

Essays on Strategic Communication and Trust in Principal-Agent Models

Emilia Oljemark

Essays on Strategic Communication and Trust in Principal-Agent Models

Emilia Oljemark

Aalto University publication series
DOCTORAL DISSERTATIONS 196/2014

© Emilia Oljemark

ISBN 978-952-60-5987-7 (printed)

ISBN 978-952-60-5988-4 (pdf)

ISSN-L 1799-4934

ISSN 1799-4934 (printed)

ISSN 1799-4942 (pdf)

<http://urn.fi/URN:ISBN:978-952-60-5988-4>

Unigrafia Oy
Helsinki 2014

Finland

Publication orders (printed book):

Unigrafia Bookstore Helsinki

<http://kirjakauppa.unigrafia.fi/>

books@unigrafia.fi



Author

Emilia Oljemark

Name of the doctoral dissertation

Essays on Strategic Communication and Trust in Principal-Agent Models

Publisher School of Business

Unit Department of Economics

Series Aalto University publication series DOCTORAL DISSERTATIONS 196/2014

Field of research Economics

Date of the defence 22 December 2014

☐ **Monograph** ☐ **Article dissertation** ☒ **Essay dissertation**

Abstract

This dissertation consists of three essays and an introductory chapter on strategic communication and trust in a principal-agent framework. The first two essays study strategic communication between a sender and a receiver under the presence of incomplete information about the receiver's preferences.

In the first essay, the interaction is static, and the sender transmits information concerning a decision-relevant state variable to the receiver who takes an action that affects the welfare of both players. The essay compares communication under two different decision-making protocols and finds that under certain conditions communication is less prone to break down if the sender communicates to a single decision maker than if he communicates to an audience of decision makers who act jointly based on the preferences of the median voter.

In the second essay, the sender and the receiver interact repeatedly. The sender learns about the type of the receiver through time by observing her actions. Cooperation ends as soon as the sender learns that the receiver's preferences are not aligned with his. The sender's uncertainty about the type of the receiver gives rise to reputation building in that a receiver who knows that her preferences oppose those of the sender seeks to be perceived as a more congruent type so as to obtain more information. This reputation building is, however, profitable only if the first period is not important to the receiver. The essay finds that there exist parameter specifications of the model under which communication is more efficient ex ante, and thereby all players strictly better off, if the sender is uncertain about the importance of the first period to the receiver.

The third essay supports the findings of essay two by studying a repeated trust game between a principal and an agent. The principal is uncertain about the agent's trustworthiness but is able to learn about it through time by observing his actions. The essay compares players' ex-ante welfare across two versions of the model which differ in their informational assumptions concerning the agent's payoffs. There is shown to always exist a set of prior beliefs of the principal about the trustworthiness of the agent for which the ex-ante welfare of players is decreasing in the amount of information that the principal holds about the agent

Keywords cheap talk, trust, reputation, asymmetric information

ISBN (printed) 978-952-60-5987-7

ISBN (pdf) 978-952-60-5988-4

ISSN-L 1799-4934

ISSN (printed) 1799-4934

ISSN (pdf) 1799-4942

Location of publisher Helsinki

Location of printing Helsinki **Year** 2014

Pages 137

urn <http://urn.fi/URN:ISBN:978-952-60-5988-4>

Acknowledgements

I owe a thank you to a number of people without whom this dissertation either would not (yet) exist or would at least look a great deal poorer.

First and foremost, I want to express my gratitude to my supervisor, Professor Juuso Välimäki for having spent much of his scarce time in guiding me with my research in all of its stages. Our regular meetings have been vital for the progress of my work. I left from each meeting with a restored clarity in what it was that I was doing, though more often than not this clarity had gradually faded by the time of the next meeting. Working with Juuso has been both inspiring and educational.

I also want to thank my pre-examiners, Professors Sidartha Gordon and Kaisa Kotakorpi whose insightful comments on my manuscript are highly appreciated and will greatly improve my essays on their way to academic journals.

During the past five years I have been fortunate to benefit from a great research community, both under Aalto University as well as under the wider construction of HECER. Several people at the department and at the Economicum have contributed to my project by giving feedback and advice. I am grateful for Pauli Murto and Hannu Vartiainen for the many helpful comments on my research. Especially the job market paper (essay 2) greatly benefited from discussions with you. As to the academic job market and research in general, discussions with Manuel Bagues, Marko Terviö, Matti Liski, and Yao Pan were very helpful.

There are a great many fellow PhD students with whom I have had the pleasure to share an office and discuss research as well as studies. I want to thank Paula in particular for the many laughs and for sharing the joy and pain, mostly joy, of micro theory and teaching assistance. Further, I want to thank Nelli, Anna, Lassi, Tuomas, Milla, James, Robin, Maija, Francesca, Riikka, Ansku, and Eero for great discussions over

lunch, whether on economics or on other important topics in life.

The department has provided me with an office and a well-functioning working environment of which I am very grateful. My gratitude in this respect is directed to the great administrative staff whose help in sorting out various practical matters is acknowledged.

Financial support from the FDPE and Yrjö Jahnsson Foundation is also gratefully acknowledged.

While writing this dissertation has been rather labor intensive, to the extent that my office has often felt like an extension to home, company and support from family and friends outside of the office has been inessential for the completion of this project.

Above all, I want to thank my parents Raija and Martin for supporting all of my academic endeavors. While my landing in a business school may have come as a surprise, I bet my landing in the field of research was more anticipated. I hope to have received even a fraction of your perseverance and creativity. Second, I want to thank my siblings Inka, Silja, Otto, Linnea, Kaius, Fiina, and Viola who make family time a great deal of fun, and who never fail to remind me not to take myself, nor life in general, too seriously. Special thanks to members L, K, and F for having been the most convivial flatmates during the job-market year.

Among people who have been important in keeping me active outside of work, I want to thank in particular Teea, Kaisa, Anna and Riikka for continuing friendship, Elina for activating my inner scout at regular intervals, and Minna for urging me for an occasional run in the forest, with or without a map.

Last but not least, my heartfelt gratitude to Frank who has been the victim of my frustration when things at work did not work out the way I wanted. Your patience, flexibility, support and love throughout the project have been unparalleled.

Konstanz, November 2014

Emilia Oljemark

List of Publications

This dissertation consists of an introductory chapter and of the following three essays.

1. Oljemark, Emilia. *“Bias Uncertainty and the Limits to Information Transmission”*, unpublished.
2. Oljemark, Emilia. *“Can You Keep a Secret? Building Reciprocal Trust in Communication”*, unpublished.
3. Oljemark, Emilia. *“Reputation and the Value of Information in a Repeated Trust Game”*, unpublished.

Contents

Contents	5
List of Figures	7
1. Introduction	9
2. Bias Uncertainty and the Limits to Information Transmission	15
2.1 Introduction	15
2.2 The Model	20
2.2.1 Preferences	21
2.2.2 Strategies and the equilibrium concept	22
2.2.3 Partition structure	23
2.3 Solving for an equilibrium	25
2.3.1 Special case: $k = 1$	26
2.3.2 Centralized decision making	27
2.3.3 Decentralized decision making	30
2.4 Comparative statics	33
2.4.1 Alternative distributions	34
2.4.2 Logarithmic utility	35
2.4.3 Non-symmetry in the receiver's preferences	36
2.5 Discussion	38
2.6 Conclusion	39
2.7 Appendix: Proofs	40
3. Can You Keep a Secret? - Building Reciprocal Trust in Communication	45
3.1 Introduction	45
3.2 Related literature	49

3.3	The Model	52
3.3.1	Payoffs	53
3.3.2	Histories and strategies	53
3.3.3	Equilibrium concept	54
3.3.4	A game without reputation	56
3.4	Public regime equilibrium	58
3.5	Private regime equilibrium	61
3.6	Welfare	67
3.6.1	Sender's welfare	69
3.6.2	Receiver's welfare	75
3.7	Discussion	76
3.8	Conclusion	78
3.9	Appendix: Proofs	78
4.	Reputation and the Value of Information in a Trust Game	93
4.1	Introduction	93
4.2	Literature	97
4.3	Public information about M	99
4.3.1	The Model	100
4.3.2	Equilibrium	102
4.4	Uncertainty about M	105
4.5	Welfare	109
4.5.1	General results	110
4.5.2	Two examples	115
4.6	Discussion and ideas for further research	117
4.7	Conclusion	119
4.8	Appendix: Proofs	120

List of Figures

2.1	Limiting biases under protocols A and A'	25
2.2	Limiting biases under different protocols	33
2.3	Difference between the limiting bias between the decentral- ized and the centralized protocol	35
3.1	Public regime strategies	61
3.2	Equilibrium cutoff of the sender	67
3.3	Reputation and screening effects as functions of y when $G(\frac{1}{2}) = \frac{2}{3}$ and $\pi^1 = \frac{2}{5}$	72
3.4	Sender's ranking of the regimes	74
3.5	Reputation and screening effects as a function of y when $G(\frac{1}{2}) = \frac{1}{4}$ and $\pi^1 = 0.3$. Dashed line: $x \in (\frac{1}{2}, \hat{x}_{Pr}]$, dotted line: $x \leq \frac{1}{2}$	85
3.6	Ex-ante expected welfare, $X \sim Beta(2, 6)$, $Y \sim Beta(1, 1)$. . .	86
3.7	Ex-ante expected welfare, $X \sim Beta(2, 8)$, $Y \sim Beta(1, 6)$. . .	87
3.8	Ex-ante expected welfare, $X, Y \sim Uniform$ ratio distribution $[0, 1000]$	88
4.1	One-shot game	102
4.2	Reputation and screening effects, $r = \delta = 1$, $M \sim U[1, 4]$. . .	113
4.3	Ex-ante expected payoffs, $r = \delta = 1$, $M \sim U[1, 4]$	116

1. Introduction

Efficient functioning of institutions and organizations often relies on cooperation among their members. For example, when decision-making in firms requires the sharing of information between an expert and a decision maker, it is necessary that the expert offer correct information to back up decisions, and that the decision maker use the expert's information appropriately. Similarly, the institution of start-up financing does not function smoothly unless investors and entrepreneurs mutually agree to the generally established rules of financing. The expert chooses to not share information, and the investor to not lend her money unless they have sufficient trust in the decision maker and the entrepreneur, respectively, to comply to the implicit rules of the transaction. The importance of trust in the economy is recognized also by Kenneth Arrow when he writes,

“Trust and similar values, loyalty or truth telling... have real, practical, economic value; they increase the efficiency of the system, enable you to produce more goods.” Arrow (1974: 23)

The objective of this thesis is to deepen our knowledge about the impact of asymmetric information on trust and thereby on the efficiency of economic interactions. In the three essays that constitute the thesis, the focus is on interactions between two individuals, a principal and an agent. The overarching theme of the essays is incomplete information about the intrinsic motivation of the agent which is a commonplace dilemma in many organizations. The first essay considers a one-time interaction whereas the other two deal with a framework in which the principal and the agent interact repeatedly which makes the evolution of trust a central issue.

The main contribution of this dissertation research is to establish novel results regarding the interplay of asymmetric information and efficiency

in principal-agent relationships. The general conclusion from the essays is that incomplete information may be beneficial for efficiency, where efficiency is measured in terms of ex-ante expected welfare of the principal and the agent.

Cooperation between an agent and a principal is often ensured by formal contracts which are enforced by courts. For instance, if a firm refuses to pay its employees or investors, these can always seek for compensation through courts by suing the firm. Given that the firm cannot escape its liabilities, following the contract is always cheaper for it, provided that the contract is designed optimally. Thus, formal contracts greatly enhance trust in the economy.

It is, however, not always possible to write formal contracts. In these settings, the efficient functioning of organizations or other economic interactions can still be ensured if contracts are implicit or relational in that they rely on mutual cooperation over a period of time under the threat that punishments, such as a termination of the relationship, are imposed if the informal contract is violated. If punishments are carefully designed they persuade all parties to comply to the implicit contract.

A myriad of collegial and hierarchical relationships in organizations involve implicit contracts, such as the trading of favors between coworkers and unwritten understandings between bosses and subordinates about task assignment, promotion, and termination decisions. Even compensation, transfer pricing, and internal auditing, which seem highly formal processes at first sight, often cannot be understood without consideration of their associated informal agreements. Moreover, relational contracts are not limited to intra-firm interactions, but are present also in relationships between firms.

The essays of this thesis all focus on these relational settings in which formal contracts do not guide the behavior of agents. In such settings, trust is essential to reach cooperative outcomes. To study trust and reciprocity, the essays of this thesis apply game theoretical models that capture the strategic elements embedded in decisions to trust or reciprocate between two individuals.

The first two essays of the thesis consider the extraction of decision-relevant information from an expert. The second essay combines this with reputation building by the decision maker. Because the expert is privately informed about the value of a decision-relevant variable, and because the transaction is not contractible, the decision maker only obtains imper-

fectly accurate information from the expert as long as their preferences are not completely aligned. The essays formalize communication using the game-theoretic framework of Crawford and Sobel (1982). Both essays amend the seminal cheap-talk model of Crawford and Sobel by assuming that the expert is uncertain about the preferences of the decision maker.

In the first essay, the focus is on a one-shot interaction. The essay seeks to formally determine whether and in which situations an expert should communicate with one decision maker as opposed to an audience of decision makers who act jointly. Or, to emphasize the concern of the decision maker, when, if ever, should she consult an expert on behalf of a larger pool of decision makers, and when, if ever, should she consult the expert individually.

To analyze this tradeoff, I formalize a static cheap-talk game in which a sender is uncertain about the conflict of interest between him and a receiver. A decision is taken either centrally by a committee of ex-ante identical receivers, or decentralized to a single receiver. The centralized decision is determined by the median voter, whereas the decentralized decision is stochastic from the sender's point of view. The sender's preferences are non-symmetric in that downward deviations from his bliss point are more costly than upward deviations. The two decision-making protocols are compared in terms of the highest variance that they support in the receiver's preferences before communication breaks down.

For the model specification under study, I find that if the receiver's decision leans to the direction preferred by the sender, decentralized decision-making sustains a higher variance in the bias, but if the decision leans to the sender's least preferred direction, centralized decision-making sustains a higher variance in the bias. This result speaks for concealing the preferences of the decision maker in certain environments as it may enhance the efficiency of communication. The amount of uncertainty that the sender has regarding the receiver's preferences can be interpreted as the level of trust that the sender has in the receiver sharing the same preferences with him. The result says, effectively, that there are environments in which less trust is needed to sustain informative communication if a decision is decentralized to a single decision maker.

The second essay focuses on a repeated interaction between an expert and one decision maker, and it seeks to determine whether and under which conditions the social value of additional information about the decision maker is positive. The repeated nature of interaction allows for

the gradual building of trust, or a sudden elimination of it if cooperation breaks out. The concern of the sender to be able to participate in decision making in the future motivates him to take costly actions by transmitting truthful information at the risk of being deceived. Repeated play thus facilitates information sharing. At the same time, this concern for future engagements works to the other direction. The concern of the receiver to obtain truthful or accurate information in the future motivates her to take costly actions in the short term by following the supposedly truthful advice of the sender.

In the model, the sender learns about the possible bias of the receiver by observing her chosen actions in consecutive decisions. The decisions differ in their importance for the two players across periods. If the receiver's stakes in the first period are low, the biased type invests in reputation by mimicking the action of the unbiased type. I characterize the most informative communication equilibrium under two alternative scenarios; one in which the stakes are players' private information, and another in which they are commonly known. Although knowledge of the stakes is beneficial for sorting the receiver in the first period, it is shown to reduce all players' in a range of parameters specifications of the model. Hence, the essay points out that more trust can sometimes be sustained ex-ante if the sender knows less about the decision maker. This result questions the pursuit of complete transparency in organizations.

The third essay considers repeated transactions between a principal and an agent in a setting without communication. Instead, the principal must simply decide whether or not to trust an agent of unknown type at the risk that the agent exploits her trust by behaving against the benefits of the principal. Naturally, if the principal knew that the agent abuses her trust, she would not engage in a transaction with the agent in the first place. The model can be applied to analyze a variety of situations that arise for instance between an investor and an entrepreneur, between a firm and its employees or suppliers, or between a buyer and a seller.

In the essay, I analyze a twice-repeated trust game where an Investor chooses in each period whether to invest or not in a project carried out by an Entrepreneur. The Investor learns about the reliability of the Entrepreneur by observing whether he repays investments or not. If the project is valuable, an unreliable Entrepreneur has a reputational incentive to repay the first investment so as to be perceived as reliable and thereby obtain financing also in the second round. The paper analyzes

two versions of the model, which differ in their informational assumptions concerning the value of the Entrepreneur's project. There is shown to always exist parameter specifications of the model under which all players' ex-ante welfare is maximized if the Entrepreneur is privately informed about the value of his project.

References

- Arrow, K. (1974) *The Limits of Organization*. New York: W.W. Norton.
- Crawford, V., and J. Sobel (1982) "Strategic Information Transmission". *Econometrica* 50 (6), 1431-1451.

2. Bias Uncertainty and the Limits to Information Transmission

2.1 Introduction

Decisions in organizations frequently rely on information provided by experts. These decisions are either made by individuals acting alone, or by a group of individuals who act based on a joint decision. In firms, for instance, financial decisions are either the responsibility of individual managers, or they are addressed in the board of directors. Suppose that the decision is about the financing of a project, and it is done based on a report from a project manager (PM) who knows the project's quality. When there is any misalignment between the interests of the project manager and the decision maker, whether an individual executive, or the board of directors, information transmission is strategic, leading the project manager to use only messages that favor his interests. At the extreme, the interests of the players are so far apart that informative communication breaks down altogether. Given that informed decisions are a prerequisite for successful business, if the firm was to optimize its decision-making protocol in a way that it maximizes the flow of information from the project manager, when should decisions be taken by individuals and when should they be brought to the board room? This is the question that this paper tries to address theoretically.

Another context in which this question seems relevant relates to purchasing decisions of consumers. More specifically, consumers in a given market need information about the quality of a new product to back up their purchasing decisions. Should they listen to local sales representatives who care about single purchases, or should they base their decision on a national advertizing campaign which is designed for the average consumer in the market?

To return to the first example, suppose that the conflict of interest between the PM and the decision maker is common knowledge among the players. Communication-wise, the decision should be taken by the person having her preferences closest to those of the PM. This is the seminal message of Crawford and Sobel (1982). In any case, as long as preferences of the PM and any given decision maker are not perfectly congruent, the equilibrium structure of communication, as shown by Crawford and Sobel, consists of noisy messages that reveal the PM's private information only partially.

Suppose instead that the PM is uncertain about the conflict of interest between himself and the decision maker. This is not at all uncommon. Many issues within firms, let alone in politics, are so complex that that there is significant heterogeneity in views even within intra-firm groups or political parties. The PM in a firm may be uncertain whether and to which extent senior employees have more optimistic views about the future of business, which would lead them to be either generous or conservative in their financing decision in comparison to what the PM prefers.

This provides a more interesting setting in which to explore the differences in the decision making protocols. When the PM is uncertain about the decision maker's bias, which can take any value from a finite interval, the equilibrium communication strategy of the PM with an individual executive is a result from the PM maximizing his expected utility, that is, maximizing the weighted sum of his payoffs for all possible levels of the decision maker's bias. The equilibrium communication strategy with a board of directors, on the other hand, results from the PM maximizing his utility given the expected decision of the board. On the face of it, these two optimization problems of the PM seem different. However, if the PM's preferences feature certainty equivalence, these two problems are strategically identical at least as long as the board implements the mean of its members' preferences. Whenever the bias of the individual decision-maker is drawn from the same distribution than the biases of individual board members, it does not matter for communication whether a decision is taken by one executive or a group of them.

Hence the second departure from the analysis of Crawford and Sobel (1982). Namely, assume that the sender's preferences are non-symmetric around his bliss point. For every quality level of the project, the bliss point of the PM is the financing decision that would maximize his payoff. When choosing what to report to the decision maker, the PM tries to minimize

the distance between the receiver's decision and his bliss point. If his preferences are non-symmetric, it matters whether deviations from the bliss point are to the left or to the right. This assumption is not difficult to motivate: economic actors are often not indifferent whether they fall short of their targets or exceed them, as for example in wage negotiations. Or, if a new technology turns out to be bad, it is better to have under-invested in it than over-invested. And when launching a new product, a higher-than-expected demand may be preferred to a lower-than-expected, ignoring capacity constraints.

Under the presence of non-symmetric preferences and the sender's uncertainty about the conflict of interests, it matters for communication who makes decisions – a single receiver or a pool of receivers. This paper studies these two decision-making protocols which are henceforth referred to as a decentralized and a centralized protocol, respectively.¹ The protocols are compared in terms of how much variance they sustain in the receiver's preferences which are assumed to differ from those of the sender by a constant which can take any value from a commonly known interval. The non-symmetry assumption imposes costs in terms of analytical rigor, which is why I concentrate in studying the limits of communication. That is, for each decision-making protocol, I solve for the highest level of conflict in interests that still allows for the existence of an informative communication equilibrium. As an example, if the variance in the receiver's bias is so high that meaningful communication fails under the centralized protocol, the results of this paper help to determine whether the decentralized protocol could be more useful for the purposes of extracting information from the sender.

The analysis shows that at least under certain assumptions on the form of the sender's preferences and on the distribution of the receiver's preferences, the spread in the receiver's bias that the sender tolerates under the decentralized protocol is more than a mean-preserving spread of the centralized protocol. In fact, when the sender prefers an over-sized budget to an under-sized one, and receivers have an upward bias, decentralized communication supports more uncertainty in a single receiver's action than centralized communication. This suggests that information transmission may improve if the board of directors could decentralize its decision to a randomly picked board member before the game begins. By

¹This terminology should not, however, be mixed with that of Dessein (2002) and related papers.

contrast, when receivers are downward biased, decentralized communication is more prone to babbling than centralized communication. In that case, decentralizing the decision would not improve communication.

More generally, the results show that the sender's uncertainty benefits those receivers who under perfect public information about the receiver's preferences would not obtain information. At the same time, it yields welfare losses to those receiver types who would obtain more precise information if the sender knew their preferences. Results obtained in the existing literature on the disclosure of biases² in cheap-talk games would suggest that non-disclosure of biases would often be ex-ante welfare enhancing if the receiver has an upward bias. A more formalized analysis of welfare is however left for future research.

Related literature

The model of this paper is closest to the seminal model of cheap talk by Crawford and Sobel (1982), hereafter referred to as CS. In their model, an informed sender transmits information about a state of the world to an uninformed receiver after which the receiver takes an action that affects the payoffs of both players. However, given a state of the world, the sender and the receiver have divergent preferences over the appropriate action. The authors show that for modest misalignment in the players' preferences, meaningful, albeit incomplete, communication can occur. An equilibrium consists of partitioning the state space into a finite number of intervals such that the sender only reports the element of a partition in which the true state lies. The receiver then chooses the action that maximizes her expected utility over the reported interval of states. CS further show that the model exhibits multiple equilibria. For a small degree of conflict in interests, there exist several alternative equilibrium partitions of the state space, and at the interim stage the sender chooses the one that maximizes his utility given the receiver's action rule. The number of equilibria decreases in the degree of conflict, and when the players' preferences are sufficiently divergent only one equilibrium remains. At that point, communication loses all information value and becomes practically useless. This babbling equilibrium is always part of the set of equilibria of the game.

This paper departs from CS in that it assumes that the sender is uncertain about the divergence in preferences between him and the receiver. In

²see in particular Li & Madarász (2008)

addition, where applications of the CS predominantly involve the use of the uniform-quadratic framework for the state space and the players' preferences, this paper applies the model of CS to non-symmetric preferences. Seidmann (1990) has considered cheap talk under uncertainty about the receiver's action, and constructed simple examples in which communication may be fully revealing despite conflicting interests. His examples, however, deal either with cases in which the receiver's actions are non-scalar, or consider a limiting case of CS in which all sender types share a common preference ordering over each of the receiver's actions. Moreover, his focus is not in comparing the equilibrium that he constructs to an equilibrium in which the receiver's action would be common knowledge.

In the subsequent literature, models with uncertainty about the sender's type are abundant. Morgan & Stocken (2003) study stock recommendations issued by a financial analyst whose incentives are uncertain to investors. With some probability, the analyst is solely concerned about inducing a stock price that is equal to the firm's true value. With a complementary probability, the analyst is additionally motivated by the benefit associated with inflating the stock price above its true value. Wolinsky (2003) analyzes a static communication game in which a decision maker is uncertain as to whether the sender is of a type that wants to minimize the magnitude of her action, or maximize it. Sobel (1985), Bénabou & Laroque (1992), and Morris (2001) analyze dynamic cheap talk where the focus is on the reputation building of the sender. All of the aforementioned papers are, however, substantially different from the current analysis in that they deal with one-sided asymmetric information or a dynamic setting.

The main question of this paper, whether more information is disclosed with one decision maker than with an audience whose decision rule is known, is essentially a question whether to disclose or not the bias of the receiver. This relates the paper to models of Li & Madarász (2008) and Rantakari (2014) who study the welfare effects of disclosing the sender's bias. Since the sender's and the receiver's biases are simply mirror images of each other, it is not surprising that the findings of Li & Madarász and Rantakari are in line with the finding of this paper in that uncertainty about biases is sometimes beneficial for communication.

Two-sided asymmetric information in cheap talk games has been studied at least by Chen (2009, 2012), Watson (1999), Lai (2013), Moreno de Barreda (2010), and Ishida & Shimizu (2012). Their papers study how

communication is affected if the receiver is partially informed about the state of the world but her preferences are common knowledge. Therefore, their focus differs from mine.

The seminal contribution to cheap talk between one sender and many receivers is Farrell and Gibbons (1989) who consider two receivers with known, distinct preferences. They analyze how welfare of players is affected by whether the sender reports his information privately or publicly. In their model, however, receivers always act as individuals. Hence, they essentially compare communication between a sender and one audience and a sender and two audiences, whereas in my paper there is always essentially just one audience. In their paper, differences between private and public communication arise because in public communication, the presence of another audience may either discipline or subvert the sender's relationship with the other audience. Hence, while communication may be credible with one receiver alone, it may unravel in the presence of another receiver, and vice versa. More recently, Goltsman and Pavlov (2011) have generalized the findings of Farrell and Gibbons in an environment with a continuum of states and actions.

When analyzing communication between a sender and a group of receivers, I abstract away from the group's internal dynamics. The aggregation of preferences in a group is not given particular attention as such. For literature on committee decision making, including communication and voting, see for example Gerardi & Yariv (2004), Krehbiel (2004), Austen-Smith & Feddersen (2009, 2006), Austen-Smith & Banks (1996), Feddersen & Pesendorfer (1997), and Li et al. (2001) among others.

The remainder of the paper is organized as follows. Section 2.2 introduces the model and points out the relevant departures from Crawford & Sobel (1982). Section 2.3 analyzes the model and establishes results under two decision-making protocols. Section 2.4 discusses some comparative statics. Discussion and conclusions are taken up in sections 2.5 and 2.6, respectively.

2.2 The Model

There are two players, a sender (S , he), and a receiver (R , she), and both have some private information. The sender, for example a project manager, needs financing for a project, say, launching a new product. He requests a budget by making a claim of the project's quality to the receiver,

such as the CEO or the board of directors, who allocates the budget. At the beginning of the game, Nature determines the quality of the project (i.e. sender's type) by drawing θ from $\Theta := [0, 1]$. The quality θ is distributed according to a uniform distribution F with density $f(\theta) > 0$ for all $\theta \in \Theta$. The sender observes θ privately without noise, and this information is neither contractible nor verifiable.

In addition, Nature draws the receiver's type, b , which measures the divergence in the preferences of the sender and the receiver, and is henceforth referred to as the receiver's bias³. Only the receiver observes b . The parameter b is drawn from a uniform distribution G with a support on $B \subset \mathbb{R}$. In the analysis later on, we consider either $B_+ := [0, \bar{b}]$, or $B_- := [-\bar{b}, 0]$. Hence, the receiver's bias is either positive or negative, and the direction is common knowledge.

After the move of Nature, the sender communicates some or all of his information about θ by sending a costless message to the receiver. Denote the message by m and let $m \in M := [0, 1]$. The message should thus be interpreted as a report about the quality which directly translates into an appropriate budget size. After observing m , the receiver takes an action by allocating a budget to the project. Assume that there is some commonly known linear mapping from the quality of the project to an appropriate budget. The action of the receiver is then essentially the choice of a quality level.⁴ Denote the receiver's action by $y \in \mathbb{R}_+$. After the receiver's decision, payoffs are realized. These are discussed in the next subsection. Finally, all aspects of the game other than b and θ are common knowledge.

2.2.1 Preferences

Both players' payoff depends on the difference between y and θ . Denote the utility of the sender by $U^S(y, \theta)$ and that of the receiver by $U^R(y, \theta, b)$. Both functions fulfill the sorting condition $U_{12}^i(y, \theta) > 0$. This condition ensures that the preferred budget for both players is strictly increasing in the quality of the project.

The sender would like to obtain a budget which accurately reflects the true quality of the project. Were the receiver's decision to differ from the

³This vocabulary is due to our focus on the sender's uncertainty. Clearly, from the receiver's point of view, the sender is biased. For later use, notice that a positive bias of the receiver is equivalent to a negative bias of the sender, and vice versa.

⁴So, if the bias of the receiver is to always allocate a thousand euros more to the project than what the sender finds optimal, if the budget function is linear in θ , and $y^S(\theta)$ is the sender's bliss point, then the receiver's action is essentially to choose a $\theta + b$ such that $y(\theta + b) = y^S(\theta) + 1000$.

true quality, then a too large budget is preferred to a too small budget. Suppose, for instance, that the costs involved with compromising on the project's goals, which may have repercussions on the project manager's reputation in the firm, are larger than the psychological costs associated with running an overly financed project. To reflect this non-symmetry in the sender's preferences, let

$$U^S(y, \theta) = \begin{cases} -k(y - \theta)^2 & , y \leq \theta \\ -(y - \theta)^2 & , y > \theta, \end{cases} \quad (2.1)$$

where $k > 1$ measures the degree of non-symmetry.⁵ The piecewise-defined utility function reflects the twofold objectives of the sender: primarily to induce the receiver to take the action $y = \theta$, and secondarily to avoid the case where $y < \theta$.

The receiver, for her part, only cares about allocating the appropriate budget to the project; whether deviations from her bliss point $y^R(\theta) = \theta + b$ are to the right or left is of no importance. Therefore, let

$$U^R(y, \theta, b) = -[y - (\theta + b)]^2,$$

where $b \in B$.

A positive bias of the receiver could be explained for instance by optimism about the prospects of the firm⁶ or by additional, private information about the future plans of the firm which bears on her evaluations of project importance today. Similar arguments can be constructed to motivate a negative bias.

2.2.2 Strategies and the equilibrium concept

A pure strategy for the sender is given by a signaling rule $\mu : \Theta \rightarrow M$, such that, for any realized state θ , $\mu(\theta)$ specifies an element m from M . A pure strategy for the receiver is given by an action rule $y : M \times B \rightarrow \mathbb{R}_+$ such that, given any message m , $y(m, b)$ specifies the budget, through the choice of a quality level.

The sender's report induces the receiver to update her prior belief about the project's quality. Let $F(\theta \mid m)$ be the distribution of the receiver's

⁵Reversed preferences, favoring a too small budget over a too large one, would be captured by setting $0 < k < 1$ or by setting $k' = 1/k$. This would simply change the direction of results.

⁶There is an abundant literature that has shown that management optimism is prevalent in the corporate world (see f.ex. Heaton, 2002 for references to other studies).

posterior beliefs about the project's quality upon observing message m .

The equilibrium concept we adopt is the Perfect Bayesian Equilibrium (PBE) which consists of a strategy profile (μ, y) and a system of beliefs $F(\theta | m)$ such that

1. For all $m \in M$, and $b \in B$, the receiver's action

$$y(m, b) \in \arg \max_y \int_{\theta \in \Theta} U^R(y, \theta, b) F(\theta | m) d\theta,$$

2. Given the receiver's beliefs, for all $\theta \in \Theta$, $\mu(\theta)$ maximizes the sender's expected payoff: if m^* is assigned by μ , then

$$m^* \in \arg \max_m \int_0^{\bar{b}} U^S(y(m, b), \theta) dG(b),$$

where $G(b)$ is degenerate under the centralized decision-making protocol.

3. The receiver's posterior belief $F(\theta | m)$ is formed using Bayes' rule whenever possible;

The definition of an equilibrium is otherwise as in CS except for the fact that the sender now maximizes his expected utility. A more detailed discussion of the equilibrium can be found in CS (p.1435).

2.2.3 Partition structure

As shown by CS, cheap-talk games have multiple equilibria, the number of which depends on the degree of conflict in players' preferences. In this paper, the parameter of interest is \bar{b} , which defines the support of the distribution for the receiver's bias. As long as $\bar{b} \neq 0$, it follows from the analysis of CS that all equilibria of the game are semi-separating, characterized by a partition structure whereby close-by sender types pool together and send the same message. The density of the largest partition equilibrium is decreasing in \bar{b} , and the least informative equilibrium, in which all sender types pool together and send the same message regardless of the state, is the only one that remains as soon as \bar{b} exceeds an endogenously determined threshold. In this babbling equilibrium, the sender's message is uncorrelated with the true state of the world, and therefore the receiver chooses her action based on her prior belief about θ .

Let $a(N) := \{a_i(N)\}_{i=0}^N$ be the set of $N + 1$ boundary points that char-

acterize a partition equilibrium of size N , where $0 = a_0(N) < a_1(N) < \dots < a_N(N) = 1$. Let m_i be the message that sender types in the interval $[a_{i-1}(N), a_i(N))$, $i \in \{1, \dots, N\}$, send in this equilibrium.⁷ Furthermore, let $N(\bar{b}, k)$ denote the size of the largest equilibrium of the game, that is, the number of intervals in the finest partition equilibrium. Thus, $N(\bar{b}, k)$ takes only integer values. As discussed in CS (Lemma 6), $N(\bar{b}, k)$ is a non-increasing step function of \bar{b} .⁸ For any pair (\bar{b}, k) , the set of equilibria of the game consists thus of partitions $\{a(1), a(2), \dots, a(N(\bar{b}, k))\}$.

The rest of the paper focuses on sets of equilibria which consist of two elements, that is, $N(\bar{b}, k) = 2$. As long as $N(\bar{b}, k) \geq 2$, the game has an informative equilibrium.

Definition 1. *An equilibrium is informative if the sender's message m in equilibrium induces the receiver to revise her prior beliefs. That is, if $F(\theta | m) \neq F(\theta)$.*

If $N(\bar{b}, k) = 1$, the unique equilibrium consists of babbling. Given some $k > 1$, denote by $| \bar{b}_1^k |$ the set of \bar{b} 's, either all positive or all negative, for which $N(\bar{b}, k) = 1$, and let $| \bar{b}_1^k |$ denote the smallest element in the set, that is, $| \bar{b}_1^k | = \inf(| \bar{b}_1^k |)$.⁹ Hence, $| \bar{b}_1^k |$ constitutes a threshold such that for all biases equal or larger than this, the unique PBE of the game consists of babbling. For all $| \bar{b} | < | \bar{b}_1^k |$, the game possesses in addition at least one informative equilibrium.

