

The Regulation of Moral Hazard
in Retail Transactions

by

Ying Xue

Business Administration
Duke University

Date: _____

Approved:

Simon Gervais, Supervisor

Xu Jiang

Vish Viswanathan

Ming Yang

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2014

ABSTRACT

The Regulation of Moral Hazard
in Retail Transactions

by

Ying Xue

Business Administration
Duke University

Date: _____

Approved:

Simon Gervais, Supervisor

Xu Jiang

Vish Viswanathan

Ming Yang

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2014

Copyright © 2014 by Ying Xue
All rights reserved

Abstract

In transactions where costly efforts from both the seller and the buyer help identify the unfit buyer, the unique optimal law that restores the first-best requires the seller to refund the buyer and pay a punitive penalty of the buyer's loss. If the social planner can infer from a loss whether efficient efforts had been exerted, the law with contingent punitive penalty always restores the first-best with a balanced actual budget. The law is harsher for one-stop than for specialized purchase, as a loss is more likely due to shirking. Higher budget surplus cost can make all traders better off and the social optimum more implementable. We also illustrate constrained optima if the first-best is unattainable or if a balanced budget is enforced. With costless efforts or unilateral moral hazard or no effort choices, no punitive penalty is needed to always restore the first-best with a balanced budget. Bilateral moral hazard rationalizes punitive penalty and determinacy of law yet complicates regulation. Our model applies widely to many goods and services.

I also study efficient disclosure by contingent non-disclosure. I show that if the sender is not always active and receivers do not know if the sender is active, then compound information can be efficiently conveyed only if some information is withheld contingently. Our theory has wide applications in fields like political economy, public policy, and law. It is optimal for the benevolent social planner to purge dissident yet useful information so as to convey the more welfare-relevant information to its citizens. Even the most liberal and transparent government should implement

undemocratic policies for the citizens' own good. Transparency or disclosure laws can hurt citizens. We contribute to the economic theory of information transmission, and identify a new situation where the sender's interest is aligned with receivers', more information always helps receivers, yet it is optimal to contingently block some information from receivers. Our theory has novel predictions and policy implications.

Contents

Abstract	iv
List of Tables	ix
List of Figures	x
Acknowledgements	xi
1 The Regulation of Moral Hazard in Retail Transactions	1
1.1 Introduction	1
1.2 The Model	7
1.2.1 Setup	7
1.2.2 The First-Best	9
1.2.3 The Need for Law	10
1.3 First-Best Implementation	10
1.4 Regulation with a Balanced Budget	13
1.5 Constrained Optimal Regulations	14
1.5.1 When the First-Best is Unattainable	15
1.5.2 Forcing a Balanced Budget	16
1.5.3 Costly Budget Surplus	17
1.6 One-Stop Seller	19
1.7 The Importance of Costly Dual Moral Hazard	21
1.8 Conclusions	24

2	Disclosure by Contingent Non-Disclosure	26
2.1	Introduction	26
2.2	The Model	30
2.3	Optimal Decision and the First-Best	33
2.4	Beliefs Under Full Disclosure	35
2.5	The Inefficiency of Full Disclosure	37
2.6	The Efficiency of Contingent Non-Disclosure	39
2.7	Predictions	42
2.8	Robustness	46
2.9	Conclusions	47
A	Proofs for Chapter 1	49
A.1	Proof of Proposition 1 (The First-Best)	49
A.2	Proof of Lemma 2 (The Need for Law)	50
A.3	Proof of Proposition 3 (Implementing the First-Best)	50
A.4	Proof of Proposition 4 (Balance the Budget)	51
A.5	Proof of Lemma 5 (The First-Best for One-Stop Seller)	52
A.6	Proof of Proposition 6 (Implementing the First-Best for One-Stop Seller)	53
A.7	Proof of Proposition 7 (The Importance of Dual Moral Hazard)	54
A.8	Motivation for Ex Post Inference	56
B	Proofs for Chapter 2	58
B.1	Derivations for Results in the First Table	58
B.2	Proof of Lemma 9 (The Social Planner's Belief)	59
B.3	Proof of Lemma 10	59
B.4	Proof of Proposition 11 (The Inefficiency of Full Disclosure)	62
B.5	Proof of Lemma 12	63

B.6	Proof of Proposition 13	64
B.7	Proof of Proposition 18	64
	Bibliography	65
	Biography	72

List of Tables

1.1	Marginal and Joint Distributions of \tilde{s} and $\tilde{\tau}$	8
1.2	Marginal and Joint Distributions of \tilde{s} and $\tilde{\tau}$ (One-Stop Seller)	20
2.1	The Agents' First-Best Decision in Each Event in Λ	34
2.2	Information Sets and Beliefs Under Full Disclosure	36
2.3	The Agents' Optimal Decision Under Full Disclosure	38
2.4	The Agents' Optimal Decision Under Contingent Non-Disclosure	40

List of Figures

1.1	The Sequence of Events	8
1.2	(a^*, e^*) traced out by changes in x or y	17
1.3	$x - p$ and $y - p$ traced out by changes in κ on $[0, \frac{1}{6}]$	18
2.1	A Simplified Game Tree	32
2.2	The Equilibrium and Agents' Information Sets Under Full Disclosure	37
2.3	The Equilibrium and Agents' Information Sets Under Contingent Non-Disclosure	41

Acknowledgements

I am extremely grateful to my advisor and the chair of my PhD dissertation committee Professor Simon Gervais and my dissertation committee members Professor Xu Jiang, Professor Vish Viswanathan, and Professor Ming Yang.

The Regulation of Moral Hazard in Retail Transactions

1.1 Introduction

Warren Buffett said: “What doesn’t work is when you start doing things that you don’t understand or because they worked last week for somebody else.” Indeed, it is crucial that investors be allocated financial products that suit them, for otherwise dire consequences like the global financial crisis may ensue. The regulation in practice,¹ the public opinion,² and the extant academic literature emphasize disciplining the supply side, but consumer protection is an insurance that the buyer may abuse. A potential trigger of the crisis may be that the homeowner did not make enough effort to effectively interact with the mortgage seller for the seller to correctly learn the suitability or affordability of certain mortgages for the homeowner. In this paper we

¹ President Obama proposed addressing the failure by directly focusing on consumers and signed the Dodd-Frank Wall Street Reform and Consumer Protection Act in 2010 that created the Consumer Financial Protection Bureau, which handled 91,000 lawsuits in 2012. Other regulatory agencies include the Federal Reserve Consumer Help, FDIC Consumer Response Center, SEC Complaint Center, Small Business Administration Consumer Affairs, etc.

