Title
Essays on Reputation and Learning in Markets and Networks

Author
Zhang, Simpson

Publication Date
2016

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Essays on Reputation and Learning in Markets and Networks

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

by

Simpson Zhe Zhang

2016

Abstract of the Dissertation

Essays on Reputation and Learning in Markets and Networks

by

Simpson Zhe Zhang

Doctor of Philosophy in Economics

University of California, Los Angeles, 2016

Professor Moritz Meyer-ter-Vehn, Chair

This dissertation addresses key topics in the economic study of reputation and learning. Many real world interactions occur among agents who are only partially informed about the qualities of other agents. Therefore the agents must learn about each other over time as they interact. Over time, as more information is learned, a reputation will develop for an agent, and this reputation will inform the decisions of other agents to interact with them. Thus understanding the nature and impact of reputational effects in dynamic settings is crucial in developing more realistic models of the world.

The first chapter, based on my paper "Reputational Learning and Network Dynamics" with Mihaela van der Schaar, considers learning and reputation in a networks model. In many real world networks agents are initially unsure of each other's qualities and must learn about each other over time via repeated interactions. This chapter provides a methodology for studying the dynamics of such networks, taking into account that agents differ from each other, that they begin with incomplete information, and that they must learn through past experiences which connections/links to form and which to break. The network dynamics in the model vary drastically from the dynamics in models of complete information. With incomplete information and learning, agents who provide high benets will develop high reputations and remain in the network, while agents who provide low benets will drop in reputation and become ostracized. We show that the information to which

agents have access and the speed at which they learn and act can have a tremendous impact on the resulting network dynamics. Using the model, we can also compute the ex ante social welfare given an arbitrary initial network, which allows us to characterize the socially optimal network structures for different sets of agents. Importantly, we show through examples that the optimal network structure depends sharply on both the initial beliefs of the agents, as well as the rate of learning by the agents. Due to the potential negative consequences of ostracism, it may be necessary to place agents with lower initial reputations at less central positions within the network.

The second chapter, based on my paper "Optimal Production Choice with Reputational Concerns", considers a model where a single firm's reputation evolves alongside continuous time as it sells a product. Consumers are initially uncertain about the firm's quality and learn about it through the purchasing the firm's product. The firm chooses its output at every point in time while taking into account that higher output levels will allow the market to know more information about its quality. Two cases are analyzed, with both informed and uninformed firms. Uninformed firms are shown to have convex value functions and always produce more than is profit maximizing. In the informed firm case, a low quality firm has an less of an incentive to produce than a high quality firm because it does not wish to give the market information.

The third chapter is based on my paper "A Dynamic Model of Certification and Reputation" with Mihaela van der Schaar, and is reprinted with permission from the journal Economic Theory. It considers learning in a market that occurs through reputation and certification simultaneously, which allows us to investigate the rich interplay and dynamics that can arise. Our work offers four main insights: (1) Without certication, market learning through reputation alone can get "stuck" at inefcient levels and high-quality agents may get forced out of the market. (2) Certication "frees" the reputation of agents, allowing good agents to keep working even after an unfortunate string of bad signals. (3) Certication can be both benecial and harmful from a social perspective,so a social planner must choose the certication scheme carefully. In particular, the market will tend to demand more certication than socially optimal because the market does not bear the certication costs. (4) Certication and reputational learning can act as complementary forces so that the social welfare produced by certication can be increased by faster information revelation.

The dissertation of Simpson Zhe Zhang is approved.

Sushil Bikhchandani

Marek G. Pycia

Mihaela van der Schaar

Moritz Meyer-ter-Vehn, Committee Chair

University of California, Los Angeles

2016

To my mother

who has always dedicated

herself to my success

Table of Contents

List of Figures

Acknowledgments

I am very grateful to my committee for all their help over the years. I am especially thankful to my chair Moritz Meyer-ter-Vehn for his invaluable comments, as well as Mihaela van der Schaar for her incredible dedication and spirit. This research would not have been possible without their guidance.

For the first chapter, which was co-authored with Mihaela van der Schaar, I am deeply indebted to Jie Xu, William Zame, and Moritz Meyer-ter-Vehn, and Peng Yuan Lai for their valuable assistance. The chapter also benefited from discussions with Matt Jackson and seminar participants at the 2015 NEGT Conference and the 2015 SWET conference. I gratefully acknowledge financial support from the National Science Foundation Grant CCF-1218136.

For the second chapter, I received very valuable comments from Moritz Meyer-ter-Vehn, Simon Board, and Ichiro Obara. I would also like to thank Sushil Bikhchandani, Joseph Ostroy, and Marek Pycia from their help.

For the third chapter, which was co-authored with Mihaela van der Schaar, I received valuable guidance from Moritz Meyer-ter-Vehn, Simon Board, Oren Louidor, and Bill Zame. The chapter also benefited from valuable discussions with Jeff Ely, Crysanthos Dellarocas, Sanjeev Goyal, Johannes Hörner, David Parkes, Yangbo Song, Jie Xu, and Kartik Ahuja. I gratefully acknowledge financial support from the ONR. This chapter is reprinted with permission from the journal Economic Theory and Springer Publishing.

VITA

2008-2009	Research Assistant, Experimental Economics Lab, California Institute of Technology.
2008-2010	Research Assistant, Social Cognition Lab, Duke University.
2010	B.S. (Economics) and B.S. (Mathematics), Duke University.
2011	M.S. (Economics), UCLA.
2011–2013	Teaching Assistant, Economics Department, UCLA.
2011–2012	Economics Assistant, Office of Financial Research, US Department of Treasury.
2013–2016	Graduate Student Researcher, Electrical Engineering Department, UCLA.

PUBLICATIONS

A Dynamic Model of Certification and Reputation (with Mihaela van der Schaar) Economic Theory, 2014

Distributed Multi-Agent Online Learning Based on Global Feedback (with Jie Xu, Cem Tekin, and Mihaela van der Schaar) IEEE Journal of Selected Topics in Signal Processing, 2015

Distributed Online Learning in Social Recommender Systems (with Cem Tekin and Mihaela van der Schaar) IEEE Journal of Selected Topics in Signal Processing, 2014

CHAPTER 1

Reputational Learning and Network Dynamics

1.1 Introduction

Networks are pervasive in all areas of society, ranging from financial networks to organizational networks to social networks. And an important feature of many real world networks is that agents do not fully know the characteristics of others and must learn about them over time. For instance a bank learns about the credit-worthiness of a new borrower, a worker in a firm learns about the ability of a coworker, and a buyer learns about the product quality of a supplier. Such learning can strongly affect the resulting shape of the network. As agents receive new information, they may revise their beliefs about other agents, update their linking decisions, and cause the network to evolve as a result. To properly analyze such network evolution, it is crucial to understand the exact mechanism by which learning impacts network dynamics.

The impact of agent learning on network evolution has not been well studied in the existing literature. A large network science literature analyzes the effect of learning on fixed networks that have already formed (see e.g. Scott [Sco12]). A smaller microeconomics literature[1] studies the formation of networks - but makes very strong assumptions (e.g., homogeneous agents/entities, complete information about other agents). Neither the network science literature nor the microeconomics literature has so far taken into account that agents behave strategically in deciding what links to form/maintain/break *and* that they also begin with incomplete information about others, so they must learn about others through their interactions. As a result, neither network science nor microeconomics provides a complete framework for understanding, predicting and guiding the

[1] See the overview in Jackson [Jac10] for instance.

formation (and evolution) of real networks and the consequences of network formation.

The overarching goal of this chapter is to develop such a framework. An essential part of the research agenda is driven by the understanding that individuals within a network are heterogeneous - some workers are more productive than others, some friends are more helpful than others, and some borrowers are more creditworthy than others. Furthermore, these characteristics are not known in advance but must be learned over time via repeated interactions. The rate of learning itself may also be strongly influenced by the network structure: agents engaged in more interactions are likely reveal more information about themselves.

As a motivating example, consider a group of financial institutions that are linked together in a financial network[2]. These financial institutions provide benefits to each other by engaging in mutually beneficial trading opportunities, such as providing each other with liquidity or engaging in joint ventures[3]. High quality firms will reliably carry through the terms of trade, but low quality firms are more likely to default and harm their neighbors. Each institution will thus only continue to link with another institution (over time) if the counterparty is believed to be of sufficiently high quality. As time progresses, the institutions *observe* the actions of their counterparties, *update* their beliefs about the quality of each counterparty, and *change* their linking decisions as a result. In this way, learning by the financial institutions causes the network topology to evolve over time. The network topology also impacts the rate of learning, as an institution with more connections interacts with more counterparties, and thus its neighbors have more observations to learn from. While having more connections opens an institution up to more beneficial opportunities, it also carries the risk of causing the institution to be shut out of the network more quickly if it starts defaulting, as in the case of Lehman Brothers due to its exposures to the subprime mortgage market during the 2008 financial crisis.

Our model takes into account the features of the previous example: agents behave strategically, begin with incomplete information about each other, and must learn through continued interactions which connections to form and maintain and which to break. We consider a continuous time model

[2]Our model can also be applied to a wide range of other networks, such as organizational networks, social networks, or expertise networks. We discuss some implications for these settings as well throughout the chapter.

[3]As in the model of Erol [Ero15].

with a group of agents who are linked according to a network and who send noisy flow benefits to their neighbors. The benefits that agents provide could be interpreted for instance as the benefits that financial institutions derive from providing liquidity to each other or from diversifying risk with each other's specialized assets. Each agent is distinguished by a fixed quality level which determines the average value of the flow benefits it produces. Agents observe all the benefits that their neighbors produce, and they update their beliefs about a neighbor's quality via Bayes rule. Neighbors with more connections will reveal more information about themselves over time. Agents will maintain links with neighbors that provide high benefits, but will cut off links with neighbors that provide low benefits. Thus the network evolves as agents learn about each other and update their beliefs. Since the amount of links an agent has determines the rate of learning about that agent, the rate of learning about an agent changes as the network changes, leading to a co-evolution of the network topology and information production.

Our model is highly tractable and allows us to completely characterize network dynamics and give explicit probabilities of the network evolving into various configurations. In addition, we are able to describe the entire set of stable networks and analytically compute the probability that any single stable network emerges. This allows for predictions regarding which types of stable networks are likely to emerge given an initial network.

We also study the implications that learning has on the social welfare and efficiency of a network. Our results show that learning has a beneficial aspect: agents that are of low quality are likely to produce low signals and will eventually be ostracized from the network. Learning also has a harmful aspect: even high quality agents may produce an unlucky string of bad signals and so be forced out of the network. Moreover, even having low quality agents leave the network can reduce overall social welfare. A marginally low quality agent may harm its neighbor slightly, but it also receives a large benefit if its neighbor is of very high quality. Thus if the low quality agent leaves the network, the overall social welfare would actually decrease. The issue here is that agents only care about the benefit their neighbors are providing them, but not the benefit they are providing their neighbors. Thus, there is a negative externality every time a link is severed[4]. In many

[4]The negative effects of ostracism can be particularly acute in financial networks during times of distress in which

situations, the negative effects of learning outweigh the positive effects, so on balance learning is actually harmful. In particular, increasing the learning rate about marginal agents whose neighbors are high quality agents is bad, because forcing the marginal quality agent out of the network sacrifices the social benefit of the link to the high quality agent. However, increasing the rate of learning about a marginal quality agent whose neighbors are also marginal quality agents is good, because more information will be revealed about that marginal quality agent, allowing its neighbors to more quickly sever their links to it. Thus the impact of learning can be either positive or negative depending on the specific network.

Our welfare results have important implications for network planning and are useful in a diverse range of settings, such as in guiding the formation of networks by the policies of a financial regulator, human resources department, online community, etc. Due to the varying effects of learning, we show that the optimal network structure will be quite different for different groups of agents. For instance, when agents all have high qualities, the optimal network will be fully connected (which allows all agents to benefit fully from their repeated interactions). On the other hand, if some agents have low initial reputations, then a fully connected network may not be optimal because it will be desirable to isolate low quality agents (or clusters of low quality agents). Because of the negative effects of learning, it may be optimal to prevent two agents from linking with each other, even if such agents have initial expected qualities higher than the linking cost. If such agents did link, they would both send more information about themselves through this link, causing them to be ostracized more quickly. Thus each agent, as well as the overall network, may become worse off through the formation of this link due to the faster learning caused by the link. In some cases, a star or a core-periphery network would generate higher social welfare than a complete network even when all agents have initial expected qualities higher than the linking cost. Such a situation arises for instance if there are two separate groups of agents, one group with very high expected quality and the other group with moderate expected quality. By placing the high quality agents in the core and the moderate quality agents in the periphery, the high quality agents are able to

banks get shut out of funding, as is the case of a liquidity freeze. Ostracism has also been demonstrated in a wide variety of social settings in the social psychology literature. We discuss this literature and our model's implications in the Literature Review section.

produce large benefits for the network, and the potential harm from the moderate quality agents is minimized. This is a new information-based reason for the benefits of a core-periphery network, in contrast to other non-informational reasons given in the networks literature.

Finally, we consider four extensions of our model that allow for even richer network dynamics and learning. In the first extension, we allow the mechanism designer to provide the agents with a subsidy that encourages linking[5]. The effect of such a subsidy is to promote the amount of experimentation done by the agents, and we show that a large enough subsidy can always improve overall social welfare because of this. In the second extension, we allow for agents with high enough reputations to form new links with each other, and we show that social welfare will be increased when the linking threshold is set high enough. In the third extension, we allow new agents to enter the network over time, and we consider the optimal time at which new agents should arrive. We show that all agents should be allowed to enter the network eventually, but delayed entry can be desirable in certain types of networks. Lastly, in the fourth extension we allow for agents that have been ostracized in the past to re-enter the network after a set period of time, and we show that the negative effects of learning can be mitigated if re-entry occurs frequently enough.

1.2 Literature Review

This chapter represents a novel contribution to the network formation literature, by being among the first to consider incomplete information and learning in networks, as well as by providing a tractable model that allows for the computation of many properties, including the *ex ante* social welfare, of different network topologies. Other papers in the network literature have usually studied network dynamics only in settings of complete information when agents perfectly know each other's qualities. For example, the papers by Jackson and Wolinsky [JW96], Bala and Goyal [BG00], Watts [Wat01], and Galeotti and Goyal [GGJ10] all consider networks where the agents have complete information. In these models, agents are aware of the exact qualities of all other agents and there is no learning. The network dynamics arise instead from externalities and indirect

[5]For instance a financial regulator could guarantee transactions within a financial network to make them less risky.

benefits between agents that are not directly linked. As one link is severed or formed, the benefits produced by other links changes as well, which causes other agents to sever or form their links in a chain reaction. For some networks, such as informational networks, these indirect benefits seem important, as an agent who has many neighbors will likely be able to produce higher quality information as well. However, in other networks such as friendship networks these indirect benefits are less relevant and it is the quality of each specific agent that matters the most. This is especially applicable when a new group of agents is meeting for the first time, as it is likely to be the mutual learning and agent beliefs that drive the evolution of the network. We argue that the network dynamics in such situations are more greatly dependent on reputational effects and mutual learning than on changes in the values of indirect benefits.

We do not assume any indirect benefits in our model and focus instead on the dynamics resulting from incomplete information and learning. Agent learning strongly influences the network formation process in a way that would not arise with complete information. Agents that send good signals will develop high reputations and remain in the network, whereas agents that send bad signals will develop low reputations and eventually become ostracized by having their neighbors cut off links. The rate of learning about an agent's quality affects how quickly the network evolves and thus has a strong effect on the resulting social welfare. With complete information however, such dynamics would not occur because agents would know each other's qualities perfectly at the onset. For instance, Watts [Wat01] considers a dynamic network formation model where agents form links under complete information. When there are no indirect benefits between agents in that paper's model, each agent would make a one time linking decision with any other agent and never update its choice later on. But with learning, agents may change their linking choices by breaking off links with neighbors that consistently produce low benefits. Thus incomplete information causes links to fluctuate dynamically over time as new information arrives and beliefs are updated, instead of staying static as in the complete information case. We argue that such effects are key and even the main driver of dynamics when a group of agents are meeting for the first time and forming a network with each other.

In addition, the tractability of our model allows us to explicitly compute the social welfare for different network structures even under incomplete information. This tractability arises from the use of continuous time in our model and the choice of Brownian motion as the information process, which allows for closed form equations of the probabilities that different networks emerge. In contrast other networks papers such as Jackson and Wolinsky [JW96] and Bala and Goyal [BG00] use discrete time models that do not allow for such clean closed form expressions. While these other papers analyze the efficiency properties of a given fixed network, our welfare results are much stronger and allow the network to evolve endogenously over time as agents learn and update their linking decisions. This enables us to compare the *ex ante* optimality of different initial network structures, as well as provide general results for when certain networks are optimal. For instance, we show that when the rate of learning in the network is either very slow or very fast, a complete network will be optimal if the agent's initial expected qualities are all higher than the cost of maintaining the links. But when learning is at an intermediate rate, it may be optimal not to have all agents connected with each other even if their expected qualities are higher than the linking cost, due to the externalities associated with learning. Such a result cannot arise under complete information, where if agent's qualities are all perfectly known it would be strictly better for all of them to be linked initially.

This chapter is also tied to the literature on observational learning in networks. Works such as Golub and Jackson [GJ10], Golub and Jackson [GJ12] and Acemoglu et al ([ADL11] analyze observational learning in social networks, in which there is a fixed exogenous network on which the agents interact, and the agents learn about an exogenous state of the world through this network by observing the actions of neighbors. These papers provide results regarding the speed and accuracy of the observational learning that can be achieved by agents connected through different types of networks. This chapter is significantly different from this literature because agents learn about other agents' *qualities* instead of an *exogenous* state of the world. As such, agents will wish to update their linking decisions over time as their beliefs about the agents with whom they are connected with change. Thus, in our paper the network and learning *co-evolve*, causing the network structure to evolve *endogenously*.

A paper that does consider the diffusion of information about agents across a network is Vega-

Redondo [Veg06]. This paper focuses on the issue of moral hazard and monitoring. It assumes that players engage in bilateral prisoner's dilemma games. Information about player actions diffuses through the network, and agents are able to sustain cooperation through punishing defectors. More densely connected networks allow for faster information transmission and thus sustain higher levels of cooperation. The paper analyzes how the structures of networks that emerge is affected by this transmission of information, and shows through simulations and mean-field analysis that the inclusion of network based information can increase network density. Our work instead focuses on adverse selection and on learning about agent types. We show that more information can be harmful for welfare because it leads to greater ostracization among agents. The tractability of our model also allows us to consider the social welfare generated across the entire path of network evolution, as opposed to the welfare of the long run average network. We are thus able to address issues of network design, and we characterize the optimal initial network structure under different environments.

There have also been recent related papers in the financial literature studying the impact of learning and information on the structures of financial networks, such as Blasques et al (2014) and Babus (2013). Many of these papers seek to explain the core-periphery structures of financial networks that have been well documented empirically. These papers show that features such as dealer heterogeneity will lead to such structures. However, the focus of these papers is either on complete information settings where the types of other agents are directly observable, or on learning through investing in information gathering (by ensuring repayment of debt) rather than on learning being affected by the network structure itself. These papers show that since networks can allow for mutual monitoring by financial institutions, they can also lead to more efficient trading. The benefits that a network provides also leads to greater stability over time. The settings of these papers are longer term than our own and are more applicable to stable financial market circumstances where informational uncertainty about counterparties is low. We view our model instead as describing a short time period with a lot of uncertainty, such as in the aftermath of a financial crisis when banks are very unsure of the solvency of other banks, due to the difficulty in assessing the quality of their assets for instance. In such situations, banks will be hesitant to trade with each other and will very carefully try to learn about the solvency of other banks through observations of repayments, which affect each bank's reputation. Banks that obtain low reputations may get shut out of the funding market entirely during liquidity runs, as was the case during the collapse of Lehman Brothers in the recent financial crisis. Thus it is important for a financial regulator to carefully structure the trading network and control the interactions so that such situations can be mitigated.

A closely related networks paper that involves learning with adverse selection is Song and van der Schaar [SS15]. Like us, this paper also considers learning by agents about the types of other agents within a network, and it shows how incomplete information and the learning process can lead to a wide variety of network structures and dynamics. However, this paper considers a discrete time model and incorporates a different type of learning process in which information is revealed *immediately*, after a single interaction. On the contrary, information is revealed gradually in our model, allowing linking decisions and learning to occur simultaneously. Since learning takes place gradually instead of instantaneously, we can derive results about how the speed of learning affects network dynamics and social welfare. Most importantly, our model assumes that the network structure itself impacts the rate of learning about agents. This assumption is more realistic as learning is often affected by an agent's locations within the network itself, and it has strong implications. We show that it necessitates the need for careful planning by a network designer to properly control the learning by agents. In addition, our use of continuous time allows our model to be more tractable and able to provide explicit characterizations of the social welfare of different network structures.

Finally, we note that the negative effects of learning in our model can be linked with the damaging impacts of ostracism found in the social psychology literature. Social ostracism is a prevalent force that has been documented in the social psychology literature in numerous settings ranging from online interactions to office workplaces. As Williams and Sommer [WS97] state, "Social ostracism is a pervasive and ubiquitous phenomenon." In this literature, ostracism can also occur when an agent's perceived quality drops too low, and will have harmful effects on the agent itself. As the paper by Wesselman et al [WWP13] notes, "Ostracism is a common, yet painful social experience...Individuals who do not fit the group's definition of a contributing member may find themselves a likely candidate for *punitive* ostracism". That paper shows the occurrence of ostracism via an online experiment, where agents differed in their ability to play a game, and agents who play badly became ostracized by the others. This is similar to our model, where agents who are learned to be of low quality are ostracized. Ostracism can also occur in workplaces, as some employees may be ostracized by their coworkers. Robinson et al [ROW12] notes that "not only are such experiences extremely painful, but under some circumstances they can have an even greater negative impact than other harmful workplace behaviors such as aggression and harassment." Thus it is important for companies to consider the harmful effects of ostracism that can occur through workplace interactions. We provide guidelines for minimizing the negative effects of ostracism through placing lower reputation agents in less central positions of the network.

1.3 Model

1.3.1 Overview

We consider an infinite horizon continuous time model with a finite set of agents denoted by $V = \{1, 2, ..., N\}$. At every moment in time, the agents choose which other agents to link with, and a link is established only through mutual consent. These choices are made subject to an underlying network constraint $\Omega = \{\omega_{ij}\}$ that specifies the pairs of agents that are able to link with each other[6]. For each pair of agents $\omega_{ij} = 1$ if agents i and j can connect with each other and $\omega_{ij} = 0$ otherwise. We call agents i and j neighbors if they can connect. Initially (time $t = 0$), agents are linked according to a network $G^0 = \{g_{ij}^0\} \subseteq \Omega$. As the network will change over time, we denote G^t as the network at time t. Moreover, we let $k_i^t = \sum_j g_{ij}^t$ be the number of links that agent i has at time t, and we let K_i^t denote the set of neighbors of agent i at time t.

Agents receive flow payoffs from each link equal to the benefit of that link minus the cost. Each agent i must pay a flow cost c for each of its links that is active. Hence, at time t, agent i pays a total

[6]This network constraint Ω may arise from the specific interests/desires of the agents regarding who they want to link with, or from potential physical/geographical constraints that limit agents from linking. It may also be planned, e.g. by the human resources department in a company for a network of employees, or through the policies of a financial regulator for a network of financial institutions.

cost of $k_i^t c$ for all its links. Agents also obtain benefits from their links, depending on their linked neighbors' qualities q_i. However each agent's true quality is initially unknown to all agents, and we do not require that agents know their own qualities. At the start of the model, each agent i's quality q_i is drawn from a commonly known normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$ with $\mu_i > c$. Both the mean and the variance are allowed to vary across agents, and several of our results below will utilize this heterogeneity. Agent i generates a different noisy benefit $b_{ij}(t)$ for each agent j that is linked to it, and these benefits follow a Brownian motion $db_{ij}(t) = q_i dt + \tau_i^{-1/2} dZ_{ij}(t)$, where the drift is the true quality q_i and the variance depends on τ_i, an exogenous parameter we call the *signal precision* of agent i[7]. $Z_{ij}(t)$ is a standard zero-drift and unit variance Brownian motion, and represents the random fluctuations in the benefits of each interaction. $Z_{ij}(t)$ is assumed to be independent over all i and j, and thus all the benefits produced by agent i are conditionally independent given q_i. We assume that all the benefits that agent i produces are observed by all the neighbors of i, which ensures that agent i's neighbors all have the same beliefs about i at any point in time (information is locally public among agent i's neighbors)[8]. For each agent i, we define the agent's benefit history as the history of all previous benefits, $\mathcal{H}_i^t = \{b_{ij}^{t'}\}_{t'=0}^{t}$.

We assume that agents are myopic, and they thus consider only the current flow benefit when making linking decisions[9]. Each agent's utility is assumed to be linear in the benefits provided by each link and the linking cost. This also implies that agents are risk neutral and so consider the expectation over neighbor qualities when there is uncertainty. The flow utility of agent i at any

[7]We can think of the signal precision as representing how much information the agent reveals about itself in each interaction, with a higher precision corresponding to more information. It could depend on the type of interaction with the agent (e.g. close partnerships or chance encounters), or factors like the agent's personality.

[8]This is an important assumption to maintain the tractability of the model. It can be interpreted, for instance, through an online expertise network where the output of agent i is public, so that all neighbors of agent i can judge the benefit that i has provided to all its links. Or in an offline setting, we could assume that the neighbors of agent i are continuously discussing the benefits they have received from i with all other neighbors of i, so that the neighbors maintain the same beliefs. For most of our results, the information does not need to be fully public; the information regarding agent i needs only be available to all the direct neighbors of agent i.

[9]Such an assumption is common within the networks literature to maintain tractability, see Jackson and Wolinsky (1996) or Watts (2001) for instance. Myopia is an appropriate assumption in financial networks where firm managers have myopic incentives. Such myopic incentives have been documented empirically in papers such as Jacobson (1993) and Mizik (2010). We relax this assumption in the extensions section where we allow for subsidies that change agent linking strategies.

time t is given by the following equation:

$$U_i = \sum_{\{j \in K_i^t\}} (E[q_j | \mathcal{H}_i^t] - c) \tag{1.1}$$

1.3.2 Reputation and Learning Speed

Since we have assumed a Brownian motion process, a sufficient statistic for all the individual link benefits is the average benefit per link produced by agent i up to time t, which we denote as $B_i(t)$. Given our above assumptions, $B_i(t)$ follows a Brownian motion $dB_i(t) = q_i dt + (k_i^t \tau_i)^{-1/2} dZ_i(t)$ where the drift rate is the true quality q_i, the instantaneous volatility rate $(k_i^t \tau_i)^{-1/2}$ depends on the number of links agent i has at time t, and $Z_i(t)$ is the standard Brownian motion with zero-drift and unit-variance. Importantly, this equation shows that the more links an agent has, the lower its volatility rate and the faster its true quality q_i is learned. This is because an agent with more links produces more individual benefits, and so the average over all benefits is more precise. Note also that an agent with no links would not send any information, and thus there would be no learning about its quality. Therefore the topology of the network strongly affects the rate of learning about each agent's quality.

If at time t all links of agent i are severed, then no benefit will be produced by agent i and this will be denoted as $b_i^t = \emptyset$. In this case no information is added and hence, the Brownian motion of agent i is stopped at the current level. As mentioned, there is a prior belief of an agent i's quality $\mathcal{N}(\mu_i, \sigma_i^2)$, and agents will update this belief in a Bayesian fashion in light of the observations of flow benefits. These observations combined with the prior quality distribution will result in a posterior belief distribution of agent i's quality $f(q_i | \mathcal{H}_i^t)$ which is also normally distributed[10]. We denote $\mu_i^t = E[q_i | \mathcal{H}_i^t]$ as the expected quality of agent i given the history $\mathcal{H}_i^t$ and call it the *reputation* of agent i at time t. The reputation represents the expected flow benefit of linking with agent i at time t.

We have assumed that agents are myopic. Therefore, to maximize flow utilities, agent i will

[10] As mentioned a sufficient statistic for the entire history is $B_i(t)$, so a neighbor only needs to know $B_i(t)$ in order to calculate this posterior.

cut off its link with agent j once agent j's reputation μ_j^t falls below the linking cost c. Since we assume all agents have homogeneous linking costs, and all neighbors have the same beliefs, any other agent that is linked to j will also decide to sever its link. From this moment on, agent j is effectively ostracized from the network; since it no longer has any links it cannot send any further information that could potentially improve its reputation[11]. While in the base model an ostracized agent cannot return to the network, we relax this assumption in the extensions section.

1.4 Network Dynamics and Stability

1.4.1 Network Dynamics

The dynamics of the model evolve as follows: all pairs of agents that are neighbors according to the network constraint Ω will choose to link at time zero, since we have assumed that all agents have initial reputations higher than the cost c (any agent with an initial reputation lower than c is immediately ostracized from the network and would not need to be considered). Thus the initial network at time 0 will be the same as the network constraint, $G^0 = \Omega$. Over time agents that send bad signals will have their reputations decrease, and once an agent's reputation hits c its neighbors will no longer wish to link with it. All its neighbors will sever their links and the agent is effectively ostracized from the network. We will show that this always happens for an agent with true quality $q_i \leq c$, and will still happen with positive probability for an agent with quality $q_i > c$. The ostracization of an agent will affect its former neighbors as well. Since they now have once less link each, they will produce information about themselves more slowly than before, and so their reputations will be updated less quickly.

The remaining agents in the network will continue to link and send signals until someone else's reputation drops too low and that agent is also ostracized. This process will continue until the qualities of all the remaining agents are known with very high precision and in the limit their

[11]Although ostracism may seem harsh, as we noted earlier ostracism is a prevelant phenomenon that has been widely studied in the social psychology literature, in settings ranging from online interactions to office workplaces. Furthermore, in financial networks low reputation institutions may get shut out of funding completely during liquidity crisis.

reputations no longer change. Since agent qualities are fixed, by the law of large numbers, any agents that remain in the network will have their qualities learned perfectly in the limit as, $t \to \infty$, and the network will tend towards a limiting structure that we call the *stable network*. The next section will explicitly characterize these stable networks, but we note that many different stable networks could potentially emerge depending on the true qualities of the agents and the signals they produce.

Figure 1.1 shows the different network dynamics that could emerge even if the initial reputations of the agents are fixed, due to the uncertainty about the true qualities of the agents as well as the randomness in the signals they send. From the same initial reputations for the red and blue agents, many different network dynamics and stable networks are possible. In the top graph the red (larger circle and bolded line) agent has a true quality less than c and so will be ostracized from the network for certain at some time, while the blue (smaller circle and thin line) agent has a true quality above c and so may or may not be ostracized from the network depending on the signals it sends. Each event leads to a different stable network, either with and one without the blue agent. In the bottom graph it is the blue agent who has a true quality lower than c and so will be ostracized for sure, whereas the red agent could potentially stay in the network indefinitely.

1.4.2 Stable Networks

As mentioned, we call the limiting network structure as t goes to infinity, denoted by G^∞, a stable network. Formally, let $G^\infty \equiv \lim_{t\to\infty} G^t$. This limiting structure always exists since agent qualities are fixed, so by the law of large numbers any agent that remains in the network will have its quality learned to an arbitrary precision over time. The probability that an agent who is still in the network at time t ever becomes ostracized must therefore tend to zero as $t \to \infty$ (we show this analytically below). Which specific stable network eventually emerges is random and depends on the signal realizations of each agent. The tractability of our model allows us to explicitly characterize the set of stable networks that could emerge given a set of agents and a network constraint Ω, as well as the impact of the rate of learning on the probability distribution over stable networks.

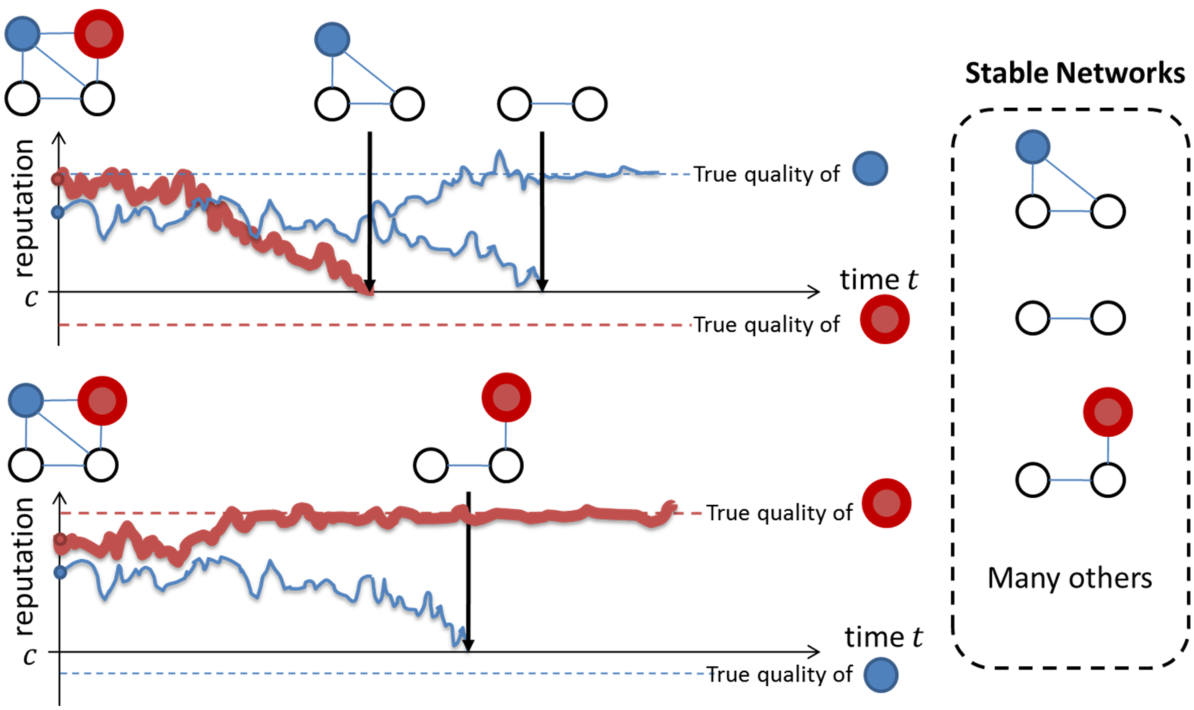

Figure 1.1: Illustration of Possible Network Dynamics

To understand which stable networks G^∞ can emerge, we investigate whether a link l_{ij} between agents i, j can exist at $t = \infty$. If two agents i and j are not neighbors (i.e. $\omega_{ij} = 0$), then it is certain that $g_{ij}^\infty = 0$. If two agents i and j are neighbors (i.e. $\omega_{ij} = 1$), then the existence of this link l_{ij} at $t = \infty$ requires that the reputations of both i and j never hit c for all finite t, which means that neither agent is ever ostracized. Hence G^∞ will always be a subset of the initial network G^0, and is composed only of agents whose reputations never hit c for all finite t.

