',\theta),$$

where $w(\ell',\theta)$ on the right side cancels and we get Inequality (2.12). Similarly, adding the first and last conditions, we have

$$w(\ell,\theta)+v(\ell,\theta)+Y(\ell',\theta')-w(\ell',\theta') \geq v(\ell,\theta')+Y(\ell,\theta'),$$

which yields Inequality (2.13). □

Proof of Proposition 2

Proof. First, the lowest types of workers $L^{(1)} = [0, \underline{\ell}^{(2)})$ are matched with the VC with lowest reputation $\theta^{(1)}$, where $\underline{\ell}^{(2)}$ is determined such that $\mathscr{Q}(\theta^{(1)}) = \mathscr{P}(L^{(1)})$. Let $\ell^{(1)}$ be the average type of worker in interval $L^{(1)}$, then in the VC-preferred stable matching, and wage transfer $w(\ell^{(1)},\theta^{(1)})$ makes $\ell^{(1)}$ indifferent between matched with $\theta^{(1)}$ and unmatched. Setting $u(\ell^{(1)},\theta^{(1)}) = 0$, we get $w(\ell^{(1)},\theta^{(1)}) = -v(\ell^{(1)},\theta^{(1)})$.

The construction of stable matching follows recursively. Given $\{\theta^{(\kappa)}\}_{\kappa=1}^{k-1}$ and $\{L^{(\kappa)}\}_{\kappa=1}^{k-1}$ have been matched with wage payment $w(\ell^{(\kappa)},\theta^{(\kappa)})$, the the lowest reputation of unmatched VCs is $\theta^{(k)}$, and the lowest type of unmatched worker is $\underline{\ell}^{(k)}$. The next interval $L^{(k)} = [\underline{\ell}^{(k)}, \underline{\ell}^{(k+1)})$ is determined such that $\mathscr{P}(L^{(k)}) = \mathscr{Q}(\theta^{(k)})$. Applying Lemma 21, with average worker type

$\ell^{(k)}$, VC-preferred stable matching requires that given $w(\ell^{(k-1)}, \theta^{(k-1)})$, wage payment makes Inequality (2.13) bind. That is,

$$
\begin{aligned}
w(\ell^{(k)}, \theta^{(k)}) = w(\ell^{(k-1)}, \theta^{(k-1)}) \\
+ Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k-1)}, \theta^{(k-1)}) - v(\ell^{(k)}, \theta^{(k)}) + v(\ell^{(k)}, \theta^{(k-1)}).
\end{aligned}
$$

Finally, the last unmatched VCs with reputation $\theta^{(K)}$ are matched with $\ell \in L^{(K)} = [\underline{\ell}^{(K)}, 1]$. This clears the market because $\{L^{(k)}\}_{k=1}^{K}$ are disjoint intervals, $\sum_{k=1}^{K} \mathscr{P}(L^{(K)}) = 1$ and $\sum_{k=1}^{K} \mathscr{Q}(\theta^{(k)}) = 1$. □

Proof of Corollary 1

Proof. For $\ell^{(k)} > \ell^{(k-1)}$, incentive conditions in stable matching requires that

$$
\begin{aligned}
w(\ell^{(k)}, \theta^{(k)}) - w(\ell^{(k-1)}, \theta^{(k-1)}) \\
= Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k-1)}, \theta^{(k-1)}) - v(\ell^{(k)}, \theta^{(k)}) + v(\ell^{(k)}, \theta^{(k-1)}).
\end{aligned}
$$

This is negative if

$$
Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k-1)}, \theta^{(k-1)}) < v(\ell^{(k)}, \theta^{(k)}) - v(\ell^{(k)}, \theta^{(k-1)}). \tag{2.14}
$$

In stable matching, $\mathscr{P}(L^{(k)}) = \mathscr{Q}(\theta^{(k)}) > 0$, so we can divide the left side of Inequality (2.14) by $\mathscr{P}(L^{(k)})$ and divide the right side of Inequality (2.14) by $\mathscr{Q}(\theta^{(k)})$. This gives us

$$
\frac{Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k-1)}, \theta^{(k-1)})}{\mathscr{P}(L^{(k)})} < \frac{v(\ell^{(k)}, \theta^{(k)}) - v(\ell^{(k)}, \theta^{(k-1)})}{\mathscr{Q}(\theta^{(k)})}.
$$

Because $v(\ell, \theta) = \theta v(\ell, 1) + (1-\theta) v(\ell, 0)$, we can further expand Inequality (2.14) to

$$
\mathscr{Q}(\theta^{(k)}) < \frac{\mathscr{P}(L^{(k)})}{Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k-1)}, \theta^{(k-1)})} (\theta^{(k)} - \theta^{(k-1)})(v(\ell^{(k)}, 1) - v(\ell^{(k)}, 0)).
$$

When $\ell^{(k)}$ and $\ell^{(k-1)}$ are close, we use first order Taylor expansion to approximate the fraction on the right side of this inequality. By Assumption 3, $\mathscr{P}(L^{(k)}) \approx p(\ell^{(k)})(\ell^{(k)} - \ell^{(k-1)})$. Therefore, rearrange terms and we have

$$\mathscr{Q}(\theta^{(k)}) < \frac{p(\ell^{(k)})}{dY(\ell^{(k)}, \theta^{(k-1)})/d\ell^{(k)}}(\theta^{(k)} - \theta^{(k-1)})(v(\ell^{(k)}, 1) - v(\ell^{(k)}, 0)).$$

□

Proof of Corollary 2

Proof. Applying Azevedo and Leshno (2016), we have the interval structure of positive assortative matching. It suffices to derive wage payment for each match. First observe that in stable matching, $r(\ell, \theta^{(k)}) = Y(\ell^{(k)}, \theta^{(k)}) - w(\ell^{(k)}, \theta^{(k)})$ for all $\ell \in f(\theta^{(k)})$, so I omit dependence of r in ℓ throughout the proof.

Consider three consecutive values of reputation $\theta^{(k-1)} < \theta^{(k)} < \theta^{(k+1)}$, then we can write $\theta^{(k)}$ as a convex combination of $\theta^{(k-1)}$ and $\theta^{(k+1)}$, that is $\theta^{(k)} = a\theta^{(k-1)} + (1-a)\theta^{(k+1)}$, for some $a \in (0,1)$.

$$\begin{aligned} & ar(\theta^{(k-1)}) + (1-a)r(\theta^{(k+1)}) - r(\theta^{(k)}) \\ = \; & a(Y(\ell^{(k-1)}, \theta^{(k-1)}) - w(\ell^{(k-1)}, \theta^{(k-1)})) \\ & + (1-a)(Y(\ell^{(k+1)}, \theta^{(k+1)}) - w(\ell^{(k+1)}, \theta^{(k+1)})) - (Y(\ell^{(k)}, \theta^{(k)}) - w(\ell^{(k)}, \theta^{(k)})). \end{aligned}$$

Substitute the conditions on $w(\ell^{(k-1)}, \theta^{(k-1)})$, $w(\ell^{(k)}, \theta^{(k)})$ and $w(\ell^{(k+1)}, \theta^{(k+1)})$, and rearrange terms to get

$$\begin{aligned} & a(Y(\ell^{(k)}, \theta^{(k-1)}) - Y(\ell^{(k)}, \theta^{(k)}) + v(\ell^{(k)}, \theta^{(k-1)}) - v(\ell^{(k)}, \theta^{(k)})) \\ & \quad + (1-a)(Y(\ell^{(k+1)}, \theta^{(k+1)}) - Y(\ell^{(k+1)}, \theta^{(k)}) + v(\ell^{(k+1)}, \theta^{(k+1)}) - v(\ell^{(k+1)}, \theta^{(k)})). \end{aligned}$$

Denote $Z(\ell,\theta) = Y(\ell,\theta) + v(\ell,\theta)$, then $Z(\ell,\theta) = \theta Z(\ell,1) + (1-\theta)Z(\ell,0)$. Therefore, this value becomes

$$\begin{aligned} a(\theta^{(k-1)}Z(\ell^{(k)},1) + (1-\theta^{(k-1)})Z(\ell^{(k)},0) \\ -\theta^{(k)}Z(\ell^{(k)},1) - (1-\theta^{(k)})Z(\ell^{(k)},0)) + (1-a)(\theta^{(k+1)}Z(\ell^{(k+1)},1) \\ + (1-\theta^{(k+1)})Z(\ell^{(k+1)},0) - \theta^{(k)}Z(\ell^{(k+1)},1) - (1-\theta^{(k)})Z(\ell^{(k+1)},0)). \end{aligned}$$

Collect terms and apply $\theta^{(k)} = a\theta^{(k-1)} + (1-a)\theta^{(k+1)}$, we get

$$\begin{aligned} & a(1-a)(\theta^{(k-1)} - \theta^{(k+1)})(Z(\ell^{(k)},1) - Z(\ell^{(k)},0)) \\ & \quad + (1-a)a(\theta^{(k+1)} - \theta^{(k-1)})(Z(\ell^{(k+1)},1) - Z(\ell^{(k+1)},0)) \\ & = a(1-a)(\theta^{(k+1)} - \theta^{(k-1)})(Z(\ell^{(k+1)},1) - Z(\ell^{(k)},1) - (Z(\ell^{(k+1)},0) - Z(\ell^{(k)},0))). \end{aligned}$$

This is positive because $\theta^{(k+1)} > \theta^{(k-1)}$ and $Z(\ell^{(k+1)},1) - Z(\ell^{(k)},1) > Z(\ell^{(k+1)},0) - Z(\ell^{(k)},0)$ by the assumption of supermodularity. The convexity of r follows from $ar(\theta^{(k-1)}) + (1-a)r(\theta^{(k+1)}) > r(a\theta^{(k-1)} + (1-a)\theta^{(k+1)})$. □

2.8.4 Proof of Lemma 18

Proof. I will show that $\{Q_t(\theta)\}$ is a Cauchy sequence for each t and apply the result that Cauchy sequence converges uniformly. First, note that for any t, $Q_{t+1}(\theta)$ and $Q_t(\theta)$ assign identical mass to $\theta \in \{\theta_0\} \cup \{\{\theta_\tau^{H(k)}, \theta_\tau^{L(k)}\}_{k=1}^{K(\tau)}\}_{\tau=0}^{t-1}$. Further, $\mathcal{Q}_t(\theta)$ is bounded above by $(1-q)^t q$. Hence $|Q_{t+1}(\theta) - Q_t(\theta)|$ is bounded above by $\max\{Q_t(\theta_t^{H(k)}), Q_t(\theta_t^{L(k)}), Q_{t-1}(\theta_{t-1}^{(k)})\} < (1-q)^{t-1}q$. For all $\varepsilon > 0$, there exists $T \equiv \lceil (\log(\varepsilon/q))/\log(1-q) \rceil + 1$, such that for all $t > T$, $|Q_{t+1}(\theta) - Q_t(\theta)| < \varepsilon$ for all $\theta \in [0,1]$. □

2.8.5 Proof of results in three cases of production technology

Proof of Proposition 3

Proof. Differentiate $D_{KL}(\theta'||\theta)$ to get

$$dD_{KL}(\theta'||\theta) = \frac{(1-\theta)\theta\log\mu(\ell)}{(1-\theta(1-\mu(\ell)))^2}d\mu(\ell) \\ + \left(\frac{\mu(\ell)\log\mu(\ell)}{1-\theta(1-\mu(\ell))} + (1-\mu(\ell))\right)\frac{d\theta}{1-\theta(1-\mu(\ell))}$$

First, for fixed $\mu(\ell)$,

$$\frac{dD_{KL}(\theta'||\theta)}{d\theta} = \left(\frac{\mu(\ell)\log\mu(\ell)}{1-\theta(1-\mu(\ell))} + (1-\mu(\ell))\right)\frac{1}{1-\theta(1-\mu(\ell))}$$

which is positive if and only if $\theta < \theta^*$. Further, because of the property of logarithm, $1-1/\mu \leq \log\mu \leq \mu - 1$, the θ^* is achieved in the interval $[0,1]$.