Definition 2. *$| \bar{b}_1^k |$ is called a 'limiting bias' if*

$$N(\bar{b}, k) = \begin{cases} 1 & \text{if } | \bar{b} | \geq | \bar{b}_1^k | \\ \geq 2 & \text{if } | \bar{b} | < | \bar{b}_1^k | . \end{cases}$$

In addition, \bar{b}_{1+}^k is the limiting bias for a positive support, B_+ , and $| \bar{b}_{1-}^k |$ is the limiting bias for a negative support, B_- .

Figure 2.1 illustrates the notion of a limiting bias for two arbitrary decision-making protocols. For some $k > 1$, $N(\bar{b}, k)$ defines the maximum number

⁷Since the boundary types are indifferent between consecutive actions, and since their probability of occurrence is zero, it does not matter which interval they belong to. Note, however, that the last interval, for $i = N$, should be closed.

⁸In light of the results of the paper, $N(\bar{b}, k)$ is nondecreasing in k if $b > 0$ and nonincreasing in k if $b < 0$.

⁹Absolute value is introduced to account for negative biases.

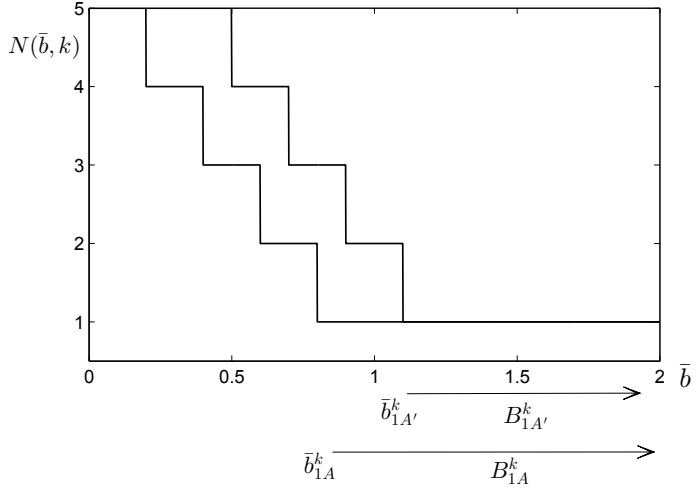


Figure 2.1. Limiting biases under protocols A and A'

of equilibria as a function of \bar{b} for protocols A and A' . The notation used should be read bearing in mind that the sender is uncertain about the actual bias, b . Thus, what is meant by saying that protocol A' has a higher limiting bias than protocol A is that the largest support of $G(b)$ that still allows for an informative equilibrium is larger under protocol A' than A . The protocol A' sustains informative communication for all distributions G with a support on $[0, \bar{b}]$ with $\bar{b} < \bar{b}_{1A'}^k$ whereas protocol A sustains informative communication only when $\bar{b} < \bar{b}_{1A}^k < \bar{b}_{1A'}^k$. Clearly then, for all $\bar{b}_{1A}^k \leq b < \bar{b}_{1A'}^k$, protocol A' is strictly superior in terms of ex-ante welfare because it allows for more precise communication which, as shown by CS, benefits all players ex ante.

2.3 Solving for an equilibrium

In any partition equilibrium of the game, the receiver's posterior belief upon observing a message m_i sent by sender types $\theta \in [a_{i-1}, a_i]$ is given by

$$F(\theta | m_i) = \frac{\mu(m_i | \theta) \cdot f(\theta)}{F(m_i)} = \frac{1}{F(a_i) - F(a_{i-1})} = \frac{1}{a_i - a_{i-1}}.$$

Hence, in equilibrium, upon receiving a message m_i , a receiver of type b chooses her action according to

$$y(m_i, b) \in \arg \max_y \int_{a_{i-1}}^{a_i} -[y - (\theta + b)]^2 dF(\theta | m_i),$$

which yields an optimal action of

$$y(m_i, b) = \frac{a_{i-1} + a_i}{2} + b.$$

Following Theorem 1 of CS, in any equilibrium of size N , with $1 \leq N \leq N(\bar{b}, k)$, the equilibrium partition is determined by an arbitrage condition which requires that a sender of boundary type $a_i(N)$, $i = 1, \dots, N - 1$, is indifferent between inducing actions $\mathbb{E}_b y(m_i, b)$ and $\mathbb{E}_b y(m_{i+1}, b)$.¹⁰ The expected action of the receiver depends on the decision-making protocol.

As discussed in the introduction, two decision-making protocols are considered. The budgeting decision is made either centrally by a committee of receivers, such as the firm's board of directors, or the decision is decentralized and made by a single receiver, such as the CEO or some board member. The decision-making protocol is known to all players before the game begins, and the receiver is committed to it. There is hence no possibility to change the protocol at the decision-making stage. The two protocols and their limiting biases are analyzed separately in the next two sections.

Without further knowledge of the functional form of $N(\bar{b}, k)$,¹¹ we are unable to tell how the number of equilibria compare for biases smaller than \bar{b}_1^k . Given some k , $N(\bar{b}, k)$ being higher under one protocol than in another for all \bar{b} would imply that the protocol ex-ante dominates the other. This is because it allows for more precise communication, which is considered jointly beneficial ex ante, that is, before the sender observes the state¹². Information about the limiting biases of different decision-making protocols is valuable, however, in determining whether another protocol could be more useful in extracting information if the support of $G(b)$ is too wide to sustain informative communication under one of them.

2.3.1 Special case: $k = 1$

To see why the assumption of the non-symmetry in the sender's preferences is necessary for the purposes of the paper, assume for a moment that $k = 1$ and the sender's preferences are given by the standard quadratic loss function used widely in applications of cheap talk. Because of cer-

¹⁰Sender types $a_0(N)$ and $a_N(N)$ only consider sending one message, m_1 and m_N , respectively.

¹¹Given the non-symmetry in the sender's preferences, an analytical expression of $N(\bar{b}, k)$ as a function of \bar{b} and k turns out to be too demanding for the purposes of the paper.

¹²See p. 1441 in CS.

tainty equivalence,

$$\max_m \mathbb{E}_b [U^S(y(m, b), \theta)] = \max_m U^S [\mathbb{E}_b(y(m, b), \theta)],$$

as the first-order conditions of the two maximization problems are equivalent. Certainty equivalence is in many ways a convenient property but it suits poorly the current context since it makes the comparison of decision-making protocols redundant. With quadratic utility, the sender's problem when faced with a single decision maker,

$$\max_m \int_{b \in B} U^S(y(m, b), \theta) dG(b)$$

is strategically equivalent to her optimization problem when facing a committee of receivers,

$$\max_m U^S(\mathbb{E}_b y(m, b), \theta),$$

where

$$\mathbb{E}_b y(m, b) = \int_{b \in B} y(m, b) dG(b) = \int_{b \in B} y(m) + b dG(b) = y(m) + \int_{b \in B} b dG(b),$$

and $dG(b) = \frac{1}{b}$. For $k > 1$, the certainty equivalence does not hold and centralized and decentralized decision making lead to distinct equilibria.

2.3.2 Centralized decision making

Assume that the decision is made centrally by a large pool of receivers that we henceforth refer to as a committee. For large firms, this could for example be the board of directors. In other applications, the committee can be thought of as the consumers or voters in a market, and the sender's payoff depends on the average reaction of the market. The committee thus consists of a large number of receivers whose types are i.i.d. Let the decision of the committee be governed by the preferences of the median voter. This is mainly to focus attention on the communication between the committee and the sender, and not on the deliberation within the committee. Moreover, all committee members are assumed to have an equal weight in the decision.¹³

Since receivers' types are drawn from a uniform distribution, the Law

¹³If some members have more weight than others the mere distribution of preferences is not enough to predict the committee's decision.

of Large Numbers (LLN) predicts that the decision rule of the committee upon observing message m_i is given as

$$y(m_i, \bar{b}) = \frac{a_{i-1} + a_i}{2} + \int_0^{\bar{b}} b \cdot \frac{1}{b} db = \frac{1}{2} [(a_{i-1} + a_i) + \bar{b}].$$

Because the action rule of the committee is deterministic, the game corresponds to a standard CS game with a commonly known bias equal to $\frac{1}{2}\bar{b}$. For any $b > 0$, an equilibrium of size two, $a(2)$, in which message m_1 is sent whenever $\theta \in [0, a)$, and m_2 is sent whenever $\theta \in [a, 1]$, is then characterized by the following arbitrage condition faced by a sender of type $a \in (0, 1)$.

$$-k \left(\frac{a}{2} + \frac{\bar{b}}{2} - a \right)^2 = - \left(\frac{1+a}{2} + \frac{\bar{b}}{2} - a \right)^2. \quad (2.2)$$

Notice that for a positive bias the existence of an equilibrium requires that $\frac{a}{2} + \frac{\bar{b}}{2} \leq a$, (i.e. $\bar{b} \leq a$), because otherwise the arbitrage condition in (2.2) would never hold. By Lemma 4 of CS, the limiting bias is obtained when $a = 1$.¹⁴ Thus, it must be that $\bar{b}_{1+}^k \leq 1$. Analogously, if the receiver's bias is negative, the equilibrium requires that $\frac{1+a}{2} - \frac{\bar{b}}{2} \geq a$, and at the limit, $a = 0$. Hence, also the negative bias must be less than one in absolute value.

To obtain the positive limiting bias, set $a = 1$ in (2.2) and solve for \bar{b} . The resulting polynomial has two roots. Since both are strictly positive for all $k > 1$, the limiting bias is given by the smaller of them.

Lemma 1. *If the decision is made centrally by a committee of receivers such that the LLN applies, and $G(b)$ is uniform, the positive limiting bias is given by*

$$\bar{b}_{1+}^k = \frac{k - \sqrt{k}}{k - 1}, \quad \forall k > 1.$$

The limiting bias is increasing in k , but since $\bar{b}_{1+}^k \leq 1$, the committee's bias

¹⁴In a game with a negative bias of the receiver (i.e. positive bias of the sender), their lemma 4 establishes that for two partition equilibria of the same size, $a(N, b)$ and $a(N, b')$, with $b' > b$, since the last threshold in both is necessarily at 1 (i.e. $a_N(N, b) = a_N(N, b') = 1$), and since partition intervals are longer in the equilibrium with a higher bias (as $4b' > 4b$), it must be that $a_i(N, b') < a_i(N, b)$ for all $i = 1, \dots, N-1$. Thus, for the negative limiting bias of the current paper, $a_1(2, \bar{b}_{1-}^k) = 0$. Analogously, for a positive bias of the receiver (i.e. negative bias of the sender), the limiting bias must satisfy $a_1(2, \bar{b}_{1+}^k) = 1$.

$\frac{\bar{b}}{2}$ must never exceed $\frac{1}{2}$. Given our assumptions about the sender's preferences, a positive bias of the receiver is henceforth called a "supporting" bias, and a negative bias is called "conflicting".

Definition 3. For any $\theta \in \Theta$, let $y^S(\theta)$ be the action that maximizes the sender's utility, and let $y^R(\theta) := y^S(\theta) + b$ be the action that maximizes the receiver's utility for some constant bias $b \in \mathbb{R}$. The bias is "supporting" if $b > 0$ and "conflicting" if $b < 0$ since

$$U^S(y^S(\theta) + b) > U^S(y^S(\theta) - b).$$

The results would be reversely identical if the sender preferred under-budgeting to over-budgeting. This could be easily captured by redefining k as $k' = 1/k$. In that case, a negative bias is supporting and a positive bias conflicting. If the receiver's bias is negative, the limiting bias is easily found by setting $a = 0$ in eq. (2.2).

Lemma 2. If the decision is made centrally by a committee of receivers such that the LLN applies, and $G(b)$ is uniform, the negative limiting bias is given by

$$|\bar{b}_{1-}^k| = \frac{\sqrt{k} - 1}{k - 1}, \quad \forall k > 1.$$

The absolute value of the limiting bias is decreasing in k . The results so far give rise to the following proposition.

Proposition 1. If the sender is uncertain about the receiver's bias, the stronger the asymmetry in the sender's preferences, the higher (lower) can the supporting (conflicting) bias be.

Naturally, the sender tolerates greater uncertainty in the receiver's preferences when her bias is to the same direction as the sender's preferred deviation from his bliss point. Hence, for the preferences of the sender that we consider, $\bar{b}_{1+}^k > |\bar{b}_{1-}^k|$ for all $k > 1$. If $k = 1$ and the sender's preferences are symmetric, $\bar{b}_{1+}^k = |\bar{b}_{1-}^k|$, and the direction of the receiver's bias plays no role.

2.3.3 Decentralized decision making

Consider now that the project's budget is decided by a single executive, for instance a firm's CEO.

Positive bias

Due to the piecewise-defined utility function, a sender of type a_i must take into account that, if $b > \frac{a_i - a_{i-1}}{2}$, the equilibrium action of the receiver upon observing message m_i , is higher than a_i and thus to the right of the sender's bliss point. Hence, in an equilibrium of size two, the arbitrage condition faced by the sender of type a , who considers whether to send message m_1 or m_2 , is given by

$$\begin{aligned} -k \int_0^{\min\{\bar{b}, \frac{a}{2}\}} \left(\frac{a}{2} + b - a\right)^2 \frac{1}{\bar{b}} db - \int_{\frac{a}{2}}^{\max\{\bar{b}, \frac{a}{2}\}} \left(\frac{a}{2} + b - a\right)^2 \frac{1}{\bar{b}} db \\ = - \int_0^{\bar{b}} \left(\frac{1+a}{2} + b - a\right)^2 \frac{1}{\bar{b}} db. \end{aligned}$$

By setting $a = 1$ and rearranging, the positive limiting bias is solved from

$$\frac{1}{3}(k-1)\bar{b}^3 - \frac{1}{2}k\bar{b}^2 + \frac{1}{4}k\bar{b} = 0, \quad (2.3)$$

if $\bar{b} \leq \frac{1}{2}$, and from

$$\frac{1}{2}\bar{b}^2 - \frac{1}{4}\bar{b} - \frac{1}{24}(k-1) = 0, \quad (2.4)$$

if $\bar{b} > \frac{1}{2}$.

The roots of the polynomial in (2.3) are given by $\frac{\frac{1}{2} \pm \sqrt{\frac{1}{3}k(1 - \frac{1}{4}k)}}{\frac{2}{3}(k-1)}$. They exist if $k \leq 4$. Moreover, for $k \leq 4$, the smaller positive root is strictly greater than $\frac{1}{2}$, a contradiction. This means that the limiting bias is larger than $\frac{1}{2}$ and determined as the smaller positive root of the polynomial in (2.4).

Lemma 3. *If the decision is decentralized and made by a single receiver whose positive bias is unknown to the sender and drawn from a uniform distribution, the limiting bias is given by*

$$\bar{b}_{1+}^k = \frac{1}{4} \left(1 + \sqrt{1 + \frac{4}{3}(k-1)} \right), \quad \forall k > 1.$$

As before, \bar{b}_{1+}^k is increasing in k . In contrast to the centralized protocol, however, the limiting bias under the decentralized protocol is larger than 1 for all $k \geq 7$.

Negative bias

If the receiver is inclined to allocate a smaller budget than the sender prefers, the limiting bias is solved in an analogous manner. The difference is, though, that we now need to take into account that if $\frac{1+a}{2} - b < a$, the sender's utility is determined by the steeper part of the utility function. To obtain the partition equilibrium of size two, the arbitrage condition faced by a sender of type $a \in (0, 1)$ is given as

$$\begin{aligned} & -k \int_0^{\bar{b}} \left(\frac{a}{2} - b - a \right)^2 \frac{1}{b} db = \\ & - \int_0^{\min\{\bar{b}, \frac{1-a}{2}\}} \left(\frac{1+a}{2} - b - a \right)^2 \frac{1}{b} db - k \int_{\frac{1-a}{2}}^{\max\{\bar{b}, \frac{1-a}{2}\}} \left(\frac{1+a}{2} - b - a \right)^2 \frac{1}{b} db. \end{aligned}$$

The negative limiting bias is obtained by setting $a = 0$ and solving for \bar{b} .

Lemma 4. *If the decision is decentralized and made by a single receiver whose negative bias is unknown to the sender and drawn from a uniform distribution, the limiting bias is given by*

$$|\bar{b}_{1-}^k| = \frac{\sqrt{\frac{1}{4} + \frac{1}{3}(k-1)} - \frac{1}{2}}{\frac{2}{3}(k-1)}, \quad \forall k > 1.$$

Proof: In the appendix. \square

Similarly to the centralized protocol, the absolute value of the negative limiting bias is decreasing in k . The lemmas obtained in this section give rise to the following result.

Theorem 1. *Let the sender's preferences be given as in (2.1) and the receiver's by a quadratic loss function. When the sender is uncertain about any single receiver's bias, and the bias is drawn from a uniform distribution,*

- *if the receiver's bias is "supporting", the decentralized protocol sustains a higher variance in b ,*
- *if the receiver's bias is "conflicting", the centralized protocol sustains a higher variance in b .*

Figure 2.2 visualizes the result of Theorem 1. As it shows, the variance in the supporting bias under the decentralized protocol can grow very large without destroying communication. There exists a large set of b 's for which informative communication is no longer feasible if the decision is made centrally, but remains feasible under the decentralized protocol. At least for these values of b , decentralizing the decision to a single receiver increases the ex-ante accuracy of communication. Moreover, the relative advantage of the decentralized protocol increases as k increases.

Exactly how this decentralization should be implemented, though, is a question of its own. The trick is that the decision should be random. Clearly, if the sender anticipates that the decision is delegated to the median voter, communication does not improve. If the decision is decentralized to a member of the committee, the decision right could be allocated only after hearing the sender, or the committee could implement rotating decision rights in such a way that the median voter's preferences are implemented on average over time.

On the other hand, centralized decision-making has an ex-ante informational advantage over decentralization when the receiver's bias is conflicting. This advantage is, however, modest. Nevertheless, when the committee members have a conflicting bias, decentralizing decisions to single committee members is useless in inducing more communication, unless $N(\bar{b}, k)$ assigns higher values under decentralization for some $b < \bar{b}_{1-}^k$.

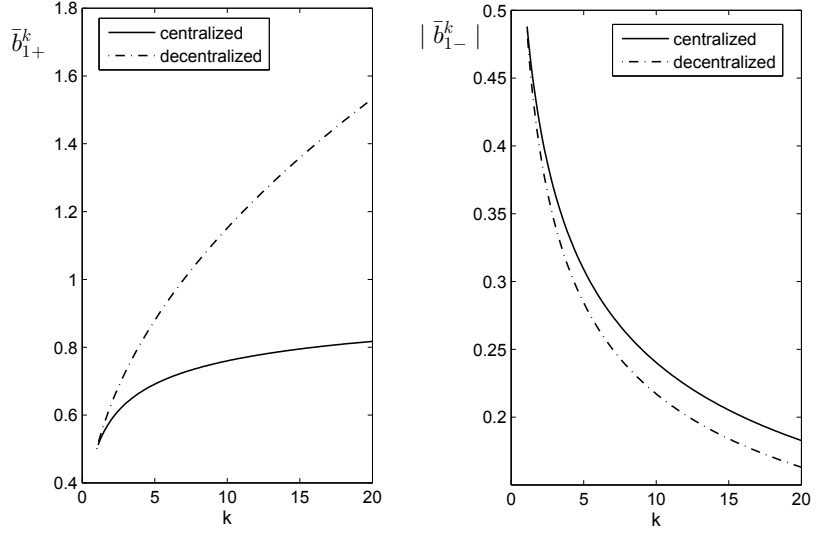


Figure 2.2. Limiting biases under different protocols

The result in Theorem 1 is rather intuitive. When the bias is supporting, the decentralized protocol can make use of the fact that the sender's expected payoff is a weighted average of payoffs for different values of b . The payoff function can thus take on values from the less steep part of the sender's utility function. That is, the supporting bias can take on values higher than $\frac{1}{2}a$ which is not possible under the centralized protocol. The expected utility from message m_1 is thus a weighted average of very low payoffs that arise when $b < \frac{1}{2}a$, and very high payoffs that arise when $b > \frac{1}{2}a$. The larger the non-symmetry in the sender's preferences the larger values above $\frac{1}{2}a$ the receiver's bias b can take to balance the average payoffs. At the same time, the relative weight of low biases decreases since probability mass is placed predominantly on biases higher than $\frac{1}{2}a$ (given that G is uniform). Naturally, this balancing of payoffs does not work to the other direction for conflicting biases.

2.4 Comparative statics

It is natural to ask to what extent the results obtained in the previous section can be generalized beyond the specifications used there. In particular, it would be useful to know if the results of section 2.3 are not just artifacts of the choice of the sender's utility function and distribution $G(b)$. This section tests whether the results of Theorem 1 continue to hold

under two changes in the model specifications. Firstly, by considering alternative distributions for the receiver's bias, and secondly, by considering logarithmic utility. The last subsection also discusses non-symmetries in the receiver's preferences.

2.4.1 Alternative distributions

Let G assign values for two points, 0 and $\bar{b} > 0$, such that $\Pr[b = \bar{b}] = p$, and $p \in (0, 1)$ is common knowledge. This section seeks to answer the following question: given p , what is the largest \bar{b} under the two protocols such that informative communication is still possible? Let us concentrate on positive biases only. By changing p we can roughly mimic alternative continuous distributions with different means. In particular, $p = \frac{1}{2}$ roughly corresponds to the uniform distribution considered in section 2.3.

The decision of the committee upon receiving message m_i is now $y(m_i, b) = \frac{a_{i-1} + a_i}{2} + p\bar{b}$. If $p = \frac{1}{2}$ the decision is identical to the centralized protocol analyzed earlier. On the other hand, if $p = 1$, the positive limiting bias must be half of what was found in the previous section. The positive limiting bias for the discrete setting can thus be solved to equal

$$\bar{b}_{1+}^k(p) = \frac{k - \sqrt{k}}{2p(k-1)}.$$

Under the decentralized protocol, for an equilibrium of size two, the arbitrage condition faced by a sender of type $a \in (0, 1)$ is now given by

$$\begin{aligned} & -pk \left(\frac{a}{2} + \bar{b} - a \right)^2 - (1-p)k \left(\frac{a}{2} - a \right)^2 \\ = & -p \left(\frac{1+a}{2} + \bar{b} - a \right)^2 - (1-p) \left(\frac{1+a}{2} - a \right)^2 \end{aligned}$$

if $\bar{b} \leq \frac{a}{2}$, and by

$$\begin{aligned} & -p \left(\frac{a}{2} + \bar{b} - a \right)^2 - (1-p)k \left(\frac{a}{2} - a \right)^2 \\ = & -p \left(\frac{1+a}{2} + \bar{b} - a \right)^2 - (1-p) \left(\frac{1+a}{2} - a \right)^2 \end{aligned}$$

if $\bar{b} > \frac{a}{2}$. After rearranging and substituting for $a = 1$, the positive limiting bias can be solved to equal

$$\bar{b}_{1+}^k(p) = \begin{cases} \frac{pk - \sqrt{p^2 k^2 - pk(k-1)}}{2p(k-1)} & \text{if } p \geq \frac{3}{4} \text{ and } k \in \left[\frac{1 - \sqrt{4p-3}}{2(1-p)}, \frac{1 + \sqrt{4p-3}}{2(1-p)} \right] \\ \frac{1}{4} \left(1 + \frac{1-p}{p} k \right) & \text{otherwise.} \end{cases}$$

The results in section 2.3 would suggest the hypothesis that the decentral-

ized protocol always guarantees higher supporting biases. To test whether this hypothesis can be maintained under other than the uniform distribution, I have analyzed numerically the difference in the limiting biases under the two protocols for different levels of p . The result of Theorem 1 (at least for the part of supporting biases) is maintained under all $p \in (0, 1)$ and for all $k > 1$. Figure 2.3 illustrates the difference when $p = \frac{1}{5}$.

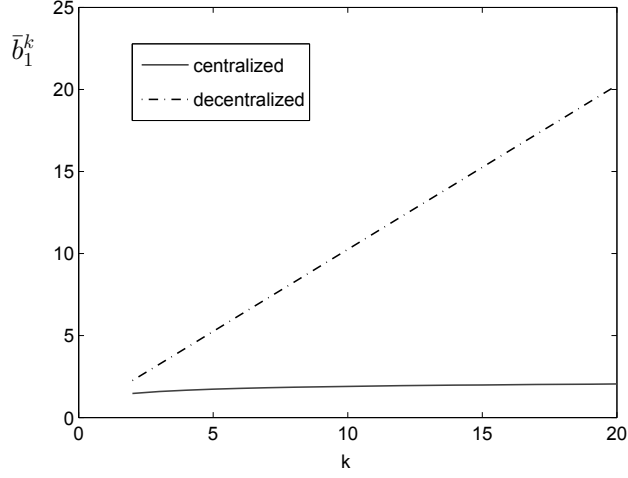


Figure 2.3. Difference between the limiting bias between the decentralized and the centralized protocol

2.4.2 Logarithmic utility

To get an idea how the results of section 2.3 depend on the form of the sender's preferences, this section considers an alternative specification. Let the preferences of the players be captured by the following logarithmic functions.

$$\begin{aligned}
 U^S(y, \theta) &= \theta \ln y - y \\
 U^R(y, \theta, b) &= \begin{cases} (\theta + b) \ln y - y & \text{w.p. } p \\ \theta \ln y - y & \text{w.p. } (1 - p), \end{cases}
 \end{aligned}$$

where $b \in \mathbb{R}$ and is common knowledge. This change in the receiver's preferences does not change her action rule since the expected utility is still maximized when her action is equal to her conditional expectation of the state θ given message m . In what follows, only the case of a positive

bias is covered.¹⁵

Under the decentralized protocol, the arbitrage condition characterizing the equilibrium partition of size 2 is given as

$$\begin{aligned} & (1-p) \left[a \ln \left(\frac{a}{2} \right) - \left(\frac{a}{2} \right) \right] + p \cdot \left[a \ln \left(\frac{a}{2} + \bar{b} \right) - \left(\frac{a}{2} + \bar{b} \right) \right] \\ &= (1-p) \cdot \left[a \ln \left(\frac{1+a}{2} \right) - \left(\frac{1+a}{2} \right) \right] + p \cdot \left[a \ln \left(\frac{1+a}{2} + \bar{b} \right) - \left(\frac{1+a}{2} + \bar{b} \right) \right], \end{aligned}$$

and when $a = 1$, the limiting bias must satisfy

$$-(1-p) \ln \frac{1}{2} + p \ln \frac{2+2\bar{b}}{1+2\bar{b}} - \frac{1}{2} = 0,$$

from where we obtain

$$\bar{b}_d(p) = \frac{1 - \left[\frac{1}{2} e^{\frac{1}{2}} \right]^{\frac{1}{p}}}{\left[\left(\frac{1}{2} \right)^{(1-p)} e^{\frac{1}{2}} \right]^{\frac{1}{p}} - 1},$$

which is positive if and only if $p \geq \frac{\ln \frac{1}{2} + \frac{1}{2}}{\ln \frac{1}{2}} \simeq 0.278$.

Under the centralized protocol, the arbitrage condition for an equilibrium partition of size 2 is given by

$$a \ln \left(\frac{a}{2} + p\bar{b} \right) - \left(\frac{a}{2} + p\bar{b} \right) = a \ln \left(\frac{1+a}{2} + p\bar{b} \right) - \left(\frac{1+a}{2} + p\bar{b} \right).$$

When $a = 1$, the limiting bias can be solved to equal

$$\bar{b}_c(p) = \frac{1 - \frac{1}{2} e^{\frac{1}{2}}}{p \left(e^{\frac{1}{2}} - 1 \right)},$$

which is strictly positive for all $p \in (0, 1)$. Further, it can be checked numerically that $\bar{b}_d(p) > \bar{b}_c(p)$ for all $p > \frac{\ln \frac{1}{2} + \frac{1}{2}}{\ln \frac{1}{2}}$. Similarly to increases in k in the piecewise-defined utility of section 2.3, the effects of changes in the concavity of the sender's preferences could be studied by taking a concave transformation of the logarithmic utility.

2.4.3 Non-symmetry in the receiver's preferences

Suppose that the receiver of any type b prefers over- to under-budgeting, that is, shares the same type of non-symmetry in preferences with the sender. This would shift her action rule to the right, towards the less

¹⁵Negative bias of the receiver is problematic since $\ln(y)$ is defined only for $y > 0$.

costly direction, but would not change the results qualitatively. The action rule of a receiver with non-symmetric preferences and no bias would be given by

$$y(m_i) \in \arg \min_y \int_{a_{i-1}}^y (y - \theta)^2 d\theta + k \int_y^{a_i} (y - \theta)^2 d\theta$$

Taking the derivative with respect to y and integrating over θ results in

$$k(y - a_i)^2 = (y - a_{i-1})^2.$$

Since $k > 1$, the identity requires that $|y - a_i| < |y - a_{i-1}|$, and hence y must be closer to a_i than to a_{i-1} . If the receiver in addition has an upward bias, her decision is leaning heavily to the right. For example, in an equilibrium partition of size 2, upon receiving message $m_1 \in [0, a)$, the receiver's action is $y(m_1) = a(k - \sqrt{k}) / (k - 1)$, where the multiplier $(k - \sqrt{k}) / (k - 1) > \frac{1}{2}$ for all $k > 1$. This means that the non-symmetry in the receiver's preferences works as a form of supporting bias, and leaves less room for the actual bias.

In fact, although both decision protocols are not affected by the same magnitude, the relative 'ranking' of the results would not change. That is, the decentralized protocol would still support a larger variance in the positive bias, but a smaller variance in the negative bias. However, as k increases, the non-symmetry in the receivers' preferences induces such a sizable shift to the right in the centralized decision that it turns the previously increasing series for \bar{b}_{1+}^k into a decreasing one. This is due to the fact that the theoretical maximum for the committee's limiting bias is one, as discussed in section 2.3.2.

On the other hand, the right shift in the action rule mitigates the effect of a conflicting bias and gives more room for it, resulting in an increase in $|\bar{b}_{1-}^k|$ for both protocols. Analogous reasoning, with reversely identical results, applies to the situation in which both players prefer underbudgeting.

If the non-symmetry in players' preferences is not of the same kind, as for instance when the sender prefers overbudgeting while the receiver prefers underbudgeting, the receiver's action rule would assign lower values than under symmetric preferences, giving all the more room for a supporting bias and even less room for a conflicting bias as compared to the results in section 2.3.

Note, however, that non-symmetries in committee members' preferences might hand more decision weight to those members who suffer more from a compromise. The members who end up on the steeper part of their utility function may induce, via negotiation power, a drift away from the median voter's decision. Therefore, a more careful formulation of the committee's inner dynamics may be in order. This problem would be absent if the committee is the pool of consumers in a market, all of whom make their individual decision but the producer only cares about the market's median reaction.

2.5 Discussion

Mainly for tractability, the decision of the committee is modeled by the median voter. To obtain more structure, the decision making in the committee could be formalized in more detail. Theoretical literature on voting in committees would offer a natural starting point. Another digression from the median voter preferences may arise in a dynamic setting. In a parliamentary committee, for instance, with conflicting interests, a joint decision hurts all but the median representative. This is likely to result in a trade of political favors whereby party A agrees to vote for the proposal of party B in return for B's support in some other decision important to party A.

The willingness of a committee member i to make compromises can be captured by adding a multiplier, k_i , to her utility function: $-k_i (y - (\theta + b))^2$. When all members have the same stakes in the decision, that is $k_i = k \forall i$, a compromise solution, in the form of a median voter outcome, can be reasonable. If the stakes are not equal across members, then those with high stakes may be able to use their bargaining power to shift the centralized decision to their preferred direction. Adding k would not change the analysis presented in this paper since the receiver's optimal action after a message from the sender remains the same as long as the symmetry in the utility function is retained.

The literature on committee decision making often assumes that each committee member gets a private or public signal, possibly with a noise. In most studies, however, the source of the signal is abstracted away. This paper points out that the signal that decision makers receive is in fact determined by the decision rule used and the composition of the committee (whether it consists of one or more receivers). Hence, this paper point

out the fact that the quality of the signal that a committee obtains about the state of nature (sent by the expert) may in some cases be significantly lower than the quality of the signal that a single decision maker would obtain to support her decision. This fact may be useful to keep in mind when designing the composition of committees or the decision-making protocol.

The results of the paper depend on the assumption that b is drawn from a uniform distribution. A distribution that is skewed to the right, having a low mean, would give rise to even higher limiting biases than what is reported now. On the other hand, the decentralized protocol may not be that attractive if the receiver's bias is drawn from a distribution skewed to the left, having a high mean. Alternative distributions combined with alternative utility functions therefore remain a necessary next step towards more general results.

Who should hide her type and who should reveal it? Naturally, receivers of low type prefer to reveal their type while high types, who benefit from the sender's uncertainty, prefer to keep their true type hidden. Hence, there is no credible way for low types to communicate their bias via cheap talk because all types have an incentive to claim that they are of the lowest type. By the same logic, the committee's attempts to encourage more information transmission by claiming to commit to a decision rule lower than the median voter would not be credible unless there was additional uncertainty about the distribution of decision weights among committee members. In such a situation, the best way to make one's low type known would be to act accordingly and use costly signaling to separate from the high types. This, however, renders the problem dynamic, whereby both low and high types optimize their paths of action over time. This seems like a prominent way forward in analyzing the effects of the sender's uncertainty about the type of his audience.

2.6 Conclusion

This paper aims to answer a question of the following type: if an expert is uncertain about the reaction of a decision maker to the information that he transmits, which protocol gives rise to more accurate communication on expectation - when the sender talks to one receiver of unknown type or when he talks to a group of receivers who act as one?

In an attempt to provide an answer to this question, this paper analyzes an otherwise standard one-shot cheap talk game but amends it in

two ways. First, the divergence of preferences between the players is only known by the receiver. Secondly, the model assumes that the preferences of the sender are non-symmetric around his bliss point. If this assumption is relaxed, the two communication environments are strategically equivalent and their comparison therefore trivial.

The paper finds that communication has the potential¹⁶ to remain informative under a higher variance of b if the receiver's bias is supporting and decision making is decentralized. If the receiver's bias is conflicting, communication has the potential to remain informative under a higher variance of b if decision making is centralized. A more general investigation would be needed to obtain results that span the whole step function $N(\bar{b}, k)$ for both decision-making protocols.

2.7 Appendix: Proofs

Proof of Lemma 4.

