UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on the Theory of Bargaining and Economics of Matching Platforms

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

by

Andrew Park

2022

ABSTRACT OF THE DISSERTATION

Essays on the Theory of Bargaining
and Economics of Matching Platforms

by

Andrew Park
Doctor of Philosophy in Economics
University of California, Los Angeles, 2022
Professor Simon Adrian Board, Chair

This thesis consists of three essays studying the theory of bargaining and learning dynamics of matching platforms. The first essay studies the role of optimism in non-cooperative bargaining, while the second essay explores how introducing bargaining incentives affect trust building process in international relations context. The final essay considers learning incentives of matching platforms that utilize their matching technology to exploit or explore the quality of their constituents.

The first essay asks a theoretic question: does exaggerated optimism benefit an agent in bargaining? The paper analyzes a two agent non-cooperative bargaining model to study if, and when, one has incentive to over-report his level of optimism. It modifies the complete information Rubinstein bargaining model to let players hold different beliefs about which player makes an offer. Defining optimism over one's perceived recognition probability, I find that an agent always "envies" a more optimistic agent, and has incentive to play optimism as strategic posture to benefit. The second part of the chapter introduces an asymmetry of information to the game, letting an agent be of a "more optimistic" type with some known

probability. I find that the less optimistic type 1) pretends to be the more optimistic type—"play optimism"—if his probability of being more optimistic is high enough, 2) reveals his type before the more optimistic type would have settled, and 3) benefits more by playing optimism the higher the probability of extreme optimism is.

The second essay studies social encounters that involve both trust building and bargaining. We show that while bargaining interferes with trust building in the sense that fully informative signaling becomes impossible, bargaining alongside trust-building actually improves welfare when initial trust is low. In contrast to the current literature, we show that actors improve welfare by building trust more slowly. Thus, windows of opportunity to build trust must be seized to prevent significant declines in expected welfare. We also characterize the evolution of stakes that lead to the best outcomes. Our analysis explains why trust building is so much more difficult than the current literature implies and illuminates the opportunities that produce the best outcomes between adversaries with something to lose.

The third essay studies how platforms can utilize its pooling ability both to generate flow output and to discover good agents at the same time. In a simple model of two types in continuous time, the paper identifies an exploration-exploitation trade-off: by only matching good agents to each other, the platform may maximize flow output while sacrificing discovery of new good agents; on the other hand, by keeping an integrated pool, the platform maximizes learning rate while sacrificing the number of good matches. We find that the optimal matching policy is bang-bang from full integration–until the discovery ratio of good agents hits a certain threshold–to full segmentation thereafter to maximize flow payoffs. We also characterize how the threshold ratio responds to parameters of the model.

The dissertation of Andrew Park is approved.

Robert Trager

Ichiro Obara

Moritz Meyer-ter-Vehn

Simon Adrian Board, Committee Chair

University of California, Los Angeles

2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply grateful for my advisor Simon Board for his constant encouragement and intellectual guidance. Between the military leave of absence and the pandemic, my graduate school life has been especially turbulent and isolated. It was with Simon's proactive supervision and reassurance that I could stay on course.

I thank Moritz Meyer-ter-Vehn for his invaluable feedback and for organizing research groups to ensure students didn't lose direction amid quarantined research; Robert Trager for the research opportunities and warmhearted advice; Ichiro Obara for his helpful suggestions; and the UCLA Economics Theory group for the feedback and support.

I would also like to thank my co-authors. Robert Trager and I wrote Chapter 2 together, and it has been a true pleasure to work with him. Chapter 3 is a version of my work with Isabel Juniewicz, who originated the idea of learning through matching.

I want to thank my friends at UCLA, especially Jay, Jingyu, Stepan, Junrok, Wookun, Keyoung, and Jaeyoung for the hangouts and soundboarding. Richard, Do Young, Minsuk, and Joe also deserve recognition for the boundless conversations that encouraged and inspired me in moments of need.

Conducting research remotely allowed me to meet a very special person. I would like to express my gratitude to Kylie Hwang, who provided emotional support for me to lean on, but when the time came, didn't shy away from giving me solid academic advice as well. Both the struggles and the achievements along the journey were made more valuable because you were there beside me.

Most importantly, I would like to thank my family. For my parents, who have always supported my decisions no matter what, thank you for believing in my endeavor to pursue PhD degree. Your encouragement kept me digging even in times of doubt. And for my sister Hailey, thank you for being the most reliable little sister I could ask for.

VITA

2011–2013	Research Assistant, Economics Department, Stanford University. Assisted Professor Fuhito Kojima in the publication of ”Double Auction with Interdependent Values: Incentives and Efficiency.”
2013	B.A.S. (Mathematics and Economics), Stanford University.
2014	M.S. (Statistics), Stanford University.
2015–2016	Teaching Assistant, Economics Department, UCLA. Received yearly TA Award for teaching Microeconomic Theory (Econ 11).
2016–2019	Seneor Researcher / Software Engineer, TmaxSoft. Developed logical optimization engine of SQL processing in Tibero, a relational database management system.
2019–2020	Teaching Assistant, Economics Department, UCLA. Received TA Award for teaching Introduction to Game Theory (Econ 106G).
2020–2022	Research Assistant, Political Science Department, UCLA. Assisted Professor Robert Trager in the working paper ”Informational Hazards in AI Races.”

CHAPTER 1

Playing Optimism: Role of Uncommon Prior in Sequential Bargaining

1.1 Introduction

Bargaining skill matters. Different buyers often settle on substantially different prices for the same product from the same supplier; defendants in civil case lawsuit pay a large sum to hire lawyers because they believe a lawyer's superior bargaining ability will more than compensate for it in settlement terms. Game theorists have developed an approach to answer what determines such bargaining power from as early as Nash [1950] and, more pertinently, Ståhl [1973]. In empirical IO literature, Grennan [2014] identifies the force of bargaining ability using empirical evidence from hospitals' purchase of medical devices, attributing as much as 79% of price variation to this force. My paper studies a more specific form of bargaining power: how optimism affects bargaining process and outcomes. In particular, it focuses on the value of optimism as a strategic posture in determining the division of surplus in a two-agent non-cooperative bargaining environment.

To formalize, I consider the celebrated model of Rubinstein [1982] as a basic framework. The model considers two agents, each with her own impatience level, seeking to divide fixed surplus under a strict procedure of infinitely repeated alternating offers. The complete information assumption drives desirable results under this setting, yielding unique solution of immediate agreement. The simple and insightful model invited many modifications to better represent uncertainty or delays that are prevalent.

Yildiz [2003] and Yildiz [2004] extend the model by letting the players hold subjective beliefs about the "recognition process." In his model, each player believes he will be recognized as proposer in period t with probability P_t^1 and P_t^2, respectively, which don't necessarily add up to 1. Such departure from common prior assumption significantly alters the bargaining process. Yildiz [2003] formulates the model, defining optimism as a social status in which the sum of each players' belief about his recognition probability exceeds 1. He further postulates certain conditions on the level and evolution of optimism under which immediate agreement still occurs. Yildiz [2004] discusses the flip side of this "immediate agreement theorem," and studies what happens when the aforementioned conditions do not hold. With added structure, he finds that when the two players' prior beliefs are different and firm enough, they wait until the realization of recognition process brings their posteriors close enough for an agreement. With strategic choice of Bayesian updating mechanism, Yildiz [2004] shows that while the players disagree about the division, they agree on the duration of delay regardless of the realization.

On another strand of research, Abreu and Gul [2000] introduces asymmetry of information to the Rubinstein framework. In their model, each player has known probability of being irrational type who requires fixed—also known—share of the surplus. The incomplete information about the other player's private type dampens the game into a war of attrition, in which each player pretends to be the irrational type and requires a larger share of the surplus until the weaker player gives in. In continuous-time setting, the authors show that a player's bargaining power is increasing in his probability of being irrational and in the opponent's impatience level. They link this result with Myerson [1991]'s observation that "influence of asymmetric information overwhelms the effect of impatience in determining the division of surplus."

This linkage is partly what motivates my paper. Abreu and Gul [2000] considers "hard" information. In their model, the difference between the types dictates whether there is any gain to bargaining, as playing against an irrational type renders any effort towards bargaining meaningless. My paper asks whether the same relationship holds over "softer" information. Does an asymmetry in softer information, such as level of optimism, also overwhelm the effect of impatience? Or was the relationship driven by the absolute strength of information?

To answer the question, my paper borrows the framework of Yildiz [2004] but asks a fundamentally different question. Rather than viewing optimism as a social status over which both players have complete knowledge, I introduce uncertainty to one of the player's level of optimism, who may use it as strategic posture to benefit in the division of surplus. Under such incomplete information setting, my paper asks whether agents have incentive to misreport their prior belief P_0^i. Unlike Abreu and Gul [2000]'s discussion over hard information, the types the agents can be, or act as to be, are still rational, and do not trivialize the game upon revelation.

To fix ideas, let me formalize my question in an example. Consider a royalty negotiation between a writer and a publisher who hold different beliefs about the expected success of the book. The publisher thinks the book will *fail* with probability P_0^1 while the writer thinks the book will *succeed* with probability P_0^2, with $P_0^1 + P_0^2 > 1$. Both agree that the up-front advance amount[1] should depend on how well the book sells. Yildiz [2004] shows that the parties will delay the settlement until enough evidence bring their beliefs closer together, or "wait to persuade." Now, suppose writer A co-authored the book with another, more optimistic writer B (with $P_0^{2'} > P_0^2$). The bargaining process of A and B against the publisher are separate but the realization of P—the gradual revelation of the book's success—is the same for both A and B. In this setting, does A envy B? If so, does A want to adopt B's optimism as a strategy against his real belief? *Can* he do it?

Note that the three questions are substantively different. The last two questions only have grounds if the answer to the first is a yes. As will be shown, however, it is not trivially

[1] Writers and publishers often agree on a royalty with advance up front, where the publisher pays a fixed amount first, and later deducts it from future royalties. I consider the advance rather than the royalty as an object of bargain because it better represents division of fixed surplus that does not depend on recognition process.

so. While B will settle at a more favorable posterior, leading her to get larger portion of the up-front advance, she will achieve the settlement at a later date than A, shrinking the size of the pie by that much. It is important to notice at this point that it is writer A who evaluates both scenarios ex-ante, which is different from purely comparing the utility of A and B at the end of each bargaining process. In section 3, the paper shows that the first effect dominates the second effect: A always envies B. In the same section, I also show that *envy* is a lower bound to the *incentive to play optimism*. In other words, if A already envies B's bargaining outcome, the utility A thinks he will get if he can use B's type optimally for himself is, almost by construction, even larger.

As noted by language, evaluation of the first two questions is contemplation: writer A is debating whether to act as B *in simulation*. I call this process "contemplating optimism", and assume that the belief A reports is adopted without a doubt to the publisher, preserving the common knowledge setting of Yildiz [2004]. An interpretation of this assumption is that writer A simulates how the bargaining game would be played out *given* that he can completely fool the publisher.

The final question of whether he can successfully "play optimism" as a strategic posture is a completely different one from the previous two: this is where I introduce asymmetry of information to the game. Since it is shown that A always has incentive to be treated as B, the publisher becomes uncertain over whether the observed optimism is B's actual belief or A's strategic posture. This uncertainty over which game he is playing—bargaining against A or against B—prompts the publisher to behave differently from either of the equilibrium strategies in which beliefs were common knowledge. This paper does not attempt to fully characterize the solution concept of all possible equilibria of this new game. Instead, by iterated elimination of conditionally dominated strategies, I show that as long as both A and B are patient enough to cause delay in the first place, whether A can play optimism only depends on his prior probability of being type B. Furthermore, I show that the second-ordered nature[2] of the difference in types results in full revelation of uncertainty before the game ends when the publisher plays against writer B.

The rest of the paper is organized as follows. Section 2 lays out the model, with some borrowed results from Yildiz [2004], and defines objects of interest such as *envy* and *conjectured utility*. Section 3 describes how writer A contemplates optimism, and presents answers to the first two research questions. Section 4 modifies the model to playing optimism game, and demonstrates two features of an equilibrium outcome that leads to the conclusion. Technical algebra and proofs are relegated to the Appendix.

1.2 Model

1.2.1 Complete Information Framework

I consider three agents $N = \{1, 2, 2'\}$ in discrete time space $T = \{0, 1, 2, \cdots\}$ with common discount factor $\delta \in (0, 1)$.[3] I take $U = \{u \in [0, 1]^2 | u^1 + u^2 \leqslant 1\}$ and $U' = \{u \in [0, 1]^2 | u^1 + u^{2'} \leqslant 1\}$ to be the set of all feasible expected utility pairs.[4]

The bargaining between agent 1 and 2 proceeds in the following way. In each period t, Nature recognizes a player as proposer. The proposer offers a utility pair $u = (u^1, u^2) \in U$. If the other player accepts, the game ends, yielding the agents $\delta^t u = (\delta^t u^1, \delta^t u^2)$. If the other player does not accept, the game moves on to period $t + 1$. If the players never agree, they both get 0. The separate bargaining between agent 1 and $2'$ would proceed in the same way, with u^2 being replaced by $u^{2'}$. The notational distinction will be preserved over the entire paper, with superscript $'$ referring to parameters in bargaining between 1 and $2'$, and naked

[2] The types are only indirectly linked to the payoffs they receive, unlike Abreu and Gul [2000] model where a type directly corresponds to the payoff.

[3] To avoid verbal confusion, I refer to agent 2 as a he and $2'$ as a she.

[4] Even though utility of agent 1 is present in both sets, agent 1 either enters bargaining game against agent 2 or $2'$ and never together. Therefore, notation works as is in the complete information case.

letters or superscript * referring to parameters in bargaining between 1 and 2.[5]

Following Yildiz [2003]'s choice, I model the players' beliefs as following beta distributions to facilitate discussion of learning process. In particular, I assume that at the beginning of time t, if agent i observes agent 1 propose m times so far, he assigns probability $\frac{\overline{m}_i+m}{n+t}$ to the event that agent 1 gets recognized at any time $s \geqslant t$. Fixing any positive integers $1 \leqslant \overline{m}_2' < \overline{m}_2 < \overline{m}_1 \leqslant n-2$ captures the desired prior belief system that there exists optimism between agent 1 and 2, and even greater optimism between agent 1 and $2'$. Note that at time 0, the agents' prior beliefs about agent 1's recognition probability share the same denominator n, positing that the agents have the same level of firmness in their differing prior beliefs. Such structure enables us to distinguish the agents by the following primitives:

- Distance in beliefs: $\Delta \equiv \overline{m}_1 - \overline{m}_2, \quad \Delta' \equiv \overline{m}_1 - \overline{m}_2'$
- Prior beliefs of one's own recognition probability:
  $P_0^1 = \frac{\overline{m}_1}{n}, \quad P_0^2 = \frac{n-\overline{m}_2}{n}, \quad P_0^{2'} = \frac{n-\overline{m}_2'}{n}$
- Posterior beliefs at t after agent 1 makes m offers:
  $P_t^1(m) = \frac{\overline{m}_1+m}{n+t}, \quad P_t^2(m) = \frac{n+t-\overline{m}_2-m}{n+t}, \quad P_t^{2'}(m) = \frac{n+t-\overline{m}_2'-m}{n+t}$
- Level of optimism:
  $y_t(m) = P_t^1(m) + P_t^2(m) - 1 = \frac{\Delta}{t+n} > 0, \quad y_t'(m) = \frac{\Delta'}{t+n} > y_t(m)$

1.2.2 Existing Results Under Complete Information

In the absence of agent $2'$, Yildiz(2004) solves the payoff-equivalent subgame-perfect equilibrium of the two-agent bargaining model. He shows, by iterated elimination of conditionally dominated strategies, that all SPEs of this model are payoff equivalent:

[5]Note that although the model consists of 3 agents, bargaining is always between two agents, so the game can be thought of as two separate two-agent bargaining game where the outcomes and procedures are disclosed to everyone. I chose to represent it as three agents in order to avoid notational verbosity, as utility of agent 1 is not an object of interest until section 4.

Lemma 1.1 (Yildiz [2004] Lemma 2)**.** *Given any* (m,t) *and* i*, there exists a unique* $V_t^i(m) \in [0,1]$ *such that, in any SPE, the continuation value of* i *at* (m,t) *is* $V_t^i(m)$*.*

Proof. See Yildiz [2003] Theorem 1. □

Utilizing the unique payoffs, Yildiz [2004] shows that agents do not settle on an agreement until the discounted value of the next period's social surplus ($\delta S_t \equiv \delta V_t^1 + \delta V_t^2$) is less than 1. While the focus of Yildiz [2004] is on the agreed duration of delay, I relay some of his intermediate findings here as they facilitate our discussion in the next section.

Lemma 1.2 (Yildiz [2004] Lemma 3)**.** *For each* t*,* $S_t(m) = S_t(k)$ *for all* m *and* k*.*

This finding that the social surplus is deterministic while continuation value or division of surplus depends on the recognition process drives his result that the agents agree on a fixed settlement date until which time both "wait to persuade." At the agreed settlement date, the recognized agent i offers δV_{t+1}^j to j, who barely accepts. The following definition and results exhibit the evolution of the continuation value, an object of interest for my results:

Definition 1.1. A proposer rent is defined as: $R_t = \max\{1 - \delta S_{t+1}, 0\}$, and is thus also deterministic.

Lemma 1.3 (Yildiz [2004] Lemma 5)**.** *Given any* (m,t) *and* i*,*

$$\begin{aligned} V_t^i(m) &= P_t^i(m) \cdot \Lambda_t, \\ S_t &= (1 + y_t) \cdot \Lambda_t, \\ \textit{where} \;\; \Lambda_t &\equiv \sum_{s=t}^{\infty} \delta^{s-t} R_s \end{aligned}$$

Lemma 1.4. *The present value of all future rents,* Λ_t*, is tightly bounded by:*

$$B_{t-1} < \Lambda_t < B_t$$

where

$$B_t \equiv \frac{1}{1 + \delta y_{t+1}}$$

Finally, I relay another approximation result of Yildiz(2004).

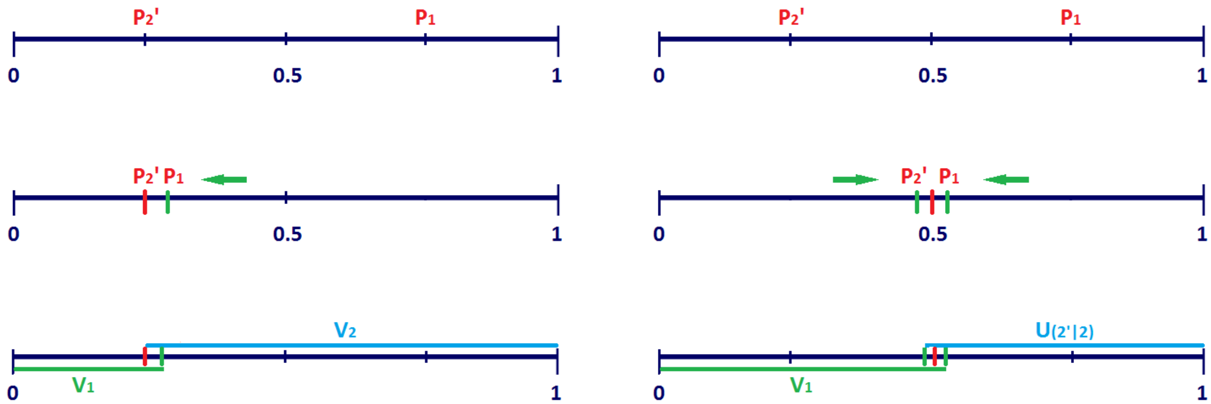
Lemma 1.5. *Settlement time* t^* *is the latest period that satisfies:*

$$t^* + n < \sqrt{\frac{\Delta\delta}{1-\delta}}$$

1.2.3 Conjectured Utility and Envy

The focus of this paper's investigation in the complete information game is not so much the outcome of the two-agent bargaining game. Rather, within the framework, this paper asks how agents involved in their own bargaining game evaluate the profitability of participating in the other bargaining game. To be specific, it studies whether agent 2, with his less optimistic beliefs, envies the bargaining outcome of the more optimistic agent $2'$ against the same agent 1. As agent 2 and $2'$ have different beliefs about the world, how they evaluate, ex-ante, the bargaining outcome of their own and the other's also differs.

In particular, I define *conjectured utility* $U(\overline{m}_2'|\overline{m}_2)$ as ex-ante evaluation of agent 2 of what would happen in his view of the world if he can completely trick agent 1 that he is agent $2'$, and mimic the strategy of $2'$ against 1. It is worth emphasizing again that this *conjectured utility* is a different object from the expected utility of agent $2'$ in her bargaining against agent 1, because her type and strategy are played out in the belief of agent 2. The following illustration demonstrates this distinction:



Evaluated by agent $2'$ Evaluated by agent 2

Figure 1.1: Illustration of Posterior Updating with $P_0^1 = \frac{3}{4}, P_0^2 = \frac{1}{2}, P_0^{2'} = \frac{3}{4}$

This is an illustration of the difference between $\mathbb{E}_{t=0}^{2'}[\delta^{t'} V_{t'}^{2'}]$ (left) and $U(\overline{m}_2'|\overline{m}_2)$ (right). The top bars show each agent's reported belief about the recognition probability of agent 1. Note that both agent $2'$ on the left and agent 2 on the right report the same prior of $P_0^{2'} = \frac{3}{4}$. The mid-left bar shows projection of agent $2'$ of how the realization will persuade agent 1 at their time of settlement t'. The division of surplus she simulates at t' is illustrated on the bottom-left. The mid-right bar shows projection of agent 2 of how the realization will persuade agent 1 at t'. Although he reported $P_0^{2'}$, he believes the realization will happen around $P_0^2 = \frac{1}{2}$, resulting in simulated posteriors that are drastic different from the ones shown in mid-left bar. Note that due to the same firmness of belief, both simulations take exactly t' periods to settle. The bottom-right bar shows the value of $U(\overline{m}_2'|\overline{m}_2)$ before discounting, which is clearly different from $V_{t'}^{2'}$ on the bottom-left.

The next lemma formally defines and calculates the conjectured utility.

Lemma 1.6. *Conjectured utility is given by:*

$$\begin{aligned} U(\overline{m}_2'|\overline{m}_2) &= \mathbb{E}_{t=0}\left[\delta^{t'} V_{t'}^{2'} \Big| \frac{\overline{m}_2}{n}\right] \\ &= \delta^{t'} \sum_{m=0}^{t'} \psi\big(m, t'\big|\frac{\overline{m}_2}{n}\big)\Big(P_{t'}^2(m)\big(1-\delta V_{t'+1}^{1'}(m)\big) + \big(1-P_{t'}^2(m)\big)\delta V_{t'+1}^{2'}(m+1)\Big) \end{aligned}$$

where

- $\psi\big(m, t'\big|\frac{\overline{m}_2}{n}\big) \equiv \binom{t'}{m}\big(\frac{\overline{m}_2}{n}\big)^m\big(\frac{n-\overline{m}_2}{n}\big)^{t'-m}$ *is the binomial probability of m occurrence in t' tries with $\frac{\overline{m}_2}{n}$ probability of success,*
- $V_{t'+1}^{1'}(m) = \frac{\overline{m}_1+m}{t'+n+1} \cdot \Lambda'_{t'+1}$
- $V_{t'+1}^{2'}(m+1)\frac{t'+n-\overline{m}_2'-m}{t'+n+1} \cdot \Lambda'_{t'+1}$ *from Lemma 1.2*

Remark. Conjectured utility is a simulated evaluation of agent 2 of what would happen if both agent 1 and himself act as if he were agent $2'$. $\psi\big(m, t'\big|\frac{\overline{m}_2}{n}\big)$ represents the probability agent 2 thinks the realization will be panned out with. In each state of the world, represented by m, agent 2 updates two versions of posteriors: one for reporting how agent $2'$ would have reacted, and another in his mind starting from his true prior. At the time of settlement, agent 2 evaluates the probability of recognition according to his true posterior, hence the terms $P_{t'}^2(m)$ and $1-P_{t'}^2(m)$ instead of $P_{t'}^{2'}(m)$ and $1-P_{t'}^{2'}(m)$. In contrast, the time of settlement and actual division of surplus follow the strategy and not the belief of agent 2, thus being t' and $1-\delta V_{t'+1}^{1'}(m)$ $\big($or $\delta V_{t'+1}^{2'}(m+1)\big)$, respectively. □

With this formal treatment of simulation, I define envy as follows.

Definition 1.2. Agent 2 with $P_0^2 = \frac{\overline{m}_2}{n}$ **envies** agent $2'$ with $P_0^{2'} = \frac{\overline{m}_2'}{n}$ if:

$$U(\overline{m}_2'|\overline{m}_2) > U(\overline{m}_2|\overline{m}_2).$$

1.3 Contemplating Optimism

In this section, I utilize the carefully defined conjectured utility to show that agent 2 always envies the more optimistic agent $2'$. The premise behind the intuition of this result is that agent 2 believes the prior belief of agent 1 is unacceptably flawed. Otherwise, agent 2 would feel no need to persuade in the first place. Therefore, I restrict the initial level of optimism to the levels in which agents fail to agree at time 0. In this case, according to Yildiz [2004], agent 2 will postpone agreement until the realization, which he believes will happen around his prior belief P_0^2, convinces agent 1 to update his posterior closer to, but not quite equal to, P_0^2.

This section shows that in my model, agent 2 contemplates a better option. By falsely reporting his prior as $P_0^{2'}$—which is bigger than P_0^2—he can convince agent 1 that agreement will not happen until posterior of agent 1 gets even closer to P_0^2. This way, even with the same realization, he can persuade agent 1 to settle at a more favorable posterior to himself, albeit at a later, and thus further discounted, settlement date.[6] The main result will show that the positive effect of more favorable posterior always dominates the time effect of discounting. Before positing the main result, the following paragraph demonstrates how the bargaining game which agent 2 contemplates pans out.

In the perfect information game between agent 1 and agent $2'$, elimination of conditionally dominated strategies dictates that the following happens. At time 0, $\delta S_1' = \delta V_1^{1'} + \delta V_1^{2'} > 1$, which means that both agents prefer waiting for the next period to offering each other his/her discounted continuation value. Such optimism decreases as evidence brings their posteriors close together, decreasing the perceived social surplus, until t' when $S_{t'+1}' \leqslant \frac{1}{\delta}$. Up to this point, prescribing any action to the proposer will constitute an equilibrium strategy as long as the opponent does not want to accept it. I will assume, without loss of generality, that

[6] Remember, we are assuming the reported prior beliefs are accepted as common knowledge. Discussion of how agent 1 reacts to this pretension possibility is postponed to next section, after showing that agent 2 has incentive to pretend in the first place.

at any time $t < t'$, the recognized agent i offers $1 - \delta V^i_{t+1}$, which is less than δV^j_{t+1}, and gets rejected. At t', the recognized agent i offers δV^j_t, which is accepted, and enjoys proposer rent $R_t = \max\{1 - \delta S_{t+1}, 0\}$. Contemplating this game leads to the following theorem.