Second, for fixed θ, the sign of

$$dD_{KL}(\theta'||\theta) = \frac{(1-\theta)\theta\log\mu(\ell)}{(1-\theta(1-\mu(\ell)))^2}d\mu(\ell)$$

is the same as the sign of $(\log\mu)d\mu$.

$$\frac{d\mu(\ell)}{d\ell} = \begin{cases} \frac{y_0(\ell)dy_1(\ell)-y_1(\ell)dy_0(\ell)}{y_0(\ell)^2} & \text{when } \mu(\ell) = \frac{y_1(\ell)}{y_0(\ell)} > 1, \\ -\frac{(1-y_0(\ell))dy_1(\ell)-(1-y_1(\ell))dy_0(\ell)}{(1-y_0(\ell))^2} & \text{when } \mu(\ell) = \frac{1-y_1(\ell)}{1-y_0(\ell)} < 1. \end{cases}$$

This then lead to the three cases in the proposition:

1. If $\frac{dy_1}{dy_0} > \frac{y_1}{y_0} > \frac{1-y_1}{1-y_0}$, then for $\mu = y_1/y_0$, $d\mu/d\ell > 0$ and $\log\mu > 0$, and for $\mu = (1-y_1)/(1-y_0))$, $d\mu/d\ell < 0$ and $\log\mu < 0$.

2. If $\frac{y_1}{y_0} > \frac{1-y_1}{1-y_0} > \frac{dy_1}{dy_0}$, then for $\mu = y_1/y_0$, $d\mu/d\ell < 0$ and $\log\mu > 0$, and for $\mu = (1-y_1)/(1-y_0))$, $d\mu/d\ell > 0$ and $\log\mu < 0$.

3. If $\frac{y_1}{y_0} > \frac{dy_1}{dy_0} > \frac{1-y_1}{1-y_0}$, then for $\mu = y_1/y_0$, $d\mu/d\ell < 0$ and $\log\mu > 0$, and for $\mu = (1-y_1)/(1-y_0))$, $d\mu/d\ell < 0$ and $\log\mu < 0$.

□

Proof of Lemma 19

Proof. Begin by finding $\frac{y_1}{y_0}$, $\frac{1-y_1}{1-y_0}$, and $\frac{dy_1}{dy_0}$. Because

$$\frac{dy_x}{d\ell} = 0.2 \times 2^{-1/\rho}(1+\ell)^{\rho-1}\left((1+x)^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho},$$

we have

$$\frac{\frac{dy_1}{d\ell}}{\frac{dy_0}{d\ell}} = \frac{\left(2^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}}{\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}}$$

and

$$\frac{y_1}{y_0} = \frac{\left(2^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}{\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}, \quad \frac{1-y_1}{1-y_0} = \frac{5\times 2^{1/\rho}-\left(2^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}{5\times 2^{1/\rho}-\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}.$$

Note that $y_1/y_0 > (1-y_1)/(1-y_0)$ because $0 < y_0 < y_1 < 1$, so I will only compare $\frac{dy_1}{dy_0}$ with the two fractions.

On one hand, $\frac{dy_1}{dy_0} > \frac{y_1}{y_0}$ if and only if

$$\frac{\left(2^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}}{\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}} > \frac{\left(2^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}{\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}.$$

Take logarithm of both sides and collect terms to get

$$\log\frac{\left(2^{\rho}+(1+\ell)^{\rho}\right)}{\left(1^{\rho}+(1+\ell)^{\rho}\right)} < 0,$$

which is equivalent to $2^{\rho} < 1$ and hence $\rho < 0$.

On the other hand, $\frac{dy_1}{dy_0} > \frac{1-y_1}{1-y_0}$ if and only if

$$\frac{\left(2^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}}{\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}} > \frac{5\times 2^{1/\rho}-\left(2^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}{5\times 2^{1/\rho}-\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}}.$$

Multiply both sides by $\left(5\times 2^{1/\rho}-\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1/\rho}\right)\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}$. Collect terms and obtain

$$\begin{aligned} 5\times 2^{1/\rho}&\left(\left(2^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}-\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}\right)\\ &> \left(2^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}\left(1-2^{\rho}\right). \end{aligned}$$

Divide both sides by $5\times 2^{1/\rho}\times\left(2^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}\left(1^{\rho}+(1+\ell)^{\rho}\right)^{-1+1/\rho}$ and get

$$\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1-1/\rho}-\left(2^{\rho}+(1+\ell)^{\rho}\right)^{1-1/\rho} > \frac{1-2^{\rho}}{5\times 2^{1/\rho}}.$$

This is always satisfied for $\rho\in(0,1]$, with the right side negative and left side positive. Define function $B(\rho,\ell)\equiv\left(1^{\rho}+(1+\ell)^{\rho}\right)^{1-1/\rho}-\left(2^{\rho}+(1+\ell)^{\rho}\right)^{1-1/\rho}-\frac{1-2^{\rho}}{5\times 2^{1/\rho}}$, When $\rho>1$, $B(\rho,\ell)$ increases in ℓ, so B is bounded above and below by $B(\rho,1)$ and $B(\rho,0)$. Therefore, a sufficient condition for $\frac{dy_1}{dy_0} > \frac{1-y_1}{1-y_0}$ for all $\ell\in[0,1]$ is $\min_{\ell} B(\rho,\ell)=B(\rho,0)\geq 0$. Further, because $B(\rho,\ell)$ increases in ρ for any fixed $\ell\in[0,1]$, this condition translates into $\rho\geq\tilde{\rho}$, where $\tilde{\rho}$ satisfies $B(\tilde{\rho},0)=0$. Similarly, a sufficient condition for $\frac{dy_1}{dy_0} < \frac{1-y_1}{1-y_0}$ for all $\ell\in[0,1]$ is $\max_{\ell} B(\rho,\ell)=B(\rho,1)\leq 0$, and this is equivalent to $\rho\leq\hat{\rho}$, where $\hat{\rho}$ satisfies $B(\hat{\rho},1)=0$. To complete the proof, it suffices to show that $1<\tilde{\rho}<\hat{\rho}<2$. One can prove it analytically, but solving equations $B(\tilde{\rho},0)=0$ and $B(\hat{\rho},1)=0$ analytically yields $\tilde{\rho}$ and $\hat{\rho}$ are approximately 1.3339 and 1.5472, which completes the proof. □

2.8.6 Proof of Lemma 20

Proof. Denote VC's continuation value starting from reputation θ as

$$G(\theta) = r(\ell,\theta) + \delta(y(\ell,\theta)G(\theta_H(\ell,\theta)) + (1-y(\ell,\theta))G(\theta_L(\ell,\theta)))$$

where $r(\ell,\theta) = Y(\ell,\theta) - w(\ell,\theta)$ and $u(\ell,\theta) = w(\ell,\theta) + v(\ell,\theta)$.

Given a positive associative equilibrium, suppose in some period, VC deviates to low worker ℓ^*, then it must be that ℓ^* is not worse off

$$u(\ell^*,\theta^*) = w(\ell^*,\theta^*) + v(\ell^*,\theta^*) \leq w(\ell^*,\theta) + v(\ell^*,\theta) = u(\ell^*,\theta)$$

wage satisfies $u(\ell^*,\theta^*) - v(\ell^*,\theta) \leq w(\ell^*,\theta)$, current rate of return becomes

$$r(\ell^*,\theta) = Y(\ell^*,\theta) - w(\ell^*,\theta) \leq Y(\ell^*,\theta) - (u(\ell^*,\theta^*) - v(\ell^*,\theta))$$

If instead hire VC θ and worker ℓ. Profit

$$Y(\ell,\theta) - Y(\ell^*,\theta) + (u(\ell^*,\theta^*) - v(\ell^*,\theta)) - (u(\ell,\theta) - v(\ell,\theta))$$

which is greater than (when $\ell^* = f(\theta)$)

$$Y(\ell,\theta) - Y(\ell^*,\theta) + (u(\ell^*,\theta) - v(\ell^*,\theta)) - (u(\ell,\theta) - v(\ell,\theta)) = 0$$

Hence $r(\ell^*,\theta) < r(\ell,\theta)$. If lower ℓ^* gives VC better continuation value, then at least continuation value gain covers current period losses

Next show that not only current period payoff, but also next period expected payoff increases in current labor input.

$$y(\ell^*,\theta)r(\theta_H(\ell^*,\theta))+(1-y(\ell^*,\theta))r(\theta_L(\ell^*,\theta))<y(\ell,\theta)r(\theta_H(\ell,\theta))+(1-y(\ell,\theta))r(\theta_L(\ell,\theta))$$

Note from analysis in section 3, we have that for $\theta<\theta'<\theta''$ and $\theta'=a\theta+(1-a)\theta''$,

$$\begin{aligned}&ar(\theta)+(1-a)r(\theta'')-r(\theta)\\&\qquad=a(1-a)(\theta''-\theta)(Z(\ell'',1)-Z(\ell,1)-(Z(\ell'',0)-Z(\ell,0))).\end{aligned}$$

On the other hand, for all ℓ and θ, we have $\theta=y(\ell,\theta)\theta_H(\ell,\theta)+(1-y(\ell,\theta))\theta_L(\ell,\theta)$ and

$$\theta_H(\ell,\theta)-\theta_L(\ell,\theta)=\frac{\theta(1-\theta)(y_1(\ell)-y_0(\ell))}{y(\ell,\theta)(1-y(\ell,\theta))}.$$

Set $a=y(\ell,\theta)$ and $a^*=y(\ell^*,\theta)$, we have

$$\begin{aligned}&y(\ell,\theta)r(\theta_L(\ell,\theta))+(1-y(\ell,\theta))r(\theta_H(\ell,\theta))-r(\theta)\\&\quad=y(\ell,\theta)(1-y(\ell,\theta))(\theta_H(\ell,\theta)-\theta_L(\ell,\theta))(Z(\ell_H,1)-Z(\ell_L,1)-(Z(\ell_H,0)-Z(\ell_L,0)))\\&\qquad=\theta(1-\theta)(y_1(\ell)-y_0(\ell))(Z(\ell_H,1)-Z(\ell_H,0)-(Z(\ell_L,1)-Z(\ell_L,0))),\end{aligned}$$

and

$$\begin{aligned}&y(\ell^*,\theta)r(\theta_L(\ell^*,\theta))+(1-y(\ell^*,\theta))r(\theta_H(\ell^*,\theta))-r(\theta)\\&\quad=y(\ell^*,\theta)(1-y(\ell^*,\theta))(\theta_H(\ell^*,\theta)-\theta_L(\ell^*,\theta))(Z(\ell^*_H,1)-Z(\ell^*_L,1)-(Z(\ell^*_H,0)-Z(\ell^*_L,0)))\\&\qquad=\theta(1-\theta)(y_1(\ell^*)-y_0(\ell^*))(Z(\ell^*_H,1)-Z(\ell^*_H,0)-(Z(\ell^*_L,1)-Z(\ell^*_L,0))).\end{aligned}$$

Further, $\ell^*<\ell$ implies $\theta_L(\ell,\theta)<\theta_L(\ell^*,\theta)<\theta_H(\ell^*,\theta)<\theta_H(\ell,\theta)$. In equilibrium, this also implies $\ell_L<\ell^*_L<\ell^*_H<\ell_H$. Sufficient condition for $y(\ell,\theta)r(\theta_L(\ell,\theta))+(1-y(\ell,\theta))r(\theta_H(\ell,\theta))>$

$y(\ell^*, \theta) r(\theta_L(\ell^*, \theta)) + (1 - y(\ell^*, \theta)) r(\theta_H(\ell^*, \theta))$ is that

1. $y_1(\ell) - y_0(\ell)$ increases in ℓ and $d(Z(\ell, 1) - Z(\ell, 1)))/d\ell$ increases in ℓ, or

2. $y_1(\ell) - y_0(\ell)$ decreases in ℓ and $d(Z(\ell, 1) - Z(\ell, 1)))/d\ell$ decreases in ℓ.

□

Acknowledgements

Chapter 2, in part is currently being prepared for submission for publication of the material. The dissertation author was the author of this material.

Chapter 3

Starting Small in Business Lending: Evidence from U.S. Small Business Data

3.1 Introduction

The bank-firm relationships often begin with asymmetric information, where the bank is unsure of the firm's riskiness, and they build their relationships through repeated loans. In the relational contract literature, the starting-small model matches the relationship lending closely. When the bank-firm relationship starts with a small loan and gradually increases only if the firm repays all loans on time, the bank can monitor and learn about the firm's private information.