We say that an agent is *included in the stable network* if their reputation never hits c for all t, so that they are never ostracized from the network. [12]

Note that being included in the stable network does not imply that an agent has any links in the stable network, as it could also be that all of the neighbors of that agent were ostracized even though the agent itself was not. We can calculate the *ex ante* probability that an agent i is included in the stable network, which we denote by $P(S_i)$ with S_i denoting the event in which agent i is included in the stable network. This can be accomplished using standard results regarding Brownian motion hitting probabilities, since $P(S_i)$ is equal to the probability that the agent's reputation never hits c for all finite t. The following proposition gives this probability.

Proposition 1. *$P(S_i)$ depends only on the initial quality distribution and the link cost and can be computed by*

$$P(S_i) = \int_c^\infty (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)))\phi\left((q_i - \mu_i)\frac{1}{\sigma_i}\right) dq_i/\sigma_i \tag{1.2}$$

Proof. See appendix. □

Proposition 1 has several important implications. Note that since $P(S_i)$ is positive and less than 1 for all i, no agent is certain to be included in or excluded from the stable network. Also note that the probability an agent is part of the stable network is independent of that agent's signal precision

[12]As a technical note, when we make the ostracization classification, we assume that an agent who has all its neighbors ostracized will continue to send information about itself at its signal precision level, with the signals sent via the same probability distribution based on its true quality. Thus we still considered the agent "ostracized" if its reputation drops to c via this information process even after all its other neighbors have been ostracized. This assumption is made for technical purposes only and has no impact on the dynamics of the model, as the agent has no links in any such case.

τ_i. Thus the rate at which the agent sends information does not affect the chance that it is in the stable network. This is because the rate at which the agent sends information only affects *when* it gets ostracized from the network, but not *if* it gets ostracized overall[13]. Furthermore, note that the probability an agent i is included in the stable network is independent of its links with other agents and the properties of those agents. Connections with other agents affect the rate at which an agent sends information but not the agent's true quality, and so will not impact whether it is eventually ostracized from the network.

Using the explicit expression above, we can also describe how $P(S_i)$ depends on an agent's initial mean and variance, μ_i and σ_i.

Corollary 1. *For each agent i, $P(S_i)$ is increasing in its initial mean quality μ_i, decreasing in the variance of its initial quality σ_i^2, and decreasing in the link cost c. Moreover,* $\lim_{\mu_i \to \infty} P(S_i) = 1$, $\lim_{\sigma_i \to 0} P(S_i) = 1$, $\lim_{c \to -\infty} P(S_i) = 1$.

Proof. See appendix. □

These properties are intuitive since an agent with a higher mean quality and smaller variance is less likely to have its reputation drop below c, and so is less likely to become ostracized. Moreover, lowering the linking cost also reduces the hitting probability since the agent's reputation would now have to fall lower to be excluded from the network.

As mentioned, G^∞ must be a subset of G^0. Further, it can contain links only amongst pairs of agents that are both included in the stable network and were linked in the initial network. Equivalently, the set of stable networks can be thought of as the set of networks that can be reached from G^0 by sequentially ostracizing agents. Let $I\{S_i\}$ denote the indicator variable of the event in which agent i is included in the stable network. Formally, a network can be stable if and only if it is a matrix with entries given by $g_{ij} = I\{S_i\}I\{S_j\}I\{g_{ij}^0 = 1\}$, for some realization of $\{S_i, \neg S_i\}_{i \in V}$. Links can

[13]To understand this intuitively, recall that reputation evolves through Bayes updating of the Brownian motion. A higher precision increases the amount of information sent at every moment in time, but the overall probability distribution of the information that is sent across all time remains the same. To see this rigorously, note that in the proof of Proposition 1 in the appendix, the survival probability of an agent depends on τ_i only through the term $t\tau_i$. Thus increasing τ_i and decreasing the time t proportionally leaves the survival probability unchanged overall.

exist only among agents that were never ostracized and were linked in the original network. Note that different realizations of $\{S_i, \neg S_i\}_{i\in V}$ could potentially correspond to the same stable network[14].

By Proposition 1, we know that the rates of learning do not affect the probability of each event S_i. Since the rate of learning has no effect at an individual level, it cannot have an effect at the aggregate level either. This is formalized in the following theorem. We can also use the equation in Proposition 1 to derive an analytic expression for the probability that any specific stable network emerges, which is presented in the corollary below.

Theorem 1. *The signal precisions of the agents, $\{\tau_i\}_{i\in V}$, do not affect the set of stable networks that can emerge or the probability that any stable network emerges.*

Proof. It is clear that a network G must be a subset of G^0 and can be stable if and only if there exists at least one combination of events $\{S_i, \neg S_i\}_{i\in V}$ such that $g_{ij} = I\{S_i\}I\{S_j\}I\{g^0_{ij} = 1\}$. Thus the set of stable networks does not depend on the learning speed. Moreover, according to Proposition 1, $P(S_i)$ is independent over the different agents and does not depend on the speed of learning. Hence the probability that any specific link exists in the stable network exists also independent of the learning speed, so the probability of any stable network emerging is also independent of the learning speed. □

Corollary 2. *The probability that a network G is a stable network is given by $\sum_{\{S_i\}} \prod_i P(S_i)$ where the summation is over all realizations of $\{S_i, \neg S_i\}_{i\in V}$ that correspond to G.*

Figure 1.2 shows an example of how the corollary can be applied to a simple network of three agents. This figure shows the five possible stable networks that could emerge given an initial network of three agents. In addition, $P(S_i)$ is given for all the agents, which allows us to calculate the exact probability of each of these networks emerging. For the first four networks, there is only one realization of $\{S_i, \neg S_i\}_{i\in V}$ that corresponds to it. For the last network, there are four possible realizations, one in which $\neg S_i$ occurs for all agents, and three in which S_i occurs for a single agent.

[14]For instance suppose that the network comprises only two agents i and j. Then the event in which S_i but not S_j occurs and the event in which both S_i and S_j occur lead to the same stable network structure: the empty network.

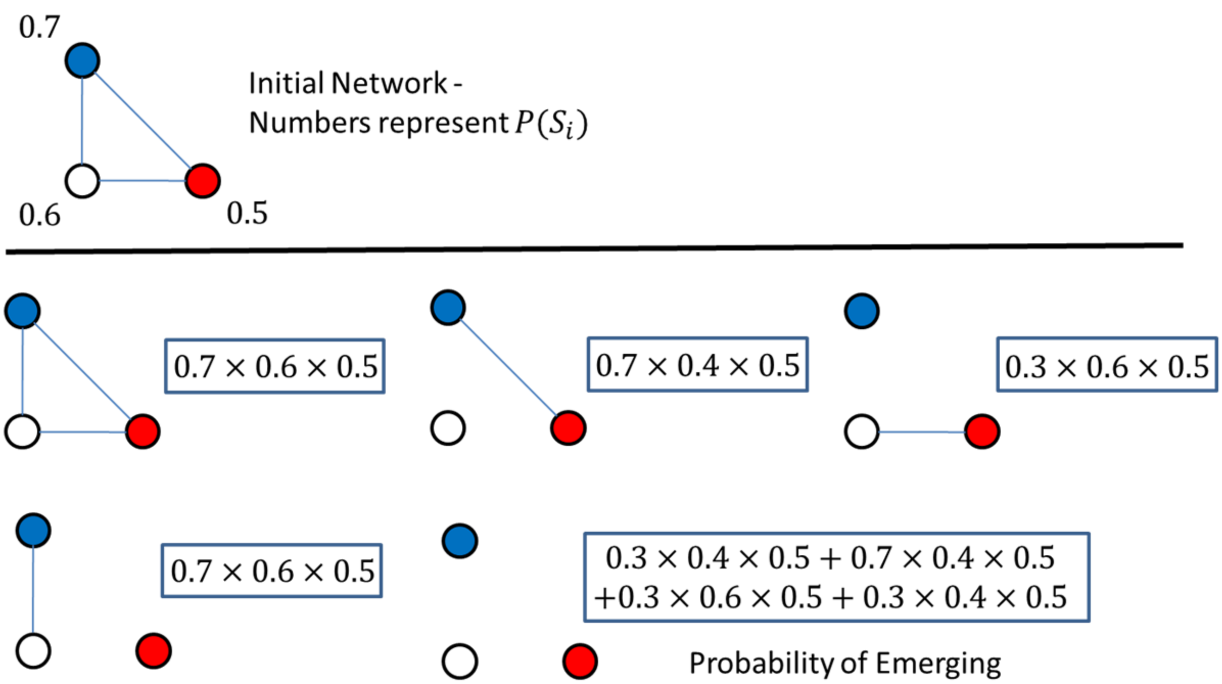

Figure 1.2: Set of Stable Networks Given an Initial Network

We have shown that the speed of learning has no impact on the probability that a network G is stable. This is intuitive since learning only affects the duration of a link but not its final state. However, learning will have a crucial role on the social welfare of a network, which directly depends on how long the agents are connected. We will consider the impact of learning on the social welfare in the next section.

1.5 Welfare Computation

We will analyze overall social welfare from an *ex ante* perspective, given only the network constraint Ω and the prior agent quality distributions. Importantly the *ex ante* welfare is calculated before the agent qualities are learned and any signals are sent. This type of welfare is the most suitable for the type of design settings we will consider later, as it requires the least knowledge on the part of the network designer. Let $P(L^t_{ij}|q, G^0)$ denote the probability that the link between agents i and j still exists at time t. Also, let the parameter ρ represent the discount rate of the network designer[15]. We can define the overall *ex ante* social welfare W formally as follows:

$$W = \int_{q_1=-\infty}^{\infty} ... \int_{q_N=-\infty}^{\infty} \sum_{i,j} \int_0^{\infty} e^{-\rho t}(q_j - c)P(L^t_{ij}|q, G^0)dt\phi(\frac{q_N - \mu_N}{\sigma_N})dq_N/\sigma_N ... \phi(\frac{q_1 - \mu_1}{\sigma_1})dq_1/\sigma_1 \tag{1.3}$$

We will show that this social welfare expression can be calculated in a tractable fashion using a somewhat indirect approach. This approach utilizes the fact that the *ex ante* social welfare is an expectation over all the possible *ex post* signal realizations. Thus we can calculate the *ex ante* welfare by integrating over all possible realizations of the *ex post* welfare, which simplifies equation 3 to a much more tractable form.

[15] We are assuming that the designer itself is more patient than the myopic agents. This can be thought of, for instance, as a company manager who is more patient than its workers who act myopically in their interactions, or a financial regulator that is more patient than the financial institutions, which have managers with myopic incentives.

1.5.1 Ex post welfare

Consider an *ex post* realization of agent hitting times $\varepsilon = \{\varepsilon_i^{t_i}\}_{i\in\mathcal{V}}$, where $\varepsilon_i^{t_i}$ denotes the event in which agent i's reputation hits c at time t_i given all the agent signals (note that $t_i = \infty$ means that agent i's reputation never hits c). In the event in which $t_i < \infty$, since the belief at time t_i is correct, the expected value of agent i's quality conditional on this event $\varepsilon_i^{t_i}$ is $E[q_i|\varepsilon_i^{t_i<\infty}] = c$. In the event with $t_i = \infty$, since the initial belief is accurate in expectation

$$\mu_i = E[q_i] = P(\varepsilon_i^{t_i<\infty})E[q_i|\varepsilon_i^{t_i<\infty}] + P(\varepsilon_i^{t_i=\infty})E[q_i|\varepsilon_i^{t_i=\infty}] \tag{1.4}$$

$$= (1 - P(S_i))c + P(S_i)E[q_i|\varepsilon_i^{t_i=\infty}] \tag{1.5}$$

and we have

$$E[q_i|\varepsilon_i^{t_i=\infty}] = \frac{\mu_i - (1 - P(S_i))c}{P(S_i)} \tag{1.6}$$

where $P(S_i)$ is given by Proposition 1 and is independent of the network and the learning speed.

According to the above discussion, given an *ex post* realization ε, an agent i obtains 0 surplus from its neighbors that have finite hitting times and obtains positive surplus from those whose reputation never hits c (and are thus included in the stable network). The exact benefit agent i receives in the second case depends on its own hitting time t_i, which determines the link breaking time with the other agents. We can calculate the *ex post* surplus that an agent i receives given ε as follows:

$$W_i(\varepsilon) = E_{q|\varepsilon}\left[\sum_{j:g_{ij}^0=1} \int_0^{\min\{t_i,t_j\}} e^{-\rho t}(q_j - c)dt\right] \tag{1.7}$$

$$= \sum_{j:g_{ij}^0=1,t_j=\infty} \int_0^{t_i} e^{-\rho t}\left(\frac{\mu_j - (1 - P(S_j))c}{P(S_j)} - c\right)dt \tag{1.8}$$

$$= \frac{1 - e^{-\rho t_i}}{\rho} \sum_{j:g_{ij}^0=1,t_j=\infty} \frac{\mu_j - c}{P(S_j)} \tag{1.9}$$

Note that this W_i is taken from the perspective of the designer as it incorporates futures payoffs at the discount rate of ρ. This equation shows that in each *ex post* realization of other agent hitting times, agent i benefits if t_i increases and it is ostracized later from the network. Summing over all agents, the social welfare given the *ex post* realization ε is therefore

$$W(\varepsilon) = \sum_i \left(\frac{1 - e^{-\rho t_i}}{\rho} \sum_{j: g^0_{ij}=1, t_j=\infty} \frac{\mu_j - c}{P(S_j)} \right) \tag{1.10}$$

By taking the expectation over the events ε, the *ex ante* social welfare can be found as $W = E_\varepsilon[W(\varepsilon)]$. In order to compute the *ex ante* social welfare, we still need to know the distribution of the t_i, which is coupled in a complicated manner with the initial network and the learning process. For instance, if the neighbor of agent i has a low hitting time and is ostracized quickly, then agent i sends information at a slower rate and its own hitting time would increase. Thus directly computing the social welfare using the above equation is still difficult. In the next subsection, we develop an indirect method to calculate the distribution of t_i.

1.5.2 Hitting time mapping

Recall that an agent's links will scale up the rate at which it sends information compared to the rate it would send information if its precision were constant at the base level of τ_i. Thus each link also scales down the time at which the agent's reputation hits c. So to calculate when the agent is ostracized, we can first find when the agent's reputation would hit c through sending signals at its signal precision level, and then scale this time downwards proportionately based on the network effect[16]. Consider an *ex post* realization of hitting times $\hat{\varepsilon} = \{\hat{\varepsilon}_i^{t_i}\}_{i \in \mathcal{V}}$ in which agent i's reputation would hit c at time t_i if its precision were fixed at τ_i at all times. Note that the events $\hat{\varepsilon}_i^{t_i}$ are independent from each other across different agents, and since the precision is fixed they also do not depend on the network structure. The probability of $\hat{\varepsilon}_i^{t_i}$ can be explicitly computed in the following lemma.

Lemma 1. *The probability density function $f(\hat{\varepsilon}_i^{t_i}), \forall t_i < \infty$ can be computed as*

$$f(\hat{\varepsilon}_i^{t_i}) = \int_{-\infty}^{\infty} \frac{\mu_i - c}{\sigma_i^2 \sqrt{\tau_i}} t^{-3/2} \phi \left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}} \right) \phi(\frac{q_i - \mu_i}{\sigma}) dq_i / \sigma_i \tag{1.11}$$

The probability mass point function $f(\hat{\varepsilon}_i^{t_i=\infty}) = P(S_i)$.

[16]Refer to footnote 12 for a justification of this type of scaling.

Proof. See appendix. □

Using Lemma 1, we can easily obtain the distribution of joint events $f(\hat{\varepsilon}) = \prod_i f(\hat{\varepsilon}_i^{t_i})$ due to the fact that the individual events are independent. This would measure the joint probability of the agents exiting the network at times $\{t_i\}_{i\in\mathcal{V}}$ if the information sending speed of the agents were not being scaled by the number of their links. If there were no network effect, the *ex ante* social welfare could be directly computed using the distribution of hitting times given by Lemma 1. However, due to the network effect, the actual hitting times may vary for each $\hat{\varepsilon}$. Let $M : [0,\infty]^N \to [0,\infty]^N$ be the hitting time mapping function, which maps the hitting times with no network effect to the actual hitting times when there is a network effect. Figure 1.3 presents an algorithm for computing M, which operates by scaling the information speed of each agent at every time t by their current number of neighbors and updating the speed at which an agent sends information when a neighbor is ostracized.

Note that if $t_i = \infty$ in the event $\hat{\varepsilon}_i^{t_i}$ then it is also ∞ in the mapped event $\varepsilon_i^{t_i}$. This means that an agent that never leaves the network with no scaling effect will not leave when the times are scaled either. Then given a realization $\hat{\varepsilon}$, the *ex post* social surplus can be computed as

$$W(\hat{\varepsilon}) = \sum_i \left(\frac{1 - e^{-\rho M_i(t)}}{\rho} \sum_{j:g_{ij}^0=1, t_j=\infty} \frac{\mu_j - c}{P(S_j)} \right) \tag{1.12}$$

Therefore, the *ex ante* social welfare is $W = E_{\hat{\varepsilon}}[W(\hat{\varepsilon})]$. We note that this is a tractable equation for the *ex ante* social welfare given any network structure and set of agents. Proposition 1 gives the explicit expression for $P(S_j)$, and Lemma 1 provides the distribution of $\hat{\varepsilon}$. Thus our model allows for easy and tractable computations of the *ex ante* social welfare of any type of network. Theorem 2 below formalizes this result.

Theorem 2. *Given Ω, the initial quality distributions, and the link cost c, the overall ex ante social welfare can be computed as follows*

$$W = E_{\hat{\varepsilon}} \left[\sum_i \left(\frac{1 - e^{-\rho M_i(t)}}{\rho} \sum_{j:g_{ij}^0=1, t_j=\infty} \frac{\mu_j - c}{P(S_j)} \right) \right] \tag{1.13}$$

Hitting Time Mapping Function
Input: $\boldsymbol{t}^B$, base precision $\tau_i, \forall i$ and initial graph G^0.
Output: new hitting time vector $\boldsymbol{t}$.
Initiate: $d_i = t_i^B \tau_i$, $v_i = k_i^0 \tau_i$
Initiate: $\mathcal{N} = \{i : t_i^B < \infty\}$, $t_i^{G^0} = 0, \forall i \in \mathcal{N}$ and $t_i^{G^0} = \infty, \forall i \notin \mathcal{N}$/
while $\mathcal{N} \neq \emptyset$ **do**
Let $i^* = \min_{i \in \mathcal{N}} d_i / v_i$.
Update $t_i^{G^0} := t_i^{G^0} + d_{i^*}/v_{i^*}, \forall i \in \mathcal{N}$.
Update $d_i := d_i - v_i \times d_{i^*}/v_{i^*}$.
Update $\mathcal{N} := \mathcal{N}/i^*$.
Update $k_i = \max\{1, k_i - 1\}$, for all i such that $G_{ii^*}^0 = 1$.
Update $v_i = k_i \tau_i$
end while

Figure 1.3: Algorithm for Hitting Time Mapping Function

where the distribution of $\hat{\varepsilon}$ is computed using Lemma 1 and the hitting time mapping function M is given in the appendix.

1.6 Impact of Information and Learning

In this section we study the impact of learning on *ex ante* welfare, both individual and overall, given an initial network G^0. In particular, we will show how the agents' signal precisions, a representation of the rate of learning, impact individual agent welfare as well as the overall social welfare.

As a benchmark, we consider the social welfare when there is no learning, which we denote by W^*. When there is no learning, no existing link will be severed. The social welfare of an agent i without learning can thus be computed by summing over the mean qualities of all agents it is connected with initially:

$$W_i^* = \sum_{j:g_{ij}^0=1} \int_0^\infty e^{-\rho t}(\mu_j - c)dt = \frac{1}{\rho} \sum_{j:g_{ij}^0=1} (\mu_j - c) \tag{1.14}$$

The *ex ante* overall social welfare without learning is given by the sum over the individual welfares:

$$W^* = \sum_i W_i^* = \frac{1}{\rho} \sum_i \sum_{j:g_{ij}^0=1} (\mu_j - c) \tag{1.15}$$

1.6.1 Overall Impact of Learning

Let $W(\tau_1, ..., \tau_N)$ be the *ex ante* social welfare when agents learn each other's true quality with the signal precisions being $\tau_1, ..., \tau_N$. We also let $W_i(\tau_1, ..., \tau_N)$ represent an agent i's *ex ante* welfare given these signal precisions. The next theorem states that in any network, the addition of learning has a negative impact on every individual's *ex ante* welfare for any value of the signal precisions. This immediately implies that it lowers the overall *ex ante* social welfare as well.

Theorem 3. *$W_i(\tau_1, ..., \tau_N) < W_i^*$ for all i and for all $\tau_1, ...\tau_N$.*

Proof. See appendix. □

There are two main factors that are at work in this result. First, the myopia of the agents causes the learning to be done inefficiently. Second, cutting off a link imposes a negative externality on the agent who is ostracized, since that agent can no longer receive benefits from its neighbors. Taken together, these factors lead to a reduction in overall social welfare. More precisely, when a link l_{ij} is severed due to agent j's reputation hitting c, agent i does not gain welfare compared to the case without learning. This is because the expected value of having a link with i from t_j^* on is 0 and thus having the link or not makes no difference[17]. However, agent j loses welfare compared to the case without learning because agent i's reputation is still above the link cost and thus having the link would benefit j over not having the link.

This result supports the damaging impacts of ostracism found in the social psychology literature, which were mentioned above in the literature review. The social psychology literature usually documents the harmful effects of ostracism from the perspective of the agents that have become ostracized and can no longer benefit from interactions with the other agents. However, our result goes further by stating that the possibility of ostracism will actually lower *every* agent's social welfare from an *ex ante* perspective. By allowing for the ostracism of others, agents open themselves up to ostracism as well, which lowers their own welfare by more than they benefit from ostracizing other agents. Theorem 3 shows that every agent is hurt *ex ante* by ostracism, even those that wouldn't themselves be ostracized in the majority of the *ex post* realizations of the network.

1.6.2 Impact of Individual Information

The previous result showed that learning is harmful on aggregate: under learning both individual and overall network welfare are lower than without learning. However, we show in this subsection that learning need not be harmful at an individual level, as the rate that a single agent sends information changes. We now investigate more closely how the information generation rate of a single

[17] Agent myopia is causing the cut-off value to be too high, and so the agent does not benefit from its learning. This feature of reputational learning is similar to that shown in van der Schaar and Zhang (2014). In Section VIII we discuss a possible solution for this problem by providing agents a subsidy to increase experimentation.

agent (i.e. an agent's signal precision) affects welfare. The faster an agent generates information about its own reputation, the faster the other agents will learn its true quality (if the link is not broken).

First we characterize the impact of an agent's signal precision on that agent's own welfare. The next proposition shows that sending more information about itself will always harm an agent.

Proposition 2. *$W_i(\tau_i, \tau_{-i})$ is strictly decreasing in τ_i.*

Proof. Consider any *ex post* realization $\varepsilon = \{\varepsilon_i^{t_i}\}_{i \in V}$. If $t_i = \infty$, then changing τ_i alone does not change the fact that agent i would stay in the network forever, as so it does not affect the hitting time realization of any other agent either. Thus agent i's welfare $W_i(\varepsilon)$ is not affected. If $t_i < \infty$, then the welfare of agent i depends on (1) the mean quality of all the neighboring agents j whose $t_j = \infty$ and (2) its own hitting time t_i. Since (1) is not affected by changing τ_i, we only need to study how τ_i affects t_i.

Intuitively t_i is decreasing in τ_i since agent i's information sending speed is faster due to a higher precision. We provide a more rigorous proof by contradiction as follows. Suppose agent i's new hitting time increases to $t_i' = t_i + \Delta > t_i$. In this new realization, consider the duration from 0 to t_i. Since $t_i' > t_i$, all other agents' information sending process and speed do not change before t_i. Hence, agent i's instantaneous precision at $t \leq t_i$ changes to $(\tau_i^t)' = \frac{\tau_i'}{\tau_i}\tau_i^t$. Hence, information sending by agent i is faster at any moment in time before t_i. Since, the stopping time t_i' is larger than t_i, the total amount of information sent by agent i given τ_i' is larger than that given τ_i. Because the total information sent should remain the same, this causes a contradiction. Therefore t_i' should be smaller than t_i for a larger τ_i'. □

This result is in accordance with Theorem 3 and shows that an agent sending information about itself will strictly decrease its own welfare. This is because in each realization in which the agent is ostracized from the network, the agent will now be ostracized sooner and hence it will enjoy less benefits from others. Since the agent already starts out with the maximal amount of links it can obtain, it in effect has nothing to gain and everything to lose by allowing its own reputation to vary. We relax this assumption in the extensions section and allow agents to form new links with those they are not connected with initially; under those circumstances an agent will be able to benefit by generating more information about itself.

Though increasing the information sending speed is always harmful for an agent itself, it can actually be helpful to its direct neighbors. The next proposition provides a sufficient condition on the initial network such that this holds.

Proposition 3. *Given an initial network G^0, for any two initially connected agents i and j that are linked through a unique path (i.e. the direct link), increasing one's precision increases the other's welfare.*

Proof. Consider any *ex post* realization $\varepsilon = \{\varepsilon_i^{t_i}\}_{i \in V}$. If $t_i = \infty$, then increasing agent i's signal precision τ_i does not change the realization $\varepsilon_i^{t_i}$. Hence t_j is not affected. If $t_i < \infty$, then according to Theorem 1, the new hitting time t_i' is sooner if agent i's signal precision is larger. This causes the link between agent i and j to be severed (weakly) sooner, leading to a (weakly) later hitting time of agent j because agent j will send information at a slower speed for a longer time. Since changing agent i's signal precision does not change the finiteness of the hitting time of all other agents, agent j's welfare increases due to a longer hitting time for itself. □

Since the information sending speed of agent j slows after agent i is ostracized, agent j's hitting time is larger. Agent j thus prefers its direct neighbor to send more information, so that it can cut off more quickly in case the neighbor is bad. After the link is broken, agent j will also be able to reveal less information about itself, which is beneficial according to Proposition 2. In this way agent j would enjoy more benefits for a longer time from its links with its other neighbors. We can extend this analysis for more distant agents when the two agents are connected through a unique path. This is summarized in the corollary to Proposition 3 below.

Corollary 3. *Given any initial network G^0, for any two agents i and j that have a unique path between them, increasing one's signal precision decreases/increases the other's welfare if they are an odd/even number of hops away from each other.*

The above result shows an odd-even effect of the distance between two agents on the agent's

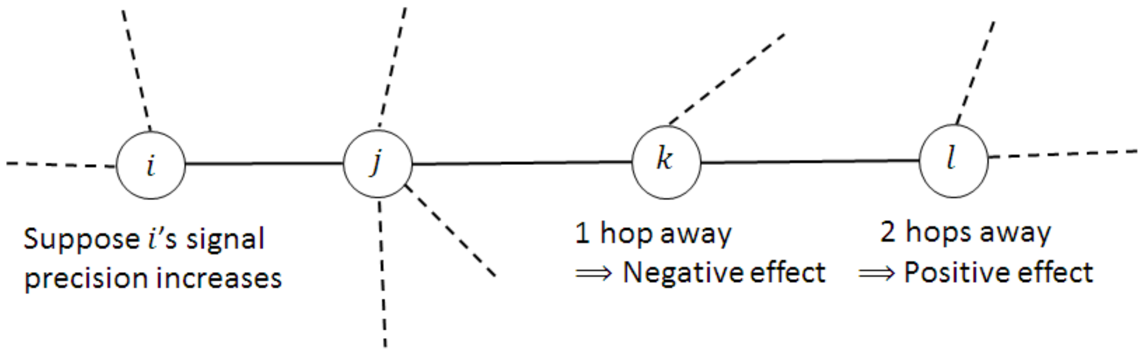

Figure 1.4: Example for Corollary 3

welfare. In all minimally connected networks (such as star, tree, forest networks), any two agents have a unique path between each other and thus the impact of any agent's information sending speed on any other agent's welfare can be completely characterized.

As an example, consider a network where four agents i, j, k, l are connected via a unique path, as depicted in Figure 2. Agent i is linked with agent j, agent j is linked with agent k, and agent k is linked with agent l. Then if agent i sends more information about itself, it stays connected with agent j for a shorter period of time. This causes agent j to send less information about itself, causing agent k to cut off its link with j more slowly if j were to be ostracized. Then agent k is able to link with its other neighbors for a shorter length of time in expectation, decreasing the *ex ante* welfare of k. Thus agent k is hurt when the neighbor of its neighbor, agent i, sends more information. However agent l now links with its own neighbors for a longer length of time, and so it benefits when i sends more information. However, when there are multiple paths between agents, which implies there are cycles in the network, the impact of the signal precision of an agent on the other agents' welfares is much less clear. The reason is that with cycles the neighbor of an agent i's neighbor may also be linked with agent i itself[18], and so the positive and negative effects of information from Corollary 3 are entangled together. The following proposition shows that even for an immediate neighbor, the impact could be totally opposite of Proposition 3 when cycles are

[18] This is known in the social network literature as triadic closure.

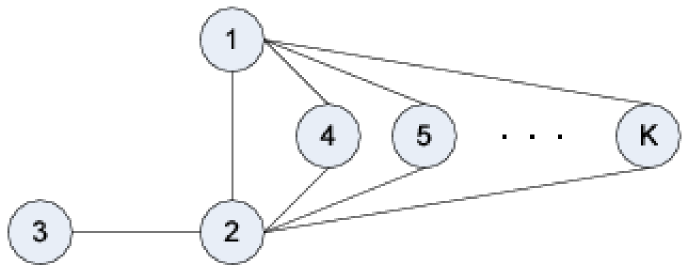

Figure 1.5: Counterexample for Proposition 4

present in the network.

Proposition 4. *If the initial network G^0 has cycles, then it is possible that increasing some agent's signal precision decreases its immediate neighbor's welfare.*

Proof. We prove by constructing a counterexample, which is shown in Figure 2. Consider a network with $K > 3$ agents. Agents 1, 2, 3 form a line and the other $K - 3$ agents connect to both and only agents 1 and 2. We assume that agent 3's quality is perfectly known and large. Hence, agent 3's reputation never hits c. We also assume that the mean qualities of agents 4 to K are close to c. Hence, agent 2 almost does not gain benefit from those agents even when $K \to \infty$.

Consider a realization in which agent 1's reputation hits c at $t_1 < \infty$ and agent 2's reputation hits c at $t_2 < \infty$. By increasing the signal precision of agent 1, its hitting time decreases to $t_1' < t_1$. If $t_1' > t_2$, then agent 2's hitting time is not affected, i.e. $t_2' = t_2$. Otherwise, the new hitting time may be different from t_2. To simplify the analysis, we consider the extreme case in which $\tau_i \to \infty$, thereby $t_1' \to 0$. Therefore, agent 2 loses the link with agent 1 from the beginning in any realization. However, since agents 3 to K also lose the link with agent 1 from the beginning, for those whose hitting time was earlier than t_2, their hitting time would increase by a factor of 2. If there are at least three agents among 4 to K whose hitting was between $[t_2/4, t_2/2]$, agent 2's information sending speed will increase sufficiently much that agent 2's hitting time is smaller. By making K large we can always making the probability of this event be large enough. Thus, agent 2's hitting time will decrease on average. □

We have seen that increasing the information sending speed of an individual agent i could be both good or bad for other agents depending on their locations in the network and their relation with agent i. We note that it could similarly be good or bad for overall social welfare. Thus in contrast with Theorem 3, increasing the amount of information about a single agent can benefit the network overall. This would happen for instance, if there are three agents, i, j, and k who are connected in a line, with links ij and jk. Suppose that the mean of agent k's quality is much higher than those of the other two agents. Then most of the welfare in this network comes through the link between agents j and k. If agent i sends more information, agent j would be able to preserve its link with agent k for a longer period of time, and overall social welfare would increase. This example highlights how critical the network structure is in determining the overall impact of more information by a single agent.

1.7 Optimal Initial Networks

In this section, we study which initial networks G^0 maximize the overall *ex ante* social welfare. Equivalently, we could think of a benevolent network planner that wishes to design the network constraint Ω by choosing which agents are able to form links with which other agents. For instance, in the financial network setting we could think of a regulator that specifies which types of financial institutions are allowed to transact with which other types of institutions in order to maximize overall social welfare[19].

1.7.1 Fully connected networks

One intuition is that a fully connected network would be the optimal initial network since it results in the largest number of links initially, and we have assumed that all agents have an initial reputation higher than the linking cost c. This intuition is accurate in certain cases, such as if the designer

[19] We note that many other types of objection functions are also possible instead of the overall *ex ante* social welfare. For instance the designer may wish to maximize network welfare generated over a certain time interval, or before a set deadline is reached. Or the designer may weigh the welfare of some agents more heavily than that of others. Given the tractability of our model, many of our results can be extended for these alternative settings.

is extremely impatient (i.e. $\rho \to \infty$). Since the designer cares only about the initial time period, and when time is short almost no new information can be learned, it is best to make the decision based on the agents' starting reputations. Surprisingly though, the fully connected network is also optimal on the other extreme, when the designer is completely patient (i.e. $\rho \to 0$). In this case, the designer cares about the social welfare of the stable network that eventually develops, and allowing all agents to be connected initially leads to the largest probability of links in the final stable network. We prove these welfare results in the following proposition.

Further, note that the designer's level of patience is inversely related with the rate of learning, as faster learning means that information is revealed sooner and thus less patience is required. Therefore a similar result holds for the rate of learning: as the rate of learning becomes extremal the fully connected network becomes optimal as well. So for instance, a financial regulator should optimally let all types of financial institutions transact with each other if it is very patient or very impatient, or the information production is extremely fast or slow.

Proposition 5. *1. If the designer is either completely impatient (i.e.* $\rho \to \infty$*) or completely patient (i.e.* $\rho \to 0$*), the optimal initial network is the fully connected network.*
2. Fix the other parameters of the model and suppose the agents' signal precisions are all multiplied by the same constant λ*. If learning becomes very fast (i.e.* $\lambda \to \infty$*) or very slow (i.e.* $\lambda \to 0$*), then the optimal initial network is the fully connected network.*

Proof. See appendix. □

When the designer is either completely patient or impatient, the social welfare depends only on the network G^0 or G^∞, respectively. The exact hitting time does not affect the social welfare. Similarly if the learning is very slow, then the network structure always remains at G^0, and if the learning is very fast then G^∞ is realized very quickly, so in both cases a fully connected network is optimal. The idea is that in both extremes, the exact path of learning is no longer critical and so the negative externalities of information are mitigated.