Theorem 1.1. *An optimistic agent always envies more optimistic agents. Equivalently:*

$$U(\overline{m}'_2|\overline{m}_2) > U(\overline{m}_2|\overline{m}_2), \quad \forall\ \overline{m}'_2 > \overline{m}_2.$$

Sketch of proof. (See Appendix for technical details)

Expanding $U(\overline{m}_2|\overline{m}_2)$ to match the form of conjectured utility, we are comparing[7]:

1) $\delta^{t'} \sum_{m=0}^{t'} \psi\big(m, t'\big|\frac{\overline{m}_2}{n}\big)\Bigg(P^2_{t'}(m)\big(1 - \delta V^{1'}_{t'+1}(m)\big) + \big(1 - P^2_{t'}(m)\big)\delta V^{2'}_{t'+1}(m+1)\Bigg)$

2) $\delta^{t^*} \sum_{m=0}^{t^*} \psi\big(m, t^*\big|\frac{\overline{m}_2}{n}\big)\Bigg(P^2_{t^*}(m)\big(1 - \delta V^{1}_{t^*+1}(m)\big) + \big(1 - P^2_{t^*}(m)\big)\delta V^{2}_{t^*+1}(m+1)\Bigg)$

Summing over different number of binomial possibilities (t' tries vs. t^* tries) complicates the comparison a lot. Since the binomial combinations share the same probability $\frac{\overline{m}_2}{n}$, we can simplify by comparing where the most mass is. Defining $m' \equiv \frac{\overline{m}_2}{n}t'$ and $m^* \equiv \frac{\overline{m}_2}{n}t^*$ simplifies our comparison to the following:

1) $\underbrace{\delta^{t'}}_{\text{A'}} \Bigg[\underbrace{P^2_{t'}(m')}_{\text{B'}} \Big(\underbrace{1 - \delta V^{1'}_{t'+1}(m')}_{\text{C'}} \Big) + \big(1 - P^2_{t'}(m')\big) \underbrace{\delta V^{2'}_{t'+1}(m'+1)}_{\text{D'}} \Bigg]$

2) $\overbrace{\delta^{t^*}}^{\text{A}} \Bigg[\overbrace{P^2_{t^*}(m^*)}^{\text{B}} \Big(\overbrace{1 - \delta V^{1}_{t^*+1}(m^*)}^{\text{C}} \Big) + \big(1 - P^2_{t^*}(m^*)\big) \overbrace{\delta V^{2}_{t^*+1}(m^*+1)}^{\text{D}} \Bigg]$

- B' = B = $\frac{\overline{m}_2}{n}$: Agent 2 thinks his real posterior will be the same as his prior.
- C' $> 1 - \delta\frac{\overline{m}'_2+m'}{n+t'+1}\frac{t'+n+1}{t'+n+1+\delta\Delta'} > 1 - \delta\frac{\overline{m}_2}{n}\frac{t^*+n+2}{t^*+n+2+\delta\Delta} > C$:
    - from Lemma 1.3 and 1.4 (See Appendix for detail)

[7] t' is the predetermined settlement time between agent 1 and $2'$ as before, and t^* is the predetermined settlement time between agent 1 and 2.

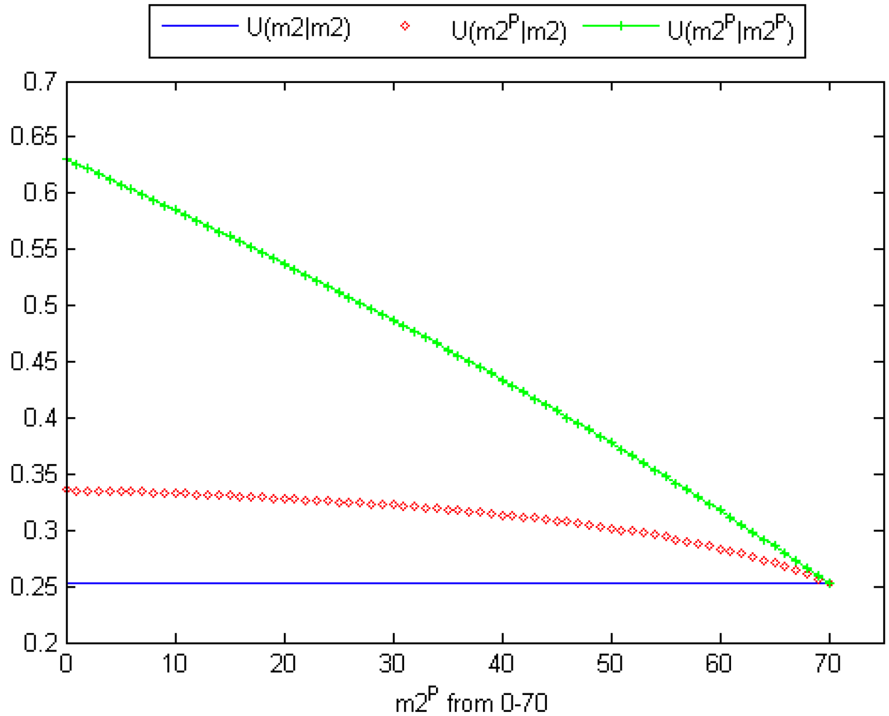
  - C' dominates C almost linearly in $\Delta' - \Delta = \overline{m}_2 - \overline{m}_2'$
  - D' dominates D in similar manner

- A > A' : $t' + n \approx \sqrt{\frac{\delta\Delta'}{1-\delta}}, \quad t^* + n \approx \sqrt{\frac{\delta\Delta}{1-\delta}}$ from Lemma 1.5

In words, t' and t^* are set such that agents are willing to suffer δ for change in posterior up to the time. For agent 2, starting from a more favorable prior facilitates movement of posterior immensely faster than the current updating so that he is willing to wait for a duration that the more optimistic agent $2'$ would have waited even with her slower updating.

□

Theorem 1.1 states that agent 2 would prefer imitating extreme optimism of agent $2'$ and receiving the conjectured utility $U(\overline{m}_2'|\overline{m}_2)$ to getting his own bargaining outcome. A simulated comparison of the three utilities—$U(\overline{m}_2|\overline{m}_2), U(\overline{m}_2'|\overline{m}_2)$, and $U(\overline{m}_2'|\overline{m}_2')$—are presented in the following figure.



$$n = 100, \quad \overline{m}_2 = 70, \quad \overline{m}_1 = 90, \quad \delta = .999$$

Figure 1.2: Illustration of Envy: $U(\overline{m}_2|\overline{m}_2) < U(\overline{m}_2'|\overline{m}_2) < U(\overline{m}_2'|\overline{m}_2')$

Based on the equilibrium strategy of agent 1 against agent $2'$, we can easily see that agent 2 can do even better. In particular, while reporting the optimism level of agent $2'$, agent 2 is actually not as optimistic about his future payoffs as agent $2'$, for there is a discrepancy between his reported posterior and his actual posterior in his mind. Therefore, if t' is set such that agent $2'$ no longer deems it worthwhile to wait another period even with her higher level of optimism, agent 2 would wish to settle earlier. Yildiz(2004) shows that if agent 2 was in a bargaining game against agent 1 truthfully, that time in which he wishes to settle is $t*$. Contemplating optimism alters this consideration in two major ways. First, agent 2 contemplates playing as agent $2'$, and therefore anticipates agent 1 to hold lower continuation value—$V_t^{1'}$. Second, the option of receiving conjectured utility in the future increases continuation value of agent 2 compared to his original bargaining—to $V_t^{2c.o.}$. The following corollaries formalize this idea.

Corollary 1.1.1. *The option of receiving conjectured utility makes agent 2 less willing to settle every period. Equivalently:*

$$V_t^{2c.o.}(m) > V_t^2(m), \quad \forall m, t \leqslant t'.$$

Proof. I prove this corollary in two steps:

i) $V_t^{2c.o.}(m) > V_t^2(m), \quad \forall m, t < t^*$.

By construction of t^*, agent 2 does not want to settle after any history (m,t) for $t < t^*$ in his original bargaining. In other words, after any history (m,t), treating his posterior belief as the new prior and applying Theorem 1.1 show that agent 2 still envies more optimistic agent $2'$ regardless of the realization m as long as $t < t^*$. Formally,

$$\begin{aligned} V_t^2(m) &= P_t^2(m)\big(1-\delta V_{t+1}^1(m)\big) + \big(1-P_t^2(m)\big)\delta V_{t+1}^2(m+1) \qquad (1.1)\\ &= P_t^2(m)\big(1-\delta S_{t+1}\big) + \delta\mathbb{E}_t\big[V_{t+1}^2|P_t^2(m)\big]\\ &= \underbrace{P_t^2(m)R_t}_{\text{A}} + \underbrace{\delta\mathbb{E}_t\big[V_{t+1}^2|P_t^2(m)\big]}_{\text{B}} \end{aligned}$$

Since he does not settle before t^* ($\delta S_{t+1} > 1$), A term is 0 in the original bargaining. Since R_t is non-negative by definition, this term can never decrease as he contemplates optimism. The option of receiving conjectured utility at time t', however, unambiguously increases expected future benefits—B term—thus showing desired property that $V_t^{2c.o.}(m) > V_t^2(m)$.

ii) $V_t^{2c.o.}(m) > V_t^2(m), \quad \forall m, t < t'$.

i) shows the mechanism through which contemplating optimism increases continuation value of agent 2. Here I show a bit more formally that it holds for all $t < t'$. Considering the aforementioned equilibrium strategy, $V_t^{2c.o.}(m)$ is the continuation value of agent 2 who evaluates his payoff of pretending to be agent $2'$ but using his belief. Formally,

$$V_t^{2c.o.}(m) = P_t^2(m)\big(1 - \delta V_{t+1}^{1'}(m)\big) + \big(1 - P_t^2(m)\big)\delta V_{t+1}^{2'}(m+1)$$

Comparing the expression with (1.1), it suffices to show that $V_{t+1}^{1'}(m) < V_{t+1}^{1}(m)$, and $V_{t+1}^{2'}(m) > V_{t+1}^{2}(m)$. Lemma 1.3 and 1.4 render this relationship easy to show, as proof of Theorem 1.1 demonstrates a similar property. The intuition is that the direct influence of increased difference in priors dominates the influence of corresponding decrease in future proposer rents. Refer to the proof of Theorem 1.1 in the appendix for the exact algebra.

□

Corollary 1.1.2. *The revised settlement date—$t^{c.o.}$—from which time agent* 2 *who contemplates optimism is willing to settle with agent* 1 *satisfies:*

$$t^{c.o.} \leqslant t'.$$

Proof. Consider the aforementioned equilibrium strategy of agent 1 in his bargaining against agent $2'$. At any time $t < t'$, he offers $1 - \delta V_{t+1}^{1'}$ if he is recognized, and only accepts offers higher than or equal to $\delta V_{t+1}^{1'}$ if $2'$ is recognized. Under this specification, $t^{c.o.}$ satisfies:

$$\delta V_{t^{c.o.}}^{1'}(m) + \delta V_{t^{c.o.}}^{2c.o.}(m) > 1 \quad \text{and} \quad \delta V_{t^{c.o.}+1}^{1'}(m) + \delta V_{t^{c.o.}+1}^{2c.o.}(m) \leqslant 1 \tag{1.2}$$

Note that the belief updating system dictates that the difference in posteriors only decrease as time passes and in the same speed for all realizations of m.[8] This property results in diminishing level of optimism and, in turn, decrease in perceived social surplus. Therefore, $t^{c.o.}$ that satisfies (1.2) always uniquely exists as long as initial delay is present. Similarly,

[8]Refer to the model section of Yildiz(2004) for detailed discussion of this property.

the settlement date t' between agent 1 and $2'$ uniquely exists and satisfies:

$$\delta V_{t'}^{1'}(m) + \delta V_{t'}^{2'}(m) > 1 \quad \text{and} \quad \delta V_{t'+1}^{1'}(m) + \delta V_{t'+1}^{2'}(m) \leqslant 1$$

Comparing the expression with (1.2), it suffices to show that $V_t^{2'} > V_t^{2c.o.}$ regardless of realization, which is straightforward from $P_t^{2'} > P_t^2$:

$$\begin{aligned} V_t^{2'}(m) &= P_t^{2'}(m)\big(1 - \delta V_{t+1}^{1'}(m)\big) + \big(1 - P_t^{2'}(m)\big)\delta V_{t+1}^{2'}(m+1) \\ &> P_t^{2}(m)\big(1 - \delta V_{t+1}^{1'}(m)\big) + \big(1 - P_t^{2}(m)\big)\delta V_{t+1}^{2'}(m+1) = V_t^{2c.o.}(m), \quad \forall m \end{aligned}$$

I pinpoint $t^{c.o.}$ here using one specific equilibrium strategy. Although different off-the-equilibrium-path strategies we imbue agent 1 will result in a less narrowly specified range of $t^{c.o.}$, we can easily check that the desired relationship holds for all strategies, as they all converge to the same equilibrium strategy at t'. □

Note that neither corollaries pinpoints the value of $V_t^{2c.o.}$ or $t^{c.o.}$. It is because the current model of contemplating optimism forces agent 2 to intrude into another equilibrium, resulting in his continuation values being dependent on the specification of off-the-equilibrium-path strategies we imbue agent 1 with. However, as proof of Corollary 3.1.2 hints, the merit of this contemplating optimism model is that the core results hold for every off-the-equilibrium-path strategies we consider as long as they constitute a subgame-perfect equilibrium in the original game. The next theorem, which naturally follows, sums up these core results into a main result that agent 2 always has an incentive to play optimism in his contemplation.

Theorem 1.2. *An optimistic agent always wants to play more optimistic.*

Proof. This theorem naturally follows from Theorem 1.1 if "envy" is the lower bound of the simulated utility of agent 2 from "playing" $\overline{m}_2'$. Notice that by definition, agent 2 only settles with agent 1 in contemplating optimism when he deems it more profitable than waiting for the conjectured utility. Therefore, $t^{c.o.} \leqslant t'$ implies the following relationship, proving the desired property:

$$\text{Utility from "playing" } \overline{m}_2' = \max\left\{ \underbrace{\delta^{t^{c.o.}+1} \cdot V^{2c.o.}_{t^{c.o.}+1}}_{A},\ U(\overline{m}_2'|\overline{m}_2) \right\} \geqslant U(\overline{m}_2'|\overline{m}_2)$$

where A corresponds to the option value of settling at $t^{c.o.}$—$\delta^{t^{c.o.}} \cdot \delta V^{2c.o.}_{t^{c.o.}+1}$. Since our choice of $\overline{m}_2$ and $\overline{m}_2'$ were only restricted by $\overline{m}_2' < \overline{m}_2$ and $t^* > 0$ over the discussion, the theorem holds for any optimistic agent who delays. □

1.4 Playing Optimism

The discussion in section 3 and Theorem 1.2 can be interpreted as a negative result that in presence of heterogeneous levels of optimism among agents, reporting his true belief is not incentive-compatible for agent 2. This, in turn, sheds doubt on the generality in assuming perfect information in games of uncommon priors such as Yildiz(2004). This section addresses the concern by imploring the full game of asymmetric information, and describes an equilibrium behavior which accommodates the uncertainty in the level of optimism agent 2 reports.

To do so, I introduce the following modifications to the model.

- Agent space is reduced to a set of two elements: $N = \{1, 2\}$
- Agent 1 can only be of one type: $\theta_1 \in \Theta_1 = \{\overline{m}_1\}$
- Agent 2 can be one of two types: $\theta_2 \in \Theta_2 = \{\overline{m}_2, \overline{m}_2'\}$
- The probability distribution over Θ_2 is summarized by: $\phi = \mathbb{P}(\theta_2 = \overline{m}_2') \in (0, 1)$

The discrete time space T and common discount factor δ, as well as the properties and related objects of $\overline{m}_2$ and $\overline{m}_2'$ are preserved. As before, superscript $'$ denotes parameters in common-knowledge bargaining game between 1 and $2'$, and naked letters or superscript

* denotes parameters in common-knowledge bargaining game between 1 and 2. Finally, I attach superscript ${}^{p.o.}$ to parameters in the current asymmetric information game and denote agent 1's subjective belief of ϕ at time t as $\hat{\phi}_t$.[9] Before I construct an equilibrium in this model, I present two results that must hold in any Perfect Bayesian Nash Equilibrium (henceforth PBE) of this game.

Lemma 1.7. *Agent 2 of type $\overline{m}_2$ is always more willing to settle than type $\overline{m}_2'$. Equivalently:*

$$V_t^{2p.o.}(m, \hat{\phi}_t) < V_t^{2'p.o.}(m, \hat{\phi}_t) \quad \forall m, t, \textit{and } \hat{\phi}_t \neq 0.$$

Proof. While equilibrium behavior of each agent has not been developed yet, we can easily predict that the continuation value of agent 1 at time t would depend not only on his posterior $P_t^1(m)$—which can be summarized by m—but also on his belief of agent 2's type, $\hat{\phi}_t$. Denoting this continuation value as $V_t^{1p.o.}(m, \hat{\phi}_t)$, we need to consider two cases:

i) If type $\overline{m}_2$ acts in a way that affects $\hat{\phi}_t$ differently from type $\overline{m}_2'$'s action, the statement naturally follows from $V_t^2(m) < V_t^{2'}(m)$.

ii) If both types' actions affect $\hat{\phi}_t$ in the same way, they would induce the same action from agent 1, and their continuation values would differ only by how they evaluate their future probabilities. Formally,

$$\begin{aligned} V_t^{2'p.o.}(m, \hat{\phi}_t) &= P_t^{2'}(m)\big(1 - \delta V_{t+1}^{1p.o.}(m, \hat{\phi}_{t+1})\big) + \big(1 - P_t^{2'}(m)\big)\delta V_{t+1}^{2'p.o.}(m+1, \hat{\phi}_{t+1}) \\ &> P_t^2(m)\big(1 - \delta V_{t+1}^{1p.o.}(m, \hat{\phi}_{t+1})\big) + \big(1 - P_t^2(m)\big)\underbrace{\delta V_{t+1}^{2'p.o.}(m+1, \hat{\phi}_{t+1})}_{\text{A}} \\ &= V_t^{2c.o.}(m, \hat{\phi}_t), \quad \forall m, \hat{\phi}_t \end{aligned}$$

where term A represents the fact that type $\overline{m}_2$ would have to act the same way as type $\overline{m}_2'$,

[9] To avoid notational verbosity, I use superscript ${}^{'p.o.}$ to denote parameters regarding agent 2 of type $\overline{m}_2'$ in the asymmetric information game. ex) $V_t^{2'p.o.}(m)$.

and thus be offered the same value. Our choice of $m, \hat{\phi}_t, \overline{m}_2$ and $\overline{m}'_2$ was arbitrary, so the statement holds generally. □

Lemma 1.8. *At $t = 0$, agent 1 is less willing to settle than he was in perfect information bargaining game against $2'$. Equivalently:*

$$V_1^{1p.o.}(m, \hat{\phi}_1) > V_1^{1'}(m) \quad \forall m,\ \hat{\phi}_1.$$

Proof. Regardless of future actions, $\hat{\phi}_0 = \phi$. Therefore, at $t = 0$, agent 1 knows that with $1 - \phi$ (> 0) probability, he is bargaining against type $\overline{m}_2$. Since Lemma 1.7 shows that type $\overline{m}_2$ is "cheaper" than type $\overline{m}'_2$, the statement naturally follows. □

Imposing sequential rationality immediately provides that the previous two lemmas hold for any PBE of this game. The following lemma is slightly less general but crucial in constructing the PBE of this game.

Lemma 1.9. *There exists a PBE in which agent 1 and agent 2 of type $\overline{m}'_2$ always make the same decision at t' as they would in their common knowledge bargaining game. Equivalently,*

$$V_{t'}^{1p.o.}(m, \hat{\phi}_{t'}) = V_{t'}^{1'}(m) \quad \forall m.$$

Proof. Note that in this sequential bargaining, history affects the decisions of agents only through their beliefs—$P_t^{ip.o.}(m)$ and $\hat{\phi}_t$. Since agents update $P_t^{ip.o.}(m)$ the same way as $P_t^i(m)$, if $\hat{\phi}_{t'} = 1$ and this belief is correct, agent 1 and agent 2 of type $\overline{m}'_2$ are essentially playing the same game as the common knowledge game of section 3 at t'. I prove that $\hat{\phi}_{t'} = 1$ by making two observations:

i) If $\hat{\phi}_s = 1$, then $\hat{\phi}_t = 1 \ \ \forall t > s$.

Suppose not: $\hat{\phi}_t \neq 1$. There are two possibilities. If $\hat{\phi}_s = 1$ was flawed, we are done. If $\hat{\phi}_s = 1$ was correct and agent 1 is indeed against type $\overline{m}'_2$, $\hat{\phi}_t \neq 1$ means that agent 2 acts sub-optimally between time s and t, contradicting the sequential rationality of agent

2 in PBE. Therefore, by proof by contradiction, if $\hat{\phi}_s = 1$ in any PBE of this model, then $\hat{\phi}_t = 1 \;\; \forall t > s$.

ii) $\hat{\phi}_s = 1$ for some $s \leqslant t'$.

Denote s as the earliest time that satisfies $\delta V_s^{2p.o.}(m, \hat{\phi}_s) + \delta V_s^{1p.o.}(m, \hat{\phi}_s) < 1$. From Lemma 1.7, we know that $V_{t'}^{2p.o.}(m, \hat{\phi}_{t'}) < V_{t'}^{2'p.o.}(m, \hat{\phi}_{t'}) \;\; \forall m$, which ensures that $s \leqslant t'$. Therefore, any strategy that leads to $\hat{\phi}_{t'} < 1$ and incentivizes agent 1 to play differently from common knowledge game is dominated by offering $\delta V_s^{2p.o.}(m, \hat{\phi}_s)$ at time $s-1$, which type $\overline{m}_2$ will always accept, thereby Bayesian updating to $\hat{\phi}_s = 1$. Similar reasoning prompts agent 2 of type $\overline{m}_2$ to always settle with agent 1 at time s, resulting in the same Bayesian updating. □

While I could not state with rigor that every possible PBE of this game satisfies Lemma 1.9, the intuition behind the relative ease with which we can solve this asymmetric information game is as follows. First of all, although the game proceeds sequentially, the payoff the two agents reap is instantaneous and not repeated. Such one-shot nature of the game strips the uninformed agent 1 of any capability to punish misreports, enabling us to focus on marginal decisions at each period. Second, not only can the types of agent 2 be used as signaling device, they also affect how agent 2 actually evaluates future outcome of the game. Therefore, even though type $\overline{m}_2$ has incentive to signal his type as $\overline{m}'_2$, he does not intend to play exactly the same as her until the end of the game. In other words, the bargaining power that type $\overline{m}_2$ can assume is somewhat bounded by his true type, saving agent 1 from being completely in the dark. The next theorem formalizes this idea by constructing a PBE of this game.

Theorem 1.3. *For all values of $\phi \neq 1$, there exists a PBE in which type $\overline{m}_2$ always settles with agent 1 at once at $t^{p.o.}$, irrespective of m.*

Proof. I prove this theorem by construction. I claim that the following constitutes a PBE:

- $\hat{\phi}_t = \begin{cases} \phi, & \forall t \leqslant t^{p.o.} \\ 1, & \forall t > t^{p.o.} \end{cases}$

- $\sigma^1_t(m, \hat{\phi}_t) = \begin{cases} \text{offer anything lower than } 1 - \delta V^{1p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t < t^{p.o.} \\ \text{and offer } \delta V^{2p.o.}_{t+1}(m, 0), & \forall t \text{ s.t. } t^{p.o.} \leqslant t < t' \\ \text{and offer } \delta V^{2'p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \text{at } t \geqslant t' \\ \text{or accept anything higher than } \delta V^{1p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t < t^{p.o.} \text{ or } \geqslant t' \\ \text{or accept anything higher than } \delta V^{1p.o.}_{t+1}(m, 0), & \forall t \text{ s.t. } t^{p.o.} \leqslant t < t' \end{cases}$

- $\sigma^2_t(\overline{m}_2, m, \hat{\phi}_t) = \begin{cases} \text{offer anything lower than } 1 - \delta V^{2p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t < t^{p.o.} \\ \text{and offer } \delta V^{1p.o.}_{t+1}(m, 0), & \forall t \geqslant t^{p.o.} \\ \text{or accept anything higher than } \delta V^{2p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t < t^{p.o.} \\ \text{or accept anything higher than } \delta V^{2p.o.}_{t+1}(m, 0), & \forall t \geqslant t^{p.o.} \end{cases}$

- $\sigma^2_t(\overline{m}'_2, m, \hat{\phi}_t) = \begin{cases} \text{offer anything lower than } 1 - \delta V^{2'p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t < t' \\ \text{and offer } \delta V^{1p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t \geqslant t' \\ \text{or accept anything higher than } \delta V^{2'p.o.}_{t+1}(m, \hat{\phi}_{t+1}), & \forall t \end{cases}$

Since agent 2 of type $\overline{m}_2$ always settles at $t^{p.o.}$, it is easy to check that the belief $\hat{\phi}_t$ is consistent with the hypothesized equilibrium strategy. The remaining pieces to be proved are i) sequential rationality of hypothesized equilibrium strategy given the consistent belief, and ii) the properties of $t^{p.o.}$.

From the proof of Lemma 1.9, we know that the given belief $\hat{\phi}_t$ leads to an optimal strategy of agent 1 and agent 2 of type $\overline{m}'_2$ playing as if they were in common knowledge game at time t'. In particular, the result of Lemma 1.9 states that $V^{1p.o.}_{t'}(m, \hat{\phi}_{t'}) = V^{1'}_{t'}(m)\ \forall m$. Starting from t' backwards, I iteratively formulate $V^{1p.o.}_t$ through the following procedure:

In the beginning of time t, assume all of the type $\overline{m}_2$ have already exited the game.[10] This leads to the belief that $\hat{\phi}_t = 1$, which, in turn, brings us back to Lemma 1.9 so that $V_t^{1p.o.}(m, \hat{\phi}_t) = V_t^{1'}(m)\ \forall m$. Move on to time $t-1$.