² Many blamed the bad incentives of finance firms and “evil” CEOs; some even occupied the Wall Street (McCann 2011).

study the optimal regulation in a market where the buyer, as well as the seller, can exert costly effort to help the seller identify the unfit buyer.

The Fed fined Wells Fargo \$85 million for inappropriate transactions of subprime mortgages and loans, and consumers got compensated what they paid for these financial products (Simon and McGrane 2011). Punitive penalty suggests the infeasibility of self-regulation and necessitates third-party regulation. It may seem puzzling why a seller actually incurred a punitive penalty given that, as we will see later, the law can effectively contain unilateral moral hazard with a balanced budget. The optimal regulation we derive in this paper justifies the Fed's legal action against Wells Fargo, and may provide guidelines for future court rulings.

Given the huge size of the retail finance market,³ our paper has substantial welfare import. Yet severe consequences due to mismatches are non-exclusive to retail finance and can occur for most goods and services. It is often the case that the seller has the expertise on matching goods to buyers while the buyer lacks such expertise. Hence by diligently interacting with the buyer, the seller can gain an informational advantage on the suitability of the good for the buyer. The buyer's effort can help the seller better understand the buyer's own specific situation and thus improve the precision of the seller's private information and reduce sales mistakes. Ideally, the seller should put himself in the buyer's shoes, the buyer should carefully listen, learn and introspect to understand the good and his own objective, needs and constraints, and the two should talk effectively.

Therefore the advising and sorting of the buyer is a joint production process where the buyer and the seller interact for the seller to learn the buyer's type. We model the situation as one of dual moral hazard with information asymmetry. The buyer and the seller each takes a costly hidden action that enhances the likelihood that

³ \$95.5 trillion of assets held by households in Q1 of 2014 according to Table B.100 of the Federal Reserve's Flow of Funds data (<http://www.federalreserve.gov/releases/z1/>).

the seller identifies the unfit buyer, and the seller decides whether to sell based on what he learns from their interaction. As examples of their costly activities, the pair need to clearly articulate and objectively analyze information, learn new knowledge and absorb information, and truthfully disclose relevant and up-to-date information. They also need to acquire good- or buyer-specific human capital, think hard and plan scrupulously, make difficult judgement, decisions and tradeoffs, and act responsibly. It is unclear what responsibility each party should take for the buyer's loss, since each may free-ride on the other's effort.

We find that the only law that restores the first-best in our model requires the seller to refund the buyer and pay the social planner a punitive penalty of the buyer's loss if the buyer experiences a bad outcome due to his mismatch to the good sold. Due to the punitive penalty and the ensuing budget surplus, self-regulation is infeasible and third-party regulation becomes necessary. Punitive penalty and hence third-party regulation are rationalized in the sense that they are necessary to restore efficiency when there is costly dual moral hazard between the buyer and the seller but not with unilateral moral hazard, no effort choices, or costless effort.

We then show that the issue of budget surplus can be resolved if the social planner has slightly more information ex post. We use contingent threat for only one player, the seller, to always restore the first-best in this double moral hazard problem with a balanced budget on the equilibrium path. The social planner uses minimal ex post binary information on whether the buyer and the seller had been sufficiently productive when the buyer suffers a loss.

We study three different constrained optimal regulations. Each approach resolves the budget surplus issue in our first-best implementation. First, we consider the case where the social planner's budget surplus exceeds the market surplus so that efficiency cannot be restored. Second, we examine the situation when the budget is forced to be always balanced so that there is no budget surplus to begin with.

Third, we analyze the setting where running a budget surplus is costly for the social planner. Surprisingly, higher budget surplus cost improves everyone's welfare in the market, and the social optimum becomes more implementable.

The seller does not need to worry about losing revenue when he can satisfy every type of buyer. For example, a highly integrated corporation is one such seller. One might suspect that such a situation calls for less regulation because there is less bad incentive against advising correctly. However, we show in an extension to our model that the law in the case of a one-stop seller is more stringent for both the seller and the buyer than the law in the case of a specialized seller. The intuition is that a loss is more likely due to the lack of effort.

We are the first to recognize, introduce, and analyze the new friction that the buyer's effort, as well as the seller's, helps the seller sort the buyer, and study legal rules that restore efficiency. Spengler (1968), Akerlof (1970), Oi (1973), Heal (1976, 1977), Spence (1977), Kronman (1978), Simon (1981) provide early economic analysis of safety, quality, buyer protection, and the relevant inefficiencies. Palfrey and Romer (1983) study buyer-seller disputes due to their imperfect judgment of product quality. Shavell (1994) analyzes traders' information acquisition and mandatory versus voluntary disclosures. Bolton, Freixas, and Shapiro (2007) examine the conflict of interest when the seller decides whether to provide information to buyers in financial markets. Inderst and Ottaviani (2009) study policy issues and firms' tradeoffs when marketing agents prospect for and advise buyers. Carlin and Gervais (2012) analyze the optimal mix of penalties that contains the dual moral hazard between producers and intermediaries.

Our introduction and creation of the dual moral hazard framework contributes to the literature on credence goods and delegated expertise. Arrow (1963) analyzes the situation in which a doctor has an informational advantage relative to his patients when he recommends a course of action. Darby and Karni (1973) introduce the

credence quality of goods and show the value monitoring. Demski and Sappington (1987) develop a principal-agent model where information acquisition and communication is costly for the expert but implementation is free, and study contracts that curb the induced moral hazard in implementation. Wolinsky (1993) shows delegated expertise may lead to vertical specialization, and the buyer's search for multiple opinions and reputation concerns may discipline the seller and determine the industrial organization. Taylor (1995) studies the insurance solution to the delegated expertise problem and concludes that imperfections in the spot insurance market rationalizes free diagnosis, buyer's procrastination, and long-term health agreements.

We contribute to the moral hazard in teams literature by restoring the first-best using penalty for only one agent contingent on minimal ex post information with a balanced budget on the equilibrium path. Seminal work on team moral hazard includes Groves (1973) and Holmström (1982).