For intermediate levels of patience or learning however, changes in individual agent hitting times due to linking could have a significant impact on the social welfare. We will show that having all agents fully connected with each other is not always the optimal choice. In the next proposition though we show that the fully connected network is optimal in the case where the agents are homogeneous and have very high initial qualities.

Proposition 6. *Suppose all agents are ex ante identical. Fixing the other parameters, there exists $\bar{\mu}$ such that if $\mu_i > \bar{\mu}\ \forall i$, then the optimal initial network is the fully connected network.*

Proof. We will prove that for $\bar{\mu}$ large enough, the social welfare of any non fully connected network will be increased through the addition of any new link. Therefore the welfare of the fully connected network will be greater than the welfare of any other network. Consider an arbitrary network constraint Ω that is not fully connected. Suppose that a link between agents i and j is added to the network, and consider the welfare of the new network constraint Ω'.

First consider the change in welfare of agent i. In any realization where agent i is ostracized, its welfare through having the extra link with j decreases by no more than $\frac{(N-2)\mu}{\rho}$, the welfare loss when it loses all its links with the other agents immediately. In any realization where agent i is not ostracized, its welfare with the additional link increases by $\frac{\mu}{\rho}$, the discounted value of the new link given the expected quality of agent j. Thus the change in welfare for agent i is bounded below by $P(S_i)\frac{\mu}{\rho} + (1 - P(S_i))\frac{(N-2)\mu}{\rho} = \mu(P(S_i)\frac{(N-1)}{\rho} - \frac{(N-2)}{\rho})$. Similarly, we can show that the change in welfare for agent j is bounded below by $\mu(P(S_j)\frac{(N-1)}{\rho} - \frac{(N-2)}{\rho})$.

Now consider the change in welfare for all the other agents in the network. In any realization where both agent i and agent j are not ostracized, the hitting times of all the agents in the network are unaffected by the new link. In any realization where either agent i or agent j are ostracized, the change in welfare for all the other agents is bounded below by $\frac{(N-2)(N-1)\mu}{\rho}$. Thus the total change in welfare for all other agents in the network is bounded below by $[P(S_i)(1 - P(S_j)) + (P(S_j)(1 - P(S_i)) + (1 - P(S_i))(1 - P(S_j))]\frac{(N-2)(N-1)\mu}{\rho}$.

Combining the above two observations, we note that the change in welfare for the whole network is bounded below by $\mu[P(S_i)\frac{(N-1)}{\rho} - \frac{(N-2)}{\rho} + P(S_j)\frac{(N-1)}{\rho} - \frac{(N-2)}{\rho} + P(S_i)(1 - P(S_j)) + (P(S_j)(1 - P(S_i)) + (1 - P(S_i))(1 - P(S_j))\frac{(N-2)(N-1)}{\rho}]$. When $\bar{\mu}$ is large, $P(S_i)$ converges to 1 by Proposition 1. Thus for $\bar{\mu}$ large enough, the lower bound for the change in welfare of agents i and j converges to

$\frac{2(N-1)\mu}{\rho}$, a positive number.

When $\bar{\mu}$ is large, $P(S_i)$ and $P(S_j)$ converge to 1 by Proposition 1. Thus the lower bound for the change in welfare converges to $\frac{2\mu}{\rho}$, a positive. □

1.7.2 Core-periphery networks

As agents become more heterogeneous in terms of their initial mean quality, different network structures other than the fully connected network can be optimal initial networks. Suppose agents are divided into two separate types, and the initial mean quality of the high type agent is μ_H while the initial mean quality of the low type agent is $\mu_L < \mu_H$. We show that when the expected qualities of the two types are sufficiently different, the optimal initial network has a core-periphery structure[20].

Theorem 4. *Suppose that there are two groups of agents, one with initial quality μ_L and one with initial quality μ_H. Fixing all other parameters, there exists $\bar{\mu}$ such that $\forall \mu_H > \bar{\mu}$, the optimal initial network is a core-periphery network where all high type agents are connected with all other agents and no two low type agents are connected. ($\bar{\mu}$ will depend on the other network parameters.)*

Proof. We first show that all high type agents should connect to all other high type agents. This is based on a similar argument as in the proof of Proposition 6. Since when $\mu_H \to \infty$, all high type agents will stay in the stable network with very high probability, adding a link between any two high type agents will strictly improve their welfare while impacting the welfare of all other agents with very low probability. Hence, there must exist a large enough value for μ_H such that the welfare of high type agents is maximized when all high type agents connect to all other high type agents in the initial network.

Next we show that all low type agents should not connect to each other in any network where each is linked to at least 1 high type agent. When $\mu_H \to \infty$, the welfare obtained by a link with any low type agent j is dominated by that a link with high type agents, i.e. we can suppose that

[20] Although this theorem, assumes there are exactly two types, a similar result holds if instead the agents are composed of two groups and within each group have parameters that are sufficiently close together.

the welfare received by a link with another low type agent is approximately zero in comparison to a link with the high type agents. Having additional links with other low-type agents reduces the hitting time of agent j, $M_j(t)$, in the event that it gets ostracized, thereby reducing agent j's welfare by more than the welfare gain of the additional link. Therefore, low type agents do not connect to each other in the optimal initial network.

Finally we show that all low type agents should connect with every high type agent. Since the probability that the high type agent is ostracized approaches zero, such a link does not affect them relative to the extra welfare that the low type agents receive. Thus we consider only the effect on the welfare of the low type agent to be connected with all high type agents. In a realization where the low type agent is not ostracized, this is optimal for all agents, as the high type agent stays in the network with very high probability when μ_H is large enough. Thus both agents have their welfare increased while not affecting the welfare of all other agents. We show that it is also optimal in realizations where the low type agent is ostracized. Again we will assume that the high type agent is not ostracized, which will hold for μ_H high enough. The low type agent receives a flow payoff of μ_H from every high type agent that it has an active link with. Note that in the hitting time mapping function the hitting time of an ostracized agent i is scaled by $1/K$, where K is the total number of high type neighbors. Thus the decrease in hitting time is exactly balanced out by the increase in flow payoff in the case without discounting, and with discounting it is strictly better for the low type agent to have an extra link.

□

The above result shows that high quality agents should be placed in the core and connected with all other agents, while low quality agents should be placed in the periphery and not connected with other low quality agents. Thus agents with lower initial reputations should be placed in less central positions within the network in order to mitigate the negative effects of ostracism. Allowing low quality agents to connect with too many other agents would increase the rate at which they send information, causing them to be ostracized sooner and hurting them more than they would gain through the direct benefits of the extra links. This core-periphery structure is commonly seen in many real-world financial networks, with large well capitalized banks in the core and smaller banks in the periphery. A reason for this could be that the greater reputation of large banks lets them withstand negative shocks more easily without being ostracized by their counterparties. Smaller banks produce less information through their lesser number of transactions, allowing them to avoid being ostracized as quickly. [21]

We note that the above result depends heavily on the type of learning environment that is present. From Proposition 6, we know that if the designer wad either very patient or impatient, or if learning was very slow or very fast, then the optimal network would be the fully connected network. It is only when the learning is at an intermediate level that the above result hold.

1.7.3 Star Networks

Star networks are common networks in the real world, where a single central agent is connected with many periphery agents. Examples include a single boss and many subordinates, the head of a political party that coordinates the disparate branches of the party, or a large trader that deals with many small traders. There are several important forces to consider when placing agents within a star network. Such networks depend greatly on the central agent, because that agent is connected with all other agents and thus has the most links. The central agent is therefore the most important agent to consider, and choosing the best agent to be in the center is crucial to the overall welfare of the network.

The initial mean and the signal precision of the central agent are two exogenous parameters that must be carefully considered when choosing the central agent. A high initial mean is beneficial because it increases the expected flow benefits that all the other agents who are connected to the central agent will receive. However, a higher signal precision is harmful because it allows for a greater probability that the central agent becomes ostracized quickly, thus causing the network to fall apart. Such an event would greatly lower social welfare. Therefore there is a trade off between

[21] We note that financial regulators have started imposing core-periphery structures on various financial networks to encourage stability. Many banks are now required to trade through a central clearing counterparty (CCP), which is a large financial institution that is ideally very stable. The idea is that trading with the CCP will mitigate the uncertainties that individual banks have about each other's qualities and thus prevent liquidity runs during financial crisis.

the initial mean and the signal precision of the central agent: it is desirable to have a central agent with a higher mean but a lower signal precision. In particular, choosing the agent based only on its initial mean expected quality is not optimal, whereas under complete information it would be optimal to always place the highest quality agent in the center.

We show these results formally in the next proposition. For concreteness, suppose that the central agent in the network is denoted by agent 1. The exogenous parameters of the agents are defined the same way as previously.

Proposition 7. *The overall social welfare is strictly increasing in μ_1 and strictly decreasing in τ_1 and σ_1^2.*

Proof. We can break social welfare into two components: the welfare of the central agent, and the welfares of each periphery agent. Notice that the welfares of the periphery agents are strictly increasing in μ_1 but do not depend σ_1 or τ_1 for the same reasons as in the proof of Theorem 3. Also, the welfare of the central agent is strictly increasing in μ_1 as that allows the central agent to stay in the network for a longer period of time. Thus overall social welfare is increasing in μ_1. The welfare of the central agent is strictly decreasing in τ_1 for the same reasons as in Proposition 2. Thus overall social welfare is decreasing in this parameter. □

For the periphery agents on the other hand, the exogenous parameters have a much less clear relationship with the overall social welfare. We can actually show through examples that social welfare can increase or decrease in each of these factors for periphery agents. The same relationships as for the central agent can hold, and a simple example would be a two person network. However a marginally higher mean or a lower signal precision by a single periphery agent can actually *decrease* overall welfare. For instance, consider a network where the central agent has an initial mean quality close to c, one periphery agent denoted by agent i also has an initial mean quality close to c, and the qualities of all other periphery agents is very high. In such a case, increasing the mean quality of agent i by a small amount, or decreasing agent i's signal precision would harm overall social welfare. These changes would result in the central agent being connected to agent i for a longer stretch of time, which is undesirable since the other periphery agents are of much higher quality, and so having the central agent send more information is harmful. Thus, in such a network it would be better for agent i to send information more quickly in order for it to exit the network sooner.

We notice again that the tradeoff denoted above matters only at intermediate values of learning, whereas if learning becomes very slow (or the designer becomes very impatient), then this tradeoff goes away. This is summarized in the following proposition.

Proposition 8. *If the rate of learning becomes very slow (i.e. $\lambda \to 0$), then the optimal star network is obtained by placing the agent with the highest initial mean in the center.*

Proof. In the limit of very slow learning, only the initial welfare generated matters, and placing the agent with the highest mean in the center generates the highest welfare. □

Proposition 8 shows that the decision to place an agent at the center depends only on each agent's initial mean in the limit of very slow learning learning (a similar result holds for very high designer impatience). This is similar to Proposition 5, although only in one direction. When the network is constrained to be a star network, the highest initial welfare is obtained by having the highest initial quality agent in the center if the learning is very slow.

1.7.4 Ring networks

In this section we focus on a special type of network: a ring network. Suppose for convenience that agents are homogeneous in terms of initial mean quality and variance. Assume that under the network constraint Ω each agent is limited to at most two neighbors. Hence, for a given number of agents, they would only be able to form one or multiple ring networks of different sizes. This could represent a work environment in which agents work in pairs on projects and can work on up to two projects at a time, or a financial network in which financial institutions seek two partners to trade with.

We study how the size of different rings affects the welfare an agent obtains and hence, we can determine the optimal size of the rings that agents should form together. Let $W(n)$ denote the welfare an agent can obtain if it is in a ring of size n under the network constraint Ω[22]. We show that networks with rings of three agents (thus there is triadic closure among the agents) will maximize both agent welfare and overall social welfare.[23]

Proposition 9. *The optimal size of a ring network is 3 agents.*

Proof. Consider a ring network consisting of three agents i, j, k. We focus on the welfare of agent i and show that it is maximized compared to rings of other sizes. Since agents are identical, this means that total social welfare is maximized as well.

Agent i obtains a positive benefit in two cases: (1) realizations in which both agents j and k's reputation never hit c; (2) realizations in which exactly one of agents j and k's reputation never hits c. The probabilities that these two cases happen are independent of the network structure by Proposition 1. In the first case, having additional agent(s) between agent j and k does not affect agent i's realization and hence, agent i's welfare is not affected. In the second case, having additional agent(s) between agent j and k will change i's realization with positive probability. Consider a realization in which agent k's reputation never hits c and agent j's reputation hits c at t_j. In the ring of size 3, agent j's direct neighbor besides i (i.e. agent k) never hits c. When there are additional agents, it is either the case that agent j's new direct neighbor never hits c or hits c before infinity. If agent j's new direct neighbor hits c before infinity, then agent j's new hitting time may increase and hence agent i's new hitting time may decrease, leading to a lower welfare for agent i. □

The intuition behind this result is similar to the reasoning of Proposition 3, in which having a direct neighbor send more information is beneficial for an agent. With only three agents in each ring, an agent learns about a neighbor that would be excluded from the stable network at a faster rate, since that neighbor remains connected with the other neighbor, when the other neighbor is included in the stable network, until the first neighbor is ostracized. This guarantees a fast rate of learning about the low expected quality neighbor, allowing the agent itself to have more time to

[22]For convenience we assume that the number of agents N is divisible by n

[23]The social networks literature views triadic closure as the result of common preferences or trust, whereas our model derives a reputational reason for such networks.

stay connected with the high expected quality neighbor that is not ostracized. With more than three agents, the neighbor that is excluded from the stable network may have its own neighbor disconnect in advance, slowing the rate of information the ostracized neighbor produces and hurting the agent itself.

We can extend this result to ring networks with more than three agents. Similar to the odd/even effect highlighted in Corollary 3, we can show that rings with an odd number of agents will always have higher expected social welfare than ring networks with an even number of agents. However, as the number of agents grows large the difference in the social welfare of an even and odd number of agents eventually goes to zero.

Corollary 4. *If n is odd, then $W(n) > W(m), \forall m > n$. If n is even, then $W(n) < W(m), \forall m > n$. Moreover, $W(n)$ converges to a limit as n approaches infinity.*

Proof. The proof is similar to that of Proposition 7 except we take into account the odd-even effect discussed in Corollary 3. We still only need to consider the case when exactly one of agents j and k's reputation never hits c. Without loss of generality assume that j is not included in the stable network. With four agents, social welfare is lower than with three because the neighbor of j, call it l, may be ostracized before than j is ostracized, causing j's information speed to slow down. With five agents, social welfare is higher than with four because in the same case, there is a chance that agent l's other neighbor is ostracized before agent l is ostracized, resulting in a decrease in agent l's information speed and an increase in agent j's information speed. This argument can be extended indefinitely for any number of agents to prove the above result. We note that the limits of the social welfares are the same, since the probability of a neighbor very far away sending a signal that affects agent j's hitting time approaches zero as the number of agents becomes very large. Such an event can only occur if all the agents in between have an ostracism time less than agent j itself, an event with probability that approaches zero as the number of agents gets large. □

1.8 Extensions

As seen above, learning can have a negative impact on social welfare in a variety of networks, and a large reason for this is the myopia of the agents. Since the agents are not experimenting for long enough, learning is inefficient and social welfare is lost. In this section, we consider four possible extensions that could alleviate this issue and allow for higher social welfare.

1.8.1 Linking Subsidy

A potential method of addressing the negative effects of learning is to give subsidies to the agents for linking with others. For instance, a company may wish to give workers awards or bonuses for collaborating with colleagues. Or in a financial setting, a regulator may give financial incentives for firms conducting mutual investments, or guarantee interbank transactions during a financial crisis to lower default risk. We model a subsidy by assuming that for every link that an agent maintains, it receives an extra flow benefit of δ from the network designer. This linking subsidy does not affect the social welfare computation since it is a direct transfer from the network designer to the agent, but it would change agents' decisions of when to break a link. Since agents are myopic, an agent i will break its link with agent j if and only if agent j's reputation drops below $c - \delta$. The linking subsidy thus causes the agents to learn more information about their neighbor's quality and break only if it is really likely to be bad. We show below that by properly choosing the linking subsidy the social welfare can improve compared with the case when there is no learning about agents' qualities. Let $W(\delta)$ denote the *ex ante* social welfare when the linking subsidy is equal to δ.

Theorem 5. *There exists $\bar{\delta}$ such that $\forall \delta > \bar{\delta}$, $W(\delta) > W^*$. Moreover, $\lim_{\delta \to \infty} W(\delta) = W^*$.*

Proof. See appendix. □

Note that by Theorem 3, this result also shows that the social welfare is higher than the standard network model with no subsidy. Thus by imparting subsidies on agents to encourage them to experiment for longer, the social welfare is higher than previously. The intuition is that when the link subsidy is high enough, any link that is broken will involve an agent that was of really bad quality. Thus although the agent that is ostracized may still hurt from being disconnected, its neighbors will benefit by a sufficiently large amount that overall social welfare increases. Thus learning is now beneficial and improves welfare overall. The second part of the theorem states that if the linking subsidy becomes too high, then the social welfare will converge to the social welfare without learning. This is because when the subsidy is too high it becomes almost impossible for a link to break, and so the network with high probability will not change, just like in the case without learning. Thus having a linking subsidy is beneficial for the network, but the subsidy cannot be set too high either in order to maximize social welfare.

1.8.2 New Link Formation

Another way that learning would be more socially beneficial is if agents were able to form new links with other agents whose reputations are very high. In this extension, we assume that a pair of agents who are not initially linked according to the network constraint Ω can form a new link by incurring an instantaneous cost $\gamma > 0$. There is no cost to forming links with agents that they are connected to under Ω. So unlike previously when there was a hard barrier between agents not connected according to Ω, agents can now break this barrier by paying an instantaneous cost. This cost could be exogenous, for instance the cost of time and energy in becoming familiar with a new agent, or the cost of reducing some physical barrier between the agents (distance or geographic barriers). The cost could also be set by the network designer such as a tax on link creation. Since we assume the formation cost is instantaneous, it is infinitesimal in the social welfare calculation and so only affects welfare through its impact on agent actions.

We assume that forming a link this way requires bilateral consent as usual. Agent i will want to form a link with agent j if agent j's reputation is higher than $c + \gamma$. Therefore a new link between agents i and j is formed at time t if and only if $\mu_i^t \geq c + \gamma$ and $\mu_j^t \geq c + \gamma$. The dynamics of our model will now feature some agents attaining high reputation levels and being able to link with other previously inaccessible agents that have also attained high reputation levels. Allowing these two high quality agents to link together will improve social welfare due to the large mutual benefits that are generated from their link.

We can compare the social welfare produced by allowing this extra link formation against the social welfare in the basic model. Let $W(\gamma)$ denote the *ex ante* social welfare when the link formation cost is equal to γ, and let W be the social welfare in the basic model without the extra link formation.

Theorem 6. *There exists $\bar{\gamma}$ such that $\forall \gamma \geq \bar{\gamma}$, $W(\gamma) > W$.*

Proof. Consider any realization ε when link formation is not allowed. The *ex post* welfare $W(\varepsilon)$ is changed only when there is some time t^* such that there exist two agents i and j, who are not initially connected, such that $\mu_i^{t^*} \geq c + \gamma$ and $\mu_j^{t^*} \geq c + \gamma$. In the original realization ε, conditional on t^*, there are two cases

- ζ_1: Both agents' reputations never hit c after t^*.
- ζ_2: At least one agent's reputation hits c after t^*.

When ζ_2 occurs, allowing link formation may change the hitting time of all agents' in the network and hence, the welfare $W(\varepsilon|\zeta_2)$ may change. However, the probability of ζ_2 occurring tends to zero as $\bar{\gamma}$ tends to infinity by Proposition 1. When ζ_1 occurs, the social welfare increases by at least $\frac{e^{-\rho t^*}}{\rho}\frac{(c+\bar{\gamma})-(1-P(\zeta_1))c}{P(\zeta_1)}$. When ζ_2 occurs, the welfare decreases by at most $B(\zeta_2)$, a function that is at most linear in $\bar{\gamma}$ as it grows large, since the set of agents and their initial qualities are fixed. Thus the overall change in welfare can be written as

$$W'(\varepsilon) - W(\varepsilon) \geq P(\zeta_1)\frac{e^{-\rho t^*}}{\rho}\frac{(c+\bar{\gamma}) - (1-P(\zeta_1))c}{P(\zeta_1)} - P(\zeta_2)e^{-\rho t^*}B(\zeta_2) \tag{1.16}$$

By choosing $\bar{\gamma}$ large enough, we can ensure that $P(\zeta_2)$ is small enough such that the change is positive in all such realizations ε. Therefore $W(\gamma) > W$.

□

This theorem states that if the link formation cost is high enough then the social welfare is improved over the base model because two agents that decide to form a new link will do so with high reputations. Thus the social welfare generated by a new link is likely to be high as well, and this dominates any potential informational externalities that the link could create. Note however that a γ that is too low may actually harm welfare, for instance if there are a group of moderate quality agents that are all linked to a very high quality agent, but separated from each other according to Ω. This is similar to the setting of Theorem 5. In such a case, allowing moderate reputation agents to link with each other would cause them to harm each other via the negative informational effects of the link. This would reduce welfare overall compared to the base model. Therefore allowing for new link formation can improve welfare, but the threshold for the link being formed must be sufficiently high as well. The optimal $\bar{\gamma}$ would depend on the specific properties of the network. If as in the example there exists a group of very high initial quality agents that the moderate quality agents are linked with, then $\bar{\gamma}$ would likely be higher as well, as it becomes more important for moderate quality agents to not be linked with each other.

1.8.3 Agent Entry

Our model can also be tractably extended to allow agents to enter into the network over time. Specifically, suppose that for the set of N agents in V there is a corresponding set of entry times $\{e_i\}_{i \in V}$, with $e_i \geq 0 \; \forall i$. Agents with $e_i = 0$ are present in the network at the beginning, while agents who have $e_i > 0$ enter later on. These entry times are fixed and known to the agents in the model. There is still the network constraint Ω over the set of all N potential agents that specifies which agents are allowed to connect to each other, including agents that arrive later. This network constraint determines where agents enter into the network at their entry times. The learning process is the same as before, with learning occurring for agents within the network based on their current amount of neighbors, and no learning occurring for an agent that has not yet entered.

Agents still make decisions myopically and will connect with a neighbor for as long as that neighbor's reputation is above the connection cost. Since we assume all agents have initial reputations above the cost, an incumbent agent with always wish to connect with a newly entering agent. However, the new agent would not want to connect with one of its neighbors that has already been ostracized previously within the network. The dynamics will evolve similarly to before, with agents connecting to neighbors until a neighbor's reputation falls too low, at which point the neighbor will be ostracized. The difference now is that new agents will arrive at certain times, and when they do they will change the benefits and amount of information produced by the network.

We can compare the model with agent entry against the base model where all agents were present in the beginning, i.e. $e_i = 0 \ \forall i \in V$. We fix a network topology Ω and perform comparative statics on the entry times of the agents. We first show that incorporating agent entry will not change either the set or the distribution of stable networks.

Proposition 10. *The set of stable networks is unchanged with agent entry. The probability of each stable network emerging is the same as that given in Corollary 2 and identical to the case without agent entry.*

Proof. First note that Proposition 1 still holds for each agent, regardless of the specific entry times. This is because the later entry of an agent only shifts the time at which it gets ostracized, but will not change the fact that it ever gets ostracized. Since the probability that each agent is ostracized is not affected, the set of stable networks and the probability that each stable network emerges does not change either. Thus the same probability distribution over stable networks as in corollary 1 will result. □

Although the properties of the final stable networks are not affected by agent entry, the overall social welfare will be affected. It is possible to calculate social welfare in a similar method as in Theorem 2, as we can account for agent entry by rescaling the hitting times of the agents in the network appropriately. Incorporating agent entry has two separate effects on social welfare: first, the links that the entering agent has are started later, so the benefits from those links are realized later as well and thus discounted more heavily; second, the neighbors of the entering agent send less information before that agent enters, and the agent itself may send information more slowly if one of its neighbors is ostracized before it enters, delaying the time at which the agent and its neighbors are potentially ostracized from the network. The first effect hurts social welfare because the benefits from any link are positive in expectation. However, the second effect can improve social welfare by delaying the agents' ostracization times and increasing the benefits that each

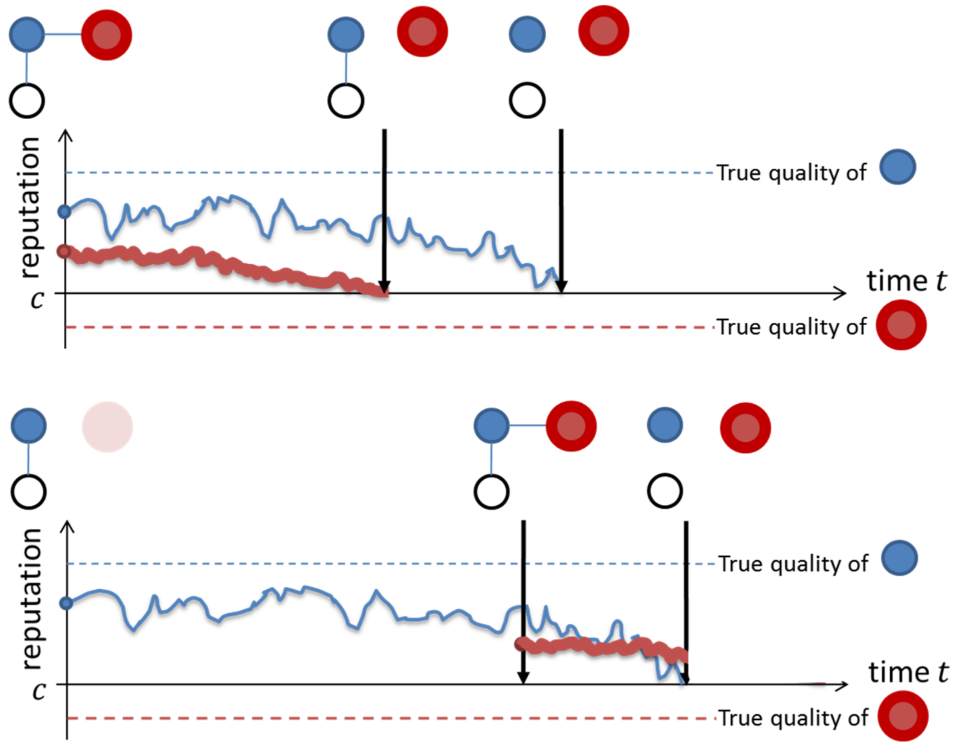

Figure 1.6: Example for Theorem 7

agent is able to extract from the network. It is possible for the second effect to dominate the first, so that delaying entry for an agent raises social welfare overall.

Theorem 7. *For some network parameters, increasing a single agent's entry time e_i can increase social welfare.*

Proof. We prove using an example, shown in Figure 3. In this network of three agents, suppose that the white agent's quality is very high and known, so there is no uncertainty about this agent. Suppose both the red (large circle and bolded line) and the blue (small circle and thin line) agents qualities are unknown and close to c. Since the white agent's quality is very high, the social welfare of the network will be completely determined by the amount of time the blue agent connects with the white agent. By delaying the entry of the red agent, the blue agent is able to stay connected for longer in each realization, and so social welfare increases. □

In this example, note that although delaying the entry of the red agent is helpful, it is still better to have the red agent enter at some finite time instead of never entering. This is because the blue agent's reputation will eventually converge to its true quality by the law of large numbers, and in the case where the blue agent has a good quality, enabling a link with the red agent will produce positive benefits. In addition, after waiting for a sufficiently large amount of time, the probability that the blue agent ever becomes ostracized if it hasn't already goes to zero, so the red agent is unlikely to impact the blue agent's connection with the white agent. Thus delaying the entry of the red agent is beneficial, but the red agent should not be excluded from the network altogether.

This argument holds in general for any network as well, all agents should have finite entry times to ensure optimal social welfare. As an implication, a financial regulator may wish to delay new firms from entering the network in times of crisis when there is a lot of uncertainty, and only allow them to enter once the crisis has ended and reputations are more stable.

1.8.4 Agent Re-entry

Our model can be tractably extended to allow agents to be forgiven and then let back into the network. For instance, suppose that a worker in a company can improve its quality after it becomes ostracized through some exogenous means, such as going back to school to increase its abilities, or taking counseling to better its personality. In financial networks, suppose that a bank can get recapitalized by the government after it gets shut out of the network, allowing its expected quality to increase. After the agent undergoes this exogenous process, the agent's reputation improves and so the other agents are again willing to link with it. We show that agent forgiveness in this manner can increase social welfare as well as mitigate the negative effects of learning. In fact, learning may now actually become beneficial.

We model agent forgiveness by assuming that when an agent is ostracized from the network, the agent can reenter the network at a later time. How long the agent must wait before reentry is an exogenous parameter, which we denote by L. When the agent reenters, its reputation is the same as the reputation that it started out with initially, $N(\mu_i, \sigma_i^2)$.[24] As mentioned above, this re-entry

[24] We make this assumption to avoid adding too many new exogenous parameters. Our results can be extended to a more general setting as well where the reputation changes upon re-entry.

could be the result of the agent undergoing additional training or preparation to improve its original quality. An alternative interpretation is also possible where this is in fact a new agent entering the network, but from the same population or background as the original agent. Thus the new agent starts out with the same reputation as the other agent.

We assume that each agent can reenter into the network as long as it has not already been ostracized in the past a total of R times. Therefore an agent can reenter the network as long as it has not already reentered $R - 1$ times in the past. R is an exogenous parameter that represents the degree to which ostracized agents are willing to undergo the process to improve themselves. A higher value of R means that the ostracized agents are willing to undergo the improvement process even if they have been ostracized multiple times in the past.

With agent re-entry, we can still compute the set of stable networks, as well as the probability that each stable network emerges. The probability that an agent is included in the stable network is now equal to the probability that an agent does not get ostracized a total of R times in a row. Since the agent's reputation is redrawn each time upon re-entry, this probability can be computed using the products of the probabilities in Proposition 1. The exact formula is given in the following proposition. Compared with the original probabilities, agent re-entry implies that each agent is more likely to be part of the stable network, since they have more chances with which to get a high quality draw.

Proposition 11. *$P(S_i)$ depends only on the initial quality distribution and the link cost and can be computed by*

$$P(S_i) = 1 - \left(1 - \int_c^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)))\phi\left((q_i - \mu_i)\frac{1}{\sigma_i}\right) dq_i\right)^R \tag{1.17}$$

‘

Proof. The probability that an agent is ostracized permanently is found by taking the 1 minus the probability in Proposition 1, and then raising that to the power of R. Thus, the probability that an agent is included in the stable network is found by taking 1 minus this probability. □

Note that since this probability is very similar to the probability given in Proposition 1, all of the relationships between this probability and the exogenous parameters (initial mean, variance, signal precision, and link cost) highlighted in Corollary 1 and Theorem 1 continue to hold. In addition, we can derive an analogue of Corollary 2 using these new probabilities. Thus we can still characterize the explicit probability that any stable network emerges as time goes to infinity and the re-entry process by all the agents has concluded.

We can also derive results about agent welfare when re-entry is possible. Specifically, we can show that if the number of periods of re-entry R is sufficiently large, and the time that an agent takes to reenter L is sufficiently small, then learning becomes beneficial. This is intuitive, because if agents are learned about faster, then bad agents can exit the network sooner to undergo improvement while good agents will stay in and are unaffected. Thus, having agent forgiveness mitigates the negative effects of learning, and makes learning a positive overall.

Theorem 8. *If R is sufficiently large compared to τ_i, and L is sufficiently small compared to τ_i, then a small increase in τ_i increases the overall social welfare of the network.*

Proof. Note that as R converges to infinity, the probability that each agent is included in the stable network goes to 1. Therefore the social welfare generated by any agent i will depend on the first time instance at which it enters and does not become ostracized. This is because L is very small, so agent i loses very little benefit when it is ostracized. The first time at which agent i reenters and does not get ostracized is strictly decreasing in its information precision τ_i, since a faster information speed implies that it gets ostracized earlier later on. Thus a larger information precision increases overall social welfare. □

We can extend the above result to show that a clique is the optimal network when the network is very forgiving and the downtime of reentry is low. A clique would allow all agents to link with each other and benefit from the resulting mutual interactions. In addition, since learning is now beneficial, the fact that each agent has many links in a clique and thus sends a lot of information also increases social welfare. This result highlights the fact that with agent forgiveness, more densely connected networks can become optimal, and the designer can allow for more links in the initial networks.

Theorem 9. *If R is sufficiently large compared to τ_i for all i, and L is sufficiently small compared to τ_i for all i, then a clique is the optimal initial network.*

Proof. Similar to the above proof, note that as R converges to infinity, the probability that each agent is included in the stable network goes to 1, and so the social welfare generated by any agent i will depend on the first time instance at which it enters and does not become ostracized. With a clique, each agent has as many neighbors as possible and sends information very quickly, and so the timing of this first time instance becomes sooner. Notice also that each agent's flow payoff is positive at any point in time that they are in the network. Since L is very small, agents are in the network almost continuously, and so having more links increases the flow benefits that each agent receives. Thus a clique is the optimal network. □

1.9 Conclusion

This paper analyzed agent learning and the resulting network dynamics when there is incomplete information. We presented a highly tractable model that explicitly characterized what the set of stable networks are for a given network, showed how learning affects both individual and social welfare depending on the specific network topology, and analyzed what optimal network structures look like for different groups of agents. Our results shed new light on network dynamics in real world situations, and they offer guidelines for optimal network design when there is initial uncertainty about the agents. When agents are sufficiently myopic in their actions, ostracism becomes harmful not just for the ostracized agents themselves, but to all agents in an *ex ante* fashion. A network designer should thus structure links appropriately in order to minimize the negative effects of ostracism.

Our results could be extended in several interesting ways. One natural extension would be to allow the qualities of agents to evolve over time. In the simplest extension, the agent's quality q_i itself change according to an exogenous stochastic process, for instance a Brownian motion. More interestingly, it would be natural to assume that the evolution of quality depends endogenously on the information the agent receives so that agents who receive better information tend to have higher quality and hence also generate better information in the future. Thus, the structure of the network and the quality of the agents in the network co-evolves. Higher quality agents may link to agents that are also of higher quality, and so their qualities would improve quickly, while lower quality agents may struggle to find good agents to link with, and their qualities would decline as a result.