For $t = t'$, the procedure runs smoothly. On the other hand, notice that as time goes backwards, $V_t^{1p.o.}(m, \hat{\phi}_t)$ increases. Since we know that both types of agent 2 refuses to settle with agent 1 at time 0, there will come a time, denoted by $t^{p.o.}$, when the assumption of the procedure causes contradiction. Formally,

$$V^{1'}_{t^{p.o.}}(m) + V^{2p.o.}_{t^{p.o.}}(m, \phi) > \frac{1}{\delta}; \quad V^{1'}_{t^{p.o.}+1}(m) + V^{2p.o.}_{t^{p.o.}+1}(m, 1) \leqslant \frac{1}{\delta}$$

At $t^{p.o.}$, agent 1 thinks that with $\hat{\phi}_{t^{p.o.}} = \phi$ probability, he is against agent 2 of type $\overline{m}_2'$ with whom he does not wish to settle, and with $1 - \hat{\phi}_{t^{p.o.}} = 1 - \phi$ probability, he is against type $\overline{m}_2$ with whom he wishes to settle. Conveniently, the incentive for settlement is mutual, so that even if there is a slightest chance that he plays against type $\overline{m}_2$, it is dominant strategy for him to offer settlement which only the desired type $\overline{m}_2$ will accept. This leads to the following property (α):

$$\begin{aligned}
&V^{1p.o.}_{t^{p.o.}}(m, \phi) \\
&\quad = \phi V^{1'}_{t^{p.o.}}(m) + (1-\phi)\big[P^1_{t^{p.o.}}(m)\big(1 - \delta V^{2p.o.}_{t^{p.o.}+1}(m+1, 0)\big) + \big(1 - P^1_{t^{p.o.}}(m)\big)\delta V^{1p.o.}_{t^{p.o.}+1}(m, 0)\big] \\
&\quad > \phi V^{1'}_{t^{p.o.}}(m) + (1-\phi)\big[P^1_{t^{p.o.}}(m)\big(1 - \delta V^{2p.o.}_{t^{p.o.}+1}(m+1, 0)\big) + \big(1 - P^1_{t^{p.o.}}(m)\big)\delta V^{1p.o.}_{t^{p.o.}+1}(m, 1)\big] \\
&\quad = V^{1'}_{t^{p.o.}}(m) + (1-\phi)\big[P^1_{t^{p.o.}}(m)\big\{\big(1 - \delta V^{2p.o.}_{t^{p.o.}+1}(m+1, 0)\big) - \big(1 - \delta V^{2'p.o.}_{t^{p.o.}+1}(m+1, 1)\big)\big\}\big] \\
&\quad = V^{1'}_{t^{p.o.}}(m) + \delta(1-\phi)P^1_{t^{p.o.}}(m)\big\{\delta V^{2'p.o.}_{t^{p.o.}+1}(m+1, 1) - \delta V^{2p.o.}_{t^{p.o.}+1}(m+1, 0)\big\}
\end{aligned}$$

The property shows that $V^{1p.o.}_{t^{p.o.}}(m, \phi)$ is bounded below by a value that is always higher

[10] "All" in the Harsanyi sense of treating probability over types as distributions. i.e. there is 0 probability that type $\overline{m}_2$ remains until this time.

than $V^{1'}_{t^{p.o.}}(m)$ and linearly decreasing in ϕ. In particular, in the extreme values of ϕ,

$$\phi \to 0 \implies V^{1p.o.}_{t^{p.o.}}(m,\phi) \to V^{1}_{t^{p.o.}}(m) \tag{1.3}$$

$$\phi \to 1 \implies V^{1p.o.}_{t^{p.o.}}(m,\phi) \to V^{1'}_{t^{p.o.}}(m) \tag{1.4}$$

This implies, in turn, that $V^{1p.o.}_{t^{p.o.}}(m,\phi)$ is so high for low enough values of ϕ that agent 2 of type $\overline{m}_2$ prefers his original outcome in common knowledge case. Since type $\overline{m}'_2$ would never mimic type $\overline{m}_2$, this case reverts the game back to settlement at t^* solution, immediately satisfying this theorem. (Remember from last section that $t^* \not\propto m$.) On the other hand, the linear dependence on ϕ of this value—together with the fact that type $\overline{m}_2$ wants to play optimism for $V^{1p.o.}_{t^{p.o.}}(m,\phi)$ close to $V^{1'}_{t^{p.o.}}(m,\phi)$—implies that there exists a threshold ϕ^* above which type $\overline{m}_2$ chooses to mimic type $\overline{m}'_2$ until $t^{p.o.}$. For those $\phi > \phi^*$, the following property pinpoints the continuation values at any $t < t^{p.o.}$, which in turn pins down the offered utilities:

Denoting $S^{p.o.}_t(\hat{\phi}_t) \equiv V^{1p.o.}_t(m,\hat{\phi}_t) + V^{2p.o.}_t(m,\hat{\phi}_t)$, $S'^{p.o.}_t(\hat{\phi}_t) = V^{1p.o.}_t(m,\hat{\phi}_t) + V^{2'p.o.}_t(m,\hat{\phi}_t)$,

$$\begin{aligned}
V^{1p.o.}_t(m,\phi) &= P^1_t(m)\big[\phi\big(1-\delta S'^{p.o.}_{t+1}(\phi)\big) + (1-\phi)\big(1-\delta S^{p.o.}_{t+1}(\phi)\big)\big] + \delta\mathbb{E}_t\big[V^{1p.o.}_{t+1}\big|P^1_t(m),\phi\big]\\
V^{2p.o.}_t(m,\phi) &= P^2_t(m)\big(1-\delta S^{p.o.}_{t+1}(\phi)\big) + \delta\mathbb{E}_t\big[V^{2p.o.}_{t+1}\big|P^2_t(m),\phi\big]\\
V^{2'p.o.}_t(m,\phi) &= P^{2'}_t(m)\big(1-\delta S'^{p.o.}_{t+1}(\phi)\big) + \delta\mathbb{E}_t\big[V^{2'p.o.}_{t+1}\big|P^{2'}_t(m),\phi\big]
\end{aligned}$$

As the perceived social surplus only depends on $\hat{\phi}_t$ and not on m, the $t^{p.o.}$ at which $\delta S^{p.o.}_t(\hat{\phi}_t)$ first shrinks below 1 does not depend on m for any $\phi > \phi^*$. This finishes the proof of existence of PBE with the property that type $\overline{m}_2$ always settles with agent 1 at once at $t^{p.o.} \not\propto m$ by construction. □

As detailed as the procedure in the proof of Theorem 1.3 is, property (α) drives the entire characterization of the equilibrium. In other words, as long as property (α) holds—or more weakly, as long as (1.3), (1.4), and the linear dependence on ϕ hold—I conjecture Theorem 1.3 will still hold. However, there are a couple of caveats in my treatment of this model that hinder me from generalizing Theorem 1.3 further.

First, the equilibrium concept is not robust. The proof of Theorem 1.3 shows that the constructed equilibrium is a PBE but not a sequential equilibrium. In order for it to be fully robust, I would like the equilibrium strategies to satisfy sequential rationality starting from *every* information set. This requires that I consider $\hat{\phi}_t = 1 - \epsilon$ for $t > t^{p.o.}$. However, Lemma 1.9 heavily depends on agent 1 and agent 2 of type $\overline{m}_2'$ playing as if their types were common knowledge. As soon as I entertain ϵ possibility that the belief is flawed, the agents' behaviors starting from t' backwards alter completely.

On this front, I still conjecture the following saving grace. The main intuition behind the Lemma 1.9 is that if agent 1 wishes to risk a rejection from type $\overline{m}_2'$ by making a greedy offer that only type $\overline{m}_2$ will accept at t', he is actually better off making that offer some time before, and making sure that he is against type $\overline{m}_2'$ when they reach t'. Even though a slight perturbation to the belief ruins all hopes for building common knowledge, an ϵ possibility that the opposing agent is still more willing to settle, and thus "cheaper", is not enough to incentivize agent 1 to risk making greedy offers at t'. Therefore, even though the constructed PBE is not robust to slight perturbation of beliefs, I believe that the equilibrium strategy laid out here can easily be modified correspondingly to construct a robust sequential equilibrium.

Furthermore, I do not consider mixed strategies from agent 2 of type $\overline{m}_2$. In constructing the equilibrium strategies in the proof of Theorem 1.3, I break the indifference towards always accepting an offer. While this is line with the literature, such practice is more restrictive in this game because there is discrepancy between $V^{1p.o.}_{t^{p.o.}+1}(m, 0)$ and $V^{1p.o.}_{t^{p.o.}+1}(m, 1)$. In other words, if I let agent 2 of type $\overline{m}_2$ play mixed strategy between accepting and rejecting when offered his discounted continuation value, there may be slight room to profit by playing optimism for one more period. Modifying $\hat{\phi}_t$ accordingly to this mixed strategy will constitute another PBE which is different from the one constructed here. However, it will not alter the core result that the time from which agent 2 of type $\overline{m}_2$ starts to accept/offer settlement with positive probability does not depend on the realization m.

Besides these caveats, this section successfully shows that there exists a PBE in which agents reveal their types at a predetermined period regardless of the realization of m. Instead, the equilibrium behaviors largely depend on ex-ante probability of type distribution ϕ. The proof of Theorem 1.3 shows that there exists ϕ^* above which type $\overline{m}_2$ always chooses to play optimism, and thus benefits by the presence of type $\overline{m}_2'$. Property (α) also shows that as ϕ increases, type $\overline{m}_2$ benefits more and reveal their type at a later time.

1.5 Conclusion

This paper studies two modifications to the Rubinstein bargaining model. On the one hand, I recognize the ease in approach and intuitive perturbation of considering uncommon prior beliefs. While entertaining the possibility of difference in prior can explain delay even in common knowledge setting, I show in section 3 that it is not likely a robust model, for agents will always have incentive to misreport their prior. Instead, I propose introducing an asymmetry of information to this model in which agents only differ in their subjective prior beliefs of recognition probability. Imbuing one side of agents an opportunity to be of more optimistic type, I find that such asymmetry of soft information still incentivizes the informed agents to signal their types as being more optimistic. However, the agents are still somewhat bound by their true types in evaluating future payoffs, resulting in full revelation before the game ends for the more optimistic type. This paper formalizes this intuition by constructing a PBE and analyzing its dependence on the parameters.

It finds that an asymmetry of information in the agents' level of optimism still governs the agents' behaviors. Section 4 shows that the decision of whether to play optimism, how long to play optimism, and how much one can gain by playing optimism all heavily depend on the initial distribution of types. The effect of impatience level only acts as an implicit threshold that allows the difference in types to manifest itself as choosing different actions (i.e. the agents must not wish to settle at time 0 for this model to have merit). It bolsters Abreu and Gul [2000]'s assessment of Myerson's quote in that the influence of asymmetry in soft information was enough to overwhelm the impact of impatience level.

The results of the paper do not depend on specific choice of parameters, preserving the generality necessary for such qualitative assessments. The paper is by no means exhaustive, however. First of all, the core results, such as independence from the realization m, takes heavy advantage of the convenience brought forth by the assumed belief structure. While the possible realizations and Bayesian updating mechanism are general, the paper depends on the beta distribution of agents' beliefs to be able to separate evolution of optimism from specific realizations. The first step in advancing this model would be to consider other belief structures and to study how the results can be accommodated to them. The paper could also be enhanced by making the solution concept more robust, or generalizing the results to all possible equilibria of the model. In the meantime, I conclude this paper by echoing its two main results: Optimistic agents always want to exaggerate their level of optimism. In the presence of enough agents with extreme optimism, the moderately optimistic can and will benefit by playing optimism.

1.6 Appendix: Addendum to Proof of Theorem 1.1

From proof of Theorem 1.1 in section 3, showing that $U(\overline{m}_2'|\overline{m}_2) > U(\overline{m}_2|\overline{m}_2)$ is equivalent to comparing the following expressions:

$$1)\ U(\overline{m}_2'|\overline{m}_2) \approx \underbrace{\delta^{t'}}_{\text{A'}} \Bigg[\underbrace{P_{t'}^2(m')}_{\text{B'}} \big(\underbrace{1 - \delta V_{t'+1}^{1'}(m')}_{\text{C'}} \big) + \big(1 - P_{t'}^2(m')\big) \underbrace{\delta V_{t'+1}^{2'}(m'+1)}_{\text{D'}} \Bigg]$$

$$2)\ U(\overline{m}_2|\overline{m}_2) \approx \overbrace{\delta^{t^*}}^{\text{A}} \Bigg[\overbrace{P_{t^*}^2(m^*)}^{\text{B}} \big(\overbrace{1 - \delta V_{t^*+1}^{1}(m^*)}^{\text{C}} \big) + \big(1 - P_{t^*}^2(m^*)\big) \overbrace{\delta V_{t^*+1}^{2}(m^*+1)}^{\text{D}} \Bigg]$$

where $m' \equiv \frac{\overline{m}_2}{n} t'$, and $m^* \equiv \frac{\overline{m}_2}{n} t^*$ denote the conjectured number of recognitions of agent 1 from the perspective of agent 2 until time t' and t^*, respectively.

Algebraically, $U(\overline{m}_2'|\overline{m}_2)$ can be thought of as a convex combination of C' and D', discounted by A', and $U(\overline{m}_2|\overline{m}_2)$ a convex combination of C and D, discounted by A. Notice, as from section 3, that $B' = B = \frac{\overline{m}_2}{n}$, which means that the weights of convex combinations are the same. Therefore, the proof of theorem amounts to showing the following two properties:

$$3)\ A' \cdot C' > A \cdot C \quad \text{and} \quad 4)\ A' \cdot D' > A \cdot D$$

Furthermore, I claim that showing 4) automatically satisfies both statements. The reasoning is as follows. As from proof of Corollary 3.1.1,

$$\begin{aligned} V_t^2(m) &= P_t^2(m)\big(1 - \delta V_{t+1}^1(m)\big) + \big(1 - P_t^2(m)\big)\delta V_{t+1}^2(m+1) \\ &= P_t^2(m)\big(1 - \delta S_{t+1}\big) + \delta \mathbb{E}_t\big[V_{t+1}^2 | P_t^2(m)\big] \end{aligned}$$

Notice that t^* and t' are determined so that $1 - \delta S_{t^*+1} \approx 1 - \delta S_{t'+1} \approx 0$—the perceived social surplus first crosses $\frac{1}{\delta}$. Therefore, if $A' \cdot D' > A \cdot D$, and the respective proposer rents are such that $1 - \delta S_{t'+1} \approx C' - D' \approx C - D \approx 1 - \delta S_{t^*+1} \approx 0$, it implies that $A' \cdot C' > A \cdot C$ as well.

This algebraic exercise also invites fitting interpretations. $A' \cdot D' > A \cdot D$ implies that after having decided the strategies based on the common knowledge priors, agent $2'$ is better off than agent 2—in utility comparison sense, not contemplating optimism—if agent 1 is recognized at the time of settlement for both games. With this interpretation, it follows naturally that under the same setting, agent $2'$ should also be better off than agent 2 even if they get recognized at the time of settlement.

From Lemma 1.3, we know that continuation values in common knowledge games surmount to who gets to take the present value of all future proposer rents, Λ_t. From positing that proposer rents come from each agent's optimistic belief about recognition probability leads to the approximation in Lemma 1.4. Utilizing the lemmas lead to the following:

$$
\begin{aligned}
A' \cdot D' = \delta^{t'} \cdot \delta V^{2'}_{t'+1}(m'+1) &= \delta^{t'} P^{2'}_{t'}(m')\Lambda'_{t'} \\
&> \delta^{t'} \cdot \frac{n+t'-\overline{m}'_2-m'}{n+t'+1} \cdot \frac{1}{1+\delta y'_{t'+1}} \qquad (1.5)\\
&= \delta^{t'} \cdot \frac{n+t'-\overline{m}'_2-m'}{n+t'+1} \cdot \frac{1}{1+\delta\frac{\Delta'}{t'+n+1}} \qquad (1.6)\\
&= \delta^{t'} \cdot \frac{n+t'-\overline{m}'_2-m'}{n+t'+1} \cdot \frac{t'+n+1}{t'+n+1+\delta\Delta'} \\
&= \delta^{t'} \cdot \frac{n+t'-\overline{m}'_2-m'}{n+t'+1+\delta\Delta'} \\
&= \delta^{t'} \cdot \frac{n+t'+\Delta'-\overline{m}_1-m'}{n+t'+\delta\Delta'+1} \qquad (1.7)
\end{aligned}
$$

where (1.5) is given by the lower bound of $\Lambda'_{t'}$ shown in Lemma 1.4, (1.6) utilizes the definition of optimism $y'_{t'+1}$, and (1.7) orders the variables nicer using the definition of Δ'. Similar procedure leads to the following upper bound:

$$
A \cdot D = \delta^{t^*} \cdot \delta V^{2}_{t^*+1}(m^*+1) < \ \delta^* \cdot \frac{n+t^*+\Delta-\overline{m}_1-m^*+1}{n+t^*+\delta\Delta+2} \qquad (1.8)
$$

Therefore, showing (1.7) > (1.8) will finalize the proof that $A' \cdot D' > A \cdot D$. In the endeavor, I utilize Lemma 1.5 to clarify equations. In particular, I multiply $\delta^{-t'}$ to equation

(1.7) and multiply $\delta^{n-(n+t')} = \delta^{-t'}$ to (1.8). Doing so gets rid of $\delta^{t'}$ term from (1.7) and transforms δ^{t^*} into $\frac{\delta^{t^*+n}}{\delta^{t'+n}} \approx \delta^{\sqrt{\frac{\Delta\delta}{1-\delta}}-\sqrt{\frac{\Delta'\delta}{1-\delta}}} = \delta^{\sqrt{\frac{\delta}{1-\delta}}(\sqrt{\Delta}-\sqrt{\Delta'})}$. Formally,

$$
\begin{aligned}
&(1.7)\cdot\delta^{-t'} > (1.8)\cdot\delta^{n-(n+t')} \\
\Longrightarrow\ & \frac{n+t^*+\delta\Delta+2}{n+t'+\delta\Delta'+1}\cdot\frac{n+t'+\Delta'-\overline{m}_1-m'}{n+t^*+\Delta-\overline{m}_1-m^*+1} > \delta^{\sqrt{\frac{\delta}{1-\delta}}\left(\sqrt{\Delta}-\sqrt{\Delta'}\right)} && (1.9)\\
\Longrightarrow\ & \frac{\sqrt{\frac{\delta\Delta}{1-\delta}}+\delta\Delta+2}{\sqrt{\frac{\delta\Delta'}{1-\delta}}+\delta\Delta'+1}\cdot\frac{\frac{n-\overline{m}_2}{n}\sqrt{\frac{\delta\Delta'}{1-\delta}}+\Delta'-\overline{m}_1+\overline{m}_2}{\frac{n-\overline{m}_2}{n}\sqrt{\frac{\delta\Delta}{1-\delta}}+\Delta-\overline{m}_1+\overline{m}_2+1} > \delta^{\sqrt{\frac{\delta}{1-\delta}}\left(\sqrt{\Delta}-\sqrt{\Delta'}\right)} && (1.10)\\
\Longrightarrow\ & \frac{\sqrt{\frac{\delta\Delta}{1-\delta}}+\delta\Delta+1}{\sqrt{\frac{\delta\Delta'}{1-\delta}}+\delta\Delta'}\cdot\frac{\frac{n-\overline{m}_2}{n}\sqrt{\frac{\delta\Delta'}{1-\delta}}+\Delta'-\overline{m}_1+\overline{m}_2}{\frac{n-\overline{m}_2}{n}\sqrt{\frac{\delta\Delta}{1-\delta}}+\Delta-\overline{m}_1+\overline{m}_2+1} > \delta^{\sqrt{\frac{\delta}{1-\delta}}\left(\sqrt{\Delta}-\sqrt{\Delta'}\right)} && (1.11)
\end{aligned}
$$

where from (1.9) to (1.10), I use the approximation of t^* and t' as in Lemma 1.5, and also the definition of m^* and m' to group with t^* and t'. Utilizing the fact that the first fraction is less than 1, I add a small simplification by decreasing both the numerator and denominator by 1 in (1.11). In effect, expression (1.11) substitutes all the variables in terms of distance in prior beliefs Δ and Δ'. The variation on the right hand side is driven by the root difference in the distances in prior beliefs between agent 2 and agent 2$'$. As this difference $\sqrt{\Delta}-\sqrt{\Delta'}$ increases, the relationship between settlement time and Δ forces the range of allowed δ, shrinking it to a small range very close to 1, containing the value of the expression in right hand side. In the meantime, the difference in the distances in prior beliefs affect the left hand side either by multiplication or in the first order difference. This effect is shown in simulation to always dominate the lower degree changes on the right hand side. Although I cannot simplify the algebra to clear-cut comparison, the generous bounds I use, as well as the strict relationship between settlement time t, discount factor δ, and distance in prior beliefs Δ, allow an ease and precision in generating unambiguous difference in simulation.

Intuitively, this result stems from the fact that both agent 2 and 2$'$ choose to delay the settlement against agent 1 in the first place. Such preferences only allow very high patience level δ and low firmness in belief n, compared to the distance in belief Δ. Therefore, as intuitive proof in section 3 delineates, agent 2 is always willing to trade vastly favorable split of the pie—driven by low firmness in belief and high Δ—for an additional damage from δ.

CHAPTER 2

Windows of Opportunity to Build Trust While Bargaining[1]

2.1 Introduction

In this chapter, we analyze how trust is built over time by rational agents. We focus on the context of international relations, but I believe that the analysis could be applied to other domains in which actors develop trust by making themselves vulnerable.

To represent the strategic problems that political actors often face, we model trust building in the context of bargaining. This is important because, in a sense, these processes are opposites. While trust building involves signaling a willingness to cooperate, bargaining involves signaling a willingness not to do so in order to achieve a favorable bargain. We study social encounters that involve both. This allows us to show that trust building is much more difficult than past literature implies.[2] When states are bargaining and trust building at the same time, fully informative signaling becomes impossible. Still more importantly, bargaining changes the types of trust building that lead to the highest expected welfare.[3]

To understand how bargaining influences welfare, we must understand the factors that influence the optimal choice of stakes over the course of the period of trust building and bargaining. This is a long-standing issue in a variety of literatures, from negotiation theory

[1]Co-authored with Robert Trager, UCLA

[2]Cf. Kydd [2000, 2001, 2007, 1997b,a].

[3]The canonical statement of the bargaining model approach to international relations is Fearon [1995].

to dispute mediation. Analysts have often wondered whether adversaries should start with reciprocal actions on lower level issues to build trust gradually or should immediately address the most important issues in "grand gestures."[4] History provides models for both. The Oslo Peace Process between the Israelis and Palestinians, which left the status of Jerusalem undetermined, is an example of putting off high stakes questions until lesser issues are dealt with. The Egyptian-Israeli peace process, which began with Egyptian President Anwar Sadat's dramatic visit to the Israeli Knesset in 1977 and culminated in the Peace Treaty in 1979, is an example of an initial agreement on with the most significant issues, like the status of the Sinai Peninsula and recognition of the state of Israel.

We show that the larger the benefits to cooperation and the larger the gains from unilateral defection while the other player cooperates, the more the optimal signaling stakes tend to start high and end low. In other words, when the benefits to cooperation and the gains from defection are high, grand gestures lead to higher expected welfare than gradual building of trust. This is one sense in which there are windows of opportunity to build trust: if the opportunity to make grand gestures in these contexts is not seized, the players are worse off on average.

Bargaining often influences the optimal stakes in a similar way. The more important bargaining is relative to trust building, the more actors have incentive to put off addressing high stakes matters early on in peace processes. This may explain the different choices and outcomes in the Israeli-Egyptian and Israeli-Palestinian peace processes. In the case of Egypt, difficulties centered more on each side convincing the other that it *wanted* to cooperate. In the case of the Palestinian peace process, there is more to bargain over. Thus, negotiators opted for high stakes issues first in the first case and down the road in the second.

But bargaining also has surprising effects on welfare. When trust is high, the more actors' need to bargain, the lower welfare. When trust is low, the opposite often occurs: welfare

[4] Berenji [2020].

tends to increase in the importance of bargaining. This is because the bargaining dynamic allows the players to signal trustworthiness using lower stakes. The risk of betrayal that the players must take on to signal trustworthiness is lower because players are already taking on another cost when they attempt to signal trustworthiness. The other cost they take on is the risk of putting themselves in a worse bargaining position. The players can demonstrate their trustworthiness by taking on this bargaining risk *in lieu* of taking on higher stakes issues before they've built up trust. This is how bargaining increases welfare when trust is low.

Thus, even though trust building and bargaining are conceptual opposites, they are not opposites when it comes to welfare implications. Bargaining can have a positive effect on welfare when initial trust is low. Yet, bargaining also restricts the range of cases over which it is possible to build trust at all.

So far, we have discussed the effect of issue sequencing on welfare. But sometimes the significant tests of trust are not chosen, but arrive when they arrive. An alliance is tested when the war arrives, for instance. There is no way to place this test ahead of lesser tests as a grand gesture. The lesser tests - a period in which trust is built or not - must come first. When we analyze such contexts, we find that it is usually better to build trust gradually. In particular, when there is a highly consequential moment on the way when trust will be tested, and the consequences of misplaced trust are high, it is always welfare improving to build trust gradually in advance. In such instances, if a chance to build trust is missed, the actors can still attempt to build trust later, and depending on context, this may or may not be successful. But welfare will always be lost. This is the second sense in which there are windows of opportunity to build trust.

The analysis below employs the framework for thinking about trust developed by Kydd [2007]. Nevertheless, we come to somewhat different conclusions. Kydd argued that trust can always be built quickly. We argue that even when trust can be built quickly, usually, it should not be. Doing so leads to lower welfare. Further, when bargaining is present, not only can trust not be built quickly, sometimes it cannot be built at all. Overall, our analysis explains why trust building is so much more difficult than the current literature implies and illuminates the opportunities that produce the best outcomes between adversaries with something to lose.

2.1.1 Literature Review

Trust has been defined and studied in a wide variety of ways across social science disciplines. Nevertheless, Ruzicka and Keating [2015] identify three primary scholarly approaches.[5] In the psychological approach, trust is a result of the psychological predispositions and emotions of actors. Though psychological approaches have had a long history in international relations theory [Jervis, 2017], much of the scholarship was produced after the development of prospect theory and behavioral economics by Kahneman and Tversky [2012]. In this line of research, Larson [1997] examines the relationship between the United States and Soviet Union during the Cold War, arguing that the psychological dispositions of the states' leadership hindered the development of trust even on issues on which the states had convergent preferences. Hall and Yarhi-Milo [2012b] concur, arguing that leaders use both costly signals and personal rapport with other leaders when deciding on the trustworthiness of another actor.[6] Yarhi-Milo et al. [2018] conduct an experiment that shows highly trusting leaders view both public signals and material action as more informative than low-trust types. While actors may still act strategically, their preferences are formed from their psychological disposition [Rathbun, 2009], which may explain dyadic rivalries [Goertz and Diehl, 1993], the democratic peace [Maoz and Russett, 1993], and resolutions to collective action problems [Mercer, 2005].

In the social constructivist approach,[7] trust develops through repeated interaction in

---

[5] See Cho et al. [2015] for a robust interdisciplinary literature review on trust.

[6] Rathbun [2012, 2009] distinguishes between the psychological notion of moral trust, which is more constant over time, and strategic trust, which is situation-dependent and more akin to the rational choice approach below.

[7] Trust is sometimes viewed as a form of social capital [Fukuyama, 1995]. In this vein of literature, scholars sometimes define trust as an actor's confidence that other actors will "do what is right" [Hoffman, 2002,

shared communities [Mitzen, 2006, Wendt, 1999, March and Olsen, 1998]. Adler and Barnett [1998], for instance, view trust as a necessary precondition for the development of a common identity among states. Similarly, [Booth and Wheeler, 2007] and [Wheeler, 2010] study the trust or distrust that develops between individual leaders. In this vein, Cronin [1999] notes that Italian principalities were able to overcome histories of distrust when they began to reconceptualize themselves as Italians instead of, for example, Tuscans or Parmans. This trusting community was then able to overcome collective action problems. Constructivists recognize the importance of an awareness of cultural norms and differences when engaging in negotiations [Buitrago, 2009].

Finally, the rationalist approach views trust as an endogenous property of situations given uncertainty about actors' intentions [Kydd, 2005]. These scholars emphasize Bayesian updating as the mechanism through which beliefs evolve. Axelrod and Keohane [1985] were some of the first scholars to view trust as tied to rational choice, arguing that actors' preferences are based on both the situation and their personal interpretation of events. Thus, mistrust is more likely to arise in bargaining over international security, in which one side can be eliminated completely, than in issues of economics, in which arrangements can continue indefinitely. Kydd [2007] defines trust as "a belief that the other side is trustworthy, that is, willing to reciprocate cooperation." Our model is in this tradition, taking Kydd's formalization of trust building processes as its starting point. Through interactions, states receive signals and update their beliefs about other's trustworthiness in Bayesian fashion.[8] Decision theory offers a strong foundation for learning from the world in this way, but this is not to say social actors always do update their levels of trust for others in this fashion [Hall and Yarhi-Milo, 2012a].[9]

2006].

[8]A related line of research focuses on enduring rivalries, dyads of states that come to mistrust one another through repeated interaction, often leading to the arms buildup and war as in traditional security dilemma models [Goertz and Diehl, 1995, Klein et al., 2006].

[9]Another strand of literature focuses on cheap talk processes as a source of trust. See Fudenberg and Tirole

In rationalist accounts, trust building and bargaining are at odds. If a security-seeking state could simply *ex ante* trust that another state was security seeking, there would be no need for further negotiation and arms buildups. Kydd [2007] notes that, during crisis bargaining, states send signals of resolve; in contrast, when trying to assure others of their trustworthiness, states send signals that they are open to cooperation. Kydd notes that these are often relatively small stakes that nonetheless "expose one to some risk of defection on [their opponent's part]." [10]

A number of models of international relations incorporate trust building. One example is the security dilemma concept of Defensive Realism. In this model, states are either trusting or fearful, which determines how aggressively they respond to the arms buildup of other states [Jervis, 2017, Kydd, 1997b]. Kydd [2010] models trust as an exogenous process in climate treaty negotiations. Initially, states are unsure whether to trust others to sign. Over time, as information about the effectiveness of the treaty increases, states learn that others are either trustworthy types who were waiting for research into the effectiveness of the policy or untrustworthy types who will not under any circumstance cooperate. A predecessor to this strand of literature is Axelrod [1984], who found that repeated Prisoner's Dilemma games can produce cooperation through an extended "shadow of the future."

Kydd [2000, 2007] first formalized the development of trust in his Assurance Game.[11] Assurance in the international relations literature can be traced back to Osgood [1962], who argued that the mistrust driving the Cold War arms race can be mitigated through a

---

[1991] in support of this idea and Bracht and Feltovich [2009] against it. Bornstein and Gilula [2003] show via lab experiments that cheap talk can improve cooperation in assurance games. Schweitzer et al. [2006] provide a potential solution to this dilemma: cheap talk can increase trust only if a player has not behaved deceptively in prior interactions. Other scholars show empirically that inter-state networks give states an opportunity to develop beliefs about the trustworthiness of other states [Maoz, 2010, Ostrom and Ahn, 2009].