The analysis is done in two parts, depending on the value of $\min\{\bar{b}, \frac{1-a}{2}\}$. Suppose that $\min\{\bar{b}, \frac{1-a}{2}\} = \bar{b}$. The arbitrage condition reduces to

$$-k \left[\frac{1}{4}a^2\bar{b} + \frac{1}{2}a\bar{b}^2 + \frac{1}{3}\bar{b}^3 \right] = -\frac{1}{2}(1-a)^2\bar{b} + \frac{1}{2}(1-a)\bar{b}^2 - \frac{1}{3}\bar{b}^3.$$

When $a = 0$, the limiting bias (in absolute value) is given by

$$|\bar{b}_{1-}^k| = \frac{-\frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{3}(k-1)}}{\frac{2}{3}(k-1)} \quad (> 0).$$

The determinant is positive for all $k \geq \frac{1}{4}$, and since the absolute value must be positive, we have

$$|\bar{b}_{1-}^k| = \frac{-\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{3}(k-1)}}{\frac{2}{3}(k-1)}.$$

Moreover, it can be checked that $|\bar{b}_{1-}^k| \leq \frac{1}{2}$, so assumption $\min\{\bar{b}, \frac{1-a}{2}\} = \bar{b}$ holds when $a = 0$. Since the limiting bias is determined already, checking the other case is not necessary. Let us, however, check still that the assumption $\min\{\bar{b}, \frac{1-a}{2}\} = \frac{1-a}{2}$ cannot hold in the limiting case. The arbi-

¹⁶as babbling is always among the equilibria of the game

trage condition would be reduced to

$$-k \left[\frac{1}{4} a^2 \bar{b} + \frac{1}{2} a \bar{b}^2 + \frac{1}{3} \bar{b}^3 \right] = -\frac{1}{24} (1-a)^3 - k \left[\frac{1}{4} (1-a)^2 \left(\bar{b} - \frac{1-a}{2} \right) - \frac{1}{2} (1-a) \left(\bar{b} - \frac{1-a}{2} \right)^2 + \frac{1}{3} \left(\bar{b} - \frac{1-a}{2} \right)^3 \right].$$

Setting $a = 0$ and rearranging yields

$$k\bar{b}^2 - k\bar{b} + \frac{1}{24} (7k - 1) = 0, \quad (2.5)$$

such that

$$|\bar{b}_{1-}^k| = \frac{k \pm \sqrt{k^2 - \frac{1}{6}k(7k-1)}}{2k}.$$

Since the determinant is negative for all $k > 1$, we conclude that there are no values of \bar{b} that would satisfy condition (2.5) for any $k > 1$. \square

References

- Austen-Smith, D. and J.S. Banks (1996) "Information Aggregation, Rationality, and the Condorcet Jury Theorem". *American Political Science Review*, 90 (1), 34-45.
- Austen-Smith, D. and T. Feddersen (2009) "Information aggregation and communication in committees". *Philosophical Transactions of the Royal Society B*, 364, 763-769.
- Austen-Smith, D. and T. Feddersen (2006) "Deliberation, Preference Uncertainty, and Voting Rules". *American Political Science Review*, 100 (2), 209-217.
- Bénabou, R., and G. Laroque (1992) "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility". *Quarterly Journal of Economics*, 107, 921-958.
- Chen, Y. (2012) "Value of public information in sender-receiver games". *Economics Letters*, 114, 343-345.
- Chen, Y. (2009) "Communication with Two-sided Asymmetric Information", mimeo, Arizona State University.
- Crawford, V., and J. Sobel (1982) "Strategic Information Transmission". *Econometrica*, 50 (6), 1431-1451.
- Farrell, J. and R. Gibbons (1989) "Cheap Talk with Two Audiences". *American Economic Review*, 79 (5), 1214-1223.
- Feddersen, T. and W. Pesendorfer (1997) "Voting Behavior and Information Aggregation in elections with private information". *Econometrica*, 65 (5), 1029-1058.
- Gerardi, D. and L. Yariv (2004) "Information acquisition in committees". *Games and Economic Behavior*, 62 (2), 436-459.
- Goltsman, M. and G. Pavlov (2011) "How to Talk to Multiple Audiences". *Games and Economic Behavior*, 72 (1), 100-122.
- Heaton, J.B., (2002) "Managerial optimism and corporate finance". *Financial Management*, 31 (2), 33-45.
- Ishida, J., and T. Shimizu (2012) "Can more information facilitate communication?". ISER Discussion Paper No. 839.

- Krehbiel, K. (2004) "Legislative Organization". *Journal of Economic Perspectives*, 18 (1), 113-128.
- Lai, E. (2013) "Expert Advice for Amateurs". mimeo, Lehigh University, January, 2013.
- Li, H., S. Rosen, and W. Suen (2001) "Conflicts and Common Interests in Committees". *American Economic Review*, 91 (5), 1478-1497.
- Li, M., and K. Madarász (2008) "When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests". *Journal of Economic Theory*, 139, 47-74.
- Moreno de Barreda, I. (2010) "Cheap Talk with Two-Sided Private Information". mimeo, London School of Economics and STICERD.
- Morgan, J., and P. Stocken (2003) "An Analysis of Stock Recommendations". *RAND Journal of Economics*, 34 (1), 183-203.
- Morris, S. (2001) "Political Correctness". *Journal of Political Economy*, 109, 231-265.
- Rantakari, H. (2014) "A Simple Model of Project Selection with Strategic Communication and Uncertain Motives". *Journal of Economic Behavior and Organization*, forthcoming.
- Seidmann, D. (1990) "Effective Cheap Talk with Conflicting Interests". *Journal of Economic Theory*, 50, 445-458.
- Sobel, J. (1985) "A Theory of Credibility". *Review of Economic Studies*, 52 (4), 557-573.
- Watson, J. (1999) "Starting Small and Renegotiation". *Journal of Economic Theory*, 85 (1), 52-90.
- Wolinsky, A. (2003) "Information transmission when the sender's preferences are uncertain". *Games and Economic Behavior*, 42, 319-326.

3. Can You Keep a Secret? - Building Reciprocal Trust in Communication

3.1 Introduction

In economic theory, models of strategic communication between an expert and a decision maker have mostly been applied to settings in which communication partners either know each other's preferences, or the expert has an informational advantage both regarding the underlying state of the world and his own payoff-type. This paper is motivated by the observation that it is commonplace that a privately informed expert is uncertain how the information he transmits will be used. I analyze how the expert's uncertainty about the preferences of the decision-maker affects strategic communication in a repeated cheap-talk game.

Consider, for instance, that a sender has confidential information that he can share with a receiver periodically. Relevant situations arise between friends sharing secrets, between a supervisor and a subordinate sharing firm-sensitive information, or between a lobbyist and a government official sharing confidential information that goes beyond mere policy consultancy. All of these relationships rely on mutual trust; the receiver must trust the sender in his report, but more importantly, the sender must trust that the receiver does not use the information against him. If the sender is uncertain about the trustworthiness or intrinsic motivations of the receiver, he must be careful in what information to share. On the other hand, by sharing information, whether truthful or not, the sender is able to test the receiver to see if she can be trusted with additional information later on. Secrets themselves come in many forms, some are highly confidential while some others cause less trouble if leaked or misused. How much the sender and the receiver value each secret need not be correlated. This paves the way for reputation building in the sense

that an intrinsically ill-motivated receiver may act like a trustworthy person in the current period to maintain the communication relationship, but betray the sender's trust as soon as he shares a scoop.

The described setup has been largely overlooked in the cheap-talk literature. In this paper, I consider a model which has two-sided asymmetric information in the sense that the sender is privately informed of a state of the world, and the receiver is privately informed of her payoff-type. The receiver uses the information transmitted by the sender to take a decision that affects the welfare of both players. Borrowing for now the notation from Sobel (1985), the receiver may be one of two types, a Friend or an Enemy, and the sender needs to learn the type in order to tailor his future reports accordingly.

The only way for the sender to learn about the receiver's type is by sending informative reports to her and by observing the action that she takes in reaction to these reports. A Friend always chooses the action that maximizes her own payoff and the sender's payoff. Therefore, communication with a Friend would always be honest. An Enemy, on the other hand, has preferences in pure conflict with those of the sender — she has a myopic incentive to always choose the action that minimizes the sender's payoff. Therefore, the sender would never report his information perfectly to an Enemy. With only two possible reports, the best the sender can do is to randomize between them independently of his private information. However, the Enemy has sometimes an inter-temporal incentive to use the sender's report in the way the sender wants it to be used. Namely, if the first stage game is not valuable to the Enemy, she would use it to build a reputation for being a Friend by mimicking the equilibrium action of the Friend, only to exploit this reputation in the second stage where the unsuspecting sender reports his private information honestly. This means that sometimes, upon observing that the receiver takes the correct action from the sender's point of view, the sender only receives an imperfect signal about the receiver's payoff-type.

How informative the receiver's actions are about her payoff-type depends on how much the sender knows about the equilibrium action of the Enemy. If the sender knows that the first-stage game is not valuable to the Enemy and that both types behave well in the first period, he does not learn anything from the receiver's first-stage action. On the other hand, if the sender knows that the first-stage game is valuable to the Enemy and that the first-stage equilibrium is separating, he learns the receiver's

type perfectly at the end of the first period. If the sender does not know how much the receiver values the first-stage game, he does not know what drives the first-stage action of the Enemy – reputation or myopic incentives. As long as the Friend behaves well with a higher probability than the Enemy, the sender always learns from the receiver's action, but only less than perfectly.

This paper studies strategic communication when the sender tries to screen the receiver in the first period while at the same time trying to benefit from the Enemy's reputation concerns. The communication relationship is, however, delicate in the sense that trust must be mutual. That is, just as the sender can punish the receiver for not honoring trust, the receiver can punish the sender for not showing trust in the first place, that is, for sending false information. To sustain mutual trust, the sender has the incentive to screen the receiver by sending honest reports more frequently than he would do without the threat of punishment.

I study the structure of equilibrium communication with a particular focus on the receiver's behavior. I build a two-stage communication game and analyze separately two versions of it. The versions differ in their informational assumptions concerning the players' stakes which measure the importance of play. In the first version, information about the stakes is public, and both players know how his or her opponent values the first stage game. In the other version, the information about the stakes is private, and none of the players knows how much his or her opponent values the first stage. By doing this, the paper contributes to the literature in two ways. Firstly, it characterizes under the two model variants the most informative equilibria of a repeated cheap talk game in which the receiver has reputational concerns. Secondly, it analyzes the welfare effects from additional public information concerning the stakes, or more generally, the conditions under which the receiver chooses her action.

My main result is that for a wide range of parameter values, all players' ex-ante expected payoffs are higher if the stakes are private information. This seems counter-intuitive at first since one would expect that information is valuable in this type of a setting. However, from all players' point of view, the reputation concerns of the biased type can be better utilized if her stakes are not known to the sender. The ex-ante expected benefits that this yields especially to the sender compensate in many cases for the ex-ante expected losses that the sender incurs in case the receiver is biased and betrays his trust immediately.

A well-suited application of the analysis of this paper would be the communicating of new ideas in an organization. Suppose you have a new and exciting project idea but since it is partly outside your area of expertise you need help from a colleague at another department. Since you and your colleague do not meet on a daily basis, you are uncertain about her trustworthiness. There is a risk that after explaining your idea in detail to her, she will steal the idea and represent it as substantially her own. To succeed, the idea also needs expertise that you have, so if the colleague is to profit from stealing your idea, she must obtain accurate information from you. The optimal scenario from your point of view would be that you reveal the project idea together with your private information to your colleague who offers to you her expertise without owning the project. If the colleague is a team-player, this is how it would proceed. If, however, the colleague is selfish, she would bluff you with flawed information and before long would proceed with the project as if it was her own. Clearly, you would never trust your ideas to her again. If you suspect that your colleague is selfish, you can test her by giving her inaccurate information about the project or about the issue under your expertise. That way, if she steals your project, the project will not succeed and she will be tagged with a failed project. Beware, however, that if you do this to a colleague who is in fact a team-player, you will be revealed, even if falsely, to be selfish and your colleague would never help you in the future.

Another application concerns the reporting of sales forecasts within a firm. Consider a market analyst who reports sales forecasts periodically to the firm's CEO to back up investment decisions. The market analyst prefers the firm to invest in new production plants counter-cyclically: investments should be made when the demand is lagging. The CEO may be of the same opinion but may also be of the opposite opinion, that investments should be made when times are good and demand is high. Based on the sales forecast, the market analyst chooses whether to recommend a new production plant to be built or not. Should he be caught lying, the CEO would replace him for good.

The rest of the paper is structured as follows. Section 3.2 discusses related literature. The model in which there is uncertainty only about the receiver's payoff-type is described in section 3.3. Section 3.4 characterizes the most informative equilibrium of the game. Section 3.5 covers a variant of the model in which there is uncertainty about the receiver's payoff-type as well as both players' stakes. Section 3.6 compares the two models in

terms *ex ante* welfare. Discussion and conclusions are covered in sections 3.7 and 3.8, respectively.

3.2 Related literature

The seminal paper on cheap talk by Crawford and Sobel (1982) has been followed by numerous applications and variations. Most relevant to the present framework are sender-receiver games that deal with reputation, two-sided incomplete information, and dynamic information transmission.

The insights from the literature on reputation in principal-agent frameworks (see e.g. Kreps & Wilson, 1982; Milgrom & Roberts, 1982; Ely & Välimäki, 2003) have been applied to communication games in various ways. In most of this literature¹, interest is on the consequences for information transmission of the sender's reputational concerns. Out of the existing literature, Sobel (1985) shares some key elements with my model. He assumes that the sender can be of two types, good and bad, where a good sender is nonstrategic and always tells the truth while a bad sender chooses between truth-telling (investing in reputation) or lying (exploiting reputation). Moreover, the bad sender's preferences directly conflict with those of the receiver. In this paper, I analyze the complete opposite of this setting, by turning the informational setup upside down. To the best of my knowledge, the present paper is the first to study the effect of the receiver's reputational concerns on information transmission. In line with existing work on reputation building in communication games, reputational concerns as such are found to have a positive effect on information transmission.

Both Sobel (1985) and Morris (2001) incorporate stakes into their models essentially to give rise to reputational incentives. A similar approach has been employed more recently by Li & Mylovanov (2008). The usual assumption has been that the stakes are common knowledge among players. However, reputational concerns are present also when information about the stakes is held privately by each player. Regardless of the informational assumptions concerning the stakes, if desirable behavior is observed, the player who is trying to learn about his or her opponent cannot

¹Reputational concerns of the sender have been studied for instance by Morris (2001), Sobel (1985), Bénabou & Laroque (1992), Wrasai & Swank (2007), Guembel & Rossetto (2009), Durbin & Iyer (2009), Frisell & Lagerlöf (2007), Ottaviani & Sørensen (2006).

be certain which type took the particular action. This paper constructs an equilibrium under both scenarios and compares them in terms of ex ante welfare. This should bring new insights to the effects of reputational concerns in communication games.

Whereas dynamic communication in the previous papers focuses on studying the role of reputation, the models of Renault et al. (2013) and Golosov et al. (2014) focus on the dynamics of communication between two long-run players. As in the present paper, this allows both the sender and the receiver to discipline the behavior of his or her opponent by threatening to resort to playing a less informative equilibrium if the opponent fails to cooperate. Apart from this similarity, their papers are substantially different from mine. The model of Renault et al. (2013) differs most notably in the assumption that the state of the world follows a Markov chain, and the preferences of the receiver are commonly known. Golosov et al. (2014) find that gradual, and eventually full, revelation of information is possible in a dynamic cheap-talk game in which, in contrast to the present set-up, information is fully persistent in the sense that the state of the world remains the same across time.

Two-sided asymmetric information in cheap talk games has been studied by Chen (2009, 2012), Watson (1999), Lai (2013), Moreno de Barreda (2010), and Ishida & Shimizu (2012). In all of these, the receiver is assumed to be partially informed about the state of the world, but her preferences remain common knowledge. Watson (1996) deals with a model in which the sender is confused about his information; he does not fully understand it because he lacks a decoding device which the receiver, in turn, possesses. Sender's confusion in Watson's model has the same implications as in my model: the sender is uncertain about the reaction of the receiver to his information. However, in Watson's analysis, the sender's uncertainty is fundamentally about the state of the world while in my model, it is about the receiver's preferred way to use the information. Many of the aforementioned papers find that incomplete information is beneficial for communication. Against this background, the result of this paper, that additional public information may not be socially optimal, is not unique, but what distinguishes it from previous work is the source of informational asymmetry.

Related to the public disclosure of information, Li & Madarász (2008) analyze the welfare consequences of mandatory versus voluntary disclosure, via cheap talk, of the sender's bias. They conclude that it never ben-

efits the receiver nor the sender to impose mandatory disclosure of the sender's bias. In this paper, the bias of the receiver is fixed and neither the bias nor the stakes are subject to voluntary disclosure.

In the model to be analyzed, players' stakes are assumed to be drawn independently across players. This differs from the assumption used in Watson (1999, 2002) that players value each project identically and jointly decide on the optimal sequencing of projects in terms of their importance. Watson shows that the equilibrium implements a "starting small" feature such that the players gradually increase the stakes as the relationship evolves and the opponent behaves in a trustworthy manner. This modeling structure, however, excludes the randomness in the stakes which is central to the results in this paper.

Blonski and Probst (2004) analyze a game closely related to that of Watson (1999, 2002). They look at a repeated prisoner's dilemma game with two-sided incomplete information in that each player is privately informed of his or her discount factor which is either high or low. In addition, they introduce stakes in the form of a parameter λ that determines the division of a cake between the two players. Hence, the stakes are jointly determined and public knowledge. They demonstrate that in some equilibria of their game with two-sided incomplete information all players do strictly better than in a corresponding game with complete information. The authors are mainly concerned of the formation of relationships, and do not account for two-dimensional types as in the present analysis. Nevertheless, their paper builds on similar type of argumentation as in the current paper regarding the welfare effects of additional public information.

Finally, analysis of the social value of additional information relates to the studies of transparency in organizations. Prat (2005), Matozzi & Merlo (2007), Levy (2007), and Bar-Isaac (2012) are but a few examples that focus on issues of transparency in principal-agent relationships, although none of them considers cheap talk. In their papers, transparency is coined as the ability of the principal to either observe how the agent behaves and/or what the consequences of such behavior are. For example, the principal may or may not observe the agent's choice of effort. Transparency in the present model should be interpreted as the expert's ability to observe the conditions under which the decision-maker takes her decisions. In line with previous findings, increased transparency is not always socially optimal.

3.3 The Model

This section formalizes a model in which both players' stakes are common knowledge. The equilibrium of the game is analyzed in section 3.4. Section 3.5 covers an alternative model in which the stakes are each player's private information. Apart from the change in the informational structure and its implications for strategies, histories and equilibrium beliefs, the basic structure remains unchanged.

There are two players, a sender (S , or he) and a receiver (R , or she) who interact twice, in time periods $t = 1, 2$. In each period, R has to make a decision that is payoff relevant to both players. S has private information concerning a state of the world that affects R 's decision, and he can report his information to R by sending a costless and unverifiable message. There is potentially a conflict of interests between the players. As a result, communication is modeled as cheap talk. Before players choose their actions, Nature moves and draws R 's payoff type and players' stakes once and for all.

The payoff-type τ is privately observed by R and it determines R 's preferences vis-à-vis S . It is drawn from the commonly known distribution

$$\tau = \begin{cases} U & \text{wp. } \pi^1 \\ B & \text{wp. } 1 - \pi^1 \end{cases},$$

where U stands for “unbiased” and B for “biased”. The unbiased type has preferences aligned with those of S , whereas the biased type has preferences in pure conflict with S . Henceforth, I apply the short-hand notation R_U and R_B for unbiased and biased type, respectively.

Furthermore, all players publicly observe y and x , which are the realizations of random variables Y and X , governed by distributions H and G , respectively, with supports on $[0, \bar{y}]$ and $[0, \bar{x}]$, $\bar{y}, \bar{x} \in \mathbb{R}_+$. These parameters measure players' *stakes* in the first period, that is, the relative importance of the first period as compared to the second period. It is assumed that the stakes are independent of the state of the world, and x is independent of R 's payoff-type, τ . As a result, R 's type is a two-dimensional vector, $(\tau, x) \in \{U, B\} \times [0, \bar{x}]$, and there is asymmetric information about the first dimension. In the repeated game, S hopes to learn τ in order to be able to use his first-best communication strategy throughout the rest of the game. With R_U , this would be perfect communication, and with R_B , this would be babbling.

After the types have been drawn, the play proceeds through two stages. At the beginning of each stage, a state of the world, θ_t , is drawn in an i.i.d. manner from the set $\Theta_t = \{0, 1\}$, with $\Pr(\theta_t = k) = \frac{1}{2}$ for $k = 0, 1$. The state of the world is privately observed by S , and after observing it, S reports it to R by sending a message $m_t \in M_t = \{0, 1\}$. After observing m_t , R chooses an action, a_t , from $A_t = \{0, 1\}$ so as to maximize her expected payoff. After R 's action, payoffs are realized and privately observed, but all players observe the realized state of the world as well as actions taken. This is a crucial assumption for the analysis and the results of the game.

3.3.1 Payoffs

The stage game payoffs for S are given by a function $u_S : \Theta \times A \rightarrow \mathbb{R}$, and for the receiver by $u_R : \Theta \times A \times \{U, B\} \rightarrow \mathbb{R}$. In addition, both players weight the first period with y and x , respectively. The stage-game payoffs of S and R_U are maximized when $a_t = \theta_t$, and the payoff of R_B is maximized when $a_t = 1 - \theta_t$. The stage-game payoff vector (u_S, u_B) as a function of the state and the action taken is characterized in the following matrix.

	$a = 0$	$a = 1$
$\theta = 0$	$\begin{array}{ c } \hline 1, -1 \\ \hline \end{array}$	$\begin{array}{ c } \hline -1, 1 \\ \hline \end{array}$
$\theta = 1$	$\begin{array}{ c } \hline -1, 1 \\ \hline \end{array}$	$\begin{array}{ c } \hline 1, -1 \\ \hline \end{array}$

Table 3.1. Stage-game payoffs

Since the players' payoffs are symmetric in θ , the explicit values of θ_t , m_t or a_t are not of importance per se which simplifies the analysis of the game. Instead, what players care about is whether or not S 's report was truthful, and whether or not R followed the report. Finally, both players maximize the sum of expected stage-game payoffs which are given as

$$U^S(\theta_1, \theta_2, a_1, a_2; y) = y u_S(\theta_1, a_1) + u_S(\theta_2, a_2)$$

$$U_\tau^R(\theta_1, \theta_2, a_1, a_2; x) = x u_\tau(\theta_1, a_1) + u_\tau(\theta_2, a_2), \quad \tau \in \{U, B\}.$$

3.3.2 Histories and strategies

At the moment when the players choose their actions in the first period, they know the public history $h^1 \in \mathcal{H}^1 = [0, \bar{y}] \times [0, \bar{x}]$. The set of period 2 public histories is given by $\mathcal{H}^2 = (\Theta \times M \times A \times [0, \bar{y}] \times [0, \bar{x}])$, with a typical element h^2 . From S 's point of view, histories can be divided into two categories, good or bad, with good histories given by the set

$\mathcal{H}_{S+}^2 = \{h^2 \mid a_1 = m_1\}$, and bad histories by the set $\mathcal{H}_{S-}^2 = \{h^2 \mid a_1 \neq m_1\}$. On the other hand, from R 's point of view, good histories are only those where S has reported the state correctly. Therefore, denote the set of good histories from R 's point of view by $\mathcal{H}_{R+}^2 = \{h^2 \mid m_1 = \theta_1\}$, and the set of bad histories by $\mathcal{H}_{R-}^2 = \{h^2 \mid m_1 \neq \theta_1\}$. What matters for the analysis of the game are the commonly perceived good or bad histories. Therefore, denote the commonly perceived set of good histories by $\mathcal{H}_+^2 \equiv \mathcal{H}_{S+}^2 \cup \mathcal{H}_{R+}^2 = \{(0, 0, 0, y, x), (1, 1, 1, y, x)\}$. All the remaining histories of play are categorized as bad, since along such a history, either S or R has misbehaved.

A behavior strategy for S in period $t = 1, 2$ is a function $\mu^t : \Theta \times \mathcal{H}^t \rightarrow [0, 1]$, where $\mu^t(\theta, h)$ gives the probability that he reports the state θ_t truthfully given the history of play up to period t . The probability that S misreports the state, or lies, is given by the complementary probability $1 - \mu^t(\theta, h)$. A behavior strategy for R of type (τ, x) is a function $\alpha^t : M \times \mathcal{H}^t \times \{U, B\} \rightarrow [0, 1]$, where $\alpha^t(m, h, \tau)$ gives the probability that R follows S 's message m_t by taking action $a_t = m_t$ given her type and the history of play. The probability that R deviates from S 's message, by not following it, is given by $1 - \alpha^t(m, h, \tau)$.

3.3.3 Equilibrium concept

The equilibrium concept applied is Perfect Bayesian equilibrium (PBE)². This requires sequential rationality, that is, for any date t and any history h^t the strategies being played from h^t onwards constitute a Bayesian equilibrium of the continuation game. Formally, given the history of play h^t , let

$$V^S((\mu, \alpha) \mid h^t, \theta)$$

be the expected continuation payoff of the sender of type θ under strategy profile (μ, α) conditional on reaching history h^t . Similarly, let

$$V^\tau((\mu, \alpha) \mid h^t, m)$$

be the expected continuation payoff of the receiver of type (τ, x) under strategy profile (μ, α) conditional on reaching history h^t .

With the continuation payoffs defined, a PBE of the two-period game consists of a strategy profile $(\mu, \alpha_U, \alpha_B)$ and posterior beliefs $\pi^t(h)$, $f(\theta_t \mid m_t, h^t)$

²Since both states of nature, 0 and 1, occur with positive probability, there are no off-path messages so that a PBE of the game coincides with a sequential equilibrium.

such that

- i) for the sender, for each state of the world θ , alternative strategy μ' , and for any history h^t ,

$$V^S((\mu, \alpha_U, \alpha_B) \mid h^t, \theta) \geq V^S((\mu', \alpha_U, \alpha_B) \mid h^t, \theta).$$

- ii) for a receiver of type τ , for each message m , alternative strategy α' , and for any history h^t ,

$$V^\tau((\alpha_\tau, \mu) \mid h^t, m) \geq V^\tau((\alpha'_\tau, \mu) \mid h^t, m).$$

- iii) beliefs about the state of the world held by R , and beliefs about R 's type held by S are updated according to Bayes' rule whenever possible. That is,

$$\begin{aligned} \pi^2(h) &= \Pr\{\tau = U \mid a_1(m) = m\} = \\ &= \frac{\pi^1 \alpha^1(m, y, x, U)}{\pi^1 \alpha^1(m, y, x, U) + (1 - \pi^1) \alpha^1(m, y, x, B)} \end{aligned}$$

$$\begin{aligned} \Pr(\theta_t = k \mid m_t = k, h^t) &= \\ &= \frac{\Pr(\theta_t = k) \mu(h^t, k)}{\Pr(\theta_t = k) \mu(h^t, k) + \Pr(\theta_t = 1 - k) (1 - \mu(h^t, k))} \end{aligned}$$

As in all cheap talk games, there always exists an uninformative babbling equilibrium in which the sender's report does not depend on θ_t , and therefore R will ignore m_t and choose a_t based on her prior belief about θ_t .

Definition 1. *An equilibrium is said to be informative if, after receiving a message m , the receiver's belief $\Pr(\theta = k \mid m = k) \neq \frac{1}{2}$, for $k \in \{0, 1\}$. As a result, in any informative equilibrium, the receiver's expected payoff, $\mathbb{E}_{\theta} u(a(m), \theta)$, is higher than in the non-informative, babbling, equilibrium.*

Among the many informative PBE of the game, I focus on the one which is obtained when all players follow grim-trigger strategies that tell to resort to the babbling equilibrium forever after whenever anyone has deviated from the equilibrium path of play. That is, if S ever lies to R , or if R ever misuses S 's trust, the friendship will end immediately and will never be restored.³ Since babbling constitutes a stage-game Nash equilibrium and

³I acknowledge, however, that R_U 's punishment is not renegotiation-proof; once

gives all players their min-max payoffs, it is the most severe punishment available to any player. The use of grim-trigger strategies gives therefore the highest incentives for mutual cooperation. In fact, the equilibrium being characterized is the most informative equilibrium of the game, giving rise to as much communication overall as possible given the asymmetries in information.

3.3.4 A game without reputation

To illustrate how the sender's incentives to screen the receiver in the first period of the game affect the informativeness of communication, consider the following example in which it is commonly known that $x, y = 1$. That is, both players put equal weight on the decisions in periods 1 and 2, and hence the stakes are redundant. This immediately implies that R_B has no incentives to invest in reputation because the immediate cost for her from mimicking R_U , -1 , is just covered by the reward of 1 from being able to exploit the reputation at $t = 2$. Instead, R_B gains strictly more by diverting from the S 's message immediately. This gives her a total payoff of 1.

Strategies in period 2 Working backwards, the second and last period of the game is played like a static game. Suppose that the history of play is good and S enters the second period with posterior beliefs $\pi^2(h_+)$. Since R_B does not have reputational concerns, it must be that $\pi^2(h_+) = 1$ and S knows that R is unbiased. An unbiased R will naturally follow m_2 , which yields her and S a payoff of 1. If R has deceived trust in the first period, $\pi^2(h_-) = 0$ and S babbles in the second period. The best response of R_B is to ignore m_2 and randomize her action based on her prior. Both players obtain an expected payoff of 0.

Strategies in period 1 Given that R_U follows all honest reports and R_B diverts from them, the incentive compatibility condition faced by S , for any realization of θ_1 , is

$$\pi^1 - (1 - \pi^1) + \pi^1 V^S(h_+) \geq -\pi^1 + (1 - \pi^1), \quad (3.1)$$

where he compares his expected payoff from honest reporting to the expected payoff from lying about the first state. Only if these two actions

S has learned R 's type via a dishonest report, both players would be strictly better off by returning to cooperation. However, especially in friendships, it is not unheard of that the sender's dishonesty hits the pride of the receiver to such an extent that the two never speak to each other again.

yield the same payoff would S consider babbling or randomizing his first message.

From (3.1) it is evident that if $\pi^1 > \frac{1}{2}$, honesty strictly dominates lying in the current period. In addition, if faced by R_U , S obtains a strictly positive continuation payoff. If $\pi^1 = \frac{1}{2}$, S obtains 0 in the current period regardless of his report. However, honesty brings with it the strictly positive continuation payoff making it the optimal strategy for S . If $\pi^1 < \frac{1}{2}$, honest reporting yields immediate expected losses as compared to lying. However, for sufficiently high π^1 , the continuation payoff, $V^S(h_+) = 1$, is enough to compensate for these losses. Proposition 3.1 states the result formally.

Proposition 3.1. *If $y, x = 1$, and $\pi^1 \geq \frac{2}{5}$,*

$$\begin{aligned} \mu^1(\theta) &= 1, & \mu^2(\theta, h^2) &= \begin{cases} 1 & \text{if } h^2 \in \mathcal{H}_+^2 \\ \frac{1}{2} & \text{otherwise} \end{cases} \\ \alpha_U^1(m) &= 1, & \alpha_U^2(m, h^2) &= \begin{cases} 1 & \text{if } h^2 \in \mathcal{H}_+^2 \\ \frac{1}{2} & \text{otherwise} \end{cases} \\ \alpha_B^1(m) &= 0, & \alpha_B^2(m, h^2) &= \begin{cases} 0 & \text{if } h^2 \in \mathcal{H}_+^2 \\ \frac{1}{2} & \text{otherwise} \end{cases}. \end{aligned}$$

Moreover,

$$\pi^2(h) = \begin{cases} 1 & \text{if } a_1 = m_1 \\ 0 & \text{otherwise} \end{cases}, \quad f(\theta_1 | m_1) = 1, \quad f(\theta_2 | m_2, h^2) = \begin{cases} 1 & \text{if } h^2 \in \mathcal{H}_+^2 \\ \frac{1}{2} & \text{otherwise} \end{cases}.$$

If $\pi^1 < \frac{2}{5}$,

$$\mu^t(\theta) = \alpha_\tau^t(m) = \frac{1}{2}, \quad \forall t, \tau.$$

$$\pi^2(h) = \pi^1, \quad f(\theta_t | m_t) = \frac{1}{2}.$$

In order to preserve communication with a potential R_U , S is willing to screen R also when it implies expected costs in the short term. Without the threat of punishment S would not send informative messages, for any stakes, if $\pi^1 < \frac{1}{2}$. The use of grim-trigger strategies thus produces an ex-ante Pareto improvement; for priors $\pi^1 \in [\frac{2}{5}, \frac{1}{2})$, all players are ex ante strictly better off when S communicates truthfully as compared to

a babbling equilibrium. In fact, honest reporting is strictly better than babbling for all $\pi^1 > \frac{1}{3}$. However, the possibility of lying undermines communication for priors $\pi^1 \in [\frac{1}{3}, \frac{2}{5})$.

Next, I proceed to analyze the game when the stakes are not degenerate at 1. In the following section, the stakes are still degenerate but not necessarily equal to 1. Section 3.5 analyzes a model in which the stakes are random and players' private information. I will henceforth refer to the former scenario as 'Public regime', and to the latter as 'Private regime'.

3.4 Public regime equilibrium

This section characterizes the most informative communication equilibrium when both players' stakes are public information. S knows not only the equilibrium strategy of R but also her equilibrium action. As a result, he knows if the first-stage equilibrium is pooling or separating. In the former case, he does not learn anything, and in the latter case he learns R 's type with certainty.

The second and last period of the game is played like a static game. Suppose that no one has misbehaved in the first period, and S holds a belief $\pi^2(h_+)$ that R is unbiased. Since R_B has no reputation to keep up, she diverts from any truthful message while R_U follows it. Given the payoffs as specified in section 3, honest reporting is strictly optimal for S if and only if $\pi^2(h_+) > \frac{1}{2}$. For any lower beliefs, given R 's proposed strategy, S would rather lie about θ_2 than report it honestly. However, since R is aware of this, she will adjust her strategy accordingly. In equilibrium, S will not be able to fool R by lying, and the unique equilibrium must consist of babbling. If $\pi^2(h) = \frac{1}{2}$, S is indifferent between being honest or dishonest.

In the first period, the best response of R_U is always to follow S 's reports if they are truthful on expectation, regardless of the continuation payoffs which are at least zero in any case. I will therefore concentrate on characterizing the first-period equilibrium strategies of S and R_B . The structure of equilibrium depends on S 's prior belief about R 's type.