Other possible extensions include having private information among the agents instead of locally public information. In this way agents would learn about their neighbors at different rates, and so they may make different decisions when connecting or disconnecting with other agents. This result could mitigate the negative effects of learning, as information is different across link, and so having more links does not increase the rate of learning. Agent preferences could also be heterogeneous, which would further increase the diversity of links and the range of linking decisions. This is a topic we are currently researching in van der Schaar and Zhang (2015).

Finally, it would be interesting to allow agents to engage in games with their linked neighbors instead of merely generating flow benefits. Games played over networks have been analyzed in several papers within the networks literature (see Jackson and Zenou (2014) for a review), but never in a dynamic setting with learning such as that considered in the current paper. The game played by agents could be a prisoner's dilemma or another type of cooperation game where the payoffs depend on the agent's types. Agents would need to seek out other agents that they can achieve high payoffs in the game with, and this process would also require learning over time about a neighbor's type. As agents are able to learn each other's type more accurately, they may achieve greater efficiency in their plays and also sustain cooperation for a longer length of time.

1.10 Appendix

1.10.1 Proof of Proposition 1

Proof. Suppose for now that agent i's reputation always evolves at the constant signal precision τ_i. Then given the true quality q_i for agent i, the probability that agent i's reputation never hits c before t can be found using standard arguments (see for example Wang and Pötzelberger (1997)) and is given by

$$P(S_i^t|q_i) = \Phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.18}$$

$$-\exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c))\Phi\left(\sqrt{t\tau_i}(q_i - c) - \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.19}$$

Therefore, given q_i, the probability that agent i stays in the network is

$$P(S_i|q_i) = \lim_{t\to\infty} P(S_i^t|q_i) \tag{1.20}$$

- If $q_i > c$, as $t \to \infty$, then we have $\Phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i-c)}{\sqrt{t\tau_i}}\right) \to 1$ and $\Phi\left(\sqrt{t\tau_i}(q_i - c) - \frac{\frac{1}{\sigma_i^2}(\mu_i-c)}{\sqrt{t\tau_i}}\right) \to 1$. Thus, $P(S_i|q_i) = 1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c))$, namely agent i stays in the network with positive probability and the probability is increasing in the true quality q_i.

- If $q_i < c$, as $t \to \infty$, then we have $\Phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i-c)}{\sqrt{t\tau_i}}\right) \to 0$ and $\Phi\left(\sqrt{t\tau_i}(q_i - c) - \frac{\frac{1}{\sigma_i^2}(\mu_i-c)}{\sqrt{t\tau_i}}\right) \to 0$, thus $P(S_i|q_i) = 0$, namely agent i's reputation hits c before $t = \infty$ for sure.

- If $q_i = c$, it is clear that $P(S_i^t|q_i) = 0$ as $t \to \infty$.

Taking the expectation over q_i, we have

$$P(S_i) = \int_c^\infty (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)))\phi\left((q_i - \mu_i)\frac{1}{\sigma_i}\right) dq_i/\sigma_i \tag{1.21}$$

From the above expression we can see that $P(S_i)$ only depends on the initial quality distribution (μ_i and σ_i) and the link cost c but does not depend on the Brownian motion precision τ_i. Since breaking links only changes the Brownian motion precision, the probability that an agent's reputation never hits c is independent of the initial network G^0 or the signal precision τ_i. □

1.10.2 Proof of Corollary 1

Proof. We first show that $P(S_i)$ is increasing in μ_i. Let $q_i - \mu_i = x$. Then $P(S_i)$ can be rewritten as

$$P(S_i) = \int_{c-\mu_i}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(\mu_i - c + x)))\phi\left(\frac{x}{\sigma_i}\right) dx/\sigma_i \tag{1.22}$$

Consider a larger mean quality $\mu_i' > \mu_i$, we have

$$P(S_i|\mu_i') = \int_{c-\mu_i'}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i' - c)(\mu_i' - c + x)))\phi\left(\frac{x}{\sigma_i}\right) dx/\sigma_i \tag{1.23}$$

$$> \int_{c-\mu_i}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i' - c)(\mu_i' - c + x)))\phi\left(\frac{x}{\sigma_i}\right) dx/\sigma_i \tag{1.24}$$

$$> \int_{c-\mu_i}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(\mu_i - c + x)))\phi\left(\frac{x}{\sigma_i}\right) dx/\sigma_i = P(S_i|\mu_i) \tag{1.25}$$

Therefore, $P(S_i)$ is increasing in μ_i.

Next we show that $P(S_i)$ is decreasing in σ_i.

$$P(S_i) = \int_{c-\mu_i}^{\infty} \phi(\frac{x}{\sigma_i})dx/\sigma_i - \int_{c-\mu_i}^{\infty} e^{-\frac{2}{\sigma_i^2}(\mu_i - c)(\mu_i - c + x))} \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{\sigma_i}\right)^2} dx \tag{1.26}$$

$$= \int_{c-\mu_i}^{\infty} \phi(\frac{x}{\sigma_i})dx/\sigma_i - \int_{c-\mu_i}^{\infty} \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2\sigma_i^2}(2(\mu_i - c)+x)^2)} dx \tag{1.27}$$

$$= \int_{c-\mu_i}^{\infty} \phi(\frac{x}{\sigma_i})dx/\sigma_i - \int_{\mu_i - c}^{\infty} \phi(\frac{x}{\sigma_i})dx/\sigma_i = \int_{c-\mu_i}^{\mu_i - c} \phi(\frac{x}{\sigma_i})dx/\sigma_i \tag{1.28}$$

Therefore, $P(S_i)$ is decreasing in σ_i.

Finally, we show that $P(S_i)$ is decreasing in c. Consider a smaller $c' < c$, we have

$$P(S_i|c) = \int_{c}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)))\phi\left((q_i - \mu_i)\frac{1}{\sigma_i}\right) dq_i/\sigma_i \tag{1.29}$$

$$< \int_{c}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c')(q_i - c')))\phi\left((q_i - \mu_i)\frac{1}{\sigma_i}\right) dq_i/\sigma_i \tag{1.30}$$

$$< \int_{c'}^{\infty} (1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c')(q_i - c')))\phi\left((q_i - \mu_i)\frac{1}{\sigma_i}\right) dq_i/\sigma_i = P(S_i|c') \tag{1.31}$$

The first inequality is because for $q_i > c$, $1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)) < 1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c))$. The second inequality is because for $c' < q_i < c$, $1 - \exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)) > 0$. □

1.10.3 Proof of Lemma 1

Proof. Since the Brownian motion precision is constant, using the survival probability

$$P(S_i^t|q_i) = \Phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.32}$$

$$-\exp(-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c))\Phi\left(\sqrt{t\tau_i}(q_i - c) - \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.33}$$

we can compute $f(\hat{\varepsilon}_i^t|q_i) = -\frac{dP(S_i^t|q_i)}{dt}$ as

$$f(\hat{\varepsilon}_i^t|q_i) = -\frac{1}{2}\left(\sqrt{\tau_i}(q_i - c)t^{-1/2} - \frac{\mu_i - c}{\sigma_i^2\sqrt{\tau_i}}t^{-3/2}\right)\phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.34}$$

$$+e^{-\frac{2}{\sigma_i^2}(\mu_i - c)(q_i - c)}\frac{1}{2}\left(\sqrt{\tau_i}(q_i - c)t^{-1/2} + \frac{\mu_i - c}{\sigma_i^2\sqrt{\tau_i}}t^{-3/2}\right)\phi\left(\sqrt{t\tau_i}(q_i - c) - \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.35}$$

$$= -\frac{1}{2}\left(\sqrt{\tau_i}(q_i - c)t^{-1/2} - \frac{\mu_i - c}{\sigma_i^2\sqrt{\tau_i}}t^{-3/2}\right)\phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.36}$$

$$+\frac{1}{2}\left(\sqrt{\tau_i}(q_i - c)t^{-1/2} + \frac{\mu_i - c}{\sigma_i^2\sqrt{\tau_i}}t^{-3/2}\right)\phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.37}$$

$$= \frac{\mu_i - c}{\sigma_i^2\sqrt{\tau_i}}t^{-3/2}\phi\left(\sqrt{t\tau_i}(q_i - c) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c)}{\sqrt{t\tau_i}}\right) \tag{1.38}$$

Taking the expectation over q_i, we obtain $f(\hat{\varepsilon}_i^{t_i})$. □

1.10.4 Proof of Theorem 3

Proof. Consider the *ex ante* surplus W_{ij} that agent i obtains from the link with a neighbor j. The *ex ante* welfare for agent i is simply the summation of this surplus over all j that i is linked with. W_{ij} can be computed as

$$W_{ij} = \int_q \int_0^\infty e^{-\rho t}P(L_{ij}^t|q)(q_j - c)dt\phi(q)dq \tag{1.39}$$

where $P(L_{ij}^t|q)$ is the probability that the link between i and j still exists at time t. Let t^* be the time at which the link between i and j is broken. Then the social welfare can be computed as

$$W_{ij} = \int_q \int_0^\infty e^{-\rho t}(q_j - c)dt\phi(q)dq - E_{t^*}[\int_{t^*}^\infty e^{-\rho t}E_{q_j}(q_j - c|t \geq t^*)dt] \quad (1.40)$$

$$= \int_0^\infty e^{-\rho t}(\mu_j - c)dt - E_{t^*}[\int_{t^*}^\infty e^{-\rho t}E_{q_j}(q_j - c|t \geq t^*)dt] \quad (1.41)$$

where the expectation is taken over the realizations in which the hitting time is t^*. The second term can be further decomposed. Let t_i^* denote the case when $t^* = t_i$, namely agent i's reputation hits c before agent j, and t_j^* be the case where $t^* = t_j$, namely agent j's reputation hits c before agent i. Then

$$W_{ij} = \int_0^\infty e^{-\rho t}(\mu_j - c)dt - E_{t_i^*}[\int_{t_i^*}^\infty e^{-\rho t}E_{q_j}(q_j - c|t \geq t_i^*)dt] - E_{t_j^*}[\int_{t_j^*}^\infty e^{-\rho t}E_{q_j}(q_j - c|t \geq t_j^*)dt] \quad (1.42)$$

In the case of t_j^*, for any $t \geq t_j^*$, since the learning has stopped, $E_{q_j}(q_c - c|t \geq t_j^*) = 0$ by the definition of t_j^*. Similarly, $E_{q_j}(q_j - c|t \geq t_i^*) > 0$ because at t_i^* the expected quality of q_j is strictly greater than c since agent j has not been ostracized. Therefore,

$$W_{ij} = W_{ij}^* - E_{t_i^*}[\int_{t_i^*}^\infty e^{-\rho t}E_{q_j}(q_j - c|t \geq t_i^*)dt] < W_{ij}^* \quad (1.43)$$

Summing over all j that i is linked with, we conclude that the agent i's *ex ante* welfare with learning is strictly less than that when there is no learning. Note that this result holds independently of the values of $\tau_1, ..., \tau_N$, as the signal precisions affect only the distribution of agent hitting times, but not the expected quality of the agents conditional on ever being ostracized. □

1.10.5 Proof of Proposition 5

Proof. From Theorem 2, we have that the *ex ante* social welfare is given by:

$$W = E_{\hat{\varepsilon}} \sum_i \left(\frac{1 - e^{-\rho M_i(t)}}{\rho} \sum_{j:g_{ij}^0=1, t_j=\infty} \frac{\mu_j - c}{P(S_j)} \right) \quad (1.44)$$

If the designer is completely impatient, it only cares about the social surplus at time 0 since with probability approaching 1 no links will be broken among the agents. Since all agents' expected qualities are above the linking cost, having all agents connected with each other yields the highest social surplus. Similarly if learning becomes very slow, then the agent's reputations are never updated and the same reasoning applies. In both cases the $e^{-\rho M_i(t)}$ term approaches zero in the above equation regardless of the network structure, and so adding more agents increases welfare.

If the designer is completely patient, only the stable networks matter. Since the stable network does not depend on the speed of learning and the probability that an agent stays in the stable network is independent of others by Proposition 1, having all agents connected with each other leads to the maximum number of links in the stable networks and hence the highest social surplus. Similarly if learning becomes very fast, the stable network will always be reached immediately and the same reasoning applies. In both cases the $e^{-\rho M_i(t)}$ term approaches one in the above equation regardless of the network structure, and so adding more agents increases welfare. □

1.10.6 Proof of Theorem 5

Proof. (1) Consider the welfare on link l_{ij}. As in the proof of Theorem 3, let t_i^* denote the event in which agent i's reputation hits $c - \delta$ at time t_i^* before agent j. The *ex ante* welfare of link l_{ij} can be computed as

$$W_{ij} + W_{ji} = \int_0^\infty e^{-\rho t}(\mu_i^0 + \mu_j^0 - 2c)dt \tag{1.45}$$

$$-E_{t_i^*}[\int_{t^*}^\infty e^{-\rho t} E_{q_i,q_j}(q_i + q_j - 2c|t \geq t_i^*)dt] \tag{1.46}$$

$$-E_{t_j^*}[\int_{t^*}^\infty e^{-\rho t} E_{q_i,q_j}(q_i + q_j - 2c|t \geq t_j^*)dt] \tag{1.47}$$

Note that the first integral in the above equation represents $W_{ij}^* + W_{ji}^*$, the social welfare of the link without learning.

In the case of t_i^*, for any $t \geq t_i^*$, since the learning has stopped, $E_{q_i,q_j}(q_i - c|t \geq t_i^*) = c - \delta - c = -\delta$ by the definition of t^*. Since agent j is not ostracized, we would have $E_{q_i,q_j}(q_j - c|t \geq t_i^*) > -\delta$. Let $h(\delta, t_i^*) = E_{q_i,q_j}(q_i + q_j - 2c|t \geq t_i^*) = E_{q_i,q_j}(q_j|t \geq t_i^*) - 2c - \delta$. This is the net change in flow payoff after the link is severed. We will show that for any t_i^*, $h(\delta, t_i^*) < 0$ if δ is sufficiently large. A symmetric argument then establishes that $h(\delta, t_j^*) < 0$, and the two together imply that the welfare of the link with learning is greater than $W^*_{ij} + W^*_{ji}$. Then, adding up over all links shows that the overall social welfare is higher than that without learning.

To prove that $h(\delta, t^*_i) < 0$ if δ is sufficiently large, we will show that $E_{q_i,q_j}(q_j|t \geq t^*_i)$ is bounded above for any t^*_i as δ tends to infinity. Consider any *ex post* realization of t^*_i, which implies that agent j's reputation does not hit $c - \delta$ before t^*_i. There are two possibilities for agent j's reputation (here we assume that agent j continues sending information at its fixed signal precision if all its other neighbors are ostracized, as in section 4):

- ζ_1: it never hits $c - \delta$ after t^*_i either.
- ζ_2: it hits $c - \delta$ at some time after t^*_i.

Clearly, $E(q_j|\zeta_1) > E(q_j|\zeta_2) = c - \delta$. Hence $E_{q_i,q_j}(q_j|t \geq t^*) < E(q_j|\zeta_1)$. The value of $E(q_j|\zeta_1)$ is given by equation (6) in the text, with c replaced by $c - \delta$. When $\delta \to \infty$, using equation (6) we can show that $\lim_{\delta\to\infty} E(q_j|\zeta_1) = \mu^0_j$ through the application of L'Hopital's rule.

Therefore, $\forall \epsilon > 0$, there exists δ'_{ij} such that $\forall \delta > \delta'_{ij}$, $E(q_j|\zeta_1) - \mu^0_j < \epsilon$. Hence, fix a value of $\epsilon > 0$ and let $\bar{\delta}_{ij} = \max\{\delta'_{ij}, \mu^0_j - 2c + \epsilon\}$, which ensures for all $\delta > \bar{\delta}_{ij}$, $E(q_j|\zeta_1) - 2c - \delta < 0$. This also implies that $h(\delta, t^*_i) < 0$ for all t^*_i and $\delta > \bar{\delta}_{ij}$. By choosing $\bar{\delta} = \max_{i,j} \bar{\delta}_{ij}$, we ensure the overall *ex ante* social welfare is greater than W^*.

(2) Define $H_{ij}(\delta) = E_{t^*_i}[\int_{t^*_i}^{\infty} e^{-\rho t} E_{q_i,q_j}(q_i + q_j - 2c|t \geq t^*_i)dt]$. We will prove $\lim_{\delta\to\infty} H_{ij}(\delta) = 0$. To prove this, we will show that for any sequence $\delta_n \to \infty$, the sequence $H_{ij}(\delta_n) \to 0$. We divide $H_{ij}(\delta)$ into two parts,

$$H_{ij}(\delta) = E_{t^*_i < \hat{t}(\delta)}\left[\int_{t^*_i}^{\infty} E_{q_i,q_j}[e^{-\rho t}(q_i + q_j - 2c|t \geq t^*_i)]dt\right] \tag{1.48}$$

$$+E_{t^*_i \geq \hat{t}(\delta)}\left[\int_{t^*_i}^{\infty} E_{q_i,q_j}[e^{-\rho t}(q_i + q_j - 2c|t \geq t^*_i)]dt\right] \tag{1.49}$$

$$= H'_{ij}(\delta) + H''_{ij}(\delta) \tag{1.50}$$

for some $\hat{t}(\delta)$. We will find a sequence $\hat{t}(\delta_n)$ such that both $H'_{ij}(\delta_n) \to 0$ and $H''_{ij}(\delta_n) \to 0$ as $\delta_n \to \infty$.

Let $\hat{t}(\delta_n) = \delta_n$. First we will show that for δ_n large enough, $P(t^*_i < \delta_n) < \frac{1}{\delta_n^2}$. Note for a given q_i, the probability that the agent is ostracized before time δ_n is equal to:

$$1 - P(S_i^{\delta_n}|q_i) = 1 - \Phi\left(\sqrt{\delta_n \tau_i}(q_i - c + \delta_n) + \frac{\frac{1}{\sigma_i^2}(\mu_i - c + \delta_n)}{\sqrt{\delta_n \tau_i}}\right) \tag{1.51}$$

$$- \exp(-\frac{2}{\sigma_i^2}(\mu_i - c + \delta_n)(q_i - c + \delta_n))\Phi\left(\sqrt{\delta_n \tau_i}(q_i - c + \delta_n) - \frac{\frac{1}{\sigma_i^2}(\mu_i - c + \delta_n)}{\sqrt{\delta_n \tau_i}}\right) \tag{1.52}$$

Note that $\lim_{x\to\infty} \Phi(x) = 1 - \frac{e^{\frac{-x^2}{2}}}{x\sqrt{2\pi}}$. Therefore the term above approaches zero faster than $\frac{1}{\delta_n^2}$ as $\delta_n \to \infty$. Integrating over all q_i shows that $P(t_i^* < \delta_n) < \frac{1}{\delta_n^2}$ for large δ_n.

Now consider $H'(\delta_n)$, it is bounded by

$$|H'(\delta_n)| < P(t_i^* < \delta_n) \sup_{t_i^* < \delta_n} |\int_{t_i^*}^{\infty} E_{q_i,q_j}[e^{-\rho t}(q_i + q_j - 2c|t \geq t_i^*)]dt| \tag{1.53}$$

$$< \frac{\sup_{t_i^* < \delta_n} |\int_{t_i^*}^{\infty} E_{q_i,q_j}[e^{-\rho t}(q_i + q_j - 2c|t \geq t_i^*)]dt|}{\delta_n^2} \tag{1.54}$$

$$< \frac{1}{\rho \delta_n^2} \sup_{t_i^* < \delta_n} |E[q_j|\zeta_1] - c + \delta_n| \tag{1.55}$$

Since as $\delta_n \to \infty, E[q_j|\zeta_1] \to \mu_j^0$, we conclude that $|H'(\delta_n)| \to 0$.

Consider $H''(\delta_n)$, it is bounded by

$$|H''(\delta_n)| < \sup_{t_i^* \geq \delta_n} |\int_{t_i^*}^{\infty} E_{q_i,q_j}[e^{-\rho t}(q_i + q_j - 2c|t \geq t_i^*)]dt| \tag{1.56}$$

$$< \frac{1}{e^{\rho \delta_n}} \sup_{t_i^* > \delta_n} |\int_{t_i^*}^{\infty} E_{q_i,q_j}[e^{-\rho(t-\delta_n)}(q_i + q_j - 2c|t \geq t_i^*)]| \tag{1.57}$$

$$< \frac{1}{\rho e^{\rho \delta_n}} \sup_{t_i^* < \delta_n} |E[q_j|\zeta_1] - c + \delta_n|t \geq t_i^*| \tag{1.58}$$

Similarly, since as $\delta_n \to \infty, E[q_j|t \geq t_i^*] \to \mu_j^0$, we conclude that $|H''(\delta_n)| \to 0$. □

References

[ADL11] Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. "Bayesian learning in social networks." *The Review of Economic Studies*, **78**(4):1201–1236, 2011.

[AO11] Daron Acemoglu and Asuman Ozdaglar. "Opinion dynamics and learning in social networks." *Dynamic Games and Applications*, **1**(1):3–49, 2011.

[BG00] Venkatesh Bala and Sanjeev Goyal. "A noncooperative model of network formation." *Econometrica*, **68**(5):1181–1229, 2000.

[Ero15] Selman Erol. "Network hazard and bailouts." Technical report, working paper, 2015.

[EV14] Selman Erol and Rakesh Vohra. "Network Formation and Systemic Risk, Second Version." 2014.

[GG10] Andrea Galeotti and Sanjeev Goyal. "The law of the few." *The American Economic Review*, pp. 1468–1492, 2010.

[GGJ10] Andrea Galeotti, Sanjeev Goyal, Matthew O Jackson, Fernando Vega-Redondo, and Leeat Yariv. "Network games." *The review of economic studies*, **77**(1):218–244, 2010.

[GGK06] Andrea Galeotti, Sanjeev Goyal, and Jurjen Kamphorst. "Network formation with heterogeneous players." *Games and Economic Behavior*, **54**(2):353–372, 2006.

[GJ10] Benjamin Golub and Matthew O Jackson. "Naive learning in social networks and the wisdom of crowds." *American Economic Journal: Microeconomics*, pp. 112–149, 2010.

[GJ12] Benjamin Golub and Matthew O Jackson. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *The Quarterly Journal of Economics*, **127**(3):1287–1338, 2012.

[GK03] Douglas Gale and Shachar Kariv. "Bayesian learning in social networks." *Games and Economic Behavior*, **45**(2):329–346, 2003.

[GRU09] Jacob K Goeree, Arno Riedl, and Aljaž Ule. "In search of stars: Network formation among heterogeneous agents." *Games and Economic Behavior*, **67**(2):445–466, 2009.

[JA93] Robert Jacobson and David Aaker. "Myopic management behavior with efficient, but imperfect, financial markets: A comparison of information asymmetries in the US and Japan." *Journal of Accounting and Economics*, **16**(4):383–405, 1993.

[Jac10] Matthew O Jackson. *Social and economic networks*. Princeton University Press, 2010.

[JW96] Matthew O Jackson and Asher Wolinsky. "A strategic model of social and economic networks." *Journal of economic theory*, **71**(1):44–74, 1996.

[JZ14] Matthew O Jackson and Yves Zenou. "Games on Networks." *Handbook of Game Theory*, p. 95, 2014.

[KB13] Steve WJ Kozlowski and Bradford S Bell. "Work Groups and Teams in Organizations: Review Update." 2013.

[Miz10] Natalie Mizik. "The theory and practice of myopic management." *Journal of Marketing Research*, **47**(4):594–611, 2010.

[ROW12] Sandra L Robinson, Jane O'Reilly, and Wei Wang. "Invisible at work an integrated model of workplace ostracism." *Journal of Management*, p. 0149206312466141, 2012.

[Sco12] John Scott. *Social network analysis*. Sage, 2012.

[SS15] Yangbo Song and Mihaela van der Schaar. "Dynamic network formation with incomplete information." *Economic Theory*, **59**(2):301–331, 2015.

[SZ14] Mihaela van der Schaar and Simpson Zhang. "A dynamic model of certification and reputation." *Economic Theory*, pp. 1–33, 2014.

[SZ15] Mihaela van der Schaar and Simpson Zhang. "From Acquaintances to Friends: Homophily and Learning in Networks." *arXiv preprint arXiv:1510.08103*, 2015.

[Veg06] Fernando Vega-Redondo. "Building up social capital in a changing world." *Journal of Economic Dynamics and Control*, **30**(11):2305–2338, 2006.

[Wat01] Alison Watts. "A Dynamic Model of Network Formation." *Games and Economic Behavior*, **34**(2):331–341, 2001.

[WP97] Liqun Wang and Klaus Pötzelberger. "Boundary crossing probability for Brownian motion and general boundaries." *Journal of Applied Probability*, pp. 54–65, 1997.

[WS97] Kipling D Williams and Kristin L Sommer. "Social ostracism by coworkers: Does rejection lead to loafing or compensation?" *Personality and Social Psychology Bulletin*, **23**(7):693–706, 1997.

[WS98] Duncan J Watts and Steven H Strogatz. "Collective dynamics of 'small-world' networks." *nature*, **393**(6684):440–442, 1998.

[WWP13] Eric D Wesselmann, James H Wirth, John B Pryor, Glenn D Reeder, and Kipling D Williams. "When do we ostracize?" *Social Psychological and Personality Science*, **4**(1):108–115, 2013.

CHAPTER 2

Optimal Production Choice with Reputational Dynamics

2.1 Introduction

The age old saying "Better to keep your mouth shut and be thought a fool than to open it and remove all doubt" reveals a proverbial truth about the management of one's reputation. An imprudent person who acts openly will risk immediately revealing his ignorance to the audience. It may be better to check one's tongue, imparting less information and merely raising suspicion. Similarly, a substandard firm could keep a low profile, minimizing product diffusion like the snake oil salesmen of old who traveled from town to town cheating uninformed customers. Information is a valuable commodity, and the less of it these quacks reveal to their customers the better.

On the flip side, exceptional firms may wish for as many people to use their product as possible, building up a solid reputation to pump up their customer base. After all, there may be no better form of advertising than a satisfied customer. Web retailers are no stranger to this idea, and programs such as Amazon Vine allow companies give away new products for free to customers simply to elicit reviews. The hope is that more positive reviews will attract more customers, offsetting the cost of these gifts. And research has shown that online product sales are positively correlated with both quality and volume of reviews. Chevalier and Mayzlin use data from Amazon and Barnes and Nobles to show that both the amount of reviews and the quality of ratings will increase sales [CM06]. Dellarocas et al. show that online user reviews of movies are a better predictor of movie success than even critic reviews [DAZ04]. Thus there is clearly an advantage for firms to engage in such loss-leading behavior. Bonatti [Bon11] also finds theoretical support for a firm to engage in introductory pricing to increase its total sales.

We may also wish to know how reputational concerns are taken into account when one does not know his true quality. Consider the case of a new employee who is uncertain about his aptitude. He learns about his true quality through the projects he takes on, but the market will also learn along with him. How active should this person be, taking into consideration both current payoffs and future reputational dividends? By analyzing this problem and the ones mentioned above through the framework of our reputational model, we can explicitly characterize this tradeoff and understand the reputational dynamics at work.

2.2 Literature Review

Reputational models are powerful and informative tools that give insight into how agents behave in the face of continuously repeated interactions. In these types of models, there is usually uncertainty over the agent's true quality, which is resolved over time through information that the market receives. Reputation models can be applied to a broad variety of circumstances. For instance, the agent could represent an academic with an unknown skill, and his true quality is revealed through the quality of the papers that he publishes. Or the agent could be a firm that sells goods of unknown quality, which the market can learn about through customer reviews.

Reputation models can be tailored to fit different situations more closely by varying the underlying assumptions inherent in the model. For example, models meant to represent a short time frame may consider the agent's quality as fixed, while models that span longer time frames may allow for either exogenous or endogenous quality changes. Likewise, the types of signals that are sent can depend on the situation. For an academic writing a scholarly paper, it may be natural to assume perfect good news signals, which means that only high quality academics could publish good papers. In contrast, for a firm producing industrial equipment, perfect bad news could be assumed meaning that only bad firms would produce defective products.

The main contribution of this chapter will be that the signals the market receives are both endogenously dependent on the agent's production choice, as well as noisy so that quality cannot be determined immediately after a signal. This is in contrast to other papers that relax these assumptions. For example, a seminal paper by Holmstrom [Hol99] considers managerial incentive problems in a discrete time setting. The manager's output is an additive function of his true quality, his effort level, and a noise term, and each period his output can be perfectly observed by the market, although it cannot be contracted upon. The manager thus has an incentive to work only for future reputational reasons since he is paid in advance. However, Holmstrom finds that reputational incentives will not be strong enough because the reputational motive can only be temporary. The paper concludes that managers produce less than the efficient output level, with effort declining to zero as the manager's quality becomes more certain. Adding in stochastic shocks to the manager's ability can result in effort being supplied indefinitely, but it will still always be at a suboptimal level.

Another important paper is by Bar-Isaac [Bar03] who considers a discrete time model where a firm produces a single unit every period. The product can either be a success or failure, with high quality firms being more likely to produce successful products. Each period, the market observes the quality of this product and updates its beliefs about the firm's type. The firm's reputation determines the price it receives, and if the firm's reputation falls too low then it may wish to exit the market. The main result is that if the firm's know their own types, high quality firms would never exit the market. This is because merely staying in the market signals strength, so by observing the continued production of the firm, the market will update the firm's reputation to a level that makes it worthwhile for the high quality firm to continue on.

Bar-Isaac's model is closely related to the one we will examine due to its assumption of an infinite horizon. This allows for him to consider Markov perfect equilibrium that depend only on the agent's current reputation. Our model will be an extension of his model, with firms allowed to produce numerous goods each period instead of just one. Instead of considering the exit decisions of high and low quality firms, we will consider production decisions. A big difference is that production cannot be directly observed by the market, whereas exit can. Production is inferred by through rate at which the market receives its signals, and thus the mere event of the signal itself allows the firm to update its belief irrespective of the signal's content. And since there is no exit, instead of a hard lower bound on reputation all reputation values will be possible. But the drift in reputation at different levels will vary based on the firm's equilibrium production choices.

Board and Meyer-ter-Vehn [BV13] consider a variant of the Bar-Isaac model that allows the firms to invest in their quality. This model is set in continuous time, and firms receive random Poisson shocks that change their quality depending on how they have been investing. If they are fully investing in quality, then after the shock they will become a high quality firm. However if they do not invest in quality, they will become a low quality firm. Their result is that if the market receives perfect good news signals, then in equilibrium firms will invest in quality at low reputation levels and not invest at high reputation levels. However, if the market receives perfect bad news signals, then in equilibrium firms do the opposite; they invest at high reputations and do not invest at low reputations. In their paper, firms can indirectly control market signals through quality investment, but this chapter will allow firms to directly control market signals through quantity choice.

One related paper that considers directly controllably endogenous signals is by Chung and Esö [CE07]. They consider a two stage game where the first stage represents the short run and the second stage represents the long run. Both the agent and the market are partially informed about the agent's true quality, but the agent has better information in that he receives an additional signal about his true quality before the game starts. In the first stage the agent chooses between two projects, which both have the same costs but differ in how informative they are about the agent's quality. In the second stage the agent must decide whether to stay in the market or exit. He receives a payoff proportional to his true quality if he stays in the market or a fixed payoff if he exits. Since the agent also has uncertainty about his own type, he has an incentive to choose the more informative action in the first stage to make a better decision in the second stage. However, an agent that is surer about his quality, whether it is good or bad, would gain less from the information in the first stage. Thus the paper finds that in equilibrium, agents that are confident that they are either good or bad choose the less informative action. Agents that are unsure of their quality will choose the more informative action. In this sense, the equilibrium is non-monotonic in the agent's expected quality.

This chapter will act as an extension to the Chung and Esö in that it frames the problem in continuous time with reputational dynamics a la Board and Meyer-ter-Vehn. Analogous to choosing between projects that provide different levels of information, firms will choose among different quantity levels that provide different amounts of information. Also, instead of choosing two projects once, the choice will be done infinitely often in continuous time. This means that there will not be a single second stage long run payoff, but instead there will be an infinite stream of payoffs. And our model assumes that higher quantities entail higher costs, which is unlike the Chung and Es model since they assumed that costs were the same for both projects.

Finally, a recent working paper by Bonatti and Horner [BH12] considers a related reputation model with perfect good news signals. The interpretation is of a new worker completing a single project, and only high quality workers can possibly be successful. Workers can choose among a continuum of effort levels, which linearly affects the arrival rate of the signal. Both the market and the agent are uninformed about the agent's quality, but workers have the advantage of observing their effort levels whereas the market cannot. In contrast to other reputation papers, this model assumes a finite horizon and workers face a penalty if they do not succeed before the time limit. Since effort cannot be observed, wages can only be based on the expected level of effort in equilibrium. The main novelty of this model is the strategic substitutability between effort today and effort tomorrow. Higher wages tomorrow will mean lower effort today, since exerting effort increases the chances of the game ending. Thus it is not possible to incentivize the agent very strongly early on since the agent will therefore exert little effort in the beginning to prolong the game. They conclude that in equilibrium wages must be single peaked, starting out low at the beginning and steadily rising until an apex, at which point it starts decreasing again once the belief of the worker's quality drops too low. Equilibrium effort is thus too low and too late compared with the efficient level.

Like this chapter, the arrival rate is linear in the worker's chosen effort. However, this chapter will allow the signal to be noisy, so a single signal will not resolve the uncertainty over the worker's true quality. Their model assumes that low quality firms cannot succeed, which implies that reputation can only drift downwards without a breakthrough. In our model, noise implies that the worker's reputations could bounce up and down from one level to the next, only converging to their true values over time. In addition, we will assume that the time horizon is infinite, allowing us to consider Markov perfect equilibrium that depend only on the agent's current reputation. And since we will assume that agents are paid directly for their output instead of their expected output, the equilibrium effort and wage in the uninformed agent case will be efficient.

2.3 Model

We consider a monopolist seller of a single good that faces an infinite amount of risk-neutral homogeneous consumers. Our model fits equally well if we assume the agent is a firm selling products or a worker selling labor, but we will stick with the first story for the sake of this chapter. The firm can either be high quality or low quality, where quality is denoted by $\theta \in \{H, L\}$. And the product sold by the firm can turn out to be either good or bad, with a high quality firm having a higher chance of producing a good product. Specifically, high quality firms produce a good product with probability g, and low quality firms produce a good product with probability $b < g$. Firms are drawn from an initial distribution with fraction x_0 being high quality and $1 - x_0$ being low quality. Consequently, x_0 is the market's prior over the firm's true quality, and it is also the firm's time zero reputation. The firm's reputation at time t is denoted by $x_t \in [0, 1]$, which represents the market's belief of the probability that the firm is high quality.