[10] Solhaug et al. [2007] note that trust is inversely related to risk and possibly directly related to the size of the stakes. Though Hoffman [2002] defines trust as more than just based on risk, arguing that common usage of the phrase generally implies uprightness and not just "a good bet." Malhotra and Murnighan [2002] show that binding contracts lead to players to become more trusting at a slower rate than non-binding contracts.

[11] Kydd [2001] applies the assurance game to NATO enlargement in Eastern Europe, showing that limited expansion of NATO can be reassuring to Russia when trust is relatively high.

strategy he called GRIT (Graduated Reciprocation in Tension-reduction). In this strategy, each state would repeatedly unilaterally offer increasingly signals of cooperation in order to reduce mistrust.[12] In the model of Kydd [2007], states first choose whether to cooperate in a low stakes encounter. If they do, this signals their willingness to cooperate in a subsequent higher stakes encounter. Formally, three types of equilibria result: non-cooperative, in which no state makes an offer in the first round; reassurance, in which only trustworthy types offer the signal; and time-will-tell, in which both states signal and then untrustworthy types defect in the second round. The costliness of signals varies inversely with level of trust, power of the sender, and the strength of first-mover advantage and varies directly with the costs of conflict. Kydd demonstrates that at any level of initial trust, security seekers may find an assurance equilibrium.[13]

This feature of Kydd's assurance model is important. It implies that when trust is low, it is easy to build. When trust is low, they can have a low stakes encounter and if the sides cooperate there, they will have no trouble cooperating in a subsequent high stakes encounter. The lower the initial level of trust, the lower the signaling stakes must be to signal trustworthiness. It is as if the U.S. and North Korea could cooperate on a relatively low stakes issue, like violent incidents near the border with South Korea, and then trust each other enough to enter high stakes cooperation on nuclear arsenals and opening to the world. Why does the model have this implication? It is due to the fact that even the low stakes international is high risk given the low level of trust. Taking such a great risk in cooperating signals a strong willingness to cooperate. Thus, signaling trustworthiness is always available - at moment's notice, or at least relatively quickly - if trustworthy types will but seize the opportunity.

This dynamic has a logic to it, but as a description of the world, it is unconvincing. If the

[12] See also Etzioni [1962], Schelling [1980], Stein [1991] for other foundational analyses of assurance.

[13] Kydd [2006] studies how mediators in bargaining processes can build trust, arguing that moderately biased mediators with single-peaked preferences are more trustworthy than mediators indifferent over outcomes.

U.S. and North Korea managed to cooperate on lower stakes issues - as they have for periods in the past - experience demonstrates that their levels of trust might still be low. They might still see the need to build trust gradually, though Kydd's model offers no explanation for why that would be so.

We show below that a key reason for this apparent discrepancy between the model and the political world is the need of adversaries to bargain while they build trust. This implies that trust building when initial trust is low may be impossible. It is no surprise then that building trust between the U.S. and North Korea has proven so difficult. Further, even when trust *can* be signaled quickly in one step, we show that it is usually welfare improving to signal trust gradually in a series of steps.

The trust-building models in international relations also generally assume a given bargaining order. The Assurance Game in Kydd [2007] assumes that smaller stakes are offered before larger ones. Sometimes, however, states can choose between equilibria in which the highest stakes are addressed early or later on in peace processes and other negotiations. Osgood [1962], for example, argued that the existing bargaining process between the US and USSR was leading to an arms race and suggested unilaterally offering smaller stakes instead to build trust. The issue remains relevant in contemporary negotiations. In the bargaining process between Israel and Palestine, should the states reach an agreement on settlements in the West Bank first or start with the final status of Jerusalem? The paper in international relations that comes closest to exploring this question is Fershtman [2000]. In this model, two sides (a and b) are deciding on two issues. However, group b is composed of two players who have conflicting preferences over the issue prioritization. The players can bargain simultaneously, issue-by-issue, or simultaneously with one player in group b acting as the group's representative. Fershtman finds that a prefers any type of negotiation with a representative to issue-by-issue bargaining to simultaneous bargaining, and each player in b prefers issue-by-issue over simultaneous bargaining.

The role of issue ordering in bargaining is explored further in the economics and management literatures.[14] Scholars have identified that varying the number, timing, and ordering of issues can affect bargaining outcomes [Sebenius, 1983, Balakrishnan et al., 1993, Geiger, 2017]. Sebenius [1983] argues that the probability of agreement depends on the number of linked issues, while Balakrishnan et al. [1993] argue that beginning with less important issues can increase the probability of a successful outcomes, though with an externally-imposed deadline, that conclusion can be reversed [Watkins, 1998].[15]

A number of scholars have sought to formalize players' preferences over agendas, varying their time preference, relative bargaining strength, and utility over the set of issues. In general, these models produce a wide variety of SPE in contrast to the single equilibrium in the single-issue game in Rubinstein [1982]. Early papers in this literature focused on how exogenous variation in bargaining structure affected SPE [Busch and Horstmann, 1997]. Inderst [2000] was the first to study fully endogenous agendas, finding that if all issues are mutually beneficial, they are settled in the first round. However, if some issues produce a net loss for one player, in equilibrium, some issues will be solved after a time delay. [16] In contrast, Fatima et al. [2002] models player time preferences, finding that if one player has an incentive to delay, agreement will be reached at the earliest deadline; otherwise, it will be reached immediately. In and Serrano [2004] study a two-issue game, finding that there always exist SPE that are decided immediately and that every bargaining process is finite. Busch and Horstmann [2002] study a two-issue game, finding that when agreements are implemented as they are reached, easy issues are negotiated before hard ones and that when agreements are implemented simultaneously, larger issues are bargaining over first. The first agenda Pareto dominates the second. Chatterjee [2005] finds that high discount rates can

[14]See Carraro et al. [2005] for a review.

[15]Thompson [1998] argues that sequential bargaining should be avoided, as simultaneous bargaining on all issues can allow for trade-offs among issues, widening the bargaining range. Shell [2006] argues that, in multi-issue bargaining, negotiators should begin with larger issues because of the norm of reciprocity.

[16]Lang and Rosenthal [2001] provide another early model showing that issue-by-issue bargaining can exist when the agenda is endogenized.

make sequential bargaining inefficient, leading players to prefer simultaneous negotiations. Finally, Banerji [2002] analyzes a game with one buyer and two sellers. He finds that when an issue is renegotiated, if the new agreement with at least one of the sellers must be strictly worse to a buyer than the preexisting agreement, the buyer will prefer sequential bargaining.[17]

Thus, there are many works that address issue sequencing in negotiations and they have a varied set of findings. These works give reasons to start bargaining processes with more or less consequential stakes, but none do so in the context of building trust. In general, they also do not examine the welfare implications of different equilibria. Those that do do not consider cases where actors benefit from signaling a willingness to cooperate at the same time as they benefit from signaling a willingness to demonstrate resolve in bargaining. Here, the international relations literature comes closest in that models of inter-state bargaining include the implicit benefit to cooperating of avoiding war.[18] Yet, in most of these models, there is no disincentive to signaling that one is the very highest resolved type, which means the type with the lowest costs conflict - that is, the type least willing to cooperate. There is therefore no benefit to signaling trustworthiness is most of these models. Such types are willing to accept concessions, if that is all cooperation requires, but they are not willing to *make concessions* when the other side does. The belief that the other will do so is the essence of trust. That is what we study below, alongside the more traditional bargaining incentives.

[17] A number of models analyze the relationship between the valuation of issues and issue ordering in the above context in which one buyer is negotiating with at least two sellers. If the buyer is able to achieve positive surplus by reaching prices below her valuation, she will negotiate with sellers with lower bargaining power first [Munster and Reisinger, 2018, Krasteva and Yildirim, 2012]. The conclusifon is reversed with negative surplus [Munster and Reisinger, 2018]. Likewise, buyers prefer to negotiate more important issues first, either in terms of valuation [Chatterjee, 2005] or anticipated surplus [Raskovich, 2007]. As in Geiger [2017], outside options weaken the bargaining power of the seller [Munster and Reisinger, 2018, Raskovich, 2007].

[18] For overviews of this literature, see Powell [2002] and Trager [2016].

2.2 Model

2.2.1 Primitives

We model trust building as a dynamic game of incomplete information between two players ($I = \{1, 2\}$) that comprises of possibly multiple stages. In each period k, both players engage in a simultaneous move game of choosing between cooperation (C) and defection (D).

The sole source of uncertainty in our model is each agent's cost to defection. We assume that each player draws cost c_i from a uniform distribution with support $[0, 1]$, and that this information is private information: c_i is known to herself but not to the opponent.[19] This payoff-relevant cost to defection also represents the agent's trustworthiness to the opponent, as agents with higher cost will tend to defect less. The game of trust dynamics revolves around the agents' incentive, or lack thereof, to communicate this private information through their actions.

Actions the agents take each period bear two consequences in payoff: a simultaneous game component that represents flow utility in the given period – henceforth "game component" – and a reputation component that represents how the signaled information from the chosen actions affect their bargaining leverage over subsequent periods – henceforth "bargaining component." We parameterize the relative importance of the bargaining payoff as μ, so that each period's stake α_k is divided into $\alpha_k \cdot \mu$ of bargaining payoff and $\alpha_k(1-\mu)$ of game payoff.

First, we summarize the game component of the payoffs, which is analogous to the setup in Kydd [2005], in Table 2.1. Each cell represents flow payoffs to the players with the corresponding action. The payoff matrix is identical for all periods, except that they are weighted by the relative importance of each period k, denoted α_k, with the last stage stake

[19] In his scholarship on trust, Kydd generally focuses on discrete types. In Kydd [2010], however, he allows for a continuous type space. In this model, trust is a function of (exogenous) increases in information. As information about a climate policy is made available, states are revealed to have types that never cooperate, initially do not cooperate until the policy is proven beneficial, initially do cooperate until the policy is proven ineffective, and never cooperate. Kydd [2003] models states with continuous costs for conflict to analyze the usefulness of a third-party mediator.

normalized to 1. When both players cooperate, they receive a cooperation benefit of $b \geqslant 0$. We do not normalize this benefit to 0 as in Kydd [2005], because we allow for more than two periods, and wish to capture the trade-offs between cooperating in one period versus another. When both players defect, they engage in a conflict over which each player i has a p_i chance of winning. Normalizing the utility from winning to 1 and the utility from losing to 0 yields the standard expected utility of $p_i - c_i$.[20] Finally, when only one of the players defects, the sole defector gains the first-mover advantage ϕ on top of the usual payoff from conflict, at the expense of the cooperating player who is caught off guard.

Table 2.1: The Game Component Payoff Matrix

	C	D
C	b, b	$p_1 - \phi - c_1, p_2 + \phi - c_2$
D	$p_1 + \phi - c_1, p_2 - \phi - c_2$	$p_1 - c_1, p_2 - c_2$

One can note from this stage game payoff that when $p_i + \phi - c_i \geqslant b$, it is a dominant strategy for player i to defect regardless of the opponent's action. On the other hand, if $p_i + \phi - c_i < b$ for both players, it would be mutually beneficial for them to communicate their private information and cooperate, as cooperation is the best response to the opponent's cooperation in this case. Therefore, denoting the threshold cost as $\bar{c}_i (\equiv p_i + \phi - b)$, we can analyze whether players with $c_i > \bar{c}_i$ can successfully communicate their trustworthiness and achieve mutual cooperation.[21]

The game component of the payoffs highlights the incentive of the high cost types to successfully communicate their trustworthiness while being wary of possibly facing the low cost type. In practice, however, broadcasting that one has high costs to a non-cooperative outcome worsens her bargaining leverage over the range of cooperative outcomes. This is

[20] $p_i + p_{-i} = 1$.

[21] In the symmetric version of $p_1 = p_2 = \frac{1}{2}$ that we often work with later, $\bar{c} \equiv \frac{1}{2} + \phi - b$ serves as a parameter that summarizes the level of trust.

so even against a cooperative partner. We capture this in a simple way in the bargaining component of the model.

We model this bargaining aspect of interactions in a simple way that is consistent with many specific bargaining protocols. In each period, there is a bargaining surplus $\mu \geqslant 0$. When players have signaled that they have high costs to a non-cooperative outcome, we assume that they receive a lower share of this surplus. For simplicity, we assume in particular that the bargaining surplus is divided in inverse proportion to the players' expected costs following their actions in a period given their equilibrium strategies.

In a threshold equilibrium, for instance, as we will soon show, in each period k, a player with cost lower than a threshold t_k defects while a player with a higher cost cooperates.[22] Therefore, when a player defects, she signals that her cost type is lower than t_k ($\frac{t_k+t_{k-1}}{2}$ in expectation) and vice versa for the cooperating player ($\frac{t_k+t_{k+1}}{2}$ in expectation). For example, consider a single period game in which costs are distributed $uniform([0,1])$ and an equilibrium in which a player with costs below $\frac{1}{2}$ defects while one with a higher cost cooperates. A defecting player signals an expected cost type of $\frac{1}{4}$ while a cooperating player signals $\frac{3}{4}$. The bargaining surplus in the second period is therefore divided $\frac{3}{4}\mu$ to the defecting player and the remaining $\frac{1}{4}\mu$ to the cooperating player.

This is consistent with many bargaining protocols. For instance, we could equally specify that the players play a Nash demand game over the bargaining surplus in each period following the trust game. This would be equivalent to the assumption we make here in the demand game equilibrium in which the players propose the equivalent division of the bargaining surplus.[23] More importantly, the bargaining outcome we assume is consistent with analyses of the many bargaining environments in which expected share of the bargaining

[22] We also allow t_k to be 0 or 1, in which case a player cooperates or defects, respectively, regardless of their private cost type.

[23] Note that one simplification we make here is the implicit assumption that the bargaining portion of the game conveys no additional information beyond what has already been conveyed in the trust building portion of the game.

surplus increases in players' value for the no-agreement outcome.[24]

Table 2.2 below summarizes the key notation of the model.

Table 2.2: Notations For Game of Trust Building While Bargaining

μ	Relative weight on bargaining payoff
α_k	kth period stake
c_i	Player i's cost from conflict, drawn from $U(0,1)$
b	Cooperation benefit
p_i	Likelihood that player i wins in conflict
ϕ	Fixed and symmetric first-mover advantage
$\bar{c}_i$	Stage game threshold cost for cooperation

2.2.2 Trust Building Equilibria

How do players build trust under this model of trust dynamics? Cooperative players with high private costs would want the opponent to cooperate as well in order to avoid getting into costly conflict. The obstacle is that the defective players with low costs would also want the opponent to cooperate, to extract the first mover advantage ϕ. In a prolonged negotiation that we represent as a multi-period game, can the cooperative players utilize earlier stages to elicit mutual cooperation? Upon observing early cooperation from the opponent, can players successfully weed out defective opponents who wish to deceive them into paying additional ϕ? Before attempting to directly answer these questions, it helps to analyze a simpler version of the game as a benchmark, that we call a one shot game of cooperation, with only one period of game component as the payoff.

Consider two players playing the game of choosing between cooperation and defection where the payoffs are given in Table 2.1 above. In this simultaneous move game without

[24] Fey and Ramsay [2011], Banks [1990].

the bargaining component, it is easy to check that everyone defecting regardless of their private cost constitutes a Nash Equilibrium. In fact, we already established in the model introduction that when the private cost c_i is less than $\bar{c}_i$, it is a dominant strategy for player i to defect regardless of the opponent's cooperation. Is there, then, a possibility of achieving a separating equilibrium that induces some cooperation?

Noting that higher cost to defection can only increase the incentive to cooperate, we can establish that the separating equilibrium would consist of a threshold strategy whereby players above a certain threshold t cooperates while the rest defects. After restricting attention to the separating equilibria in threshold strategies, one can easily check that such t solves the following equality condition, and always exists between $\bar{c}$ and 1 as long as $\bar{c} \leqslant (\sqrt{\phi} - 1)^2$.

$$\underbrace{(1-t)(\bar{c}-t)}_{\substack{\text{offensive cost/gain in switching from C to D} \\ \text{against the cooperator}}} \qquad + \qquad \underbrace{t\phi}_{\substack{\text{defensive gain in switching from C to D} \\ \text{against the defectors}}} \qquad = 0$$

With this separating equilibrium in the benchmark game in mind, we can modify the original analyses to the following sets of questions. When we extend the one shot game of cooperation to multiple periods, can we still achieve separating equilibria with equal or larger cooperation range? What does trust building look like? Do all signaling happen in the first stage and players play the game of certain actions from the second period onward? Or do the eventual defectors gradually reveal their types over the course in the order of their private costs?

To formalize ideas, we define the parameters that describe the K-period game and an action sequence of players within the game as follows:

Definition 2.1. A K-period game is summarized by the parameter set $(\bar{c}, \phi, \mu, \boldsymbol{\alpha})$, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_K(\equiv 1))$. In each period k, player i takes action $a_{i,k} \in \{C, D\}$ given the parameters and previous actions of the two players. Such actions over K periods yield an action sequence $\boldsymbol{a_i} = (a_{i,1}, a_{i,2}, ..., a_{i,K})$.

Similar to how we restricted our attention to the equilibria of threshold strategies in the benchmark model, in our main model, we will focus on strategies in which no player starts with defection followed by cooperation. As we established earlier, in our model, it only helps a player if the opponent trusts her early and cooperates, regardless of her private cost. Therefore, our restriction to strategies that preclude signaling defection only to cooperate later is an intuitive one.

Assumption 2.1 (No Perverse Signaling)*.* We assume away any perverse signaling where $a_{i,m} = D$ and $a_{i,n} = C$ for any m, n such that $m < n$.

One could also interpret this assumption as a property of our model that once the players miscoordinate, they can never restore trust and obtain mutual cooperation later. With the assumption of no perverse signaling and the definition of an action sequence, we are now ready to introduce strategies and the different forms of trust building equilibria. First, in parallel to the reassurance equilibrium of Kydd[2005], we introduce a separating equilibrium in which the defecting players reveal themselves by defecting from the first stage of the game, and the cooperating players in the first stage cooperate throughout the entire game.

Definition 2.2 (Immediate Revelation Equilibrium)**.** In the K period symmetric game of trust building with $(\bar{c}, \phi, \mu, \boldsymbol{\alpha})$, the **Immediate Revelation Equilibrium** $(\boldsymbol{s(c)}, \boldsymbol{b})$ is a strategy profile paired with consistent belief system as follows:[25]

$$s_i(c_i) = \begin{cases} (C, (C \mid_{a_{-i,1}=C}, D \mid_{a_{-i,1}=D}), ..., (C \mid_{a_{-i,K-1}=C}, D \mid_{a_{-i,K-1}=D})) & \text{if } c_i > t \\ (D, ..., D) & \text{if } c_i \leqslant t \end{cases}$$

$$b_i = \begin{cases} \mathbb{P}(c_{-i} > t \mid_{a_{-i,l}=C}) = 1 \\ \mathbb{P}(c_{-i} \leqslant t \mid_{a_{-i,l}=D}) = 1 & \quad \forall l \leqslant K \\ \mathbb{P}(\text{anything else}) = 0 \end{cases}$$

[25]The sequential rationality conditions that the threshold t needs to satisfy are detailed in the appendix.

The Immediate Revelation Equilibrium in which the eventual defectors reveal themselves in the earliest stage is reassuring in the sense that from the second period onward, players of all cost types proceed with complete knowledge of the ensuing actions. This particular equilibrium corresponds well with the reassurance equilibrium of Kydd[2005]. Kydd's reassurance equilibrium also constitutes a separating equilibrium with binary types, with trust level that is exogenously given as a parameter. Our Immediate Revelation Equilibrium can be thought of as its generalization which not only deals with continuum of types, but also endogenizes the trust level a player holds against the opponent.

Players may also build trust gradually. Since players across all cost types prefer that the opponent cooperates, some proportion (with high costs, as will soon be shown) of the eventual defectors may start with cooperation. Knowing this incentive, cooperative players do not update their belief about the opponent's trustworthiness as dramatically as in the Immediate Revelation Equilibrium. We call such equilibrium a Gradual Revelation Equilibrium, as is formalized below.

Definition 2.3 (Gradual Revelation Equilibrium)**.** In the K period symmetric game of trust building with $(\bar{c}, \phi, \mu, \boldsymbol{\alpha})$, the **Gradual Revelation Equilibrium** $(\boldsymbol{s(c)}, \boldsymbol{b})$ is a strategy profile paired with consistent belief system as follows:

$$s_i(c_i) = \begin{cases} (C, ..., (C \mid_{a_{-i,K-1}=C}, D \mid_{a_{-i,K-1}=D})) & \text{if } c_i > t_K \\ (C, ..., (C \mid_{a_{-i,K-2}=C}, D \mid_{a_{-i,K-2}=D}), D) & \text{if } t_{K-1} < c_i \leqslant t_K \\ (C, ..., (C \mid_{a_{-i,K-3}=C}, D \mid_{a_{-i,K-3}=D}), D, D) & \text{if } t_{K-2} < c_i \leqslant t_{K-1} \\ ... & ... \\ (D, ..., D) & \text{if } c_i \leqslant t_1 \end{cases}$$

$$b_i = \begin{cases} \mathbb{P}(c_{-i} > t_l \mid_{a_{-i,l}=C}) = 1 \\ \mathbb{P}(t_{l-1} < c_{-i} \leqslant t_l \mid_{a_{-i,l-1}=C \ \& \ a_{-i,l}=D}) = 1 \\ \mathbb{P}(c_{-i} \leqslant t \mid_{a_{-i,l}=D}) = 1 \\ \mathbb{P}(\text{anything else}) = 0 \end{cases} \qquad \forall l \leqslant K$$

The composition of which cost types start their defection in each period depends not only on the parameters but also on the stakes of the game α_k. The more back-loaded the stakes are, the longer the defectors would postpone deviation not only to capture the bigger ϕ but also because they risk less in earlier periods from the probability of losing ϕ by being defected upon. In later sections, we analyze which stakes evolution is conducive of which types of Gradual Revelation Equilibrium, and the welfare implications of various issue sequencing.

The final variant of the equilibrium we introduce is similar to the Gradual Revelation Equilibrium except that no player of any cost types reveal themselves in the first period. We introduce this equilibrium more for a technical reason that when the trust level is too high, the Gradual Revelation Equilibrium may fail to exist because the belief system cannot sustain defection in the first period.

Definition 2.4 (Cheating (Time-will-tell) Equilibrium)**.** In the K period symmetric game of trust building with $(\bar{c}, \phi, \mu, \boldsymbol{\alpha})$, the **Cheating Equilibrium** $(\boldsymbol{s(c)}, \boldsymbol{b})$ is a strategy profile paired with consistent belief system as follows:

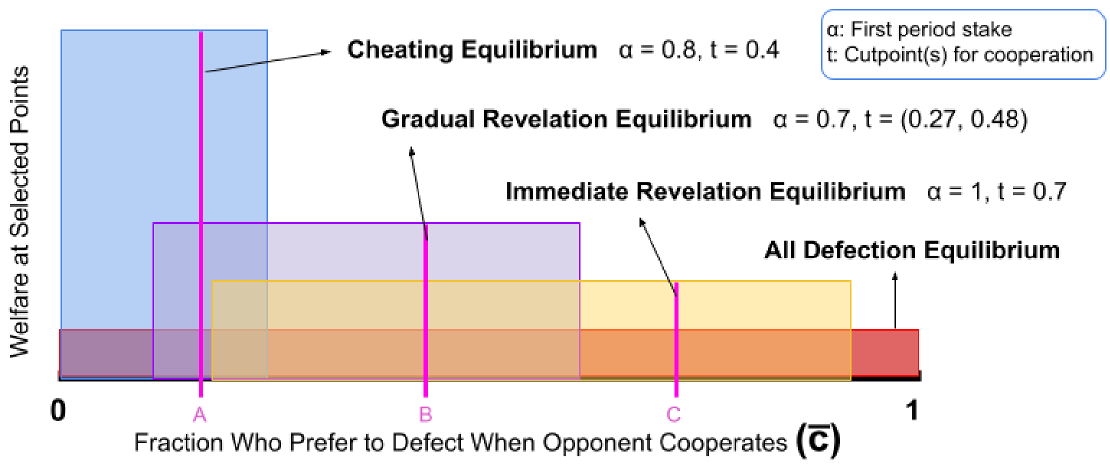
$$s_i(c_i) = \begin{cases} (C, ..., (C \mid_{a_{-i,K-1}=C}, D \mid_{a_{-i,K-1}=D})) & \text{if } c_i > t_K \\ (C, ..., (C \mid_{a_{-i,K-2}=C}, D \mid_{a_{-i,K-2}=D}), D) & \text{if } t_K \geqslant c_i > t_{K-1} \\ ... & ... \\ (C, C, D, ..., D) & \text{if } t_2 < c_i \leqslant t_3{}^{26} \\ (C, D, ..., D) & \text{if } c_i \leqslant t_2 \end{cases}$$

[26] In the 2-period case ($K = 2$), t_3 becomes 1, and there are only two strategies (C, D) and (C, C)

$$b_i = \begin{cases} \mathbb{P}(c_{-i} > t_l \mid {}_{a_{-i,l}=C}) = 1 \\ \mathbb{P}(t_{l-1} < c_{-i} \leqslant t_l \mid {}_{a_{-i,l-1}=C \ \& \ a_{-i,l}=D}) = 1 \\ \mathbb{P}(c_{-i} \leqslant t_2 \mid {}_{a_2=D}) = 1 \\ \mathbb{P}(\text{anything else}) = 0 \end{cases} \qquad \forall l \leqslant K$$

In all equilibria introduced thus far, the proportion of types that play a certain strategy determines where the thresholds get drawn, which in turn determine the proportion of each action types. Given the endogenous nature, there may exist more variants of equilibria especially for games of longer periods. As we restricted attention to the equilibria with no perverse signaling, we focus on the Immediate Revelation Equilibrium and the Gradual Revelation Equilibrium in the analysis as they bear the most pertinent interpretation of building trust in our model. We do, however, have the following completeness result for the two period game, which is often the entire focus of analysis in the literature, including Kydd[2005].

Observation. *In the* 2 *period game of trust building without perverse signaling, the Cheating Equilibrium, Gradual Revelation Equilibrium, Immediate Revelation Equilibrium, and the All Defection Equilibrium exhaust the set of possible equilibria, and their domain of existence is in increasing order in* $\bar{c}$.



In the 2 *period game of trust, the three types of equilibria defined above–the Cheating, the Gradual Revelation, and the Immediate Revelation Equilibrium–together with All Defection Equilibrium, exhaust the set of possible equilibria. The height of each equilibria sketches the expected welfare, to be discussed later.*

A: At $\bar{c} = 0.2$*, the Cheating Equilibrium requires* α *of* 0.8*, and yields expected welfare of* 0.428
B: At $\bar{c} = 0.4$*, the Gradual Revelation Equilibrium requires* α *of* 0.7*, and yields* 0.184
C: At $\bar{c} = 0.6$*, the Immediate Revelation Equilibrium requires* α *of* 1*, and yields* 0.041

Figure 2.1: Existence of Equilibrium Types by Trust Level

2.3 Analysis

2.3.1 Maximal Cooperation

We grouped the types of equilibria in our game of trust building by how the eventual defectors reveal themselves under each equilibrium type, and how the revelation pattern induces corresponding changes in the belief system. A cooperation in any period under the Immediate Revelation Equilibrium signals that a player is sure to cooperate in all periods, whereas the signaling effect of a cooperation under Gradual Revelation Equilibrium is incomplete and gradual over repeated cooperations.

In analyzing the implications of each equilibrium, as important as the revelation pattern is the eventual cooperation range–how often the opponent is expected to be an eventual cooperator as opposed to a defector. In the binary type model of Kydd[2005], in which the actions each type wants to take is exogenously given, the Immediate Revelation Equilibrium represents an immediate inception of trust as the cooperative players know they will cooperate with each other throughout the game. In our continuous type model in which action types are endogenous, however, evaluating how reassuring an equilibrium is to the cooperative type is a circular endeavor as the equilibrium simultaneously decides who the cooperative types are. In particular, in the Immediate Revelation Equilibrium where the players finish signaling in the first period, the threshold t that solves the incentive compatibility equation unilaterally decides the cooperation region. In the Gradual Revelation Equilibrium, analysis of welfare and the notion of trust is more involved than the position of a single threshold, but the corresponding notion of the eventual cooperation range is a relevant measure nonetheless. We propose the following definition to formalize this concept.