Proposition 4.1. *If $\pi^1 \geq \frac{1}{2}$, the sender reports θ_1 honestly for all $y \geq 0$.*

R_B plays according to the following cutoff strategy.

$$\alpha_B^1(m, h) = \begin{cases} 1 & \text{if } x \leq \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases}$$

Moreover,

$$\pi^2(h) = \begin{cases} \pi^1 & \text{if } x \leq \frac{1}{2} \\ 1 & \text{if } x > \frac{1}{2} \text{ and } a_1 = m_1 \\ 0 & \text{otherwise.} \end{cases}$$

Proof: In the appendix. \square

If S is confident that R is unbiased (i.e. $\pi^1 \geq \frac{1}{2}$), S 's equilibrium strategy is driven by his concern to maximize payoffs against R_U . This involves reporting the state truthfully regardless of any player's stakes. Irrespective of R_B 's action, S 's payoff is higher in both periods if he reports the first state honestly instead of lying. Given the strategy of S , if R 's stakes are lower than $\frac{1}{2}$, or put differently, if the second period is at least twice as important as the first period, the optimal strategy is to invest in reputation by mimicking R_U because it guarantees the maximal payoff at the second stage. If R 's stakes are higher than $\frac{1}{2}$, the myopic incentives dominate and R_B deceives S 's trust immediately.

Proposition 4.2. *If $\pi^1 \in [\frac{1}{4}, \frac{1}{2})$, and if*

$x \leq \frac{1}{2}$, $\mu^1(\theta, h) = 1$ for all $y \geq 0$, R_B follows the first report with probability $\alpha_B^1(m, h) = \frac{\pi^1}{1-\pi^1}$. If $h^2 \in \mathcal{H}_+^2$, $\mu^2(\theta, h) = \frac{1}{2} + x$;

$x > \frac{1}{2}$, S plays according to the following cutoff strategy,

$$\mu^1(\theta, h) = \begin{cases} 1 & \text{if } y \leq \hat{y}_{Pub} \\ \frac{1}{2} & \text{otherwise,} \end{cases}$$

where $\hat{y}_{Pub} = \min \left\{ \frac{\pi^1}{2(1-2\pi^1)}, \bar{y} \right\}$. R_B 's strategy is given by

$$\alpha_B^1(m, h) = \begin{cases} 0 & \text{if } y \leq \hat{y}_{Pub} \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Proof: In the appendix. \square

If S 's prior is less than $\frac{1}{2}$ he cares about R_B 's action, and all the more so the lower the prior is. Consider the case in which R_B has reputational concerns. To ensure that the investment in reputation pays off in the future, R_B must mimic R_U with a probability less than one so as to induce a high enough posterior $\pi^2(h)$. In equilibrium, all types $x \leq \frac{1}{2}$ randomize their first-stage action by mimicking R_U with a probability that just induces a posterior belief of $\frac{1}{2}$. To ensure that all types $x \leq \frac{1}{2}$ are indifferent between investing in reputation and deceiving S immediately, S must decrease the benefits from reputation by promising an honest report at $t = 2$ with a probability less than one. One way to achieve this is to randomize between honest and false reporting.⁴ Since $\pi^2(h_+) = \frac{1}{2}$, consistency of S 's mixed-strategy is ensured.

When S knows that R_B has only myopic concerns ($x > \frac{1}{2}$) and the first-stage equilibrium is separating, his prior is too low to screen R for all stakes y . For priors less than $\frac{1}{2}$, honest reporting yields expected losses in period 1. Therefore, S is willing to report the first state truthfully and thereby screen R if and only if the first period is not very important, that is, if and only if $y \leq \hat{y}_{Pub}$. For all stakes higher than \hat{y}_{Pub} , since S cannot gain from lying in the Public regime, communication can only consist of babbling.

Proposition 4.3. *If $\pi^1 < \frac{1}{4}$, there exists an informative equilibrium such that for all $y \leq \hat{y}_{Pub}$ S reports θ_1 honestly, R_U follows m_1 and R_B diverts from it. Otherwise the unique equilibrium consists of babbling.*

If S 's prior is lower than $\frac{1}{4}$ the mixed-strategy equilibrium cannot be supported any longer because S prefers to lie about θ_1 . This is because to induce a high enough posterior belief, R_B must mimic R_U with such a low probability that S does not find it worthwhile to seek benefits from R_B 's reputational concerns. Because R_B has no further ways to profitably invest in reputation, the equilibrium for low priors is separating. As discussed under Proposition 4.2, S is willing to screen R if and only if his stakes are low enough.

The equilibrium of the game is visualized in Figure 3.1 (Figure 1). For mere expositional purposes, \bar{y} is set equal to 4. The lower graph displays R_B 's strategy in the first period in the nontrivial case when $x \leq \frac{1}{2}$. For in-

⁴Another way would be to randomize between honest reporting and babbling.

intermediate priors $\pi^1 \in [\frac{1}{4}, \frac{1}{2})$, R_B randomizes her action in the first period by following with probability $\alpha_B^1(m) \in [\frac{1}{3}, 1)$. In this interval, the realization of x also affects the equilibrium second-stage action of S .

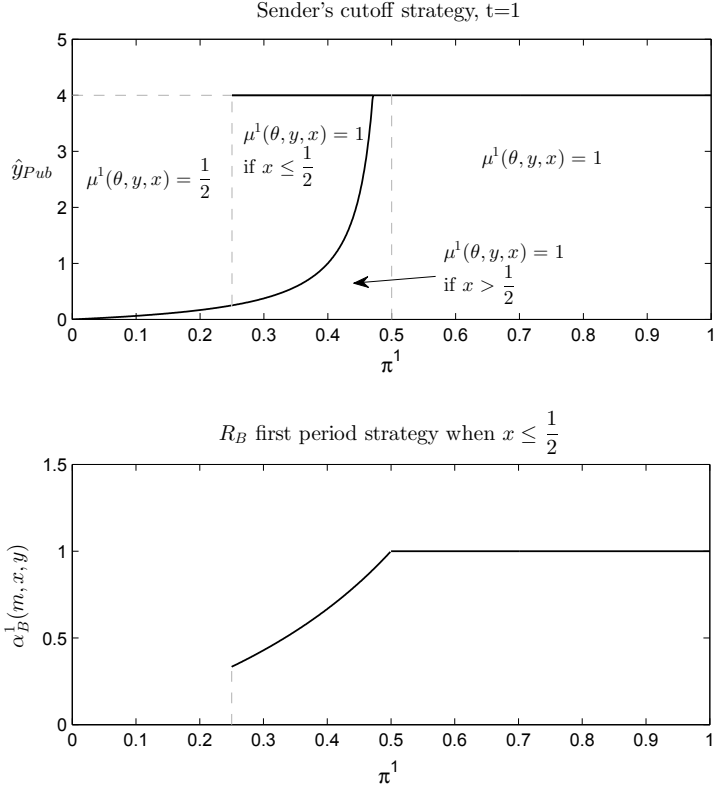


Figure 1. Public regime equilibrium

Figure 3.1. Public regime strategies

3.5 Private regime equilibrium

The assumption that players always know each other's stakes seems somewhat optimistic. Interactions in which none of the players actually knows how valuable the stage game is to their opponent are so common that it is important to study how this affects strategic communication. In this section, I characterize the most informative communication equilibrium of a game in which both players observe their stakes privately at the beginning of the game. That is, asymmetric information relates not only to

the state of the world and R 's payoff-type but also to the stakes of one's opponent. The distributions H and G from which the stakes are drawn from are common knowledge. In other respects the model is as formalized in section 3.3.

The change in the informational assumptions affects histories, strategies and beliefs. The public history at $t = 1$ is now $h^1 = \{\emptyset\}$, and therefore omitted in notation for the rest of the section. The public history of play at the beginning of the second period is $\mathcal{H}^2 = (\Theta \times M \times A)$. Players' strategies now only depend on one's own stakes, and equilibrium beliefs are characterized by the following Bayes' rules.

$$\begin{aligned}\pi^2(h_+) &= \Pr \{ \tau = U \mid a_1(m) = m \} \\ &= \frac{\pi^1 \int_{[0, \bar{x}]} \alpha_U^1(m, x) dG(x)}{\pi^1 \int_{[0, \bar{x}]} \alpha_U^1(m, x) dG(x) + (1 - \pi^1) \int_{[0, \bar{x}]} \alpha_B^1(m, x) dG(x)} \\ \Pr(\theta_t = k \mid m_t = k, h^t) &= \\ &= \frac{\Pr(\theta_t = k) \int_{[0, \bar{y}]} \mu^1(k, y, h^t) dH(y)}{\Pr(\theta_t = k) \int_{[0, \bar{y}]} \mu^1(k, y, h^t) dH(y) + \Pr(\theta_t = 1 - k) \int_{[0, \bar{y}]} 1 - \mu^1(k, y, h^t) dH(y)}\end{aligned}$$

As in the previous section, the equilibrium will be characterized by cutoff strategies for both S and R_B , with equilibrium cutoffs denoted by \hat{y}_{Pr} and \hat{x}_{Pr} , respectively. Now that y is not observed by R , the equilibrium cutoff strategy of S will be of different nature than under the Public regime. Because R is not able to tell whether y is below or above \hat{y}_{Pr} , S is able to lie about the state in equilibrium. Therefore, the structure of equilibrium strategies that I look at is the following. For S ,

$$\mu^1(\theta, y) = \begin{cases} 1 & \text{if } y \leq \hat{y}_{Pr} \\ 0 & \text{otherwise,} \end{cases}$$

and for R_B ,

$$\alpha_B^1(m, x) = \begin{cases} 1 & \text{if } x \leq \hat{x}_{Pr} \\ 0 & \text{otherwise.} \end{cases}$$

In the second and last period of the game, R_B has no reputation to keep up and will divert from any message which is truthful on expectation. R_U

will continue to follow if the history of play is good. The equilibrium will thus be identical to the Public regime as characterized in section 3.4.

In the first period, since R_U has only two available actions, she follows S 's report as long as it is truthful on expectation, that is, as long as $H(\hat{y}_{Pr}) > \frac{1}{2}$. For the rest of the section, I concentrate on characterizing the equilibrium strategies of S and R_B . Proposition 5.1 is the Private regime analog of Proposition 4.1 with the difference that the lower bound for priors supporting the equilibrium is given by $\tilde{\pi} \leq \frac{1}{2}$.

Proposition 5.1. *If $\pi^1 \geq \tilde{\pi} \equiv \max\{\tilde{\pi}_{SR}, \tilde{\pi}_{IC}\}$, S reports θ_1 honestly for all $y \geq 0$, and $\mu^2(\theta, h) = 1$ iff $h^2 \in \mathcal{H}_+^2$. R_B follows a cutoff strategy with $\hat{x}_{Pr} = \frac{1}{2}$.*

Proof: In the appendix. \square

The condition $\pi^1 \geq \tilde{\pi}_{IC}$, where ' IC ' is short-hand for incentive compatibility, ensures that honest reporting is optimal at $t = 1$ without accounting for continuation payoffs. It follows from the inequality

$$[\pi^1 + (1 - \pi^1)G(\hat{x}_{Pr}) - (1 - \pi^1)(1 - G(\hat{x}_{Pr}))]y \geq 0$$

that is,

$$\pi^1 \geq \frac{1 - 2G(\hat{x}_{Pr})}{2(1 - G(\hat{x}_{Pr}))} \equiv \tilde{\pi}_{IC}. \quad (3.2)$$

The condition $\pi^1 \geq \tilde{\pi}_{SR}$, where ' SR ' is short-hand for sequential rationality, ensures that honest reporting is optimal at $t = 2$ after a good history of play, and it is derived from the condition

$$\pi^2(h_+) := \frac{\pi^1}{\pi^1 + (1 - \pi^1)G(\hat{x}_{Pr})} \geq \frac{1}{2}, \quad (3.3)$$

that is,

$$\pi^1 \geq \frac{G(\hat{x}_{Pr})}{1 + G(\hat{x}_{Pr})} \equiv \tilde{\pi}_{SR}.$$

Since S has a myopic incentive to report honestly in both periods, R_B is free to invest in reputation whenever $x \leq \frac{1}{2}$. If S 's prior is lower than $\tilde{\pi}$, either he or R_B must adjust the equilibrium strategy to guarantee informative communication.

Proposition 5.2. *If $\pi^1 \in [\frac{1}{4}, \tilde{\pi}_{SR})$, S reports θ_1 honestly for all $y \geq 0$. R_B invests in reputation if and only if $x \leq \hat{x}_{Pr} = G^{-1}\left(\frac{\pi^1}{1-\pi^1}\right)$. Moreover, in equilibrium, $\mu^2(\theta, h_+) = \frac{1}{2} + \hat{x}_{Pr}$.*

Proof: In the appendix. \square

If S 's prior is lower than $\tilde{\pi}_{SR}$ but higher than $\tilde{\pi}_{IC}$, which occurs whenever $G(\frac{1}{2}) \geq \frac{1}{3}$, S finds it optimal to report θ_1 honestly, but if R_B follows the first report with probability 1, S would babble in period 2, a contradiction. An equilibrium with reputation building thus requires that \hat{x}_{Pr} adjust to ensure a high enough posterior $\pi^2(h_+)$. In equilibrium, the highest type who still invests in reputation is the type \hat{x}_{Pr} for whom the sequential rationality condition (3.3) holds as equality. If there was slack, implying that S would report the second state truthfully after a good history of play, there would always exist some type $\hat{x} + \epsilon$ who would strictly prefer to invest in reputation than not. To make the cutoff type \hat{x}_{Pr} indifferent, and in particular, to make sure that all types higher than the cutoff prefer to forgo the investment in reputation, S must make the investment less appealing by randomizing his second-stage action after a good history of play. Unlike in the Public regime where each type $x \leq \frac{1}{2}$ must be made indifferent, in the Private regime S only needs to make sure that the cutoff type is indifferent. Finally, the equilibrium requires that S want to report θ_1 honestly for all stakes y . Since S 's expected payoff from period 2 is zero, the lower bound for priors that support the equilibrium is obtained by substituting for $G(\hat{x}_{Pr}) = \frac{\pi^1}{1-\pi^1}$ in (3.2).

Proposition 5.3. *If $\pi^1 \in [\hat{\pi}, \tilde{\pi}_{IC})$, where $\hat{\pi} = \max\{\hat{\pi}_x, \hat{\pi}_y\}$, and $\hat{\pi}_x$ is the lowest prior such that $G(\hat{x}_{Pr}) \leq \frac{1-2\pi^1}{2(1-\pi^1)}$, and $\hat{\pi}_y$ is the lowest prior such that $H(\hat{y}_{Pr}) > \frac{1}{2}$, the equilibrium first-period strategies of S and R_B are characterized as*

$$\mu^1(\theta, y) = \begin{cases} 1 & \text{if } y \leq \hat{y}_{Pr} \\ 0 & \text{otherwise} \end{cases}, \quad \alpha_B^1(m, x) = \begin{cases} 1 & \text{if } x \leq \hat{x}_{Pr} \\ 0 & \text{otherwise} \end{cases},$$

where \hat{y}_{Pr} and \hat{x}_{Pr} are determined simultaneously by

$$\hat{y}_{Pr}(\pi^1, \hat{x}_{Pr}) = \frac{\frac{1}{2} [\pi^1 - (1 - \pi^1)G(\hat{x}_{Pr})]}{[(1 - \pi^1)(1 - 2G(\hat{x}_{Pr})) - \pi^1]}$$

$$\hat{x}_{Pr}(\hat{y}_{Pr}) = \frac{H(\hat{y}_{Pr})}{2(2H(\hat{y}_{Pr}) - 1)}.$$

Moreover, in equilibrium,

$$\pi^2(h_+) = \frac{\pi^1}{\pi^1 + (1 - \pi^1)G(\hat{x}_{Pr})}.$$

Proof: In the appendix. \square

If S 's prior is less than $\tilde{\pi}_{IC}$ but higher than $\tilde{\pi}_{SR}$, which occurs whenever $G(\frac{1}{2}) < \frac{1}{3}$, honest reporting yields S an immediate expected loss and therefore, taking into account the positive continuation payoff if S is truthful, honest reporting of the first state occurs only for $y \leq \hat{y}_{Pr}$. The equilibrium cutoff \hat{y}_{Pr} is higher than under the Public regime because S uses truthful reporting not only to screen R but also to benefit from the reputational concerns of R_B that exist with probability $G(\hat{x}_{Pr})$.

As a reaction to the fact that S lies about the first state with probability $1 - H(\hat{y}_{Pr})$, R_B 's cutoff \hat{x}_{Pr} adjusts upwards. Namely, R_B has now a higher incentive to follow the first report, not to invest in reputation but to reap as much as she can from the first period in case S lies and communication turns into babbling in period 2 irrespective of R_B 's action. The higher the probability that S lies, the higher the R 's cutoff. Therefore, in the equilibrium \hat{x}_{Pr} is a decreasing function of \hat{y}_{Pr} .

Finally, notice that the lower bound supporting the equilibrium is determined by the interplay of \hat{x}_{Pr} and $\tilde{\pi}_{IC}$ in the following sense. Since \hat{x}_{Pr} is decreasing in π^1 , the probability $G(\hat{x}_{Pr})$ is also decreasing in π^1 . This means that S 's incentive to report θ_1 truthfully is also decreasing in π^1 . The equilibrium, however, requires that the first-period yields S an expected loss, that is, that $\pi^1 < \tilde{\pi}_{IC}$. This again requires that $G(\hat{x}_{Pr}) \leq \frac{1-2\pi^1}{2(1-\pi^1)}$. Denote by $\hat{\pi}_x$ the prior at which the curves $G^{-1}\left(\frac{1-2\pi^1}{2(1-\pi^1)}\right)$ and \hat{x}_{Pr} as defined in the proposition intersect. At the same time, the equilibrium requires that $H(\hat{y}_{Pr}) > \frac{1}{2}$. Denote by $\hat{\pi}_y$ the lowest π^1 for which this holds. The lower bound is then given by $\max\{\hat{\pi}_x, \hat{\pi}_y\}$.

Proposition 5.4. *If π^1 is lower than $\frac{1}{4}$ or $\hat{\pi}$, depending on the value of*

$G(\frac{1}{2})$, there exists a cutoff

$$\hat{y}_{Pr} = \frac{\pi^1}{2(1 - 2\pi^1)}$$

such that for each π^1 , if $H(\hat{y}_{Pr}) > \frac{1}{2}$, S reports θ_1 truthfully for all $y \leq \hat{y}_{Pr}$ and R_B diverts from all m_1 . Therefore, $\mu^2(\theta, h_+) = 1$. For each π^1 , if $H(\hat{y}_{Pr}) \leq \frac{1}{2}$, the unique equilibrium consists of babbling.

Proof: In the appendix. \square

Similarly to the Public regime, R_B does not invest in reputation for low priors. Since π^1 is low, S engages in screening R by reporting θ_1 truthfully if and only if his stakes are low. While under the Public regime, a low cutoff \hat{y} poses no problem for the existence of the informative equilibrium, the existence is restricted by the condition $H(\hat{y}_{Pr}) > \frac{1}{2}$ under the Private regime.

Before turning to the welfare analysis, a short look at the difference in S 's cutoffs between the regimes is in order as it will in part determine the relative ranking of the regimes in terms of ex ante welfare. Notice, firstly, that \hat{y}_{Pub} is independent of distributions G and H . Secondly, $\hat{y}_{Pub} < \hat{y}_{Pr}$ for all $\pi^1 \geq \hat{\pi}$. The reason is that under the Public regime, the cutoff only ensures that honest reporting is optimal when R_B is known to deviate immediately, which imposes a more demanding constraint for truthful communication. At the limit, when $G(\frac{1}{2}) = 0$, the cutoffs are identical for all π^1 .

Figure 3.2 (figure 2) plots the sender's cutoff \hat{y}_{Pr} for two alternative distributional assumptions. The dashed line assumes that both players' stakes are distributed according to a uniform ratio distribution⁵. The support of the distribution is \mathbb{R}_+ but the vertical axis is limited to 4 for expositional purposes. The dotted line assumes that both players' stakes are distributed uniformly on $[0, 2]$. Both of the considered distributions are analytically convenient and impose the assumption that $\mathbb{E}(X) = \mathbb{E}(Y) = 1$. Moreover, both distributions imply that the equilibrium condition $H(\hat{y}_{Pr}) > \frac{1}{2}$ holds if $\hat{y}_{Pr} > 1$. Therefore, if $\pi^1 < \frac{1}{4}$, an informative equilibrium fails to exist under the Private regime because the equilib-

⁵This assumption seems reasonable if the stakes are derived as the ratio of the stakes of periods 1 and 2. Suppose that both are draws from $U[0, 1]$. The ratio $Z = \frac{X_1}{X_2}$ follows a uniform ratio distribution with a pdf $p(z) = \begin{cases} \frac{1}{2} & \text{if } z \leq 1 \\ \frac{1}{2z^2} & \text{if } z > 1. \end{cases}$

rium cutoff, equal to \hat{y}_{Pub} as plotted in the figure, is too low. This explains why the Public regime yields all players strictly higher ex ante welfare for low priors; there is a positive probability that information is transmitted which gives all players strictly higher expected payoffs than babbling. We turn to the welfare comparison next.

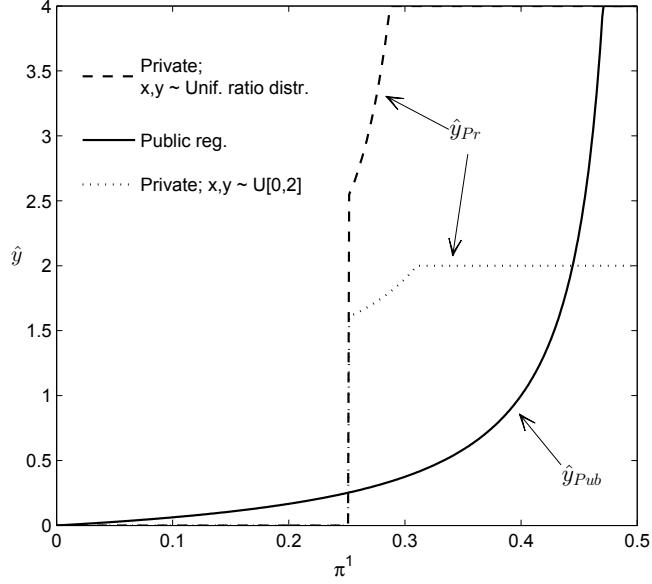


Figure 2. Equilibrium cutoffs \hat{y}

Figure 3.2. Equilibrium cutoff of the sender

3.6 Welfare

This section examines the social value of public information concerning the stakes. This is done by comparing the ex ante expected payoffs from the two models analyzed in section 3.4 and 3.5. The results show that, for a range of parameter values, public information about the stakes hurts at least some and often all players. The result derives mainly from the differences in the informational assumptions about R 's stakes. Whether S 's stakes are known or not affects welfare to a lesser extent. In fact, the result would remain qualitatively unchanged if we set $y = 1$ and ignore the sender's stakes altogether. The result that additional information about R is detrimental for ex-ante welfare seems surprising because in a single-player game information is always valuable in the sense of Blackwell.

To see this, consider a single-player game which replicates the commu-

nication model but replaces R with a machine. A player does not know the type of the machine and can play it at most twice. A good machine always yields a prize whereas a bad machine follows a cutoff rule in the first period by returning a prize if the realization of a random variable X is below an exogenously given threshold \hat{x} , and otherwise returns nothing. There is a prior probability of π that the machine is good. In this setting, a player is always at least as well off by knowing x than not knowing it because he always has the option to ignore this additional information and play as if he did not know x if that results in higher expected payoffs. When the machine is replaced by a strategic R who, if biased, behaves well only if it ensures a high enough continuation payoff, more information may make not only R but also S strictly worse off. The fact that private information smooths the posterior turns out to be a blessing for all players.

The analysis is done in two parts, separately for S and R . I limit attention to priors in the range $[\frac{1}{4}, \frac{1}{2}]$. The welfare analysis for all lower and all higher priors is straightforward. More specifically, as long as $H(\hat{y}_{Pr}) < \frac{1}{2}$, the Public regime strictly dominates the Private for all $\pi^1 < \frac{1}{4}$ because the Private regime has a unique equilibrium in babbling. If $H(\hat{y}_{Pr}) > \frac{1}{2}$, the Private regime has an informative equilibrium in which S lies if $y > \hat{y}_{Pr}$. The only player who benefits from this is S because both types of R prefer babbling to being deceived. A numerical example of such a case is included in the appendix and it shows that lying hurts R significantly more than S ever benefits from it.

On the other hand, for all $\pi^1 \geq \frac{1}{2}$, the regimes are identical in equilibrium strategies and therefore also in expected payoffs. Since S is confident enough in facing R_U , he ignores any information about the stakes, his own or R 's, and therefore the regime does not play any role in determining the welfare.

For the remainder of the section, to obtain a better view of what is driving the results, I focus on the ex-ante welfare at the point in time where the realization of Y is known (to the researcher) but the realization of X remains unknown. This should ease the exposition of the results and makes the composition of each player's welfare more transparent than if we only look at the expectation of Y . There exist parameter environments in which all players strictly prefer the Private regime over the Public, but also those in which the reverse holds. The actual ex-ante welfare is obtained by integrating over Y . A few numerical examples are provided in the appendix when relevant. The next subsection concentrates on the

sender's welfare, and the receiver's welfare is covered separately in subsection 3.6.2.

3.6.1 Sender's welfare

The sender's welfare is a result of two counteracting effects. On the one hand, he benefits in the short run from R_B 's reputational concerns by obtaining his maximum payoff whenever R_B mimics R_U . On the downside, this involves the risk of incurring losses at the second period where a potential R_B exploits her reputation at the expense of S . However, the more important the first period is to S the less the expected losses in the second period matter. Therefore, the higher S 's stakes are the more he stands to gain from R_B 's reputational incentives. When R_B 's motive in the first period is to establish a reputation, S maximizes his gains from this by allowing R_B to maximize the probability with which she mimics R_U . For this reason S strictly prefers the Private regime where R_B has more room to invest in reputation than under the Public regime where S cannot update his beliefs when R_B is known to pool with R_U . Furthermore, the optimality of the Private regime increases monotonically in y .

On the other hand, S incurs expected losses from trusting R if R_B does not have reputational concerns and deceives S 's trust immediately. Suppose that $x > \frac{1}{2}$. If $y \leq \hat{y}_{Pub}$, both regimes have an equilibrium in which S trusts R at $t = 1$, and hence they yield identical expected payoffs. If $y > \hat{y}_{Pub}$, the unique Public regime equilibrium consists of babbling. The informative equilibrium in the Private regime dominates babbling as long as S 's stakes are not too high ($\hat{y}_{Pub} \leq y \leq 2\hat{y}_{Pub}$) because the positive continuation payoff if R turns out to be unbiased outweigh the short-term expected cost from being deceived by R_B . For all stakes higher than $2\hat{y}_{Pub}$, however, the equilibrium of the the Private regime, involving high trust, is inferior in expected payoffs to the babbling equilibrium of the Public regime. Moreover, the difference to the benefit of the Public regime increases in y as S has more to lose if exploited by R_B . Hence, for all high stakes, S would like to know if R_B is going to divert from m_1 so as to tailor the reporting strategy accordingly. In fact, the optimal strategy for those high stakes is to babble immediately and forgo the possibility to learn R 's type altogether.

For the analysis of the sender's welfare, it is useful to distinguish between two effects which I call the reputation effect and the screening effect. The reputation effect favors the Private regime and occurs with

probability $G(\hat{x})$, and the screening effect favors the Public regime and occurs with probability $1 - G(\hat{x})$.

Definition 2. *Reputation effect (RE) and screening effect (SE) are defined as follows.*

$$RE := \mathbb{E}_\tau \mathbb{E}_X \left[V^{Pr}(\mu, \alpha) - V^{Pub}(\mu, \alpha) \mid X \leq \hat{x} \right]$$

$$SE := \mathbb{E}_\tau \mathbb{E}_X \left[V^{Pub}(\mu, \alpha) - V^{Pr}(\mu, \alpha) \mid X > \hat{x} \right],$$

where V^{Pr} is S 's total expected payoff under the Private regime, and V^{Pub} is S 's total expected payoff under the Public regime.

With the notation introduced, the condition for the Private regime to dominate ex-ante can be expressed as

$$G(\hat{x})RE - (1 - G(\hat{x}))SE \geq 0.$$

That is, the reputation effect outweighs the screening effect. In fact, all players' ex-ante expected payoffs can be expressed in terms of these two effects. However, it is not as illustrative for the receiver who benefits from more communication regardless of the value of x . As to the ex-ante welfare of the sender, Proposition 6.1 gives the cleanest result which holds if the reputation concerns arise with a substantial probability.

Proposition 6.1. (Sender) *If $G(\frac{1}{2}) \geq \frac{1}{3}$ and $\pi^1 \in [\frac{1}{4}, \frac{1}{2})$, the Private regime yields a strictly higher ex-ante expected payoff for all $y > \min \{\hat{y}_{Pub}, \frac{1}{2}\}$.*

Proof: In the appendix. \square

For all stakes higher than $\min \{\hat{y}_{Pub}, \frac{1}{2}\}$, the reputation effect is large enough to outweigh the screening effect. For stakes lower than this, the benefits from R_B 's reputational concerns are low; the first period has little value to S and he has little to gain even if both types follow his report, but has a lot to lose in the second period if R deceives his trust. If the first period is not valuable, S would optimally use it for screening R , and the Public regime offers him a better environment to do this. To see this, recall that under the Public regime, R_B must randomize her first-stage action. Hence there is a strictly positive probability that S learns R 's type immediately which allows S to design his reporting strategy optimally in

the second period.

The scale of the payoff difference between the reputation and the screening effects is represented in figure 3.3. It visualizes S 's expected payoff at an arbitrarily chosen point $\pi^1 = \frac{2}{5}$ as a function of the stakes x and y . It is assumed that $\frac{2}{5} > \tilde{\pi}_{SR}$ so that the Private regime consists of an equilibrium in pure strategies (characterized in Proposition 5.1). This in turn requires that G is any distribution such that $G(\frac{1}{2}) \geq \frac{2}{3}$.

As is apparent from the figure, the superiority of the Private regime in case R_B has reputational concerns ($x \leq \frac{1}{2}$) greatly exceeds its inferiority in case R_B has myopic concerns. Moreover, since $G(\frac{1}{2}) > \frac{1}{2}$, the ex-ante value of RE is always higher than the ex-ante value of SE . The wedge $(\hat{y}_{Pub}, 2\hat{y}_{Pub}]$ where the Private regime is strictly optimal both when $x \leq \frac{1}{2}$ and when $x > \frac{1}{2}$ (as SE is negative) arises because under the Public regime, when $x > \frac{1}{2}$, the sender would like to lie about θ_1 . Since lying cannot be part of an equilibrium in which y is publicly known, the only equilibrium must consist of babbling which, however, yields S strictly less than if he was able to commit to truthful reporting at $t = 1$.

When $\pi^1 \in [\frac{1}{4}, \tilde{\pi}_{SR})$, the intuition behind the result remains unchanged even though the Private regime consists of an equilibrium in mixed strategies (see Proposition 5.2). In the case where $x \leq \frac{1}{2}$ the Private regime is still superior because it allows R_B to invest in reputation with a higher probability than in the Public regime. Regarding the ex-ante welfare before the realization of y , a few numerical examples where both y and x follow a $\text{Beta}(\alpha, \beta)$ distribution with some, not necessarily the same, parameters α and β , are provided in the appendix (figures 3.6, 3.7, and 3.8).

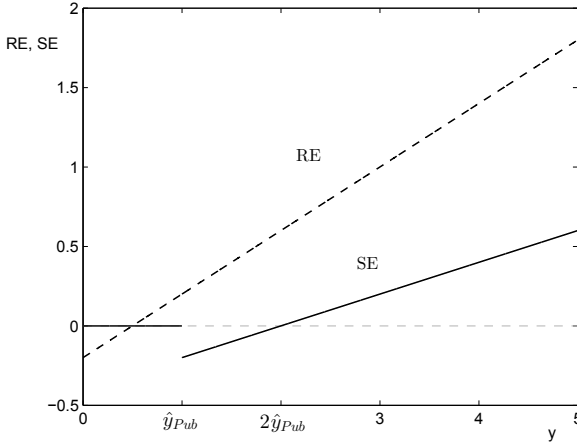


Figure 3.3. Reputation and screening effects as functions of y when $G(\frac{1}{2}) = \frac{2}{3}$ and $\pi^1 = \frac{2}{5}$.

Things are slightly different if R_B has reputational concerns with a less substantial probability, that is, if the first period is likely to be important to R_B . Proposition 6.2 formalizes the welfare result in such a case.

Proposition 6.2. (*Sender*) If $G(\frac{1}{2}) < \frac{1}{3}$, and $\pi^1 \in [\hat{\pi}, \frac{1}{2})$,

(i) if $y \leq \min \{ \hat{y}_{Pub}, \frac{1}{2} \}$: the *Public regime* yields a strictly higher expected payoff, with indifference at $y = \min \{ \hat{y}_{Pub}, \frac{1}{2} \}$;

(ii) if $y \in (\min \{ \hat{y}_{Pub}, \frac{1}{2} \}, \min \{ y^*, \hat{y}_{Pr} \}]$, where

$$y^* = \frac{\pi^1 - (1 - \pi^1)G(\hat{x}_{Pr})}{(1 - 2\pi^1)(1 - 3G(\frac{1}{2})) - 2(1 - \pi^1)(G(\hat{x}_{Pr}) - G(\frac{1}{2}))},$$

the *Private regime* yields a strictly higher expected payoff;

(iii) if $y > \min \{ y^*, \hat{y}_{Pr} \}$, and if

$G(\hat{x}_{Pr}) > Z(\pi^1)$: the *Public regime* yields a strictly higher expected payoff,

$G(\hat{x}_{Pr}) \leq Z(\pi^1)$: the *Private regime* yields a strictly higher expected payoff for all $y > \hat{y}_{Pr}$, where

$$Z(\pi^1) := \frac{1 - 2\pi^1 - (4\pi^1 - 1)G(\frac{1}{2})}{2(1 - \pi^1)}.$$

Proof: In the appendix. \square

The fact that R_B is not likely to have reputational concerns affects the Private regime equilibrium and hence welfare in two notable ways. First, the reputation effect only outweighs the screening effect if S 's stakes are lower than $y^* < \hat{y}_{Pr}$. The reason is that the ex ante probability with which the reputation effect takes place is so low that the ex ante dominance of the Private regime is not enough to compensate for its ex-ante inferiority in case R_B is myopic.

The value of y^* is, however, very different if S 's prior is above $\tilde{\pi}_{IC}$ than what it is for priors lower than this. Recall from proposition 5.1 that for $\pi^1 \geq \tilde{\pi}_{IC}$, $\hat{y}_{Pr} = \bar{y}$ which implies that $H(\hat{y}_{Pr}) = 1$ and therefore $\hat{x}_{Pr} = \frac{1}{2}$. Substituting this into the expression of y^* in the proposition results in a threshold which is increasing in π^1 . Without changes in $G(\frac{1}{2})$, an increase in π^1 implies above all that the dominance of the Public regime in case R_B is myopic decreases since the risk of obtaining a negative expected payoff is lower under both regimes. The Private regime benefits from this more through the continuation payoff of π^1 which S obtains if R turns out to be unbiased and the communication is preserved.