To effectively examine the relational contract dynamics between borrowers and lenders, it is essential to have diversity among lenders and borrowers while ensuring comparability of loans. Utilizing Small Business Administration (SBA) datasets proves advantageous for such analysis as it meets these criteria: small businesses of similar profiles avail loans through the SBA 7(a) program from lenders nationwide. Moreover, the focus extends to understanding the relationship dynamics amidst interruptions. The SBA datasets offer an opportunity to investigate the effects of interruptions, such as those induced by the COVID-19 pandemic and the implementation of the Paycheck Protection Program (PPP), on the borrower-lender relationship. Notably, there exists overlap between borrowers and lenders participating in both programs, facilitating a comprehensive examination of relational dynamics across loan contexts.

The main findings of this paper are three impacts of repeated loans. First, to test the model predictions about credit risks, using the SBA 7 (a) data, I implement the survival analysis and obtain evidence that repeated borrowers are less riskier than one-time borrowers, which is consistent with the model.

Second, this project tackles the effect of repeated borrowing on interest rates. The linear regression shows that the repeated loans on average receive loans with lower interest rates. However, according to synthetic control and difference-in-difference methods, such correlation is caused by the selection effect. The borrowers with lower risks choose to have the repeated loans, and refinances it when interest rates are low. During low-interest periods, the new borrowers can get lower interest rates than existing borrowers, because bank uses extracted information to take advantage of existing borrowers.

Third, during the COVID-19 pandemic, the Paycheck Protection Program provides larger loan amounts to new borrowers than to existing 7 (a) borrowers. Finally, in the analysis of the loans, I examine the impact of distance between borrower and lender as a proxy of the severity of information asymmetry. The empirical results show that the farther borrower and lender are, the more severe the information asymmetry is, and the results are consistent with theory.

This paper contributes to the literature in three ways. First, the theory model extends the starting-small model in an institutional setting, and the internally consistent renegotiation provides a thorough analysis of loan renegotiation under crisis. Second, the empirical analysis characterizes the impact of repeated lending using loan-level data, including during an economic downturn. Third, the results shed light on the ratchet effect in the business lending setup. When the loan starts small, the firm's private information is gradually revealed, which allows the bank to extract more surplus during follow-up loans and economic downturns.

The related literature includes starting-small theory, relationship lending, and SBA loan empirics. The model of this paper is mainly based on Hua and Watson (2022) and Watson (1999), and the related literature on starting small can be found in these papers. The empirical literature on relationship lending is rich. For example, the dynamics of collateral (Boot, Thakor, and Udell 1991), initial loan amount (Chang and Sundaresan 2001), control rights (Freudenberg et al. 2017), and credit lines (Sannikov 2007) have been well-documented. Craig, Jackson III, and Thomson (2009) reviewed recent literature on SBA guaranteed loans, and most papers study the economic efficiency of the SBA lending program, such as economic growth, employment enhancement, allocation efficiency, government intervention, discrimination, and the impact of the financial crisis. Two papers that study credit risks using this dataset are Glennon and Nigro (2005), which follows a survival analysis approach, and DeYoung, Glennon, and Nigro (2008), which incorporates the distance between borrower and lender as a factor of the severity of information asymmetry.

This paper is related to SHARPE (1990) by investigating how banks capitalize on the informational advantages accrued from historical bank-firm interactions to extract economic rents, resembling the implicit contracts prevalent in relationship banking. However, my study extends this analysis by introducing unanticipated productivity shocks and formalizes the renegotiation dynamics within the relationship, thus enhancing the understanding of how banks navigate asymmetric information and adapt to exogenous shocks. Moreover, my work shares similarities with RAJAN (1992) in its examination of lending contract structures. While Rajan primarily explores the ramifications of bank control rights on firm effort, my model delves into the role of loan contracts in efficiently selecting and screening borrowers with private information.

This paper is organized as follows. Section 3.2 outlines the model predictions and their intuition. Section 3.3 describes the dataset and empirical approach, and Section 3.4 presents the main results on the impact of repeated loans. The starting-small model, definition of variables, tables, and figures are provided in the Appendix.

3.2 Model Overview

In this section, I summarize the model of repeated loans and its predictions and leave the detailed analysis in Appendix 3.6.1. Applying the starting-small concept in the relational contract literature, I model the bank-firm relationship with a discrete-time partnership game with one-sided private information, wherein an endogenous sequence of levels establishes the scale of the game.

Consider a bank lending to a firm with private information about its riskiness. The bank and firm expect a potential long-term relationship, and the firm hopes to expand its operation to a size that requires a substantial loan amount. Instead of granting the entire loan, the bank proposes a scale of a loan α_t, which is a multiplicative factor that applies to the loan amount, collateral, and project return. In each period, after the bank proposes the loan size, the firm either cooperates by investing in the project properly or shirks by taking the money and not investing. At the end of the period, the firm uses the realization of the project return to repay the loan, where the default event serves as a signal about the firm's type.

A firm's type is defined as its riskiness. I assume that the firm's return distribution is symmetric regardless of the type. I further adopt the assumption of Stiglitz and Weiss (1981) that the high-risk firm's return is a mean-preserving spread of a lower-risk firm's return. The assumption is meant to simplify the analysis to a one-to-one correspondence of riskiness and probability of default.

In this model, the loan sequence proposed by the bank separates the firm's different types over time. For a given history of the game, the bank learns about the firm's riskiness θ. In period t, the loan sequence incentivizes the firm with risk lower than θ_t to cooperate, so that the types between θ_t and θ_{t-1} shirks in period t. For this incentive condition to hold in all periods, the level sequence should increase over time.

The probability of default increases in a firm's type, which makes shirking more attractive for the higher-risk firms. When the level is lower than α_t, a firm with type θ_t gains by waiting until the level gets higher; when the level is higher than α_t, the firm of this type shirks and stops investing.

The incentive-compatible sequence of loans and the bank's belief updating result leads to the first hypothesis: riskier types shirk early and repeated borrowers tend to have lower risks, corresponding to the lower default risk. That is,

Hypothesis 1. *Repeated loans have a lower hazard rate.*

Furthermore, the speed of learning (α_t increasing and θ_t decreasing) is determined by the severity of information asymmetry. I adopt the distance between the firm and the bank as a proxy of the severity of information asymmetry: the farther away the firm and bank are apart, the more severe the information asymmetry is. The following hypothesis summarizes this result. While it is conceivable that banks may adopt a more stringent selection criterion for firms situated at a distance (favoring only the most creditworthy), it remains plausible that high-quality firms could have opted to borrow from nearby banks, especially considering that SBA lenders encompass a significant portion of the US market, potentially offering more favorable interest rates. Additionally, firms located farther away from lenders might exhibit diminished relational and reputational concerns in the event of default. [1]

Hypothesis 2. *The hazard rate positively correlates with the distance between the firm and the bank.*

In the baseline model, the interest rate is not treated as a free variable to be chosen by the bank. Instead, it is assumed to be pinned down by the exogenous economic conditions and the loan contract terms. This simplification is with loss but has little impact on the result of this paper. First, in the 7 (a) program of SBA, the bank has only little, if any, control over interest rates. The 7 (a) program allows lenders to determine interest rates subject to one of three base rates and an upper bound on spread. [2]

Second, even though the model does not explicitly describe the determination of interest rates, it is still reasonable to have the interest rates correlated with the pool of borrowers. When t increases, the decreasing information severity of the riskiness of borrowers should cause the

1. Tian (2011) uses the same approach to measure the information asymmetry between venture capitalists and entrepreneurs in a stage financing contract. Moreover, Liberti and Petersen (2018) also argues that the smaller geographical distance between borrower and lender helps collect soft information.

2. According to SBA (2021c), the 3 Acceptable Base Rates are Prime rate published in a daily national newspaper, London Interbank One Month Prime plus 3%, and SBA Peg Rate. The Maximum Allowable Spread is 2.25% for Maturity less than 7 years and 2.75% for Maturity greater than 7 years.

variance of of interest rates to decrease. Third, because this model assumes the bank to have full bargaining power, then with the revelation of types, the bank can extract more surplus of the firms, so the average interest rate should be higher.

Hypothesis 3. *After repetition, the mean of interest rate increases.*

Following the baseline model, an extension is considered. During an economic downturn, even though the original contract is renegotiation-proof, an unexpected productivity shock forces the borrower and lender to renegotiate. According to an internally consistent renegotiation, the existing borrowers will get a loan with a lower level than past loans, while new borrowers will obtain a larger loan amount than existing borrowers. The following two hypotheses summarize this result.

Hypothesis 4. *Under an economic downturn, existing contracts stall at a low level.*

Hypothesis 5. *The first-time borrowers can obtain a larger loan amount than repeated borrowers during an economic downturn.*

Hypothesis 3 and 5 are counter-intuitive but consistent with the ratchet effect literature (Freixas, Guesnerie, and Tirole (1985), Ickes and Samuelson (1987), Charness, Kuhn, and Villeval (2011), Bhaskar et al. (2014), and Gerardi and Maestri (2020)). When the relationship of borrower and lender grows, the lender can learn about the lender's type and hence extract more surplus. Given that borrowers place considerable value on their relationship with lenders, they often exhibit reluctance to switch lenders or be grouped with first-time borrowers. Recognizing this dynamic, banks can strategically leverage their relational capital to prioritize other first-time borrowers, thereby enticing new business and fostering stronger lender-borrower relationships.

3.3 Empirical Setup

3.3.1 SBA 7 (a) loan and PPP

To help small businesses, the SBA offers the 7(a) Loan Program, with a maximum loan amount of $5 million, monthly repayment, fixed or variable interests, maturity up to 25 years, and an SBA guarantee of $75\% - 85\%$ of the loan amount. SBA provides borrowers two alternatives of 7 (a): the 504 loans and microloans.[3] The 7 (a) program is "SBA's most common loan program"; the use of 7 (a) loans is less restrictive: the 504 loans cannot be used as working capital and the microloans cannot be used to repay existing debts, while the 7 (a) loans do not have such restrictions.

After the outbreak of the Covid-19 pandemic, the Coronavirus Aid, Relief, and Economic Security Act, or the CARES Act, was passed on March 27, 2020.[4] Included in the CARES Act are the Paycheck Protection Program (PPP), COVID-19 Economic Injury Disaster Loans (EIDL), Shuttered Venue Operators Grant (SVO Grant Program), SBA Express Bridge Loan, and SBA Debt Relief. [5] According to the SBA COVID Relief Program Report (2021b), PPP and EIDL take the most Covid relief loans; PPP loans are short-term, low-interest, un-collateralized, whereas the EIDL loans are long term, high interest and collateralized.

As the model studies the rolling short-term loans, the data used to test the hypothesis are from the 7 (a) program and PPP.

3.3.2 Data

The U.S. Small Business Administration (SBA) discloses 1999 to present 7(a) loan data, which contains information about the terms and status of each loan, as well as borrower and lender characteristics. I retrieved post-2010 SBA 7(a) loan data, including 633,298 loans approved between Oct 1, 2008, and Dec 31, 2020 (146 months).

3. Details see https://www.sba.gov/funding-programs/loans.
4. Details see https://home.treasury.gov/policy-issues/cares.
5. Details see https://www.sba.gov/funding-programs/loans/coronavirus-relief-options

According to the 7 (a) data, 50,242 of the 574,453 borrowers repeatedly borrowed via the 7(a) program, and times of repeated dealing range from 2 to 31. Furthermore, 26,089 out of all 633,298 loans are charged off, and 2,691 out of 109,087 repeated loans are charged off conditional on repeated borrowing. Summary statistics show that the default rate conditional on repeated interactions (2.47%) is less than the unconditional default rate (4.12%). The summary statistics of the dataset are presented in Table 3.1.

The SBA also discloses all Paycheck Protection Program (PPP) data. I retrieved PPP data with borrowers in California and loan amounts up to $\$150,000$, including 602,088 loans approved from April 3, 2020, to January 31, 2021. According to the PPP data, 777 borrowers borrow from both loan programs.