For simplicity, we assume buyers have a value of 1 for a successful good and 0 for an unsuccessful good, and we assume they bid up the good until the market price equals the probability that the good is successful. These two assumptions imply that the price the firm receives at every moment in time is a linear combination of the firm's reputation:

$$p_t = x_t g + (1 - x_t)b \tag{2.1}$$

A firm can sell as much quantity as it wants during time t at price p_t, but it faces an increasing, twice continuously differentiable, and strictly convex cost function. We assume that the product is infinitely divisible for simplicity, so that the firm's production choice can be taken from a continuum of values. The instantaneous profit the firm receives is $\pi_t(q_t) = pq_t - C(q_t)$. We make the assumption that $C(0) < b$, so the inferior firm never has to exit even if the market knows with certainty that it is low quality. Note that convexity of the cost function is not required for the force of our arguments. We need only that $C'(q) \to \infty$ as $q \to \infty$ so that production remains bounded. But this could introduce discontinuities into our value function which would complicate the computations, so we will thus use the assumption of strict convexity.

The market receives signals at random intervals that indicate if the product was successful or unsuccessful. These signals can be thought of as individuals posting reviews online or new consumer reports being published. The higher the quantity that the firm produces, the faster these signals will arrive and the more information the market will get about the firm's true quality. Specifically, these signals follow a Poisson process with arrival rate $\lambda(q_t) = q_t$. We assume that the arrival rate is linear in the production quantity for simplicity. The probability that the signal is good equals the probability that the firm produces a successful good (which is equivalent to the reviews posted being honest). So if the firm is of high quality, the signal will be good with probability g and be bad with probability $1 - g$. Conversely if the firm is of low quality, the signal will be good with probability b and be bad with probability $1 - b$. Importantly, the signal arrival rate does not depend on the actual quality of the firm outside of its production choice. This will be crucial for the uninformed firm, implying that the market will not update its beliefs without a signal arriving.

This chapter will examine two separate cases regarding the firm's *a priori* knowledge. First, we will assume that the firm is unaware of its true quality and has no private information. Thus the firm must learn by observing the same signals that the market receives. Then, we will assume that firm's know with certainty their own quality at the beginning. In this case, firm's will never update their own beliefs but will need to take into account both the market's beliefs about its own production choice and the market's beliefs about the other type of firm's production choice. The dynamics of reputation and firm production choices will be quite distinct across the two models.

2.4 Analysis of Uninformed Firm

In the first case, the firm starts out uninformed about its true quality, having only the same prior as the market. It receives no information about its true quality outside of the signals that the entire market receives. Although the firm can directly observe the history of its production decisions and the market cannot, our assumption that the signal arrival rate does not depend on firm quality implies that this history cannot provide any information beyond the realizations of the signals themselves. So because the firm and the market start off with the same prior, their beliefs will evolve in exactly the same way through observing the signals. Thus the firm's own belief that it is a high quality firm at any time t will be the same as the market reputation x_t.

Consequently, if no signal arrives then the firm's reputation does not change. Suppose that the firm chooses to produce q_t at time t. Since the market and firm share the same beliefs, the probability that the firm is high quality given that no signal arrives in the time period $[t, t + dt]$ is

$$x_{t+dt} = \frac{(1 - q_t dt)x_t}{(1 - q_t dt)x_t + (1 - q_t dt)(1 - x_t)} = x_t \tag{2.2}$$

If a signal does arrive, then the firm's reputation will evolve according to Bayes rule. We will let x_g denote the new reputation after a good signal and x_b denote the new reputation after a bad signal. Applying Bayes rule gives the following set of formulas:

$$x_g = \frac{gx}{gx + b(1 - x)} = x + \frac{(g - b)x(1 - x)}{gx + b(1 - x)} \tag{2.3}$$

$$x_b = \frac{(1 - g)x}{(1 - g)x + (1 - b)(1 - x)} = x - \frac{(g - b)x(1 - x)}{(1 - g)x + (1 - b)(1 - x)} \tag{2.4}$$

These equations imply that reputation change is greatest for middle values of reputation, and that reputation will not change at the extremes $x = 0$ and $x = 1$. When the market is very ambivalent about the firm's quality, a signal will have a strong influence. On the other hand, if the market is very sure about the quality of the firm, a signal will have little effect.

The firm must choose q_t to maximize its value at every time t:

$$V(x_0) = E_{x_t,q_t,\tilde{q}_t}\Big[\int_{t=0}^{\infty} e^(- rt)(q_t p_t - C(q_t))dt\Big] \tag{2.5}$$

We look for Markov perfect equilibrium where the firm's actions and the market's beliefs depend only on the firm's current reputation. Thus the firm's past history of realizations is not relevant except for its impact on the firm's current reputation. A Markov perfect equilibrium will be defined by a set of quantity decisions by the firm $q(x)$ and market beliefs $q\hat{(x)}$ such that market beliefs are correct, and the firm's quantity choice maximizes its value function given these beliefs: $q \in argmax_q V(x_0)$.

We can rewrite the value function above into a continuous time Bellman equation:

$$V(x) = \max_q (pq - C(q))dt + qdt(pV(x_g) + (1-p)V(x_b) - V(x)) + (1 - rdt)V(x) \tag{2.6}$$

The three terms in this expression represent the firm's instantaneous profit, the change in firm value after the arrival of a signal, and the firm's value without a signal respectively. Here r represents the discount rate. Rearranging the above equation gives the following:

$$V(x) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(pV(x_g) + (1-p)V(x_b)) \tag{2.7}$$

This can be seen as a weighted sum of the flow payoff and the jump payoff. The higher the value of r, the more weight is placed on the current flow value. Now, we turn to solving for the firm's optimal production choice. Taking the first order condition of the equation 2.6 gives

$$C'(q) = p + (pV(x_g) + (1-p)V(x_b) - V(x)) \tag{2.8}$$

This says that marginal cost is equal to the price, plus the expected change in reputation after a signal arrives. The optimal q is therefore continuous as long as V is continuous. If the value function is convex, the firm stands to gain after a signal comes and will produce strictly more than the profit maximizing quantity. Taking the FOC of equation 2.7 allows a solution for the optimal quantity choice independent of the current value:

$$(q + r)C'(q) - rp - C(q) = r(pV(x_g) + (1 - p)V(x_b)) \tag{2.9}$$

This equation sets the change in the current payoff from increasing production, given by the first term in equation 2.7, against the change in the future payoff from increasing production, given by the second term in equation 2.7.

We can derive several properties about the value function defined above:

Proposition 1. *1. $V(x)$ is unique and continuous.*

2. $V(x)$ is strictly increasing in x.

3. $V(x)$ is strictly convex in x.

Proof. We prove these properties by applying the contraction mapping theorem. First, we define the operator $T : B[0, 1] \to B[0, 1]$ over the set of bounded continuous real-valued functions on the domain $[0, 1]$. We also restrict the range of these functions, $V : [0, 1] \to [0, \frac{\pi(1)}{r}]$, and $V(0) = 0, V(1) = \frac{\pi(1)}{r}$. This means that the value functions must be bounded below by zero and bounded above by the present discounted value of having reputation one. This follows directly from the fact that reputation can never change at these two extremes. The form of our operator is:

$$T(V(x)) = \max_q \frac{1}{r + q}(pq - C(q)) + \frac{q}{r + q}(pV(x_g) + (1 - p)V(x_b))$$

T takes bounded continuous real-valued functions on the domain $[0, 1]$ to bounded continuous real-valued functions on the domain $[0, 1]$ since both x_g and x_b are in $[0, 1]$ and the cost function is convex. The function will be continuous if V is continuous because $q(x)$ is continuous for continuous V. Note that if the range of V was $[0, \frac{\pi(1)}{r}]$, then the new range will also be $[0, \frac{\pi(1)}{r}]$ since the value can never be lower than zero or be above the value of having reputation one forever, and we will achieve equality at the endpoints $x = 0$ and $x = 1$.

Next, we prove Blackwell's sufficiency conditions for a maximum. First, we show that T satisfies monotonicity, which means that for any two functions $V, W \in B[0,1]$ with $V(x) \geq W(x)$ for all $x \in [0,1]$, we must have $T(V(x)) \geq T(W(x)$. This can be seen by setting the optimal quantity choice under function V equal to the optimal quantity choice under function W, or $q_V(x) = q_W(x)$. Then $T(V(x)) \geq T(W(x)$ because $V(x) \geq W(x)$ for all $x \in [0,1]$. But since $q_V(x)$ may be different from $q_W(x)$ in general, we will have $T(V(x)) \geq T(W(x)$.

Now we prove that T satisfies discounting, or that there exists some $\alpha \in (0,1)$ such that $T(V(x)) + \alpha a \geq T(V(x) + a)$ for all constants a. We have:

$$T(V(x) + a) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(pV(x_g) + (1-p)V(x_b) + a)$$
$$\leq \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(pV(x_g) + (1-p)V(x_b)) + \alpha a = T(V(x)) + \alpha a$$

We can choose $\alpha \in \max\left\{\frac{q(x)}{r+q(x)} \;\middle|\; x \in [0,1]\right\}$ to satisfy the above equation. The convexity of the cost function together with the fact that the value function is bounded ensures that the optimal quantity is bounded above. Thus, discounting is satisfied and our operator T is a contraction mapping. So by the contraction mapping theorem, $V(x)$ is unique and continuous.

Now we prove that the operator will take strictly increasing functions to strictly increasing functions. To prove that the value function is increasing, we consider two reputation levels with $x_1 > x_2$ and compare the value at those points. Let q_1, q_2 be the optimal quantities at x_1 and x_2 respectively. Set $q_1 = q_2$ and note that the actual value at x_1 will be weakly higher than choosing q_2. We have

$$T(V(x_1)) - T(V(x_2)) = \frac{q_2}{r+q_2}(p_1V(x_{1g}) + (1-p_1)V(x_{1b}) - p_2V(x_{2g}) - (1-p_2)V(x_{2b}))$$

Since $pV(x_g) + (1-p)V(x_b)$ is strictly increasing if V is strictly increasing, this implies that $(V(x_1)) \geq T(V(x_2))$ with equality holding only at $q_2 = 0$. But the only time a zero quantity would be produced is at $x = 0$, since we assume that a firm can make a positive profit at any reputation higher than zero. This implies that the value at any reputation higher than zero is also positive. Thus we have $T(V(x_1)) \geq T(V(x_2))$.

Finally, we prove that the operator takes strictly convex functions to strictly convex functions. Suppose that we have two reputation levels x_1 and x_2, and a convex combination of these two points $x_\lambda = \lambda x_1 + (1-\lambda)x_2$ for some $\lambda \in (0,1)$. Then we want to show that $V(x') < \lambda V(x_1) + (1-\lambda)V(x_2)$. Denote the optimal production choice at x_λ by q_λ. Now we calculate the convex combination of the value at x_1 and x_2 with production choice q_λ, and we note that in general the value at x_1 and x_2 must be even higher. Then we have:

$$\begin{aligned}
&\lambda T(V(x_1)) + (1-\lambda)T(V(x_2)) \\
&= \frac{1}{r+q_\lambda}((\lambda p_1 + (1-\lambda)p_2)q' - C(q')) \\
&+ \frac{q_\lambda}{r+q_\lambda}(\lambda p_1 V(x_{1g}) + (1-\lambda)p_2 V(x_{2g}) + \lambda(1-p_1)V(x_{1b}) + (1-\lambda)p_2 V(x_{2b}))
\end{aligned}$$

Note that $\lambda p_1 + (1-\lambda)p_2 = p_\lambda$ by the linearity of p. Also, note that the function $pV(x_g) + (1-p)V(x_b)$ is convex if V is convex. This can be proven by looking at the second derivative (where $S(x)$ is used to denote the above function)

$$\begin{aligned}
S'(x) &= (g-b)V(x_g) + pV'(x_g)\frac{bg}{p^2} - (g-b)V(x_b) \\
&\quad + (1-p)V'(x_b)\frac{(1-b)(1-g)}{(1-p)^2} \\
&= (g-b)(V(x_g) - V(x_b)) + \frac{bg}{p}V'(x_g) + \frac{(1-b)(1-g)}{1-p}V'(x_b)
\end{aligned}$$

$$\begin{aligned}
S''(x) &= (g-b)\left(V'(x_g)\frac{bg}{p^2} - V'(x_b)\frac{(1-b)(1-g)}{(1-p)^2}\right) - \frac{bg}{p^2}(g-b)V'(x_g) + \frac{bg}{p}V''(x_g)\frac{bg}{p^2} \\
&\quad + \frac{(1-b)(1-g)}{(1-p)^2}(g-b)V'(x_b) + \frac{(1-b)(1-g)}{(1-p)}V''(x_b)\frac{(1-b)(1-g)}{(1-p)^2} \\
&= \frac{b^2g^2}{p^3}V''(x_g) + \frac{(1-b)^2(1-g)^2}{(1-p)^3}V''(x_b) > 0
\end{aligned}$$

Putting the above two facts together, we have that

$$\begin{aligned}
\lambda T(V(x_1)) + (1-\lambda)T(V(x_2)) - T(V(x_\lambda)) = \\
\frac{q_\lambda}{r+q_\lambda}\Big[\lambda(p_1 V(x_{1g}) + (1-p_1)V(x_{1b})) + (1-\lambda)(p_2 V(x_{2g}) \\
+(1-p_2)V(x_{2b})) - p_\lambda V(x_{\lambda g}) - (1-p_\lambda)V(x_{\lambda b})\Big] \\
> 0
\end{aligned}$$

□

Thus our operator takes strictly convex functions to strictly convex functions and so the actual value function itself must be strictly convex. Note that this proof is using a revealed preference argument because we fix q_λ, the optimal production choice at x_λ, and assume that the firm chooses this q_λ at the two endpoints x_1 and x_2 as well.

The fact that our value function is strictly convex is very important, because it implies that experimentation is good for the firm. This means that the firm wishes to produce more than is profit maximizing in order to learn more information. Specifically, we can prove the following set of facts about the optimal production choice.

Proposition 2.

1. *The optimal quantity is weakly higher than the myopic profit maximizing quantity.*
2. *The difference between the optimal quantity and the profit maximizing quantity is zero only at the extremes* $x = 0$ *and* $x = 1$.
3. *The optimal quantity is strictly increasing and continuous.*

Proof. The first part can be proven simply by using equation 2.8 above:

$$C'(q) = p + \big(pV(x_g) + (1-p)V(x_b) - V(x)\big)$$

Since the value function is convex, the second term is positive and the convexity of the cost function implies that the optimal quantity will be higher than where marginal cost equals price.

The second part can also be proven by using the above equation. Since the reputation no longer changes at the extremes $x = 0$ and $x = 1$, we have the familiar condition that price equals marginal cost, which leads to the same amount of production as in the no reputation case. At all other reputational values the reputation will update after a signal, so the convexity of the value function implies that the quantity is strictly greater than the profit maximizing quantity.

The third part can be proven by rearranging equation 2.9 above:

$$(q + r)C'(q) - C(q) = r\Big[p + \big(pV(x_g) + (1 - p)V(x_b)\big)\Big]$$

The right hand side of this equation is strictly increasing in x, and the left hand side is strictly increasing in q by the convexity of the cost function. Thus as x increases, q must increase. In addition, the continuity of the value function guarantees that the intersection point will be continuous in x.

□

The fact that the optimal quality is strictly increasing shows that although there are two forces impacting the increase of q above the myopic optimum, a positive force from additional information acquisition and a negative force from myopic profit maximization, the total effect on q balances out so that q is still strictly increasing in reputation, as in the myopic case. Note that this is true for any convex costs, even if they are very weakly convex and almost linear. Although information acquisition has a very strong positive force at intermediate levels of x and diminishes at high x, the negative force of myopic profit optimization also decreases since the myopically optimal quantity rises at higher reputations. This balances out so that the optimal quantity is strictly increasing.

2.5 Analytic Example

By making some simplifying assumptions on the signals that the market receives, we can compute an analytic solution for both the optimal quantity choice and the firm's value function. Specifically, we will assume that the signal the market receives is perfectly informative. This means that the high quality firm always sends the good signal and the low quality always sends the bad signal. Thus upon a high signal reputation jumps up to 1 and a low signal brings reputation down to 0. Quality will be perfectly learned after the signal arrives, so the firm merely needs to decide how fast it wants this signal to come. Since reputation no longer changes, we can simplify our above value function to the following:

$$V(x) = \max_q \frac{1}{r+q}(xq - C(q)) + \frac{q}{r+q}(xV(1) + (1-x)V(0)) \tag{2.10}$$

Now if we assume a cost function, we can explicitly solve for both the optimal quantity choice and the value function directly. Let us assume that cost is quadratic and equal to $C(q) = q^2$. This means that our value function is:

$$V(x) = \max_q \frac{1}{r+q}(xq - q^2) + \frac{q}{r+q}\left(\frac{x}{4r}\right)$$

Solving for the optimal quantity choice using equation 2.9 gives:

$$q^2 + 2rq = r\left[x + \frac{x}{4r}\right]$$

This implies that the optimal quantity choice is

$$q(x) = \frac{-2r + \sqrt{4r^2 + 4rx + x}}{2}$$

Note that at $x = 0$, the optimal quantity is 0, and at $x = 1$, the optimal quantity is $\frac{1}{2}$. Thus the optimal quantity equals the profit maximizing quantity at those two reputation extremes. In

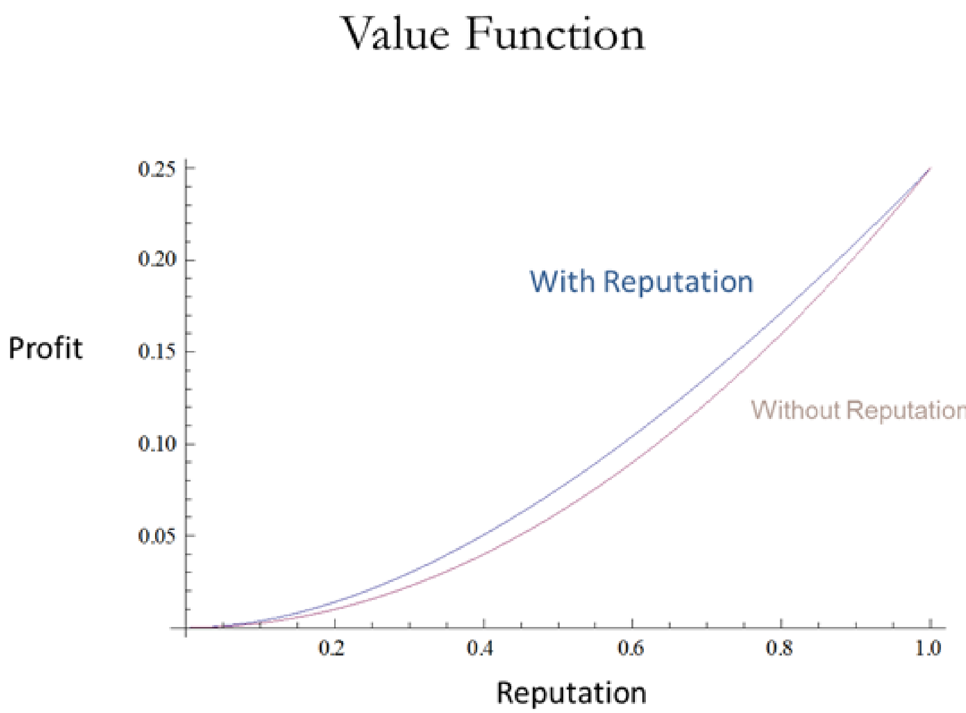

Figure 2.1: Value Function of the Firm with and without Reputational Concerns

general, this function is strictly increasing and concave in reputation x. By plugging this optimal quantity into the value function above, we can get an analytic solution for the value function.

$$V(x) = 2r + x + \frac{x}{4r} - \sqrt{x + 4r(r + x)}$$

Notice that at $x = 0$ the value is equal to 0 and at $x = 1$ the value is equal to $\frac{1}{4r}$, which is the profit maximizing level at both points. The value function is strictly convex in reputation, indicating that the firm has an incentive to overproduce at every level of reputation.

Figure 2.1 shows the value at every level of reputation. Here, it is assumed that $r = 1$. The blue line shows the value function above, and the purple line shows the value if there were no reputational dynamics and the firm stayed at the same reputation forever (so profit equals $\frac{x^2}{4}$). The value function is strictly above the value without reputation, showing how the ability to overproduce increases the firm's value.

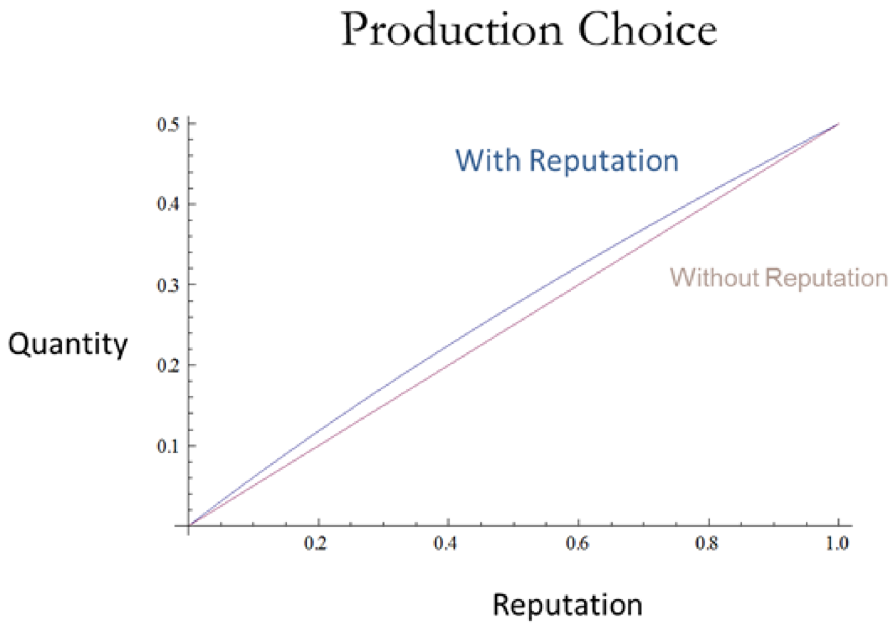

Figure 2.2: Production Choice of the Firm with and without Reputational Concerns

Figure 2.2 shows the production choice of the firm in this case. The quantity is always above the myopic profit maximizing quantity and is concave and strictly increasing. We note that the quantity does not have to be strictly concave, as it could be convex for low values of reputation as long as it equals the profit maximizing production choice at the endpoints.

2.6 Analysis of Informed Firm

Now we consider the model with informed firms that know their true quality with certainty. We now need to consider the market's beliefs about the firm's production choice since high quality and low quality firms may choose different quantities and thus send signals at different rates. How the market updates will depend both on the signal realization as well as the fact that the signal was actually received. For instance, if a high quality firm chooses to produce a higher quantity than a low quality firm, a high quality firm would be more likely to send a signal. Thus if the market receives a signal, this would indicate that the firm is likely to be high quality.

Suppose that the market believes that the high quality firm is producing a quantity $\tilde{q}_H$ and the low quality firm is producing a quantity $\tilde{q_L}$. Thus $\tilde{q_H}$ and $\tilde{q_L}$ will be the market's belief about the arrival rates of the signals from the two types. Using Bayes rule, we can find how reputation is updated depending on the signal realization. After the receipt of a good signal, the reputation will be updated according to

$$x_g = \frac{g\tilde{q_H}x}{g\tilde{q_H}x + b\tilde{q_L}(1-x)} = x + \frac{(g\tilde{q_H} - b\tilde{q_L})x(1-x)}{g\tilde{q_H}x + b\tilde{q_L}(1-x)}$$

And after the receipt of a bad signal, the reputation will be updated according to

$$x_b = \frac{(1-g)x\tilde{q_H}}{((1-g)x\tilde{q_H} + (1-b)(1-x)\tilde{q_L}} = x + \frac{((1-g)\tilde{q_H} - (1-b)\tilde{q_L})x(1-x)}{(1-g)\tilde{q_H}x + (1-b)\tilde{q_L}(1-x)}$$

It is important to note that the firm's reputation could actually rise after the receipt of a bad signal if

$$\tilde{q_L} < \tilde{q_H}\frac{1-g}{1-b}$$

In this case, the bad firm is expected to produce so little that receiving any signal is better than receiving no signal.

In the case that no signal arrives, we can again use Bayes rule and the markets beliefs to find the change in reputation. Without a signal, reputation is updated according to

$$x' = \frac{(1-\tilde{q_H}dt)x}{((1-\tilde{q_H}dt)x + (1-\tilde{q_L}dt)(1-x)} = x + \frac{(\tilde{q_L} - \tilde{q_H})x(1-x)dt}{(1-\tilde{q_H}dt)x + (1-\tilde{q_L}dt)(1-x)}$$

$$\Rightarrow dx \approx (\tilde{q_L} - \tilde{q_H})x(1-x)dt \quad \text{as} \quad dt \to 0$$

The value function for both types will still be the expected discounted value of the profit streams. Both types of firms must choose q_t to maximize its value at every time t given the market beliefs:

$$V_{\theta_0}(x_0) = E_{\theta_0,x_0,q_\theta,\tilde{q_H},\tilde{q_L}}\Big[\int_{t=0}^{\infty} e^{-rt}(q_t p_t - C(q_t))dt\Big]$$

We will again solve for the Markov perfect equilibrium of this model. A Markov perfect equilibrium will be defined by a set of quantity decisions by the firm $q(x)$ and market beliefs $\tilde{q}(x)$ such that:

1. Both firm's quantity choice maximizes their value functions:

$$q_\theta \in argmax_q\{V_\theta(x_0)\}$$

2. Market beliefs are correct: $\tilde{q_H}(x) = q_H(x), \tilde{q_L}(x) = q_L(x)$ for all x

We can write the continuous time Bellman equation for the high quality firm as

$$V_H(x) = \max_q (pq - C(q))dt + qdt(gV_H(x_g) + (1-g)V_H(x_b) - V_H(x+dx)) + (1-rdt)V_H(x+dx) \quad (2.11)$$

Here the three parts of the value function are the same as in the uninformed firm case, but there is an extra drift term in the value without signal. The value function of the low quality firm is similar and is given by

$$V_L(x) = \max_q (pq - C(q))dt + qdt(bV_L(x_g) + (1-b)V_L(x_b) - V_L(x+dx)) + (1-rdt)V_L(x+dx) \quad (2.12)$$

We can rewrite both value functions to get the following set of equations

$$V_H(x) = \max_q \frac{1}{r+q}(pq - C(q) + V_H'(\tilde{q_L} - \tilde{q_H})x(1-x)) + \frac{q}{r+q}(gV_H(x_g) + (1-g)V_H(x_b)) \quad (2.13)$$

$$V_L(x) = \max_q \frac{1}{r+q}(pq - C(q) + V_L'(\tilde{q_L} - \tilde{q_H})x(1-x)) + \frac{q}{r+q}(bV_L(x_g) + (1-b)V_L(x_b)) \quad (2.14)$$

These equations express the firm's value as a weighted sum of the flow payoff and the jump payoff. Increasing q shifts more weight to the jump term since signals become more likely. And increasing r shifts more weight on the flow term since the future is discounted more heavily. Now we can take the first order condition to get the equations for the optimal profit of the high quality and the low quality firm

$$(q_H + r)C'(q_H) - C(q_H) = r\Big[p + \Big(gV_H(x_g) + (1-g)V_H(x_b)\Big)\Big] + V_H'(\tilde{q_H} - \tilde{q_L})x(1-x)$$
$$(q_L + r)C'(q_L) - C(q_L) = r\Big[p + \Big(bV_L(x_g) + (1-b)V_L(x_b)\Big)\Big] + V_L'(\tilde{q_H} - \tilde{q_L})x(1-x)$$

Note that these are a pair of differential equations which may not be possible to solve analytically. We will not attempt to directly solve these equations. Instead, we will make some simplifying assumptions on the signal types to gain analytical tractability and derive certain results.

2.7 Noiseless Signals

We will first solve the above problem first for the case where the signals are perfectly precise. This is equivalent to saying that only high type firms can produce a good product and only low type firms can produce a bad product. Without noise, all the uncertainty will be resolved after the first signal arrives, so all the dynamics must take place before then. Thus, all the interior reputation movements will be determined solely by reputational drift. The two quantity expressions above can now be rewritten as:

$$(q_H + r)C'(q_H) - C(q_H) = r[p + (V(1))] + V_H'(\tilde{q_H} - \tilde{q_L})x(1-x)$$
$$(q_L + r)C'(q_L) - C(q_L) = rp + V_L'(\tilde{q_H} - \tilde{q_L})x(1-x)$$

We can prove that with noiseless signals, the high quality firm must produce a strictly higher quantity than the low quality firm.

Proposition 3. *For noiseless signals, at every reputation level in* $(0, 1)$ *and any set of market beliefs over production choices* $\{\tilde{q_H}, \tilde{q_L}\}$*, it must be the case that* $q_H > q_L$ *and* $V_H > V_L$*.*

Proof. Consider an arbitrary reputation level x_0 and an arbitrary set of market beliefs $\{\tilde{q_H}, \tilde{q_L}\}$. Note that fixing the market beliefs determines reputational drift at all reputational values. We compare the production incentives of the high quality firm and the low quality firm. First we write the value function at time t as the present discounted value of all future profits.

$$V_\theta(x_0; q_t, \tilde{q_t}) = E_{\theta, x_0, q_t, \tilde{q_t}}\left[\int_{t=0}^{\infty} e^{-rt}(\pi_t(q_t))dt\right]$$

Consider the stream of profits $\{\pi_\tau(q_\tau)\}_{\tau=t}^{\tau=\infty}$ that both types of firms receive from time t all the way into the future. Note that until the first signal arrives for both types, this stream of profits is exactly the same. After the first signal comes, the high firm gets $\pi(q(1))$ forever, while the low firm gets 0 forever. Thus the high firm has a strictly higher incentive to hasten the signal's arrival at every level of reputation.

To show this explicitly, let $x^\varnothing$ denote the path of reputation in the absence of a signal. Specifically, it is the deterministic solution to the ODE governing reputation change given by

$$d(x_t) = -(\tilde{q_H} - \tilde{q_L})x_t(1 - x_t)$$

We can then break down the value function for both types. Let Δ denote a small time increment, and let $j_s(x_t)$ be the new reputation after a signal, which with no noise will be 1 for a good firm and 0 for a bad firm.

$$\begin{aligned} V_\theta(x_0) &= E_{\theta, x_0}\left[\int_{t=0}^{\infty} e^{-rt}(\pi_t(q_t))dt\right] \\ &= \int_0^\Delta \pi_t(q_\theta(x_t^\varnothing))dt + e^{-r\Delta}\int_0^\Delta q_\theta(x_t^\varnothing)e^{-q_\theta(x_t^\varnothing)t}V(j_s(x_t))dt + e^{-r\Delta}e^{-\int_0^\Delta q_\theta(x_t^\varnothing)dt}V_\theta(x_\Delta^\varnothing) \\ &\approx \int_0^\Delta \pi_t(q_\theta(x_t^\varnothing))dt + e^{-r\Delta}\int_0^\Delta q_\theta(x_t^\varnothing)V(j_s(x_t))dt + e^{-r\Delta}e^{-\int_0^\Delta q_\theta(x_t^\varnothing)dt}V_\theta(x_\Delta^\varnothing) \end{aligned}$$

We have broken the value function up into three separate terms: the profit in the period $[0, \Delta]$, the continuation value after a signal arrives, and the continuation value without a signal arriving. Now we show that $V_H(x_0) - V_L(x_0)$ is strictly increasing in $q(x_t^\varnothing)$ for all values of $x_t^\varnothing$, which will be sufficient to prove that the high type has a higher optimal value of q than the low type.

$$\begin{aligned} & V_H(x_0) - V_L(x_0) \\ = e^{-r\Delta} \int_0^\Delta q_\theta(x_t^\varnothing)\Big(V\big(j_g(x_t)\big) - V\big(j_b(x_t)\big)\Big)dt &+ e^{-r\Delta} e^{-\int_0^\Delta q_\theta(x_t^\varnothing)dt}\big(V_H(x_\Delta^\varnothing) - V_L(x_\Delta^\varnothing)\big) \\ \approx e^{-r\Delta} \int_0^\Delta q_\theta(x_t^\varnothing)\Big(V(1) - V(0) - \big(V_H(x_\Delta^\varnothing) - V_L(x_\Delta^\varnothing)\big)\Big)dt &+ e^{-r\Delta}\big(V_H(x_\Delta^\varnothing) - V_L(x_\Delta^\varnothing)\big) \end{aligned}$$

This must be increasing in q because the difference between $V(1)$ and $V(0)$ must be greater than the difference at all other values of reputation since those are the maximum and minimum values respectively. This is sufficient to show that $q_H > q_L$ at all values of reputation. To show that $V_H > V_L$, we can iterate the formula above recursively to get the following infinite sum:

$$V_H(x_0) - V_L(x_0) = \sum_{n=0}^{\infty} e^{-r(n+1)\Delta} \prod_{j=0}^{j=n} e^{-\int_j^{(j+1)\Delta} q_\theta(x_t^\varnothing)dt} \int_0^\Delta q_\theta(x_{t+n\Delta}^\varnothing)(V(1) - V(0))dt$$

Since this is positive, our result holds for all values of x_0. □

2.8 Noisy Signals

Now we will introduce some noise into our problem. First, we will consider the case of partially noisy signals. Suppose that one of the signals is perfectly precise, while the other one is noisy. This could be the case if only high type firms can produce a good product, while low type firms could produce either good or bad products. Then low type firms are the only ones who can send the bad signal, which means the bad signal perfectly distinguishes the two types. But a good signal would still leave some ambiguity. Conversely, we could also assume that only the low type firm can produce bad signals, while the high type firm can produce good signals and bad signals. In that case the good signal is perfectly distinguishing while the bad signal is not.