We show the seller's valuable expertise goes wasted unless the buyer and the seller are policed by well-designed laws, which predicts a positive relationship among retail market law, quality matching and advising services, reliable seller and buyer behavior, and the productivity of their interaction. The seller and the buyer act as each other's agent and the law motivates both to behave responsibly on behalf of the other. We thus contribute to the literature on agency law and its economics (Sykes 1984, Sykes 1988, Rasmusen 2004).

Our framework applies to myriads of goods and services for individuals, organizations and nations alike. It covers all credence goods like repairs and maintenance, education and advising, and counseling and technological services. It encompasses sellers who provide consulting services before selling a good or service such as doctors, technicians, lawyers, and consultants. Transactions between two institutions often necessitate the buyer's personnel or technology support for the seller to determine the suitability of the deal. Similarly, collaboration at multiple levels of businesses

and governments are usually called for before a nation deems it fit to sell to another. The seller can work in primary or secondary markets and can be a direct seller or an intermediary. The one-stop purchase extension of our model applies to customized designs like architecture, software, and decoration that require the seller to process each order in consultation with the buyer.

The buyer's exerting effort can mean overcoming the unpleasant and the difficult to comply and cooperate. The patient needs to strictly stick to the restrictions and instructions set by the doctor to exclude confounding factors that interfere with diagnosis, especially before medical exams or decision points at the start or change of treatment plans. Effort may also mean adequate disclosure and high transparency. Despite the miserable memories, the patient's detailed recall and clear articulation of his feeling and full disclosure of his family and medical history definitely facilitate the diagnosis. It is thus unclear whether the failure to cure due to wrong diagnosis or treatment is the doctor's misjudgment and malpractice or the patient's noncooperation.

Costly effort can be investment in the economic (Prendergast 1993, Nunn 2007), marketing (Dwyer et al 1987, Doney and Cannon 1997), and financial (Puri and Rocholl 2008, Drucker and Puri 2009, Iyer and Puri 2012) relationships, the development of social ties and familiarity, and the production of information (Diamond 1984, 1991; Stein 2002), all of which enable the seller to better judge whether the buyer is suited for transactions like relationship lending. We shed light on the optimal regulation for transactions with these costly collaborative inputs.

The buyer's loss includes damages to third-parties. To prevent not only the potential harm to the buyer but also the immense externality on the society, weapons, harmful substances, addictive intakes and activities, R-rated movies and disturbing material should be kept away from terrorists and criminals, the addicted, the intoxicated, the mentally ill, and the underage. The seller's effort in adhering to the

rule to verify the buyer’s identity, condition, purpose, and background is costly, as well as the buyer’s effort to cooperate, acquiesce and follow. Gun and liquor related violence and accidents like atrocious shooting massacres and severe traffic collisions or pileups could have been avoided had arms dealers been conscientious enough to not sell to criminals or the mentally disordered, and bar tenders judiciously guided away the drunk. Consequences from shirking can cause far-reaching repercussions beyond what can be compensated by money, which endows the seller and the buyer with important duties.

The rest of this paper is organized as follows. The next section presents our model. Section 1.3 implements the first-best. Section 1.4 shows how ex post binary information on joint productivity helps. Section 1.5 studies three constrained optimal regulations. Section 1.6 derives the optimal law in the case of one-stop purchase. Section 1.7 shows the importance of dual moral hazard. The final section concludes.

1.2 The Model

1.2.1 Setup

A risk-neutral seller can sell a good or provide a service that is worth 0 to himself to a buyer with unit demand and 0 reservation utility. The buyer’s payoff from the good is $\tilde{\tau} \in \{l, h\}$, where $h > 0 > l$. No one knows the buyer’s type ex ante but its distribution is common knowledge: $Pr(\tilde{\tau} = l) = \phi \in (0, 1)$ and $Pr(\tilde{\tau} = h) = 1 - \phi$. When the buyer and the seller interact, the buyer chooses hidden effort $e \in [0, 1]$ at cost $c(e)$, and the seller chooses hidden effort $a \in [0, 1]$ at cost $k(a)$. Only the seller has the expertise to observe a private non-verifiable signal $\tilde{s} \in \{l, h\}$. The seller bases his decision to sell on \tilde{s} , which is generated from type by $Pr(\tilde{s} = l | \tilde{\tau} = h) = 0$ and $Pr(\tilde{s} = l | \tilde{\tau} = l) = f(a, e) \in [0, 1]$. The functions k, c, f are increasing in a and e with $k(0) = c(0) = f(0, 0) = 0$. Table 1.1 shows the marginal and joint distributions of \tilde{s} and $\tilde{\tau}$. The seller’s belief given \tilde{s} is $Pr(\tilde{\tau} = l | \tilde{s} = l) = 1 > \phi$ and

$$Pr(\tilde{\tau} = l | \tilde{s} = h) = \frac{1-f}{1-\phi f} \phi \leq \phi.$$

Table 1.1: Marginal and Joint Distributions of \tilde{s} and $\tilde{\tau}$

		$Pr(\tilde{\tau} = l)$	$Pr(\tilde{\tau} = h)$
		ϕ	$1 - \phi$
$Pr(\tilde{s} = l)$	ϕf	ϕf	0
$Pr(\tilde{s} = h)$	$1 - \phi f$	$(1 - f)\phi$	$1 - \phi$

The legal environment is the following. We assume the social planner can perfectly verify whether or not a buyer has experienced l for a good that he purchased. Such verification implies that the good was sold to a buyer who was unfit for it and who suffered from a bad outcome. To protect buyers, the legal system set by the social planner imposes a penalty of x on the seller and a compensation of y to the seller if the buyer experiences l . A buyer who experiences l sues and gets compensated at no cost. In setting the legal environment (x, y) , the government's objective is to maximize total welfare in the economy.

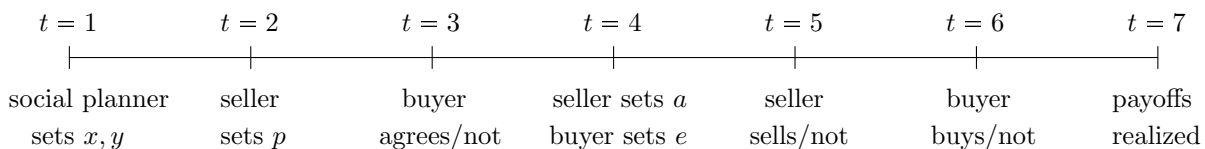


FIGURE 1.1: The Sequence of Events

Figure 1.1 shows the timing of the game. At $t = 1$ the social planner sets the law (x, y) . At $t = 2$ the seller sets the publicly observable price p , which is a take-it-or-leave-it offer. At $t = 3$ the buyer decides whether to agree to the offer and interact with the seller. If the buyer agrees, then at $t = 4$ the pair interact, during which the seller sets a and the buyer sets e . At $t = 5$ the seller decides whether or not to sell given his privately observed \tilde{s} . At $t = 6$ the buyer decides whether or not to buy. Finally, at $t = 7$ payoffs are realized.