3.3.3 Empirical Strategy

In this project, I will answer three questions based on the empirical tests: (I) How does repeated lending affect credit risk? (ii) How does the borrower's history affect the loan interest rates? and (iii) What is the role of the borrower-lender relationship in the pandemic? Based on the starting small theory, repeated dealing decreases credit risk, and repeated borrowing is associated with lower risk, which implies a lower interest rate.

The first question takes the loan contracts as given, and tackles the production realization of the last stage. The second question then considers the negotiation stage of the lender and borrower. In a subgame perfect equilibrium, the first stage decision is made in anticipation of the second stage realizations. The third question tackles the extension of the model: the renegotiation under unexpected productivity shock.

To answer the first question, I adopt the survival analysis approach and implement the Kaplan-Meier estimation and Cox Proportional Hazard Model. In the Kaplan-Meier estimation, the effect of repeated dealing on credit risk is captured by the gap between the estimated survival rate of the full sample and the conditional survival rate for the repeated borrower. The parameter of interest in the Cox Proportional Hazard Model is the coefficient estimate of the indicator variable of the repeated borrower (IndRepeat). As predicted by the theory, the hazard rate of loans conditional on repeated borrowing should be lower than the loans in population, and the coefficient estimate of IndRepeat should be significantly negative. Section 3.4.1 answers this question and tests Hypotheses 1 and 2.

The second question is intended to study the effect of repeated interactions on borrowing costs and interest rates. I implement two approaches to investigate the second question. The first approach is to regress the interest rates on the indicators for repeated borrowers and the borrowers repeat more than twice. The second approach is the difference-in-difference regression, and the parameter of interest is the coefficient estimate of the cross-term of time indicator and treatment indicator, which is predicted to be negative by theory. Section 3.4.2 answers this question and tests Hypothesis 3.

The third question uses both 7 (a) and PPP datasets. I merged the two datasets and filtered them to get the data for California. I regress the loan amount in PPP on an indicator variable of whether the borrower is in the 7 (a) program. Section 3.4.3 answers this question and tests Hypotheses 4 and 5.

3.4 Results

3.4.1 The Impact of Repeated Lending on Credit Risks

From a bank's perspective, credit risk is one of the biggest concerns in evaluating loan portfolio performance. I adopt the survival analysis approach to study the credit risk of SBA loans. In this section, I will test Hypothesis 1 and 2 by the Kaplan-Meier Estimate and Cox Proportional Hazard Model. In this part of the analysis, the loan contracts are taken as given, and we focus on the default probability conditional on the loan covariates. Because the time range of this dataset is limited, and many loans are neither paid in full nor charged off at the end of the period of this study, I will use the censoring method for both approaches.

Kaplan-Meier Estimate

I first use the Kaplan-Meier estimate to find the non-parametric estimates of the survival rate. Figure 3.1 shows the hazard rate estimates for the full sample, a subsample of all repeated borrowers (IndRepeat = 1), and a subsample of the repeated loans of repeated borrowers (NumRepeat $\geq$2).

Based on this non-parametric estimate, firms with repeated dealings are significantly less risky than the full population, and the subsample with NumRepeat $\geq$2 is only slightly less risky than the subsample of IndRepeat = 1. This observation implies that borrowers with different risk levels are screened for different types of contracts, and there is not enough evidence to conclude the repeated transaction itself has a significant impact on the credit risks.

Cox Proportional Hazard Model

To obtain a more quantitative result, I then use the Cox Proportional Hazard Model to estimate the credit risk with covariates. I run a short regression (3.1) and three long regressions to investigate the effect of repeated loans and distance between borrower and lender.

In the short regression, I use loan-specific characteristics to predict the hazard rate of each loan. The regression equation is

$$h^{(1)}(X_i,t) = h_0^{(1)}(t)\exp(\hat{\gamma}^{(1)}X_i)), \tag{3.1}$$

where $h^{(1)}(X_i,t)$ is the hazard rate for loan i at time t, $h_0^{(1)}(t)$ is the baseline hazard rate, where all covariates of being 0, and X_i is loan specific covariates, including gross approval amount, percentage guaranteed, term in months, initial interest rate and jobs supported.

In the first and second long regression, I add one explanatory variable to the short regression, *IndRepeat* and *ZDis* respectively, and the regression equations become

$$h^{(2)}(X_i,t) = h_0^{(2)}(t)\exp(\hat{\beta}_1 IndRepeat_i + \hat{\gamma}^{(2)}X_i)), \tag{3.2}$$

and

$$h^{(3)}(X_i,t) = h_0^{(2)}(t)\exp(\hat{\beta}_2 ZDis_i + \hat{\gamma}^{(3)} X_i)), \tag{3.3}$$

where $IndRepeat_i$ is the indicator variable of the repeated borrower, *ZDis* is the distance between borrower and lender, and the rest of the notations are the same as (3.1).

Finally, I include both *IndRepeat* and *ZDis* as explanatory variables and get regression (4). The regression results are presented in Table 3.2. The parameter of interest $\hat{\beta}_1$ and $\hat{\beta}_2$ are significant as shown in Table 3.2. There are three main findings.

The main findings are threefold. First, negative coefficient estimates of $\hat{\beta}_1$ in both regression (2) and regression (4) in Table 3.2 provide evidence of Hypothesis 1 that repeated borrowers are significantly less risky. Second, positive coefficient estimates of $\hat{\beta}_2$ in both regression (3) and regression (4) in Table 3.2 provide evidence of Hypothesis 2.

Third, even though all coefficient estimates are significant, *IndRepeat* has more explanatory power than *ZDis*. The log-likelihood tests, comparing the short regression and long regressions, show that the regression (2) and (3) in Table 3.2 fit data better, which implies that both the indicator of repeated loans and distance between borrower and lender have explanatory power for hazard rate. However, comparing the log-likelihood of regression (2) and (3) in Table 3.2, *IndRepeat* adds more explanatory power than *ZDis*.

3.4.2 The Impact of Repeated Lending on Interest Rates

From the perspective of a borrower, the cost of borrowing is the most important factor in evaluating a loan. From the perspective of a lender, interest rates reflect the lender's belief about the riskiness of the borrower. The result of the previous section also motivates the analysis of the influence of repeated borrowing on the interest rate.

In the first approach, I analyze the impact of repeated lending on interest rates by a simple linear regression. Next, I construct the loan amount weighted average of interest rates at the issuance month as a proxy of market interest rates and include the market rates as an additional control variable that captures the time-fixed effects.

In the second approach, I first use the synthetic control method to construct counterfactual control observation for each repeated loan observation. Then use the difference in difference method to estimate the effect of treatment (repeated borrowing). I also use the difference in difference regression with the constructed control group and treated group with repeated transactions to study the effect of treatment. Details of synthetic control and difference-in-difference results are presented in section 3.4.2 and 3.4.2. In this section, I restrict attention to the California subset of the SBA 7 (a) program.

Market Rate and Determination of Interest Rates

To study the determination of interest rates, I first construct a proxy for 7 (a) program average interest rates: the interest rate in month t is the loan amount weighted average interest rate, using the loans approved in month t. Figure 3.2 visualizes the cross-sectional data of the 7 (a) program interest rates.

Table 3.3 presents the regression results using the loan interest rates as dependent variables. In this table, regressions (1), (2), and (5) use the full sample, while regressions (3) and (4) use the subsample from the sample before the second interest rate hike. The coefficients of interest are for the indicator variable of the repeated borrower, $IndRepeat$, and the indicator variable of the borrower repeats more than twice, $NumRepInd$. The baseline model is when these two indicators are both zero, which is the one-time borrowers. When $IndRepeat = 1$ and $NumRepInd = 0$, the model is for the loan of two-time borrowers; when $IndRepeat = 1$ and $NumRepInd = 1$, the model is for the repeated borrowers with at least three loans.

The first result from Table 3.3 is that the coefficient of $IndRepeat$ is negative in all regressions, which indicates that repeated borrowers receive lower interest rates compared with one-time borrowers. Moreover, because the coefficient of $NumRepInd$ is negative in all regressions, Table 3.3 shows that the one-time borrowers get the highest interest rates, and the multiple-time borrowers get the lowest interest rates.

Second, the farther the distance between borrower and lender is, the higher the interest rates are. This result further provides evidence of the relationship between information asymmetry and geographical distance. Finally, the higher interest rate is correlated with a shorter term, fewer jobs supported, revolved loans, smaller amounts of SBA guarantees, smaller loan amounts, and higher market interest rates.

To sum up, the repeated borrowers receive lower interest rates on average. However, this correlation is not necessarily causality: such a result can be due to the selection. In the following analysis, the causal effect of repetition on interest rates is extracted by the synthetic control and difference-in-difference method.

Synthetic Control

Because the dataset contains large cross-sectional observations and very limited time series data for each borrower, loan or borrower fixed effects suffer from the incidental parameter problem. However, the structure of this dataset enables the construction of synthetic control.

The economics behind the synthetic control construction is based on the relative valuation method. This method of valuation using multiples consists of identifying comparable assets and using the price of assets in the peer group to find valuations. The peer group construction sheds light on the econometric method of synthetic control. In the practice of negotiating loan interest rates, the interest rates of comparable loans play an important role in determining the finalized loan contract, and a natural benchmark of interest rate for a new loan is the weighted average interest rates of peer loans.

The purpose of constructing the synthetic control group is to study the treatment effect of repeated interactions. If repeated interaction is regarded as a treatment, the pre-and post-treatment outcome of the treatment group (repeated borrowers) is easily measured by the interest rate of the first and second loans of the same borrower. However, the borrowers without repeated loans, the control group, only have the pre-treatment interest rates. Each repeated borrower has a different treatment time, and in order to construct the outcome of the control group to match the timeline and other characteristics of the treatment group, a synthetic control group is constructed for each borrower in the treatment group.

More specifically, four steps are as follows construct a synthetic control loan for each repeated loan i. First of all, before constructing the control group, the one-time borrowers in the full sample are filtered by the most fundamental characteristics, including borrower-lender distance, delivery method, and business type, such that the peer group and treatment individual have the exact same fundamental characteristics. Secondly, because only the loans that exist in the current market are informative for initiating a new loan, the filtered sample is further reduced to the loans that are approved and not matured before the approval of i's first loan month. The last step is to find the loans in the candidate peer group that best match the i's first loan in the other loan characteristics, including time approval, term in months, SBA guaranteed approval, and jobs supported. The steps of constructing the post-treatment period control group are similar to the pre-treatment period control group, except that the post-treatment period loans are matched with the second loan of repeated borrower i.

Difference-in-Difference Approach

After the construction of treatment and control groups, the observations of interest rates of repeated borrowers are separated into the first loan (pre-treatment) and second loan (post-treatment), and the interest rates for the synthetic control group are calculated using a weighted average of the control group loan interest rates. Figure 3.3 plots of means of control and treatment groups for both pre-and post-treatment periods, with the synthetic control group in red and the repetition treatment group in blue. Figures 3.3 shows that the repetition treatment effect is negative on the first moment of interest rates, which supports Hypothesis 3.

I then run the following difference-in-difference Regression (3.4)

$$InitialInterestRate_i = \hat{\beta}_0 + \hat{\beta}_1 Time_i + \hat{\beta}_2 Treatment_i + \hat{\beta}_3 Time_i \times Treatment_i, \tag{3.4}$$

where *Time* and *Treatment* are time and group indicators. The parameter of interest is $\hat{\beta}_3$, which is predicted to be positive by theory. Based on the regression result in Table 3.4, significantly negative for each subsample and full sample, verifying the treatment of repeated borrowing has a negative impact on initial loan interest rates. Furthermore, the negative time trend effect common to control and treatment groups implies the negative trend of market interest rates, and the negative treatment group-specific effect implies different types of borrowers are screened for different contracts (with or without starting small components). The last observation from the regression table is that the adjusted R^2 is low, which is due to a lack of covariates.

The interpretation of this regression is that the negative correlation between repetition and interest rates is not caused by the treatment but due to selection. When the interest rates decrease over time time (the significantly negative coefficient estimates of *Time*), firms have more incentive to borrow; on the other hand, when learning about the firm's type after repetition, the bank extracts more surplus from the repeated borrowers. Because the existing borrowers value the relationship with their current lender, they would be reluctant to return to the credit market and start a new relationship with another lender, which causes the ratchet effect.