If $\pi^1 < \tilde{\pi}_{IC}$, the cutoffs \hat{x}_{Pr} and \hat{y}_{Pr} are determined simultaneously, and \hat{x}_{Pr} is decreasing in π^1 . While a decrease in π^1 induces a decrease in the numerator of y^* the effect of a decrease in π^1 to the denominator is ambiguous. It increases both of its terms but the total effect remains unclear without a closed-form solution for \hat{x}_{Pr} . In fact, the effect of a decrease in π^1 to y^* depends on how it affects the reputation effect and the screening effect driving the welfare result. The condition for the Private regime to dominate when $\pi^1 \in [\tilde{\pi}, \tilde{\pi}_{IC})$ can be expressed as

$$\begin{aligned} G(\frac{1}{2}) [(1 - 2\pi^1) (2y - 1)] + \left(G(\hat{x}_{Pr}) - G(\frac{1}{2}) \right) [y - (1 - 2\pi^1)] \\ \geq (1 - G(\hat{x}_{Pr})) [(1 - 2\pi^1) y - \pi^1], \end{aligned}$$

where the LHS measures the ex-ante reputation effect, that is, the superiority of the Private regime in case R_B has reputational concerns, and the RHS measures the ex-ante screening effect, that is, how much the Private regime costs to S relative to the Public regime in case R_B has only myopic incentives. Now, the effect of a change in π^1 is a balancing act between three effects. First, a decrease in π^1 increases the superiority of the Private regime in case $x \leq \frac{1}{2}$; both regimes' expected payoffs decrease but they decrease less in the Private regime because S benefits more from R_B 's reputational concerns. Second, it increases \hat{x}_{Pr} thereby increasing the probability $(G(\hat{x}_{Pr}) - G(\frac{1}{2}))$, although at the same time,

the superiority of the Private regime, as measured by the bracketed term after it, decreases. This is because the expected payoffs of the Public regime consist of babbling payoffs which are not affected by changes in π^1 whereas the Private regime payoffs decrease if π^1 decreases. The third effect comes through the RHS where the probability $(1 - G(\hat{x}_{Pr}))$ decreases at the same time as the inferiority of the Private regime increases because S is at a greater risk of incurring costs which he avoids in the Public regime due to babbling. The relative strengths of these effects determine how y^* reacts to changes in the prior.

One numerical result is illustrated in figure 3.4 which visualizes proposition 6.2 for the scenario in which both players' stakes are drawn from a uniform ratio distribution. This implies that $G(\frac{1}{2}) = 0.25$, which produces $\tilde{\pi}_{IC} = \frac{1}{3}$, $\hat{\pi} = 0.25$, and the \hat{y}_{Pr} as plotted in the figure.

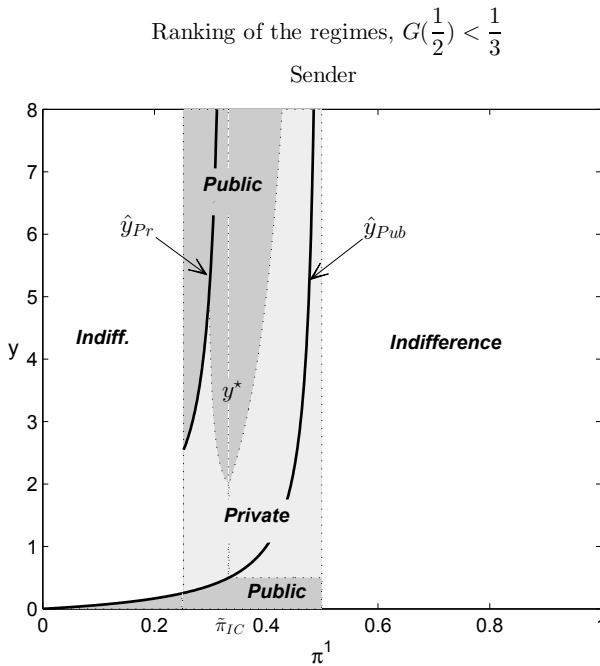


Figure 3.4. Sender's ranking of the regimes

At least in this specification, the threshold y^* is decreasing in S 's prior which means that as \hat{x}_{Pr} increases, it benefits the Private regime relatively more by increasing the frequency of the reputation effect. Moreover, since for this specification $G(\hat{x}_{Pr}) > Z(\pi^1)$, the Public regime strictly dominates for all $y > \min \{y^*, \hat{y}_{Pr}\}$.

The low probability of reputational concerns has another implication for the welfare result through the introduction of \hat{y}_{Pr} . In the Private regime, S lies about θ_1 for all stakes higher than \hat{y}_{Pr} . What the proposition establishes is that S 's ability to lie actually harms him ex ante unless the probability of reputation concerns is very low. Because the Private regime cutoffs for S and R are determined simultaneously, and \hat{x}_{Pr} is decreasing in \hat{y}_{Pr} , S only gets to lie for very high stakes. Therefore, if S lies about θ_1 but R_B actually invests in reputation by following the dishonest message, it is very costly to S . As a result, ex ante, S only gains from lying if R_B misbehaves frequently enough (if the probability $G(\hat{x}_{Pr})$ is small enough). Figure 3.5 in the appendix reproduces the breakdown of S 's welfare in the same manner as in figure 3.4, but in the case where $G(\frac{1}{2}) = \frac{1}{4}$ and $\pi^1 = 0.3$. The losses that S incurs from lying in case R_B actually invests in reputation are so high, and $G(\hat{x}_{Pr})$ is not low enough to mitigate these ex ante, that the babbling equilibrium under the Public regime is superior in terms of welfare.

3.6.2 Receiver's welfare

The general principle guiding the welfare analysis of the receiver is that both receiver types strictly prefer the regime which minimizes the probability of a babbling equilibrium since it always yields them strictly lower payoffs than any informative equilibrium. Proposition 6.3 establishes a result for the higher end of intermediate priors where the welfare comparison is straightforward and unambiguous. This Proposition is enough to make the claim that, for any distribution G , there always exists prior beliefs of the sender under which both receiver types strictly prefer the Private regime.

Proposition 6.3. *(Receiver) If $\pi^1 \in [\tilde{\pi}, \frac{1}{2})$, where $\tilde{\pi}$ is equal to $\tilde{\pi}_{SR}$ if $G(\frac{1}{2}) \geq \frac{1}{3}$, and equal to $\tilde{\pi}_{IC}$ if $G(\frac{1}{2}) < \frac{1}{3}$, the Private regime yields strictly higher ex-ante expected payoffs to both types of R .*

Proof: In the appendix. \square

What makes the Public regime inferior in terms of welfare is that R_B can only invest in reputation partially, by randomizing her first-stage action if $x \leq \frac{1}{2}$. This hurts R_B but not yet R_U . Both types are, however, hurt by the fact that in the Public regime S rewards good behavior with less

than certainty. Notice also that both types of R prefer the Private regime regardless of the value of x . Therefore, the distinction between the reputation and the screening effect is not as valuable for intuition as it is for the sender.

Graphical analysis reveals that the result of Proposition 6.3 holds also for many priors smaller than $\tilde{\pi}$. However, showing this analytically is left for future work.⁶ Figures 3.6, 3.7, and 3.8 in the appendix provide a few examples, though.

3.7 Discussion

The main result of the paper is that under the presence of incomplete information about the payoff-type of the receiver, public information about the stakes is often not welfare improving. The analysis builds on a simple two-period model with binary state, message and action spaces. While this seems a rather stylized setting, the main results of the paper would go through at least in a model where θ remains binary but R 's action y is continuous and all players have a strictly concave utility over y . To translate the assumption of pure conflict in interests into such a model, let there be a constant bias b that measures the divergence in the preferences of S and R_B . As long as b is so large that S would choose to babble with R_B , the results would go through without significant changes. If θ is continuous as well, the presence of stakes y complicates the construction of partition equilibria which require there to exist fixed cutoff types who are indifferent between any two consecutive messages.

Regarding the time horizon, if it was longer than two periods, as in Sobel (1985), the equilibrium cutoffs would be characterized by functions that decrease in time. When the remaining horizon is very long, the continuation value for S from learning that R is unbiased is so high that it incentivizes him to screen R at high stakes even for low priors. However, since the equilibrium cutoffs would be recursive functions of the continuation payoffs, welfare analysis would be a lot more complicated.

Regarding the motivating assumption that R 's payoff-type is fixed over time, this assumption is reasonable only if the two decisions taken up in periods 1 and 2 are similar in nature. It could of course be that S is advising R in two very different issues where R is biased in one but

⁶Results depend mainly on the relationship between $\mu_{P_r}^2(\theta)$ and $\mu_{P_{ub}}^2(\theta)$, that is, between \hat{x}_{P_r} and $\mathbb{E}[X \mid X \leq \hat{x}]$. Without further knowledge about G , it is hard to determine whether or not $\hat{x}_{P_r} > \mathbb{E}[X \mid X \leq \hat{x}]$.

unbiased in the other. However, assuming R 's bias-type to be i.i.d across time would reduce the model to a sequence of one-shot games thereby removing the reputation and learning aspects of the current game. An intermediate approach could be to assume that there is, in every period, a constant probability $p \in [0, 1]$ that R is unbiased. If S is faced by R_U , $p = 1$, and if he is faced by R_B , $p \in (0, 1)$. That is, rather than learning about $\tau \in \{U, B\}$, S is learning about p . This variation would not change the qualitative results of the model. The current model represents a special case with $p = 0$ for R_B .

While private information about R 's stakes makes players better off, private information about S 's stakes is harmful if S 's initial prior is low. In other respects, whether S 's stakes are public or private information has little effect on the social welfare. In general, not much is lost of the analysis if S 's stakes are public information in both regimes, and the only dimension in which the regimes differ is in information about R 's stakes.

Throughout the analysis, I abstract away from the mechanism of making the stakes public knowledge. The paper merely analyzes the difference in ex ante welfare under two different regimes, where the regime is exogenously imposed and pre-existing. A question of its own would be to consider the transition from the Private regime to the Public when beneficial. How much information about the stakes could be transmitted via cheap-talk, by adding a prior stage of communication to the existing model or by allowing S to use multidimensional messages, seems a prominent way forward. In some settings, on the other hand, the disclosure could be obtained via legislation. A case in point is the disclosure of campaign finances by elected politicians which in many countries has been made compulsory by law. Information about the sources and amounts of campaign contributions may serve as a proxy for politicians' stakes in various policy issues. In light of the results from this paper, the reporting of campaign finances may reduce welfare through distortions in the flow of information from lobbyists to politicians.

Finally, an interesting question arising out of the present analysis concerns the optimal sequencing of decisions. Under the Public regime, once the stakes have been realized, and given the equilibrium strategy of R , if S is given the option to choose in which order the two decisions are taken up, how would the optimal solution be characterized? Watson (1999, 2002) has addressed the issue of "starting small" in a framework in which two players with equal stakes play prisoner's dilemma against each other in

continuous time. Starting small is not necessarily optimal when the players' stakes are not perfectly correlated and the sender must balance between riding on the biased receiver's reputation concerns and the benefit from a quick revelation of the receiver's type.

3.8 Conclusion

This paper has analyzed strategic information transmission in a repeated model of communication in which the sender is uncertain about the preferences of the receiver. There are two effects taking place simultaneously. On the one hand, S tries to learn about R 's type by sending her informative messages and observing her actions. Due to the grim-trigger strategy followed by all players, S is more inclined to learn about R via truthful messages than via lying. This alone benefits informative communication and produces ex ante gains to all players. On the other hand, S is wary of the incentives of the biased type of R to invest in reputation by mimicking the action of the unbiased type when her stakes are low. Whether S knows if these reputation concerns exist or not matters for the ex ante welfare of all players. In particular, the higher S 's stakes are, the more he gains if R_B invests in reputation in the current period, by following his message. The ex ante value of these gains is often enough to outweigh the ex ante costs that this trusting strategy incurs to S in case the biased type betrays his trust immediately. The higher frequency of communication that this results to benefits also both types of R . The result suggests that preserving informational asymmetries in organizations is sometimes justifiable on the grounds of higher ex ante welfare which is manifested through an increase in the frequency of transmission of valuable information.

3.9 Appendix: Proofs

Proof of Proposition 4.1. Consider first R_B . Given S 's strategy, investing in reputation is optimal if and only if $-x + V^B(h_+^2) \geq x$, that is,

$$x \leq \frac{1}{2}V^B(h_+^2), \quad (3.4)$$

where $V^B(h_+^2) = 1$. Consider then S . Given R 's strategy, if $x \leq \frac{1}{2}$, honest reporting is optimal if and only if $y + V^S(h_+^2) \geq -y$. Since $V^S(h_+^2) = \pi^1 - (1 - \pi^1) \geq 0$, honest reporting is optimal for all $y \geq 0$. If $x > \frac{1}{2}$, honest reporting is optimal if and only

if

$$[\pi^1 - (1 - \pi^1)] y + \pi^1 V^S(h_+^2) \geq -[\pi^1 - (1 - \pi^1)] y. \quad (3.5)$$

When $\pi^1 \geq \frac{1}{2}$, honest reporting is not only optimal at $t = 1$ but yields also positive continuation payoffs, and therefore S is honest for all $y \geq 0$. \square

Proof of Proposition 4.2. Consider first the case $x \leq \frac{1}{2}$. In the equilibrium, $\alpha_B^1(m, h)$ must satisfy

$$\pi^2(h_+^2) \equiv \frac{\pi^1}{\pi^1 + (1 - \pi^1)\alpha_B^1(m, h)} = \frac{1}{2}. \quad (3.6)$$

To see this, consider that the last equality is replaced by a strict inequality, such that $\pi^2(h_+^2) > \frac{1}{2}$. In that case, S would reward for good behavior with certainty. Since following m_1 would be strictly optimal for all R_B of type $x \leq \frac{1}{2}$, they could increase $\alpha_B^1(m, h)$ slightly without violating the condition $\pi^2(h_+) > \frac{1}{2}$. From condition (3.6) it follows that $\alpha_B^1(m, h) = \frac{\pi^1}{1 - \pi^1}$.

In equilibrium, $\mu_{Pub}^2(\theta, h_+)$ must satisfy the incentive compatibility condition of R_B to randomize between her pure actions at $t = 1$. By replacing $V^B(h_+^2)$ with

$$\mu_{Pub}^2(\theta, h_+) - (1 - \mu_{Pub}^2(\theta, h_+))$$

in (3.4), R_B 's incentive condition holds as equality for each type $x \leq \frac{1}{2}$ if and only if

$$\mu_{Pub}^2(\theta, h_+) = \frac{1}{2} + x.$$

Finally, given R 's strategy, S 's expected payoff from honesty at $t = 1$,

$$[\pi^1 + (1 - \pi^1)\alpha_B^1(m, h) - (1 - \pi^1)(1 - \alpha_B^1(m, h))] y + [\pi^1 + (1 - \pi^1)\alpha_B^1(m, h)] V^S(h_+^2),$$

must exceed his expected payoff from dishonesty,

$$-[\pi^1 + (1 - \pi^1)\alpha_B^1(m, h) - (1 - \pi^1)(1 - \alpha_B^1(m, h))] y.$$

Since $\pi^2 = \frac{1}{2}$, $V^S(h_+^2) = 0$, and honesty is optimal for all $y \geq 0$ iff

$$\pi^1 + (1 - \pi^1)\alpha_B^1(m, h) - (1 - \pi^1)(1 - \alpha_B^1(m, h)) \geq 0,$$

which, after substituting for $\alpha_B^1(m, h) = \frac{\pi^1}{1 - \pi^1}$, holds iff $\pi^1 \geq \frac{1}{4}$.

Consider then the case $x > \frac{1}{2}$. Since $\pi^1 < \frac{1}{2}$, honest reporting yields S an expected loss in period 1 but a continuation payoff of 1 if R is unbiased. S 's incentive compatibility condition for reporting θ_1 honestly is

$$[\pi^1 - (1 - \pi^1)] y + \pi^1 \geq -[\pi^1 - (1 - \pi^1)] y.$$

The condition holds as equality if

$$y = \frac{\pi^1}{2(1 - 2\pi^1)} \equiv \hat{y}_{Pub},$$

and holds as a strict inequality for all $y < \hat{y}_{Pub}$. Finally, in case $x > \frac{1}{2}$ and $y > \hat{y}_{Pub}$, the unique equilibrium must consist of babbling where both types of R randomize their actions independently of S 's reports in both periods. Given R 's strategy, S cannot do better than to randomize his reports independently of the state of the world. \square

Proof of Proposition 5.1. Given S 's strategy, R_B invests in reputation iff

$$-x + 1 \geq -x \Leftrightarrow x \leq \frac{1}{2}.$$

Given the strategy of R , for any θ_1 and y , S reports the first state honestly iff

$$\begin{aligned} & [\pi^1 + (1 - \pi^1)G(\tfrac{1}{2}) - (1 - \pi^1)(1 - G(\tfrac{1}{2}))] y + [\pi^1 + (1 - \pi^1)G(\tfrac{1}{2})] V^S(h_+^2) \geq \\ & - [\pi^1 + (1 - \pi^1)G(\tfrac{1}{2}) - (1 - \pi^1)(1 - G(\tfrac{1}{2}))] y, \end{aligned} \quad (3.7)$$

where the terms including y are identical but of opposite sign. If

$$\pi^1 + (1 - \pi^1)G(\tfrac{1}{2}) - (1 - \pi^1)(1 - G(\tfrac{1}{2})) \geq 0,$$

that is, if $\pi^1 \geq \frac{1 - 2G(\frac{1}{2})}{2(1 - G(\frac{1}{2}))} = \tilde{\pi}_{IC}$, the incentive compatibility conditions holds for all $y \geq 0$ given that $V^S(h_+^2) \geq 0$. Finally, to ensure that S reports θ_2 honestly after a good history of play, we need

$$\pi^2(h_+) := \frac{\pi^1}{\pi^1 + (1 - \pi^1)G(\frac{1}{2})} \geq \frac{1}{2},$$

that is,

$$\pi^1 \geq \frac{G(\frac{1}{2})}{1 + G(\frac{1}{2})} = \tilde{\pi}_{SR}. \quad (3.8)$$

Both $\pi^1 \geq \tilde{\pi}_{IC}$ and $\pi^1 \geq \tilde{\pi}_{SR}$ hold whenever $\pi^1 \geq \tilde{\pi} \equiv \max\{\tilde{\pi}_{IC}, \tilde{\pi}_{SR}\}$. \square

Proof of Proposition 5.2. S is indifferent between honest and false reporting at $t = 2$ iff $\pi^2 = \frac{1}{2}$. From the Bayes' rule, we obtain that this requires that $G(\hat{x}_{Pr}) = \frac{\pi^1}{1 - \pi^1}$, that is, $\hat{x}_{Pr} = G^{-1}\left(\frac{\pi^1}{1 - \pi^1}\right)$. R_B of type \hat{x}_{Pr} is indifferent between following m_1 and diverting from it if and only if

$$-\hat{x}_{Pr} + \mu_{Pr}^2(\theta, h_+) - (1 - \mu_{Pr}^2(\theta, h_+)) = \hat{x}_{Pr},$$

that is, iff $\mu_{P_r}^2(\theta, h_+) = \frac{1}{2} + \hat{x}_{P_r}$. Finally, S reports θ_1 honestly iff

$$\begin{aligned} & [\pi^1 + (1 - \pi^1)G(\hat{x}_{P_r}) - (1 - \pi^1)(1 - G(\hat{x}_{P_r}))] y \\ & + [\pi^1 + (1 - \pi^1)G(\hat{x}_{P_r})] V^S(h_+^2) \geq \\ & - [\pi^1 + (1 - \pi^1)G(\hat{x}_{P_r}) - (1 - \pi^1)(1 - G(\hat{x}_{P_r}))] y, \end{aligned} \quad (3.9)$$

where $G(\hat{x}_{P_r}) = \frac{\pi^1}{1 - \pi^1}$, and $V^S(h_+^2) = 0$ since $\pi^2(h_+) = \frac{1}{2}$. The condition reduces to

$$(4\pi^1 - 1) y \geq - (4\pi^1 - 1) y,$$

which holds for all $y \geq 0$ iff $\pi^1 \geq \frac{1}{4}$. \square

Proof of Proposition 5.3. To see that S 's and R_B 's cutoffs are consistent with each other, fix first \hat{x}_{P_r} and consider S . He reports θ_1 honestly if and only if condition (3.9) holds, where

$$V^S(h_+^2) = \pi^2(h_+) - (1 - \pi^2(h_+)) = \frac{\pi^1 - (1 - \pi^1)G(\hat{x}_{P_r})}{\pi^1 + (1 - \pi^1)G(\hat{x}_{P_r})},$$

where the last equality follows from Bayes' rule. As long as $G(\hat{x}_{P_r}) \leq \frac{1 - 2\pi^1}{2(1 - \pi^1)} \leq \frac{\pi^1}{1 - \pi^1}$, where the last inequality holds for all $\pi^1 \geq \frac{1}{4}$, and which ensures that S 's expected first-stage payoff is negative after honest reporting but expected continuation payoff is positive so that S reports θ_2 honestly after a good history of play, the equilibrium cutoff type \hat{y}_{P_r} is solved from (3.9) to equal

$$\hat{y}_{P_r}(\pi^1, \hat{x}_{P_r}) = \frac{\frac{1}{2} [\pi^1 - (1 - \pi^1)G(\hat{x}_{P_r})]}{[(1 - \pi^1)(1 - 2G(\hat{x}_{P_r})) - \pi^1]}. \quad (3.10)$$

Given S 's cutoff strategy, R_B invests in reputation if and only if

$$H(\hat{y}_{P_r})(-x + 1) + (1 - H(\hat{y}_{P_r}))x \geq H(\hat{y}_{P_r})x - (1 - H(\hat{y}_{P_r}))x,$$

that is, iff

$$x \leq \frac{H(\hat{y}_{P_r})}{2(H(\hat{y}_{P_r}) - 1)} \equiv \hat{x}_{P_r}(\hat{y}_{P_r}). \quad (3.11)$$

The closed-form solutions for the cutoffs are obtained by solving the system of equations consisting of (3.10) and (3.11). Since $\hat{x}_{P_r}(\hat{y}_{P_r})$ is decreasing in \hat{y}_{P_r} and $\hat{y}_{P_r}(\pi^1, \hat{x}_{P_r})$ is increasing in \hat{x}_{P_r} , a solution to the system always exists. Finally, the equilibrium requires that $H(\hat{y}_{P_r}) > \frac{1}{2}$ and $G(\hat{x}_{P_r}) < \frac{\pi^1}{1 - \pi^1}$. Denote by $\hat{\pi}_y$ the lowest prior for which $H(\hat{y}_{P_r}) > \frac{1}{2}$, and by $\hat{\pi}_x$ the lowest priors for which $G(\hat{x}_{P_r}) < \frac{\pi^1}{1 - \pi^1}$. The equilibrium is supported by all $\pi^1 \geq \hat{\pi} \equiv \max\{\hat{\pi}_x, \hat{\pi}_y\}$. \square

Proof of Proposition 5.4. Given that the first-stage equilibrium is separating, S screens R via an honest report if and only if

$$[\pi^1 - (1 - \pi^1)] y + \pi^1 \geq - [\pi^1 - (1 - \pi^1)] y,$$

that is, if and only if

$$y \leq \frac{\pi^1}{2(1-2\pi^1)} \equiv \hat{y}_{Pr}.$$

Given \hat{y}_{Pr} , R_B diverts from all m_1 iff

$$H(\hat{y}_{Pr})x - (1 - H(\hat{y}_{Pr}))x \geq -H(\hat{y}_{Pr})x + (1 - H(\hat{y}_{Pr}))x,$$

that is, iff $H(\hat{y}_{Pr}) > \frac{1}{2}$. An analogous condition applies for R_U to follow all m_1 . If $H(\hat{y}_{Pr}) \leq \frac{1}{2}$, all players would deviate from their proposed strategies, and the only equilibrium must consist of babbling. \square

Proof of Proposition 6.1. Analyze first the interval $\pi^1 \in [\tilde{\pi}_{SR}, \frac{1}{2})$. Consider separately cases (i) $y \leq \hat{y}_{Pub}$, and (ii), $y > \hat{y}_{Pub}$.

Case (i). Private regime maximizes the sender's ex-ante expected payoff if and only if

$$\begin{aligned} \pi^1(y+1) + (1-\pi^1) \left[G\left(\frac{1}{2}\right)(y-1) - \left(1 - G\left(\frac{1}{2}\right)\right)y \right] \geq \\ G\left(\frac{1}{2}\right)(4\pi^1-1)y + \left(1 - G\left(\frac{1}{2}\right)\right) \left[(2\pi^1-1)y + \pi^1 \right] \end{aligned}$$

which, after rearranging, can be shown to hold for all $y \geq \frac{1}{2}$.

Case (ii). Private regime is optimal iff

$$\begin{aligned} \pi^1(y+1) + (1-\pi^1) \left[G\left(\frac{1}{2}\right)(y-1) - \left(1 - G\left(\frac{1}{2}\right)\right)y \right] \geq G\left(\frac{1}{2}\right)(4\pi^1-1)y \\ \Leftrightarrow \left[2 \left(1 - 3G\left(\frac{1}{2}\right)\right) \pi^1 + \left(3G\left(\frac{1}{2}\right) - 1\right) \right] y \geq - \left[\pi^1 - (1-\pi^1)G\left(\frac{1}{2}\right) \right], \end{aligned}$$

which holds for all $y \geq 0$ when $\pi^1 \in [\tilde{\pi}_{SR}, \frac{1}{2})$ since the bracketed terms on the LHS and the RHS are both positive.

Consider then the interval $\pi^1 \in [\frac{1}{4}, \tilde{\pi}_{SR})$. Consider separately cases (i) $y \leq \hat{y}_{Pub}$, and (ii), $y > \hat{y}_{Pub}$.

In case (i), Private regime yields a higher expected payoff iff

$$(4\pi^1-1)y \geq G\left(\frac{1}{2}\right)(4\pi^1-1)y + \left(1 - G\left(\frac{1}{2}\right)\right) \left[(2\pi^1-1)y + \pi^1 \right],$$

where the LHS can be written as a convex combination

$$G\left(\frac{1}{2}\right) \left[(4\pi^1-1)y \right] + \left(1 - G\left(\frac{1}{2}\right)\right) \left[(4\pi^1-1)y \right],$$

such that the inequality reduces to

$$(4\pi^1-1)y \geq (2\pi^1-1)y + \pi^1,$$

which holds if $y \geq \frac{1}{2}$.

In case (ii), Private regime yields S a higher ex ante expected payoff iff

$$(4\pi^1 - 1)y \geq G(\frac{1}{2})(4\pi^1 - 1)y,$$

which holds for all $y \geq 0$. The proposition follows from observing that the Private regime is optimal for all $y > \hat{y}_{Pub}$, and $\hat{y}_{Pub} < \frac{1}{2}$ when $\pi^1 < \frac{1}{3}$. \square

Proof of Proposition 6.2. Consider first $\pi^1 \in [\tilde{\pi}_{IC}, \frac{1}{2})$. Consider separately cases (i) $y \leq \hat{y}_{Pub}$, and (ii), $y > \hat{y}_{Pub}$.

Case (i). Private regime maximizes the ex-ante expected payoff iff

$$\begin{aligned} \pi^1(y+1) + (1-\pi^1) \left[G(\frac{1}{2})(y-1) - (1-G(\frac{1}{2}))y \right] \geq \\ G(\frac{1}{2})(4\pi^1-1)y + (1-G(\frac{1}{2})) \left[(\pi^1 - (1-\pi^1))y + \pi^1 \right] \end{aligned}$$

which, after rearranging, can be shown to hold for all $y \geq \frac{1}{2}$.

Case (ii). Private regime is optimal iff

$$\begin{aligned} \pi^1(y+1) + (1-\pi^1) \left[G(\frac{1}{2})(y-1) - (1-G(\frac{1}{2}))y \right] \geq G(\frac{1}{2})(4\pi^1-1)y \\ \Leftrightarrow \left[(2\pi^1-1)(1-3G(\frac{1}{2})) \right] y \geq - \left[\pi^1 - (1-\pi^1)G(\frac{1}{2}) \right], \end{aligned}$$

where the LHS is negative since $G(\frac{1}{2}) < \frac{1}{3}$. The inequality holds for all $y \leq y^{**}$ where

$$y^{**} = \frac{[\pi^1 - (1-\pi^1)G(\frac{1}{2})]}{[(1-2\pi^1)(1-3G(\frac{1}{2}))]}.$$

Consider then $\pi^1 \in [\hat{\pi}, \tilde{\pi}_{IC})$. Consider separately cases (i) $y \leq \hat{y}_{Pub}$, (ii) $y \in (\hat{y}_{Pub}, \hat{y}_{Pr}]$, (iii) $y > \hat{y}_{Pr}$

Case (i). Under the Private regime, where R_B follows m_1 if $x \leq \hat{x}_{Pr}$, the sender's expected payoff can be written as

$$\begin{aligned} \pi^1(y+1) + (1-\pi^1) \left[G(\frac{1}{2})(y-1) + (G(\hat{x}_{Pr}) - G(\frac{1}{2}))(y-1) - \right. \\ \left. (1-G(\frac{1}{2}))y + (G(\hat{x}_{Pr}) - G(\frac{1}{2}))y \right]. \end{aligned}$$

His expected payoff under the Public regime is given by

$$G(\frac{1}{2})(4\pi^1-1)y + (1-G(\frac{1}{2}))[\pi^1(y+1) - (1-\pi^1)y].$$

After rearranging, the Private regime is shown to dominate iff

$$2[(1-\pi^1)G(\hat{x}_{Pr}) - \pi^1 G(\frac{1}{2})]y \geq (1-\pi^1)G(\hat{x}_{Pr}) - \pi^1 G(\frac{1}{2}),$$

from where we obtain that the Private regime yields higher payoffs iff $y \geq \frac{1}{2}$.

Case (ii). The sender's ex-ante expected payoff under the Private regime is

$$\pi^1(y+1) + (1-\pi^1) \left[G(\hat{x}_{Pr})(y-1) - (1-G(\hat{x}_{Pr}))y \right],$$

and under the Public regime, $G(\frac{1}{2})(4\pi^1 - 1)y$. Rewriting the Private regime payoffs as in case (i) and rearranging terms yields that the payoff under the Private regime is higher if and only if

$$- \left[(1 - 3G(\frac{1}{2})) (1 - 2\pi^1) - 2(1 - \pi^1) (G(\hat{x}_{Pr}) - G(\frac{1}{2})) \right] y \geq \\ - [\pi^1 - (1 - \pi^1)G(\hat{x}_{Pr})],$$

where the term in the brackets on the RHS is positive since in the equilibrium $G(\hat{x}_{Pr}) < \frac{\pi^1}{1-\pi^1}$, and the term in the brackets on the LHS is positive if

$$G(\hat{x}_{Pr}) \leq \frac{(1 - 3G(\frac{1}{2})) (1 - 2\pi^1) + 2(1 - \pi^1)G(\frac{1}{2})}{2(1 - \pi^1)}$$

which holds by the equilibrium requirement $G(\hat{x}_{Pr}) \leq \frac{1-2\pi^1}{2(1-\pi^1)}$. Hence, the Private regime yields a higher ex-ante expected payoff iff $y \leq y^*$, where y^* is given in the Proposition. Notice that for priors, $\pi^1 \geq \tilde{\pi}_{IC}$, $\hat{x}_{Pr} = \frac{1}{2}$, and $y^* = y^{**}$.

Case (iii). The sender's expected payoff under the Private regime is

$$- [\pi^1 + (1 - \pi^1)G(\hat{x}_{Pr}) - (1 - \pi^1)(1 - G(\hat{x}_{Pr}))] y,$$

and under the Public regime,

$$G(\frac{1}{2})(4\pi^1 - 1)y.$$

Payoff under the Private regime is higher if and only if

$$G(\hat{x}_{Pr}) \leq \frac{1 - 2\pi^1 - (4\pi^1 - 1)G(\frac{1}{2})}{2(1 - \pi^1)} := Z(\pi^1).$$

The proposition follows from combining the results of cases (i) through (iii). \square

Proof of Proposition 6.3. Proceed by looking at two cases separately: case (i):

$y \leq \hat{y}_{Pub}$, and case (ii): $y > \hat{y}_{Pub}$.

Case (i). $y \leq \hat{y}_{Pub}$. Consider R_U : Private regime is optimal if

$$E(X) + 1 \geq E(X) + 1 - G(\frac{1}{2})(1 - r_{Pub}),$$

which holds for all $y \geq 0$, and for all $r_{Pub} < 1$. Consider then R_B : Private regime is optimal if

$$G(\frac{1}{2})(-E(X | X \leq \frac{1}{2}) + 1) + (1 - G(\frac{1}{2}))E(X | X > \frac{1}{2}) \geq \\ G(\frac{1}{2})[\mu(-E(X | X \leq \frac{1}{2}) + r_{Pub}) + (1 - \mu)E(X | X \leq \frac{1}{2})] + (1 - G(\frac{1}{2}))E(X | X > \frac{1}{2}),$$

which can be reduced to

$$1 - E(X | X \leq \frac{1}{2}) \geq \mu(-E(X | X \leq \frac{1}{2}) + r_{Pub}) + (1 - \mu)E(X | X \leq \frac{1}{2}),$$

where $r_{Pub} = \frac{1}{2} + E(X | X \leq \frac{1}{2})$ so that we are left with

$$1 - E(X | X \leq \frac{1}{2}) \geq \frac{1}{2}\mu + (1 - \mu) E(X | X \leq \frac{1}{2}).$$

The LHS $\geq \frac{1}{2}$ and the RHS $\leq \frac{1}{2}$ so the condition holds.

Case (ii). $y > \hat{y}_{Pub}$. Consider first R_U : Private regime is optimal if

$$E(X) + 1 \geq G(\frac{1}{2}) (E(X | X \leq \frac{1}{2}) + r_{Pub}),$$

which holds for all distributions G since the Public regime yields 0 with probability $(1 - G(\frac{1}{2}))$.