The seller sells with non-verifiable probabilities $\lambda_l, \lambda_h \in [0, 1]$ to a buyer with $\tilde{s} = l$ and to a buyer with $\tilde{s} = h$ respectively. Let q_s, q_h, q_l be the respective

probabilities that the buyer gets the good, gets a payoff of h , and gets a payoff of l . If the buyer buys, we have $q_s = Pr(\tilde{s} = l)\lambda_l + Pr(\tilde{s} = h)\lambda_h = \phi f\lambda_l + (1 - \phi f)\lambda_h$, $q_h = Pr(\tilde{s} = \tilde{\tau} = h)\lambda_h = (1 - \phi)\lambda_h$, and $q_l = q_s - q_h = [f\lambda_l + (1 - f)\lambda_h]\phi$.

The buyer's ex ante expected utility if he participates and buys is

$$\begin{aligned} u(a, e, \lambda_h, \lambda_l, p, y) &= q_h h + (y + l)q_l - q_s p - c \\ &= [(1 - \phi)h - (1 - \phi f)p + (1 - f)(y + l)\phi]\lambda_h + (y + l - p)\phi f\lambda_l - c. \end{aligned} \quad (1.1)$$

The seller's expected profit is

$$\begin{aligned} \pi(a, e, \lambda_h, \lambda_l, p, x) &= q_s p - q_l x - k \\ &= [(1 - \phi f)p - (1 - f)\phi x]\lambda_h + (p - x)\phi f\lambda_l - k. \end{aligned} \quad (1.2)$$

With probability q_l the social planner collects x from the seller and issues y to the buyer. The social planner's expected budget surplus is thus $(x - y)q_l$. The budget surplus is redistributed at no cost so it directly adds to total welfare. The total welfare is thus

$$\begin{aligned} w(a, e, \lambda_h, \lambda_l) &= q_h h + q_l l - c - k \\ &= [(1 - \phi)h + (1 - f)\phi l]\lambda_h + \phi f l\lambda_l - c - k. \end{aligned} \quad (1.3)$$

1.2.2 The First-Best

In the first-best problem the social planner aims to maximize total welfare by solving

$$\max_{a, e, \lambda_h, \lambda_l \in [0, 1]} w(a, e, \lambda_h, \lambda_l). \quad (1.4)$$

Proposition 1. (The First-Best) *Let (a^r, e^r) maximize $w(a, e, 1, 0)$.*

If $w(a^r, e^r, 1, 0) < 0$, then $(a, e, \lambda_h, \lambda_l) = (0, 0, 0, 0)$ is the first-best with zero social welfare. The market breaks down and no law is needed.

If $w(a^r, e^r, 1, 0) \geq 0$, then $(a^r, e^r, 1, 0)$ is the first-best.

We call $\lambda_h = \lambda_l = 0$ the No-sale mode and $(\lambda_h, \lambda_l) = (1, 0)$ the Regular mode. No effort should be exerted in the No-sale mode since effort does not help. No-sale and

zero efforts are optimal if the good fails to create value overall even when optimal efforts are exerted in the Regular mode. For example, risky or inefficient goods with high ϕ or low h and l , as well as goods with low f but high k and c so that mismatches are hard to identify unless huge amounts of efforts are exerted. Such products are best left out of the market. Since the market attains the first-best by itself without law if No-sale is optimal, hitherto assume the more interesting case of $w(a^r, e^r, 1, 0) > 0$ so that $(a^r, e^r, 1, 0)$ is the first-best.

1.2.3 *The Need for Law*

Since $a, e, \lambda_h, \lambda_l$ are not contractible, shirking and the lack of commitment become central concerns. We next show that inefficiency exists if there is no law or if the law is weak.

Lemma 2. (The Need for Law) *If $x < p$ then $(a, e, \lambda_h, \lambda_l) = (0, 0, 0, 0)$. The outcome is inefficient as the market breaks down.*

Therefore modest penalty fails to sort the buyer: zero efforts and inefficiency ensue. So in general potentially beneficial products and the seller's knowledge are wasted under insufficient regulation. The transaction is inefficient and potential gains from effective seller-buyer interaction are lost. The penalty on the seller should be sufficiently high to deter indiscriminate selling. Only appropriate law (x, y) can potentially correct incentives, induce sensible sales and efforts, and attain efficiency. Regulation must be carefully designed for the buyer and the seller's incentives to be re-aligned with the social planner's.

1.3 First-Best Implementation

In this section we study how appropriate regulation can realign incentives and restore efficiency.

Proposition 3. (Implementing the First-Best) *If $w(a^r, e^r, 1, 0) + (1 - f^r)\phi l \geq 0$ where $f^r \equiv f(a^r, e^r)$, then $(x^r, y^r) = (p - l, p)$ is the unique law that restores the first-best $(a^r, e^r, 1, 0)$ and does so with an expected budget surplus of $(f^r - 1)\phi l$; otherwise the first-best cannot be restored and $a = e = q_s = 0$ in the No-sale mode prevails instead.*

The condition for the seller to choose the Regular mode $w(a^r, e^r, 1, 0) + (1 - f^r)\phi l \geq 0$ is more stringent than $w(a^r, e^r, 1, 0) \geq 0$, the condition for the Regular mode to be the first-best. So the seller acts more conservatively than the social planner due to the punitive penalty which creates the budget surplus.

The law (x^r, y^r) requires that the seller issue a full refund of the transaction price to the buyer who experiences l and pay the social planner a punitive penalty that equals the buyer's loss. Thus a good that is worse for the type l buyer leads to more punitive penalty for the seller and a less balanced budget. This makes sense since the punishment for the seller is in lockstep with the buyer's misery, which effectively disciplines the seller. A compensation to the buyer that guarantees his money back yet relates less to his loss helps him relax yet still disciplines him. The law gives both parties incentives to act responsibly. The punitive penalty fits the empirically observed incidents such as the Wells Fargo case (Simon and McGrane 2011).