Definition 2.5 (Maximal Revelation)**.** Under trust building equilibria of the K period game, we say the players **maximally reveal** when all but those for whom it is a dominant strategy to defect in a stage game cooperate in the penultimate period. In other words, $t_{K-1} = \bar{c}$.

Remember from our discussion of the benchmark one shot game of cooperation that $\bar{c}$ is the threshold cost type below which the players have the dominant strategy to defect. In context of our K period model, any cost type below $\bar{c}$ can never be made to cooperate in the final period. Therefore, our definition of maximal revelation can be thought of as setting the most conservative line in the continuum of cost types so that all types who *can* cooperate *do* signal cooperation from the penultimate round of the game.

We are now ready to present the first main result of our analysis. Kydd[2005] finds that in a two period game of trust building with exogenously given binary types, a separating equilibrium in which the two types play the same action over both periods always exists with appropriately sized first period stake compared to the second. We provide a generalization of the result with continuum of cost types that endogenously decide which type plays which action.

Proposition 2.1. *In the 2 period game of trust building without bargaining* $(\bar{c}, \phi, (\mu = 0))$*, there always exists an Immediate Revelation Equilibrium that supports maximal revelation. In particular, maximal revelation can be achieved when* $\alpha = \frac{1-\bar{c}}{\bar{c}}$

Proof. Proofs of the propositions can be found in the appendix. □

The two period game of trust building is particularly interesting if we interpret the first stage as a signaling gesture for the second main stage. Under such interpretation, the Immediate Revelation Equilibrium in our model of trust building with continuous cost types describes a negotiation in which the players choose to play the same action in the main stage as they signaled in the first stage. Our first result then states that when the ratio of the signaling stake to the main game is equal to the ratio of the higher cost types to the lower with respect to $\bar{c}$–the stage game dominant strategy threshold–the players choose to signal and play cooperation in both stages as long as their private costs are above $\bar{c}$.

We could also interpret the knife-edge α that supports maximal revelation as players choosing the size of the signaling stake before their types get realized. That there exists a unique size of the signaling round that induces maximal truthful signaling establishes grounds for the players to choose the appropriate signaling stake with respect to the relevant parameters–the cooperation benefit b and the first mover advantage ϕ in case of conflict. Proposition 2.1 states that the more of the "high cost" types there are, the larger the signaling stake that induces maximal revelation becomes.

As astonishing as it is to find that the separating equilibrium result of binary action type model generalizes to our model with continuous cost types, we acknowledge the limitations in applying this result to trust dynamics in the real world. In particular, thus far, the highest cost type has no reason to hesitate broadcasting his private cost and achieving mutual cooperation whenever possible. In the current model without bargaining, the opponent's high cost only serves to signal guaranteed cooperation and thus encourages mutual cooperation instead of providing incentive to take advantage of such high cost. In practice, however, it is hard to imagine that advertising one's vulnerability to defection would be a solid negotiation tactic.

We introduce bargaining to the game of trust building to address this gap between traditional trust building models and real world negotiation tables. With bargaining payoffs introduced in the model section, a player receives the inverse of her advertised cost type in comparison to her opponent. Therefore, communicating her very high cost may incentivize the opponents to defect earlier to enjoy bargaining edge for longer periods, including some lower cost types among the eventual cooperators of the original game. This introduction of bargaining can be thought of as another augmentation of Kydd[2005] as he suggests it would be "interesting to modify [the fully stationary model] in future research to allow for gestures that weaken a state in subsequent disputes."

While we detail the impact of bargaining in later sections, we note that bargaining adds incentive to defect across all types as described above. In particular, as the marginal type near the threshold $\bar{c}$ defects, bargaining precludes maximal revelation in the Immediate Revelation Equilibrium in trust building game of any length.

Proposition 2.2. *In the trust building game* $(\bar{c}, \phi, (\mu \neq 0))$ *of any length, Immediate Revelation Equilibrium can never support maximal revelation.*

On the other hand, under Gradual Revelation Equilibrium, defection strategy and the corresponding belief system evolve gradually over multiple periods. To fix ideas, take the two period model. Under the Gradual Revelation Equilibrium, some of the eventual defectors with cost below $\bar{c}$–who would have defected from the first period under the Immediate Revelation Equilibrium–may signal cooperation in the first period to extract a larger ϕ in later period. Rationally expecting this behavior, some of the cooperators in the first period do not believe the opponent's cooperation as truthful signal and defensively defect in the second period to protect against the cheating. As higher cost represents lower defection benefit to the defectors and higher cost to defection to the cooperators, it is naturally the high costs among the eventual defectors that cheat (play (C, D)), and the low costs among the original cooperators that defensively play (C, D).

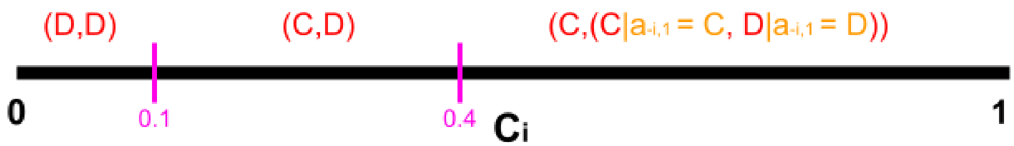


Figure 2.2: Strategies under GRE2 when $(\bar{c}, \phi) = (0.3, 0.2)$ and $\alpha = 0.6$

As players with costs marginally above $\bar{c}$ defensively react to the cheating players with costs marginally below $\bar{c}$, the mass of players who play strategy (C, D) have private cost range that always encompasses $\bar{c}$. In context of maximal revelation, it not only means that the Gradual Revelation Equilibrium is not compatible with maximal revelation, but it also implies that the cooperation range is always short of maximal in the last period.

Proposition 2.3. *In the 2 period game of trust building $(\bar{c}, \phi, \mu)$, Gradual Revelation Equilibrium can never support maximal revelation. Moreover, the last period cooperation is also never maximal under Gradual Revelation Equilibrium, even without bargaining.*

Does this mean that as we allow players cheat their signal and defect on later periods, we necessarily sacrifice the eventual cooperation range? We find that as we make the game longer, maximal revelation in the penultimate period and maximal cooperation in the last period are both possible under Gradual Revelation Equilibria. In particular, in the three period game, we find that maximal revelation and maximal cooperation always coincide under Gradual Revelation Equilibrium, meaning that no player starts defection on the third period.

Corollary 2.3.1. *In the three period game of trust building $(\bar{c}, \phi, \mu)$, maximal cooperation under the Gradual Revelation Equilibrium always coincides with maximal revelation. In other words, $t_{K-1} = \bar{c}$ whenever $t_K = \bar{c}$.*

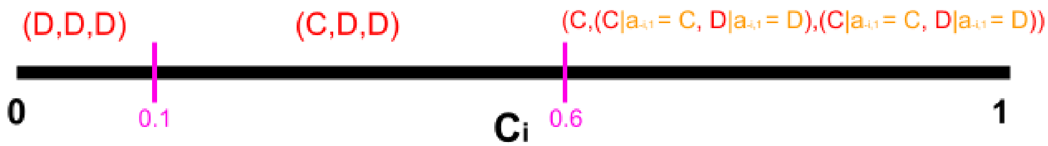
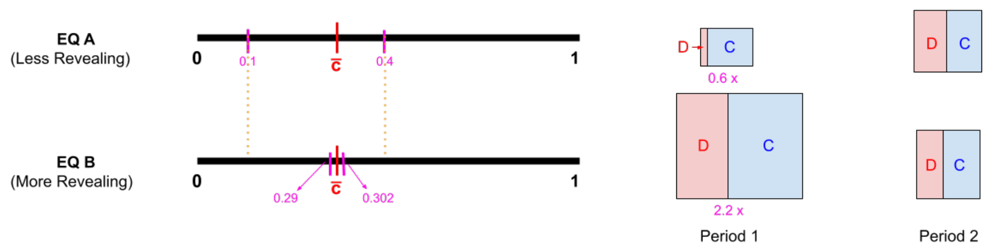


Figure 2.3: Strategies under GRE3 when $(\bar{c}, \phi) = (0.6, 0.3)$ and $(\alpha_1, \alpha_2) = (0.2, 0.8)$

2.3.2 Welfare

Thus far, we have concentrated on the cooperation range in the context of building trust, and how various equilibria achieve maximal cooperation in the ultimate stage of the game. We also looked into the concept of revelation in the penultimate period, with the interpretation of providing reassurance in the final stage as all actions are truthfully signalled for. While cooperation range and truthful signaling are natural metrics in analyzing the Immediate Revelation Equilibria, the analysis of trust gets more involved under the Gradual Revelation Equilibria, in which players of different cost types cooperate for varying duration of stages.

To illustrate, let us take as examples two Gradual Revelation Equilibria in the two period model without bargaining when the parameters are: $(\bar{c}, \phi, b) = (0.3, 0.2, 0.4)$. When the signaling stake $\alpha = 0.6$, the two thresholds that determine an equilibrium as fixed points are $t_1 = 0.1$ and $t_2 = 0.4$. On the other hand, when $\alpha = 2.2$, the two thresholds become $t_1 = 0.2933$ and $t_2 = 0.3027$. Remember, under the Gradual Revelation Equilibrium of two periods, a player of a cost type below t_1 defects for both periods, of type between t_1 and t_2 cooperates in the first period then defects in the second, and of types higher than t_2 cooperates for both periods as long as the opponent cooperates in the first period. The following figure illustrates the two examples.



Equilibrium B exhibits a bigger eventual cooperation range $[0.302, 1]$*, but at the cost of much bigger signaling stake. Under Equilibrium A, players not only utilize much smaller signaling stake to build trust but also cooperate more in the first period.*

Figure 2.4: Two examples of the Gradual Revelation Equilibrium

As shown in figure 2.4, Equilibrium B elicits a bigger eventual cooperation range than Equilibrium A. If the objective of the game were to achieve maximum likelihood of cooperation, the players would unambiguously prefer the signaling stake be 2.2. Notice, however, that under the second equilibrium, there is 29.3% chance that the opponent defects from the first stage. And in our model, the first stage not only serves as a signaling round but also incurs a game component of the payoff. Therefore, while the second equilibrium induces more cooperation in the final round, it also requires a much bigger signaling stake at which period the opponent defects with much bigger probability.

As the example illustrates, the cooperation range is an incomplete metric in analyzing the Gradual Revelation Equilibrium, in which each period represents utility generating rounds on top of serving the signaling purpose. In order to address the issue, we introduce a welfare notion to evaluate an equilibrium. For each equilibrium, given a player's cost type, we can calculate her expected utility against a random opponent. We calculate such expected utility for all cost types that are willing to cooperate if the opponent cooperates; note that this range of cost types correspond to region above $\bar{c}$. Taking the expectation of these expected utilities would give us a welfare notion that represents how reassuring an equilibrium is to the cooperative types. We formally define this concept as follows.

Definition 2.6 (Ex-ante Expected Utility)**.** In evaluating welfare under an equilibrium in the K period game of trust building, an **Ex-ante Expected Utility** is the expected utility a "cooperative" player is expected to gain from an equilibrium before both her and her opponent's type has been realized:

$$\mathbb{E}_{c_i \geqslant \bar{c}}\left[\mathbb{E}_{c_{-i}}\left[\frac{1}{\sum_{k=1}^{K}\alpha_k}\sum_{k=1}^{K}\alpha_k \cdot u_{i,k}\big(a_k(c_i, c_{-i})\big)\right]\right]$$

Evaluating the two equilibria in the example above with the ex-ante expected utility, we can assess that, prior to realization of the private cost types, players are actually better off under Equilibrium A.[27]

The applicability of ex-ante expected utility as a welfare measure is not limited to the Gradual Revelation Equilibria. We could also evaluate the utility a player is expected to gain out of the Immediate Revelation Equilibrium using the same measure. We recognize that comparing welfare implications of two separate equilibria of a game is a precarious endeavor, and we do not argue that the players choose to be in one equilibrium versus another. However, just as we established that maximal cooperation is only possible under the Immediate Revelation Equilibrium and not under the Gradual Revelation Equilibrium in the two period game, it is worth noting that the ex-ante expected utility in the Gradual Revelation Equilibrium is always higher than under the Immediate Revelation Equilibrium given the same parameters.

Proposition 2.4. *In the* 2 *period game of trust building* $(\bar{c}, \phi, \mu)$*, for any*[28] *Immediate Revelation Equilibrium, there exists a Gradual Revelation Equilibrium that is welfare-improving in the ex-ante expected utility sense.*

The main mechanism by which a Gradual Revelation Equilibrium welfare-improves on

[27] $\mathbb{E}U_A = 0.2013, \mathbb{E}U_B = 0.1976$. Riemann sum approximation method is detailed in the appendix.

[28] There can be up to two Immediate Revelation Equilibria per parameter set.

the Immediate Revelation Equilibrium in the 2 period game is twofold: 1) Gradual Revelation Equilibrium is supported by smaller signaling stake than the Immediate Revelation Equilibrium, and 2) interim cooperation in the first stage by the eventual defectors benefits the high cost types. While the second force gets stronger in games of longer periods, unlike the 2 period case, the Gradual Revelation Equilibria in games of longer periods often require bigger signaling stakes than the Immediate Revelation Equilibrium. Yet, the computations of welfare for the two equilibrium types show that the welfare result generalizes at least to the three period model without bargaining.

Corollary 2.4.1. *Proposition 2.4 generalizes to trust building games of* 3 *periods without bargaining.*

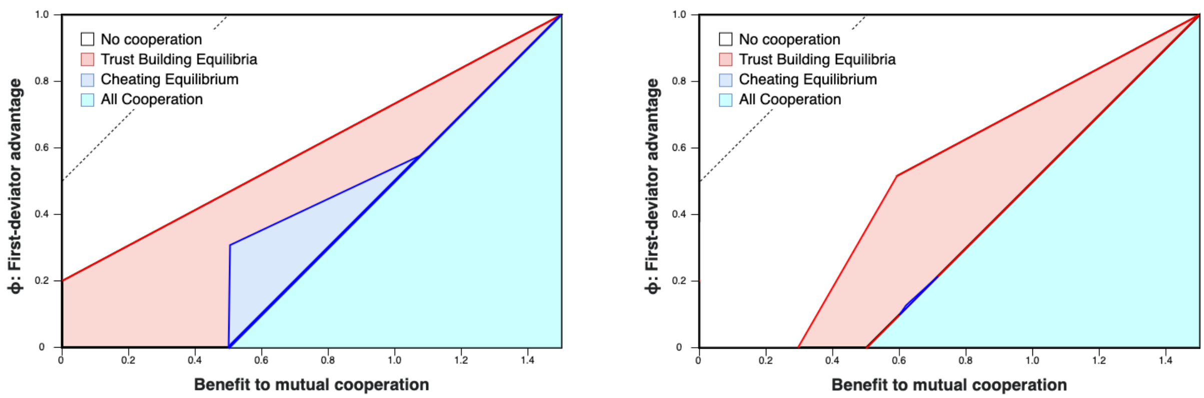
We need to note that the welfare notion we utilize is an ex-ante concept in terms of the realization of private cost types. One of the implications is that with low $\bar{c}$ or high cooperation benefit, the welfare notion will favor an equilibrium that exhibits the highest mutual cooperation in either period of the game. In particular, because the Cheating Equilibrium requires that all cost types play cooperation in the signaling period, even though players fail to gain any information out of the first round, it outperforms the trust building equilibria under which meaningful screenings occur.

Proposition 2.5. *In the* 2 *period game of trust building* $(\bar{c}, \phi, \mu)$*, when it exists, the Cheating Equilibrium always outperforms the Gradual Revelation Equilibrium and the Immediate Revelation Equilibrium.*

While the unanimous cooperation in the signaling period allows the Cheating Equilibrium to generate the highest welfare, it also constrains its existence. As $\bar{c}$ increases (trust level lowers), the Cheating Equilibrium is no longer supported by any signaling stake as the increasing defection bonus ϕ compared to the cooperation benefit b lures the low cost types to defect in the first period.[29] In the plots below, we show that as ϕ increases, the Cheating

[29]As $\bar{c} = \frac{1}{2} + \phi - b$, ϕ and b jointly determine $\bar{c}$.

Equilibrium is only supported by a narrow band of cooperation benefits, compared to other trust building equilibria on the left scatter plot. Figure 2.5 [30]also shows that the existence of the Cheating Equilibrium is also much more sensitive to bargaining than other trust building equilibria, as bargaining provides defection incentives across all types.



(a) Support of Equilibria when bargaining is unimportant ($\mu = 0$)

(b) Support of Equilibria when bargaining is more important ($\mu = 0.4$)

Figure 2.5: Support of the Cheating Equilibrium is Narrow and Highly Sensitive to Bargaining

2.3.3 Issue Sequencing

With the appropriate measure to assess an equilibrium's welfare implications in place, we are ready to analyze which game induces better outcome for the players. In particular, we could interpret the trust building games under different parameters as alternative negotiation sequences. Given the same parameters $(\bar{c}, \phi, \mu)$, do parties prefer to sequence bargaining deals so that the most consequential issue gets settled first? Or do they prefer to test waters with less significant issues, and then proceed to the most important issue after building necessary trust?

Notice that the two Gradual Revelation Equilibria of two periods given in Figure 2.4

[30]Recall that $\bar{c} = \frac{1}{2} + \phi - b$. Thus, restricting the range of $\bar{c}$ between 0 and 1 bounds b from above by $\phi + \frac{1}{2}$

describes these two alternatives. In the given example–in which it was more likely that a player is above the dominance threshold $\bar{c}$, and the cooperation benefit b was twice as big as the first mover defection payoff ϕ–the players enjoyed higher welfare when they sequenced issues from low stakes deal to the high stakes. Does this pattern generalize, so that players always prefer that issues be sequenced in increasing importance? If not, what determines when it is optimal to test the waters first? How does such optimal stakes evolution look like when we extend the game to more than two periods?

We start with a result that echoes our first proposition regarding the Immediate Revelation Equilibrium. In the Immediate Revelation Equilibrium, the threshold in cost type that divides the players who cooperate from those who defect summarizes the trust dynamic in the game. The lower the threshold, the higher the probability of mutual cooperation for both periods. Therefore, the optimal sequencing of stakes is those that result in the maximal revelation. Recall that in the two period game of trust building the maximal revelation becomes the unique Immediate Revelation Equilibrium when the relative size of the signaling stake exactly matches the odds ratio of being above dominance-defection threshold $\bar{c}$ to being below ($\alpha = \frac{1-\bar{c}}{\bar{c}}$).

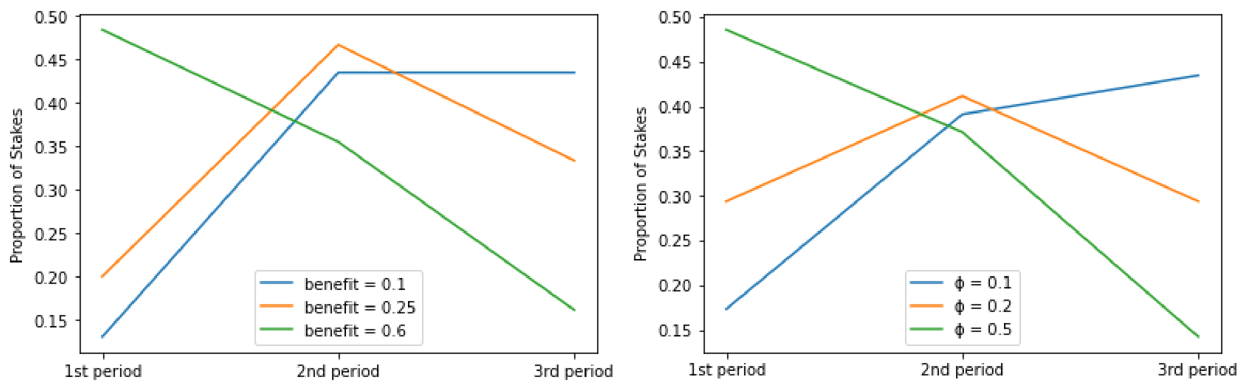
As players with cost type below $\bar{c}$ are sure to defect in the last period, $\bar{c}$ represents the overall trust level in the system. The higher the $\bar{c}$, the more people defect in the last period, and the smaller the signaling stake has to become to induce maximal revelation for the high cost types to willingly risk cooperation in the first stage. We find that this pattern of optimal signaling stake decreasing as $\bar{c}$ increases–or trust level decreases–generalizes to the General Revelation Equilibrium and the Cheating Equilibrium as well.

Note that the Gradual Revelation Equilibrium is summarized by two thresholds instead of one, and we need to utilize the welfare metric we established earlier in assessing which sequencing lead to better outcome for the participants.

Proposition 2.6. *In the 2 period game of trust building, the optimal signaling stakes increase as trust increases: As $\bar{c}$ decreases, α_{opt}–the signaling stake that maximizes the ex-ante* *expected utility–increases.*

In context of the 2 period game, this result about the size of the signaling round provides a complete answer to the issue sequencing problem. The higher the overall trust level, the more confident the cooperative players become in risking earlier rounds to make sure that they get truthful signaling. On the other hand, in games of longer periods, in addition to the ratio of the ultimate round to the rest of the game, we are also interested in how to sequence the multiple issues leading up to the final round. If the players postpone the settlement on the most important issues to the very last round, how do they stack the issues leading up to the finale? When they choose to open up with a grand gesture, do they follow it up by smoothly decreasing the stakes? Or do they switch to minor issues in the interim, and raise the stake back up in the final round?

The simplest model to investigate how players choose to allocate different stakes within the signaling rounds is to study optimal stakes structure in the Gradual Revelation Equilibrium of three periods. Normalizing the final round as 1, we could ask how the optimal stakes evolution looks like under different parameters of the model. Although we do not have an analytic result that is as definitive as Proposition 2.6, we do find a few patterns that bear important interpretations. First, we find that for high $\bar{c}$ (low trust level), which implies that the cooperation benefit b is smaller than the first defection bonus ϕ, the optimal stakes form a decrescendo followed by a grand gesture. In other words, under low trust, the optimal Gradual Revelation Equilibrium happens when both parties start with the most important issue, and then gradually decrease the stakes. Second, in the lower $\bar{c}$ range that is still above a half so that ϕ remains bigger than b, the players enjoy the highest welfare under the Gradual Revelation Equilibrium where the stakes are hump-shaped. That is, the parties start with a minute issue, address the most pressing concern in the interim, and cools down in the final stage. Finally, when $\bar{c}$ is in the medium range, and both the cooperation benefit b and the first defection payoff ϕ are small, the optimal stakes evolves in s-shape: the parties test the water with the smallest issue, and increases the stakes dramatically in the interim period, and then increases the stakes more slowly toward the final round. The figure below illustrates, by example, how the optimal stakes evolution responds to the fundamental parameters of the model: b and ϕ. The graphs are produced by holding the other fundamental parameters constant.



(a) Optimal Stakes w.r.t. changing b (b) Optimal Stakes w.r.t. changing ϕ

As cooperation benefit increases, the optimal stakes put higher emphasis on the signaling stakes.

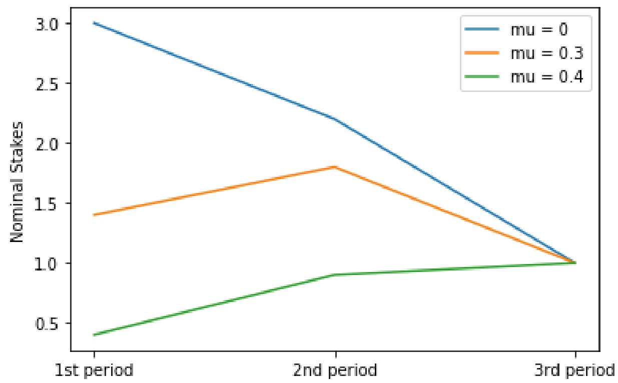
Figure 2.6: How Optimal Stakes Respond to Fundamental Parameters, ϕ and b

We do find an overarching condition that dictates the three different patterns of the optimal stakes above. It is that the growth factor from the first stakes to the second is always higher than the growth factor from the second stakes to the final. Not only is the condition met in the increasing stakes evolution–the s-shape and the hump-shape–but the condition also holds when the stakes are decreasing: in the grand gesture pattern, the drop in importance is less dramatic between the first and the second issue than between the second and the final. We formalize the finding as a proposition.

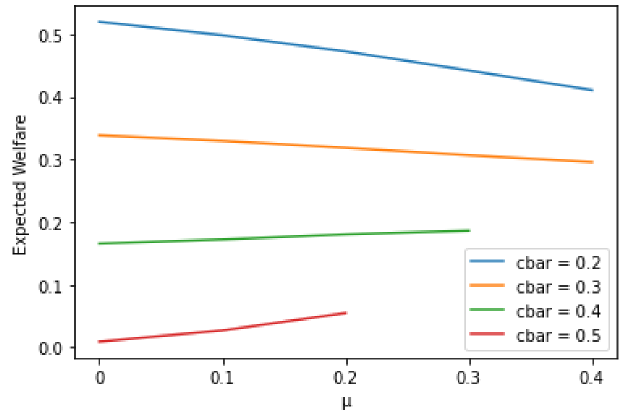
Proposition 2.7. *In the* 3 *period game of trust building without bargaining* $(\bar{c}, \phi, (\mu = 0))$, *the stakes,* α_1 *and* α_2 *that result in the Gradual Revelation Equilibrium which induces the* *highest ex-ante expected utility satisfy the following inequality:*

$$g_1 \equiv \frac{\alpha_2}{\alpha_1} \geqslant \frac{1}{\alpha_2} \equiv g_2$$

So far, we have investigated the issue sequencing problem in games without bargaining. How does introduction of bargaining to the trust building game affect the optimal choices of stakes evolution? As noted earlier, bargaining induces players across all cost types to defect more. Such additional defection incentives limit the range of parameters under which the Gradual Revelation Equilibria can be sustained. On the other hand, negative prospect from bargaining payoffs also brings the effect of lowering overall trust in the game. In light of Proposition 2.6 and the three patterns introduced above, lowering trust results in incentivizing higher stakes in the last period. The graphs below shows how the optimal stakes in the example above responds to introduction of bargaining (with $\mu = 0.2$). We can notice that bargaining induces more s-shaped evolution of stakes as the optimal choice under the same parameter set.



(a) Optimal nominal stakes w.r.t. changing μ



(b) Welfare on μ for various trust levels

Increasing bargaining weights shows two effects on the optimal signaling stakes:
1) The sum of the signaling stakes decrease, and
2) The evolution shifts toward an s-shape; bargaining reduces the first stage stake more dramatically than the second signaling stake.

The welfare graph on the right shows that as trust decreases, bargaining tends to increase welfare in the aforementioned $\alpha - \mu$ *trade-off.*

Figure 2.7: How Optimal Stakes Respond to μ and Welfare Implications

2.3.4 How Bargaining Affects Trust Building

In our earlier analysis of maximal cooperation, we found that as bargaining introduces unilateral incentive to defect across all types, it precludes the Immediate Revelation Equilibrium from maximally revealing, as the marginal types above $\bar{c}$ are pushed away from cooperating. It is important to stress, however, that it is not necessarily true that unilaterally providing defection incentives should result in all Perfect Bayesian Equilibria demonstrating higher defection. In fact, in certain Immediate Revelation Equilibria, as players rationally expect the opponents to defect more, there will be less cooperative opponents to take advantage of, which, in turn, decreases the incentive to defect for some cost types. Such effect is more stark in the Gradual Revelation Equilibria, where reputation affects trust building in more intricate ways as cost types reveal gradually over the periods.

We do find that introducing bargaining unambiguously limits how many equilibria can be supported given the parameters. While bargaining may not shift all equilibria toward more defection, it is still true that the predominant force bargaining has on our model of trust building is that bargaining provides further incentives to defect across all types. Such defection incentive proves incompatible with all three types of trust building equilibria, and while introduction of bargaining finds new equilibria under certain sets of parameters, it limits the parameters under which trust building equilibria may be supported in general.