Let us assume for now that the good signal is perfectly distinguishing while the bad signal leaves some doubt. The other case will be similar. In this case, the two value function expressions found in section 2.6 can be rewritten as:

$$V_H(x) = \max_q \frac{1}{r+q}(pq - C(q) + V_H'(\tilde{q}_L - \tilde{q}_H)x(1-x)) + \frac{q}{r+q}(gV(1) + (1-g)V_H(x_b))$$
$$V_L(x) = \max_q \frac{1}{r+q}(pq - C(q) + V_L'(\tilde{q}_L - \tilde{q}_H)x(1-x)) + \frac{q}{r+q}V_L(x_b))$$

For the high quality firm, a good signal boosts reputation up to one, giving the firm a value of $V(1) = \frac{\pi(1)}{r}$. The subscript is not written here because a low quality firm has the same payoff at one since its reputation does not change. A low quality firm will get a bad signal for sure and jump to reputation x_b.

The main difficulty in analyzing these pairs of equations is the derivative term in the equations. This derivative corresponds to the change in firm value with no signal arrival, and is multiplied by the drift without a signal arrival, $(\tilde{q}_L - \tilde{q}_H)x(1-x)$. With the derivative term, these equations become a pair of differential equations that are hard to solve. However, we can use the fact that this derivative term disappears at values of x where the drift is zero to calculate the production incentives at those points. Specifically, we will show that it cannot be possible to have zero drift at any point, because that would imply strictly higher production incentives for the high quality firm.

We will now impose a restriction on the types of beliefs held by the market. Specifically, we will assume that both $\tilde{q}_H$ and $\tilde{q}_L$ are continuous in x. This restriction is necessary to ensure that firm incentives do not change too much as x changes, since a big switch in the relative magnitudes of $\tilde{q}_H$ and $\tilde{q}_L$ would greatly change the amount that reputation jumps by after a signal. Although this restriction is somewhat strict, it is plausible because flow profits are continuous in reputation, so it seems reasonable for the market to believe that firms change their production choices smoothly as well. And we showed previously that $q(x)$ is indeed continuous in the uninformed firm case.

With continuous beliefs, we can prove a similar result to Proposition 3, that the high quality firm will always produce more than a low quality firm in any equilibrium.

Proposition 4. *For partially noisy signals, given any set of market beliefs $\tilde{q_H}, \tilde{q_L}$, we have $V_H > V_L$ for all $x \in [0, 1)$. In addition, in any possible Markov perfect equilibrium with continuous beliefs, it must be the case that $q_H > q_L$ for all $x < 1$.*

Proof. First consider an arbitrary reputation level $x \in [0, 1)$ and an arbitrary set of market beliefs $\tilde{q_H}, \tilde{q_L}$. As before, fixing market beliefs means that reputational drift is determined at all reputational values. Consider again the stream of profits $\pi_\tau(q_\tau)_{\tau=t}^{\tau=\infty}$ that both types of firms receive from time t all the way into the future. Now suppose that the high quality firm mimics the optimal strategy $q_L(x)$ of the low quality firm, and note that the two firms will have the same distribution over signal arrival times (but not signal types). Although a bad signal will not end the updating process, both types of firms will still have the same reputation after a bad signal. Thus the only difference in this profit stream occurs after the first good signal for the high firm. And since this boosts the high firm's reputation up to 1 giving it the greatest profits forever, the high firm must have a strictly higher value function at every level of reputation.

Now we prove that in any Markov perfect equilibrium, it must be the case that $q_H q_L$ for all $x < 1$. First, we show that at $x = 0$, the high quality firm will produce strictly more than the low quality firm. Note that drift is 0 at $x = 0$. The high quality firm and low quality firm have value functions given by

$$V_H(0) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(gV(1) + (1-g)V_H(0))$$
$$V_L(0) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}V_L(0)$$

Now since $gV(1) + (1-g)V_H(0) > V_L(0)$ by the fact that $V_H > V_L$, the high quality firm has a strictly higher incentive to increase q. Thus we have $q_H(0) > q_L(0)$. By continuity, this also implies that there exists $\varepsilon > 0$ such that $q_H(x) \geq q_L(x)$ for all $x \in [0, \varepsilon]$.

Now suppose for contradiction that $q_H(x) < q_L(x)$ for some x. Let $\hat{x} = \inf\left(x \mid q_H(x) < q_L(x)\right)$. Since $q_H(x) \geq q_L(x)$ for small x and the beliefs are continuous, we must have that $q_H(\hat{x}) = q_L(\hat{x})$, which means that $\tilde{q_H}(\hat{x}) = \tilde{q_H}(\hat{x})$ since we are in a Markov perfect equilibrium. Thus, the drift at $\hat{x}$

is equal to zero, and we can write the value functions for the firms at $\hat{x}$ as

$$V_H(\hat{x}) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(gV(1) + (1-g)V_H(\hat{x}_b))$$
$$V_L(\hat{x}) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}V_L(\hat{x}_b)$$

Since $gV(1) + (1-g)V_H(\hat{x}_b) > V_L(\hat{x}_b)$, the high quality firm has a strictly higher production incentive at $\hat{x}$, which means that $q_H(\hat{x}) > q_L(\hat{x})$, a contradiction. Thus there cannot be any x at which $q_H(x) < q_L(x)$.

□

Finally, we focus on the most complex case with fully noisy signals. In this case, we have $g, b \in (0,1)$ with $g > b$. Since the signals are fully noisy, the reputation of both types of firms will not change at 0 or 1, and both firms will have the same value at these reputations. Once again, we will impose restrictions on the beliefs to make sure that the firm's incentives are tractable. First, we will maintain our previous assumption that both $\tilde{q}_H$ and $\tilde{q}_L$ are continuous in x. In addition, we will also assume that the reputation jumps after a signal are increasing in x. From above, these jumps are equal to

$$x_g = \frac{gx}{gx + b(1-x)\tilde{q}_L/\tilde{q}_H}, \qquad x_b = \frac{(1-g)x}{((1-g)x + (1-b)(1-x)\tilde{q}_L/\tilde{q}_H}$$

Assuming that these jumps are increasing means that the ratio $\tilde{q}_L/\tilde{q}_H$ is not increasing too quickly in x. With this assumption, we can prove that the value functions $V_\theta(x)$ of both types of firms is strictly increasing in x. This is done with a revealed preference argument; a firm at a higher reputation level can mimic the strategy of a firm at a lower reputation level and receive strictly higher flow payoffs, while still maintaining a higher reputation after every history. Thus its value will be strictly higher than that of the lower reputation firm.

Given these two assumptions, we can prove that a high type firm must produce more than a low type firm at any reputation level.

Proposition 5. *For fully noisy signals, in any possible Markov perfect equilibrium with continuous beliefs and increasing reputation jumps, it must be the case that* $V_H > V_L$ *and* $q_H > q_L$ *for all* $x \in (0,1)$.

Proof. First we prove that the high quality firm has a strictly higher value for all $x \in (0,1)$. Consider an arbitrary reputation level$x \in (0,1)$ and focus on the stream of profits $\{\pi_\tau(q_\tau)\}_{\tau=t}^{\tau=\infty}$ that both types of firms receive from time t all the way into the future. Now suppose that at every time t the high quality firm mimics the strategy q_t of the low quality firm, and note that the two firms will have the same distribution over signal arrival times (but not signal types). After any signal, a high quality firm is more likely to jump to a higher level of reputation than the low quality firm, and thus receive strictly higher flow payoffs. By the monotonicity of the reputation jumps, the reputation of the high quality firm will never fall below the reputation of the low quality firm. Thus all the future flow payoffs of the high quality firm will be strictly higher in expectation, and thus $V_H > V_L$.

Since $V_H(0) = V_L(0)$, $V_H(1) = V_L(1)$, and $V_H > V_L$ for all $x \in (0,1)$, there exists some $\tilde{x} \in (0,1)$ such that $V'_H(\tilde{x}) < V'_L(\tilde{x})$. We write down the value functions for the high and low quality firm at $\tilde{x}$, and we suppose for contradiction that $\tilde{q_L}(\tilde{x}) \geq \tilde{q_H}(\tilde{x})$. Then

$$V_H(\tilde{x}) = \max_q \frac{1}{r+q}\Big(pq - C(q) + V'_H(\tilde{q_L} - \tilde{q_H})\tilde{x}(1-\tilde{x})\Big) + \frac{q}{r+q}\Big(gV_H(\tilde{x_g}) + (1-g)V_H(\tilde{x_b})\Big)$$
$$V_L(\tilde{x}) = \max_q \frac{1}{r+q}\Big(pq - C(q) + V'_L(\tilde{q_L} - \tilde{q_H})\tilde{x}(1-\tilde{x})\Big) + \frac{q}{r+q}\Big(bV_L(\tilde{x_g}) + (1-b)V_L(\tilde{x_b})\Big)$$

Note that the high quality firm has a strictly higher incentive to increase q, because

$$V'_H(\tilde{q_L} - \tilde{q_H})\tilde{x}(1-\tilde{x}) < V'_L(\tilde{q_L} - \tilde{q_H})\tilde{x}(1-\tilde{x})$$
$$\text{and} \qquad gV_H(\tilde{x_g}) + (1-g)V_H(\tilde{x_b}) > bV_L(\tilde{x_g}) + (1-b)V_L(\tilde{x_b})$$

Thus we must have that $q_L(\tilde{x}) < q_H(\tilde{x})$. Then by continuity of the beliefs, there must be some range $[a,b]$ such that the $q_H > q_L$ for all values in this range. Suppose for contradiction that there exists some x such that $q_L(x) > q_H(x)$. Again by continuity of beliefs there exists some range of beliefs $[c,d]$ such that: $q_H < q_L$for all $x \in (c,d)$, $q_H = q_L$ for $x \in \{c,d\}$, and either $c > 0$ or $d < 1$. Then let $\hat{x} = \max\{c, 1-d\}$, and note that $q_H(\hat{x}) = q_L(\hat{x})$ and $\hat{x} \in (0,1)$.

Now we can proceed in a similar fashion as in the proof of Proposition 4. Since $q_H(\hat{x}) = q_L(\hat{x})$, we have that $\tilde{q_H}(\hat{x}) = \tilde{q_L}(\hat{x})$ since we are in a Markov perfect equilibrium. Thus, the drift at $\hat{x}$ is equal to zero, and we can write the value functions for the firms at $\hat{x}$ as

$$V_H(\hat{x}) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(gV_H(\hat{x_q}) + (1-g)V_H(\hat{x_b}))$$
$$V_L(\hat{x}) = \max_q \frac{1}{r+q}(pq - C(q)) + \frac{q}{r+q}(bV_L(\hat{x_q}) + (1-b)V_L((\hat{x_b}))$$

Since $gV(1) + (1-g)V_H(\hat{x_b}) > V_L(\hat{x_b})$, the high quality firm has a strictly higher production incentive at $\hat{x}$, which means that $q_H(\hat{x}) > q_L(\hat{x})$, a contradiction. Thus there cannot be any x at which $q_H(x) < q_L(x)$.

□

2.9 Conclusion

This chapter has analyzed the optimal production choices of firms in a continuous time setting. By allowing the quantity to have a direct effect on the signals that the firm sends, firms must balance their current profit against future gains from reputational effects. In the uninformed case, firms must have convex value functions, which imply that it is optimal to always produce more than is profit maximizing. This can be interpreted as firms choosing to experiment in order to get information more quickly, and paying the cost out of current profits to do so. The implication is that new firms who are unsure of their true quality should try to produce at higher levels initially just so they and the market can learn as quickly as possible. Similarly, new workers should try to assume as many projects as possible to learn their true aptitude.

In the informed firm case, we placed some restrictions on the type of signals that were possible and found that the high quality firms would indeed produce more than the low quality firms. This confirms our intuition that low quality firms may wish to hide their true nature, while high quality firms want to "advertise" by getting more consumers using the product. In the cases of partially noisy and fully noisy signals, we also placed additional restrictions on the types of beliefs that the market could hold, and again found that high quality firms must produce more than low quality firms in any equilibrium.

For future research, we wish to verify our results under less restrictive assumptions on the beliefs of the market. The conjecture is that even with milder assumptions than continuity or monotonicity of the jump functions, low quality firms will still be producing strictly less than high quality firms because they want to hide their true quality from the market. The more signals a low quality firm sends, the faster its reputation should drop and the more future profits it will lose. And a high quality firm has the other incentive, to send more signals and get to a high reputation quickly. The problem with this line of argument though, is that the solution is highly dependent on the market's beliefs. It is certainly possible to create arbitrary beliefs whereby the high type firm wishes to work less with fully noisy signals. For instance, it could be the case that at an arbitrary reputation level x, the market believes that low type firms work very little compared to high type firms so that even a bad signal raises reputation significantly. However, suppose that at a reputation $x' > x$, low type firms are believed to work much more relative to high type firms, such that a good signal decreases reputation significantly to below even x. And suppose that at an intermediate reputation level $\hat{x} \in (x, x')$, the beliefs are such that a good signal increases reputation to x' and a bad signal decreases reputation to x. Then the good firm may wish to work less than the bad firm, since being at reputation x' is less desirable than being at reputation x. Indeed, it may even be the case that the low quality firm has a higher value at this point than the high quality firm! This type of scenario would need to be ruled out with a detailed look at what market beliefs are constrained to be in equilibrium. More research can be performed to characterize these reputational effects with fully noisy signals.

References

[Bar03] Heski Bar-Isaac. "Reputation and Survival: learning in a dynamic signalling model." *The Review of Economic Studies*, **70**(2):231–251, 2003.

[BH12] Alessandro Bonatti and Johannes Horner. "Career concerns with coarse information." 2012.

[Bon11] Alessandro Bonatti. "Menu pricing and learning." *American Economic Journal: Microeconomics*, **3**(3):124–163, 2011.

[BV13] Simon Board and Moritz Meyer-ter Vehn. "Reputation for quality." *Econometrica*, **81**(6):2381–2462, 2013.

[CE07] Kim-Sau Chung and Péter Eső. "Signalling with career concerns." Technical report, Discussion paper//Center for Mathematical Studies in Economics and Management Science, 2007.

[CM06] Judith A Chevalier and Dina Mayzlin. "The effect of word of mouth on sales: Online book reviews." *Journal of marketing research*, **43**(3):345–354, 2006.

[DAZ04] Chrysanthos Dellarocas, Neveen Awad, and Xiaoquan Michael Zhang. "Using online reviews as a proxy of word-of-mouth for motion picture revenue forecasting." *Available at SSRN 620821*, 2004.

[Hol99] Bengt Holmström. "Managerial incentive problems: A dynamic perspective." *The Review of Economic Studies*, **66**(1):169–182, 1999.

CHAPTER 3

A Dynamic Model of Certification and Reputation

Econ Theory (2015) 58:509–541
DOI 10.1007/s00199-014-0836-9

RESEARCH ARTICLE

A dynamic model of certification and reputation

Mihaela Van Der Schaar · Simpson Zhang

Received: 30 April 2014 / Accepted: 28 July 2014 / Published online: 5 October 2014

Abstract Markets typically have many ways of learning about quality, with two of the most important being reputational forces and certification, and these types of learning often interact with and influence each other. This paper is the first to consider markets where learning occurs through these different sources simultaneously, which allows us to investigate the rich interplay and dynamics that can arise. Our work offers four main insights: (1) Without certification, market learning through reputation alone can get "stuck" at inefficient levels and high-quality agents may get forced out of the market. (2) Certification "frees" the reputation of agents, allowing good agents to keep working even after an unfortunate string of bad signals. (3) Certification can be both beneficial and harmful from a social perspective, so a social planner must choose the certification scheme carefully. In particular, the market will tend to demand more certification than socially optimal because the market does not bear the certification costs. (4) Certification and reputational learning can act as complementary forces so that the social welfare produced by certification can be increased by faster information revelation.

The authors are deeply indebted to Moritz Meyer-ter-Vehn, Simon Board, Oren Louidor, and Bill Zame for their guidance. This paper also benefitted from valuable discussions with Jeff Ely, Crysanthos Dellarocas, Sanjeev Goyal, Johannes Hörner, David Parkes, Yangbo Song, Jie Xu, and Kartik Ahuja. We thank several anonymous referees for their feedback. We gratefully acknowledge financial support from the National Science Foundation Grant CCF-1218136.

M. Van Der Schaar (✉)
Department of Electrical Engineering, UCLA, Los Angeles, CA, USA
e-mail: mihaela@ee.ucla.edu

S. Zhang
Department of Economics, UCLA, Los Angeles, CA, USA
e-mail: simpsonzhang@ucla.edu

Keywords Certification · Reputation · Learning · Market Design · Social welfare

JEL Classification D82 · D83 · D47 · M51

1 Introduction

A market typically has several sources from which to learn about the quality of an agent, such as the agent's prior work history, or reports from trusted outside sources like a certification board. An agent's work results provide a steady, but noisy, flow of information to the market, allowing the agent's reputation to be continuously updated. Certification boards also exist in many industries such as healthcare and accounting, and certification verifies that a minimum level of quality has been met. The previous literature has explored markets where quality is learned through each source of information separately, but the combined effects of the two have not yet been well analyzed. Given that these different informational sources often coexist, important questions are raised regarding how they can, and should, interact in a marketplace. As an example, physicians can become board certified to demonstrate proficiency in a medical specialty, and patients also learn about physicians from the experiences of previous patients. While certification helps increase a physician's reputation, it also places significant costs on the physician. In such a market, will the physician choose to perform board certification immediately after entering, or only after some time depending on his performance results? And how stringently should the medical board set its certification requirements if it wishes to maximize social welfare? By considering a general model which incorporates both types of learning, this paper is able to address such questions.

Our main results show that certification remains an important component of a marketplace even when learning can be done through other channels. The key observation is that certification provides a very different kind of information to the market, a type of information that is socially beneficial due both to its informativeness as well as its dependability. Without certification, even high-quality agents may experience an unfortunate string of bad reports and so get forced out of the market. Once out of the market, agents can no longer work and thus cannot produce more information that would change the market's opinion. Certification, however, remedies this problem by acting as a safety net, letting these socially beneficial agents continue to work. The type and amount of information revealed by the agent's work history can have a strong effect on the social welfare provided by certification, and we show that contrary to the case without certification, faster information revelation can be socially beneficial. Thus, certification and the agent's work history can act as complements.

In this paper, we focus attention on markets with pricing frictions that result in fixed wages for the agent. With pricing frictions, information revelation can get "stuck" at inefficient levels of reputation,[1] so that the learning via reputation may not actually increase social welfare. Since principals are myopic and have limited incentive to

[1] See for instance the discussion in Sect. 3 of Bar-Isaac and Tadelis (2008).

experiment, the market does not learn in a socially efficient manner. For instance, patients who see physicians for special procedures are unlikely to return and thus do not internalize the value produced through their own observations. Thus, they may refuse to see newer and less known physicians, even though the visit would produce information beneficial to future patients. If wages are fixed for all physicians, these less known physicians may not be able to get employed altogether, even if their true quality was actually very high. We show in Theorem 1 that under certain conditions, reputational learning alone cannot increase social welfare due to this inefficiency of the learning process. Thus, the markets that we analyze are the ones in which certification can be the most important, by reducing asymmetric information and keeping high-quality agents in the market.

While our fixed wage assumption may represent a departure from the literature, which usually assumes completely flexible wages like in Holmström (1999), we argue that fixed wages are a reasonable assumption in many real-world markets where certification is present. For instance, doctors who see Medicare patients are paid according to a fixed fee for service schedule. This payment schedule is decided by the government every year and remains the same for all doctors regardless of their reputations. In this market, certification plays an extremely important role as doctors can voluntarily choose to get board certified in order to verify their expertise. Although certification does not increase the wage a doctor receives, it helps ensures a steady flow of patients by increasing the doctor's reputation. Outside the US, many European markets also have prices for medical services that are completely fixed by the government, and there are similar informational issues.

In addition to a mandated fixed wage, our model is also applicable to markets with binding price floors, such as a minimum wage or union negotiated price. Especially for lower skilled workers, such as florists or custodians, these price floors can often be binding. In such markets, the wage will be "fixed" at the level of the price floor because the agent cannot work for less and the agent does not have enough market power to demand more. Agents still benefit from good reputations because it helps them find work, while agents that develop bad reputations will not get hired. Certification can be quite important to such markets both to guarantee the quality of workers for employers, as well as to boost the reputation of high-quality workers and help them to continue working. The regulation and licensing of lower skilled workers have been on a sharp rise in recent years,[2] so the proper use of certification in these markets represents a very timely issue.

Our model features reputational dynamics related to those in Holmström (1999), with an agent of uncertain quality working for a market of homogenous short-lived principals. The agent knows its own quality, but the market does not and must infer quality through the agent's work history, which generates a stochastic process that is publically observable. The agent's work history thus provides a steady stream of information to the market, continuously updating the agent's reputation. Our model imposes few assumptions on the stochastic process itself, or the agent's initial quality distribution. In contrast, most papers in the reputation literature make much stronger

[2] See for example the New York Times article "Why License a Florist" (http://nyti.ms/1wnLUZk, Kleiner (2014)).

assumptions such as binary quality types or Brownian motion signals[3]. Conditional on its work history, the agent chooses when (or whether) to certify, revealing additional information to the market. Certification has a cost to the agent of k and verifies that true quality lies above some standard $\underline{q}$, but does not reveal the exact quality level. This type of "imprecise" certification is one of the most common in practice, with examples such as board certification for doctors, pass/fail exams for accountants, and security seals for websites.

We first consider a benchmark model without certification and learning only through the agent's work history. Even in this simpler setting, we find surprising results. The key driver of these results is that high-quality agents can be forced out of the market after a string of bad signals, after which learning stops and the agent's reputation has no chance to improve. Since learning may be halted prematurely, having access to the agent's work history alone does not guarantee a socially efficient outcome. In fact, our result shows that for many types of information processes,*work history information does not provide any social benefit at all*. This result is reminiscent of work on "bad reputation" (Ely and Välimäki 2003; Ely et al. 2008) where the ineptness of reputation is due to the myopia of the short-lived principals. But in our model, the failure of reputational forces alone results not because of moral hazard, but because of an informational friction between principals and agents. Principals are not willing to hire once the agent's expected quality falls to the price level, but there is still positive social value from hiring at this because an agent with a low reputation might still be of high quality—hiring has an informational value, but since the principal only interacts once, it does not appropriate this value. Thus, principals would force agents, even agents who face no moral hazard issues, out of the market inefficiently due to their short-sightedness.

Given the ineffectiveness of learning through the work history alone, we then introduce certification into the model to see if welfare can be improved through a combination of the two information sources. The decision to become certified or not at any given moment is a strategic one, which depends on the results of the agent's previous work history. There are many equilibria as in most signaling games. However, we show that all equilibria share a common feature: They are type independent, meaning that all types of agents could get certified will choose to certify after the same work histories. This implies that every equilibrium is characterized by a single certification stopping time strategy that is used by all types of agents regardless of quality. But there are many different stopping times strategies that could constitute an equilibrium, with some requiring the agent to certify very early on, or even immediately, and others requiring the agent to certify later. We characterize the specific equilibria that maximize principal, agent, and social welfare and rank them based on how quickly certification occurs: Agents prefer later certification than is socially optimal, and principals prefer earlier certification than is socially optimal.

After characterizing these three different types of equilibria, we then perform comparative statics to find the optimal certification standard and price level for each type of equilibrium. We find that certification increases welfare if the standards are set

[3] Such as Holmström (1999), Bar-Isaac (2003), and Bonatti (2011).

appropriately, and certification costs are low. In addition, the information produced by the agent's work history and certification act in a complementary fashion, with faster information revelation increasing the welfare generated by certification. However, certification will also decrease welfare if certification costs are very high or the standards are set inappropriately. When costs are too high, certification will destroy more social welfare than it generates. But even when costs are low, certification may not increase welfare if the certification standards are set incorrectly. For instance, in some equilibrium, the market belief will expect too much certification. Such a case can occur in the principal optimal equilibrium, since principals do not internalize certification costs. Because agents are forced to certify at inefficient reputation levels, overall welfare could actually be reduced. Thus, it is important to set certification standards appropriately depending on the specific market beliefs.

2 Literature review

2.1 Reputation papers

Our work is closely related to several papers in the reputation literature, such as Holmström (1999), Mailath and Samuelson (2001), Bar-Isaac (2003), and Ely and Välimäki (2003). All of these papers have some form of reputational mechanism that follows from work history, but ours is the first to allow the agent to certify and send information through this channel. Holmström (1999) presents the classic "signal jamming" model where an agent of unknown quality can exert effort to bias the market's perception. Contingent contracts are not possible and without reputational incentives the agent would exert no effort. Holmström finds that reputation can provide work incentives in the short run while the agent's quality is unknown, but not in the long run once quality does become known. Crucially, there is no exit in this model and so the agent's quality will become known perfectly over time. We show that when there is an exit point, perhaps due to market frictions resulting in fixed prices, then agents may be forced out before their true quality is revealed. In this case certification is necessary as a form of insurance, so that high quality agents can stay in the market and continue working. Mailath and Samuelson (2001) also considers permanent reputations with moral hazard like Holmström, but with firms that know their own qualities. They show that good firms will build up reputations and exert effort in order to separate themselves from low-quality firms. In our paper, there is no moral hazard and good agents will instead certify to separate themselves. Also related is Board and Meyer-ter-Vehn (2013), in which a firm invests in order to improve its quality and reputation, while in this paper the quality is fixed, but can be verified through certification. Finally, Ordoñez (2013) analyzes a situation where managers can take actions such as scapegoating after failures or successes in order to manage their reputations, which is similar to our model where agents can choose to certify in order to increase their reputations.

Unlike these previous papers that do not consider exit, a reputational paper that does focus on exit is Bar-Isaac (2003). This model arrives at the striking conclusion that high-quality agents never exit the market, because staying in the market is a signal of quality that increases reputation. Even an agent who receives a string of bad signals can demonstrate resolve by refusing to quit, which boosts the market's perceptions. Importantly, Bar-Isaac assumes that the wage varies at every moment in time to equal the agent's expected quality, so good agents can internalize the future benefits of reputation and are thus willing to sustain a period of negative payoffs. While flexible wages are reasonable for some markets, we argue that there are also some markets where fixed wages are more reasonable. Flexible wages require the agent to have market power (in fact the agent is extracting full surplus from principals), which can be unrealistic in many types of markets where agents have unknown qualities. Without market power (e.g., Bertrand competition), agents may be unable to internalize the benefits of high reputation and thus unwilling to endure negative flow payoffs. Second, good agents may be forced to sustain numerous periods of negative payoffs in equilibrium, which may not be feasible for markets where they have liquidity constraints. With liquidity constraints, both good and bad agents may be forced to exit at the same time once their reputation drops below cost. Finally, in some markets factors such as market customs and menu costs can make changing wages quite difficult. When wages are fixed, we show opposite results: High-quality agents are forced to exit the market with positive probability and learning may not be socially efficient.

While most papers in the literature show that reputation can increase social welfare, the papers by Ely and Välimäki (2003) and Ely et al. (2008) are noteworthy in showing cases where reputation is potentially harmful. Ely and Välimäki (2003) consider a car repair framework, with both honest and dishonest mechanics. Honest mechanics wish to recommend the best repair for a car, be it a cheap tune-up or an expensive engine repair, whereas dishonest mechanics always want to recommend engine repairs. The mechanic's reputation represents the probability of being honest, which will change over time depending on how many tune-ups or engine repairs are recommended. The paper shows that a reputational mechanism may actually destroy social surplus: Honest mechanics have an incentive to recommend tune-ups even when engine repairs are truly needed in order to boost reputation. But such an action hurts the consumer, who is short lived and so does not internalize the benefit of higher reputation for the mechanic. In equilibrium, the market may break down as consumers are not willing to go to either type of mechanic. Ely et al. (2008) extends this model to a more general framework and shows that similar results will hold in models that feature "temptation" actions for the long-lived player. These "temptations" are actions that boost reputation but are socially inefficient, like recommending tune-ups in place of engine repairs. The assumption of short-lived consumers is critical in both papers, as Ely and Välimäki (2003) shows that a long-lived consumer could devise a mechanism that benefits from the information process. The results in these papers are related to those of our model, which also considers a market of short-lived players. Because of myopia, the market is unable to internalize the benefits of experimentation and so learning may stop early at an inefficient level. But our model does not require moral hazard in the way of "temptations," as agents face no moral hazard in their actions. Even without moral hazard, we show that the market myopia by itself can be sufficient to undercut the value of reputation if prices are fixed and principals unwilling to experiment. The inefficiency thus results from an informational friction rather than a strategic friction.

Finally, there are some related papers that consider how entry costs may be necessary to complement the gains from reputation in markets where there is free entry. Klein and Leffler (1981) shows that if there are moral hazard issues, firms need to earn price premiums in order to incentivize them not to cheat. With free entry, such price premiums cannot be sustained in equilibrium, and so it is important to have fixed entry or investment costs to support these premiums. Atkeson et al. (2012) also consider a reputational model with free entry. In their paper, producers make quality decisions at the time of entry, and these qualities are then fixed afterward. The market learns over time about quality via signals as in our paper, and the government can impose fixed entry costs for entry into the market. These costs do not produce information as in certification, but do allow for higher-quality firms to enter as in Klein and Leffler, leading to greater social welfare. While we do not explicitly model free entry in our paper, certification is costly and can be thought of as an entry barrier to a market. Indeed, we show that in some markets certification will act as a *de facto* license, with non-certifying agents unable to ever find work. However, there are two main differences between certification and the entry costs in these other papers. The first is that certification provides information to the market, and so serves more than just a signaling role since not all agents can pass certification. The second is that certification can also happen after an agent enters the market, with the agent choosing to certify only if it sends bad signals. Thus, very high-quality agents may not need to pay the certification cost if they produce good results, leading to a potentially more efficient outcome than one where all agents must pay an entry cost.

2.2 Certification papers

Our work represents a novel approach to the certification literature, where the agent's certification choice has never been combined with a similar reputational mechanism in a dynamic setting. Much of the focus of the theoretical certification literature has instead been on the certifier's actions and revolves around studying the decisions made by a strategic certifier who can control the type of information that it releases about agents or the payments that it charges (Lizzeri 1999; Stahl and Strausz 2011; Farhi et al. 2013). Some papers also allow the certifier to collude with agents and assign false ratings as in Strausz (2005). In contrast, our paper focuses more on the agent's decision process and how its reputation affects its certification decision and the principals' beliefs. This allows us to analyze the strategic aspects of certification for the agent, and show how even the possibility of certification will affect an agent's reputation. We do not explicitly model the certifier, instead delegating it to the role of the mechanism designer and analyzing the comparative statics of our model with regard to the certification standard. This implicitly assumes that the certifier is not strategic, is always accurate in its judgments, and is not allowed to cheat. Thus, our work is best suited to markets where the certifier is a government agency, which would be trustworthy and sets certification standards in a benevolent fashion. Finally, we note that while papers by Shapiro (1983) and Panzar and Savage (2011) have considered the welfare implications of minimum standards when agents have reputational concerns, the focus of our work is different because these models consider reputation in terms of moral hazard instead of adverse selection, and so there is no learning about the agent's type.

Although our work on combining agent reputation and certification is new from a theoretical perspective, there have been interesting empirical papers published on this subject. For instance, Xiao (2010) tests the value of certification in the childcare industry. In childcare, certification is voluntary and usually chosen by only a small fraction of firms. The paper shows that the social value generated by certification is positive, albeit small. Further, the costs of certification are not negligible, such as administrative and personnel costs. Thus, certification on the whole is a negative for the industry, as the benefits are outweighed by the costs. In our model, we show that this is certainly a possibility when certification is present, in particular when market beliefs are pessimistic and require certification often. In such a situation, the informational value of certification is outweighed by the inherent costs, resulting in lower social welfare.

Jin and Leslie (2003) analyzes the impact of the introduction of government issued hygiene grade cards on restaurant cleanliness in Los Angeles. They find that cleanliness did in fact improve significantly after the government started using these grade cards. Restaurant revenue is significantly affected by the posted scores, with restaurants that received an "A" grade getting a 5.7 percent increase in revenue from mandatory disclosure, while restaurants that received a "B" got only a .7 percent increase, and restaurants that received a "C" got a 1 percent decrease. This provides evidence that consumers do pay attention to this additional source of information, and that market demand was changing as a result. Jin and Leslie (2009) extends this study by considering whether reputational effects alone are adequate to explain this restaurant behavior. The paper assumes that restaurants affiliated with chains have an incentive to free ride on the chain's reputation. Thus, franchised stores tend to have lower hygiene scores than company-owned stores. After the introduction of the grade cards, the authors assume that such incentives would go away, because consumers are able to infer hygiene directly from the posted scores. From the data, the authors find evidence that the introduction of grade cards causes hygiene scores to increase more for non-chain stores than for chain stores, lending support to the idea that chain stores care more about reputation in the absence of grading cards. And in comparing the effect of reputation versus the effect of the grade cards, the authors find that hygiene improvements due to reputation were 70 % as large as hygiene improvements due to the posted grade cards. The significant changes in hygiene after grade cards were introduced thus shows that certification can send a clearer output to the market than mere reputational effects, and that reputation alone may not be sufficient to solve adverse selection problems. Our model echoes this result, by showing that certification can increase social welfare even if the market can learn via the agent's reputation.

3 The model

We consider an infinite horizon continuous time model with a single long-lived agent and a marketplace of principals.[4] The agent has a fixed quality q that is determined at the start of the game according to a commonly known continuous distribution

[4] Alternatively, we can think of the agent as representing a firm and the principals as representing consumers.

$q \sim f_0(q)$, which has a finite mean μ_0. Denote the support of this distribution by D, which can be any measurable non-null set in $\mathbb{R}$. The agent is privately informed about its own quality, but the market must learn this quality over time. We assume the agent has a reservation value of $c < \mu_0$, which represents the agent's disutility of work and outside options. The agent also has a discount rate of ρ. At each moment in time the agent meets a different short-lived principal, who can make an offer to the agent at a fixed price p. We will assume throughout the paper that the price satisfies $p \geq c$ so that agents are always willing to accept an offer.

If the agent accepts the offer, it will work for one infinitesimal instant of time and return output with true quality q. The market's observations (or reports) of q are noisy, and the evolution of these observations follows a càdlàg[5] stochastic process $R_q(t)$. This stochastic process will depend on the agent's true quality q, with a different stochastic process $R_q(t)$ corresponding to each q. We define the agent's *work history* as the history of all previous observations by the market, $\mathcal{H}_t = \left\{R(t')\right\}_{t'=0}^{t}$, which is a continuous set of realizations of the stochastic process up to time t. We also call the set $\mathcal{L} \equiv \left\{R_q(t)\right\}_{q\in D}$ an *information process* for the market. The set $\mathcal{L}$ is common knowledge, so the market knows what kind of stochastic process to expect from each type of agent. Thus, given any history $\mathcal{H}_t$, the market can update its prior beliefs about the agent via Bayes' rule. We define the agent's time t *work history rating* as the market's belief of the agent's quality distribution given only information from the work history, $f_t(q|\mathcal{H}_t)$.