The fact that the social planner expects to run a budget surplus to implement the first-best implies that it is very difficult if not impossible for the seller to restore efficiency by self-regulation. Beyond a refund, it is hard for a seller to commit to punishing himself by throwing money away just to attain the social optimum. A seller would only compensate the buyer via a limited warranty and set $x = y$ or $x = p$ but not $x^r > y^r = p$ as the optimal law dictates. Even if the firm donates the money it helps improve the firm's image, reputation and public relations which is different from a pure penalty.

Because of the unbalanced budget we are able to implement the first-best only conditionally when the seller can still make profit. Otherwise the seller quits under (x^r, y^r) and No-sale ensues instead.

It is interesting and somewhat surprising that product or security design is irrelevant up to one degree of freedom for our first-best equilibrium. The three parameters that characterize the good's payoff are ϕ , h , and l . Suppose we give up one degree of freedom and only keep the unconditional expected loss $\phi l \equiv L < 0$ fixed. Since the first-order conditions for the first-best efforts only depend on L , we have the same first-best efforts a^r and e^r . Then the expected budget surplus $(f^r - 1)L$ remains the same.

The first-best social welfare $(1 - \phi)h + (1 - f^r)L - c^r - k^r$ and hence the seller's profit $w(a^r, e^r, 1, 0) + (1 - f^r)L$ change with the product design in the same way that $(1 - \phi)h$ does. If $(1 - \phi)h \equiv H > 0$ is fixed and independent of ϕ , then these two quantities are independent of product design.

Now suppose the seller has some discretion over product design and is free to choose any $\phi \in [\delta, 1 - \delta]$ where $\delta \in (0, \frac{1}{2})$. We still keep $L = \phi l$ fixed. If $(1 - \phi)h$ is decreasing in ϕ , say if h is fixed and independent of ϕ , then it is both socially optimal and best for the seller to choose the minimal $\phi = \delta$ so as to capture the upside gain h most of the time. However, this means when the buyer loses he loses big: $l = L/\delta = \max_{\phi \in [\delta, 1 - \delta]} L/\phi$. This is consistent with some explanations that have been offered for the recent credit crisis. Warren Buffett (2010) testified at the Financial Crisis Inquiry Commission that "the whole American public was caught up in a belief that American housing couldn't fall dramatically" and Paul Volcker (2010) concurred. With such belief the economy including the regulator allowed the mortgages to be designed and sold in a way that homeowners win most of the time, but with a small likelihood that house prices keep falling they lose and lose big.

If $(1 - \phi)h$ is increasing in ϕ , say if the upside potential h is increasing fast

in the probability of mismatch ϕ , then it is both socially optimal and best for the seller to give up matching probability for higher upside potential and choose the maximal $\phi = 1 - \delta$. Then with high probability the product does not work for the buyer but when it does it wins big. When the buyer loses he does so moderately: $l = L/(1-\delta) = \min_{\phi \in [\delta, 1-\delta]} L/\phi$. This could be a fair characterization of the securities that come with limited liability in the asset market.

1.4 Regulation with a Balanced Budget

The issue of budget surplus and conditional implementation can be resolved if the social planner has a bit more information ex post. In particular, suppose that if the buyer gets l then the social planner can tell $f < f^r$ from $f \geq f^r$ ex post.⁴ Still, $1\{f < f^r\}$ can not be inferred if the buyer gets h or before payoffs are realized.

Suppose the buyer gets l . If ex post the social planner infers $f \geq f^r$, then measured in terms of f the seller and the buyer exerted at least as much joint efforts as (a^r, e^r) . It would seem that the social planner should intervene less in this case since the agents are already working hard enough to contain the chance of the buyer's loss below its first-best level. In this section we show that when the buyer incurs a loss, as long as the social planner can in some way infer ex post whether the joint productivity of the buyer and the seller had been satisfactorily high, then the actual penalty for the seller equals the compensation to the buyer so that the social planner runs a balanced budget on the equilibrium path and implements the first-best unconditionally by contracting with the seller on $1\{f < f^r\}$ in a natural and reasonable way.

The solution we propose is a slight modification to (x^r, y^r) by making the punitive penalty on the seller contingent on $1\{\tilde{\tau} = l, f < f^r\}$. If ex post $\tilde{\tau} = l$ and $f < f^r$

⁴ Motivations for such possibility, as well as alternative modeling methodology, are given in the Appendix.

then the punitive penalty is imposed to deter the seller; otherwise there is no penalty other than the full refund to the buyer. This regulation can both correct incentives to restore efficiency and balance the budget on the equilibrium path. The imperfection under (x^r, y^r) can now be fixed.

Proposition 4. (Balance the Budget) *If the social planner can infer $1\{\tilde{\tau} = l, f < f^r\}$ ex post, then $(x^b, y^r) = (p - 1\{\tilde{\tau} = l, f < f^r\}l, p)$ always restores the first-best $(a^r, e^r, 1, 0)$ with a balanced budget on the equilibrium path.*

Under (x^b, y^r) , the seller chooses the Regular mode over the No-sale mode whenever the Regular mode is the first-best because the social planner no longer expects to collect any budget surplus so the seller faces precisely the same incentive that the social planner did in its first-best solution. Recall that (x^r, y^r) implements the first-best only when the seller makes profit but now (x^b, y^r) implements the first-best unconditionally. (x^b, y^r) is a simple, practical, and reasonable rule to use.

Our result contributes to the literature on moral hazard in teams. We discipline both agents and restore the first-best by contingent penalty for only one agent with a balanced budget on the equilibrium path. We can in general do no better than (x^b, y^r) in the sense of less threat off the equilibrium path or less refund on the equilibrium path. For one thing, y^r is the unique compensation that restores the buyer's effort to e^r given a^r from the seller in the Regular mode. For another, under a less severe contingent penalty the seller may find an effort level lower than a^r optimal instead.

1.5 Constrained Optimal Regulations

We investigate three different constrained optimal regulations. In each case the problem is due to the budget surplus in our first-best implementation, so none of these is necessary if $1\{\tilde{\tau} = l, f < f^r\}$ can be inferred ex post so that its budget can be balanced on the equilibrium path.

First, we study the situation when the first-best is unattainable. Second, we examine the case when the budget is forced to be always balanced so that there is no punitive penalty. Third, we analyze the setting when the budget surplus is costly. At such generality we do not get explicit solutions, so we focus on an analytical example for the economic intuition.