Consider then R_B . Private regime is optimal if

$$G(\frac{1}{2}) [-E(X | X \leq \frac{1}{2}) + 1] + (1 - G(\frac{1}{2})) E(X | X > \frac{1}{2}) \geq G(\frac{1}{2}) [\mu (-E(X | X \leq \frac{1}{2}) + r_{Pub}) + (1 - \mu) E(X | X \leq \frac{1}{2})].$$

It was shown under case 1 that R_B obtains higher expected payoffs under the Private regime in case $x \leq \frac{1}{2}$. Since the Public regime yields 0 in case $x > \frac{1}{2}$, the Private regime is clearly strictly preferred. The Proposition follows from combining results from cases (i) and (ii). \square

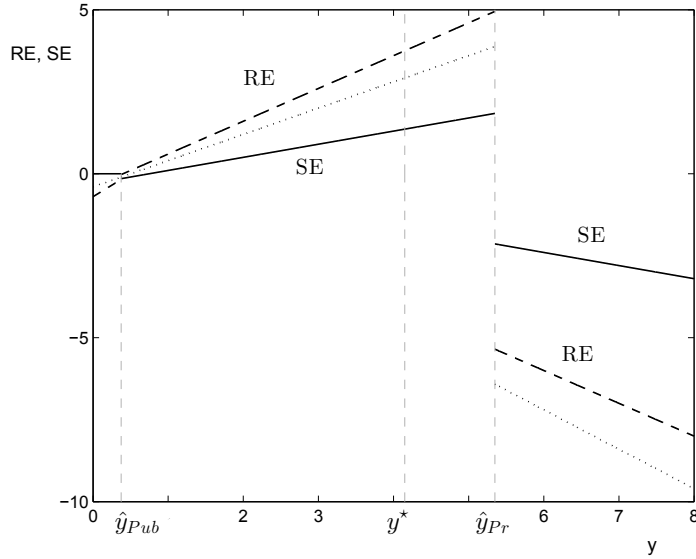


Figure 3.5. Reputation and screening effects as a function of y when $G(\frac{1}{2}) = \frac{1}{4}$ and $\pi^1 = 0.3$. Dashed line: $x \in (\frac{1}{2}, \hat{x}_{Pr}]$, dotted line: $x \leq \frac{1}{2}$.

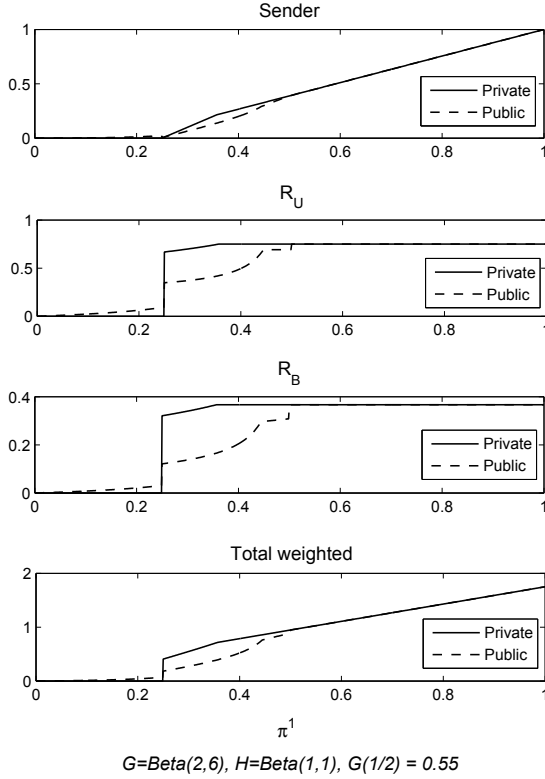


Figure 3.6. Ex-ante expected welfare, $X \sim \text{Beta}(2, 6)$, $Y \sim \text{Beta}(1, 1)$

Figure 3.6 plots the ex-ante expected payoffs of all players separately and the sum of payoffs weighted by π^1 . It provides an example in which $G(\frac{1}{2}) \geq \frac{1}{3}$. If the distribution of y is heavily skewed to the right, such that $H(\hat{y}_{Pr}) > \frac{1}{2}$, where $\hat{y}_{Pr} = \frac{\pi^1}{2(1-2\pi^1)}$, then informative communication can still be sustained under the Private regime for priors less than $\frac{1}{4}$. In that case, S is able to fool R_B by lying if $y > \hat{y}_{Pr}$ which benefits S at the expense of both types of R . This benefit is not very sizable, though, given the fairly low probability that R_B actually deviates from S 's dishonest message. All players still get higher expected payoffs than under babbling due to the high probability that S is actually truthful.

Figure 3.7 represents one possible situation. Since the probability mass in H is so concentrated on low values of y , the Public regime ex ante dominates for the sender for all intermediate priors (Proposition 6.1). To the other direction, if S 's stakes are very high on expectation it benefits the Private regime because the probability of babbling increases in the Public regime due to $\Pr[Y \leq \hat{y}_{Pub}]$ being low.

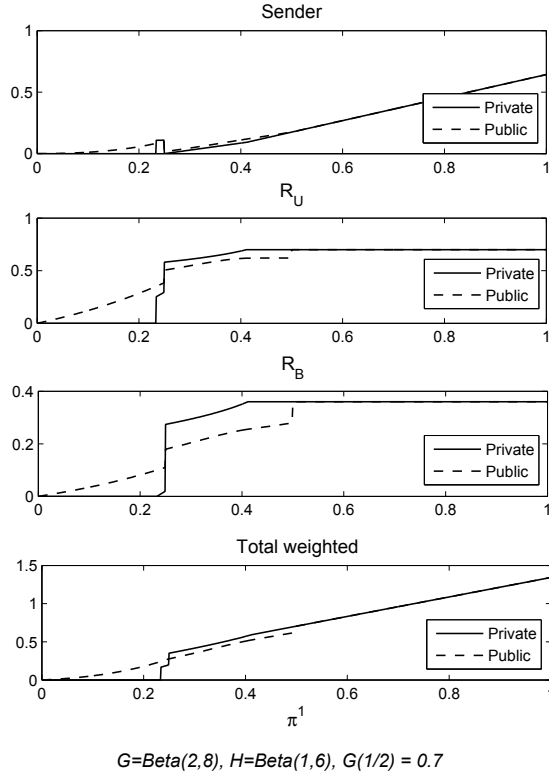


Figure 3.7. Ex-ante expected welfare, $X \sim \text{Beta}(2,8)$, $Y \sim \text{Beta}(1,6)$

Ex ante welfare in case $G(\frac{1}{2}) < \frac{1}{3}$

Figure 3.8 plots the ex ante welfare for each player when each player's stakes are drawn from a uniform ratio distribution. The support of the distribution is \mathbb{R}_+ which this allows also for very high realizations of Y . This largely explains why S obtains higher welfare under the Public regime for all $\pi^1 \in [\frac{1}{4}, \frac{1}{2})$. Restricting the support on a finite interval, for example to $[0, 10]$ alleviates the difference, and at least if $y \in [0, 2]$, the Private regime is strictly optimal also for the sender.

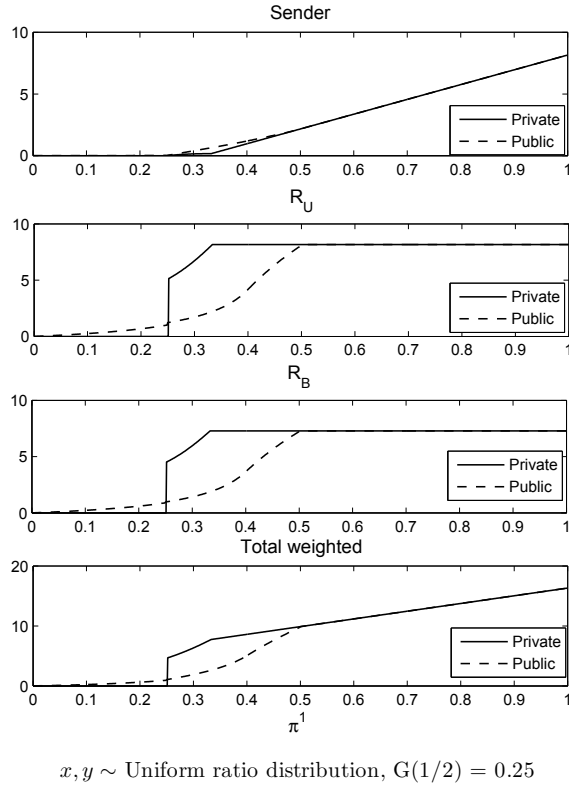


Figure 3.8. Ex-ante expected welfare, $X, Y \sim \text{Uniform ratio distribution } [0, 1000]$

References

- Bar-Isaac, H. (2012) "Transparency, Career Concerns, and Incentives for Acquiring Expertise". *The B.E. Journal of Theoretical Economics*, 12 (1), 1-13.
- Bénabou, R., and G. Laroque (1992) "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility". *Quarterly Journal of Economics*, 107, 921-58.
- Blonski, M. and D. Probst (2004) "The Emergence of Trust". unpublished manuscript, Goethe University, Frankfurt am Main.
- Chen, Y. (2012) "Value of public information in sender-receiver games". *Economics Letters*, 114, 343-345.
- Chen, Y. (2009) "Communication with Two-sided Asymmetric Information". mimeo, Arizona State University.
- Crawford, V., and J. Sobel (1982) "Strategic Information Transmission". *Econometrica*, 50 (6), 1431-1451.
- Durbin, E., and G. Iyer (2009) "Corruptible Advice". *American Economic Journal: Microeconomics*, 1 (2), 220-242.
- Ely, J., and J. Välimäki (2003) "Bad Reputation". *Quarterly Journal of Economics*, 118 (3), 785-814.
- Frisell, L., and J. Lagerlöf (2007) "A model of Reputation in Cheap Talk". *Scandinavian Journal of Economics*, 109 (1), 49-70.
- Golosov, M., V. Skreta, A. Tsyvinski, and A. Wilson (2014) "Dynamic Strategic Information Transmission". *Journal of Economic Theory*, 151, 304-341.
- Guembel, A., and S. Rossetto (2009) "Reputational Cheap Talk with Misunderstanding". *Games and Economic Behavior*, 67 (2), 736-744.
- Ishida, J., and T. Shimizu (2012) "Can more information facilitate communication?". ISER Discussion Paper No. 839.
- Kreps, D., and R. Wilson (1982) "Reputation and Imperfect Information". *Journal of Economic Theory*, 27 (2), 253-279.
- Lai, E. (2013) "Expert Advice for Amateurs". mimeo, Lehigh University, January, 2013.

- Levy, G. (2007) "Decision-Making Procedures for Committees of Careerist Experts". *American Economic Review*, 97 (2) 306-310.
- Li, M., and K. Madarász (2008) "When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests". *Journal of Economic Theory*, 139, 47-74.
- Li, M., and T. Mylovanov (2008) "Credibility for Sale - the Effect of Disclosure on Information Acquisition and Transmission". Concordia University Working Paper no. 09-008.
- Matozzi, A., and A. Merlo (2007) "The Transparency of Politics and the Quality of Politicians". *American Economic Review*, 97 (2), 311-315.
- Milgrom, P., and J. Roberts (1982) "Predation, Reputation, and Entry Deterrence". *Journal of Economic Theory*, 27 (2), 280-312.
- Moreno de Barreda, I. (2010) "Cheap Talk with Two-Sided Private Information". mimeo, London School of Economics and STICERD.
- Morgan, J., and P. Stocken (2003) "An Analysis of Stock Recommendations". *RAND Journal of Economics*, 34 (1), 183-203.
- Morris, S. (2001) "Political Correctness". *Journal of Political Economy*, 109, 231-265.
- Ottaviani, M., and P. Sørensen (2006) "Reputational Cheap Talk". *RAND Journal of Economics*, 37 (1), 155-175.
- Prat, A. (2005) "The Wrong Kind of Transparency". *American Economic Review*, 95 (3), 862-877.
- Renault, J., E. Solan, and N. Vieille (2013) "Dynamic Sender-Receiver Games". *Journal of Economic Theory*, 148 (2), 502-534.
- Sobel, J. (1985) "A Theory of Credibility". *Review of Economic Studies*, 52 (4), 557-573.
- Watson, J. (1996) "Information Transmission When the Informed Party is Confused". *Games and Economic Behavior*, 12, 143-161.
- Watson, J. (1999) "Starting Small and Renegotiation". *Journal of Economic Theory*, 85 (1), 52-90.
- Watson, J. (2002) "Starting Small and Commitment". *Games and Economic Behavior*, 38, 176-199.

Wolinsky, A. (2003) "Information transmission when the sender's preferences are uncertain". *Games and Economic Behavior*, 42, 319–326.

Wrasai, P., and O. Swank (2007) "Policy makers, advisers, and reputation". *Journal of Econ. Behavior and Organization*, 62, 579-90.

4. Reputation and the Value of Information in a Trust Game

4.1 Introduction

Many transactions between economic actors, whether between a firm and its employees, between a firm and its suppliers, or between a buyer and a seller, require some trust to realize. A firm would not commit to pay wages to employees who are anticipated to exert no effort; similarly, no employee would choose to work for a firm that has a reputation of reneging on its wage payments; furthermore, no buyer would trade with an online seller who is believed to offer low quality for a high price. In general, where a mutually beneficial cooperative outcome requires one party to take an action that the opponent may exploit, a minimal level of trust must exist for cooperation to arise.

As in the examples mentioned, the principal must often take a potentially costly action without being certain if it will pay off. Because the agent who moves second and is assumed to be strategic will betray trust if given the opportunity, cooperation unravels in a finite game unless the agent's behavior is disciplined via some form of external control.¹ In many relationships, such as the one between an Investor and an Entrepreneur analyzed in this paper, it is not possible to write formal contracts. Instead, the agent is motivated by reputational concerns. If a fraction of agents are honest, in a sense of always rewarding trust, then a strategic agent may want to be perceived as trustworthy. This way, the strategic type is able to ensure a positive payoff by building up a reputation for being the non-strategic type and thereby reassuring the principal to keep investing.

¹In an *infinite* game this problem would be resolved, though, provided that the agent is patient enough.

Reputation building by the strategic type is welfare enhancing since it encourages more trade. In a finitely repeated game, this requires that the agent who invests in reputation find it worthwhile to forgo high short-term profits in return for a continued interaction. Thus, the agent's choice between an immediate exploitation of the principal and a postponed exploitation depends on his current and future payoffs. Suppose that this payoff information may or may not be known to the principal. It is then natural to ask how the probability of trade depends on the information available to the principal. Is welfare increasing in the information that the principal holds about the agent? If the principal could control the agent and monitor his payoffs, or downright spy him, would she be better off?

To address these questions, this paper studies two versions of a twice-repeated, binary trust game. The versions differ in their informational assumptions concerning the type of the agent. In the stage-game, an Investor first decides whether to invest or not an amount x to a project led by an Entrepreneur. In the hands of the Entrepreneur, an investment worth of x translates into an output worth of Mx , with $M \geq 0$. After the production, the Entrepreneur decides whether to honor trust by returning x with a fixed interest rate r , or to abuse trust by keeping all of the proceeds to himself.

The Entrepreneur is privately informed about his payoff-type. A reliable Entrepreneur is nonstrategic and always returns the investment regardless of the realized output. Assume for example that his payoffs are driven by reciprocity or fairness which imply high psychological costs from deceiving someone. Alternatively, this may be a reflection of the probability of another trading opportunity with another principal, say in the case where past repayments are a matter of public record. The reliable agent is one that has always a sufficiently large continuation surplus to make exploitation excessively costly. An unreliable Entrepreneur is strategic and only returns an investment if it is optimal for his total payoff. Therefore, in a one-shot game, the unreliable Entrepreneur would always abuse trust, and, anticipating this, the Investor would invest only if she has a high enough belief that the Entrepreneur is reliable. In a repeated game, however, the unreliable type sometimes repays the Investor in the early periods so as to mimic the reliable type's behavior and be perceived as one. This reputation building is optimal only if the discounted future cost for the Entrepreneur from the punishment that ensues if trust is not re-

turned is larger than the immediate gain from abusing trust.

The value of the future from the perspective of the Entrepreneur is determined by δM where δ is a common discount factor. For high enough M , the unreliable type finds it worthwhile to mimic the reliable type and return the investment in the first period. The value of his project is, however, assumed to be stochastic, drawn by Nature at the beginning of the game. Two alternative assumptions can be imposed on the observability of M . First, M may be publicly observed. This is the assumption applied generally in experimental studies of the trust game (where usually M is deterministic and equal to 3).

On the other hand, it is rather easy to assume that the Entrepreneur is better informed about the profitability of his project, or that the profitability is determined only after the investment has been made². Therefore, the Investor has to make her investment decisions based on imperfect information about M . This would not pose a problem if the Entrepreneur was known to be reliable. The unreliable type, however, repays an investment if his business is profitable enough. Therefore, not observing M exposes the Investor to an additional uncertainty concerning the equilibrium action of the unreliable type of Entrepreneur.

While trust games with incomplete information about the type of the Entrepreneur have been subject to numerous studies in the literature, they all deal with situations in which M is common knowledge, that is, the Entrepreneur's type is one-dimensional. The main contribution of this paper is to formalize a variant of the game in which the agent's payoffs are his private information and to compare the ex-ante expected equilibrium payoffs across the two models. The main result from this analysis is that, for any distributional assumptions about M , and for any parametrizations of r and δ , there always exists a non-empty set of prior beliefs of the Investor about the reliability of the Entrepreneur for which all players are ex-ante strictly better off if and only if M is private information of the Entrepreneur.

The result means that less information about the Entrepreneur gives rise to more trust ex-ante. More trust, in turn, benefits all players, even the Investor. In fact, under incomplete information, more trust hurts the Investor if the bad Entrepreneur is only concerned of immediate gains. Naturally, the lower the prior belief of the Investor, the more risky it is

²though still remaining private information about the Entrepreneur. Note also that x and M are assumed to be independent.

to trust the Entrepreneur. However, this ex-ante exposure to risk is compensated by the ex-ante possibility that the bad type has reputational concerns. If this is the case, then more trust is beneficial to the extent that these ex-ante expected benefits outweigh the ex-ante expected costs from being exploited immediately.

The result would suggest that the cost of control in organizations (for instance, the monitoring of employees' workload which determines the performance of the unreliable worker types) may not only be borne by a decreased intrinsic motivation of the agent who feels being under surveillance, but may also be borne by the reduction of trust, such as delegation, between the principal and the agent. To increase trusting behavior, it helps if the principal remains ignorant about an aspect related to the agent which determines whether the strategic type is more concerned of the future of cooperation or of the immediate gains. To return to the model studied in this paper, from an ex-ante welfare point of view, the owner of a start-up may do a favor not only to himself but also to the Investor if his business plan and revenue forecast is not very detailed.

The trust game is stylized enough to be readily applied to varying settings. An example of the importance of trust is the case of Lincoln Electric (see e.g. Miller, 2001: 316). The firm has publicly committed to no wage cuts, which has increased the workers' trust in their employment and increased their motivation to put voluntary effort in their work. In light of the results of this paper, if there is less than full trust among the employees about the ability or willingness of the firm to not cut their wages or lay off workers, it may be advisable that Lincoln Electric keep its employees on their toes by not releasing them information about the firm's financial condition. Through the increased trust that this induces, the employees give the firm a chance to prove its loyalty to them, and thus the unreliable firm has more room to mimic the reliable type which in the end benefits everyone.

The outline of the paper is as follows. Section 4.2 takes a look at the relevant literature. Section 4.3 formalizes the model with common knowledge about M and solves for its equilibrium. Section 4.4 solves for the equilibrium of the model with private information about M . Section 4.5 formalizes the main results of the paper which are obtained by comparing the ex-ante welfare of players across the two models. Section 4.6 discusses extensions and outlines ideas for further research, and section 4.7 concludes.

4.2 Literature

This paper is directly related to the existing literature on trust games, studied both in the theoretical and experimental literature about trust and reciprocity. In addition, the stylized model of the paper is indirectly related to a variety of more applied frameworks where trust plays a role, such as relational contracting and delegation in agency settings. In addition, the key feature in this paper, the comparison of different information structures regarding the value of the Entrepreneur's project, can be linked to the literature on transparency in principal-agent models.

A vast literature in behavioral economics has studied variants of the trust game between a Trustor and a Trustee in the lab³. Camerer (2003) offers a comprehensive coverage of these studies which have consistently found that, against theoretical predictions, people tend to reciprocate trust even in one-shot interactions. More importantly, numerous studies⁴ have investigated the behavior of subjects in trust games with incomplete information about the Trustee, both under static and under repeated interactions. In the latter case, the focus is on reputation building in finite games that build on the seminal work by Kreps and Wilson (1982). Generally, trust is maintained for some periods in the beginning of the game but it declines as the game approaches its end.

A somewhat different theoretical contribution to the literature is that of Colombo & Merzoni (2006) who consider a two-period game with incomplete information, but their focus is on the length of contracts: they show that sometimes trust requires that the principal commits to a two-period game.

The game analyzed in this paper is a version of the loan model of Sobel (1985). However, he assumes that the Investor chooses the level of investment which directly determines players' stage-game payoffs. The payoff uncertainty which is key to the current paper is thus not present. Sobel shows that the amount invested in equilibrium increases with each successful loan, which is consistent with findings from other papers studying the phenomenon of "starting small" (see. e.g. Watson, 1999, 2002).

All of the aforementioned papers deal with a setting where the only uncertainty concerns whether the Trustee is strategic or not. The novelty of the current paper is to analyze a trust game where the Trustee's type

³see e.g. Cocharde et al. (2004) for a repeated trust game with perfect information

⁴see e.g. Anderhub et al. (2002), Neral & Ochs (1992), Camerer & Weigelt (1988), Brandts & Figueras (2001)

is two-dimensional, and study how this scenario compares in terms of expected welfare to the more standard version of one-dimensional types.

Trust-games are readily transformed to analyze various worker-firm settings in which a worker must trust the firm in paying the bonuses it has promised in exchange for high effort, or the firm must trust the worker to exert the effort required after a hiring decision. This takes the model closer to the burgeoning literature on relational contracts, studied recently among others by Levin (2003) and Halac (2014). The results of my paper would suggest a new angle to explore in the theory of relational contracts. Relational contracting between a principal and an agent with a fixed, unknown type have been studied for instance by Halac (2014), Watson (1999, 2002), Yang (2013), and Lukas & Schöndube (2012). In all of them, uncertainty concerns the fixed strategy type of the agent. In Lukas and Schöndube (2012), for instance, the agent has some initial trust in the firm, denoted by γ , which gives the probability that the firm is a non-strategic type who always pays a bonus. With complementary probability the firm is strategic and only pays a bonus when it is optimal. Whether the welfare results of the current paper continue to hold in such settings may be worthwhile to explore formally.

As to models of relational contracting with uncertainty about the pay-offs, a recent paper by Li and Matouschek (2013)⁵ considers a firm and an employee who interact infinitely, and the firm has private information about its opportunity costs to pay a bonus to the worker. More specifically, the firm sometimes prefers an alternative use for the funds intended for bonus payments, such as an exceptional investment opportunity. The firm is known to be strategic, and what causes conflict is that since the worker does not observe whether the firm is hit by a shock or not, he does not know if the nonpayment of a bonus was warranted or not. This imperfect monitoring, and the fact that paying the bonus is efficient in some periods, also makes possible the gradual decline in trust instead of a sudden termination of the relationship. In contrast to the setting in my model, the setting of Li and Matouschek is not concerned with learning about the firm. In addition, they assume that the firm's opportunity costs, and hence payoffs, vary over time.

The question of whether the value of the project should be publicly observed or not relates to theoretical research conducted on transparency in

⁵see also a related paper by Englmaier & Segal (2013) in which the worker's choice is binary, and the firm's opportunity cost of paying the worker in a bad state of the world is infinite.

principal-agent settings. In previous literature (e.g. Prat, 2005; Matozzi & Merlo, 2007, Levy, 2007a,b), transparency has been coined as the ability of the principal to either observe how the agent behaves and/or what the consequences of such behavior are. I consider a somewhat different version of transparency, namely the ability of the principal to observe one dimension of the agent's type.

Literature on delegation and authority⁶ in organizations is relevant in terms of applications in that the decision of a principal to delegate a task to an agent of unknown type is essentially analogous to the principal deciding to trust the agent in completing a task on her behalf. Literature on for instance delegated portfolio management⁷, and delegation models in political science⁸ provide interesting applied settings in which the results of this paper could be applied. As an example, it may be socially optimal to have the compensation structure of a portfolio manager concealed from investors.

4.3 Public information about M

This section introduces the model in which the Entrepreneur is privately informed of his reliability, but the value of his project is common knowledge. In addition, both players discount the future at the rate of δ . In the short game that we consider, though, discounting does not really play a role, and one could as well set $\delta = 1$. A δ less than 1 can be interpreted as there being a risk that the interaction stops after the first period due to sudden changes in either the Investor's or the Entrepreneur's environment. For instance, Entrepreneur may not need financing anymore after one round, perhaps due to a bequest from a suddenly deceased relative, or due to health problems leading to an early retirement. In what follows, the model of this section is sometimes referred to as 'Public regime', and the terms investment and loan are used interchangeably despite subtle differences in their etymologies.

⁶see e.g. Baker et al (1999), Marino et al. (2010), Bester & Kräbmer (2008), Armstrong & Vickers (2010)

⁷see e.g. Huberman & Kandel (1993), Huddart (1999), Chevalier & Ellison (1999), Stracca (2006)

⁸see e.g. survey by Bendor et al. (2000)

4.3.1 The Model

There are two players, an Investor and an Entrepreneur who interact in periods $t = 1, 2$, to play a stage-game which is identical over time⁹. At the beginning of the game, Nature moves and draws once and for all the Entrepreneur's type (τ, M) .¹⁰ The type vector consists of his strategy-type, $\tau \in \{g, b\}$, and the value of his project, M . The strategy type τ is privately observed by the Entrepreneur. With a common prior probability $p \in (0, 1)$ the Entrepreneur is reliable, or "good", g , and with the complementary probability he is unreliable, or "bad, b ". A good Entrepreneur is non-strategic and always repays investments. Therefore, with a good type, the Investor would always choose to invest. A bad Entrepreneur is strategic: he maximizes his total payoff which is given by the sum of stage-game payoffs. The value of the project, M , is publicly observed. For the purposes of ex-ante welfare, discussed in section 4.5, let M be drawn from a commonly known continuous¹¹ distribution F with a support $\mathcal{M} \subset \mathbb{R}_+$. After Nature's move, the following stage-game is repeated twice.

1. Investor first chooses whether to "Invest, I " or "Not invest, NI " a fixed amount $x \geq 0$ to the Entrepreneur's project. For simplicity, normalize x to 1. If her action is "Not invest", the stage game ends.
2. If an investment is made, the Entrepreneur can either "Repay, R " it or "Default, D " on it.
3. Payoffs are realized and privately observed¹².

The payoffs are displayed in Figure 4.1 which sketches the stage-game between the Investor and the bad Entrepreneur. If the Investor chooses

⁹the stage game is an adaptation of the trust game introduced by Berg et al. (1995)

¹⁰Since the Entrepreneur does not have the choice to turn down investments, it may be more reasonable to have M drawn only after an investment is made. This would not change the dynamics of the game as long as $\mathbb{E}(M) > 1 + r$.

¹¹This is required for the construction of the Private regime equilibrium for priors in the range $[\hat{p}, \tilde{p}_{SR})$. Were F discrete, the determination of \hat{M}_{adj} would not be possible. In such a case, the equilibrium would consist of the bad type betraying trust for all values of M . This would result in changes in the lower bound \hat{p} .

¹²This is crucial only in the Private regime where the Investor would be able to learn that the Entrepreneur is good if he repays a loan when $M < \hat{M}$.

not to invest, all players get a payoff of 0, that is, no one's initial wealth is affected. If the Investor chooses to invest and the Entrepreneur repays it the Investor gets a net payoff of $r > 0$ which is the interest accrued to the investment. The Entrepreneur gets the output less the repayment of the investment, $M - (1 + r)$. If the Entrepreneur defaults the Investor simply loses her investment and gets a payoff of -1 . The bad type of Entrepreneur obtains a payoff of M whereas the good type incurs an infinitely large cost from betraying the Investor. All players maximize the sum of stage-game payoffs, discounted by the commonly known discount factor $\delta \in (0, 1]$.

Denote the behavior strategy of the Investor in period t by $\sigma_{It} : H_t \rightarrow [0, 1]$ such that $\sigma_I(h_t)$ gives the probability that she invests in period t given the realized history of play. Similarly, denote the behavior strategy of the Entrepreneur of type $\tau \in \{g, b\}$ by $\sigma_{\tau t} : H_t \rightarrow [0, 1]$ such that $\sigma_{\tau}(h_t)$ gives the probability that he repays an investment made in period t .

Notice that if the realization M is less than $1 + r$, the good type already knows upon entering the game that his payoff will be negative. While it is reasonable to assume that in such environments he may choose to not seek financing, the possibility of declining loans is not considered here. In fact, it is not uncommon that long-term projects yield deficits in their early stages. In a longer game, this issue may need to be addressed for example by assuming that $\delta < 1$ and F is supported on $M \geq 1 + r$, or by introducing some dynamics to M . In the current setup where the Entrepreneur cannot decline investments, if it occurs that $M < 1 + r$, the good Entrepreneur would prefer the Investor to not invest. That is, for some combinations of r, δ and F the good type prefers an equilibrium which minimizes the frequency of investments. This will be reflected in the welfare results discussed in section 4.5.

One-shot game

Suppose a game of perfect information between an Investor and the bad Entrepreneur. Being the second mover, a bad Entrepreneur has a dominant strategy in a one-shot game to default. Anticipating this, the Investor will not invest, and the only sub-game perfect Nash equilibrium of the game with perfect information is $\{NI, D\}$. Notice, however, that both players would prefer the cooperative outcome $\{I, R\}$ to the no-trust outcome.

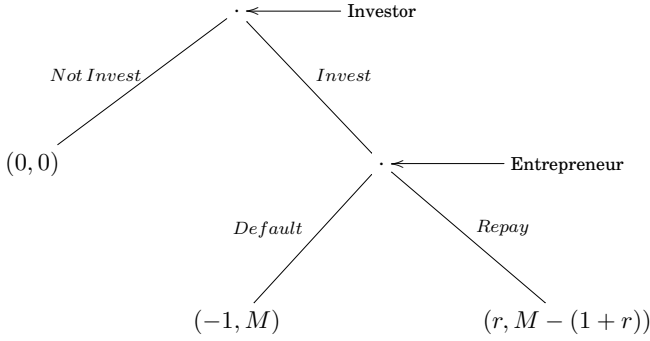


Figure 4.1. One-shot game

In a game with imperfect information about the reliability of the Entrepreneur, the one-shot game has a Bayesian Nash equilibrium in which the Investor invests if she is confident enough that the Entrepreneur is good.

Proposition 1. *In a one-shot game, the Investor invests if and only if her belief about the Entrepreneur being good satisfies*

$$p \geq \frac{1}{1+r}.$$

4.3.2 Equilibrium

Assume that the Investor plays according to the following trigger strategy:

In the first period, play “invest”. Thereafter, if all moves in all previous periods have been “invest” and “repay”, play “invest”; otherwise, play “not invest”.

Given the assumption on the strategy of the Investor, this section analyzes when it is supported as part of an equilibrium of the game. In a Perfect Bayesian Equilibrium (PBE) of the game, no player wants to deviate from his or her equilibrium strategy after any possible history of the game, given the equilibrium strategies of other players. In addition, beliefs held by the Investor about the payoff type of the Entrepreneur are updated according to Bayes rule whenever possible.

Second period: Solving backwards, in the second and last period the Investor invests if and only if $p_2(h_2) \geq \frac{1}{1+r}$, where $h_2 \in H_2 \equiv \{NI, (I, D), (I, R)\}$. More specifically, given that the good Entrepreneur always repays the Investor, after a history including a default, $p_2(I, D) = 0$ leading to $\sigma_{I2}(I, D) = 0$. If the Investor has chosen to not invest in the first project, $p_2(NI) = p$, and $\sigma_{I2}(NI) = 1$ if $p \geq \frac{1}{1+r}$. In case the Entrepreneur has repaid the Investor, the revised beliefs about the Entrepreneur's type depend on the equilibrium first-period strategy of the bad Entrepreneur, to which we turn next.

First period: When choosing their actions in the first period, both the Investor and the Entrepreneur take into account the value of M which determines the equilibrium first-period action of the bad Entrepreneur. If the business is profitable enough, the bad type has a reputational incentive to repay the first-period loan so as to convince the Investor of his trustworthiness. Suppose that the Investor continues investing as long as the Entrepreneur has repaid earlier loans. For any δ , the bad Entrepreneur chooses to repay the first-period loan if and only if $M - (1+r) + \delta M > M$, that is, if

$$M > \frac{1+r}{\delta} \equiv \hat{M}.$$

Let $M > \hat{M}$. If $\sigma_{b1}(M) = 1$, the Investor's posterior $p_2(I, R) = p$ and repayment leads to a new investment only if $p \geq \frac{1}{1+r}$. Any lower belief than this contradicts $\sigma_{b1}(M) = 1$ and hence it must be that $\sigma_{b1}(M) < 1$.

Assume that $0 < \sigma_{b1}(M) < 1$. If this leads to $\sigma_{I2}(I, R) = 1$, then $\sigma_{b1}(M) = 1$ would be optimal, a contradiction. Thus, the requirement for $0 < \sigma_{b1}(M) < 1$ is that $0 < \sigma_{I2}(I, R) < 1$. This means that the Investor must be indifferent between investing or not in period 2, that is, $p_2(I, R) = \frac{1}{1+r}$. Hence, for all priors $p < \frac{1}{1+r}$, $\sigma_{b1}(M)$ must satisfy

$$\frac{p}{p + (1-p)\sigma_{b1}(M)} = \frac{1}{1+r},$$

where the LHS gives the updated belief of the Investor that the Entrepreneur is good after a repayment of the first loan. Moreover, the proposed mixed strategy for the Entrepreneur requires him to be indifferent between repaying or not, given the proposed strategy of the Investor. That is,

$$M - (1+r) + \delta\sigma_{I2}(I, R)M = M,$$

from where we can solve for $\sigma_{I2}(I, R)$. The following proposition summarizes the equilibrium strategies of the Investor and the bad Entrepreneur.

Proposition 2. *Let $M > \hat{M}$.*

- *If $p \geq \frac{1}{1+r}$, the Investor invests in the first period and continues investing as long as investments are repaid. The equilibrium of the first stage game is pooling: the bad type of Entrepreneur repays the first investment with probability 1.*
- *If $p \in [\frac{1}{r^2+2r+1}, \frac{1}{1+r})$, the Investor invests in the first period and, if the investment is repaid, invests in period 2 with probability $\sigma_{I2}(h_2) = \frac{1+r}{\delta M}$. The equilibrium of the first stage game is semi-separating: the bad type of Entrepreneur repays the first investment with probability $\sigma_{b1}(M) = \frac{rp}{1-p}$.*
- *If $p < \frac{1}{r^2+2r+1}$, the unique equilibrium of the game consists of no investments being made. If an investment is made, the bad type of Entrepreneur defaults with probability 1.*

In proposition 2, the lowest prior for which investment is supported in period 1 is solved from the incentive compatibility condition of the Investor given the proposed equilibrium strategy of the Entrepreneur. That is, investment in the first period is optimal if and only if

$$p(1 + \delta\sigma_{I2})r + (1 - p)[\sigma_{b1}(r - \delta\sigma_{I2}) - (1 - \sigma_{b1})] \geq 0.$$

Assume now that $M \leq \hat{M}$. The bad Entrepreneur is only concerned about maximizing his payoff in the first period, and the first-period equilibrium is separating, leading to $p_2(I, R) = 1$ or $p_2(I, D) = 0$.