The treatment group selection effect is significant (the significantly negative coefficient estimates of *Treatment*), but the selection effect is not high enough to compensate for the ratchet effect. As Figure 3.3 exhibits, the repeated borrowers refinance and start a new loan when interest rates decline, and the one-time borrowers are more sensitive to the decreasing market interest rates. In the before treatment period, all borrowers are first time borrowers, and lower risk borrowers are selected to loans with lower interest rates. As the market interest rates decrease, low risk borrowers either pay off existing loans and refinances from either current lender or borrows from another lender. The after treatment period interest rates show that those borrowers could have acted as a first time borrower and obtain a loan with lower interest rates, as their current lender takes advantage of the existing relationship and credit history to charge a higher interest rate.

3.4.3 The Impact of Repeated Lending on the Paycheck Protection Program

From the dataset, almost all of the borrowers that had 7(a) loans get smaller amounts of PPP loans compared to existing 7 (a) loans, which supports Hypothesis 4. Note that the PPP loan amount in the dataset is bounded above, which could cause bias in this result. However, most of the existing borrowers in the 7 (a) program choose not to apply for the PPP loan, so on average, the data supports Hypothesis 4.

To study the impact of repeated loans on PPP, I introduce a new regressor, *IsRep*, the indicator of whether the PPP loan is an existing borrower in the 7 (a) program or not. Table 3.5 shows the regression results. The first three short regressions examine the explanatory power of each regressor, and the last long regression includes all control variables.

First, including the control of loan and borrower characteristics leads to a very high adjusted R^2, that is, most of the variation of PPP loan amount is explained by the control variables. The reason of this result is that during the pandemic, the Covid relief program does not use the soft information very much.

Second, the coefficient estimate of *IsRep* is significantly negative in all regressions, which supports Hypothesis 5 in that the borrowers that had 7(a) loans get smaller amounts of PPP loans compared to new borrowers. Intuitively, when an existing borrower has already obtained a large credit line, they do not need to apply for another relief loan. Even if the existing borrower needs a larger amount of PPP, the lender is able to use the updated information about them to extract more surplus: the lender would prefer to attract new borrowers if existing borrowers value the relationship.

Third, residual analysis shows that residuals clustered around 0, but the residual plot suggests that there may be missing variables that cause a negative correlation between fitted loan amounts and residuals. The linear pattern in the Residual plot is caused by indicator variables. Q-Q plot of the residuals shows that the residuals behave nicely approximately Normal around the mode, but tails are heavy.

3.5 Conclusion

This paper presents a starting small model of business loan, provides an extension of renegotiation under productivity shock, and tests the model predictions using the SBA database on the 7 (a) loan program and Paycheck Protection Program.

The surprising results related to Hypothesis 3 and 5 rely crucially on the structure of the game. The game assumes the bank to have full bargaining power in both contract design and renegotiation, whereas the literature, such as Stiglitz and Weiss (1981), usually assumes that the capital market is competitive. The assumption of lenders having bargaining power is reasonable in the SBA environment; the lenders with the SBA guarantee have a comparative advantage so that the lenders can extract surplus until the borrower is indifferent between borrowing from SBA-backed loans or competitive market. Nevertheless, it is interesting to explore the model where the borrower has bargaining power.

There are other extensions of this model. First, the lender could offer a menu of loans in each period. Second, when considering the bank's design of interest rates and covenant terms, the model would involve a multi-dimensional level sequence. For example, in a model with imperfect monitoring, the more restrictive covenant term could be modeled as a higher accuracy of the monitoring signal. Third, the borrower's outside option could be endogenous. When a borrower shirks or defaults, it will return to the borrower pool with a credit history, which in turn endogenously determines the market composition of borrowers. [6]

In the empirical section, the sample selection may cause biased results. As discussed in 3.3.1, there are many options for small business loans and COVID relief loans, but this paper only focuses on the 7 (a) program and California PPP with loan amounts up to $\$150,000$. Because this study requires loan-level data, it is reasonable to expect a large dataset, which causes the need for sample selection.

6. Boone and Watson (2007) has a similar model of an endogenous composition of the agent pool.

The evidence of the impact of repeated lending on interest rates is surprising and deserves more investigation. Even if the repeated loans have lower interest rates on average, the treatment effect of repetition is actually the opposite. The explanation of the ratchet effect and the value of the existing relationship needs further analysis.

3.6 Appendix

3.6.1 Model

This section presents the starting small model of business lending. As discussed in Section 3.2, the model includes a baseline model and an extension of an economic downturn. The baseline model, Section 3.6.1 - 3.6.1, can be interpreted as the model of SBA 7 (a) program, while Section 3.6.1 is a model of PPP.

Baseline Model

In this subsection, I outline the baseline starting small model in business lending. Consider a bank and a firm interact for a potentially infinite number of periods; in period $t = 1, 2, ...$, the bank writes a loan to the firm and the firm has private information about its riskiness. Denote δ as their common discount factor.

Follow most of the assumptions and notations in Stiglitz and Weiss (1981) for now, and assume that the bank is writing a loan to a project of a continuum of types of risks. Denote the project with the highest risk as $\theta = \bar{\theta}$ and with the lowest risk as $\theta = \underline{\theta}$, and riskiness is in the sense of mean preserving spreads of return R distribution over the same domain. Further, assume that the mean risk of the population is θ^m. Denote the cumulative distribution function of R as $F(R|\theta)$ and the cumulative distribution function of θ as $G(\theta)$.

Assumption 6. *For each type θ, the common domain of return R is a compact domain $[\underline{R}, \bar{R}]$. Conditional on θ, the probability density function, $f(\cdot|\theta)$, is well-defined, continuous, everywhere positive over the domain of R, and symmetric around $(\bar{R}+\underline{R})/2$.*

Assumption 7. *The type of firms* θ *has a compact domain* $[\underline{\theta}, \bar{\theta}]$. *For any two ranked types* $\theta' \geq \theta''$, $R|\theta'$ *is a symmetric mean preserving spread of* $R|\theta''$. *The probability density function,* $g(\cdot)$, *is well-defined, continuous, and everywhere positive over the domain of* θ,

Assumptions 6 and 7 rule out the discussion of good projects with high variance but low downside risks. With all firms having common mean and symmetric return distribution, the firm's type θ only characterizes the second-order moment of return. Denote the mean and median of return distribution as $R_0 = (\bar{R} + \underline{R})/2$, which is the expected return for all types of firms. Further, Assumption 6 and 7 imply that $F(R_0|\theta) = 0.5$ for all θ. By single-crossing of the cumulative distribution function, the assumptions also imply that $F(R|\theta'') < F(R|\theta') < 0.5$ for $\theta'' \leq \theta'$ and $R \leq R_0$.

In the Stiglitz and Weiss (1981) model, the static loan contract includes the following terms: loan amount B, collateral $C(\leq B)$, and per period interest rate r. The main assumption that deviates from Stiglitz and Weiss (1981) is the divisibility of the loan. I follow the starting small literature and assume that the loans are perfectly dividable and denote $\alpha_t \in [0,1]$ as t period level (size) of the entire loan. [7] To simplify the model, I ignore the within-period discounting and treat all cash flow and payoff occurring at the end of each period. Denote the flow payoff vector $u_t = (\rho_t, \pi_t)$, where ρ_t is bank's period t payoff and π_t is firm's period t payoff. [8]

The timeline of this game is as follows. Bank and firm meet at the beginning of the game, or period $t = 0$, with B and C fixed for the entire game. The firm's type θ is privately drawn from the population distribution. At the beginning of period $t = 1, 2, ...$, bank writes loan with principal $\alpha_t B$ and collateral $\alpha_t C$. The loan amount is the bank's cash outflow and the firm makes investment decisions. If a firm chooses to not invest, then the game ends with payoff $(\alpha_t(C-B), \alpha_t(B-C))$. Otherwise, at the end of period t, return R_t is drawn, independent of all other periods, from conditional distribution $F(\cdot|\theta)$. If $R_t < B(1+r_t) - C$, then the loan is

7. Covenant terms are ignored for now, but can be analyzed later.
8. If I were to consider within-period discounting, then I need to also model the bank and firm's reinvestment return, which complicates the model.

in default and the game ends with payoff $(\alpha_t(R_t+C-B), -\alpha_t C)$. If $R_t \geq B(1+r_t)-C$, then the loan is repaid in full, and the game proceeds to the next period with current flow payoff $(\alpha_t B r_t, \alpha_t(R-Br_t))$. Bank learns firm's type through firm's binary decision and whether project return is sufficiently high to repay the loan; I assume that bank cannot verify the realization of R_t and the firm cannot choose to hide R_t to default. [9]

Assumption 8. *For each loan amount B and collateral C, the interest rate is bounded above, that is* $r < (R_0+C)/B-1$.

Assumption 8 exogenously defines an upper bound on interest rates, which is with loss in general. However, in the environment of the SBA 7 (a) program, this assumption is consistent with the maximum allowable spread restrictions. Assumption 8 implies that the probability of default for all types of firm, $F(B(1+r)-C|\theta)$, is bounded above by $F(R_0|\theta) = 0.5$ and that default events only happen at the left tail of return distribution. Denote $D_t \equiv B(1+r_t)-C$ as the threshold that determines default events and probability of default then is $F(D_t|\theta)$. [10]

Firm's Incentive Condition

Denote firm θ's continuation value at the beginning of period t as U_t. If the firm invests the loan in the project, then the firm's expected continuation value conditional on investing is

$$F(D_t|\theta)(-\alpha_t C) + (1-F(D_t|\theta))(\alpha_t(R_0-Br_t)+\delta U_{t+1}).$$

The firm chooses to invest if the continuation value from investing is no smaller than the value from shirking, $\alpha_t(B-C)$. If the bank incentivizes the firm to not shirk, then the firm's

9. Assume for now that bank's only decision is to offer the contract and an extension of the baseline model studies bank's monitoring choice.

10. In future research, this assumption should be relaxed and the bounds on interest rate should be endogenous.

continuation value is obtained by

$$\begin{cases} U_t & = F(D_t|\theta)(-\alpha_t C) + (1 - F(D_t|\theta))(\alpha_t(R_0 - Br_t) + \delta U_{t+1}) \\ U_t & \geq \alpha_t(B - C) \end{cases}$$

for all t.

Furthermore, if firm with type θ shirks at period $t+1$, then the firm with type θ must weakly prefer to invest in period t, that is

$$-F(D_t|\theta)C + (1 - F(D_t|\theta))\left(R_0 - Br_t + \delta\frac{\alpha_{t+1}}{\alpha_t}(B - C)\right) \geq B - C,$$

which yields the incentive condition

$$\delta\frac{\alpha_{t+1}}{\alpha_t} \geq 1 + \frac{B/(B-C)}{1 - F(D_t|\theta)} - \frac{R_0 - B(1 + r_t)}{B - C} \tag{3.5}$$

By Assumption 6-8, the right side of inequality (3.5) increases in θ. Denote θ_t as the bank's belief updates about the highest type of firm that invests the loan in period t, then the types that shirk in period t is $(\theta_t, \theta_{t-1}]$. By a similar analysis in Hua and Watson (2022), θ_t satisfies indifference condition

$$\delta\frac{\alpha_{t+1}}{\alpha_t} = 1 + \frac{B/(B-C)}{1 - F(D_t|\theta_t)} - \frac{R_0 - B(1 + r_t)}{B - C} \tag{3.6}$$

The sequence $\{\alpha_t\}$ increases in t if

$$1 - F(D_t|\theta_t) < \frac{B}{-B(1 + r_t) + R_0 - (1 - \delta)(B - C)}.$$

The right side, the survival function, decreases in the probability of default, so risky firm deserves increasing levels to invest the loan into the project.

The model implies that high-risk borrowers default early. Define hazard rate as density divided by survival rate, then we have the Hypothesis 1.