1.5.1 When the First-Best is Unattainable

We study the best the social planner can do when the first-best is not attainable. Namely, we look for the optimal regulation when the seller's profit under (x^r, y^r) is negative: $w(a^r, e^r, 1, 0) + (1 - f^r)\phi l = (1 - \phi)h + 2(1 - f^r)\phi l - c^r - k^r < 0$. This can happen if the good has low payoffs h or l , or if ϕ is high so that it suits few, or if the sorting technology is crude with low f^r , or if effort costs c^r, k^r are high. The seller quits the market so the first-best cannot be implemented and positive social surplus $w(a^r, e^r, 1, 0)$ is lost.

However, mutually beneficial transactions may still occur with gains from trade realized if we do not insist on first-best efforts, say if we settle for less than the first-best efforts so that c and k drops much while f changes little. Positive total surplus can still be implemented if the punitive penalty and hence the budget surplus is not too big.

For concreteness suppose

$$f(a, e) = \frac{a + e}{2}, \quad c(e) = \frac{e^2}{2}, \quad k(a) = \frac{a^2}{2}, \quad \phi = \frac{1}{2}, \quad h = 2, \quad l = -2. \quad (1.5)$$

The first-best efforts $a^r = e^r = \frac{1}{2}$ yield positive total welfare $w(\frac{1}{2}, \frac{1}{2}, 1, 0) = f^r - k^r - c^r = \frac{1}{4} > 0$. The only law that stands a chance of implementing the first-best is $(x^r, y^r) = (p + 2, p)$ but it fails to work because the seller makes negative profit: $w(\frac{1}{2}, \frac{1}{2}, 1, 0) + (1 - f^r)\phi l = -\frac{1}{4} < 0$. The optimal law in this case is $(x, y) = (p + \frac{4}{3}, p + \frac{2}{3})$ and it implements less than the first-best efforts: $a^* = e^* = \frac{1}{3}$. Since $x^r > p + \frac{4}{3}$ and

$y^r < p + \frac{2}{3}$, the optimal law is more lenient than (x^r, y^r) for both the seller and the buyer. The maximal implementable social welfare is $w(a^*, e^*, 1, 0) = \frac{2}{9} > 0$. Hence positive total surplus is realized even though the first-best is not attainable.

The first-best is implementable when the joint surplus in the market is positive. However, if the seller and the buyer cannot break even under (x^r, y^r) , then the market's participation constraint is violated and the first-best is not implementable. But if we maximize total surplus given such binding constraint then some positive total surplus may still be realized. Such constrained optimal regulation and (x^r, y^r) deal with the two possible market statuses in the first-best implementation under (x^r, y^r) , and hence together they provide a more comprehensive solution to the dual moral hazard problem.

1.5.2 Forcing a Balanced Budget

We saw the first-best can not be restored if the social planner's budget surplus drives the seller out of the market though the total surplus is still positive. Now we impose $y = x$ from the outset so that the budget is always balanced and the market gets all the surplus. This is appealing because the seller's sales decision is always efficient and the total welfare is always realized as long as it is positive. We look for the constrained optimal law x , the amount that the seller compensates the buyer when the buyer suffers a loss.

Continue from the previous analytic example. The constrained optimal regulation when the budget is always balanced is $x = p - l/2 = p + 1$. Since $x^r > p + 1 > y^r$, this is less stringent than (x^r, y^r) for both the seller and the buyer. The implemented efforts are below the first-best level: $a^* = e^* = \frac{1}{4}$. Positive social welfare is realized: $w(\frac{1}{4}, \frac{1}{4}, 1, 0) = \frac{3}{16} > 0$.

Figure 1.2 shows a continuum of implementable equilibria that realize positive social surplus with an enforced balanced budget. The first-best $a^r = e^r = \frac{1}{2}$ is above

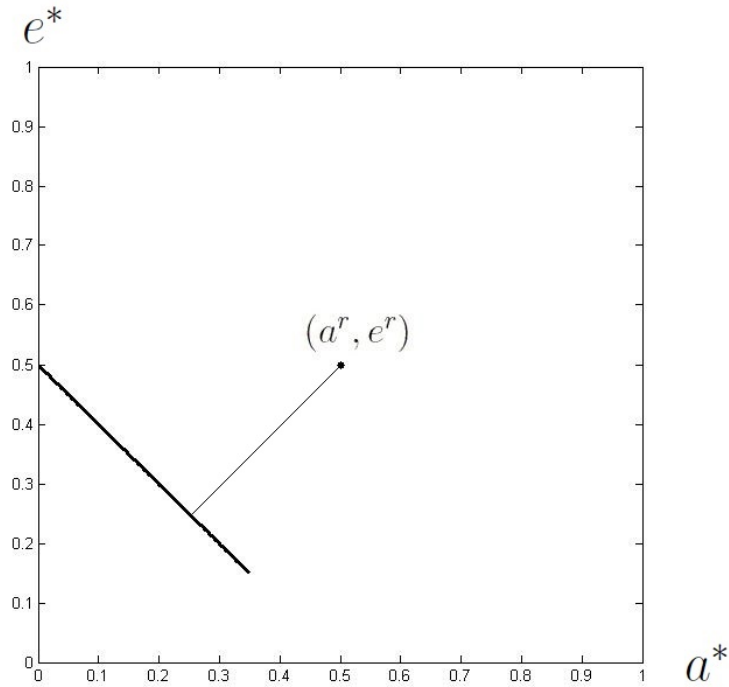


FIGURE 1.2: (a^*, e^*) traced out by changes in x or y

all of them. The second-best equilibrium we have identified is the one closest to the first-best.

1.5.3 Costly Budget Surplus

We have assumed that the budget surplus is fully utilized by the social planner and hence costless. If not, then the first-best can not be implemented due to the dead weight loss from budget surplus cost. We now derive the optimal law when the budget surplus is costly.

For concreteness suppose the budget surplus costs $(x - y)\kappa$ in our analytic example. The optimal regulation $x = p + 2 - \frac{2\kappa}{1-4\kappa} < x^r$ and $y = p + \frac{2\kappa}{1-4\kappa} > y^r$ is less harsh than (x^r, y^r) for both the seller and the buyer. The budget surplus given the buyer's loss drops from 2 to $2 - \frac{4\kappa}{1-4\kappa}$, which makes sense since the social planner uses less budget surplus when it is costly. The budget surplus being positive means $\kappa \in [0, \frac{1}{6}]$.