Proposition 3. *Let $M \leq \hat{M}$. The equilibrium of the first stage game is separating; the bad type of Entrepreneur defaults with certainty if an investment is made. The Investor invests in the first period if and only if*

$$p \geq \frac{1}{1 + (1 + \delta)r}.$$

The proof is straightforward. Given that the bad Entrepreneur does not

compensate the Investor, the threshold outlined in proposition 3 is solved from the Investor's incentive compatibility constraint to invest in the first period. This is given by

$$p(1 + \delta)r - (1 - p) \geq 0.$$

Notice that if $p \geq \frac{1}{1+r}$, the Investor would invest in period 2 even after no investment in period 1. This, however is suboptimal and hence not an equilibrium. The Investor gets a higher expected payoff by investing already in period 1 and reinvesting in period 2 if the outcome of the first period is (I, R) . We now turn to analyze the game in which M is the Entrepreneur's private information.

4.4 Uncertainty about M

This section solves for the PBE of a trust game which is otherwise the same as the one in the previous section, but M is now private information to the Entrepreneur. As a result, private information of the Entrepreneur concerns the whole type vector (τ, M) . The Investor's belief about M is given by the commonly known prior distribution F . In what follows, this model is sometimes referred to as the 'Private regime'.

Since M does not affect the strategy of the bad type in the last period, the second period is played as in the Public regime. That is, the Investor invests if and only if $p_2(h_2) \geq 1/(1+r)$. Suppose that this holds and $\sigma_{I2}(I, R) = 1$. As under the Public regime, there exists a cutoff \hat{M} such that all bad types $M \leq \hat{M} = (1+r)/\delta$ do not repay and all higher types invest in reputation in the first period by repaying the first loan.

Let us first look for an equilibrium in which the Investor invests in period 1, and $\sigma_{b1}(M) = 1$ if $M > \hat{M}$. The strategy of the bad type is optimal only if $\sigma_{I2}(I, R) = 1$. Given the proposed strategy of the bad type, the Investor continues investing in period 2 if and only if her belief at the beginning of period 2 is high enough. That is, if

$$\frac{p}{p + (1-p)(1 - F(\hat{M}))} \geq \frac{1}{1+r}. \quad (4.1)$$

Denote by \tilde{p}_{SR} the lowest prior p for which the condition above still holds (as an equality). The subscript 'SR' stands for sequential rationality. Notice that \tilde{p}_{SR} is increasing in the probability $(1 - F(\hat{M}))$: the more likely

it is that the bad type is concerned for reputation and repays the first investment, the less the Investor's posterior beliefs react if an investment is repaid and hence the more limited the bad Entrepreneur's scope for mimicking the good type.

Suppose that $p \geq \tilde{p}_{SR}$. To check that the Investor wants to invest in period 1, the following incentive constraint must hold. The LHS gives her expected payoff if she invests, given the proposed strategy of the Entrepreneur, and the RHS is her payoff if she does not invest.¹³

$$p(1 + \delta)r + (1 - p) \left[(1 - F(\hat{M})) (r - \delta) - F(\hat{M}) \right] \geq 0. \quad (4.2)$$

The set of priors which satisfy the incentive-compatibility constraint is larger the more likely it is that the bad type is concerned for reputation and repays the first investment. In this case, if $r > \delta$ it is always optimal to invest in the first period, and if $\delta > r$ investing remains optimal for quite long. If the bad type is only concerned for short-term gain, investing is optimal only for priors higher than $1 / (1 + (1 + \delta)r)$. Since the value of M is unknown to the Investor, investing remains optimal also for priors less than this, provided that it is sufficiently likely that $M > \hat{M}$. Call the lowest prior that satisfies the incentive-compatibility constraint by \tilde{p}_{IC} . By the previous discussion, \tilde{p}_{IC} is decreasing in the probability $(1 - F(\hat{M}))$.

Since \tilde{p}_{IC} is decreasing and \tilde{p}_{SR} increasing in $(1 - F(\hat{M}))$, there exists a unique value $Z(r) \equiv \frac{1}{2+r}$ such that $\tilde{p}_{SR} = \tilde{p}_{IC}$ if $(1 - F(\hat{M})) = Z(r)$. If $(1 - F(\hat{M})) \geq Z(r)$, the reputation concern is so likely that $\tilde{p}_{IC} \leq \tilde{p}_{SR}$. If the reverse holds, the bad Entrepreneur is likely to have a concern for short-term profits, and $\tilde{p}_{SR} < \tilde{p}_{IC}$.

Proposition 4. *For all $p \geq \tilde{p} \equiv \max \{ \tilde{p}_{SR}, \tilde{p}_{IC} \}$, where*

$$\tilde{p}_{SR} = \frac{(1 - F(\hat{M}))}{(1 - F(\hat{M})) + r},$$

¹³In fact, if $p > \frac{1}{1+r}$, and the Investor does not invest, she would have the option to still invest in period 2. This would place a more demanding constraint for the equilibrium that we are looking at. However, it is easy to check that the Investor gets always a higher expected payoff by investing immediately and not postponing investments. This way, she may have two profitable trades instead of one.

and

$$\tilde{p}_{IC} = \begin{cases} \frac{F(\hat{M}) - (1 - F(\hat{M}))(r - \delta)}{(1 + \delta)r + F(\hat{M}) - (1 - F(\hat{M}))(r - \delta)} & \text{if } (1 - F(\hat{M})) \leq \frac{1}{1 + r - \delta} \\ 0 & \text{otherwise.} \end{cases}$$

the unique PBE of the game consists of the Investor investing as long as earlier investments have been repaid. The bad Entrepreneur's strategy in period 1 is characterized by

$$\sigma_{b1}(M) = \begin{cases} 1 & \text{if } M > \hat{M} \\ 0 & \text{otherwise} \end{cases},$$

where $\hat{M} = \frac{1+r}{\delta}$.

Proof: In the appendix. \square

For the equilibrium for priors lower than \tilde{p} , it matters what \tilde{p} is. Suppose that $(1 - F(\hat{M})) \geq Z(r)$, and consider priors $p < \tilde{p}_{SR}$. If the bad Entrepreneur is concerned for his reputation, it must be that $\sigma_{b1}(M) < 1$. On the other hand, for there to exist an equilibrium in which the bad type randomizes in period 1, it must be that $\sigma_{I2}(I, R) < 1$, which again requires that $p_2(I, R) = 1/(1+r)$. Thus, $\sigma_{b1}(M)$ must satisfy

$$\frac{p}{p + (1 - p)(1 - F(\hat{M}))\sigma_{b1}(M)} = \frac{1}{1 + r}. \quad (4.3)$$

But recall from the analysis of the Public regime that any equilibrium which consists of mixing by the bad type requires that the Investor's mixed strategy in period 2 is a function of M . When M is not observed by the Investor, the mixed strategy of the bad type is equivalent to there being a cutoff type \hat{M}_{adj} such that all types higher than this invest in reputation with probability 1. The adjusted cutoff is implicitly defined by

$$\begin{aligned} (1 - F(\hat{M}_{adj})) &= (1 - F(\hat{M}))\sigma_{b1}(M) \\ &= \frac{rp}{1 - p}, \end{aligned}$$

where the last equality follows from eq. (4.3). The above equality is valid as long as $r < \frac{1-p}{p}$.¹⁴ To ensure that the bad Entrepreneur of type \hat{M}_{adj}

¹⁴Since the term $\frac{1-p}{p}$ is decreasing in p , and the condition holds for all $r \geq 0$

is indifferent between repaying or not in period 1, the Investor needs to lower the benefit from an established reputation. Let $\sigma_{I2}^{Pr} \in (0, 1)$ give the probability that she invests in period 2 if the earlier investment has been repaid. In the equilibrium, this probability has to decrease in \hat{M}_{adj} so as to discourage all types lower than the cutoff to forgo investing in reputation.

Finally, the equilibrium requires that the Investor find it optimal to invest in period 1 given the strategy of the Entrepreneur. Denote the lowest prior belief for which the equilibrium is supported by \hat{p} . For priors below it, the expected gains to the Investor from the reputation building of the bad type are so low that they are outweighed by the expected costs that occur in case the bad type is concerned for short-term gains only. Thus, the only equilibrium of the game is a no-trade equilibrium.

Proposition 5. *If $(1 - F(\hat{M})) \geq Z(r)$, and if*

- $p \in [\hat{p}, \tilde{p}_{SR})$, where $\hat{p} = \frac{1}{r^2 + 2r + 1}$, the unique PBE consists of the Investor investing in period 1 and the bad Entrepreneur repaying it with certainty if $M > \hat{M}_{adj}$. In period 2, if the first investment was repaid, the Investor randomizes her action by investing with probability

$$\sigma_{I2}^{Pr}(p) = \frac{1 + r}{\delta \hat{M}_{adj}}.$$

- $p < \hat{p}$, the unique PBE of the game consists of no investments, and $\sigma_{b1}(M) = \sigma_{b2}(M, h_2) = 0$.

Proof: In the appendix. \square

Suppose now that $(1 - F(\hat{M})) < Z(r)$, and consider priors lower than \tilde{p}_{IC} . The mixed-strategy equilibrium of Proposition 5 cannot be sustained if $p \in [\tilde{p}_{SR}, \tilde{p}_{IC})$ because the posterior of the Investor is high enough to have $\sigma_{I2}(I, R) = 1$, and therefore $\sigma_{b1}(M) = 1$. However, since $p < \tilde{p}_{IC}$, the Investor does not have the incentive to make the first investment. On the other hand, if $p < \tilde{p}_{SR}$, the mixed-strategy equilibrium in Proposition 5 cannot be sustained since it can be checked that $\tilde{p}_{SR} < \hat{p}$. Finally, since $\tilde{p}_{IC} < \frac{1}{1 + (1 + \delta)r}$ for all $r \geq 0$ and for all $\delta \in [0, 1]$, a separating first-period equilibrium cannot be sustained either.

when $p = \tilde{p}_{SR}$, \hat{M}_{adj} is always defined.

Proposition 6. *If $(1 - F(\hat{M})) < Z(r)$, and if $p < \tilde{p}_{IC}$, the unique PBE consists of no investments being made and $\sigma_{b1}(M) = \sigma_{b2}(M, h_2) = 0$.*

Proposition 6 states that in case the project does not seem to be very valuable and the Investor's prior belief is rather pessimistic, the only equilibrium of the game is a no-trade equilibrium in which the Investor does not invest and therefore the bad Entrepreneur does not repay for any M . We now turn to comparing the equilibria under the Public and the Private regime in terms of their welfare.

4.5 Welfare

This section presents the main results of the paper which concern the ex-ante welfare of the players. Ex-ante welfare of any player i and for any prior p is given as the expected payoff conditional on any realized value of M . In what follows, the focus is on priors in the intermediate range of $[\hat{p}, \frac{1}{1+r})$. For all other prior beliefs, the equilibria as well as expected payoffs of the two regimes are identical. Recall that for priors below \hat{p} , both regimes feature no investments. On the other hand, for priors higher than $\frac{1}{1+r}$, the Investor is sufficiently confident in facing a good Entrepreneur to not pay attention to the actions of the bad type. Information about M is therefore redundant. Ex-ante expected payoffs of each player in both of the models are reported in the appendix.

In general, whether any given player prefers the Private or the Public regime depends on the form of F . It is a somewhat complicated exercise to state general conditions that hold for all possible distributions F , and for all r and δ . Therefore, the results of this section are not exhaustive; they do not cover all possible scenarios for all players¹⁵. Rather, the purpose of this section is to point out that there exist, under fairly general conditions, parameter specifications of the model under which all players obtain a strictly higher ex-ante expected payoff if information about M is privately held by the Entrepreneur. In fact, the result holds also to the other direction. In particular, this section shows that for any $r, \delta \in (0, 1]$ and for any distribution F , there always exists at least some priors such that the Private regime is optimal for all players. The results are expressed in terms of the probability $(1 - F(\hat{M}))$ instead of $F(\hat{M})$ because the former gives directly the probability of reputational concerns which is a central measure for the analysis.

¹⁵In particular, the welfare analysis of the Entrepreneur is not comprehensive for all possible priors.

4.5.1 General results

The main result concerning the ex-ante welfare of the Investor across regimes is summarized in Theorem 1. The welfare of the Entrepreneur is covered shortly at the end of this subsection. Theorem 1 holds for any $r, \delta \in (0, 1]$ and covers all possible distributions F . It establishes that the range of priors for which the Private regime dominates is increasing in the ex-ante probability that the bad Entrepreneur has reputational concerns. Moreover, as long as there is a strictly positive probability that the bad type has reputational concerns, there is shown to exist a set of priors for which the Private regime dominates. This result is restated in Corollary 1.

Theorem 1. *(Investor) For any $r, \delta \in (0, 1]$, the Private regime yields the Investor strictly higher ex-ante expected payoff if and only if*

- $Z(r) \leq (1 - F(\hat{M})) < 1$ and $p \in (\hat{p}, \frac{1}{1+r})$, or
- $X(p, r, \delta) \leq (1 - F(\hat{M})) < Z(r)$ and $p \in (\tilde{p}_{IC}, \frac{1}{1+r})$, or
- $0 < (1 - F(\hat{M})) < X(p, r, \delta)$ and $p \in (\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$, where

$$Z(r) \equiv \frac{1}{2+r}, \text{ and } X(r, \delta, p) \equiv \frac{\Delta_1}{\Delta_1 + \Delta_2}, \text{ with } \Delta_1 = (1 - p) - r(1 + \delta)p, \text{ and } \Delta_2 = 1 + r - \delta - (r^2 + (2 - \delta)r + (1 - \delta))p.$$

Proof: In the appendix. \square

The more probable it is that the bad Entrepreneur has reputational concerns the larger is the set of priors for which the Private regime dominates in terms of ex-ante expected payoffs. Notice that all of the intervals for the prior beliefs are nonempty as long as $r, \delta > 0$. Therefore, given that Theorem 1 spans all values of $F \in (0, 1)$, it directly implies Corollary 1 below.

Corollary 1. *For any non-degenerate distribution F and for any $r, \delta \in (0, 1]$, there exists a non-empty set of priors p such that the Investor is strictly better off under the Private regime.*

It is rather intuitive that the Entrepreneur of any type always prefers

the regime which maximizes the probability of investments, both in the first and in the second period.¹⁶ Why the the Investor holds a fairly general preference for the Private regime is less clear. The remainder of this subsection tries to shed light on what drives the welfare result of the Investor.

The general principle is that whenever $M \leq \hat{M}$, the Investor would prefer the Public regime so as to avoid expected losses due to the bad type defaulting on all loans. Under the Private regime, investments are made in equilibrium as long as the Investor believes that there is a high enough probability that the bad type repays the first loan. For some priors¹⁷, were she able to tell that this probability is zero, she would be better off by not investing.

On the other hand, whenever $M > \hat{M}$, the Investor is better off under the Private regime where, given that the bad type of Entrepreneur is concerned for his reputation and repays the first loan, the Investor obtains her maximum payoff in the first period as long as both types of the Entrepreneur repay the investment with certainty. Because the Investor always learns from the equilibrium action of the Entrepreneur, the bad type has to rely less on mixed strategies under the Private regime. Thus, the Private regime benefits from having more room for reputation building, and this benefit is increasing in the difference $(r - \delta)$. With this in mind, let us introduce the following notation.

Definition 1. *Reputation effect, R , and the screening effect, S are defined as*

$$\begin{aligned} R &\equiv \mathbb{E}_\tau \mathbb{E}_M \left[V^{Pr}(\sigma_I, \sigma_\tau) - V^{Pub}(\sigma_I, \sigma_\tau) \mid M > \hat{M} \right] \\ S &\equiv \mathbb{E}_\tau \mathbb{E}_M \left[V^{Pub}(\sigma_I, \sigma_\tau) - V^{Pr}(\sigma_I, \sigma_\tau) \mid M \leq \hat{M} \right], \end{aligned}$$

where $V^\kappa(\sigma)$ is the Investor's total payoff under the regime $\kappa \in \{Pub, Pr\}$.

Using the reputation and the screening effect, the Private regime ex-ante dominates the Public regime if and only if

$$(1 - F(\hat{M})) R - F(\hat{M}) S \geq 0. \quad (4.4)$$

¹⁶as long as $\mathbb{E} \left[M \mid M \leq \hat{M} \right] > 1 + r$. Ref. section 4.5.2. for more on this.

¹⁷ $p < \frac{1}{1+(1+\delta)r}$

Both the reputation and the screening effect are positive. The reputation effect occurs when $M > \hat{M}$ and, as discussed, it works for the benefit of the Private regime. The screening effect occurs if $M \leq \hat{M}$ and it works for the benefit of the Public regime. Notice that the inequality in (4.4) merely reformulates the condition that the ex-ante expected payoff is higher under the Private regime than under the Public regime. Theorem 2 below is therefore essentially another way to express Theorem 1 using the reputation and the screening effects.

Theorem 2. *(Investor) Take any $r, \delta \in (0, 1]$, and any non-degenerate distribution F such that $\left(1 - F(\hat{M})\right) \geq Z(r)$. The ex-ante reputation effect strictly dominates the ex-ante screening effect for all priors $p \in (\hat{p}, \frac{1}{1+r})$.*

Proof: In the appendix. \square

What Theorem 2 shows, which is not evident from the discussion under Theorem 1, is that the reason for the ex-ante dominance of the Private regime is the fact that the ex-ante value of the reputation effect outweighs the ex-ante value of the screening effect. A similar result can be constructed for other distributions as well. The less likely it is that the bad Entrepreneur has reputational concerns, however, the more the screening effect weighs for ex-ante welfare and therefore the higher the Investor's prior belief has to be for the Private regime to dominate. In particular, if $X(p, r, \delta) \leq \left(1 - F(\hat{M})\right) < Z(r)$, the ex-ante reputation effect strictly dominates the ex-ante screening effect for all priors $p \in (\tilde{p}_{IC}, \frac{1}{1+r})$. If $0 < \left(1 - F(\hat{M})\right) < X(p, r, \delta)$, the strict dominance of the Private regime can only be ensured for $p \in (\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$.

Figure 4.2 illustrates the reputation and the screening effects in one numerical example where $r = \delta = 1$, and M is drawn from a uniform distribution on $[1, 4]$. These imply that $\hat{M} = 2$ and $\left(1 - F(\hat{M})\right) = \frac{2}{3} > Z(r) = \frac{1}{3}$. Thus, reputation concerns are very likely. By Theorems 1 and 2, the Private regime is strictly optimal for all intermediate priors (now all priors between $\frac{1}{4}$ and $\frac{1}{2}$). To ease the interpretation of the figure, note in addition that $\tilde{p}_{SR} = \frac{2}{5}$, and $\frac{1}{1+(1+\delta)r} = \frac{1}{3}$. The values behind the figure are reported in the appendix. Notice that the figure compares the ex-ante effects, that is, R and S weighted by their respective probabilities.

The figure shows that the reputation effect remains relatively stable for priors below \tilde{p}_{SR} , and decrease linearly as p increases towards $\frac{1}{2}$. Consider

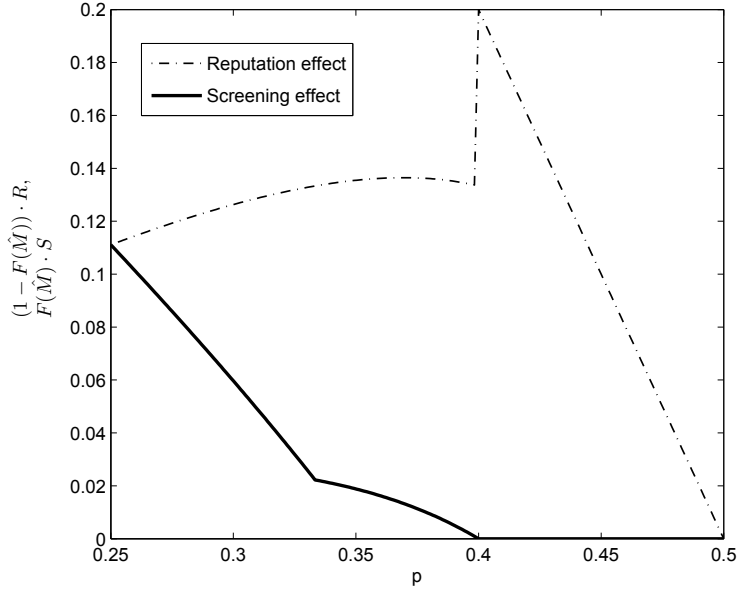


Figure 4.2. Reputation and screening effects, $r = \delta = 1$, $M \sim U[1, 4]$

first the interval $[\frac{2}{5}, \frac{1}{2}]$. The Private regime equilibrium consists of the bad Entrepreneur repaying the first loan with probability 1, whereas the Public regime equilibrium is in mixed strategies and the bad type repays the first loan with a probability which is increasing in the prior. This is why the reputation effect dies out gradually as p approaches $\frac{1}{2}$.

For priors lower than $\frac{2}{5}$, the Private regime features mixed strategies as well. This is why the reputation effect is more modest though still positive. This is because the bad type is still able to repay the first investment with a higher probability if M is private knowledge. The source of this difference in the repayment rates is explained in more detail under the proof of Theorem 2. Basically, when conditioning on the event that $M > \hat{M}$, the probability $\Pr[M > \hat{M}_{adj} \mid M > \hat{M}]$ is given by

$$\frac{(1 - F(\hat{M}_{adj}))}{(1 - F(\hat{M}))} = \frac{\frac{rp}{1-p}}{(1 - F(\hat{M}))} = \frac{\sigma_{b1}^{Pub}(M)}{(1 - F(\hat{M}))} > \sigma_{b1}^{Pub}(M),$$

where $\sigma_{b1}^{Pub}(M)$ is the probability that the bad type repays the investment in period 1 under the Public regime.

Turning to the screening effect, it is zero for $p \in [\frac{2}{5}, \frac{1}{2})$ because both regimes involve investments as long as they are repaid. Since the first investment is not repaid if the Entrepreneur is bad, the Investor learns

this under both regimes and is hence in the same position at the beginning of period 2 regardless of the regime. For $p \in [\frac{1}{3}, \frac{2}{5}]$, the screening effect is strictly positive and decreasing in the prior because the Public regime allows the Investor to learn if the Entrepreneur is good in which case she invests in period 2 with certainty. The Private regime, on the other hand, consists of mixed strategies by which the Investor invests in period 2 with a probability strictly less than 1 even though, would she know M , she would know that the Entrepreneur is good for sure if he has repaid.

For $p \in [\frac{1}{4}, \frac{1}{3}]$, there is no trust under the Public regime. This means, on one hand, that the good Entrepreneur gets no chance to reveal his type, but more importantly, that the bad type has no chance to deceive the Investor which for low priors weighs more than the forgone opportunity to invest to the good type's project. Under the Private regime, the equilibrium features mixed strategies which means that the good type gets financing with certainty in period 1 but with less than certainty in period 2. More importantly, though, the lower the prior is the higher is the Investor's expected loss from investing in period 1.

Let us concentrate briefly on the welfare of the Entrepreneur. It is sensitive to the shape of F , as the mixed strategy equilibria of both the Private and the Public regime involve randomization which depends either on $\hat{M}_{adj} = F^{-1}\left(\frac{1-p-rp}{1-p}\right)$ or on $\mathbb{E}(M \mid M > \hat{M})$. However, the following Lemmas 4 and 5 can be readily established. They are enough to provide the result of Theorem 3 that follows.

Lemma 4. (Entrepreneur)

Take any $r, \delta \in (0, 1]$ and any F such that $\mathbb{E}(M \mid M \leq \hat{M}) > 1 + r$. Then an Entrepreneur of any type is strictly better off under the Private regime for all $p \in (\tilde{p}, \frac{1}{1+r})$, where \tilde{p} is either \tilde{p}_{SR} or \tilde{p}_{IC} , depending on the value of $F(\hat{M})$.

Proof: In the appendix. \square

Lemma 5. (Entrepreneur)

Take any $r, \delta \in (0, 1]$ and any F such that $\mathbb{E}(M \mid M \leq \hat{M}) < 1 + r$. Then an Entrepreneur of any type is strictly better off under the Private regime for all $p \in (\max\left\{\tilde{p}_{SR}, \frac{1}{1+(1+\delta)r}\right\}, \frac{1}{1+r})$.

Proof: In the appendix. \square

Lemmas 4 and 5 only provide a partial truth in the sense that the results may hold for a larger set of priors. This remains future work for now. However, since the lemmas span all $F \in (0, 1)$ Theorem 3 can be established.

Theorem 3. *(Entrepreneur) For any non-degenerate distribution F and for any $r, \delta \in (0, 1]$, there exists a non-empty set of priors p such that the Entrepreneur of any type is strictly better off under the Private regime.*

The following subsection covers two simple numerical examples that illustrate the more general results in this subsection. The first example is to illustrate Lemma 5. Namely, it matters for the welfare of the good Entrepreneur whether the Entrepreneur's project may be unprofitable, that is, whether M may take on values less than $1 + r$. The second example assumes a discrete uniform distribution to show that the results of this section continue to hold. This, however, requires small changes in the equilibrium of the Private regime.

4.5.2 Two examples

Example 1. *Let $r = \delta = 1$, and let F be uniform on $[1, 4]$. For $p \in (\frac{1}{3}, \frac{1}{2})$, all players are strictly better off under the Private regime. For $p \in (\frac{1}{4}, \frac{1}{3})$, the Investor and the bad Entrepreneur are strictly better off under the Private regime, but the good Entrepreneur is strictly better off under the Public regime. For all other priors, all players are indifferent between the regimes.*

In this example¹⁸, because an Entrepreneur who always repays investments incurs ex-ante expected losses in case $\mathbb{E}[M \mid M \leq \hat{M}] < 1 + r$, he would prefer no investments to be made if $M < \hat{M}$. This is the case under the Public regime where, for $p < \frac{1}{1+(1+\delta)r}$, the Investor refuses to invest knowing that the bad type of the Entrepreneur would default with certainty. This non-investment equilibrium is a blessing for the good type who avoids the negative payoff associated with a low realization of M . The bad type obviously prefers that an investment be made because he would not repay it in any case. Also, given that M does not directly affect the Investor's payoff, as long as she expects investments to be repaid, she prefers the regime under which the investments are more likely to be

¹⁸The ex-ante expected payoffs are reported in the appendix.

made. The welfare in Example 1 is illustrated in Figure 4.3.

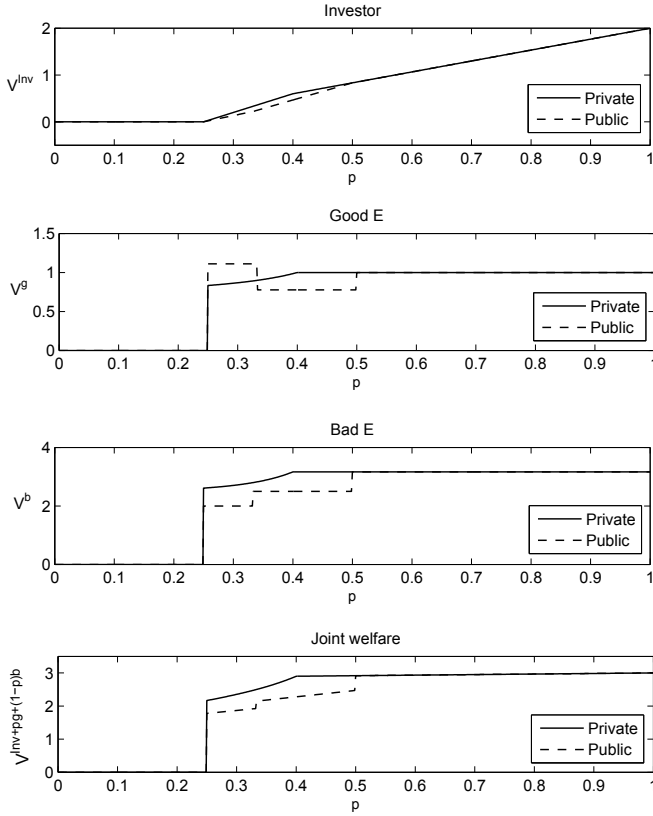


Figure 4.3. Ex-ante expected payoffs, $r = \delta = 1$, $M \sim U[1, 4]$

Discrete uniform distribution Let $r = \delta = 1$ such that $\hat{M} = 2$, and let $M \in \{2, 4\}$ with $F(2) = F(4) = \frac{1}{2}$. A discrete distribution implies small changes in the equilibrium under the Private regime for priors $p < \tilde{p}_{SR}$. Namely, since we are unable to define \hat{M}_{adj} , which would require M to be continuous, the only Perfect Bayesian Equilibrium consists of the bad type betraying trust with certainty in both periods. The equilibrium is supported for all priors $p_1 \geq \frac{1}{1+(1+\delta)r}$, as solved in section 4.3.2 for the Public regime.

Example 2.¹⁹ Let $r = \delta = 1$, and $M \in \{2, 4\}$ with equal probability. For all priors in the range $(\frac{1}{3}, \frac{1}{2})$, all players obtain strictly higher ex-ante payoffs under the Private regime. For all other priors, all players are indifferent between the regimes.

¹⁹The ex-ante expected payoffs are reported in the appendix.

4.6 Discussion and ideas for further research

The current model is binary by construction. It does not allow the Investor to adjust the level of her investment as the game proceeds. Since the Investor's stage-game payoff is increasing in x , focusing on an all or nothing set-up is without loss of generality in a one-shot game. In a repeated game, though, allowing x to adjust between periods would seem a more realistic assumption. However, if the Investor can choose the future value of the relationship by committing to an increasing sequence of investments as trust evolves, she could incentivize a bad Entrepreneur to repay investments. By handing this power to the Investor we however lose the distinction between the Private and the Public regime, as the Investor would always be informed about the bad Entrepreneur's equilibrium action in the first period. The logic of "starting small" has been studied by many authors (see e.g. Watson, 1999 and 2002; Watson and Rauch, 2003; and Halac, 2014), and it has been shown to resolve challenges in cooperation games in which players can choose to change partners at any period (Datta, 1996; Ghosh & Ray, 1996; Kranton, 1996). When relationships start small, the temptation to switch partners at every period and betray consecutively is removed because a change of partner will lead to a period of low payoffs during new phase of trust building.

Suppose that the Investor is able to design a contract which specifies the interest rate that the Entrepreneur must use for repayments and an investment x as a function of the interest rate r . The Investor must design the contract in such a way that the Entrepreneur accepts it. Certain incentive compatibility conditions must hence be satisfied. The good type of Entrepreneur would never breach the contract but the bad type has an incentive to set $r = 0$ to maximize his immediate gain although at the cost of not obtaining financing in period 2. By designing the contract optimally, the Investor is able to dictate the type of equilibrium to be implemented. However, since the optimal level of r likely depends on M , if the Investor is uncertain about M , she may end up implementing a suboptimal equilibrium. This model, which is reminiscent of the relational contracting models discussed in section 2, may serve as an example where at least the Investor is hurt by private information about M .

The current set-up features one-sided imperfect information in that the Entrepreneur has all the available information in both regimes. Two-sided imperfect information could be introduced for instance by assum-

ing that the Investor is privately informed of her investment horizon. She may know with certainty that she is only going to invest once in the project, and look for other investment opportunities in the second period. Alternatively, since the availability of substitutes may not be clear at the beginning of the game, or finding them may turn out to be too costly, she may only know the probability δ that the game continues after the first period. The Entrepreneur only knows the distribution from which δ is drawn. Tentative analysis suggests that introducing this type of additional information asymmetry would still allow for the existence of parameter environments under which all players strictly prefer the Private regime. However, the existence of those environments seems to be more limited than in the present model. Moreover, interim welfare would suggest that after learning that her investment horizon is short, the Investor prefers to not reveal it. Conversely, after learning that the investment horizon is long (two periods), the Investor prefers to have it publicly known. Since the short-horizon type would always mimic the reports of the long-horizon type, cheap talk regarding the investment horizon does not seem credible.

Just as the Investor may be privately informed about her investment horizon, the Entrepreneur may be privately informed about his future plans. The analysis in this paper would go through without significant changes²⁰ if, instead of M , the Entrepreneur may be privately informed about the discount factor δ which could be interpreted as the probability that he continues to do business in the second period. It could happen, for example, that due to sudden changes in his health, he will retire after the first period. Since he is better informed of his health, he knows the probability that a sudden retirement occurs, but the Investor holds only a belief about the true value of δ .

Currently, there is only one Investor offering financing for the Entrepreneur. In the presence of a pool of Investors, each uninformed about the history of play in other games but the ones they are involved in, the Entrepreneur's incentives to repay loans would be weakened. If the cost of finding a new investor is low, the bad Entrepreneur would never repay his loans. The harder it is to replace the Investor, for instance in economic downturns,

²⁰The only change in the analysis would arise from the fact that δ , being the common discount factor, enters the incentive compatibility constraint of the Investor. Hence, under the Private regime, the equilibrium of the game depends on how the Investor's prior falls between thresholds which are functions of the conditional belief about δ . Uncertainty about δ will shift the lower bound \hat{p} up such that there will be priors for which all players strictly prefer the Public regime because the Private regime no longer supports trust.

the more likely it is that the bad Entrepreneur has concerns for reputation. Instead, publicly observed history of play between any Investor and Entrepreneur would bring us back to the model of this paper.

Another simplifying assumption relates to the interest rate which is fixed. It therefore excludes the possibility for the good Entrepreneur to attract the Investor by optimizing the interest rate it offers to pay in return for financing. As long as the cooperative outcome Pareto dominates the no-trust outcome, the good Entrepreneur could, by adjusting r , try to induce an equilibrium in which the bad type would default on the loan immediately if it was preferred to an equilibrium in which the bad type's action remains uncertain. By merely setting the interest rate high, the good type is not able to separate from the bad type because the latter could promise also a high rate. However, the Investor would know after seeing a high r that it is at least very unlikely that the bad type would actually repay in the first period.

Finally, the results of the theoretical model could be tested empirically in the lab by comparing the frequency of trust across two versions of the standard trust game used in experiments: one in which all payoffs are common knowledge and another in which the Trustee is privately informed of his payoffs.

The model features both adverse selection (hidden information) (which the Principal tries to mitigate by learning from the agent's observed actions), and moral hazard (hidden action) in the sense that once an investment is made, the agent's action is still unknown.