Bank's Contract Design

Denote V_t as bank's continuation value at period t, then bank's payoff when firm shirks is $\alpha_t(C-B)$ and when firm invests the loan properly is

$$F(D_t|\theta \leq \theta_t)\alpha_t(R_0+C-B)+(1-F(D_t|\theta \leq \theta_t))(\alpha_t Br_t+\delta V_{t+1}),$$

where $F(D_t|\theta \leq \theta - t)$ is the probability of default conditional on the firm being less risky than θ_t. So bank's continuation value is

$$V_t = \left(1-\frac{G(\theta_t)}{G(\theta_{t-1})}\right)\alpha_t(C-B)+\frac{G(\theta_t)}{G(\theta_{t-1})}$$
$$(F(D_t|\theta \leq \theta_t)\alpha_t(R_0+C-B)+(1-F(D_t|\theta \leq \theta_t))(\alpha_t Br_t+\delta V_{t+1})) \quad (3.7)$$

Bank's contract design problem then is to solve for sequences $\{\alpha_t\}$ and $\{r_t\}$ subject to (3.6). There are multiple equilibria, and one can select an equilibrium based on equilibrium selection criteria, such as maximizing a bank's profit. In particular, Hua and Watson (2021) characterizes the renegotiation-proofness contract in a similar principal-agent problem. The renegotiation-proofness is internally consistent in that only on-path renegotiation is allowed and the principal is incentivized to stay at the current equilibrium rather than speeding up or slowing down. Further, when the period length shrinks to 0, this renegotiation-proofness contract makes the principal indifferent between the current equilibrium and stalling for one period.

In the current model, I adopt the same equilibrium selection criteria. Instead of working through the alteration-proofness analysis, I use the property of the continuous-time renegotiation-proofness contract to narrow down the bank's value function. Specifically, the bank is indifferent between the current equilibrium and stalling for one period. The continuation value should be equal to the payoff of stalling for one period (playing the period $t-1$ equilibrium in period t), that is

$$V_t = F(D_{t-1}|\theta \leq \theta_{t-1})\alpha_{t-1}(R_0 + C - B) + (1 - F(D_{t-1}|\theta \leq \theta_{t-1}))(\alpha_{t-1}Br_{t-1} + \delta V_t),$$

which yields the bank's continuation value

$$V_t = \alpha_{t-1}\frac{F(D_{t-1}|\theta \leq \theta_{t-1})(R_0 + C - B(1 + r_{t-1})) + Br_{t-1}}{1 - \delta(1 - F(D_{t-1}|\theta \leq \theta_{t-1}))} \geq 0 > \alpha_t(C - B) \tag{3.8}$$

Evaluate this one period forward and combine (3.7) we get

$$\begin{aligned} &\frac{\alpha_{t-1}}{\alpha_t}\frac{F(D_{t-1}|\theta \leq \theta_{t-1})(R_0 + C - B(1 + r_{t-1})) + Br_{t-1}}{1 - \delta(1 - F(D_{t-1}|\theta \leq \theta_{t-1}))} \\ &\quad = \left(1 - \frac{G(\theta_t)}{G(\theta_{t-1})}\right)(C - B) + \frac{G(\theta_t)}{G(\theta_{t-1})}\frac{F(D_t|\theta \leq \theta_t)(R_0 + C - B(1 + r_t)) + Br_t}{1 - \delta(1 - F(D_t|\theta \leq \theta_t))} \end{aligned} \tag{3.9}$$

Combining Equations (3.6), (3.8) and (3.9), the equilibrium is determined up to a given sequence $\{r_t\}$. If the bank can choose r_t freely, then the equilibrium is selected by the bank maximizing its value. However, in general, the interest rate is determined by base rate, credit record, and loan characteristics, so I treat $\{r_t\}$ as exogenous and common knowledge to players.

The speed of learning and increasing levels is driven by the severity of information asymmetry. In particular, as in Hua and Watson (2021), the more dispersed the distribution of the firm is, the slower the level increases. Motivated by Liberti and Petersen (2018), I use the distance between borrower and lender as a proxy for information asymmetry, then we have the Hypothesis 2. [11]

Furthermore, even though this paper does not model interest rates explicitly, interest rates are determined by loan, borrower, and time-fixed effects. However, the model still has two predictions regarding the loan interest rates. First, repeated borrowers reveal their private

11. Note that this hypothesis is not straightforward to get from the above derivation, so I state the conclusion using the comparative statics result in Hua and Watson (2021) without proof.

information over time, and hence banks are able to extract more surplus by charging a higher interest rate. Second, as the information asymmetry is reduced after the first loan, second-time borrowers should face a less dispersed interest rate. The predictions regarding interest rate are summarized as Hypothesis 3.

Productivity Shock

Existing Borrowers

Consider an unexpected productivity shock, which causes a period of economic downturn. During the crisis, the firm's return distribution is temporarily multiplied by a known factor $\rho \in (0,1)$. In this case, the original repeated lending process is disrupted and the bank then chooses to stall for one period. Denote the period of crisis is τ, then firm's incentive condition at period τ becomes

$$\delta \frac{\alpha_{\tau+1}}{\alpha_\tau} \geq 1 + \frac{B/(B-C)}{1-F(D_\tau/\rho|\theta_\tau)} - \frac{\rho R_0 - B(1+r_\tau)}{B-C}$$

which is violated in the original equilibrium, because for all $\rho \in (0,1)$, $F(D_t|\theta_t) < F(D_t/\rho|\theta_t)$ and

$$\frac{B/(B-C)}{1-F(D_\tau|\theta_\tau)} - \frac{R_0 - B(1+r_\tau)}{B-C} < \frac{B/(B-C)}{1-F(D_\tau/\rho|\theta)} - \frac{\rho R_0 - B(1+r_\tau)}{B-C}.$$

Therefore, the bank and firm will both be better off after a proper renegotiation.

I adopt the definition of alteration-proofness in Watson (1999) and Hua and Watson (2021) to study renegotiation. In period τ, the bank selects $\tilde{\alpha}$ to play, and in the periods after τ, the original equilibrium is played. That is, in period τ, the new $\tilde{\alpha}$ is selected to incentivize firm not shirking, and in period $\tau + j$, $j = 1,2,...$, the original $\alpha_{\tau+j-1}$ is played. Further, the internally consistent alteration should cause no type of firm shirks, so the period τ incentive condition is

$$\delta \frac{\alpha_\tau}{\tilde{\alpha}} = 1 + \frac{B/(B-C)}{1-F(\tilde{D}/\rho|\theta_{\tau-1})} - \frac{\rho R_0 - B(1+\tilde{r})}{B-C}, \tag{3.10}$$

where $\tilde{r}$ is the interest rate during economic downturn and $\tilde{D} = B(1+\tilde{r}) - C$. In the game without productivity shock, a stalling alteration would have $\tilde{\alpha} = \alpha_{\tau-1}$ and $\tilde{r} = r_{\tau-1}$. Note that this condition implies for $\rho \in (0,1)$, $\tilde{\alpha} \leq \alpha_\tau$, if $\tilde{r}$ is close to $r_{\tau-1}$. The result summarizes into Hypothesis 4. [12]

After the alteration, the bank's continuation value is

$$\tilde{V}_\tau = F(\tilde{D}/\rho|\theta \leq \theta_{\tau-1})\tilde{\alpha}(\rho R_0 + C - B) + (1 - F(\tilde{D}/\rho|\theta \leq \theta_{\tau-1}))(\tilde{\alpha}B\tilde{r} + \delta V_\tau),$$

where V_τ satisfies Equation (3.8).

New Borrowers

For the new borrowers, the bank's belief of their type is the prior distribution $[\underline{\theta}, \bar{\theta}]$. If the first-time borrowers are one-time borrowers, then these loans are static contracts, and hence bank's contract pools all first-time borrowers. In this case, the incentive condition of one-time borrowers is

$$0 = 1 + \frac{B/(B-C)}{1 - F(\tilde{D}/\rho|\bar{\theta})} - \frac{\rho R_0 - B(1+\tilde{r})}{B-C}. \tag{3.11}$$

Because there is no need to start small for the static contract, any level α_τ^{new} could satisfy the firm's incentive condition with the rest of the contract terms subject to (3.11).

On the other hand, if the first-time borrowers repeat the interaction with the same bank, then α_τ becomes their initial loan level. Denote $\{\alpha_{\tau+j}^{new}\}_{j=0}^{\infty}$ as the contracts started by the first time borrower in period τ, then period τ incentive condition is

$$\delta \frac{\alpha_1}{\alpha_\tau^{new}} \geq 1 + \frac{B/(B-C)}{1 - F(\tilde{D}/\rho|\bar{\theta})} - \frac{\rho R_0 - B(1+\tilde{r})}{B-C} = 0$$

which is always satisfied.

For the existing borrowers to stay at the current contract, the incentive condition of

12. The requirement of $\tilde{r}$ being relatively close to $r_{\tau-1}$ is not too restrictive: it only needs the effect of productivity shock ρ to dominate.

existing borrowers not to be pooled with new borrowers is

$$\tilde{\alpha}(B-C) \geq \alpha_{\tau}^{new}(B-C) - cost,$$

where *cost* denotes the cost of switching lenders. Note that bank's payoff increases in α_τ, when $\alpha_{\tau+1} = 0$, so the bank wants to set α_{τ}^{new} as high as possible:

$$\alpha_{\tau}^{new} = \min\left\{1, \tilde{\alpha} + \frac{cost}{B-C}\right\}.$$

Therefore, for non-negative cost, the first-time borrowers get a weakly higher loan amount than existing borrowers. Hypothesis 5 summarizes this result.

3.6.2 Definition of Variables

- Borrower characteristics.
  - **JobsSupported**: Total Jobs Created + Jobs Retained as reported by lender on SBA Loan Application. SBA does not review, audit, or validate these numbers - they are simply self-reported, good faith estimates by the lender.
  - **IndRepeat**: Indicator of whether a borrower repeatedly borrow from the same bank (0=One time borrower, 1=Repeated borrower).
  - **NumRepeat**: Counting variable. For each loan, NumRepeat equaling k means this loan is the k-th time of repeated interaction, ordered by approval date.
  - **ZDis**: Distance between borrower's Zip code and lender's code (in miles), calculated using Sami (2021)
- Loan characteristics.
  - **GrossApproval**: Total loan amount.

- **GrossChargeOffAmount**: Total loan balance charged off (includes guaranteed and non-guaranteed portion of loan).
- **SBAGuaranteedApproval**: Amount of SBA's loan guaranty.
- **InitialInterestRate**: Initial interest rate - total interest rate (base rate plus spread) at time loan was approved
- **PercentDefaulted**: GrossChargeOffAmount divided by GrossApproval.
- **PercentGuaranteed**: SBAGuaranteedApproval divided by GrossApproval.
- **RevolverStatus**: Indicator of whether a loan is a term loan or revolving line of credit (0=Term, 1=Revolver).

• Loan Status.

- **IsCOMMIT**: Indicator for loan status (1=Undisbursed, 0=Otherwise).
- **IsPIF**: Indicator for loan status (1=Paid In Full, 0=Otherwise).
- **IsCHGOFF** : Indicator for loan status (1= Charged Off, 0=Otherwise).
- **IsCANCLD** : Indicator for loan status (1= Cancelled, 0=Otherwise).
- **IsEXEMPT** : Indicator for loan status (1= Exempt, 0=Otherwise). The status of loans that have been disbursed but have not been cancelled, paid in full, or charged off are exempt from disclosure under FOIA Exemption 4.

• Time in months.

- **TermInMonths**: Length of loan term.
- **TimeApproval**: Month the loan was approved.
- **TimeChargeOff**: Month SBA charged off loan (if applicable).
- **TimeToDefault**: Number of months between TimeApproval and TimeChargeOff (if applicable).