There is another important implication that lowering trust has on our model of trust building. Recall from Proposition 2.1 that lower trust level in the game corresponds to smaller signaling stake that achieve maximal revelation in Immediate Revelation Equilibria. Following the comparative statics, one could conjecture that introducing bargaining that lowers overall trust level would then result in shrinking the optimal signaling stake. In fact, we find through simulation that the statement holds true for almost all optimal equilibria in the trust building game of two periods, as long as they are still supported.

Proposition 2.8. *In the* 2 *period game of trust building, increasing weight on bargaining decreases the size of signaling stake needed to support the most optimal trust building equilibria.*

We could interpret the proposition as follows. In the process of building trust, in order to get a more truthful signaling leading up to the ultimate stage of the negotiation, the players may choose to increase the signaling stake in order to incentivize the eventual defectors to reveal themselves sooner. Under such interpretation, bargaining may serve as an alternative device that provides incentive to defect earlier in the game, thus making the decision in the final stage more reassuring. We could thus understand Proposition 2.8 as demonstrating a trade-off between employing higher signaling stakes–investing in α–and introducing higher bargaining payoffs–investing in μ–in order to sustain trust building equilibria.

How does this $\alpha - \mu$ trade-off affect welfare? In particular, given the fundamental parameters $(\bar{c}, \phi)$, when are players better off building trust through more bargaining as opposed to through bigger signaling stakes? Since bargaining and signaling jointly affect trust building equilibria in an intricate manner, we do not claim that we have an analytic result that dictates when it is welfare-improving to build trust through bargaining. Instead, we share a computational result that reveals a consistent pattern. Our computed equilibria show that in supporting the Gradual Revelation Equilibria in games of lower trust, employing higher bargaining weights results in higher social welfare in the ex-ante expected utility sense than utilizing higher signaling stakes.

Proposition 2.9. *Among the optimal Gradual Revelation Equilibria in games of* 2 *and* 3 *periods, bargaining increases expected welfare when trust level is low. In other words, there is a threshold in the trust level under which increased weight on bargaining payoffs unambiguously increases maximal expected utility.*

As such, introducing bargaining to the game of building trust over multiple periods not only serves to make earlier results more realistic, but it also gives us insight about the relationship between defection incentives and welfare. In particular, the proposition supports empirical wisdom that when trust level is low, the common knowledge of unilateral motive to deviate increases the overall welfare by having agents buckle up the safety belts.

2.3.5 Building Trust Under Fixed Stakes

So far in our comparative analyses of welfare under various trust building equilibria, we have endogenized the signaling stakes $\boldsymbol{\alpha}$ with the interpretation of issue sequencing or employment of signaling versus bargaining as means of supporting defection incentives. We found that building trust gradually requires smaller signaling stakes and thus outperforms the immediate inception of trust in ex-ante welfare sense. On the other side of the coin, however, building trust over a fixed set of negotiation issues is often of interest as well. In this section, we study the welfare implications of various trust building equilibria in both 2 and 3 period games with fixed stakes. Through comparison, we also analyze the conditions under which there is opportunity to build trust over longer duration and what factors may close this window of opportunity.

In the welfare section, we established that between the two forms of trust building equilibria in 2 period games, the Gradual Revelation Equilibrium welfare-improves on the Immediate Revelation Equilibria by employing a smaller signaling round to build trust (Proposition 2.4). We find that this pattern–that gradual building of trust elicits higher social welfare than the immediate counterpart–holds more robustly in 3 period games as it also holds true even with fixed stakes. In the 3 period game of trust building, when the first two stages serve as signaling rounds for the main contention in the last period, we find that the Gradual Revelation Equilibrium always induces higher expected welfare than the Immediate Revelation Equilibrium. We formalize the finding in the following proposition.

Proposition 2.10. *In the* 3 *period game of trust building* $(\bar{c}, \phi, \mu)$*, with fixed signaling stakes* $\alpha_1, \alpha_2 < 1$*, for any Immediate Revelation Equilibrium, there exists a Gradual Revelation Equilibrium that is welfare-improving.*

This is a stronger result than Proposition 2.4, where the main force through which we found a Gradual Revelation Equilibrium that outperforms all Immediate Revelation Equilibria was by endogenizing the size of the stake. It is both intuitive and reassuring to find that as we analyze trust building games of longer duration, welfare favors the gradual building of trust more readily. The more stages there are over which to build trust, the better the players utilize each opportunity to establish trust rather than rushing all signaling in the first round and playing the sure game for the rest.

We can also perform different sets of welfare comparisons to address more directly this concept of windows of opportunity to gradually build trust. Consider planning for negotiation with endogenous number of signaling rounds as follows. There is a set issue of fixed stake in the final stage. The players may wait until the final stage to play the one shot game of cooperation, or they may open up one more negotiation round before the main stage and utilize the additional stage to communicate their private types. This would be the comparison between the Nash Equilibrium in the benchmark one shot game of cooperation and the trust building equilibria in the 2 period games that we analyzed. We found that the players elicit higher cooperation from each other by utilizing the signaling round to separate the types. We now analyze the extension of this comparison: given the set issues of fixed stakes in the final two periods, when can the players benefit from additional signaling? In other words, knowing the stakes of the main issue and the size of penultimate signaling round, when is there a window of opportunity to start earlier?

Given the parameters of the 2 period game $(\bar{c}, \phi, \mu)$ and the fixed stake of the penultimate signaling round, α_2, we first document which equilibrium elicits the highest expected welfare. We already established in Proposition 2.5 that when trust level is high enough to support the Cheating Equilibrium, it outperforms both the immediate and the gradual version of trust building equilibria that require some players to play defection in the signaling round. While we also established that with endogenous α_2, the Gradual Revelation Equilibrium welfare-improves on the Immediate Revelation Equilibrium, when we require that both equilibria be supported by the same fixed stake, we could not find a distinctive pattern under which one equilibrium elicits higher welfare than the other. Then, we compare the highest welfare achieved by any equilibrium under $(\bar{c}, \phi, \mu, \alpha_2)$ with the welfare generated by the Gradual Revelation Equilibrium of the 3 period game with the same $(\bar{c}, \phi, \mu, \alpha_2)$, where the size of the earliest signaling round α_1 is endogenously chosen to maximize welfare. The welfare comparison bears the following pattern.

Proposition 2.11. *For any equilibrium in the 2 period game $(\bar{c}, \phi, \alpha_2, \mu)$, if there exists a Gradual Revelation Equilibrium in the 3 period game with the same parameters plus a welfare-maximizing α_1, it induces higher welfare when:*

1. *The final period stake is sufficiently higher than the signaling stakes,*

2. And the cooperation benefit and the first mover advantage are sufficiently high.

In words, we find that the sufficient condition that the players may benefit from earlier signaling stage is that 1) the signaling stage stakes are low relative to the final stage, and 2) the per-period swing between mutual cooperation benefit and unilateral defection advantage is high.

2.4 Conclusion

In the introduction to the article, we argued that trust has been studied without taking bargaining dynamics into account. The reverse is perhaps even more true, particularly in the international relations literature. It is important to understand signaling a willingness to cooperate, to exist together in peace, alongside signaling a willingness to resort to force. Not that this has not been addressed,[31] but the literature devoted to signaling a willingness to cooperate is a vanishingly small fraction of the literature devoted to the other. The focus on the credibility of commitments to use force is perhaps a result of the Cold War problems of deterrence. Even there, however, as other authors have shown,[32] problems of trust loomed large. Trust is the facilitator of all aspect of social relations because it relates to the resources devoted to productive activities as opposed to conflict and to the risks that individuals and groups are willing to take to reduce the dangers of conflict. Trust building has been studied, but there is reason to think it should be examined more.

In this chapter, we make a start in this direction. We confirm some past results and question others. We question the view that trust building is relatively easy and can be efficiently accomplished quickly. In a rationalist framework, it cannot. Usually–always when the stakes are building to final consequential encounter–it must take place in a series of

[31] On the need for the credibility of the promise to cooperate if demands are met to be credible in order to coerce, see George et al. [1991] and Jervis [2003].

[32] See, for instance, Larson [1997], Kydd [2005].

interactions. The need to bargain is a crucial factor inhibiting trust building, and this factor can also be welfare improving. It makes actors less willing to signal trustworthiness, but this also means that actors can signal trustworthiness with lower stakes, which means lower risk overall.

We have also demonstrated results on the difficult questions of issue sequencing. When the bargaining aspect of the interaction is important or the consequences of misplaced trust are high, grand gestures–starting with high stakes issues–maximize welfare. When trust is more important that bargaining or the consequences of misplaced trust are low, it is better to begin with the easier, less consequential issues. We also show that if the optimal stakes are increasing, they increase more in the early or middle stages than in the latter ones. This is because of the dynamics of the need to screen out the actors who would be tempted to pretend to be trustworthy until the latter stages.

One of the important implications of the analysis is that there are windows of opportunity to build trust. When they are missed, there may not be enough time to recover. This decreases the welfare of all.

One reason this may be important is that international relations today are more adversarial than they were in the period after the Cold War ended. Several state dyads of great consequence (Russia - U.S., China - U.S., and North Korea - U.S.) are more fraught with conflicts than they were in the recent past. Rather than building trust, it is arguable that trust among these powers has diminished. At the end of the Cold War, for instance, Russia allowed Soviet Republics to break away and Russian influence to diminish greatly, even in its traditional spheres of influence. The United States did not initially fully exploit Russia's pulling back; in fact U.S. Secretary of State James Baker promised that NATO "there would be no extension" of NATO's jurisdiction "one inch to the east."[33] Subsequent history gave the lie to that. Arguably, Russia showed itself willing to cooperate until the United States

[33] Trachtenberg [2020].

showed itself unwilling to reciprocate. The relationship today–trust–has suffered greatly as a result.

If predictions about the strategic implications of changing technologies are borne out, these declines in trust may be extremely consequential. If AI technologies destabilize the nuclear balance, for instance, trust will be the key factor determining the risks of a catastrophic–potentially civilization ending–nuclear accident that countries take on. We would then greatly regret the current missed opportunities to build trust over time.

The case of North Korea is an interesting one because the bargaining and trust dynamics are each on full display. Both sides are worried that the other is willing to behave cooperatively in response to cooperation, and both sides are bargaining. Consider Lindsey Graham's statement, which he said characterized U.S. President Trump's thinking: "If there's going to be a war to stop [Kim Jong Un], it will be over there. If thousands die, they're going to die over there. They're not going to die here."[34] On the one hand, the statement can be understood in terms of bargaining: the U.S. would like to signal its willingness to bear the costs of war in order to improve its bargaining position. On the other hand, however, if the U.S. is able to signal that it sees low costs to conflict, it is a less trustworthy partner. This, in turn, potentially places the U.S. in great danger because a distrustful North Korea is more likely to take actions itself that might impose enormous costs on the U.S.

The findings suggest many fascinating paths for future research. Empirical evidence suggests, for instance, that the psychological and rationalist approaches to belief updating are often used in tandem [Yarhi-Milo et al., 2018]. Sometimes, however, psychological dispositions cause actors to choose suboptimal actions, ignoring Bayesian updating of trust in favor of psychological judgments [Hall and Yarhi-Milo, 2012b, Larson, 1997]. One promising area of future research is elucidating the precise nature of the interaction between psychological

[34]NBC News, Sen. Lindsey Graham: Trump Says War With North Korea an Option, 2 August, 2017, available here: https://www.nbcnews.com/news/north-korea/sen-lindsey-graham-trump-says-war-north-korea-option-n788396.

and strategic trust updating. Another direction for future work is to consider trust building in the context of unequal or changing balances of power. How should actors build trust when power is slowly or rapidly shifting in favor of one side? How would the dynamics change if there were uncertainty over the direction of the power shift? We hope other researchers will continue to investigate these topics through these questions and many others. In spite of the insightful work that has already been done in the area, we believe it remains dramatically understudied, particularly in relation to its importance in social relations.

2.5 Appendix

2.5.1 Sequential Rationality of 2-Period Equilibria

In the main text, we formally defined the three types of equilibria in the 2-period game of trust building as strategy profiles paired with consistent belief systems. The equilibrium concept we employ is the Perfect Bayesian Equilibrium; the strategies and the beliefs need to satisfy two conditions: sequential rationality and consistency of beliefs. In this first section of the appendix, we complete the definitions of the trust building Perfect Bayesian Equilibria by laying out the sequential rationality conditions that pin down where the thresholds get drawn in each equilibrium.

1. **Immediate Revelation Equilibrium**

$$s_i(c_i) = \begin{cases} (C, (C \mid_{a_{-i,1}=C}, D \mid_{a_{-i,1}=D}), ..., (C \mid_{a_{-i,K-1}=C}, D \mid_{a_{-i,K-1}=D})) & \text{if } c_i > t \\ (D, ..., D) & \text{if } c_i \leqslant t \end{cases}$$

$$b_i = \begin{cases} \mathbb{P}(c_{-i} > t \mid_{a_{-i,l}=C}) = 1 \\ \mathbb{P}(c_{-i} \leqslant t \mid_{a_{-i,l}=D}) = 1 & \forall l \leqslant K \\ \mathbb{P}(\text{anything else}) = 0 \end{cases}$$

where t is a threshold that satisfies the following conditions[35]:

(a) $\mathbb{E}(U(C,(C|C,D|D))) < \mathbb{E}(U(D,D))$ when $c_i < t$ and

$\mathbb{E}(U(C,(C|C,D|D))) \geqslant \mathbb{E}(U(D,D))$ when $c_i \geqslant t$:

$$(1-\mu)((\alpha+1)t-1)\phi + (1-t)(\alpha+1)(\bar{c}-t)) + \mu(\alpha+1)(\tfrac{2-t}{3-2t} - \tfrac{1}{2}) = 0$$

(b) $\mathbb{E}(U(C,D)) < \mathbb{E}(U(D,D))$ when $c_i < t$:

$$(1-\mu)((\alpha+1)t-1)\phi + (1-t)\alpha(\bar{c}-t)) + \mu(\alpha+1)(\tfrac{2-t}{3-2t} - \tfrac{1}{2}) \geqslant 0$$

---

[35] $\mathbb{E}(U(C,(C|C,D|D))) \geqslant \mathbb{E}(U(C,D))$ when $c_i \geqslant t$) is always satisfied: 1. $b > a$ and 2. $b > \frac{1}{2} + \phi - c_i$ when $c_i \geqslant t \geqslant \bar{c}$)

2. **Gradual Revelation Equilibrium**

$$s_i(c_i) = \begin{cases} (C, ..., (C \mid_{a_{-i,K-1}=C}, D \mid_{a_{-i,K-1}=D})) & \text{if } c_i > t_K \\ (C, ..., (C \mid_{a_{-i,K-2}=C}, D \mid_{a_{-i,K-2}=D}), D) & \text{if } t_{K-1} < c_i \leqslant t_K \\ (C, ..., (C \mid_{a_{-i,K-3}=C}, D \mid_{a_{-i,K-3}=D}), D, D) & \text{if } t_{K-2} < c_i \leqslant t_{K-1} \\ ... & ... \\ (D, ..., D) & \text{if } c_i \leqslant t_1 \end{cases}$$

$$b_i = \begin{cases} \mathbb{P}(c_{-i} > t_l \mid_{a_{-i,l}=C}) = 1 \\ \mathbb{P}(t_{l-1} < c_{-i} \leqslant t_l \mid_{a_{-i,l-1}=C \ \& \ a_{-i,l}=D}) = 1 \\ \mathbb{P}(c_{-i} \leqslant t \mid_{a_{-i,l}=D}) = 1 \\ \mathbb{P}(\text{anything else}) = 0 \end{cases} \qquad \forall l \leqslant K$$

where t_1 and t_2 are the thresholds that satisfy the following conditions:

(a) $\mathbb{E}(U(C,(C|C,D|D))) \geqslant \mathbb{E}(U(C,D))$ when $c_i \geqslant t_2$ and

$\mathbb{E}(U(C,(C|C,D|D))) < \mathbb{E}(U(C,D))$ when $t_2 \geqslant c_i \geqslant t_1$:

$$(t_2-t_1)((1-\mu)\phi+\mu(\tfrac{2-t_1-t_2}{3-t_1-2t_2}-\tfrac{1}{2}))+(1-t_2)((1-\mu)(\bar{c}-t_2)+\mu(\tfrac{2-t_1-t_2}{3-t_1-2t_2}-\tfrac{1}{2}))=0$$

(b) $\mathbb{E}(U(C,D)) \geqslant \mathbb{E}(U(D,D))$ when $c_i \geqslant t_1$ and

$\mathbb{E}(U(C,D)) < \mathbb{E}(U(D,D))$ when $c_i < t_1$:

$$t_1 * ((1-\mu)\alpha\phi + \mu(\alpha+1)(\tfrac{2-t_1}{3-2t_1} - \tfrac{1}{2}))$$
$$+ (t_2-t_1)((1-\mu)\alpha(\bar{c}-t_1) + \mu(\alpha+1)(\tfrac{2-t_1}{3-2t_1} - \tfrac{1}{2}))$$
$$+ (1-t_2)((1-\mu)(\alpha(\bar{c}-t_1)-\phi) + \mu(\alpha(\tfrac{2-t_1}{3-2t_1} - \tfrac{1}{2}) + \tfrac{2-t_1}{3-2t_1} - \tfrac{2-t_1-t_2}{3-t_1-2t_2}) = 0$$

3. **Cheating Equilibrium**

$$s_i(c_i) = \begin{cases} (C, ..., (C \mid_{a_{-i,K-1}=C}, D \mid_{a_{-i,K-1}=D})) & \text{if } c_i > t_K \\ (C, ..., (C \mid_{a_{-i,K-2}=C}, D \mid_{a_{-i,K-2}=D}), D) & \text{if } t_K \geqslant c_i > t_{K-1} \\ ... & ... \\ (C, C, D, ..., D) & \text{if } t_2 < c_i \leqslant t_3{}^{36} \\ (C, D, ..., D) & \text{if } c_i \leqslant t_2 \end{cases}$$

$$b_i = \begin{cases} \mathbb{P}(c_{-i} > t_l \mid_{a_{-i,l}=C}) = 1 \\ \mathbb{P}(t_{l-1} < c_{-i} \leqslant t_l \mid_{a_{-i,l-1}=C \ \& \ a_{-i,l}=D}) = 1 \\ \mathbb{P}(c_{-i} \leqslant t_2 \mid_{a_2=D}) = 1 \\ \mathbb{P}(\text{anything else}) = 0 \end{cases} \quad \forall l \leqslant K$$

where t is a threshold that satisfies the following conditions:

(a) $\mathbb{E}(U(C,(C|C,D|D))) < \mathbb{E}(U(C,D))$ when $c_i < t$ and

$\mathbb{E}(U(C,(C|C,D|D))) \geqslant \mathbb{E}(U(C,D))$ when $c_i \geqslant t$:

$(1-\mu)((1-t)(\bar{c}-t)+t\phi)+\mu(\frac{2-t}{3-2t}-\frac{1}{2})=0$

(b) $\mathbb{E}(U(C,D)) > \mathbb{E}(U(D,D))$ when $c_i < t$:

$(1-\mu)(\alpha\bar{c}-(1-t)\phi)+\mu(\alpha(\frac{2}{3}-\frac{1}{2})+\frac{2}{3}-(\frac{1}{2}t+(1-t)\frac{2-t}{3-2t}))\leqslant 0$

2.5.2 Maximal Revelation Proofs

In this section of the appendix, we provide analytic proofs for the first three main results along with a corollary. The driving force in deciding whether a game could result in maximal revelation is the incentives of marginal types sustaining the threshold around the stage game dominance threshold $\bar{c}$.

[36] In a 2-period case ($K=2$), t_3 becomes 1, and there are only two strategies (C,D) and (C,C)

[36] $\mathbb{E}(U(C,(C|C,D|D))) \geqslant \mathbb{E}(U(C,D))$ when $c_i \geqslant t$) is always satisfied: 1. $b > a$ and 2. $b > \frac{1}{2}+\phi-c_i$ when $c_i \geqslant t \geqslant \bar{c}$)

Proof of Proposition 2.1. [37] Maximal revelation strategy is "Play $(C, (C|C, D|D))$ if $c_i > \bar{c}$, play (D, D) otherwise." Consider the following incentives to deviate:

i. $c_i = \bar{c} - \epsilon$ can profitably deviate to (C, D) when $F(t) < \frac{1}{\alpha+1}$

   Compared to the revelation strategy, $c_i < t$ may risk losing ϕ in the first stage to the low cost types, for gaining $\phi - \epsilon$ from the high cost types in the second stage. Specifically, by changing from (D, D) to (C, D),

   – From the low types $(F(\bar{c}))$: lose ϕ in stage 1,

   – From the high types $(1 - F(\bar{c}))$: $(\frac{1}{2} + \phi - c_i, \frac{1}{2} - c_i) \rightarrow (b, \frac{1}{2} + \phi - c_i)$: lose $\bar{c} - c_i$ in stage 1, and "lose" $-\phi$ in stage 2

   In expectation, the gain from deviation exceeds the loss when

$$c_i > \bar{c} - \frac{1}{\alpha}\big(1 - \alpha\frac{F(\bar{c})}{1 - F(\bar{c})}\big)\phi$$

   Since $c_i = \bar{c} - \epsilon$, this profitable deviation is possible when $F(\bar{c}) < \frac{1}{\alpha+1}$

ii. $c_i = \bar{c} + \epsilon$ can profitably deviate to (D, D) when $F(\bar{c}) > \frac{1}{\alpha+1}$

   For $c_i > \bar{c}$, loss from first stage ϕ could be larger than the possible cooperation on both stages. Specifically, by changing from $(C, (C|C, D|D))$ to (D, D),

   – From the low types $(F(\bar{c}))$: gain ϕ in stage 1,

   – From the high types $(1 - F(\bar{c}))$: $(b, b) \rightarrow (\frac{1}{2} + \phi - c_i, \frac{1}{2} - c_i)$: "gain" $\bar{c} - c_i$ in stage 1, "gain" $\bar{c} - \phi - c_i$ in stage 2

   In expectation, the gain from deviation exceeds the loss when

$$c_i < \bar{c} - \frac{1}{\alpha + 1}\big(1 - \alpha\frac{F(\bar{c})}{1 - F(\bar{c})}\big)\phi$$

[37] In most of the main text, we work with the case in which the cost types are drawn from uniform distribution $F(c) = c$. The proof is on more general space of any continuous cumulative density function $F(c)$.

Since $c_i = \bar{c} + \epsilon$, this profitable deviation is possible when $F(\bar{c}) > \frac{1}{\alpha+1}$

Since $c_i < \bar{c}$ can profitably deviate when $F(\bar{c}) < \frac{1}{\alpha+1}$, and $c_i > \bar{c}$ can profitably deviate when $F(\bar{c}) > \frac{1}{\alpha+1}$, maximal revelation is only possible when $F(\bar{c}) = \frac{1}{\alpha+1}$.

□

Proof of Proposition 2.2. In order to sustain maximal revelation with bargaining, the threshold t that satisfies both sequential rationality condition of the Immediate Revelation Equilibrium should equal the dominance threshold $\bar{c}$.

$$(1-\mu)\cdot((\alpha+1)t-1)\phi + \mu\cdot(\alpha+1)(W_{CD}-0.5) = 0 \tag{2.1}$$

$$(1-\mu)\cdot((\alpha+1)t-1)\phi + \mu\cdot(\bar{c}(W_{CD}-0.5) + (1-\bar{c})(W_{CD}-W_D) \geqslant 0 \tag{2.2}$$

where $W_{CD} = \frac{1-t}{1-t+\frac{1}{2}-\frac{t}{2}} = \frac{2}{3}$, and $W_D = \frac{1-\frac{t}{2}}{1-\frac{t}{2}+\frac{1}{2}-\frac{t}{2}} = \frac{2-t}{3-2t} > \frac{1}{2}$.

Substituting $t = \bar{c}$, and subtracting (2.1) from (2.2) gives:

$$\begin{aligned}(2)-(1):\ \ \bar{c}\cdot(W_{CD}-\frac{1}{2}) + (1-\bar{c})(W_{CD}-W_D) &\geqslant (\alpha+1)(W_{CD}-\frac{1}{2})\\ \Rightarrow \bar{c}\cdot\frac{1}{6} + (1-\bar{c})\cdot(\frac{2}{3}-W_D) &\geqslant (\alpha+1)\cdot\frac{1}{6} \end{aligned} \tag{2.3}$$

Since $\bar{c} \in [0,1]$ and $W_D > \frac{1}{2}$, the LHS of (2.3) is a convex combination between $\frac{1}{6}$ and a number smaller than $\frac{1}{6}$, while the RHS is clearly larger than $\frac{1}{6}$ as $\alpha > 0$. Therefore, (2.3) can never be satisfied. □

Proof of Proposition 2.3. The sequential rationality condition of the Gradual Revelation Equilibrium *without bargaining* is given by:

$$(t_2 - t_1)\phi + (1-t_2)(\bar{c} - t_2) = 0$$

1. In order to achieve maximal revelation ($t_1 = \bar{c}$), and still support Gradual Revelation

Equilibrium ($t_2 > t_1$), there must exist a range of types strictly above $\bar{c}$ who employ (C, D) strategy in vacant fear of facing defection in the second period despite knowing for sure that the opponent who played C in the first period has cost type strictly above $\bar{c}$. It is easy to check that players will universally prefer to be in the maximally revealing Immediate Revelation Equilibrium, and continue to cooperate in the second period when faced with cooperation in the first period.

2. As for the maximal cooperation in the last period, consider $t_2 = \bar{c}$. The maximal cooperation would bring the second term to 0, leaving the first term to also be 0. This is only possible when $t_1 = t_2 = \bar{c}$, which violates the definition of the Gradual Revelation Equilibrium ($t_2 > t_1$). In fact, this case exactly corresponds to the maximal revelation in the Immediate Revelation Equilibrium. We can see that the Gradual Revelation Equilibrium converges to the Immediate Revelation Equilibrium as the revelation region gets close to maximal.

3. When we add bargaining, it would only add positive value to the left hand side, corresponding to the interpretation of adding unilateral incentive to defect. This would not only preclude maximal revelation in the Gradual Revelation Equilibrium, but also the Immediate Revelation Equilibrium, as $t_1 > t_2$ is the only way to support the incentive compatibility condition, which is a contradiction.

□

Proof of Corollary 2.3.1. The sequential rationality condition of the Gradual Revelation Equilibrium of 3 period games *without bargaining* is given by:

$$\begin{aligned} t_1 \cdot \phi + (1-t_1)(\bar{c}-t_1) &= (1-t_2)g_1\phi \\ (t_2-t_1) \cdot \phi + (1-t_2)(\bar{c}-t_2) &= (1-t_3)g_2\phi \\ (t_3-t_2) \cdot \phi + (1-t_3)(\bar{c}-t_3) &= 0 \end{aligned} \tag{2.4}$$

Maximal cooperation would mean that $t_3 = \bar{c}$. One could easily see from 2.4 that this is only possible when $t_2 = t_3$. □

2.5.3 Welfare Results

In the main text, we constructed a welfare notion as formally given in definition 2.6. The metric is an ex-ante welfare metric as it expectations out the agents' own cost type as well as the cost type of the opponent. As we interpret the metric as average utility that an agent expects to receive before her type or the opponent's type gets realized, we compare this ex-ante expected utility under different types of equilibria to gain insight on welfare.

This section of the appendix provides justification of the three results about ex-ante expected utility comparisons among trust building equilibria. First, on dense set of parameters [38], we calculate the thresholds t that characterize trust building equilibria by solving the system of sequential rationality conditions provided in the first section of the appendix. Then, we calculate ex-ante expected utility under each of the equilibrium characterization by taking Riemann sum approximation on both the own agent's and the opponent's cost type (c_i, c_{-i}). With the data frame of welfare calculations, we compare and contrast the three types of equilibria, applying necessary data processing with python pandas package. We defer the details constituting proofs of Proposition 2.4, Corollary 2.3.1, and Proposition 2.5 to the online appendix. The solutions of the equilibrium constructing conditions, as well as python scripts for welfare calculations and processing are also provided in the online appendix.