4.7 Conclusion

The paper analyzes a twice-repeated trust game with incomplete information about the reliability of an Entrepreneur under two model variants and compares them in terms of ex-ante expected payoffs. Based on the results, the paper proposes a rationale for concealing information about payoffs or about the length of the horizon in principal agent interactions. Namely, there is shown to always exist parameter specifications under which all players are strictly better off if and only if the Investor remains ignorant about the payoffs of the Entrepreneur. The intuition for the result is that private information about the Entrepreneur's payoffs allows for more room for reputation building by the bad type which again results in more trust on average. From the Investor's point of view, the benefits

from increased trust due to reputation building by the bad type exceed the expected losses that the Investor incurs in case the bad type is only concerned about immediate gains. The model analyzed in the paper, although highly stylized, is reminiscent of models analyzed in the literature on delegation and on relational contracts, and the analysis of more nuanced settings could bring new results on transparency in organizations.

4.8 Appendix: Proofs

Proof of Proposition 4. First, \tilde{p}_{SR} is solved from condition (4.1) in the text. To solve for the \tilde{p}_{IC} , notice that the incentive-compatibility condition of the Investor in (4.2) can be rewritten as

$$p \left[(1 + \delta) r - (1 - F(\hat{M})) (r - \delta) + F(\hat{M}) \right] \geq F(\hat{M}) - (1 - F(\hat{M})) (r - \delta). \quad (4.5)$$

There are two cases to consider. First, if $(1 - F(\hat{M})) \leq \frac{1}{1+r-\delta} \equiv Z(r, \delta)$, both sides of inequality (4.5) are positive, and the incentive condition holds for all

$$p \geq \frac{F(\hat{M}) - (1 - F(\hat{M})) (r - \delta)}{(1 + \delta) r + F(\hat{M}) - (1 - F(\hat{M})) (r - \delta)}. \quad (4.6)$$

Second, if $(1 - F(\hat{M})) > Z(r, \delta)$, then either the RHS of (4.6) is negative and the LHS is positive, or both sides are negative. In either case, the incentive condition would hold for all positive priors so that $\tilde{p}_{IC} = 0$. Combining condition (4.6) with the condition that $p \geq \tilde{p}_{SR}$ implies that there is a pooling first-period equilibrium for all priors $p \geq \max \{\tilde{p}_{SR}, \tilde{p}_{IC}\}$.

To check for $Z(r)$, notice that if $(1 - F(\hat{M})) \leq Z(r, \delta)$, then $\tilde{p}_{SR} \geq \tilde{p}_{IC}$ if

$$(1 - F(\hat{M})) \geq \frac{1}{2+r} \equiv Z(r). \quad (4.7)$$

Furthermore, $Z(r) < Z(r, \delta)$ for all r and δ . Thus, $\tilde{p}_{SR} \geq \tilde{p}_{IC}$ if $(1 - F(\hat{M})) \geq Z(r)$, and otherwise the reverse holds. Following the discussion in the text, it is clear that none of the players want to deviate from the proposed equilibrium as long as others follow their equilibrium strategies. \square

Proof of Proposition 5. To ensure that the bad Entrepreneur of type \hat{M}_{adj} is indifferent between repaying investments and not, the Investor needs to lower the benefit from an established reputation. Then σ_{I2}^P is

solved from

$$M - (1 + r) + \hat{M}_{adj} \sigma_{I2}^{Pr} M = M.$$

The incentive-compatibility condition for the Investor to invest in period 1 given the strategy of the bad type is given by

$$F(\hat{M}_{adj}) [p(1 + \sigma_{I2}^{Pr} \delta)r - (1 - p)] + \\ (1 - F(\hat{M}_{adj})) [p(1 + \sigma_{I2}^{Pr} \delta)r + (1 - p)(r - \sigma_{I2}^{Pr} \delta)] \geq 0,$$

which holds whenever

$$p \geq \frac{1}{r^2 + 2r + 1} \equiv \hat{p}.$$

Finally, it can be checked that if $(1 - F(\hat{M})) \geq Z(r)$, $\tilde{p}_{SR} \geq \tilde{p}_{IC} \geq \hat{p}$. When the conditions stated in the Proposition hold, there are no profitable unilateral deviations for any player. \square

Ex-ante welfare

Public regime

$$\text{If } p \geq \frac{1}{1+r},$$

Investor:

$$F(\hat{M}) [p(1 + \delta)r - (1 - p)] + (1 - F(\hat{M})) [p(1 + \delta)r + (1 - p)(r - \delta)]$$

$$E_g: \quad (\mathbb{E}[M] - (1 + r))(1 + \delta)$$

$$E_b: \quad F(\hat{M})\mathbb{E}[M \mid M \leq \hat{M}] + (1 - F(\hat{M})) [(1 + \delta)\mathbb{E}[M \mid M > \hat{M}] - (1 + r)]$$

$$\text{If } p \in \left[\frac{1}{1+(1+\delta)r}, \frac{1}{1+r} \right),$$

$$\text{Investor:} \quad F(\hat{M}) [p(1 + \delta)r - (1 - p)] + (1 - F(\hat{M})) [(r^2 + 2r + 1)p - 1]$$

$$E_g:$$

$$F(\hat{M})(1 + \delta) [\mathbb{E}[M \mid M \leq \hat{M}] - (1 + r)] + \\ (1 - F(\hat{M})) (1 + \sigma_{I2}^{Pub} \delta) [\mathbb{E}[M \mid M > \hat{M}] - (1 + r)]$$

E_b :

$$\begin{aligned}
 & F(\hat{M})\mathbb{E}[M \mid M \leq \hat{M}] + \\
 & \left(1 - F(\hat{M})\right) \left[\sigma_{1b} \left(\mathbb{E}[M \mid M > \hat{M}] (1 + \delta \sigma_{I2}^{P_{ub}}) - (1 + r) \right) + \right. \\
 & \left. (1 - \sigma_{1b})\mathbb{E}[M \mid M > \hat{M}] \right] \\
 & = F(\hat{M})\mathbb{E}[M \mid M \leq \hat{M}] + \left(1 - F(\hat{M})\right) \mathbb{E}[M \mid M > \hat{M}] \\
 & = \mathbb{E}[M]
 \end{aligned}$$

If $p \left[\frac{1}{r^2+2r+1}, \frac{1}{1+(1+\delta)r} \right)$,

Investor: $\left(1 - F(\hat{M})\right) \left[(r^2 + 2r + 1) p - 1 \right]$

E_g : $\left(1 - F(\hat{M})\right) (1 + \sigma_{I2}^{P_{ub}} \delta) \left[\mathbb{E}[M \mid M > \hat{M}] - (1 + r) \right]$

E_b : $\left(1 - F(\hat{M})\right) \mathbb{E}[M \mid M > \hat{M}]$

If $p < \frac{1}{r^2+2r+1}$: each player gets a payoff of 0.

Private regime

Case I: If $\left(1 - F(\hat{M})\right) \geq Z(r)$

If $p \geq \tilde{p}_{SR} = \frac{(1-F(\hat{M}))}{(1-F(\hat{M}))+r}$,

Payoffs for each player identical to those under the Public regime for

$p \geq \frac{1}{1+r}$.

If $p \in \left[\frac{1}{r^2+2r+1}, \tilde{p}_{SR} \right)$,

Investor:

$$\begin{aligned}
 & F(\hat{M}_{adj}) \left[p (1 + \sigma_{I2}^{Pr} \delta) r - (1 - p) \right] + \\
 & \left(1 - F(\hat{M}_{adj})\right) \left[p (1 + \sigma_{I2}^{Pr} \delta) r + (1 - p) (r - \sigma_{I2}^{Pr} \delta) \right] \\
 & = p (1 + \sigma_{I2}^{Pr} \delta) r - (1 - p) + rp + rp (r - \sigma_{I2}^{Pr} \delta) \\
 & = (r^2 + 2r + 1) p - 1
 \end{aligned}$$

E_g : $\left[\mathbb{E}[M] - (1 + r) \right] (1 + \sigma_{I2}^{Pr} \delta)$

E_b : $F(\hat{M}_{adj})\mathbb{E}[M \mid M \leq \hat{M}_{adj}] + \left(1 - F(\hat{M}_{adj})\right) \left[(1 + \sigma_{I2}^{Pr} \delta) \mathbb{E}[M \mid M > \hat{M}_{adj}] - (1 + r) \right]$

If $p_1 < \frac{1}{r^2+2r+1}$: each player gets a payoff of 0.

Case II: If $\left(1 - F(\hat{M})\right) < Z(r)$

If $p \geq \tilde{p}_{IC}$,

Payoffs as in Case I for $p \geq \tilde{p}_{SR}$.

If $p < \tilde{p}_{IC}$,

Investor, E_g , E_b : each player gets a payoff of 0.

Proof of Theorem 1. Let us go through the proof in three parts, depending on the value of $(1 - F(\hat{M}))$. Lemmas 1-3 below and their proofs establish necessary conditions for the dominance of the Private regime. Since the lemmas span all possible distributions F and hold for all $r, \delta \in (0, 1]$, they also establish the sufficient conditions for the dominance of the Private regime.

Lemma 1. *Take any $r, \delta \in (0, 1]$ and any F such that $(1 - F(\hat{M})) \geq Z(r)$. Then the Investor is strictly better off under the Private regime for all $p \in (\hat{p}, \frac{1}{1+r})$.*

Proof of Lemma 1.

Depending on the value of $F(\hat{M})$, either $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$ or the reverse holds. We shall see that the result of the Lemma holds for both cases. Suppose first that $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$. We must check separately cases (i) $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$, (ii) $p \in [\tilde{p}_{SR}, \frac{1}{1+(1+\delta)r})$, and (iii) $p \in [\hat{p}, \tilde{p}_{SR})$. In case (i), Private regime dominates iff

$$p(1 + \delta)r + (1 - p)(r - \delta) \geq (r^2 + 2r + 1)p - 1,$$

which holds if and only if

$$p \leq \frac{1 + (r - \delta)}{r^2 + (2 - \delta)r + (1 - \delta)}. \quad (4.8)$$

Since the expression in the RHS is always weakly higher than $\frac{1}{1+r}$, the condition (4.8) holds for all priors $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$. In case (ii), the Private regime yields a negative expected payoff if $M \leq \hat{M}$ since for priors lower than $\frac{1}{1+(1+\delta)r}$, $p(1 + \delta)r - (1 - p) < 0$. In the Public regime, this expected loss is avoided because the equilibrium consists of no investments in case $M \leq \hat{M}$. Hence, the dominance of the Private regime requires that the expected payoff in case $M > \hat{M}$ is sufficiently higher than under the

Public regime. Thus, what is required is that the ex-ante superiority of the Private regime in case $M > \hat{M}$ is larger than the ex-ante inferiority of the Private regime in case $M \leq \hat{M}$. In case $M \leq \hat{M}$, the expected payoff under the Public regime is higher than under the Private regime by an amount

$$(1 - p) - r(1 + \delta)p \equiv \Delta_1 > 0. \quad (4.9)$$

In case $M > \hat{M}$, the expected payoff under the Private regime is higher than under the Public regime by an amount

$$1 + r - \delta - (r^2 + (2 - \delta)r + (1 - \delta))p \equiv \Delta_2 > 0. \quad (4.10)$$

Equation (4.10) is positive by case (i) earlier. The difference of eq. (4.10), multiplied by the probability $(1 - F(\hat{M}))$, and eq. (4.9), multiplied by the probability $F(\hat{M})$, is positive, and the Private regime dominates if and only if

$$(1 - F(\hat{M})) \geq \frac{\Delta_1}{\Delta_1 + \Delta_2} \equiv X(p, r, \delta). \quad (4.11)$$

It can further be shown, using some algebra, that $X(p, r, \delta) < Z(r)$ for all $p \geq \hat{p}$, which means that condition (4.11) is satisfied by the assumption that $(1 - F(\hat{M})) \geq Z(r)$. In case (iii), Private regime dominates iff

$$(r^2 + 2r + 1)p - 1 \geq (1 - F(\hat{M}))[(r^2 + 2r + 1)p - 1],$$

which clearly holds for all F , r , and δ .

Suppose then that $\tilde{p}_{SR} > \frac{1}{1+(1+\delta)r}$. We must check separately cases (i) $p \in [\tilde{p}_{SR} \frac{1}{1+r})$, (ii) $p \in [\frac{1}{1+(1+\delta)r}, \tilde{p}_{SR})$, and (iii) $p \in [\hat{p}, \frac{1}{1+(1+\delta)r})$. Cases (i) and (iii) are identical to cases (i) and (iii) earlier in the proof. In case (ii), Private regime dominates iff

$$(r^2 + 2r + 1)p - 1 \geq p(1 + \delta)r - (1 - p),$$

which can easily be shown to hold for all $r, \delta \in [0, 1]$. The result of the lemma follows directly after combining the results from all cases studied.

□

Lemma 2. Take any $r, \delta \in (0, 1]$ and any F such that

$$X(r, \delta, p) \leq (1 - F(\hat{M})) < Z(r).$$

Then the Investor is strictly better off under the Private regime for all $p \in (\tilde{p}_{IC}, \frac{1}{1+r})$.

Proof of Lemma 2.

If $(1 - F(\hat{M})) < Z(r)$, it can be shown that $\tilde{p}_{IC} < \frac{1}{1+(1+\delta)r}$. Hence, to show that the Private regime dominates for all $p \geq \tilde{p}_{IC}$, we must consider separately two cases, (i) $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$, and (ii) $p \in [\tilde{p}_{IC}, \frac{1}{1+(1+\delta)r})$. Since for all $p \geq \tilde{p}_{IC}$, the payoff of the Investor is identical to what she obtains for $p \geq \tilde{p}_{SR}$ which was covered in the proof of Lemma 1, the analysis in case (i) is identical to that in case (i) in the proof of Lemma 1. The analysis in case (ii) is also identical to that in case (ii) in the earlier proof, but now that $(1 - F(\hat{M})) < Z(r)$, referring to the proof of Lemma 1, the Private regime dominates only if $X(p, r, \delta) \leq (1 - F(\hat{M})) < Z(r)$. \square

Lemma 3. Take any $r, \delta \in (0, 1]$ and any F such that $0 < (1 - F(\hat{M})) < X(r, \delta, p)$. Then the Investor is strictly better off under the Private regime for all $p \in (\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$.

Proof of Lemma 3.

The proof is analogous to that of Lemma 2, with the difference that, by the assumption that $(1 - F(\hat{M})) < X(r, \delta, p)$, the Private regime does not dominate the Public regime in case (ii) of the proof of Lemma 2. \square

Proof of Theorem 2. This proof refers to some extent to the proof of Theorem 1, and some of the overlapping notation is omitted here. Depending on the exact value of $F(\hat{M})$, either either $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$ or the reverse holds. Suppose first that $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$. We must check separately cases (i) $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$, and (ii) $p \in [\tilde{p}_{SR}, \frac{1}{1+(1+\delta)r})$, and (iii) $p \in [\hat{p}, \tilde{p}_{SR})$. In case (i), the reputation effect, given as the difference in the ex-ante expected payoffs under the Private and the Public regime in case $M > \hat{M}$, is given as

$$R_1 \equiv \underbrace{p(1+\delta)r + (1-p)(r-\delta)}_{\text{Private reg.}} - \underbrace{(r^2 + 2r + 1)p - 1}_{\text{Public reg.}}$$

Referring to the proof of Theorem 1, the reputation effect is positive in the range of priors that we are interested in. The screening effect, given as the difference in the ex-ante expected payoffs under the Public and the Private regime in case $M \leq \hat{M}$, is equal to zero. Hence, the ex-ante reputation effect strictly dominates the ex-ante screening effect for all $r, \delta \in (0, 1]$ and for all relevant F .

In case (ii), the reputation effect is the same as in case (i). For later use, rewrite it as

$$R_2 \equiv 1 + r - \delta - (r^2 + (2 - \delta)r + (1 - \delta))p \equiv \Delta_2$$

The screening effect is given by

$$S_2 \equiv 0 - [p(1 + \delta)r - (1 - p)],$$

where 0 is the Investor's payoff under the Public regime, and the latter term her expected payoff under the Private regime which is negative for priors less than $\frac{1}{1+(1+\delta)r}$. Hence, the screening effect is positive and rewritten as $S_2 = (1 - p) - p(1 + \delta)r \equiv \Delta_1$. Since the ex-ante dominance of the Private regime is determined by whether, ex-ante, the value of the reputation effect outweighs the value of the screening effect, the condition for the Private regime to dominate is given by

$$(1 - F(\hat{M}))R_2 - F(\hat{M})S_2 \geq 0.$$

This condition was shown under the proof of Lemma 2 to hold whenever $(1 - F(\hat{M})) \geq X(p, r, \delta)$. Since for the priors that we are considering, $Z(r) > X(p, r, \delta)$, the result holds.

In case (iii), the expected payoff of the Investor conditional on $M > \hat{M}$ can be expressed as a weighted average of the expected payoffs in cases $\hat{M} < M \leq \hat{M}_{adj}$ and $M > \hat{M}_{adj}$:

$$(1 - F(\hat{M})) \left\{ \frac{F(\hat{M}_{adj}) - F(\hat{M})}{(1 - F(\hat{M}))} [p(1 + \sigma_{I2}^{Pr} \delta)r - (1 - p)] + \frac{(1 - F(\hat{M}_{adj}))}{(1 - F(\hat{M}))} [p(1 + \sigma_{I2}^{Pr} \delta)r + (1 - p)(r - \sigma_{I2}^{Pr} \delta)] \right\}$$

where, notice that

$$\frac{F(\hat{M}_{adj}) - F(\hat{M})}{(1 - F(\hat{M}))} = 1 - \frac{(1 - F(\hat{M}_{adj}))}{(1 - F(\hat{M}))}.$$

To save on notation, define

$$\sigma_{b1}^{Pr} \equiv \frac{(1 - F(\hat{M}_{adj}))}{(1 - F(\hat{M}))} = \frac{rp}{(1 - p)(1 - F(\hat{M}))}.$$

The ex-ante expected payoff can be rewritten as

$$(1 - F(\hat{M})) \{p(1 + \delta\sigma_{I2}^{Pr})r + (1 - p)[\sigma_{b1}^{Pr}(r - \delta\sigma_{I2}^{Pr}) - (1 - \sigma_{b1}^{Pr})]\}, \quad (4.12)$$

where $\sigma_{I2}^{Pr} = \frac{1+r}{\delta\hat{M}_{adj}}$.

Before proceeding with reformulating expression (4.12), it is instructive to point out the relation of it to the ex-ante expected payoff under the Public regime, given as

$$(1 - F(\hat{M})) \{p(1 + \delta\sigma_{I2})r + (1 - p)[\sigma_{b1}(r - \delta\sigma_{I2}) - (1 - \sigma_{b1})]\}, \quad (4.13)$$

where $\sigma_{b1} = \frac{rp}{1-p}$, and $\sigma_{I2} = \frac{1+r}{\delta E(M|M>\hat{M})}$. Since, by definition, $\sigma_{b1}^{Pr} > \sigma_{b1}$, the Private regime yields a higher expected payoff in case the Entrepreneur is bad. If in addition $\hat{M}_{adj} < E(M|M>\hat{M})$, $\sigma_{I2}^{Pr} > \sigma_{I2}$, and the reputation effect is certainly positive. If the last inequality does not hold, then the Public regime yields a higher expected payoff in case the Entrepreneur is good. For the Private regime to dominate, it must be that the payoff against a bad type is sufficiently high to compensate for the lower payoff against a good type.

By substituting in both (4.12) and (4.13) for σ_{b1} , σ_{b1}^{Pr} , σ_{I2} , and σ_{I2}^{Pr} , and rearranging, one obtains that the ex-ante reputation effect is positive iff

$$pr(1 + r)F(\hat{M}) \left(1 - \frac{1}{\hat{M}_{adj}}\right) \geq 0, \quad (4.14)$$

which holds for all $r \geq 0$ since $\hat{M}_{adj} > \hat{M} = \frac{1+r}{\delta} > 1$.

The ex-ante screening effect is given as

$$F(\hat{M}) \{0 - [p(1 + \delta\sigma_{I2}^{Pr})r - (1 - p)]\},$$

where 0 is the payoff under the Public regime, and the second term the expected payoff in the Private regime. Notice that it is negative. After substituting for σ_{I2}^{Pr} , one can check that the reputation effect outweighs

the screening effect iff

$$F(\hat{M}) [(r^2 + 2r + 1)p - 1] \geq 0,$$

which holds for all $p \geq \hat{p}$, and is exactly the condition for the Private regime to dominate ex-ante, as seen in the proof of Theorem 1.

Suppose then that $\tilde{p}_{SR} > \frac{1}{1+(1+\delta)r}$. We need to check cases (i) $p \in [\tilde{p}_{SR}, \frac{1}{1+r})$, (ii) $p \in [\frac{1}{1+(1+\delta)r}, \tilde{p}_{SR})$, and (iii) $p \in [\hat{p}, \frac{1}{1+(1+\delta)r})$. Cases (i) and (iii) are identical in analysis to their counterparts earlier in the proof. In case (ii), the reputation effect is identical to case (iii) earlier in the proof, given in expression (4.14). The screening effect is given by

$$\begin{aligned} F(\hat{M}) \{p(1+\delta)r - (1-p) - [p(1+\delta\sigma_{I2}^{Pr})r - (1-p)]\} \\ = pr\delta F(\hat{M}) (1 - \sigma_{I2}^{Pr}) \cdot \parallel \sigma_{I2}^{Pr} = \frac{1+r}{\delta\hat{M}_{adj}} \end{aligned}$$

The ex-ante reputation effect outweighs the ex-ante screening effect iff

$$prF(\hat{M}) (1+r-\delta) \geq 0,$$

which holds for all $r, \delta \in (0, 1]$. The result of the theorem follows directly after combining all the cases covered. \square

Proof of Lemma 4. Consider first the case that $(1 - F(\hat{M})) < Z(r)$, and $\tilde{p} = \tilde{p}_{IC}$. Then, one can compare the Entrepreneur's expected payoffs separately for cases (i) $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$, and (ii) $p \in [\tilde{p}_{IC}, \frac{1}{1+(1+\delta)r})$ to see that the payoff under the Private regime is strictly higher in both cases for all relevant values of r and δ . In particular, since the payoff under the Private regime is the same in both cases, but that of the Public regime is lower in case (ii)²¹ where the unique equilibrium in case $M \leq \hat{M}$ consists of no investments, it is enough to show that the payoff under the Private regime is higher in case (i). For the good type, this requires that

$$\begin{aligned} (1+\delta) (\mathbb{E}[M] - (1+r)) &\geq \\ F(\hat{M}) (1+\delta) [\mathbb{E}[M \mid M \leq \hat{M}] - (1+r)] &+ \\ (1 - F(\hat{M})) (1 + \sigma_{I2}\delta) [\mathbb{E}[M \mid M > \hat{M}] - (1+r)] & \end{aligned}$$

which is easily checked to hold, since the Public regime equilibrium involves mixing by the Investor in period 2, and $\sigma_{I2}\delta < \delta$. For the bad type,

²¹Note, this requires that $\mathbb{E}[M \mid M \leq \hat{M}] \geq (1+r)$

Private regime is optimal if

$$F(\hat{M})\mathbb{E}[M \mid M \leq \hat{M}] + (1 - F(\hat{M})) \left[(1 + \delta)\mathbb{E}[M \mid M > \hat{M}] - (1 + r) \right] \geq \mathbb{E}(M),$$

where the LHS can be rewritten as $\mathbb{E}(M) + (1 - F(\hat{M})) \left[\delta\mathbb{E}[M \mid M > \hat{M}] - (1 + r) \right]$, where the last term is positive since $\mathbb{E}[M \mid M > \hat{M}] \geq \hat{M} \equiv \frac{1+r}{\delta}$. Thus, the condition holds for all $F(\hat{M}) \in (0, 1)$ and $r, \delta \in (0, 1]$.

Consider then the case that $(1 - F(\hat{M})) \geq Z(r)$, and $\tilde{p} = \tilde{p}_{SR}$. Depending on the value of $F(\hat{M})$, either $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$ or the reverse holds. Suppose first that $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$. We must check separately cases (i) $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$, and (ii) $p \in [\tilde{p}_{SR}, \frac{1}{1+(1+\delta)r})$. Since the Entrepreneur's expected payoff is identical for all $p \geq \tilde{p}$, whether \tilde{p} is \tilde{p}_{SR} or \tilde{p}_{IC} , the analysis of these two cases is identical to that in cases (i) and (ii) earlier in the proof. Suppose then that $\tilde{p}_{SR} > \frac{1}{1+(1+\delta)r}$. The only case to check is that of $p \in [\tilde{p}_{SR}, \frac{1}{1+r})$, which is identical in analysis to that in case (i) earlier in the proof. The result of the theorem follows directly after combining all the cases covered. \square

Proof of Lemma 5. Consider first the case that $(1 - F(\hat{M})) < Z(r)$, and $\tilde{p} = \tilde{p}_{IC}$. Then, the only case to consider is that of (i) $p \in [\frac{1}{1+(1+\delta)r}, \frac{1}{1+r})$. This is identical to the analysis under the proof of Lemma 4. The fact that $\mathbb{E}[M \mid M \leq \hat{M}] < 1 + r$ does not change the results there since the expressions including the term $\mathbb{E}[M \mid M \leq \hat{M}]$ cancel out from the payoff comparison of the good type, and the relationship of $\mathbb{E}[M \mid M \leq \hat{M}]$ and $1 + r$ does not concern the bad type who never repays the investment in case $M \leq \hat{M}$. Consider then that $F(\hat{M}) \leq Z(r)$, and $\tilde{p} = \tilde{p}_{SR}$. Depending on the value of $F(\hat{M})$, either $\tilde{p}_{SR} < \frac{1}{1+(1+\delta)r}$ or the reverse holds. In the former case, the analysis is identical to the discussion earlier in the proof. In case $\tilde{p}_{SR} > \frac{1}{1+(1+\delta)r}$, we need only to check the case (i) $p \in [\tilde{p}_{SR}, \frac{1}{1+r})$, which was shown to hold under the proof of Lemma 4. Finally, to check that the function $\max \left\{ \tilde{p}_{SR}, \frac{1}{1+(1+\delta)r} \right\}$ selects the correct input, notice that if $(1 - F(\hat{M})) < Z(r)$, $\tilde{p}_{SR} < \tilde{p}_{IC} < \frac{1}{1+(1+\delta)r}$, and in case $(1 - F(\hat{M})) \geq Z(r)$, the maximum of the two thresholds is determined by the exact value of $F(\hat{M})$. \square

Reputation and screening effects for Figure 4.2.

Reputation effect: For $p \in [\frac{2}{5}, \frac{1}{2}) : 1 - 2p$. For $p \in [\frac{1}{4}, \frac{2}{5}) : 2pF(\hat{M}) \left(1 - \frac{1}{\hat{M}_{adj}}\right)$.

Screening effect: For $p \in [\frac{2}{5}, \frac{1}{2}) : 0$. For $p \in [\frac{1}{3}, \frac{2}{5}) : pF(\hat{M}) \left(1 - \frac{2}{\hat{M}_{adj}}\right)$. For $p \in [\frac{1}{4}, \frac{1}{3}) : F(\hat{M}) [(1 - p) - p(1 + \sigma_{I2}^{Pr}(M))]$.

Expected payoffs in Example 1.

By applying the numerical values to the ex-ante expected payoffs detailed earlier in the appendix, one obtains the following payoffs. The proposition is obtained by straightforward comparison of these.

Public		$p \geq \frac{1}{2}$	$p \in [\frac{1}{3}, \frac{1}{2})$	$p \in [\frac{1}{4}, \frac{1}{3})$	$p < \frac{1}{4}$
Investor		$\frac{7}{3}p - \frac{1}{3}$	$\frac{11}{3}p - 1$	$\frac{8}{3}p - \frac{2}{3}$	0
E_g		1	$\frac{7}{9}$	$\frac{10}{9}$	0
E_b		$\frac{19}{6}$	$\frac{5}{2}$	2	0

Private		$p \geq \frac{2}{5}$	$p \in [\frac{1}{4}, \frac{2}{5})$	$p < \frac{1}{4}$
Investor		$\frac{7}{3}p - \frac{1}{3}$	$4p - 1$	0
E_g		1	$\frac{6-9p}{8-14p}$	0
E_b		$\frac{19}{6}$	$\frac{5}{2} + \frac{3p^2}{(1-p)(4-7p)}$	0

Expected payoffs in Example 2.

Notice that the bad Entrepreneur repays the first investment only if $M > \hat{M} = 2$. Thus, $F(\hat{M}) = \frac{1}{2}$. By applying the numerical values to the ex-ante expected payoffs detailed earlier in the appendix, one obtains the following payoffs. The proposition is obtained by straightforward comparison of these.

Public		$p \geq \frac{1}{2}$	$p \in [\frac{1}{3}, \frac{1}{2})$	$p \in [\frac{1}{4}, \frac{1}{3})$	$p < \frac{1}{4}$
Investor		$2p - \frac{1}{2}(1 - p)$	$\frac{7}{2}p - 1$	$2p - \frac{1}{2}$	0
E_g		2	$\frac{3}{2}$	$\frac{3}{2}$	0
E_b		4	3	2	0

Private		$p \geq \frac{1}{3}$	$p < \frac{1}{3}$
Investor		$2p - \frac{1}{2}(1 - p)$	0
E_g		2	0
E_b		4	0

References

- Alonso, R. and N. Matouschek (2007) "Relational Delegation". *RAND Journal of Economics*, 38 (4), 1070-1089.
- Alonso, R. and N. Matouschek (2008) "Optimal Delegation". *Review of Economic Studies*, 75, 259-293.
- Anderhub V., D. Engelmann, W. Güth (2002) "An experimental study of the repeated trust game with incomplete information". *Journal of Economic Behavior & Organization*, 48, 197-216.
- Armstrong, M., and J. Vickers (2010) "A Model of Delegated Project Choice". *Econometrica*, 78 (1), 213-244.
- Baker, G., R. Gibbons, and K. Murphy (1999) "Informal Authority in Organizations". *Journal of Law, Economics & Organization*, 15 (1), 56-73.
- Bendor, J., A. Glazer, and T. Hammond (2000) "Theories of Delegation". *Annual Review of Political Science*, 4, 235-269.
- Berg, J., J. Dickhaut, and K. McCabe (1995) "Trust Reciprocity, and Social History". *Games and Economic Behavior*, 10, 122-142.
- Bester, H., and D. Krähmer (2008) "Delegation and incentives". *RAND Journal of Economics*, 39 (3), 664-682.
- Brandts, J, and N. Figueras (2001) "An exploration of reputation formation in experimental games". *Journal of Economic Behavior and Organization*, 50 (1), 89-115.
- Camerer, C. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Camerer, C. and K. Weigelt (1988) "Experimental tests of the sequential equilibrium reputation model". *Econometrica*, 56, 1-36.
- Chevalier, J. and G. Ellison (1999) "Risk taking by mutual funds as a response to incentives". *Journal of Political Economy*, 105 (6), 1167-1200.
- Colombo, F., and G. Merzoni (2006) "In praise of rigidity: The bright side of long-term contracts in repeated trust games". *Journal of Economic Behavior and Organization*, 59, 349-373.

- Cochard, F., P.N. Van, and M. Willinger (2004) "Trusting behavior in a repeated investment game". *Journal of Economic Behavior & Organization*, 55 (1), 31-44.
- Datta, S. (1996) "Building Trust". Discussion paper, STICERD, London School of Economics.
- Englmaier, F. and C. Segal (2013) "Unions, Communication, and Cooperation in Organizations". mimeo, July 2013.
- Ghosh, P., and D. Ray (1996) "Cooperation in community interaction without information flows". *Review of Economic Studies*, 63, 491-519.
- Halac, M. (2014) "Investing in a Relationship". mimeo, Columbia University.
- Huberman, G. and S. Kandel (1993) "On the incentives for money managers: a signalling approach". *European Economic Review*, 37, 1065-1081.
- Huddart, S. (1999) "Reputation and performance fee effects on portfolio choice by investment advisers". *Journal of Financial Markets*, 2 (3), 227-271.
- Kranton, R. (1996) "The formation of cooperative relationships". *Journal of Law, Economics, and Organization*, 12, 214-233.
- Kreps, D. and R. Wilson (1982) "Reputation and imperfect information". *Journal of Economic Theory*, 27, 253-279.
- Levin, J. (2003) "Relational Incentive Contracts". *American Economic Review*, 93 (3), 835-850.
- Levy, G. (2007a) "Decision Making in Committees: Transparency, Reputation, and Voting Rules". *American Economic Review*, 97 (1) 150-168.
- Levy, G. (2007b) "Decision-Making Procedures for Committees of Careerist Experts". *American Economic Review*, 97 (2) 306-310.
- Li, J. and N. Matouschek (2013) "Managing Conflicts in Relational Contracts". *American Economic Review*, 103 (6), 2328-2351.
- Lukas, C., and J. R. Schöndube (2012) "Trust and adaptive learning in implicit contracts". *Review of Management Science*, 6, 1-32.

- Marino, A., J. Matsusaka, and J. Zabochnik (2010) "Disobedience and Authority". *Journal of Law, Economics and Organization*, 26 (3), 427-459.
- Matozzi, A. and A. Merlo (2007) "The Transparency of Politics and the Quality of Politicians". *American Economic Review*, 97 (2), 311-315.
- Miller, G. (2001) "Why is Trust Necessary in Organizations? The Moral Hazard of Profit Maximization". in: Cook, K. ed. *Trust in Society*, New York: Russell Sage Foundation, pp. 307-331.
- Neral, J. and J. Ochs (1992) "The sequential equilibrium theory of reputation building: a further test". *Econometrica*, 60, 1151-1169.
- Prat, A. (2005) "The Wrong Kind of Transparency". *American Economic Review*, 95 (3), 862-877.
- Sobel, J. (1985) "A Theory of Credibility". *Review of Economic Studies*, 52 (4), 557-573.
- Stracca, L. (2006) "Delegated Portfolio Management: A Survey of the Theoretical Literature". *Journal of Economic Surveys*, 20 (5), 823-848.
- Xie H. and Y.J. Lee (2012) "Social Norms and Trust Among Strangers". *Games and Economic Behavior*, 76, 548-555.
- Yang, H. (2013) "Nonstationary Relational Contracts with Adverse Selection". *International Economic Review*, 54 (2), 525-547.
- Watson, J. (2002) "Starting Small and Commitment". *Games and Economic Behavior*, 38, 176-199.
- Watson, J. (1999) "Starting Small and Renegotiation". *Journal of Economic Theory*, 85, 52-90.
- Watson, J. and J.E. Rauch (2003) "Starting small in an unfamiliar environment". *International Journal of Industrial Organization*, 21, 1021-1042.



ISBN 978-952-60-5987-7 (printed)
ISBN 978-952-60-5988-4 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Business
Department of Economics
www.aalto.fi

**BUSINESS +
ECONOMY**

**ART +
DESIGN +
ARCHITECTURE**

**SCIENCE +
TECHNOLOGY**

CROSSOVER

**DOCTORAL
DISSERTATIONS**