3.6.3 Figures and Tables

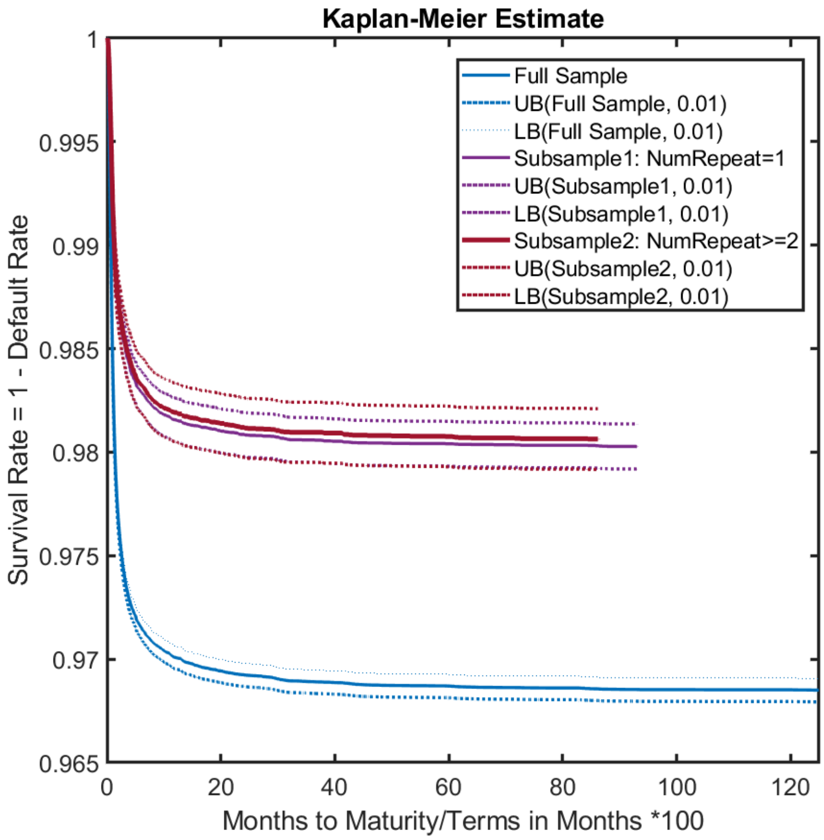


Figure 3.1. Nonparametric estimate of the survival rate. Solid curves show Kaplan Meier estimates of survival rates and dashed curves present 99% confidence intervals of the estimates.

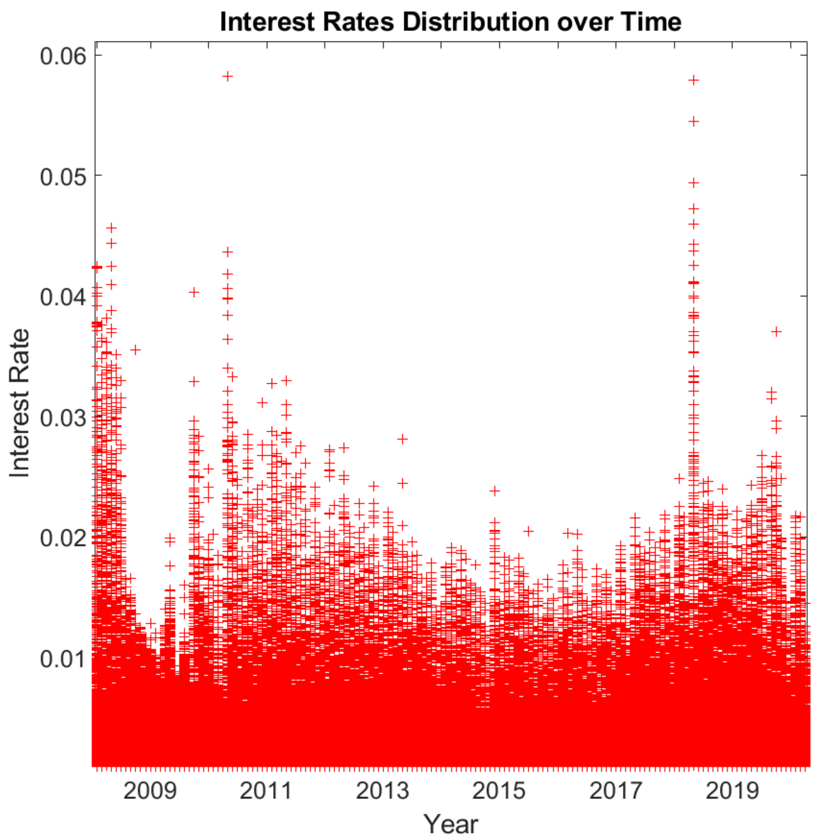


Figure 3.2. Interest rates Distribution Over Time

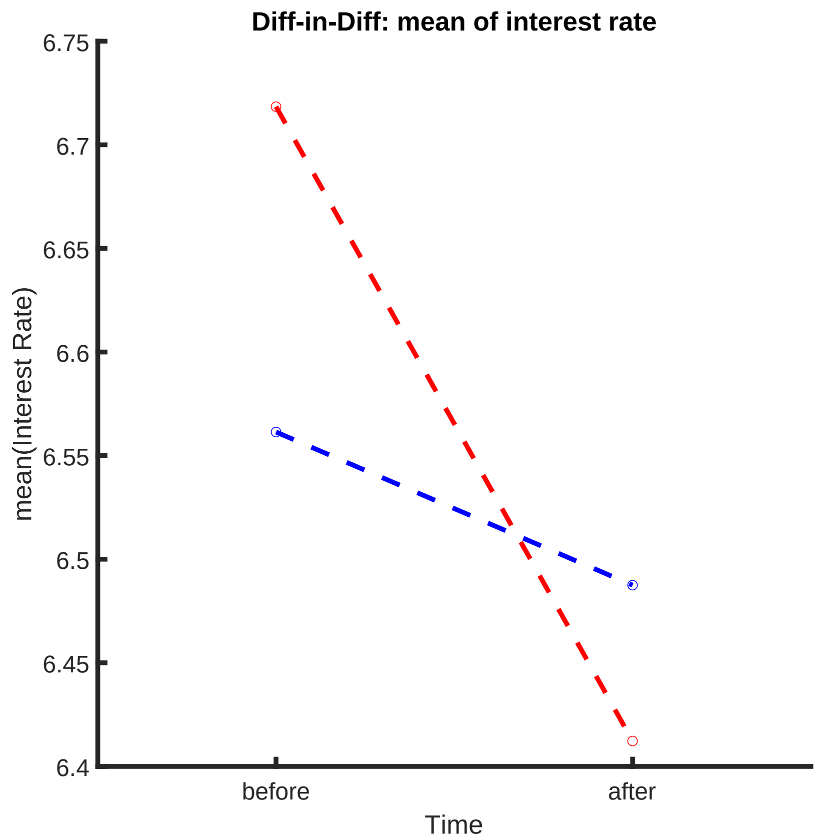


Figure 3.3. Difference-in-Difference: Mean of Interest Rates. Red line shows the changes in average interest rates in the synthetic control group, and blue line shows the changes in average interest rates in the repetition treatment group. Existing borrowers have higher incentive to refinance existing loans when interest rates are low, during which period, existing borrowers could have borrowed at lower interest rates if they were first time borrowers.

Table 3.1. Summary Statistics. Summary statistics for TimeChargeOff and TimeToDefault are conditional on defaulted bonds.

7(a) Program	mean	median	mode	std	min	max
Borrower Characteristics						
ZDis	606.27	207.75	0	1012.64	0	7953.33
JobsSupported	10.7364	5	0	21.4960	0	4504
IndRepeat	0.1723	0	0	0.3835	0	1
NumRepeat	1.1146	1	1	0.4334	1	31
Loan Characteristics (first three rows are in 10^3)						
GrossApproval	385.72	130500	50	684.90	1	5000
GrossChargeOffAmount	4.96	0	0	53.35	0	4706.18
SBAGuaranteedApproval	286.61	85	25	526.56	0.5	6175
InitialInterestRate	6.5148	6	6	1.5543	0	56
PercentDefaulted	0.0306	0	0	0.1557	0	1
PercentGuaranteed	0.6534	0.75	0.5	0.1532	0.1	1
RevolverStatus	0.3118	0	0	0.4632	0	1
Loan Status						
IsCANCLD	0.1139	0	0	0.3177	0	1
IsCHGOFF	0.0412	0	0	0.1987	0	1
IsCOMMIT	0.0217	0	0	0.1458	0	1
IsEXEMPT	0.4137	1	1	0.4925	0	1
IsPIF	0.4094	0	0	0.4917	0	1
Time in months						
TermInMonths	122.78	90	84	78.84	0	847
TimeApproval	76.16	79.63	17.37	39.99	0.033	147
TimeChargeOff	96.43	102.40	121	33.10	10.87	147
TimeToDefault	48.52	44.20	32.43	24.34	5.3	143.6

PPP	mean	median	mode	std	min	max
Borrower Characteristics						
ZDis	1435.60	1760.33	1119.4	1012.64	0	10534.55
JobsReported	5.5645	3	1	13.2571	0	500
IsRep	0.0062	0	0	0.0788	0	1
Loan Characteristics (first row is in 10^3)						
CurrentApprovalAmount	35.26	20.83	20.83	34.137	0.001	150
Term in months	30.61	24	24	13.94	22	60
IsRural	0.0175	0	0	0.1312	0	1
IsHubzone	0.1895	0	0	0.3919	0	1

Table 3.2. Cox Proportional Hazard Model. Numbers in parentheses are standard errors for each coefficient estimate. Baseline specification is compared with 0. All coefficient estimates are significant at 1% level. The last row reports the log-likelihood difference between each model and the first model.

	(1)	(2)	(3)	(4)
IndRepeat		-0.1617		-0.1595
		(0.0057)		(0.0057)
ZDis/1000			0.0236	0.0209
			(0.0022)	(0.0022)
GrossApproval/10000	-0.0050	-0.0050	-0.0050	-0.0050
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
PercentGuaranteed	1.1151	1.0935	1.1145	1.09321
	(0.0149)	(0.0149)	(0.0149)	(0.0149)
TermInMonths	-0.0225	-0.0227	-0.0225	-0.0227
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
InitialInterestRate	-0.0342	-0.0392	-0.0362	-0.0410
	(0.0015)	(0.0015)	(0.0015)	(0.0015)
JobsSupported	0.0011	0.0011	0.0011	0.0011
	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Observations	633297	633297	633297	633297
Log Likelihood Diff	0	421.8532	56.4898	466.0726

Table 3.3. Interest Rate Determination. Numbers in parentheses are standard errors for each coefficient estimate. All coefficient estimates are significant at 1% level. The unit of Term is the year. NumRepInd is indicator for NumRepeat$> 2\%$. MarketRate is the loan amount-weighted interest rate at approval time. Control variables include loan status (IsCANCLD, IsCHGOFF, IsCOMMIT, IsEXEMPT). Time-fixed effects are indicator variables for the year of approval.

	(1)	(2)	(3)	(4)	(5)
Intercept	6.9156	8.0059	6.9816	8.0720	0.9202
	(0.0110)	(0.0299)	(0.0112)	(0.0298)	(0.0203)
IndRepeat	-0.3909	-0.4004	-0.3936	-0.3953	-0.3987
	(0.0049)	(0.0045)	(0.0051)	(0.0048)	(0.0045)
NumRepInd	-0.2180	-0.1981	-0.1879	-0.1809	-0.1968
	(0.0158)	(0.0145)	(0.0167)	(0.0156)	(0.0145)
ZDis/1000	0.17481	0.1641	0.185	0.1746	0.1661
	(0.0018)	(-0.123)	(0.0018)	(0.0017)	(0.0016)
Term	-0.1510	-0.1164	-0.1436	-0.1123	-0.1251
	(0.0034)	(0.0031)	(0.0036)	(0.0034)	(0.0031)
JobsSupported/1000	-4.44035	-4.0545	-4.4394	-4.0229	-4.1698
	(0.0903)	(0.0829)	(0.0927)	(0.0869)	(0.083)
RevolverStatus	0.68949	0.6637	0.6096	0.6168	0.6720
	(0.0049)	(0.0045)	(0.0050)	(0.0047)	(0.0045)
PercentGuaranteed	-0.9022	-0.9752	-0.9893	-1.0735	-0.9919
	(0.0152)	(0.0144)	(0.0155)	(0.015)	(0.0140)
GrossApproval	-0.2895	-0.2836	-0.2959	-0.3015	-0.2794
	0.0032	(0.0029)	(0.0034)	(0.0032)	(0.0029)
MarketRate					1.0656
					(0.0031)
Loan status	Yes	Yes	Yes	Yes	Yes
Time fixed effects	No	Yes	No	Yes	No
Observations	633294	633294	559670	559670	633294
Adj R^2	0.1839	0.3141	0.1865	0.2861	0.3107

Table 3.4. Diff-in-Diff: interest rate. Numbers in parentheses are standard errors for each coefficient estimate. All coefficient estimates are significant at 5% level.