The interpretation is that the social planner is willing to run a budget surplus only if it is not too costly. As long as there is a budget surplus, the implemented efforts $a^* = e^* = \frac{1}{2} \left(1 - \frac{\kappa}{1-4\kappa}\right)$ are below the first-best, the social surplus is positive, and the seller's profit $\frac{3}{4} \left(\frac{\kappa}{1-4\kappa}\right)^2 + \frac{\kappa}{2(1-4\kappa)} - \frac{1}{4}$ is increasing in κ the budget surplus cost. Thus if the budget surplus costs more, the social planner keeps less of it and the seller gets more. The seller makes profit and thus is willing to participate if $\kappa \geq \frac{1}{7}$, so low budget surplus cost does not help. In particular, if $\kappa = 0$, then the seller makes negative profit $-\frac{1}{4} < 0$ which is the reason the first-best can not be implemented in the first place.

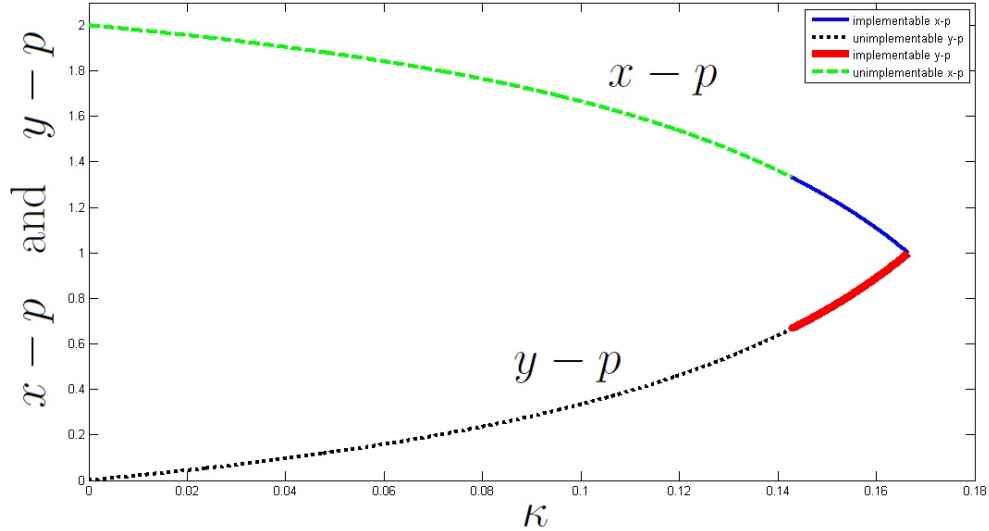


FIGURE 1.3: $x - p$ and $y - p$ traced out by changes in κ on $[0, \frac{1}{6}]$

Let κ gradually increase from 0. Without budget surplus cost the first-best can not be implemented because the seller makes negative profit. As the budget surplus cost κ rises, however, the social planner decreases x and increases y which reduces the budget surplus. As κ reaches $\frac{1}{7}$, the social optimum becomes implementable because the social planner cuts the budget surplus to the extent that the seller collects enough to be willing to participate. This continues all the way to $\kappa = \frac{1}{6}$

where the budget is always balanced. Therefore the social optimum is implementable for $\kappa \in [\frac{1}{7}, \frac{1}{6}]$. At $\kappa = \frac{1}{6}$ the optimal regulation is $x = y = p + 1$ which concurs with the constrained optimal law when the budget is forced to be balanced. Figure 1.3 plots the optimal regulation against κ with the implementable part on $[\frac{1}{7}, \frac{1}{6}]$ solid and the unimplementable part on $[0, \frac{1}{7})$ broken. x drops and y rises with κ until eventually they converge at $\kappa = \frac{1}{6}$. It is surprising and ironic that higher budget surplus cost makes market participants better off, the market surplus higher, and the social optimum more implementable.

1.6 One-Stop Seller

In the process of sorting the buyer, the seller effectively gives up the opportunity to sell or earn income when he advises the buyer away from his product, and thus may cheat if being honest means losing customers. Hence an incentive for the seller to sell indiscriminately is the fear of losing revenue when the good does not appear to suit the buyer.

A natural question is whether the seller's bad incentive would be mitigated if he does not have to worry about losing revenue. If yes, then presumably a less aggressive regulation can restore efficiency—or can it? In this section we investigate such issues by extending our model so that the seller has different products or services to address the need of each type of buyer. The comparison of this extension with the standard model is a comparison between a highly integrated one-stop seller versus a specialized one.

Suppose the seller sells two types of good, type l and type h , which suit a buyer of type l and type h respectively. The buyer gets utility h if he buys a good that matches his type and gets l if he buys a good that does not match his type. The buyer is ex ante equally likely to be of either type.⁵ The seller's private signal \tilde{s} is

⁵ Asymmetric type distribution does not bring more economic insight but brings unnecessary

generated from type $\tilde{\tau}$ by $Pr(\tilde{s} = \tilde{\tau}|\tilde{\tau}) = \frac{1+f}{2}$ and $Pr(\tilde{s} \neq \tilde{\tau}|\tilde{\tau}) = \frac{1-f}{2}$. Upon learning \tilde{s} , the seller updates his belief to $Pr(\tilde{\tau} = \tilde{s}|\tilde{s}) = \frac{1+f}{2}$ and $Pr(\tilde{\tau} \neq \tilde{s}|\tilde{s}) = \frac{1-f}{2}$. Table 1.2 shows the marginal and joint distributions of \tilde{s} and $\tilde{\tau}$.

Table 1.2: Marginal and Joint Distributions of \tilde{s} and $\tilde{\tau}$ (One-Stop Seller)

		$Pr(\tilde{\tau} = l)$	$Pr(\tilde{\tau} = h)$
		$\frac{1}{2}$	$\frac{1}{2}$
$Pr(\tilde{s} = l)$	$\frac{1}{2}$	$\frac{1+f}{4}$	$\frac{1-f}{4}$
$Pr(\tilde{s} = h)$	$\frac{1}{2}$	$\frac{1-f}{4}$	$\frac{1+f}{4}$

Now when the seller sells he must in addition decide which type of good to sell. Let λ_{lh} be the probability of selling the type l good when $\tilde{s} = h$. The definitions of λ_{ll} , λ_{hl} , and λ_{hh} are similar. We impose the feasible sales constraints $\lambda_{ll} + \lambda_{hl} \leq 1$ and $\lambda_{hh} + \lambda_{lh} \leq 1$ so that the sales probability given the signal never exceeds 1. Let the social welfare be $w(a, e, \lambda_{hh}, \lambda_{hl}, \lambda_{lh}, \lambda_{ll})$.