We make some restrictions on the information process $\mathcal{L}$ to ensure that the agent's work history rating will evolve in a sufficiently continuous fashion over time. Let $\mu_t \equiv E[q|\mathcal{H}_t]$ be the time t expected mean of the agent's quality, and let $\mu_t^{\underline{q}} \equiv E\left[q|\mathcal{H}_t, q < \underline{q}\right]$ be the time t expected mean of the agent's quality given that it lies below some cutoff $\underline{q}$. We say that the information process is *admissible* if for any $q \in D$, the generated paths of μ_t and $\left\{\mu_t^{\underline{q}}\right\}_{\underline{q}\in\mathbb{R}}$ are almost surely right-continuous and upper semi-continuous in time. That is, at any point of discontinuity these means can jump upwards but not downwards. The reason that our continuity requirement needs only to be one-sided will become apparent in the next section. We will only use admissible information processes in this paper.

Note that a restriction on $\mathcal{L}$ entails a restriction on the types of stochastic signals $R_q(t)$ that can be allowed. One way to satisfy our continuity requirement is for the market to only be able to infer a little bit of information from $R(t)$ at any point in time. For instance suppose that for every q, $R_q(t)$ is the diffusion $dR_q(t) = q\,dt + \sigma_r dZ(t)$, with drift q and variance σ_r^2 (and precision τ_r), and with $Z(t)$ a standard Brownian motion. This type of process would result if the market observes each unit of output with some normal noise, and we note that this is the continuous time non-moral hazard version of Holmström (1999). In this case, the agent's work history rating would evolve continuously in time, and so our admissibility requirement is satisfied. Even more general diffusion processes are admissible, such as $dR_q(t) = f(q, \mathcal{H}_t)dt + g(\mathcal{H}_t)dZ(t)$ so that the drift and variance do not have independent

[5] Right-continuous with left limits.

or stationary increments. However, stochastic processes such as $dR_q(t) = qdt + q\sigma_r dZ(t)$, where the variance depends on the quality, are not admissible because the variance would be learned immediately through Bayes' rule, causing the means to jump downwards with positive probability. The variance may depend on the history of previous signals or any other public information, but not on the agent's true quality.

Stochastic processes with jumps can also be considered. For instance, suppose we have a Poisson process with an arrival rate $\lambda(q)$, where λ is increasing in q. This is a good news Poisson process, where an arrival indicates a positive event like being mentioned in an article or winning an award. Without a signal arrival, the means drift slowly and continuously downwards, and at a signal arrival, the expected mean and all truncated means would jump upwards. Thus, our admissibility condition would be satisfied. More generally the arrival rate could depend on the history as well, $\lambda(q, \mathcal{H}_t)$, as long as it was still increasing in q at all t. And if the stochastic process was a combination of diffusions and good news Poisson processes, the result could still hold. However, if the stochastic process was a bad news Poisson process, where $\lambda(q)$ was decreasing in q, then admissibility would be violated because at a signal arrival the means would jump downwards. So with bad news, the path of μ_t may not be right-continuous and upper semi-continuous. Thus, jump stochastic processes are admissible as long as they indicate good news, but not if they indicate bad news.

Finally, note that the stochastic process of reports about the agent only runs if the agent is actually working. If at a time t the principal does not hire the agent, no output gets produced and the stochastic process is stopped at that value of $R(t)$. Since the agent is not working, no further information gets sent. If the agent does get hired again in the future, then the stochastic process can once again proceed at that time. We will show that this can only happen if the agent passes certification; that is unless the agent certifies he will be kicked out of the market forever.

At every moment of time t, the agent can choose to certify if it has not done so already by paying a one-time certification cost of $k > 0$ and getting certified if it has a quality level of at least $\underline{q}$, which is a fixed exogenous standard. We assume that $\underline{q} \geq p$, which will imply that certified agents never get forced out of the market. We also assume that $k < \frac{p-c}{\rho}$ so that the cost is low enough for the agent to want to choose certification (this inequality implies that the net present value of staying in the market forever is greater than the certification cost). We denote the time t certification status by $\theta_t \in \{[0, t] \cup \phi\}$, where a number represents the time at which that agent became certified and ϕ means that the agent has not yet certified. Once the agent becomes certified, all future principals will know for certain that the agent quality is at least $\underline{q}$. Since this information is permanent and public knowledge to all future principals, an agent only needs to be certified once.

We also assume principals do not know if an agent attempts to certify and fails, so an agent that cannot pass certification has no incentive to certify. Such an assumption is relevant when $\underline{q} > p$. If the market could observe agents that certify and fail, an agent with quality $q < \underline{q}$ may have an incentive to attempt certification to increase its reputation. For instance, if the market believed that attempting certification was more likely by agents with qualities in the range $[p, \underline{q}]$, then attempting certification, even when the agent cannot pass, would increase the agent's reputation.

There are thus two sources of information in our market: The first being the history of observations from work history, $\mathcal{H}_t = \left\{R(t')\right\}_{t'=0}^{t}$, and the second being the certification status of the agent θ_t. Upon observing these sources of information, the market uses Bayes' rule in updating its beliefs about agent quality. Together, these two signals combined with the prior quality distribution will result in a posterior belief distribution of agent quality $f_t(q|\theta_t, \mathcal{H}_t)$, which we call the *reputation* of the agent. This is different from the work history rating that was defined previously because it also takes into account the agent's certification status. Note that the reputation can be calculated from the work history rating by using Bayes' rule together with θ_t on the work history rating $f_t(q|\mathcal{H}_t)$. We will use $F_t(q|\theta_t, \mathcal{H}_t)$ to denote the cdf of the reputation. Note that we sometimes suppress the notation and write $f_t(q)$ and $F_t(q)$ for the reputation and cdf of the reputation, respectively. We also use $F_t^-(q)$ to denote $\lim_{q \nearrow \underline{q}} F_t(q)$.

We assume principals have identical linear utilities given by $U\,(p, q) = q - p$. Since principals are short lived, they care only about maximizing their current utility. Given the above linear utility function, we assume principals will hire uncertified agents if and only if $p < E[q|\theta_t = \phi, \mathcal{H}_t]$, where the expectation is taken with respect to the agent's time t reputation. Principals will hire an agent that certified at t' if and only if $p \leq E[q|\theta_t = t', \mathcal{H}_t]$.[6] Note that once the agent's expected quality falls below p the agent will not be able to work again unless it certifies, because principals would not be willing to hire at any quality level less than p, so no outputs will be produced and no further information sent. Thus, the market's beliefs will be stuck unless the agent improves its reputation through certification.

If the agent works at time t, it receives a flow payoff of $\pi_t = p - c$. Otherwise, the agent receives a payoff of $\pi_t = 0$. And the agent must pay a one-time cost of k when it chooses to certify, which can be incorporated by subtracting the flow payoffs after certification, changing them to $\pi_t = p - c - \rho k$ (our assumption that $\underline{q} \geq p$ implies that a certified agent will always keep working). Thus, the agent decides whether to certify at time t through maximizing its expected discounted value, $\int_{t'=t}^{\infty} e^{-\rho(t'-t)} E\left[\pi_{t'}|\theta_t, \mathcal{H}_t, q\right] dt'$, taking into account the state variables: the agent's certification status $\theta_t = \phi$, the history of outputs $\mathcal{H}_t$, and the agent's true quality q. The agent's certification strategy can be represented by an optimal stopping time that is measurable with respect to the filtration generated by $R_q(t)$. We denote the stopping time strategy for an agent of type q by $\tau(q)$, which depends on the specific history of signals $\mathcal{H}_t$, and we note that $\tau(q) = \infty$ for all $q < \underline{q}$ since these agents cannot pass certification. The market must also have beliefs about the certification strategy of agents of each possible quality, and we represent these beliefs using $\tilde{\tau}(q)$. In equilibrium, these beliefs must be correct so that $\tilde{\tau}\,(q) = \tau(q)$.

4 Benchmark case: no certification

In order to analyze the welfare benefits of certification, we will first derive some welfare results for a benchmark setting with no certification. This can be considered

[6] For technical reasons, we assume principals who are indifferent will hire certified agents, but not uncertified agents. The distinction ensures that certified agents will be hired regardless of the market's beliefs, while uncertified agents have a well-defined time at which they leave the market.

a specialized case of the above model with $\underline{q} = \infty$. A principal will hire at time t if and only if $p < E[q|\theta_t, \mathcal{H}_t]$, with the expectation taken with respect to the agent's reputation, $f_t(q|\theta_t, \mathcal{H}_t)$. Since the agent will never certify, its reputation is always equal to the work history rating, $f_t(q|\mathcal{H}_t)$. To analyze this model we thus need to determine how the agent's reputation evolves as a function of its work history. Working generates a stochastic process $R_q(t)$, and without certification the market will update its beliefs only through this stochastic process itself. Through Bayes' rule the agent's expected quality level, $\mu_t \equiv E_t[q|\mathcal{H}_t]$, will be continuously updated as the market receives information. The reputational dynamics in the benchmark are as follows: The agent starts out with a quality $\mu_0 > p$ (for $\mu_0 \leq p$ no hiring ever takes place) and continues working as long as its expected quality stays higher than p, but the first time that its expected quality drops to or below p it will be forced to stop. At this point, no principals are willing to hire the agent, and thus, the stochastic process is stopped forever causing no future principals to hire either (Fig. 1).

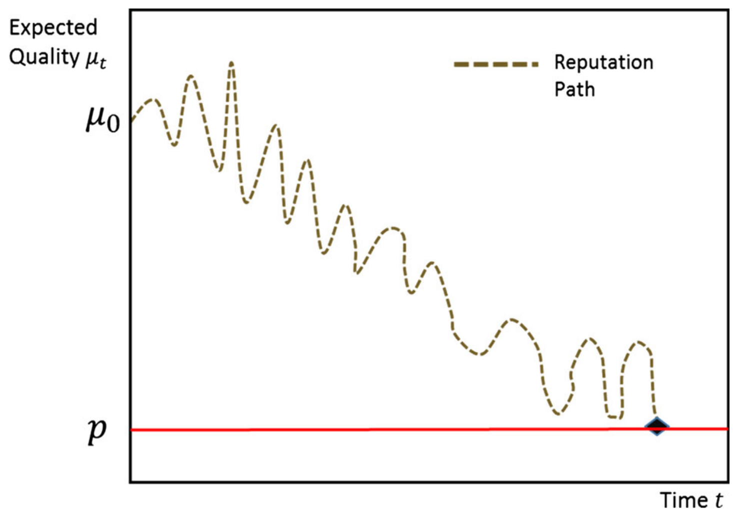

Fig. 1 *Reputational dynamics in benchmark*. The agent starts out with an expected quality greater than the price, and the expected quality then evolves according to the stochastic process generated by the outputs. The agent continues getting hired until its expected quality falls to the price, at which point it stops working forever. The probability that the agent stops working depends on its true quality and the price and may be strictly positive even for very high-quality agents

Given this characterization of the dynamics, once the agent stops working it can never start again. Thus, the probability that the agent is still working at time t, denoted by $P(S_t|q, p, \mathcal{L})$, is the same as the probability that the hitting time of μ_t against the price p or any lower value is greater than t. Since we only consider admissible information processes, at the first time that μ_t falls below p it must be that $\mu_t = p$. This is due to the fact that the agent's expected quality can never jump downwards. Let $t^* \equiv \inf\{t|\mu_t = p\}$ denote this (stochastic) hitting time.[7] So $P(S_t|q, p, \mathcal{L})$, the probability that the agent is still working at time t, is equal to the probability that $t^* > t$. Now we can write out the *ex ante* expected social surplus (sum of principal and agent surplus) for a price p as:

[7] Note that when $p \geq \mu_0$, we have $t^* = 0$ and the agent is never hired.

$$W(p,\mathcal{L}) = \int_{-\infty}^{\infty}\left[\int_{0}^{\infty} e^{-\rho t}\,(q-c)\,P\,(S_t|q,p,\mathcal{L})\,dt\right] f_0(q)dq$$

For an agent with quality q, the social welfare at every moment will be given by $q-c$ if the agent is hired and 0 otherwise. Thus, hiring is socially optimal if and only if $q \geq c$. There are two sources of inefficiency in our model: Perceived quality may be above p even though true quality is below c so that principals are hiring when they should not (bad agents are working), or perceived quality is below p even though true quality is above c so that principals are not hiring when they should (good agents are not working).

Although in general the working probability may not have an analytic expression, and so the social surplus cannot be directly computed, we can nonetheless still prove results about welfare. We show a perhaps surprising result: For all information processes that are admissible, the *ex ante* social welfare will be less than if no information were generated at all. That is, the information provided through the reputational mechanism can actually improve social welfare.

We define a *blind process* to be any information process such that $R_q(t) = R_{q'}(t), \forall q, q' \in D$. Under a blind process, the information provided to the market is completely uninformative of the agent's true quality, and so no learning occurs. At all times t, the agent's work history reputation is thus the same as the initial prior, $f_t(q|\mathcal{H}_t) = f_0(q)$.

Theorem 1 (Blind Boundedness) *Let $\mathcal{L}^b$ denote any blind process. If $\mathcal{L}$ is an admissible information process, then $W(p,\mathcal{L}) \leq W\left(p,\mathcal{L}^b\right)$ for all $p \in \mathbb{R}$.*

Proof See Appendix. □

The intuition for this theorem is that since the market is forcing out agents when their expected quality hits p, the social welfare impact must be weakly negative for any $p \geq c$ because beliefs are correct, and so the expected quality of an agent that leaves *must* be higher than the social cost of that agent. If, at the time the agent is getting kicked out, we could instead choose to let that agent stay in and work forever, we would wish to do so. Then since under the blind process no agent ever gets kicked out, the blind process must be weakly better than any other admissible information process. So all admissible information processes give a social welfare that is bounded above by the blind process.

It is important to emphasize that the principals are *short lived*, which creates a source of inefficiency in the marketplace because principals do not internalize the social benefits of experimentation. Once the agent's expected quality level reaches the price, it would still be socially beneficial for the agent to work, because working sends information to the market and this information is valuable. Unlike a short-lived principal, a long-lived principal that could make future purchases would wish to continue hiring the agent because there is a chance the agent's reputation could improve and thus a positive value from experimentation.

In addition, from the proof we can see that the social welfare for any information process is the same as for a blind process if $p = c$, but can be strictly less if $p > c$. This implies that for *any* admissible information process, the socially optimal price is equal to the agent's reservation value. When price is exactly equal to the reservation value, there is equal social welfare from letting agents with $\mu_t = p$ stay and work forever, or leave immediately. But when $p > c$ it is strictly better to have agents that are being kicked out stay in because they are still generating positive flow payoffs in expectation. Thus, it is socially optimal to set the price as low as possible in order to make the exiting process more efficient. Note that if it was possible to subsidize agents to work below their reservation values, this would increase social welfare because an agent with expected quality equal to c still generates positive value to society through the information it provides by working.

5 General model

5.1 Belief updating with certification

Having shown the inefficiency of information revelation when certification is not possible, we now turn to the general model with both the work history and certification. If certification is available, agents will have a decision to make regarding the exact time to undergo certification. For an agent whose true quality is above $\underline{q}$ and has not undergone certification before time t, its expected discounted value if it chooses certification at time t will be $\int_t^{\infty} e^{-\rho(t'-t)} E\left[\pi_{t'} | \theta_t = t, \mathcal{H}_t, q\right] dt'$. We can analyze $\pi_{t'}$ by solving the principals' problem. Principals update the agent's reputation with Bayes' rule, and they will hire an uncertified agent if the expected quality level is above the price given the certification status and history of outputs, or $p < E[q_t | \theta_t = \phi, \mathcal{H}_t]$. Recall that they will hire an agent that certified at time t' if $p \leq E[q_t | \theta_t = t', \mathcal{H}_t]$, so principals will hire certified agents when indifferent but not uncertified agents.

The introduction of certification will result in posterior beliefs that are truncated distributions, even for agents that do not get certified. Suppose for instance that in equilibrium all types of agents with quality at least $\underline{q}$ are expected to certify at time 0. Given this fact, an agent that becomes certified at time 0 will be believed to have a quality higher than $\underline{q}$. If the prior distribution is $f_0(q)$, then the posterior distribution given certification will be $f_0(q)$ truncated to be over the interval $[\underline{q}, \infty)$. Since all types of agents in this quality region choose to certify, the relative density inside the support is unchanged. Likewise, the reputation for an agent that does not certify will also be updated. The market will believe that the agent has a quality less than $\underline{q}$, and so the posterior becomes $f_0(q)$ truncated to be over the region $(-\infty, \underline{q})$. An agent that certifies will continue to work forever (since we have assumed $\underline{q} \geq p$), and an agent that does not certify will stop working if the truncated mean drops below the price.

The key with certification, and the reason that it can improve welfare, is that high-quality agents who certify will be able to keep working even if they send an unlucky string of bad signals. Without certification such agents would instead be inefficiently kicked out of the market. In this way, certification is "freeing" the reputation of good agents which could otherwise get "stuck." Another way to look at it is from the perspective of the agents that are getting kicked out because they do not certify. Since their posterior gets truncated downwards, the expected mean of such agents can fall below the price level. Thus, unlike the benchmark case where agents got kicked out when $\mu_t = p \geq c$, here we can kick out agents with an expected mean less than c, and so kicking out these agents is socially efficient. This aspect of certification is welfare improving, but since certification also has a cost, welfare may decrease as well if certification is not implemented correctly.

5.2 Agent's certification strategy

We now analyze the agent's optimal certification stopping time strategy $\tau(q)$. In general this will also depend on what the market's beliefs $\tilde{\tau}(q)$ are, but we will show that for any beliefs the agent's strategy will have a common structure.

Throughout this section it is important to keep in mind two assumptions that we are making. For convenience, we assume that the certification standard $\underline{q} \geq p$, which implies that the agent will never stop working if it certifies. Using this assumption reduces the set of equilibria allowing for greater tractability, and implies that once certified the principals no longer need to doubt the agent's quality. With this assumption, the payoff after certifying is deterministic and given by $\frac{p-c}{\rho} - k$. For technical reasons, we assume that even if the agent's reputation had zero density on $[\underline{q}, \infty)$, and it passes certification, then the market will still be willing to hire the agent forever. This assumption matters only off the equilibrium path, where it ensures that the certification incentives for an agent remain the same. The second important assumption is that $k < \frac{p-c}{\rho}$, so that the certification cost is low enough for the agent to be able to benefit from certification. This can also be thought of as a restriction on the price, specifically that $p > k\rho + c$. This condition then means that the price is high enough to cover the cost of certification. Thus, an agent that is being forced out of the market would choose to certify and continue working instead of exiting forever. Without this assumption, the only possible equilibria would involve no certification.

We show in the following theorem that under our assumptions the only possible equilibria are *type independent equilibria*, where after any given work history either all types of agents with quality above $\underline{q}$ will choose to certify, or none will. Thus, in equilibrium all types of agents will use the same stopping time strategy τ^* and get certified after the same work histories.[8]

Theorem 2 (Type Independence) *In each equilibrium ε, all types of agents with quality $q \geq \underline{q}$ have the same certification stopping time strategy τ^*_ε.*

This type independence result arises from the fact that all agents who are able to certify face the same costs and benefits from certification. Note that since reputation depends only on the observed history, fixing a work history also fixes the agent's reputation independently of its true quality for any (not necessarily type independent) market strategy beliefs $\tilde{\tau}(q)$. Given the agent's reputation, either it can still work without having to certify, or it must certify to work. In the first case, since agents do not receive a higher flow payoff once certified, they would like to delay certification in

[8] Recall that the agent's certification strategy τ^* represents a certification time given the work history. Therefore, saying that all agent's use the same certification strategy is equivalent to saying that they certify after the same work histories.

order to delay the certification cost. Thus, no types of agents will choose to certify, and instead wait until their reputation falls lower. In the second case, since all agents prefer to certify rather than remain unemployed (given our assumption that the certification cost is sufficiently low), all agents would choose to certify regardless of their true quality.

This result does not mean that an agent's quality is irrelevant. For instance, agents with higher qualities may be more likely to send better signals and have better work histories. Since their reputations are more likely to be higher *ex ante*, they will be less likely to need to certify. Thus, the *ex ante* probability that an agent certifies at a given time t will in general depend heavily on its true quality. The type independence result says instead that *given* a history of signals, agents of all types will make the same certification decision. But the probability distribution over the history of signals will vary depending on agent quality.

Due to our type independence result, we will maintain the following notation in this paper: For any time t, let μ_t^{NC} represent the mean of the non-certifying agents and μ_t^C represent the mean of certifying agents, if the market were to believe that all types of agents with quality $q \geq \underline{q}$ would choose to certify at that t. Formally, $\mu_t^{NC} \equiv E\left[q|\mathcal{H}_t, q < \underline{q}\right]$ and $\mu_t^C \equiv E\left[q|\mathcal{H}_t, q \geq \underline{q}\right]$. These two expectations will be helpful for characterizing the equilibria of our model. We also keep with our previous notation by letting $\mu_t \equiv E_t\left[q|\mathcal{H}_t\right]$ be the agent's expected quality using only information from the work history. This is equal to the agent's expected quality given that it has not certified, $E_t\left[q|\theta_t = \phi, \mathcal{H}_t\right]$, if the market were to believe that no agents would certify before time t (i.e., $\tilde{\tau}^* > t$).

Now that we have shown the equilibria are all type independent, we characterize the possible equilibrium optimal stopping time strategies τ^*. The corollary below shows that the agent's equilibrium certification time is related to the agent's work history and reputation through the values of μ_t^{NC} and μ_t. If $\mu_t^{NC} \leq p$, then agent's that do not certify if the market expects all types of agents to certify cannot keep working, and so all types of agents will choose to certify with these market strategy beliefs. Further, if $\mu_t \leq p$ then no type of agent can keep working without certifying regardless of market strategy beliefs, and so all types of agents that can certify will certify for any market strategy beliefs.

Corollary *Fix any time t and history $\mathcal{H}_t$ such that $\mu_{t'} > p$ for all $t' < t$. Suppose that τ^* is an equilibrium stopping time strategy such that $\tau^* \geq t$ given this history. Then it is possible for $\tau^* = t$ if and only if $\mu_t^{NC} \leq p$. Further, it must be the case that $\tau^* = t$ if and only if $\mu_t = p$.*

Proof This follows from the Proof of Theorem 2. □

Note that there are some histories in which no types of agents certifying or all types of agents certifying are both supportable in equilibrium. The above corollary implies that as long as $\mu_t > p$ no agents certifying is supportable in equilibrium, and as long as $\mu_t^{NC} \leq p$, all types of agents certifying is supportable in equilibrium. The first condition indicates that if no agents are believed by the market to certify, all types of agents can keep working without certifying. The second condition indicates that if all types of agents with qualities in $\left[\underline{q}, \infty\right)$ are believed to certify, then agents that do not certify will be forced to stop working. As long as $\mu_t < p \leq \mu_t^{NC}$, we could have either type of behavior in equilibrium at time t. Figure 2 below highlights the range of possible equilibrium if the work histories follow Brownian motions.[9]

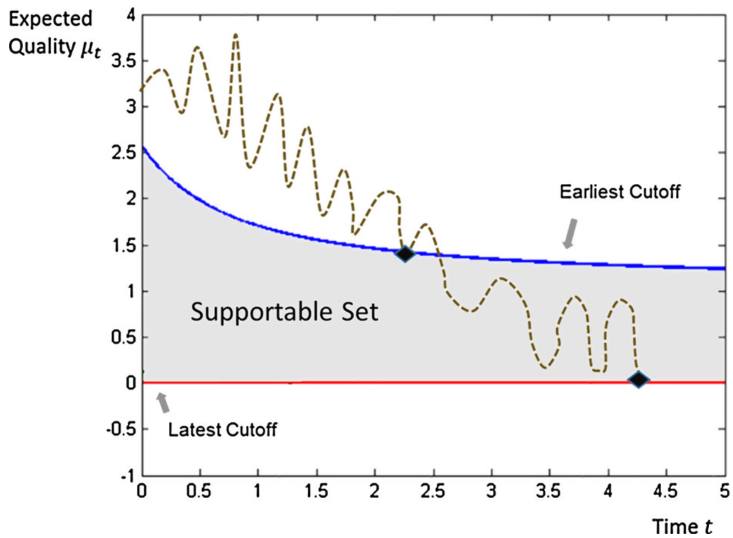

Fig. 2 *Range of possible equilibria* The agent starts with an expected quality above the price of $p = 0$, and continues getting hired without certifying until its expected quality falls into the shaded region. Within this shaded region, if the market expects the agent to certify, then the agent will be forced to certify or exit. But if the market expects the agent not to certify, then the agent will choose not to certify. The higher curve represents the earliest supportable certification time for the agent and the lower curve represents the latest time. Above the higher curve, certification is not supportable in equilibrium, and at or below the lower curve, certification must occur

5.3 Equilibrium characterization

As we have shown, and as is common in signaling games, many different equilibria are possible. Thus, we narrow down the set of equilibria by focusing on the specific equilibria that give the highest payoffs to the different players in our model. In order to find out which equilibrium maximizes principal, agent, or social surplus, we find the specific equilibrium beliefs and strategies that will result in the highest payoffs to each party. Theorem 2 implies that every equilibrium features a single stopping time strategy used by all types of agents. We denote the agent optimal, socially optimal, and principal optimal equilibrium stopping times by τ_a^*, τ_s^*, and τ_p^*, respectively. The next section characterizes the optimal equilibria for these three different cases. We show that the different types of optimal equilibria can be ranked according to the

[9] With Brownian motion signals and a normal prior, the posterior will always be normal and so the expected mean μ_t will be strictly increasing in the truncated mean μ_t^{NC}. Thus, we can transform the condition $\mu_t \leq p \leq \mu_t^{NC}$ into just a condition on p and μ_t, allowing us to depict the stopping times explicitly via cutoffs for the expected mean as in Fig. 2.

equilibrium certification times. Theorem 3 summarizes our main findings: We see that principals always prefer earlier certification than is socially optimal and that agents always prefer later certification than is socially optimal.

Theorem 3 *The optimal stopping times for the principal, agent, and social welfare cases must satisfy*

$$\tau_p^* \leq \tau_s^* \leq \tau_a^*$$

Proof The result follows from a comparison of the stopping times in Propositions 1, 2, and 3 below. □

5.3.1 Agent optimal equilibrium

We start with characterizing the best equilibrium from the agent's point of view. Since the agent is long lived, the notion of optimality we will use is the one that maximizes the agent's expected value at time 0, given by $\int_0^\infty e^{-\rho t}(p-c)\,dt - E_{\tau_a^*}\left[e^{-\rho\tau_a^*}k\right]$ if $q \geq \underline{q}$ or $E_{\tau_a^*}\left[\int_0^\infty e^{-\rho t}(p-c)\,dt\right]$ if $q < \underline{q}$. In general, the agent's expected discounted value will depend on its own quality level, but we show that the specific equilibrium that maximizes the agent's value does not.

Proposition 1 *The equilibrium that maximizes the agent's payoff is for* $\tau_a^* = inf\{t|\mu_t \leq p\}$.

Note that τ_a^* is the absolute latest that the agent can certify in equilibrium because at this point the agent must stop working no matter what the market certification beliefs are. This proposition thus shows that the best equilibrium from the agent's perspective will always delay certification for as long as possible, with certification occurring just as it would be forced to exit the market. Since agents bear the full cost of certification and do not directly benefit in terms of flow payoffs, they are much less willing to certify than principals would want them to. In this equilibrium the beliefs expect certification at the latest possible time, and so we will call them *optimistic* beliefs.

If $\mu_o \leq p$ then agents cannot work initially and must certify to get hired. In this case certification act as a *de facto license*, with non-certifying agents never being able to work. Certification thus functions as an entrance key into the market itself. But as long as $\mu_0 > p$, in any agent optimal equilibrium certification will not be a de facto license. Instead the equilibrium will feature *delayed certification*, with agents working and sending information initially before needing to certify if their reputation drops too low. In this case the particular information process will make a big difference on the timing of the agent's certification decision, as well as the overall social welfare that is generated by certification.

5.3.2 Social welfare optimal equilibrium

Next, we characterize the equilibrium beliefs that maximize the total social welfare for the market. We will use the same notion of *ex ante* social welfare that was defined in Sect. 4, meaning that we want to maximize the working probability of good agents while minimizing the working probability of bad agents. The formula is given by:

$$W = \int_{\underline{q}}^{\infty} \left[E \left[\int_{0}^{\infty} e^{-\rho t} (q - c)\, dt - e^{-\rho \tau_s^*} k | q \right] \right] f_0(q) dq$$
$$+ \int_{-\infty}^{\underline{q}} \left[E \left[\int_{0}^{\tau_s^*} e^{-\rho t} (q - c)\, dt | q \right] \right] f_0(q) dq$$

Early certification is beneficial because it verifies that an agent's quality is high more quickly. However, since certification is also costly, we may want agents to delay certification if their reputation is sufficiently high, since that saves the discounted flow cost of certification, as well as keeps the option value of certifying for the future. The following proposition characterizes a *necessary condition* for certification to be socially optimal: The agent should certify only if its reputation is lower than the following cutoff.

Proposition 2 *At the socially optimal stopping time* τ_s^* *, if the expected quality of the agent is strictly higher than the price,* $\mu_{\tau_s^*} > p$ *, then the following condition must also be satisfied:*

$$\mu_{\tau_s^*}^{NC} \leq c + \rho k - \frac{\rho k}{F_{\tau_s^*}^{-}\left(\underline{q}\right)} \tag{1}$$

We note that this bound can also be a sufficient condition for some information processes. With a blind process (1) gives the exact value of the truncated mean below which it becomes optimal to have the agent certify. Certification is socially beneficial because it gets rid of "bad" agents with qualities less than c, but it is socially harmful because it can get rid of "good" agents with qualities above c but below $\underline{q}$ and it also carries a cost of k. Certification at any time t must balance these two aspects together, and (1) gives a condition on how low the mean of non-certifying agents (i.e., the agents that would be kicked out by certification) must be for certification to be beneficial.

However, it is possible for (1) not to be satisfied at any $t \leq t^* = \inf\{t | \mu_t \leq p\}$. In this case we would have $\tau_s^* = t^*$, because when the expected quality falls below the price, the unique equilibrium is for an agent to certify. Consequently, the agent optimal stopping time and the socially optimal stopping time would coincide. In general, however, the socially optimal stopping time can be strictly earlier than the agent optimal stopping time due to the beneficial information that certification provides.

5.3.3 Principal optimal equilibrium

Finally, we analyze the equilibrium that maximizes the utility of the principals. The right notion of maximizing principal utility is a bit tricky since each principal is short lived and only cares about itself. The equilibrium that gives the time t principal the highest payoff will in general be different than the equilibrium that gives the time t' principal the highest payoff. One way to define principal welfare would be to aggregate all principal utilities *ex ante*, similar to the way we computed the social welfare

$$PW = \int_{\underline{q}}^{\infty} E\left[\int_0^{\infty} e^{-\rho t}\,(q-p)\,dt|q\right] f_0(q)dq$$
$$+\int_{-\infty}^{\underline{q}} E\left[\int_0^{\tau_s^*} e^{-\rho t}\,(q-p)\,dt|q\right] f_0(q)dq$$

For such a definition of principal welfare, the result of Theorem 3 still holds. However, such coordination among principals may seem unrealistic given that all the principals are short lived. We resolve this tension in the following way: Principals that arrive earlier should have a greater say in the equilibrium than principals who arrive later, because presumably the later principals may not be able to interact with or even know of the agent at earlier times. Thus, we define a *principal optimal equilibrium* to be the equilibrium in which the agent certifies at the first (stochastic) time t such that given the agent's reputation, the principal at t prefers hiring the agent only if it could certify over always hiring the agent.[10] Formally, the agent certifies at the first t such that $\left(\mu_t^C - p\right)\left(1 - F_t^-\left(\underline{q}\right)\right) \geq \mu_t - p$. We show in the proof of the following proposition that such an equilibrium always exists and is unique. In this equilibrium, the earlier principals have more influence over the agent, with later principals having an influence only if the agent has not certified early on. This notion of principal optimality is not the same as the net present value of principal surplus defined above, since the myopic time t principal does not consider the welfare of future principals.

Proposition 3 *The principal optimal equilibrium stopping time is*

$$\tau_p^* = inf\{t|\mu_t^{NC} \leq p\} \tag{2}$$

This Proposition gives a cutoff in terms of μ_t^{NC}, with principals wanting the agent to certify the first time its truncated expected quality falls below the price. Certification is good for principals because it gets rid of "bad" agents with qualities less than p, but it may also hurt principals because it can get rid of "good" agents with qualities between p and $\underline{q}$. The cutoff $\mu_t^{NC} = p$ is the quality such that these two effects exactly balance out. We see from (2) that the best equilibria for principals are those where the certification strategy beliefs τ^* expect agents to certify at the highest possible truncated quality that is still an equilibrium, since $\mu_t^{NC} \leq p$ is necessary for certification to occur. Since this is the highest possible supportable value, we call such types of beliefs *pessimistic beliefs*. Because principals do not bear the cost of certification, they want agents to certify at the fastest possible time that is supportable in an equilibrium. We will denote the principal optimal equilibrium stopping time by τ_p^*.

Note that if $\underline{q} = p$ then $\mu_t^{NC} \leq p$ for all t, and so principals will want certification by all types of agents immediately. In such a case, certification acts as a de facto license because any agent that does not certify will be immediately excluded from the market and unable to work at all. With a de facto license, the actual information process being

[10] If $\underline{q} = p$, then principals would always wish to hire the agent only if it could certify. However, if $\underline{q} > p$, and if the principal believes it is very likely that the agent's quality falls between $(p, \underline{q})$, the principal may prefer to hire all types of agents instead of hiring only when the agent's quality was above $\underline{q}$.

used by the market is irrelevant since agents certify before any learning takes place, and they then stay in the market forever regardless of the signals that are sent.

However, for $\underline{q} > p$ it could be possible that $\mu_0^{NC} > p$, and so principals may not want agents to certify immediately. In this case, if the agent's initial reputation is high enough, then the principals will want the agent to work for a while and only certify if its reputation drops too low, with the specific threshold given by Eq. (2). Thus, the equilibrium will be a delayed certification equilibrium. In this type of equilibrium, the information process itself matters because it affects how quickly the agent's reputation changes.

Putting together all our results for the optimal certification times from Theorem 3 allows us to show on the graph below what the optimal stopping time thresholds are for the special case of Brownian motion reports. Figure 3 below highlights the different optimal certification thresholds, which assume the same type of Brownian motion as in Fig. 2. The principal optimal threshold is the highest supportable threshold, the agent optimal threshold is the lowest supportable threshold, and the socially optimal threshold is in between.