---

[38] $\bar{c} : [0, 1]$ in steps of 0.05
$\phi : [0 : 1 - \bar{c}]$ in steps of 0.05
$\mu : [0 : 1]$ in steps of 0.1
$\alpha_1 \& \alpha_2 : [0.1 : 3.8]$ in steps of 0.1

2.5.4 Issue Sequencing Results

This section of the appendix provides proofs for Proposition 2.6 and Proposition 2.7. Proposition 2.6 posits a computational pattern that as trust level falls off–or $\bar{c}$ increases–the size of the signaling stake under the optimal equilibrium decreases. The online appendix details the computations. We utilize the same data frame that we detailed in the welfare analysis section. Then, within each type of equilibria[39], we document the highest welfare yielding equilibrium. Iteratively checking how the welfare under each of the optimal equilibrium moves with increasing $\bar{c}$ demonstrates the pattern that as trust level decreases, so does the signaling stake needed to support the equilibrium that yields the highest welfare for given parameters.

Proof of Proposition 2.7. For this proof, we first provide an analytic result that convinces us to claim that optimal Gradual Revelation Revelation should always demonstrate $g_1 \geqslant g_2$. Then, we back up our claim with computations that confirms that optimal equilibria do demonstrate this rule given the parameters.

The following constitute sequential rationality conditions for Gradual Revelation Equilibria in a 3 period game:

$$t_1 \cdot \phi + (1 - t_1)(\bar{c} - t_1) = (1 - t_2)g_1\phi \tag{2.5}$$

$$(t_2 - t_1) \cdot \phi + (1 - t_2)(\bar{c} - t_2) = (1 - t_3)g_2\phi \tag{2.6}$$

$$(t_3 - t_2) \cdot \phi + (1 - t_3)(\bar{c} - t_3) = 0 \tag{2.7}$$

Suppose, by contradiction, that $g_2 > g_1$. Function $t\phi + (1 - t)(\bar{c} - t)$ is decreasing in t for all $t \leqslant \frac{\bar{c}-\phi+1}{2}$ which, in turn, implies that it holds for all $t \leqslant \bar{c}$ by optimality condition

[39] There are 5 types of equilibria in total: The Cheating, Gradual Revelation, and the Immediate Revelation equilibria for the game of 2 periods, and the Gradual Revelation and the Immediate Revelation Equilibria for the game of 3 periods.

$(1-\bar{c} \geqslant \phi)$. Thus, utilizing $t_2 > t_1$, we get the following inequality from (2.5) and (2.6):

$$(t_3 - t_2)g_2 \geqslant (t_3 - t_2)g_1 \geqslant (1 - t_2)g_1 - (1 - t_3)g_2 > t_1 \tag{2.8}$$

Now we have all ingredients for our claim. For optimal Gradual Revelation Equilibria, t_3 cannot be too much higher than $\bar{c}$, for that would suggest a considerable mass of types playing defection in the final round when it only hurts themselves. On the other hand, by (2.7), the tight $(t_3 - \bar{c})$ binds $(t_3 - t_2)$ as well. The only possibility to support tight $(t_3 - t_2)$ while sustaining (2.8) would be to have g_1 be large and t_1 be minimal at the same time. However, such combination would contradict (2.5).

Our bounds on parameters, as well as our definition of optimal Gradual Revelation Equilibria, are not tight enough to warrant proof of proposition. While our analytic portion of the proof results in a claim, in the online appendix, we provide actual computations of optimal Gradual Revelation Equilibria in the ex-ante welfare sense that indeed demonstrates that $g_1 \geqslant g_2$. □

2.5.5 How Bargaining Affects Welfare Under Trust Building Equilibria

In the main text, we discuss how bargaining affects trust building. We posit that the main force bargaining brings to the trust building game is to provide unilateral incentive to deviate across all cost types. While this force has implications on range of parameters that support Perfect Bayesian Equilibria as well as the thresholds that characterize actions under each equilibrium, we summarize the unambiguous effect bargaining has on welfare as two main propositions. In this section of the appendix, we explain the computational approach we took in order to detect such unambiguous pattern in welfare, and guide the readers to the online appendix for details.

We use the same data frame that documents solutions of Perfect Bayesian Equilibria under the fine grid of parameters as in the welfare section. While the welfare section dealt with comparisons on *all* equilibria that exist, in this section, we focus on the equilibrium that yields the highest expected utility per each parameter set. In the online appendix, under sections Proposition 2.8 and Proposition 2.9, we first provide this table of "optimal" equilibria. Then, we proceed to iterate over each trust parameters $(\bar{c}, \phi)$, checking that as we increase μ, the pattern we try to prove persists without an instance of reversal.

For Proposition 2.8, the pattern is that α_{opt}–the size of the signaling stake in the equilibrium with highest welfare–decreases as we increase the weight on bargaining. We do find less than 1% of outliers that do reverse this pattern, but the size of the welfare reversal is so small in those instances that we posit the general pattern as a proposition. Proposition 2.9 finds that as we increase the bargaining weight, the expected welfare increases unambiguously when the trust level is low enough. For the Gradual Revelation Equilibrium in the 2 period game, the threshold trust level is $\bar{c} = 0.5$, whereas in the 3 period game, the pattern persists when $\bar{c} \geqslant 0.55$.

The table of "optimal" equilibria under each parameter set, as well as the procedure for finding the pattern as we increase the bargaining weight μ is well documented in the online appendix with detailed captions.

2.5.6 Computations of Welfare Under Fixed Stakes

In this last section of the appendix, we provide computational justifications for Propositions 2.10 and 2.11. As in the welfare section, the propositions constitute comparisons of expected utility among trust building equilibria; only this time, the size of each stages is fixed. Proposition 2.10 states that when stakes are fixed, agents are better off on average under the gradual building of trust, as opposed to immediate inception of trust. The online appendix details this simple and strong welfare comparison.

Proposition 2.11 considers a more complicated comparison. As the main text explains, we consider optimally elongating a 2 period game by choosing the size of an additional signaling round, and supporting a Gradual Revelation Equilibrium of 3 periods. This comparison exercise represents adding one more opportunity to build trust. The comparison allows us to investigate when it is unambiguously beneficial to have one more opportunity to build trust. The online appendix demonstrates that the condition is two fold: 1) $\alpha_1, \alpha_2 < 0.3$, and 2) $b, \phi > 0.5$. The online appendix discloses the full python code that generates the comparisons as well as the result table and captions for explaining each step.

CHAPTER 3

Learning Through Matching[1]

3.1 Introduction

From successful transactions in TaskRabbit to tips in Uber Rides, many platforms derive benefit from matching one good agent to another. With thousands of new users per day, however, big matching platforms cannot ensure the quality of all members.

Based on these observations, this paper proposes a model of learning through matching whereby the platform dynamically maximizes output by controlling which agents to recruit and how to group them into pools. At each point in continuous time, the platform makes recruitment decisions over agents who have high or low pizzazz. Agents derive utility from matching with a partner of high pizzazz regardless of her own, whereas the platform wants to maximize the number of matches between the high pizzazz agents. In particular, we assume a supermodular match output whereby only the matches between the high types may create value to the platform. This feature applies to tipping behavior in Uber Rides, in which tip realizes only when a good driver takes a frequent tipper.

The platform in our model perfectly observes all output and its producers. Therefore, once an agent creates output, the platform learns of her high pizzazz without error; our model features the "perfect good news" signal structure. Within this framework, we study how the platform utilizes its pool of signaled agents to optimize output over time. We recognize the *exploration-exploitation trade-off* as main force: by exclusively matching these *proven*

[1]Co-authored with Isabel Juniewicz, UCLA

members to each other, the platform may maximize flow output while sacrificing discovery of other high pizzazz agents who are yet to signal; on the other hand, by keeping a general pool, the platform sacrifices instantaneous output but facilitates learning, which in turn, accelerates future output.

In the first part of the paper, we analyze how the platform balances this tradeoff when they can only maintain one matching pool. Essentially, this sub-model examines the optimal exploration decision of how to mix unknown type agents with the pool of *proven members*. We find that the solution is bang-bang: the optimal exploration policy prescribes full exploration followed by full exploitation. Along optimal schedule, the platform begins with full recruitment of unknown agents as *candidate members*. As matches realize, some candidate members with high pizzazz generate output and become proven members. Over the candidate members who don't, the platform grows more and more pessimistic until it drops them at once, and thereafter only keeps the agents who have signaled thus far.

At the heart of the analysis is the shuffling argument. Candidate members lose reputation as a function of their time spent in the platform. Therefore, given a choice between two unknown agents, the platform recruits one that has been recruited for less time. Since the platform in our model makes decisions in continuous time with discounting, optimal recruitment demonstrates two features: 1) constant shuffling of candidates to hold all unknown agents at equal posterior, and 2) the earlier the learning the better. We combine the two features to prove the bang-bang solution to the optimal exploration problem, and characterize the optimal stopping time of exploration.

The paper then turns to the optimal two pool control problem. This sub-model endows the platform the ability to segment its recruits into two pools. The platform may now keep different mixtures of *proven members* and its *candidate members* in two pools and find balance between exploitation and exploration. We surprisingly find that even with the flexibility to mix, the optimal allocation schedule prescribes full integration – recruiting all agents into one pool – until full segmentation – putting together all unknown agents into one pool, and constantly promoting signaled agents to the other "premium" pool.

The rest of the paper is organized as follows. Section 2 introduces the general model with arbitrary number of pools. Section 3 fixes the number of pools to one, and characterizes optimal exploration policy. Section 4 then turns to the two pool model, and describes the optimal allocation schedule of full integration to full segmentation. Section 5 concludes.

3.2 Model

3.2.1 Agents

We consider a one-to-one matching market of two sides. There is unit mass of agents on both sides, who are heterogeneous in their pizzazz θ. In particular, we assume a dichotomy in which p_0 of the agents are attractive ($\theta = H$) and the rest are not ($\theta = L$). Pizzazz in our model represents value to the other side. When an agent matches up with an attractive counterpart, she receives utility of 1 regardless of her own pizzazz, whereas she enjoys 0 utility from matching with an unattractive partner, also regardless of her type. We assume that the proportion of attractive agents in the market are symmetric. Agents are assumed to be memoryless and ignorant of their own pizzazz. In fact, the scope of the current paper takes agents as automata that only make participation decisions. Rather, the focus of the paper is on the matching platform as a designer that dynamically controls learning by its recruitment and matching decisions – subject to the participation constraints.

3.2.2 Matching

We consider a dynamic model in continuous time with discounting. For each $t \in [0, \infty)$, the platform symmetrically allocates $\{N_{i,t}\}_{i=1}^{m}$ agents from both sides into m matching pools.[2]

[2]We could generalize to allow asymmetric recruitment in future research. For the scope of the current article, we limit the platform's ability to symmetric recruitment, and ask how the platform optimally chooses to explore agents with unknown quality.

We will discuss how the platform populates multiple pools in "the platform's design problem" section of the model. Until then, to introduce notations with more ease, we suppose one set pool with N_t agents on each side and drop the i subscripts.

Matching is random and one-to-one with constant return to scale (CRS) technology.[3] In particular, over a time interval $[t, t + dt]$, $\mu N_t dt$ matches realize, where $\mu \in (0, 1)$ represents match friction. Matching is random in an independent and identical manner both across time and across agent types: in a given period, agents of high and low pizzazz face the same probability of matching regardless of their matching history.

We assume that matches produce output in stochastic and supermodular fashion. Only the matches between high-pizzazz agents independently have $\lambda > 0$ probability of producing output. Other combinations can never produce output. It is this match output that the platform dynamically maximizes through its pool allocation decision.

3.2.3 Information Structure

Here we lay out the information structure, which is at the heart of the study. At time 0, the platform believes that all agents in the market have p_0 probability of being the high type. In subsequent time, we assume a powerful platform; it observes all matches, and identifies ones that produce output. Along with the supermodularity of the match output function, this means that the platform completely learns the high quality of an agent the moment she contributes to an output. We assume that the platform keeps track of the identity of these *proven members* and we denote their mass n_t.

On the other hand, over the agents who have been in the platform but have not yet

[3] We discuss a generalization to increasing returns to scale in the concluding section.

generated output the platform updates the posterior belief p_t according to Bayes' rule[4]:

$$p_t \equiv \mathbb{P}(\theta = H | \notin n_t) = \frac{p_0 - n_t}{1 - n_t} \tag{3.1}$$

Implicit in notation 3.1 are two assumptions. First is that this agent has to be active in the platform for the entire $[0, t]$ time frame. We show that this always holds in Theorem 3.1: once the platform begins "exploring" an agent (recruits a *candidate member* from pool of unknown agents), it keeps him in the pool until p_t falls below a certain threshold, or he produces output, at which point he becomes a *proven member*. The second, more important, assumption is an anonymity assumption. While the platform follows the identity of the *proven members*, we assume that it does not exploit it in updating the reputation of the *candidate members* who match up with them differently from those who don't:

Assumption 3.1. The platform does not utilize the identity of *proven members* in updating the reputation of the *candidate members*.

This compromise we make to keep our model both tractable and realistic is standard in matching literature.[5]

3.2.4 The Platform's Design Problem

We are now ready to formally state the platform's design problem across m pools. Combining the CRS matching function and the supermodular output function, we can summarize the law of motion of expected output in pool i as:

$$\dot{M}_{i,t} = \lambda \cdot \mu N_{i,t} \cdot q_{i,t}^2 \tag{3.2}$$

[4]Theoretically, notation in the demonstrated updating is only sound in a 1-pool setting; formally, we would need at least one $p_{i,t}$ per pool in which the platform explores unknown agents. As we will show in the analysis section, however, the optimal policy dictates that the platform keeps exploration to at most one pool.

[5]See Anderson and Smith [2010] Assumption 2.

where $N_{i,t}$ denotes the total number of agents in pool i and $q_{i,t}$ denotes its fraction of high pizzazz agents.

At $t = 0$ when the platform only has impartial information about its agents, it cannot yield any discriminatory power over any pool. In other words, $q_{i,0}$ is expected to be p_0 in any matching pool. However, as matches realize and the platform learns the identity of the *proven members*, it can utilize the information to maximize output.

Within this framework, we fix m – the total number of pools – and analyze the platform's optimal allocation decision. In particular, for each time t and each pool i, we allow the platform to allocate $\alpha_{i,t}$ fraction of its *proven members* and $\beta_{i,t}$ fraction of its unknown agents to constitute $N_{i,t}$ to dynamically maximize output:

$$N_{i,t} = \alpha_{i,t} \cdot n_t + \beta_{i,t} \cdot (1 - n_t) \tag{3.3}$$

Formally, the platform in our model chooses an allocation schedule $(\{\alpha_{i,t}\}_{i \leqslant m}, \{\beta_{i,t}\}_{i \leqslant m})_{t \geqslant 0}$ to maximize total output, namely:

$$\mathcal{W}(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \sum_{i \leqslant m} \int_{t \geqslant 0} e^{-rt} \cdot \dot{M}_i(n(t), \alpha_i(t), \beta_i(t)) dt \tag{3.4}$$

where $\{\dot{M}_{i,t}\}_{i \leqslant m}$ and $\{N_{i,t}\}_{i \leqslant m}$ follow the allocation and the required law of motion: equation (3.2) and (3.3). The allocation choice will also dynamically determine the evolution of n_t.

Even with fixed number of pools, the continuous time nature of our model renders general optimal control analysis relatively complicated. Instead, we focus on the one-pool and the two-pool case, in which we are able to utilize a few key insights to fully characterize the platform's optimal allocation policy.

3.3 One Pool Control: Optimal Recruitment Problem

We now characterize the platform's optimal allocation policy in the one pool problem. We first observe that the "perfect good news" nature of our model dictates that the platform always recruits *proven members* into their one pool; that is, $\alpha_t \equiv 1$. This follows from the fact that a *proven member* both increases flow output $\dot{M}_t$ – having more high pizzazz agents increases expected number of high-high matches – *and* facilitates faster learning – a high type *candidate member* has higher chance of signaling through matching with a *proven member* than another *candidate member*. We can thus fix $\alpha_t \equiv 1$ and focus on the platform's optimal exploration policy β_t over the unknown agents.

The platform expects high pizzazz agents to signal their high type through matching at a certain rate. Therefore, over its *candidate members*, the platform grows more and more pessimistic. In particular, the platform updates its p_t over these agents of unknown pizzazz as a function of their time spent in the platform. Even with assumption 3.1, this means that for every agent, the platform needs to keep track of when and for how long they have been recruited. While characterizing a reputation vector over the continuum of agents as a function of the platform's recruitment policy is possible, analyzing the platform's incentive provides us insight that, in the optimal exploration policy, we need only to update p_t as a function of time.

To see why, first consider the platform's incentive to recruit one *candidate member* versus another. As we noted that the platform grows more pessimistic of its *candidate members* the more time they spent in the platform, if the platform has to choose certain agents over the others in constituting its β_t, it would be ones that have been recruited for less time. Bring this discrete choice concept to continuous time, and we have a platform that constantly shuffles through the unknown agents to constitute β_t; i.e. p_t and q_t both evolve at the same rate for all of unknown agents, and they only depend on $\{\beta_t\}_{t\geqslant 0}$.

So far, the platform's optimal exploration problem becomes:

$$\max_{\beta_t} \quad \int_{t\geqslant 0} e^{-rt} \cdot \lambda\mu \cdot \frac{\big(n_t + \beta_t(p_0 - n_t)\big)^2}{n_t + \beta_t(1 - n_t)} dt \tag{3.5}$$

$$\text{s.t.} \quad \dot{n}_t = \lambda\mu \cdot \beta_t(p_0 - n_t) \tag{3.6}$$

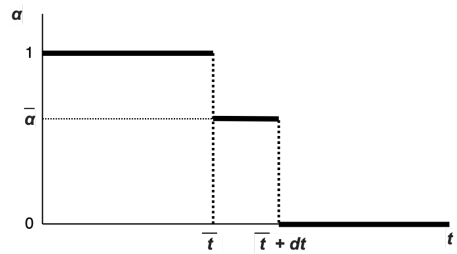
where we substituted in $N_t = 1 \cdot n_t + \beta_t \cdot (1 - n_t)$, and $q_t = \frac{n_t + \beta_t(p_0 - n_t)}{N_t}$. The law of motion for n_t – mass of *proven agents* – highlights the two factors that make up a discovery; the platform *discovers* a high type agent when 1) output realizes and 2) it involves at least one member not already identified as the high type. It also shows the two countervailing effects of recruiting from unknown agents: recruiting a *candidate member* hurts flow output $\dot{M}_t$, but facilitates discovery, which, in turn, helps future output.

With the updated platform problem, consider the platform's decision to recruit an arbitrary $\bar{\beta} < 1$ proportion of unknown agents for *dt* period of time. Since the platform constantly shuffles the unknown type agents to keep p_t the same for all, we can imagine an alternative exploration strategy of recruiting full 1 mass of unknown agents for shorter period of time to elicit the same learning. The law of motion given in (3.6) reveals that boosting $\bar{\beta}$ to 1 would facilitate learning exactly by proportion $\frac{1}{\bar{\beta}}$, thus requiring $\bar{\beta}dt$ time to elicit the same learning. Between the two exploration strategies, discounting implies that, regardless of p_t, it is always beneficial to front-load exploration. In other words, if the platform chooses to explore at all, the optimal policy mandates full exploration until full exploitation. We formalize the finding as theorem.

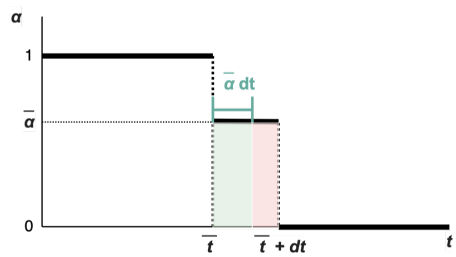
Theorem 3.1. *Optimal Exploration Problem becomes Optimal Stopping Problem.*

Proof. We prove this theorem in three steps.

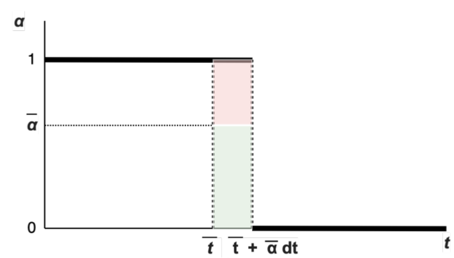
First, we formalize the shuffling argument. Suppose, by contradiction, that the platform chooses $\bar{\beta} < 1$ agents to explore for $[\bar{t}, \bar{t} + dt]$ interval, for arbitrary $\bar{\beta}$ and $\bar{t}$:



Among its *candidate members*, the platform is more pessimistic of those who have been in the platform the longest. Therefore, the platform can improve on exploring set $\bar{\beta}$ agents for dt by exploring one set of $\bar{\beta}$ agents for $\bar{\beta}dt$, and switching the exploration set to yet unexplored agents for the remaining $(1 - \bar{\beta})dt$ time:



Then, it is obvious that with time discounting, the platform can explore the two separate cohorts simultaneously and improve:



This improvement is always applicable to any $\bar{\beta}$ at any time $\bar{t}$, so the platform can iron any gradual exploration evolution into jumps.

Second, we check that the flow output does not suffer from this expedited learning. It is clear from evolution of n_t in equation (3.6) that faster learning benefits the entire output series that follow. In addition, we find that even within the instantaneous output, the optimal recruitment policy of unknown agents prescribes bang-bang: the platform either explores everyone or none at all.

We check this property by analyzing the second order derivative of instantaneous output. Differentiating the integrand of equation (3.5) twice gives:

$$\frac{\partial^2 I}{\partial \beta_t^2} = \frac{2n_t^2(1-p_0)^2}{(n+\beta_t(1-n_t))^3} > 0$$

which confirms that the platform's optimal recruitment policy for maximal instantaneous output is also convex.

Finally, we reiterate that faster learning benefits future output to conclude that the all-or-nothing exploration policy has be be all exploration in the beginning until no exploration and full exploitation stage. Therefore, the optimal exploration problem of (3.5) and (3.6) reduces to an optimal stopping problem of when to stop exploration. □

Theorem 3.1 reduces the optimal control problem of choosing an infinite series $\{\beta_t\}_{t\geqslant 0}$ to a univariate optimization problem of at which time τ the platform should halt exploration. Formally, the optimal stopping time problem is:

$$\max_{\tau} \quad \int_0^{\tau} e^{-rt}\cdot\lambda\mu\cdot p_0^2 dt + \int_{t\geqslant\tau} e^{-rt}\cdot\lambda\mu n_\tau dt \tag{3.7}$$

$$\text{s.t.} \quad n_t = p_0\big(1-e^{-2\lambda\mu p_0\cdot min\{t,\tau\}}\big) \tag{3.8}$$

The first integral of (3.7) refers to total output generated in the full exploration stage until τ. Since the platform keeps all agents – signaled and unsignaled – q_t is constant at p_0. The second term shows output in the exploitation stage. From τ onward, the platform only keeps *proven members*, and thus N_t is n_t and q_t is 1.

In words, the platform begins with full recruitment to maximally facilitate learning when it is most valuable. This point is trivial in the current model when the platform starts with 0 mass of *proven members*, but it also holds true when it begins with $n_0 > 0$ as long as it is small enough compared to the projected mass of high type agents p_0. We will formally lay out the optimal exploration policy in the next theorem, which would also specify the threshold n_0 that divides the platform's initial exploration policy from full exploration to no exploration. For now, we introduce a notation that will aid our description of the optimal stopping time. At each time t, we denote the discovery ratio – proportion of *proven members* within the total mass of high type agents – as $\rho_t := \frac{n_t}{p_0}$.

To formally characterize optimal stopping time τ is to solve a univariate maximization problem of (3.7) and (3.8). Combining the equations and taking the first order condition gives:

$$\frac{\partial}{\partial \tau} = \lambda \mu e^{-r\tau} \cdot \Big(\big(p_0 + \frac{\lambda \mu p_0}{r}\big) p_0 - \big(1 + \frac{\lambda \mu p_0}{r}\big) n_\tau \Big) \tag{FOC}$$

As n_t naturally evolves in a non-decreasing manner while the first term in the main parenthesis remains fixed, our maximization problem is convex and solution obtains when the two terms in the parenthesis equal. We can better understand the trade-off when we express it in terms of ρ_t:

$$\text{Value from exploration increases as long as } (1 + \frac{\lambda \mu p_0}{r}) \cdot (1 - \rho_t) > 1 - p_0 \tag{3.9}$$

Intuitively, when ρ_t is low, there are many high type agents left to explore, increasing the incentive to explore, and vice versa for high ρ_t. In particular, the optimal policy would mandate further exploration if the total value of possibly discovering a yet unknown high type agent (LHS) exceeds the expected damage of recruiting a low type agent (RHS). The added $\frac{\lambda \mu p_0}{r}$ represents an additional continuation value of a high type agent; she not only increases total number of matches, but might also contribute to output, at which point she becomes a *proven member* to stay in the platform forever after.

We encompass the results so far to characterize the platform's optimal allocation policy in the following theorem.

Theorem 3.2. *The platform's optimal allocation policy is given by:*

$$\alpha^* \equiv 1$$

$$\beta^*(\rho_t) = \begin{cases} 1 & \text{if } \rho_t \leqslant \rho^* \\ 0 & \text{if } \rho_t > \rho^* \end{cases}$$

where ρ^* solves:

$$(1 + \frac{\lambda \mu p_0}{r}) \cdot (1 - \rho^*) > 1 - p_0$$

3.4 Two Pool Control: Optimal Segmentation Problem

Thus far, in light of the platform's general allocation problem, we have limited the number of pools to one, with the interpretation that platform often decides how to optimally recruit agents with learning value. In the process, we have essentially reduced the optimal allocation problem of two types – the *proven members* and the unknown agents – into a single-armed decision of optimally recruiting *candidate members* (which we then reduced to an univariate optimization problem with respect to time.) While such simplicity of the one pool control problem enabled us to fully characterize the optimal allocation policy, we recognize that many matching platforms keep more than one pool to differentiate their users. For example, various dating apps utilize different versions of a ranking system whereby agents of a rank are matched with a partner of similar rank.

Does the platform keep every *proven member* in a premium pool to maximize matching between high type agents? Or does it keep pools of different mixtures? In this section, we study the next simplest version of our general allocation model – the two pool control problem – to answer these questions.

Formally, the platform in our two pool control model chooses an allocation schedule

$(\{\alpha_0, \alpha_1, \alpha_2\}_{t\geqslant 0}, \{\beta_0, \beta_1, \beta_2\}_{t\geqslant 0})$ to maximize total output, given by equation (3.4). Remember, $\alpha_{i,t}$ and $\beta_{i,t}$ refer to the proportion of *proven members* and unknown agents that the platform allocates to the ith pool at time t, where the subscript 0 refers to non-recruitment. For example, at arbitrary time $\hat{t}$, an allocation policy $(\{0, 1, 0\}, \{0, 0, 1\})$ would represent the following:

(i) full recruitment of *proven members* $(\alpha_{0,\hat{t}} = 0)$,

(ii) pool 1 purely populated by *proven members* $(\beta_{1,\hat{t}} = 0, \alpha_{1,\hat{t}} > 0)$

(iii) full mass of *candidate members* into pool 1 $(\alpha_{1,\hat{t}} = 1, \alpha_{2,\hat{t}} = 0)$,

(iv) full recruitment of unknown agents as *candidate members* $(\beta_{0,\hat{t}} = 0)$,

As in the one pool model, (i) is intuitive. A *proven member* contributes to both instantaneous output and learning by matching, and thus should be fully recruited at any point in time. On the other hand, the other features (ii) - (iv) of the given allocation policy may seem overly simplistic and for expository purpose only. In fact, however, this given allocation policy – populating one pool with the entire mass of unknown agents and constantly promoting *proven members* to the other "premium" pool – turns out to be one of the only two[6] possible policies in any optimal allocation schedule.