	(1)
Intercept	6.7184
	(0.05715)
Time	-0.30615
	(0.0808)
Treatment	-0.15703
	(0.0808)
Time× Treatment	0.23223
	(0.11429)
Observations	4064
Adj R^2	0.00312

Table 3.5. PPP Loan Amount Regression (California Dataset). The unit of PPPTerm is year. Numbers in parentheses are standard errors for each coefficient estimates. All coefficient estimates are significant at 1% level. Control variables include business type, jobs reported, payroll proceed, rural indicator, hubzone indicator, business age.

	(1)	(2)	(3)	(4)
Intercept	0.4051	2.7642	2.89590	2.8970
	(0.0007)	(0.1897)	(0.1898)	(0.1898)
IsRep	-0.1365	-0.2049		-0.2026
	(0.0055)	(0.0692)		(0.0691)
ZDis/1000	-0.0360		-0.0743	-0.0742
	(0.0004)		(0.0050)	(0.0050)
PPPTerm		-0.1164	-0.1112	-0.1112
		(0.0050)	(0.0050)	(0.0050)
Controls	No	Yes	Yes	Yes
Observations	602088	602088	602088	602088
Adj R^2	0.0145	0.9847	0.9847	0.9847

Bibliography

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.

Azevedo, Eduardo M, and Jacob D Leshno. 2016. “A supply and demand framework for two-sided matching markets.” *Journal of Political Economy* 124 (5): 1235–1268.

Battigalli, Pierpaolo, and Joel Watson. 1997. “On "Reputation" Refinements with Heterogeneous Beliefs.” *Econometrica* 65 (2): 369–374.

Becker, Gary S. 1973. “A theory of marriage: Part I.” *Journal of Political economy* 81 (4): 813–846.

Bengtsson, Ola, and Berk A Sensoy. 2011. “Investor abilities and financial contracting: Evidence from venture capital.” *Journal of financial intermediation* 20 (4): 477–502.

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

Bernstein, Shai, Kunal Mehta, Richard R Townsend, and Ting Xu. 2022. *Do Startups Benefit from Their Investors' Reputation? Evidence from a Randomized Field Experiment.* Technical report. National Bureau of Economic Research.

Bhaskar, Venkataraman, et al. 2014. "The ratchet effect re-examined: A learning perspective." *Working Paper.*

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

Boone, Johnathan, and Joel Watson. 2007. "Testing and the Composition of the Labor Pool." *The Scandinavian Journal of Economics* 109 (4): 883–903.

Boot, Arnoud W. A., Anjan V. Thakor, and Gregory F. Udell. 1991. "Secured Lending and Default Risk: Equilibrium Analysis, Policy Implications and Empirical Results." *The Economic Journal* 101 (406): 458–472.

Bottazzi, Laura, Marco Da Rin, and Thomas Hellmann. 2008. "Who are the active investors?: Evidence from venture capital." *Journal of Financial Economics* 89 (3): 488–512.

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.

Chang, Ganlin, and Suresh M. Sundaresan. 2001. "A Model of Dynamic Sovereign Borrowing Effects of Credit History and Sanctions1." *Working Paper* (January).

Charness, Gary, Peter Kuhn, and Marie Claire Villeval. 2011. "Competition and the ratchet effect." *Journal of Labor Economics* 29 (3): 513–547.

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

Chemmanur, Thomas J, Karthik Krishnan, and Debarshi K Nandy. 2011. "How does venture capital financing improve efficiency in private firms? A look beneath the surface." *The Review of Financial Studies* 24 (12): 4037–4090.

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

Coles, Melvyn G, and Dale T Mortensen. 2016. "Equilibrium labor turnover, firm growth, and unemployment." *Econometrica* 84 (1): 347–363.

Cong, Lin William, and Yizhou Xiao. 2022. "Persistent Blessings of Luck: Theory and an Application to Venture Capital." *The Review of Financial Studies* 35 (3): 1183–1221.

Craig, Ben R., William E. Jackson III, and James B. Thomson. 2009. "The Economic Impact of the Small Business Administration's Intervention in the Small Firm Credit Market: A Review of the Research Literature." *Journal of Small Business Management* 47 (2): 221–231.

Cripps, Martin W, and Jonathan P Thomas. 2003. "Some asymptotic results in discounted repeated games of one-sided incomplete information." *Mathematics of Operations Research* 28 (3): 433–462.

Cripps, Martin W., Klaus M. Schmidt, and Jonathan P. Thomas. 1996. "Reputation in Perturbed Repeated Games." *Journal of Economic Theory* 69 (2): 387–410.

DeYoung, Robert, Dennis Glennon, and Peter Nigro. 2008. "Borrower-lender distance, credit scoring, and loan performance: Evidence from informational-opaque small business borrowers." Financial Contracting and Financial System Architecture, *Journal of Financial Intermediation* 17 (1): 113–143. ISSN: 1042-9573.

Dinlersoz, Emin M, Henry R Hyatt, and Hubert P Janicki. 2019. "Who works for whom? Worker sorting in a model of entrepreneurship with heterogeneous labor markets." *Review of Economic Dynamics* 34:244–266.

Ewens, Michael, Alexander Gorbenko, and Arthur Korteweg. 2022. "Venture capital contracts." *Journal of Financial Economics* 143 (1): 131–158.

Ewens, Michael, and Matthew Rhodes-Kropf. 2015. "Is a VC Partnership Greater than the Sum of its Partners?" *The Journal of Finance* 70 (3): 1081–1113.

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

Freixas, Xavier, Roger Guesnerie, and Jean Tirole. 1985. "Planning under incomplete information and the ratchet effect." *The review of economic studies* 52 (2): 173–191.

Freudenberg, Felix, Björn Imbierowicz, Anthony Saunders, and Sascha Steffen. 2017. "Covenant violations and dynamic loan contracting." *Journal of Corporate Finance* 45:540–565.

Fudenberg, Drew, and David K. Levine. 1989. "Reputation and Equilibrium Selection in Games with a Patient Player." *Econometrica* 57 (4): 759–778.

Gale, David, and Lloyd S Shapley. 1962. "College admissions and the stability of marriage." *The American Mathematical Monthly* 69 (1): 9–15.

Gerardi, Dino, and Lucas Maestri. 2020. "Dynamic contracting with limited commitment and the ratchet effect." *Theoretical Economics* 15 (2): 583–623.

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

Glennon, Dennis, and Peter Nigro. 2005. "Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach." *Journal of Money, Credit and Banking* 37 (5): 923–947.

Gompers, Paul, Anna Kovner, Josh Lerner, and David Scharfstein. 2010. "Performance persistence in entrepreneurship." *Journal of financial economics* 96 (1): 18–32.

Gompers, Paul, and Josh Lerner. 1999. "An analysis of compensation in the US venture capital partnership." *Journal of Financial Economics* 51 (1): 3–44.

Gompers, Paul A. 1996. "Grandstanding in the venture capital industry." *Journal of Financial economics* 42 (1): 133–156.

———. 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.

Hart, Sergiu. 1985. "Nonzero-Sum Two-Person Repeated Games with Incomplete Information." *Mathematics of Operations Research* 10 (1): 117–153.

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu. 2007. "Whom you know matters: Venture capital networks and investment performance." *The Journal of Finance* 62 (1): 251–301.

Hochberg, Yael V, Alexander Ljungqvist, and Annette Vissing-Jørgensen. 2014. "Informational holdup and performance persistence in venture capital." *The Review of Financial Studies* 27 (1): 102–152.

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

Horstmann, Ignatius, and James Markusen. 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.

Hsu, David H. 2004. "What do entrepreneurs pay for venture capital affiliation?" *The journal of finance* 59 (4): 1805–1844.

Hua, Xiameng, and Joel Watson. 2021. "Starting Small and Renegotiation in Discrete-Time Relationships with a Continuum of Types." *Working Paper* (January).

———. 2022. "Starting small in project choice: A discrete-time setting with a continuum of types." *Journal of Economic Theory* 204:105490. ISSN: 0022-0531. https://doi.org/https://doi.org/10.1016/j.jet.2022.105490. https://www.sciencedirect.com/science/article/pii/S0022053122000801.

Ickes, Barry W, and Larry Samuelson. 1987. "Job transfers and incentives in complex organizations: Thwarting the ratchet effect." *The Rand Journal of Economics,* 275–286.

Kaplan, Steven N, and Antoinette Schoar. 2005. "Private equity performance: Returns, persistence, and capital flows." *The journal of finance* 60 (4): 1791–1823.

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.

Krishnan, CNV, Vladimir I Ivanov, Ronald W Masulis, and Ajai K Singh. 2011. "Venture capital reputation, post-IPO performance, and corporate governance." *Journal of Financial and Quantitative Analysis* 46 (5): 1295–1333.

Liberti, José María, and Mitchell A Petersen. 2018. "Information: Hard and Soft." *The Review of Corporate Finance Studies* 8, no. 1 (November): 1–41.

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

Nahata, Rajarishi. 2008. "Venture capital reputation and investment performance." *Journal of financial economics* 90 (2): 127–151.

Pei, Harry. 2021. "Trust and betrayals: Reputational payoffs and behaviors without commitment." *arXiv preprint arXiv:2006.08071.*

RAJAN, RAGHURAM G. 1992. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *The Journal of Finance* 47 (4): 1367–1400. https://doi.org/https://doi.org/10.1111/j.1540-6261.1992.tb04662.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1992.tb04662.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1992.tb04662.x.

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.

Sami. 2021. "Zip Code Latitude Longitude City State County." *MATLAB Central File Exchange.*

Sannikov, Yuliy. 2007. "Agency Problems, Screening and Increasing Credit Lines." *Working paper.*

SBA. 2021a. "Paycheck Protection Program (PPP) Data." *U.S. Small Business Administration,* https://www.sba.gov/funding-programs/loans/coronavirus-relief-options/paycheck-protection-program/ppp-data.

———. 2021b. "SBA COVID Relief Program Report as of 03/11/21." *U.S. Small Business Administration,* https://www.sba.gov/sites/default/files/2021-03/2021.03.12%5C%20-%5C%20SBA%5C%20COVID%5C%20Relief%5C%20Program%5C%20Report%5C%20-%5C%20vFINALSHARE-508.pdf.

———. 2021c. "Small Business Administration $7(a)$ Loan Data." *U.S. Small Business Administration,* https://www.sba.gov/about-sba/open-government/foia#section-header-32.

Shalev, Jonathan. 1994. "Nonzero-sum two-person repeated games with incomplete information and known-own payoffs." *Games and Economic Behavior* 7 (2): 246–259.

Shapley, Lloyd S, and Martin Shubik. 1969. "On market games." *Journal of Economic Theory* 1 (1): 9–25.

SHARPE, STEVEN A. 1990. "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships." *The Journal of Finance* 45 (4): 1069–1087. https://doi.org/https://doi.org/10.1111/j.1540-6261.1990.tb02427.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1990.tb02427.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1990.tb02427.x.

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

Sørensen, Morten. 2007. “How smart is smart money? A two-sided matching model of venture capital.” *The Journal of Finance* 62 (6): 2725–2762.

Stiglitz, Joseph E., and Andrew Weiss. 1981. “Credit Rationing in Markets with Imperfect Information.” *The American Economic Review* 71 (3): 393–410.

Terviö, Marko. 2009. “Superstars and mediocrities: Market failure in the discovery of talent.” *The Review of Economic Studies* 76 (2): 829–850.

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

Watson, Joel. 1993. “A "Reputation" Refinement without Equilibrium.” *Econometrica* 61 (1): 199–205.

———. 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.

———. 2013. “Contract and game theory: Basic concepts for settings with finite horizons.” *Games* 4:457–496.

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

———. 2021. “Convergence of Discrete-Time Models with Small Period Lengths.” *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.