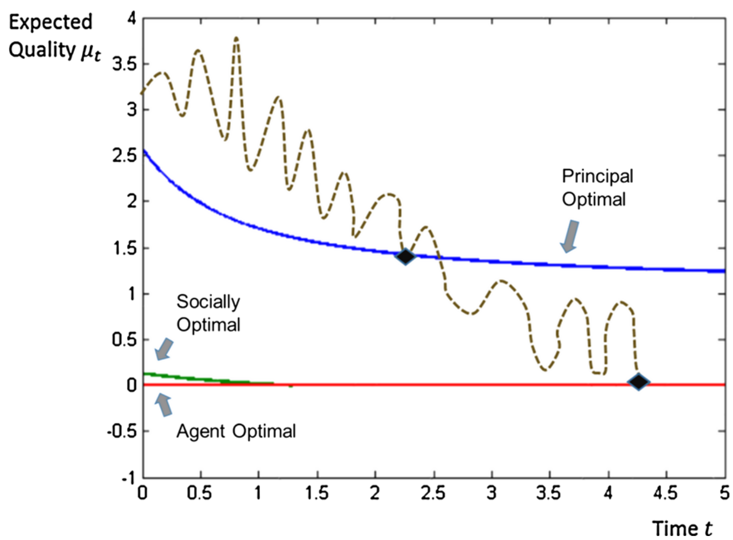

Fig. 3 *Optimal certification cutoffs*. This figure shows several thresholds that represent optimal certification cutoff qualities for principals and agents, as well as the socially optimal cutoff quality. The principal and agent thresholds represent optimal stopping rules: The agent should certify the first time that its expected quality hits each threshold. The socially optimal threshold represents a necessary condition: The agent should certify only if its expected quality is below the threshold. The principal optimal threshold lies far above the other two, whereas the socially optimal and agent optimal thresholds are close together and coincide after a while. The first point shows the optimal certification time from the principal's perspective, and the second point shows the optimal certification time from the agent's perspective, as well as the socially optimal certification time

6 Optimal certification standards

6.1 Certification standards for socially optimal beliefs

Now that we have characterized what can happen in the various types of equilibrium, we will perform some comparative statics in order to analyze the socially optimal certification standards and prices. This analysis will tell us how certification should be implemented depending on what type of equilibrium is expected. Thus, we are now analyzing a mechanism design framework, where the designer chooses the $\underline{q}$ and p in order to maximize social welfare. In doing so we maintain our assumption that the certification standard must be set higher than the price, $\underline{q} \geq p$. However, we no longer require that $\frac{p-c}{\rho} < k$, so it is possible to set the price sufficiently low that agents would never wish to get certified. Note also that the price can be chosen such that $\frac{p-c}{\rho} = k$, in which case the agent is indifferent between certifying and leaving the market[11]. Since there are many possible equilibrium, and the optimal standard and price will differ across equilibrium, we focus specifically on the socially optimal, agent optimal, and principal optimal equilibria that we analyzed above.

We start by analyzing socially optimal equilibrium because it is the most straightforward case. We assume that for any choice of $p, \underline{q}$ by the designer, the market will hold the socially optimal beliefs that were discussed in the previous section. Since these beliefs are part of an equilibrium, the agent will play according to these beliefs. With this restriction on the resulting game equilibrium, we find the values of $p, \underline{q}$ that maximize *ex ante* social welfare.

Although finding the socially optimal equilibrium stopping time strategy is challenging in general, characterizing how the price and standard should be set is much simpler. The reason is that, unlike in the principal and agent optimal cases, the beliefs will always be chosen in a socially optimal fashion, and we do not need to alter the price and standard in order to counteract possible "bad beliefs" that expect certification at inefficient reputation levels. Thus, it is socially optimal to set the standard and price as low as possible (but still above c) to allow as many good firms to certify as necessary. Let $p^* = c + k\rho$ and $\underline{q}^* = p$, and let $W^*(k, p^*, \underline{q}^*, \mathcal{L})$ be the corresponding *ex ante* social welfare given this certification standard and price assuming a socially optimal equilibrium. This corresponds to the lowest possible price that still allows for certification, and the lowest possible standard given this price. The next theorem states that the price and standard should be set at these lowest possible levels as long as the resulting social welfare is higher than that without certification.

Theorem 4 *Assuming that a socially optimal equilibrium will occur, the price and certification standard that maximize ex ante social welfare are given by:*

1. $p = c + k\rho$ *and* $\underline{q} = p$ *if* $W^* \geq \frac{\mu_0 - c}{\rho}$.
2. $p = c$ *and* $\underline{q} = \infty$ *if* $W^* \leq \frac{\mu_0 - c}{\rho}$.

[11] When this equation holds with equality, there may now be equilibria that are type dependent, because agents are indifferent between certifying and exiting the market, so some types could choose to certify and other types could choose to exit. But all the type independent equilibria characterized by the corollary to theorem 2 will still exist as well, and they will generate higher social welfare than the corresponding type dependent equilibria. This is because it is socially optimal for an agent with quality $q \geq \underline{q}$ to certify instead of exiting the market, so the type dependent equilibria that feature some types exiting will be suboptimal.

6.2 Socially optimal certification standards for pessimistic beliefs

We now characterize the social welfare maximizing certification standards and prices for the pessimistic beliefs defined in Sect. 5.3.3, which result in a principal optimal equilibrium. Unlike with the socially optimal beliefs, these beliefs are not aimed at maximizing the social welfare and so we may need to manipulate the standards and prices in order to increase welfare. We analyze the comparative statics of our model for the parameters $\underline{q}$ and p, assuming that given any choice of the parameters a principal optimal equilibrium will occur. We present a perhaps surprising result: Certification should only be implemented in a principal optimal equilibrium as a de facto license. In this equilibrium, the principals should never get to learn from an agent's work history at all before certification occurs.

The next theorem tells us exactly what the optimal certification standards and prices are as a function of the certification cost. When the certification cost is sufficiently high, certification should not be allowed, and so we end up with our benchmark model. When the certification cost is low, certification should be allowed and implemented as a de facto license where agents certify immediately. In no situation is it optimal to allow delayed certification, where agents do not certify immediately, but only after working and sending some signals.

Theorem 5 *Assuming a principal optimal equilibrium will occur, the price and certification standard that maximize ex ante social welfare are given by:*

1. $p = c + k\rho$ *and* $\underline{q} = p$ *if* $k \leq \frac{\mu_0^c - c}{\rho} - \frac{\mu_0 - c}{\rho(1 - F_0^-(c + \rho k))}$.
2. $p = c$ *and* $\underline{q} = \infty$ *if* $k \geq \frac{\mu_0^c - c}{\rho} - \frac{\mu_0 - c}{\rho(1 - F_0^-(c + \rho k))}$.

When $\underline{q} = p$, principals will want certification immediately, and when $\underline{q} = \infty$ no agents will ever be able to certify. Thus, there is no case where agents are allowed to certify but not forced to certify immediately. The reason why delayed certification is never optimal is that if it were implemented the pessimistic beliefs would require certification by agents very early and often, and in fact so early and often that certification destroys social welfare. To understand this intuitively, note that under pessimistic beliefs the mean of a non-certifying agent will be exactly equal to the price level. Thus, we have the same result as in the benchmark, where we are always kicking out agents who have expected qualities above c. And since $p > c$ in order to allow for certification, kicking out these agents will strictly reduce social welfare. Even worse, high-quality agents are also forced to pay for the cost of certification, destroying additional social value. Since the myopic principals would want the agent to certify at inefficient reputation levels, the value of certification gets entirely undercut.

Because certification should act as a de facto license under these types of beliefs, learning through the work history will be irrelevant and does not affect the social welfare. Agents that do not certify can never work, and agents that do certify never exit, so it does not matter what type of information the work history is sending. For any information process, the ex ante social welfare will be the same. When certification costs are low enough, de facto licenses will provide a higher social welfare than the no certification benchmark and thus should be implemented. In fact, for $k = 0$ we can achieve the first best outcome through setting $\underline{q} = p = c$, which results in all socially beneficial agents working forever, and all other agents exiting immediately. And as $k \to 0$, we can asymptotically achieve first best by setting the price and certification standard lower and lower. However, if certification costs are too high, then too much social welfare is lost through certification itself, and it is thus better to not allow for certification.

6.3 Socially optimal certification standards for optimistic beliefs

Now we will analyze the agent optimal equilibrium that features the optimistic beliefs mentioned earlier in Sect. 5.3.1. Since optimistic beliefs never result in de facto licenses if $\mu_0 > p$, the social welfare generated by these beliefs can be heavily dependent on the specific information process that the market has access to. In order for certification to occur, the price needs to be high enough to compensate for the cost of certification, so we need $p \geq c + k\rho$ or the result is no certification. Given that agents with qualities higher than $c+k\rho$ contribute positively by certifying instead of exiting, the certification standard $\underline{q}$ should be set equal to p no matter what the optimal p is. The inequality $p \geq c + k\rho$ need not be binding, because it may be optimal for the price to be raised so that certification occurs sooner. For instance, if the certification cost k is very low, then it is optimal to have p higher so that we can get certification earlier.

Also note that a de facto license can be implemented if the price is set at $p \geq \mu_0$, since then principals would never buy from non-certifying agents. The optimal way to implement a de facto license is to set $p = \underline{q} = \max(c + \rho k, \mu_0)$, for similar reasons as in the principal optimal case. However, a de facto license would result in all types of agents with qualities less than μ_0 but greater than c getting forced out of the market. Thus, it may be better not to have a de facto license, but delayed certification instead.

In general then, for k sufficiently low it will be optimal to have $\underline{q} = p \in [c+\rho k, \mu_0]$, where the exact value depends on the information process itself. This results in either delayed certification if $p < \mu_0$ or a de facto license if $p \geq \mu_0$. If k is too high, then we should set $p = c$ and $\underline{q} = \infty$ to ensure that no certification occurs. The exact cutoff value for k, as well as the exact values of the price and certification standard, will depend on the information process because that will influence the social welfare of any delayed certification scheme.

Given that delayed certification can be desirable in an agent optimal equilibrium, one may think that the social welfare under optimistic beliefs will always be higher than the social welfare under pessimistic beliefs, where delayed certification could never be optimal. However, optimistic beliefs also have a downside – de facto licenses are now much harder to implement because the price needs to be set at a very high level to do so. Under pessimistic beliefs, we could asymptotically get first best as $k \to 0$ by implementing a de facto license. But that is not the case with optimistic beliefs, as the next theorem shows. With optimistic beliefs, a de facto license cannot achieve first best, and delayed certification cannot achieve first best either for any positive k[12]. Given any value of k, we define $W_a(k, \mathcal{L})$ as the optimal *ex ante* social welfare in the

[12] When $k = 0$, agents do not mind certifying at an expected quality above the price, and so first best can be achieved in an agent optimal equilibrium.

agent optimal equilibrium. That is, this is the highest social welfare that can result in an agent optimal equilibrium from any choice of p and $\underline{q}$.

Theorem 6 *Suppose that the initial quality distribution $f_0(q)$ has positive density on the interval $[c-\epsilon, c+\epsilon]$ for some $\epsilon > 0$. Then for all $k > 0$, the value of the ex ante social welfare in an agent optimal equilibrium is strictly bounded away from the first best value,* $\sup_{k>0} W_a(k, \mathcal{L}) < \int_c^\infty \frac{q-c}{\rho} f_0(q)dq$.

Proof See Appendix. □

This theorem shows that given any information process, the *ex ante* social welfare is strictly bounded away from the optimal as costs become low, as long as it is initially possible for the agent to have a quality close to c. However, we note that if the information process itself become very informative, then the optimal social welfare can be asymptotically achieved. Define a *fully revealing* information process to be such that almost surely $\tau_a^* < \infty$ for all types of agents with $q < c$. This means that bad agents will for sure be forced to certify at some point, and thus will always get kicked out of the market in the long run. Such a property holds as long as the information process allows bad agents to be distinguished from good agents over time. For fully revealing information processes, faster information revelation will allow for the first best social welfare to be asymptotically achieved when certification costs are low. Formally, we say that an information process $\mathcal{L}'$ is *faster* than another process $\mathcal{L}$ if for some $n > 1$, $R_q'(t) = R_q(nt)$ for all t, q. The parameter n tells us how much faster one process is at sending information than another. For some information processes, faster information has a natural interpretation. For example, with the Brownian motion $dR_q(t) = qdt + \sigma_r dZ(t)$, a faster information process would correspond to dividing the variance of the Brownian motion by n. For Poisson processes, faster information would correspond to multiplying the arrival rates by n.

The reason that faster information revelation is helpful is with certification is because low-quality agents will get kicked out of the market very quickly, while high-quality agents can still stay in.[13] As the following proposition shows, first best can be asymptotically achieved if the certification costs become very low and the information arrives at a very fast speed. Thus, information revelation and certification can act in a complementary fashion, with more information increasing the welfare provided by certification.

Proposition 4 *Suppose that $\mathcal{L}$ is a fully revealing information process, and let $\{\mathcal{L}_n'\}$ be a sequence of information processes that are faster than $\mathcal{L}$ by the factor n. Then we have*

$$\lim_{n\to\infty} \sup_{k>0} W_a\left(k, p, \underline{q}, \mathcal{L}_n'\right) = \int_c^\infty \frac{q-c}{\rho} f_0(q)dq$$

Proof See Appendix. □

[13] Note that without certification, faster information cannot increase welfare as shown by Theorem 1.

On the other hand, if certification costs are high and a delayed certification scheme were attempted, faster information revelation could actually lower social welfare. When costs are high, the price and certification standard may need to be much higher than c to get agents to certify. But setting high certification standards will make certification kick out agents less efficiently than before because exiting agents will have higher quality levels. A faster information process compounds this problem by getting to the kick out point faster, thus lowering social welfare. In this case, it may be better not to implement certification.

7 Conclusion

This paper analyzed how two separate avenues of information, certification and work history, can interact to affect learning about the quality of an agent. We showed that the information provided by the work history alone cannot raise social surplus, making certification necessary even when the market can learn through the agent's work history. With certification, all equilibria will feature type independent certification strategies, and the various equilibria that maximize principal, agent, and social welfare can be ordered according to the equilibrium certification stopping time strategies. Principals will prefer the agent to certify at the earliest possible time, and thus, delayed certification by the agent will not be socially efficient. With pessimistic beliefs, only a de facto license type of certification can be socially beneficial. On the other hand, in the agent optimal case, the agent will delay certification for as long as possible. This is suboptimal if the certification costs are low, and it is socially beneficial to have the agent certify quickly. In this case, faster information revelation can increase social welfare, creating a complementarity between the reputational forces and certification.

There are several possible extensions for future research that could have interesting implications. One important case is to allow for variable prices that depend partially or even wholly on the agent's reputation. Such a change complicates the exit decision for the agent, because it may wish to keep working even if the price falls below its own reservation value. In extreme cases such as Bar-Isaac (2003), higher-quality agents may never exit, which means that there is no loss in efficiency in the long run. Thus, certification would only be useful to flush bad agents out of the market sooner, and so certification would be less valuable than in the current model.

Another interesting extension is to allow for moral hazard. For instance, the agent could increase the quality of its work by exerting effort, which would also send more positive information to the market. The agent may thus choose greater effort to increase its reputation instead of certifying. And after the agent does certify, its work incentives may drop quite significantly. Principals would anticipate this and therefore prefer the agent to wait longer instead of certifying quickly. So with moral hazard, delayed certification could now be socially beneficial in a principal optimal equilibrium. Likewise, certification times in socially optimal equilibrium may get delayed as well. Certification may thus become less beneficial with moral hazard, because it undercuts an agent's incentive to signal through effort.

8 Appendix

8.1 Proof of Theorem 1

First note that if the price is less than c, the agent would refuse to accept any offers, and so the social welfare is equal to 0 regardless of the information process. Thus, we only need to consider prices $p \geq c$.

Note that under any blind process, the agent's expected quality is never updated, and so since we assume that $\mu_o > p$ the agent will never stop working. The *ex ante* social surplus can thus be calculated as $\frac{\mu_0 - c}{\rho}$. For a general information process the market continues to hire any agent until its expected quality drops below p, and for an admissible process this happens at the first time that $\mu_t = p$. We can write out the *ex ante* expected social welfare for any information process as

$$\begin{aligned} W(p, \mathcal{L}) &= \int_{-\infty}^{\infty} \left[\int_{0}^{\infty} e^{-\rho t}(q-c)\,dt \right] f_0(q)\,dq - E_{q,t^*}\left[\int_{t^*}^{\infty} e^{-\rho t}(q-c)\,dt \right] \\ &= \int_{-\infty}^{\infty} \left[\int_{0}^{\infty} e^{-\rho t}(q-c)dt \right] f_0(q)\,dq - E_{t^*}\left[\int_{t^*}^{\infty} E_q\left[e^{-\rho t}(q-c)|t^* \right] dt \right] \\ &= \frac{\mu_0 - c}{\rho} - \frac{p-c}{\rho} E_{t^*}\left[e^{-\rho t^*} \right] \leq \frac{\mu_0 - c}{\rho} \end{aligned}$$

□

8.2 Proof of Theorem 2

Suppose the market believes that agents are following some (not necessarily type independent) certification strategy $\tilde{\tau}(q)$. Consider any time t' and history $\mathcal{H}_{t'}$, and let the expected quality for an agent that has not certified by t' be given by $\mu_{t'}^N \equiv E[q|\mathcal{H}_{t'}, \theta_{t'} = \phi]$. Note that for any market strategy beliefs, the value of $\mu_{t'}^N$ depends only on the work history $\mathcal{H}_{t'}$ and certification status $\theta_{t'}$, and not the agent's true quality. Thus, fixing the work history and certification status, all agents will have the same $\mu_{t'}^N$ regardless of their true quality. There are two possible cases: $\mu_{t'}^N > p$ or $\mu_{t'}^N \leq p$, and we show that for both cases, either all types of agents with qualities $q \geq \underline{q}$ will choose to certify at t' or no types of agents will choose to certify at t'.

First consider the case where $\mu_{t'}^N > p$. Then agents can still work even without certifying. The payoff of an agent that chooses to certify is given by $\frac{p-c}{\rho} - k$. This payoff is identical for all agents with quality $q \geq \underline{q}$ because of our assumption that $\underline{q} \geq p$, and so certified agents will never stop working. Now consider the alternate strategy of waiting until the time $\hat{t} \equiv \inf\left\{t|\mu_t^N \leq p\right\}$ and then certifying. This strategy gives a payoff of

$$\int_0^{\hat{t}} e^{-\rho t}(p-c)\,dt - \left(e^{-\rho\hat{t}}\right)\left(\frac{p-c}{\rho} - k\right)$$

This alternate strategy gives a payoff higher than certifying immediately by $(1 - e^{-\rho\hat{t}}) * k$. So certifying at time t' is not optimal, and with this $\mu_{t'}^N$ all types of agents would choose not to certify.

Next consider the case where $\mu_{t'}^N \leq p$. An agent that does not certify will not be able to work at time t', and so receives a maximum payoff of

$$(1 - \rho dt) \left(\frac{p - c}{\rho} - k \right)$$

This is the payoff the agent would receive if it certified at time $t' + dt$. Since no observations are made, the mean at a later time cannot be greater than p unless the agent were to certify at some time $t' + dt$. If the agent instead chooses to certify immediately at time t', it would get a payoff of

$$\frac{p - c}{\rho} - k > (1 - \rho dt) \left(\frac{p - c}{\rho} - k \right)$$

Thus, certifying in the current time step would increase the payoff by $\rho dt \left(\frac{p-c}{\rho} - k \right)$. Therefore, all types of agents would choose to certify given this value of $\mu t'\ N$.

Since all agents with quality $q \geq \underline{q}$ would choose the same certification decision for any value of $\mu_{t'}^N$, every equilibrium must feature all agents with quality $q \geq \underline{q}$ utilizing the same certification strategy τ_ε^*. □

8.3 Proof of Proposition 1

Consider any equilibrium that requires the agent to certify at a time t' where $\mu_{t'} > p$. Now let us compare the payoffs to the agent against the equilibrium where the agent certifies at the first time t such that $\mu_t \leq p$. In the second equilibrium, at time t' the agent would instead delay certification until the time $t^* = \inf \{t | \mu_t = p\}$. Since the payoff from certification is the same regardless, the agent would be able to improve its payoff by the amount $(e^{-\rho t'} - e^{-\rho t^*})k$. Thus, no equilibrium that requires the agent to certify at a $\mu_t > p$ can be agent optimal. □

8.4 Proof of Proposition 2

We fix an arbitrary equilibrium and compute the social welfare flow payoff difference at the equilibrium certification time τ^* between certification and no certification. Without certification, the social welfare flow payoff is given by

$$E_{\tau^*} [q - c] = \int_{-\infty}^{\infty} (q - c)\ f_{\tau^*}(q) dq$$

With certification this would become

$$E_{\tau^*} [q-c] = \left(1 - F_{\tau^*}^- \left(\underline{q}\right)\right) \int_{\underline{q}}^{\infty} (q-c)\ f_{\tau^*}(q | q \geq \underline{q}) dq = \int_{\underline{q}}^{\infty} (q - c)\ f_{\tau^*}(q) dq$$

Thus, the difference in the two expectations is given by

$$\lim_{q' \nearrow \underline{q}} \int_{-\infty}^{q'} (q - c)\, f_{\tau^*}(q) dq$$

Certification has a higher flow payoff if and only if the difference in expectations plus the flow cost of certification is less than zero:

$$\lim_{q' \nearrow \underline{q}} \int_{-\infty}^{q'} (q - c)\, f_{\tau^*}(q) dq + \rho k \left(1 - F_{\tau^*}^{-}\left(\underline{q}\right)\right) \leq 0$$

Or equivalently $\left(\mu_{\tau^*}^{NC} - c - \rho k\right) F_{\tau^*}^{-}(\underline{q}) \leq -\rho k$, which leads to (1).

Now suppose for the sake of contradiction that (1) does not hold at the certification time τ^*. We will consider the following strategy that can be an equilibrium, and we show that it provides a higher social welfare than certifying at time τ^* if (1) does not hold at τ^*. Suppose instead that the agent keeps working until 1: the first $t > \tau^*$ such that (1) holds, or until 2: the first $t > \tau^*$ such that $\mu_t \leq p$, and then the agent certifies.

Note that the agent keeps working without certifying as long as the social welfare flow payoff from not certifying is greater than the flow payoff from certifying at τ^*. Once the agent certifies, the flow payoffs are the same as with certifying at τ^*. Thus, under this alternate strategy the flow payoffs can never be less than under certifying at τ^*. Therefore, the total social welfare generated must be higher as well. □

8.5 Proof of Proposition 3

The short run time t principal's utility is given by $\mu_t - p = E_t[q - p]$, so the principal prefers certification if and only if this expectation with certification is higher than the expectation without certification. Without certification, we have

$$E_t\left[q - p\right] = \int_{-\infty}^{\infty} (q - p)\, f_t(q) dq$$

With certification this would become

$$E_t\left[q - p\right] = \left(1 - F_t^{-}\left(\underline{q}\right)\right) \int_{\underline{q}}^{\infty} (q - p)\, f_t(q | q \geq \underline{q}) dq = \int_{\underline{q}}^{\infty} (q - p)\, f_t(q) dq$$

Thus, the difference in the two expectations is given by

$$\lim_{q' \nearrow \underline{q}} \int_{-\infty}^{q'} (q - p)\, f_t(q) dq$$

Certification is preferred if and only if the above term is less than zero.

$$\int_{-\infty}^{p} (q - p)\, f_t(q) dq + \lim_{q' \nearrow \underline{q}} \int_{p}^{q'} (q - p)\, f_t(q) dq \leq 0$$

This means that the benefit of removing bad agents (qualities below p) from the market outweighs the costs of removing the good agents (qualities above p). Or equivalently:

$$\left(\mu_t^{NC} - p\right) F_t^{-}(\underline{q}) \leq 0$$
$$\mu_t^{NC} \leq p$$

This results in Eq. (2) in the Proposition. Since at each time t, that time t principal wishes for the agent to certify if and only if this equation holds, the resulting certification strategy will feature the agent certifying at the first moment that this equation holds. By the corollary to theorem 2, we know that such a certification strategy can be an equilibrium. □

8.6 Proof of Theorem 4

First note that if $p < c + \rho k$ there can be no certification in equilibrium. The optimal way to implement no certification is to set $p = c$, $\underline{q} = \infty$ from Theorem 1. This results in a social welfare of $\frac{\mu_0 - c}{\rho}$ for any admissible information process. Next, suppose that we wish to allow certification in equilibrium. Thus, we need to set $p \geq c + \rho k$. We show that for any $p \geq c + \rho k$ and any $\underline{q} \geq p$, we can achieve at least as high of a welfare by setting $\underline{q} = p = c + \rho k$. Given the first set of parameters, denote the socially optimal certification stopping time by τ_s^*. But under the second set of parameters, τ_s^* can also be implemented because $\mu_{\tau_s^*}^{NC} \leq p$ will also hold at any τ_s^* (recall that $\mu_t^{NC} \leq p$ for all t if $\underline{q} = p$), and $\mu_t > p$ for all $t < \tau_s^*$. In addition, once implemented the social welfare provided by certification will be at least as high, because all types of agents with qualities $\underline{q} > c + \rho k$ contribute positively to social welfare by certifying instead of exiting. Thus, the social welfare with $\underline{q} = p = c + \rho k$ must be at least as high as with any other standard. □

8.7 Proof of Theorem 5

This proof will proceed in several steps. Note that certification can be broken up into three possible types: immediate certification at $t = 0$, delayed certification that takes place at some $t > 0$, and no certification for all times. Which type of certification results will depend on the specific values of $\underline{q}$ and p. We prove that under pessimistic principal beliefs, delayed certification is never optimal. Then, we characterize the social welfare generated by immediate and no certification. We prove that immediate certification and no certification can both be optimal depending on how high the certification cost is.

First we show that implementing delayed certification is never socially optimal. If the price is set lower than $c + \rho k$, then certification can never occur because agents would prefer to exit than certify. Thus, for any type of certification to be implemented, we must have $p \geq c + \rho k$. Next, note that under pessimistic principal beliefs, the certification standard $\underline{q}$ must be set high enough such that $\mu_0^{NC} > p$ because otherwise agents would be expected and thus forced to certify immediately in a principal optimal equilibrium by Proposition 3. In particular, this requires that $\underline{q} > p$.

Now we analyze the social welfare generated by any delayed certification scheme, and we show that the welfare is strictly less than under no certification. Let $\hat{t} = \inf\left\{t|\mu_t^{NC} \leq p\right\}$. In a principal optimal equilibrium, $\hat{t}$ is the time at which certification would occur. Note that admissibility implies that at t, the truncated expected mean $\mu_{\hat{t}}^{NC} = p$. For any p and $\underline{q}$ that satisfy the above conditions, we can compute the resulting social welfare as:

$$\begin{aligned} W(p) &= \int_{-\infty}^{\infty}\left[\int_0^{\infty} e^{-\rho t}(q-c)\,dt\right] f_0(q)\,dq - E_{\hat{t}}\left[ke^{-\rho\hat{t}}\left(1-F_{\hat{t}}^{-}\left(\underline{q}\right)\right)\right. \\ &\quad \left. + F_{\hat{t}}^{-}\left(\underline{q}\right)\int_{\hat{t}}^{\infty} e^{-\rho t}\left(\mu_{\hat{t}}^{NC}-c\right)dt\right] \\ &= \frac{\mu_0-c}{\rho} - \left(kE_{\hat{t}}\left[e^{-\rho\hat{t}}\left(1-F_{\hat{t}}^{-}\left(\underline{q}\right)\right)\right] + \frac{p-c}{\rho}E_{\hat{t}}\left[e^{-\rho\hat{t}}F_{\hat{t}}^{-}\left(\underline{q}\right)\right]\right) \\ &\leq \frac{\mu_0-c}{\rho} \end{aligned}$$

This proves that the social welfare of any delayed certification scheme is bounded above by setting $p = c$ and $\underline{q} = \infty$, which results in no certification. From the blind boundedness theorem, we know that the payoff of such a scheme is exactly $\frac{\mu_0-c}{\rho}$.

Now, fixing a $\underline{q}$, the welfare provided by a de facto license is given by the expression $\int_{\underline{q}}^{\infty}(\frac{q-c}{\rho}-k)f_0(q)\,dq$. Given a k, the optimal certification standard is $\underline{q} = c+\rho k$. The reason is that any agent that certifies will give a social welfare of $\frac{q-c}{\rho}-k$, and this is the quality where this expression is equal to zero. Any agent with a quality higher than this amount contributes positively to welfare by certifying. Since we require that $p \leq \underline{q}$ and we need $p \geq c+\rho k$ for certification to occur, this implies that we need to set $p = c+\rho k$ in order to implement immediate certification. Thus, the highest *ex ante* surplus generated by any immediate certification scheme is $\int_{c+\rho k}^{\infty}(\frac{q-c}{\rho}-k)f_0(q)\,dq = \left(\frac{\mu_0^C-c}{\rho}-k\right)\left(1-F_0^{-}(c+\rho k)\right)$.

Thus, to see whether immediate certification is better, or whether no certification is better, we need to see which of the two surpluses is higher. This depends on the value of k, and specifically the cutoff value will be given by $k^* = \frac{\mu_0^C-c}{\rho} - \frac{\mu_0-c}{\rho\left(1-F_0^{-}(c+\rho k)\right)}$.
□

8.8 Proof of Theorem 6

First note that if either no certification or a de facto license is implemented, the social welfare will always be bounded away from the social optimal. With no certification, the welfare always equals the benchmark welfare for any k, and with a de facto license, the price and standard have to be set to at least μ_0. The social welfare of the de facto license is thus equal to $\int_{\mu_0}^{\infty}(\frac{q-c}{\rho}-k)f_0(q)\,dq$, which is bounded away from the first best welfare, $\int_c^{\infty}(\frac{q-c}{\rho})f_0(q)\,dq$, for any value of k.

Thus, in order to get asymptotic efficiency as $k \to 0$, we need to do it through a delayed certification scheme. We now show that the social welfare of any delayed certification scheme will also be bounded away from first best. First note that if the standard $\underline{q} \nrightarrow c$, then first best cannot be achieved. The reason is that before certification occurs, we are losing welfare from letting bad agents work, and after certification occurs we also lose welfare since not all good agents are working. Then assume that $p, \underline{q} \to c$. Fix any path of the expected mean for the agent. Let $t_c^* = \inf\{t|\mu_t \leq p; \overline{p} = c\}$ be the stopping time of this path in the limit as the price approaches c. In order to achieve first best as $k \to 0$, we must have $t_c^* = 0$ or else bad agents will be working for some stretch of time. But for any admissible information process this is impossible since $\mu_0 > p$. Thus, t_c^* is strictly above 0, and so delayed certification cannot achieve first best. □

8.9 Proof of Proposition 4

From the proof of Theorem 6, we see that immediate certification and no certification cannot asymptotically achieve first best as the information speed increases, because the speed of the reputational mechanism does not affect social welfare in either case. So we must show that the social welfare of a delayed certification scheme approaches first best. We wish to show that as the speed becomes very high, $t_c^*(q) \to 0 \, \forall q < c$, because doing so means that all agents who have socially inefficient qualities will be kicked out extremely quickly, and the good agents will be able to stay in forever (perhaps paying the certification cost that asymptotically approaches 0). Since the process is fully revealing, almost surely $t_c^*(q) < \infty$ for agents with quality $q < c$. Then as n gets large, agents will be kicked out at time $\frac{t_c^*(q)}{n}$ instead, which approaches 0 for all finite $t_c^*(q)$. Thus, $t_c^*(q) \to 0 \, \forall q < c$ and so delayed certification asymptotically achieves the first best outcome. □

References

Atkeson, A., Hellwig, C., Ordonez, G.: Optimal regulation in the presence of reputation concerns. In: NBER WP 17898 (2012)
Bar-Isaac, H.: Reputation and survival: learning in a dynamic signalling model. Rev. Econ. Stud. **70**, 231–251 (2003)
Bar-Isaac, H., Tadelis, S.: Seller Reputation. Now Publishers Inc (2008)
Board, S., Meyer-ter-Vehn, M.: Reputation for quality. Econometrica **81**, 2381-2462 (2013)
Bonatti, A.: Menu pricing and learning. Am. Econ. J. Microecon. **3**, 124–163 (2011)
Ely, J., Fudenberg, D., Levine, D.K.: When is reputation bad? Game. Econ. Behav. **63**, 498–526 (2008)
Ely, J.C., Välimäki, J.: Bad reputation. Q. J. Econ. **118**, 785–814 (2003)
Farhi, E., Lerner, J., Tirole, J.: Fear of rejection? Tiered certification and transparency. RAND J. Econ. **44**, 610–631 (2013)
Holmström, B.: Managerial incentive problems: a dynamic perspective. Rev. Econ. Stud. **66**, 169–182 (1999)
Jin, G.Z., Leslie, P.: The effect of information on product quality: evidence from restaurant hygiene grade cards. Q. J. Econ. **118**, 409–451 (2003)
Jin, G., Leslie, P.: Reputational incentives for restaurant hygiene. Am. Econ. J. Microecon. **1**, 237–267 (2009)

Klein, B., Leffler, K.B.: The role of market forces in assuring contractual performance. J. Polit. Econ. **89**, 615–641 (1981)

Kleiner, M.: Why license a florist? The New York Times. http://nyti.ms/1wnLUZk (2014). Accessed 26 July 2014

Lizzeri, A.: Information revelation and certification intermediaries. RAND J. Econ. **30**, 214–231 (1999)

Mailath, G.J., Samuelson, L.: Who wants a good reputation? Rev. Econ. Stud. **68**, 415–441 (2001)

Ordoñez, G.: Reputation from nested activities. Econ. Theory **52**, 915–940 (2013)

Panzar, J.C., Savage, I.: Does a minimum quality standard always reduce the price of high quality products? BE J. Econ. Anal. Policy **11**, 1–30 (2011)

Shapiro, C.: Premiums for high quality products as returns to reputations. Q. J. Econ. **98**, 659–679 (1983)

Strausz, R.: Honest certification and the threat of capture. Int. J. Ind. Organ. **23**, 45–62 (2005)

Stahl, K., Strausz, R.: Who should pay for certification? In: ZEW WP 11–054 (2011)

Xiao, M.: Is quality accreditation effective? Evidence from the childcare market. Int. J. Ind. Organ. **28**, 708–721 (2010)