To illustrate why that is, we will examine the features one by one. Lemma 3.1 will establish that in any non-trivial two-pool allocation, it is optimal to keep one pool entirely consisting of *proven members*. This shows that the optimal allocation policy always prescribes feature (ii), and also reduces the dimension of the general allocation problem to choosing (x_t, y_t) of $(\{0, x_t, 1 - x_t\}, \{y_t, 0, 1 - y_t\})$. Theorem 3.3 then constructs the solution to the reduced problem, displaying the bang-bang nature of feature (iii) and (iv). As the

[6]Disregarding trivial isomorphic alternatives. With CRS technology, full exploration policy $(\{0, 1, 0\}, \{0, 1, 0\})$ can be represented by an infinite isomorphic copies of $(\{0, l, 1 - l\}, \{0, l, 1 - l\})$ $\forall l$.

optimal policy of the one-pool model mandated full exploration followed by full exploitation, the optimal two-pool policy prescribes full integration followed by full segmentation.

Lemma 3.1. *If* $q_{1,t} \neq q_{2,t}$, *then* $\max\{q_{1,t}, q_{2,t}\} = 1$. *In words, if both pools are utilized in a non-trivial manner, one of the pools purely consists of proven members.*

Proof. We already established full recruitment of *proven members* ($\alpha_0 \equiv 0$). Therefore, the following two cases exhaust possible outcomes: 1) entire n_t *proven members* are concentrated in one pool ($\boldsymbol{\alpha_t} = \{0, 1, 0\}$), or 2) they are dispersed into two pools ($\boldsymbol{\alpha_t} = \{0, x, 1-x\}$. In both cases, we show that it is optimal to keep one pool purely populated by *proven members.*

1) $\boldsymbol{\alpha_t} = \{0, 1, 0\}$ case.

   Without loss of generality, we call the pool that only houses *proven members* as pool 1. We show that, at any time t, regardless of $\hat{\beta}$:= (how many unknown agents are recruited), $\boldsymbol{\beta_t} = \{1-\hat{\beta}, 0, \hat{\beta}\}$ is the optimal allocation.

   Suppose, by contradiction, that for interval $[t, t+dt]$, $\boldsymbol{\beta_t} = \{1-\hat{\beta}, b, \hat{\beta}-b\}$ for arbitrary $b \in (0, \hat{\beta})$. Then, flow output generated in the interval for both pools as a function of b become:

$$dM_t^1(b) = \lambda\mu(n_t + b) \cdot \Big(\frac{n_t + b \cdot p_t}{n_t + b}\Big)^2 dt$$
$$dM_t^2(b) = \lambda\mu(\hat{\beta} - b) \cdot p_t^2 dt$$

   Differentiating with respect to b and combining the two to calculate the effect of unit mixture of an unknown agent into pool 1 gives:

$$\frac{\partial dM_t(b)}{\partial b} = -\lambda\mu(1-p_t)^2 dt < 0$$

   Therefore, regardless of the size of n_t, allocating any unknown agent into the premium pool unambiguously hurts output.

Likewise, total learning (i.e. newly discovered *proven members*) generated in the interval as a function of b is:

$$dn_t(b) = \lambda\Big(dM_t^1(b) \cdot \frac{b \cdot p_t}{n_t + b \cdot p_t} + dM_t^2(b)\Big)$$

$dM_t^1(b) > 0, dM_t^2(b) < 0$ and $dM_t^1(b) + dM_t^2(b) < 0$. Since $\frac{b \cdot p_t}{n_t + b \cdot p_t} < 1$, $dn_t(b)$ – the impact of mixing an unknown agent into pool 1 on total learning generated – is also negative.

Since adding b mass of unknown agents into pool 1 decreases both output and rate of learning, we have shown that when $\boldsymbol{\alpha_t} = \{0, 1, 0\}$, optimal policy prescribes $\beta_{2,t} = 0$.

2) $\boldsymbol{\alpha_t} = \{0, x, 1 - x\}$ case.

We go through the same steps as case 1)[7] and find that when $\boldsymbol{\alpha_t} = \{0, x, 1 - x\}$, optimal allocation of unknown agents becomes $\boldsymbol{\beta_t} = \{1 - \hat{\beta}, 0, \hat{\beta}\}$ or $\{1 - \hat{\beta}, \hat{\beta}, 0\}$ or $\{1 - \hat{\beta}, x \cdot \hat{\beta}, (1 - x) \cdot \hat{\beta}\}$. In the last case, the ratio of *proven members* to the unknown agents in both pools are the same – $q_{1,t} = q_{2,t}$. With CRS matching technology, this allocation is isomorphic to recruiting all agents into one-pool, and thus is a trivial use of the two pools.

Since case 1) and 2) exhaust all possibilities, we have proven that optimal allocation schedule always keeps one of the two pools purely consisting of *proven members*. □

Lemma 3.1 shows that at each point in time along the optimal allocation schedule, the platform keeps pool 1 entirely populated by *proven members*. In effect, pool 1 becomes the premium pool, with $q_{1,t} \equiv 1$, designed for full exploitation of the high type agents. Therefore, not unlike our approach in the one-pool model, we can focus on the platform's optimal exploration policy over pool 2.

[7] We detail the calculations in the appendix.

Formally, Lemma 3.1 reduces the set of possible optimal allocation policy to $(\{0, x_t, 1 - x_t\}, \{1 - y_t, 0, y_t\})$ for $x_t \in [0, 1], y_t \in [0, 1]$. We now show that, analogous to the bang-bang result of the one-pool model, the optimal allocation schedule prescribes full integration followed by full segmentation.

Theorem 3.3. *Optimal allocation schedule always prescribes full integration – effectively recruiting and pooling all agents into one pool*[8] *– followed by full segmentation.*

Formally, there exists n^* by which the platform's optimal allocation policy prescribes:

$$\boldsymbol{\beta_t^*} \equiv \{0, 0, 1\}$$

$$\boldsymbol{\alpha_t^*}(n_t) = \begin{cases} \{0, 0, 1\} & \textit{if } n_t \leqslant n^* \\ \{0, 1, 0\} & \textit{if } n_t > n^* \end{cases}$$

Proof. The reduced problem is to optimally control $x_t \in [0, 1]$ and $y_t \in [0, 1]$ to dynamically maximize positive match outcome. This is a canonical form of optimal control problem, with admissible control variables (x_t, y_t) jointly affecting the evolution of the state variable n_t.

Formally, the optimal control problem is:

$$\begin{aligned} \max_{\{x(\cdot), y(\cdot)\}} \quad & \int_0^\infty e^{-rt} \cdot \dot{M}(n(t), x(t), y(t)) dt \\ \text{s.t.} \quad & \dot{n}(t) = f(n(t), x(t), y(t)) \\ & n(0) = 0; \quad x(t) \in [0, 1]; \quad y(t) \in [0, 1] \end{aligned}$$

We solve the infinite horizon Hamilton–Jacobi–Bellman equation in the appendix. The solution prescribes $x_t \in \{0, 1\}$ for all values of y_t. Using insight from the one pool solution, we conclude that the platform front-loads learning $(x = 0)$ until the diminishing exploration value crosses the growing exploitation value – both due to growing n_t – and switch to full

[8] With CRS matching technology, we can always mimic one pool setting in two pool model by keeping q_t the same in both pools.

segmentation ($x = 1$). Since the platform keeps all recruited unknown type agents to pool 2, the decision to fully recruit them ($y_t = 1$) follows from the one pool solution. □

3.5 Conclusion

This paper integrates insights from literature on experimentation and random matching, and asks a new question that involves both: How do platforms utilize their matching power to strike balance between optimal experimentation and maximal flow output? We propose a general allocation model that covers both the one-pool recruitment problem and the two-pool segmentation problem, and find that solution to both resolves the exploration-exploitation tradeoff in a bang-bang manner.

In the one pool model, we found that the platform never wants to sample for assuring quality. The optimal exploration policy prescribes front-loading exploration by recruiting everyone in the market. And when the platform discovers sufficient mass of agents with assured quality, it drops agents of still unknown quality all at once. The finding provides two takeaways. First, the value function describing the balance between exploration and exploitation is convex even in the dynamic setting. Second, the threshold at which the platform's strategy switches from full exploration to full exploitation only depends on the discovery ratio.

That we found similar result for the two pool model is surprising. With additional ability to segment its users into two pools and keep different mixtures, one could imagine a more nuanced resolution of the exploration-exploitation tradeoff. Instead, the optimal allocation schedule prescribes that the platform keeps one pool for each extreme, due to the aforementioned convex nature of the value function. We think it would be fruitful research to extend our general model to more than two pools, and see whether the bang-bang pattern breaks.

While this paper provides some answers to the matching platform's optimal policy, we recognize the doors left open for further research. We assumed a constant-return-to-scale matching technology for technical reasons, but in application, there are scaling involved in many matching industries. We conjecture that introducing scale would further incentivize the platform to keep an integrated pool, but fully characterizing optimal policy under increasing return to scale technology would be interesting. Breaking the symmetry assumption and better characterizing the matching friction would also enrich the analysis.

3.6 Appendix

Addendum to Proof of Lemma 3.1. Here we lay out the proof that when $\boldsymbol{\alpha_t} = \{0, x, 1-x\}$, the optimal *candidate member* allocation becomes $\boldsymbol{\beta_t} = \{1-\hat{\beta}, 0, \hat{\beta}\}$ or $\{1-\hat{\beta}, \hat{\beta}, 0\}$ or $\{1-\hat{\beta}, x\cdot\hat{\beta}, (1-x)\cdot\hat{\beta}\}$

Suppose without loss of generality that $\boldsymbol{\beta_t} = \{1-\hat{\beta}, y\cdot\hat{\beta}, (1-y)\cdot\hat{\beta}\}$. Then, flow output generated in the interval $[t, t+dt]$ for both pools become:

$$dM_t^1(y) = \lambda\mu(x\cdot n_t + y\cdot\hat{\beta})\cdot q_{1,t}^2 dt$$
$$dM_t^2(y) = \lambda\mu(1 - x\cdot n_t - y\cdot\hat{\beta})\cdot q_{2,t}^2 dt$$

Combining the two, we examine the first-order and the second-order differentiation with respect to y:

$$\frac{\partial dM_t(y)}{\partial y} = q_{2,t}^2 - 2p_0 q_{2,t} + 2p_0 q_{1,t} - q_{1,t}^2$$
$$\frac{\partial^2 dM_t(y)}{\partial y^2} = \frac{2}{N_{2,t}}(p_0 - q_{2,t})^2 + \frac{2}{N_{1,t}}(p_0 - q_{1,t})^2 > 0$$

Therefore, regardless of $n_t, x, or\hat{\beta}$, full concentration of candidate members into either pool 1 or pool 2 maximizes flow output.

Likewise, learning (i.e. newly discovered *proven members*) generated in both pools as a function of y become:

$$dn_{1,t}(y) = dM_{1,t} \cdot \frac{p_{1,t} \cdot y \cdot \hat{\beta}}{x \cdot n_{1,t} + p_{1,t} \cdot y \cdot \hat{\beta}} = \lambda\mu q_{1,t} p_{1,t} y \cdot \hat{\beta} dt$$
$$dn_{2,t}(y) = dM_{2,t} \cdot \frac{p_{2,t} \cdot (1-y) \cdot \hat{\beta}}{(1-x) \cdot n_{2,t} + p_{2,t} \cdot (1-y) \cdot \hat{\beta}} = \lambda\mu q_{2,t} p_{2,t} (1-y) \cdot \hat{\beta} dt$$

Combining the two, we examine the first-order and the second-order differentiation with respect to y:

$$\frac{\partial dn_t(y)}{\partial y} = \frac{p_0(1-y) \cdot \hat{\beta}}{N_{2,t}}(q_{2,t} - p_0) - p_0 \cdot q_{2,t} - \frac{y \cdot \hat{\beta} p_0}{N_{1,t}}(q_{1,t} - p_0) + p_0 \cdot q_{1,t}$$
$$\frac{\partial^2 dn_t(y)}{\partial y^2} = \frac{2p_0}{N_{2,t}}(\frac{(1-y) \cdot \hat{\beta}}{N_{2,t}} - 1)(q_{2,t} - p_0) + \frac{2p_0}{N_{1,t}}(\frac{y \cdot \hat{\beta}}{N_{1,t}})(q_{1,t} - p_0) < 0$$

Solving the first order condition gives: $y = x$. Therefore, regardless of $n_t, x, or\hat{\beta}$, full integration of all members maximizes total learning.

This completes the calculations for the $\boldsymbol{\alpha_t} = \{0, x, 1-x\}$ case.

□

Optimal Control Problem in Theorem 3.3. Here we reiterate the optimal control problem, and characterize the optimal path of the control variables $\{x(t), y(t)\}$.

The optimal control problem is:

$$\max_{\{x(\cdot),y(\cdot)\}} \quad \int_0^\infty e^{-rt} \cdot \dot{M}(n(t), x(t), y(t))dt$$
$$\text{s.t.} \quad \dot{n}(t) = f(n(t), x(t), y(t))$$
$$n(0) = 0; \quad x(t) \in [0,1]; \quad y(t) \in [0,1]$$

We can write the the current-value Hamiltonian as:

$$\hat{H}\big(n(t), x(t), y(t), \phi(t)\big) = \dot{M}\big(n(t), x(t), y(t)\big) + \phi(t) \cdot f\big(n(t), x(t), y(t)\big)$$

According to Pontryagin's Maximum Principle, The optimal pair $\big(\hat{n}(t), \hat{x}(t)\big)$ satisfies the necessary conditions:

(I) $\hat{H}_x\big(n(t), x(t), y(t), \phi(t)\big) = 0 \quad \forall t \in \mathbb{R}_+$

(II) $\hat{H}_y\big(n(t), x(t), y(t), \phi(t)\big) = 0 \quad \forall t \in \mathbb{R}_+$

(II) $\hat{H}_n\big(n(t), x(t), y(t), \phi(t)\big) = r\phi(t) - \dot{\phi}(t) \quad \forall t \in \mathbb{R}_+$

Checking (I):

$$n - \frac{2n(p_0 - x \cdot n)}{1 - x \cdot n} + \frac{n(p_0 - x \cdot n)^2}{(1 - x \cdot n)^2} + \phi(t) \cdot \Big(- \frac{n(p_0 - n)}{1 - x \cdot n} + \frac{n(p_0 - n)(p_0 - x \cdot n)}{(1 - x \cdot n)^2} \Big) = 0$$
$$\Rightarrow \quad \phi(t) = \frac{1 - p_0}{p_0 - n(t)}$$

However, plugging the $\phi(t)$ into (III) yields:

$$\hat{H}_n\big(n(t), x(t), y(t), \phi(t)\big) < r\phi(t) - \dot{\phi}(t) \quad \text{regardless of n(t)}$$

which implies that the optimal path of the control variable x_t is always corner solution for all t. Since $x_t \in [0, 1]$, by insight from one-pool model that learning is the most valuable in the beginning, we get that regardless of $y(t)$, the optimal allocation policy of the *potential members* become:

$$\text{Optimal } \hat{x}(t) = \begin{cases} 0 & \text{if } n(t) < \nu \\ 1 & \text{if } n(t) \geqslant \nu \end{cases}$$

This completes the proof needed to establish that the optimal path prescribes $x_t \in \{0, 1\}$ for all values of y_t. $\square$

Bibliography

Dilip Abreu and Faruk Gul. Bargaining and reputation. *Econometrica*, 68(1):85–117, 2000. ISSN 1468-0262. doi: 10.1111/1468-0262.00094.

Emanuel Adler and Michael Barnett. *Security Communities*. Cambridge University Press, October 1998.

Axel Anderson and Lones Smith. Dynamic matching and evolving reputations. *The Review of Economic Studies*, 77(1):3–29, 2010.

Robert Axelrod. *The Evolution of Cooperation*. Basic Books, 1984. ISBN 978-0-465-02121-5.

Robert Axelrod and Robert O. Keohane. Achieving Cooperation under Anarchy: Strategies and Institutions. *World Politics*, 38(1):226–254, October 1985.

P. V. (Sundar) Balakrishnan, Charles Patton, and Phillip A. Lewis. Toward a Theory of Agenda Setting in Negotiations. *Journal of Consumer Research*, 19(4):637–654, March 1993.

A. Banerji. Sequencing strategically: wage negotiations under oligopoly. *International Journal of Industrial Organization*, 20(7):1037–1058, September 2002.

Jeffrey S. Banks. Equilibrium behavior in crisis bargaining games. *American Journal of Political Science*, 34(3):599–614, 1990.

Shahin Berenji. Sadat and the road to jerusalem: Bold gestures and risk acceptance in the search for peace. *International Security*, 45(1):127–163, 2020.

Ken Booth and Nicholas Wheeler. *The Security Dilemma: Fear, Cooperation, and Trust in World Politics*. Springer Nature, November 2007. ISBN 978-0-333-58745-4.

Gary Bornstein and Zohar Gilula. Between-Group Communication and Conflict Resolution in Assurance and Chicken Games. *Journal of Conflict Resolution*, 47(3):326–339, June 2003.

Juergen Bracht and Nick Feltovich. Whatever you say, your reputation precedes you: Observation and cheap talk in the trust game. *Journal of Public Economics*, 93(9):1036–1044, October 2009.

Sybille Reinke de Buitrago. What Role Can Trust Play? Conflicts and Trust-Building in East Asia. *Asian Politics & Policy*, pages 739–761, 2009.

Lutz-Alexander Busch and Ignatius J. Horstmann. Bargaining Frictions, Bargaining Procedures and Implied Costs in Multiple-Issue Bargaining. *Economica*, 64(256):669–680, 1997.

Lutz-Alexander Busch and Ignatius J. Horstmann. The game of negotiations: ordering issues and implementing agreements. *Games and Economic Behavior*, 41(2):169–191, November 2002.

Carlo Carraro, Carmen Marchiori, and Alessandra Sgobbi. *Advances In Negotiation Theory : Bargaining, Coalitions, And Fairness*. Policy Research Working Papers. The World Bank, June 2005.

K. Chatterjee. Strategic Choice of Sequencing of Negotiations : A Model with One-Sided Incomplete Information, 2005.

Jin-Hee Cho, Kevin Chan, and Sibel Adali. A Survey on Trust Modeling. *ACM Computing Surveys*, 48(2):28:1–28:40, October 2015.

Bruce Cronin. *Community Under Anarchy: Transnational Identity and the Evolution of Cooperation*. Columbia University Press, August 1999.

Amitai Etzioni. *The Hard Way to Peace: A New Strategy*. Collier Books, 1962.

Shaheen S. Fatima, Michael Wooldridge, and Nicholas R. Jennings. Multi-issue negotiation under time constraints. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, AAMAS '02, pages 143–150, New York, NY, USA, July 2002. Association for Computing Machinery. ISBN 978-1-58113-480-3.

James D Fearon. Rationalist explanations for war. *International Organization*, 49(3):379–414, 1995.

Chaim Fershtman. A note on multi-issue two-sided bargaining: bilateral procedures. *Games and Economic Behavior*, 30(2):216–227, 2000.

Mark Fey and Kristopher W. Ramsay. Uncertainty and incentives in crisis bargaining: Game-free analysis of international conflict. *American Journal of Political Science*, 55(1): 149–169, 2011.

Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, August 1991. ISBN 978-0-262-30376-7.

Francis Fukuyama. *Trust: Human Nature and the Reconstitution of Social Order*. Free Press, 1995. ISBN 978-0-684-82525-0.

Ingmar Geiger. A model of negotiation issue–based tactics in business-to-business sales negotiations. *Industrial Marketing Management*, 64:91–106, July 2017.

Alexander L George et al. *Forceful persuasion: Coercive diplomacy as an alternative to war*. US Institute of Peace Press, 1991.

Gary Goertz and Paul F. Diehl. Enduring Rivalries: Theoretical Constructs and Empirical Patterns. *International Studies Quarterly*, 37(2):147–171, June 1993. Publisher: Oxford Academic.

Gary Goertz and Paul F. Diehl. Taking "enduring" out of enduring rivalry: The rivalry approach to war and peace. *International Interactions*, 21(3):291–308, November 1995.

Matthew Grennan. Bargaining ability and competitive advantage: Empirical evidence from medical devices. *Management Science*, 60(12):3011–3025, December 2014. ISSN 0025-1909. Publisher: INFORMS.

Todd Hall and Keren Yarhi-Milo. The personal touch: Leaders' impressions, costly signaling, and assessments of sincerity in international affairs1. *International Studies Quarterly*, 56 (3):560–573, 2012a.

Todd Hall and Keren Yarhi-Milo. The Personal Touch: Leadersâ Impressions, Costly Signaling, and Assessments of Sincerity in International Affairs1: Leaderâs Impressions, Costly Signaling, and Assessments of Sincerity. *International Studies Quarterly*, 56(3):560–573, September 2012b.

Aaron M. Hoffman. A Conceptualization of Trust in International Relations. *European Journal of International Relations*, 8(3):375–401, September 2002.

Aaron M. Hoffman. *Building Trust: Overcoming Suspicion In International Conflict*. SUNY Press, January 2006. ISBN 978-0-7914-6635-3.

Younghwan In and Roberto Serrano. Agenda restrictions in multi-issue bargaining. *Journal of Economic Behavior & Organization*, 53(3):385–399, March 2004.

Roman Inderst. Multi-issue Bargaining with Endogenous Agenda. *Games and Economic Behavior*, 30(1):64–82, January 2000.

Robert Jervis. *Perception and Misperception in International Politics: New Edition*. Princeton University Press, May 2017. ISBN 978-1-4008-8511-4.

Robert L Jervis. The confrontation between iraq and the us: Implications for the theory and practice of deterrence. *European Journal of International Relations*, 9(2):315–337, 2003.

Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision Under Risk. In *Handbook of the Fundamentals of Financial Decision Making*, volume Volume

4 of *World Scientific Handbook in Financial Economics Series*, pages 99–127. WORLD SCIENTIFIC, June 2012. doi: 10.1142/9789814417358$_0$006.

James P. Klein, Gary Goertz, and Paul F. Diehl. The New Rivalry Dataset: Procedures and Patterns. *Journal of Peace Research*, 43(3):331–348, May 2006.

Silvana Krasteva and Huseyin Yildirim. Payoff uncertainty, bargaining power, and the strategic sequencing of bilateral negotiations. *The RAND Journal of Economics*, 43(3):514–536, 2012.

Andrew Kydd. Sheep in sheep's clothing: Why security seekers do not fight each other. *Security Studies*, 7(1):114–155, 1997a.

Andrew Kydd. Game Theory and the Spiral Model. *World Politics*, 49(3):371–400, April 1997b.

Andrew Kydd. Trust, Reassurance, and Cooperation. *International Organization*, 54(2):325–357, 2000.

Andrew Kydd. Trust Building, Trust Breaking: The Dilemma of NATO Enlargement. *International Organization*, 55(4):801–828, 2001.

Andrew Kydd. Which Side Are You On? Bias, Credibility, and Mediation. *American Journal of Political Science*, 47(4):597–611, 2003.

Andrew Kydd. *Trust and Mistrust in International Relations*. Princeton University Press, Princeton, NJ, 2005.

Andrew H. Kydd. When Can Mediators Build Trust? *The American Political Science Review*, 100(3):449–462, 2006.

Andrew H. Kydd. *Trust and Mistrust in International Relations*. Princeton University Press, August 2007. ISBN 978-0-691-13388-1.

Andrew H. Kydd. Learning together, growing apart: Global warming, energy policy and international trust. *Energy Policy*, 38(6):2675–2680, June 2010.

Kevin Lang and Robert W. Rosenthal. Bargaining Piecemeal Or All At Once? *The Economic Journal*, 111(473):526–540, 2001.

Deborah Welch Larson. Trust and Missed Opportunities in International Relations. *Political Psychology*, 18(3):701–734, 1997.

Deepak Malhotra and J. Keith Murnighan. The Effects of Contracts on Interpersonal Trust. *Administrative Science Quarterly*, 47(3):534–559, September 2002.

Zeev Maoz. *Networks of Nations: The Evolution, Structure, and Impact of International Networks, 1816–2001*. Cambridge University Press, December 2010. ISBN 978-1-139-49249-2.

Zeev Maoz and Bruce Russett. Normative and Structural Causes of Democratic Peace, 1946-1986. *The American Political Science Review*, 87(3):624–638, 1993.

James G. March and Johan P. Olsen. The Institutional Dynamics of International Political Orders. *International Organization*, 52(4):943–969, 1998.

Jonathan Mercer. Rationality and Psychology in International Politics. *International Organization*, 59(1):77–106, 2005.

Jennifer Mitzen. Ontological Security in World Politics: State Identity and the Security Dilemma. *European Journal of International Relations*, 12(3):341–370, September 2006.

Johannes Munster and Markus Reisinger. Sequencing bilateral negotiations with externalities. page 44, April 2018.

Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA, 1991.

John F. Nash. The bargaining problem. *Econometrica*, 18(2):155–162, 1950. ISSN 0012-9682. doi: 10.2307/1907266. Publisher: [Wiley, Econometric Society].

Charles E. Osgood. *An alternative to war or surrender.* An alternative to war or surrender. Univer. Illinois Press, Oxford, England, 1962.

Elinor Ostrom and Toh-Kyeong Ahn. The meaning of social capital and its link to collective action. *Handbook of social capital: The troika of sociology, political science and economics*, 17, 2009.

Robert Powell. Bargaining theory and international conflict. *Annual Review of Political Science*, 5(1):1–30, 2002.

Alexander Raskovich. Ordered bargaining. *International Journal of Industrial Organization*, 25(5):1126–1143, October 2007.

Brian C. Rathbun. It Takes All Types: Social Psychology, Trust, and the International Relations Paradigm in Our Minds. *International Theory*, 1:345, 2009.

Brian C. Rathbun. From vicious to virtuous circle: Moralistic trust, diffuse reciprocity, and the American security commitment to Europe. *European Journal of International Relations*, 18(2):323–344, June 2012.

Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109, 1982. ISSN 0012-9682. doi: 10.2307/1912531.

Jan Ruzicka and Vincent Charles Keating. Going global: Trust research and international relations. *Journal of Trust Research*, 5(1):8–26, January 2015.

Thomas C. Schelling. *The Strategy of Conflict.* Harvard University Press, 1980. ISBN 978-0-674-84031-7.

Maurice E. Schweitzer, John C. Hershey, and Eric T. Bradlow. Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, 101(1):1–19, September 2006.

James K. Sebenius. Negotiation arithmetic: adding and subtracting issues and parties. *International Organization*, 37(2):281–316, 1983.

G. Richard Shell. *Bargaining for Advantage: Negotiation Strategies for Reasonable People*. Penguin, 2006. ISBN 978-0-14-303697-5.

B. Solhaug, D. Elgesem, and K. Stolen. Why Trust is not Proportional to Risk. In *The Second International Conference on Availability, Reliability and Security (ARES'07)*, pages 11–18, April 2007.

Ingolf Ståhl. Bargaining as a decision process. a comment. *The Swedish Journal of Economics*, 75(4):406–409, 1973.

Janice Gross Stein. Reassurance in International Conflict Management. *Political Science Quarterly*, 106(3):431–451, 1991.

Leigh L. Thompson. *The Mind and Heart of the Negotiator*. Prentice Hall, 1998. ISBN 978-0-13-270950-7.

Marc Trachtenberg. The united states and the nato non-extension assurances of 1990: New light on an old problem? *International Security*, 45(3):162–203, 2020.

Robert F Trager. The diplomacy of war and peace. *Annual Review of Political Science*, 19: 205–228, 2016.

Michael Watkins. Building Momentum in Negotiations: Time-Related Costs and Action-Forcing Events. *Negotiation Journal*, 14(3):241–256, 1998.

Alexander Wendt. *Social Theory of International Politics*. Cambridge University Press, October 1999.

Nicholas J. Wheeler. "I Had Gone to Lahore With a Message of Goodwill But in Return We Got Kargil": The Promise and Perils of "Leaps of Trust" in India-Pakistan Relations. *India Review*, 9(3):319–344, August 2010.

Keren Yarhi-Milo, Joshua D. Kertzer, and Jonathan Renshon. Tying Hands, Sinking Costs, and Leader Attributes. *Journal of Conflict Resolution*, 62(10):2150–2179, November 2018.

Muhamet Yildiz. Bargaining without a common prior—an immediate agreement theorem. *Econometrica*, 71(3):793–811, 2003. ISSN 1468-0262. doi: 10.1111/1468-0262.00426.

Muhamet Yildiz. Waiting to persuade*. *The Quarterly Journal of Economics*, 119(1):223–248, February 2004. ISSN 0033-5533. doi: 10.1162/003355304772839579.