Lemma 5. (The First-Best for One-Stop Seller)

Let (a_2^r, e_2^r) maximize $w(a, e, 1, 0, 0, 1)$.

If $w(a_2^r, e_2^r, 1, 0, 0, 1) < 0$, then $(a, e, \lambda_{hh}, \lambda_{hl}, \lambda_{lh}, \lambda_{ll}) = (0, 0, 0, 0, 0, 0)$ is the first-best with zero social welfare. The market breaks down and no law is needed.

If $w(a_2^r, e_2^r, 1, 0, 0, 1) \geq 0$, then $(a_2^r, e_2^r, 1, 0, 0, 1)$ is the first-best.

We assume the more interesting case of $w(a_2^r, e_2^r, 1, 0, 0, 1) \geq 0$ since otherwise No-sale is optimal and the market attains the first-best by itself without law. Let $f_2^r \equiv f(a_2^r, e_2^r)$, $c_2^r \equiv c(e_2^r)$, $k_2^r \equiv k(a_2^r)$. We summarize the result on first-best implementation below.

Proposition 6. (Implementing the First-Best for One-Stop Seller) *If $f_2^r h + (1 - f_2^r)l > c_2^r + k_2^r$ then $x_2^r \equiv h - l > x^r$ restores the first-best $(a_2^r, e_2^r, 1, 0, 0, 1)$ with algebra. Thus we assume symmetric type distribution in this section only.*

an expected budget surplus $(1 - f^r)(h - l)/2$. If $1\{\tilde{\tau} = l, f < f_2^r\}$ can be inferred ex post, then $x_2^b \equiv 1\{f < f_2^r\}(h - l)$ always restores the first-best with no action required from the social planner on the equilibrium path.

Since $x_2^r > x^r$, now the seller has to pay heavier penalty and there is no compensation to the buyer. The law that restores the first-best is harsher on both the seller and the buyer for one-stop purchase than for specialized purchase since the social planner is more convinced of the interaction failure due to lower efforts if there is always something suitable for the buyer. The fact that the law becomes more stringent is consistent with the Chinese government's heavy penalty on Walmart for wrong selling when it had the ability to cater to everyone, yet the regulator did not order compensation for consumers (Bradsher 2011). Without additional information the social planner implements the first-best conditionally due to the budget surplus. If the social planner can tell $f < f_2^r$ from $f \geq f_2^r$ after loss then it can always restore efficiency with a balanced budget.

When the purchase is from a universal seller, what the social planner actually does is more sensitive to the amount of its information than if the purchase is from a specialized seller. If a bit more information is available ex post on whether the first-best levels of effort were exerted, for a one-stop seller the social planner can restore the first-best as if it never interfered. Thus the law is harsher when regulating a universal firm that operates at a highly consolidated level, but the actual action from the regulator is more polarized depending on the amount of its information.

1.7 The Importance of Costly Dual Moral Hazard

In this section we show the importance of the assumption of costly efforts from both the seller and the buyer. If efforts were costless or moral hazard were unilateral or there were no effort choices, then the first-best can always be restored with a balanced

budget. Bilateral moral hazard thus rationalizes punitive penalty and complicates policy-making.

To see the relevance of the costly efforts assumption, suppose $c = k \equiv 0$ so that the buyer and the seller can exert efforts at no cost. This perturbation also has its practical significance and addresses questions such as whether the social planner can do better if efforts were costless or nearly so. Zero effort costs approximate cases where information exchange, learning, and decision making for both parties require little interaction time or hard thinking, such as when the transaction is simple. It may depict the tradeoff of disclosing versus withholding information where a and e stand for transparency or integrity.

To check the importance of the buyer's moral hazard, we examine the difference that the existence of the buyer's effort choice makes. This also addresses the practical concern of what the optimal regulation would be if the the buyer only decides whether to buy. We remove the buyer's effort e from our model so that the detection probability f only depends on the seller's effort a . $f(a)$ is increasing and $f(0) = 0$.

Finally, we look at the role of the buyer and the seller's effort choices altogether. This also addresses the pragmatic curiosity of what the situation would be if the buyer and the seller only decide whether to trade. We thus remove their efforts a and e from our model: $f \in (0, 1]$ becomes a constant.

Proposition 7. (The Importance of Dual Hidden Action) *If efforts are costless, or if there is no moral hazard from the buyer, or if there are no effort choices by the buyer and the seller, then a continuum of laws restore the first-best, and some of them always do so with a balanced budget.*

If any one of the relevant assumptions in costly dual moral hazard is removed, then a contingent transfer from the seller to the buyer can always restore the first-best, and legislation would be simple with no concern for budget surplus or conditional im-

plementation. In fact, since efficiency can always be restored with a balanced budget and no punitive penalty is needed, self-regulation becomes possible and third-party regulation is not strictly required. Therefore costly dual moral hazard is important in explaining punitive penalty and third-party regulation observed in practice. We expect markets that require dual costly efforts in the buyer-product matching process to more likely have third-party regulators and punitive penalty that breaks the balanced budget. Dual moral hazard complicates regulation because the government has to face complex policy issues and tradeoffs.

Costly bilateral hidden action rationalizes the determinacy rather than the arbitrariness of law. In any one of the three less complex cases, the optimal law is indeterminate. However, in the dual moral hazard problem, there is a definite regulation that restores efficiency. Given that retail market and consumer protection regulations have been written into the laws in different countries and regions and a unique policy for compensation and punitive penalty is often repeatedly observed in practice, the predictive power of a retail transaction model is greatly enhanced if in sorting the buyer there is costly hidden action from the buyer as well as from the seller.

Having fully presented our model and its relevance, we observe towards the end that our theory naturally predicts that for markets for specialized products and their sorting services to thrive, for expertise and knowledge on buyer-good matching to be acquired and put to use, and for product specialization to flourish, it is crucial to have well-designed legal rules that govern the retail market and especially the advising and sorting of buyers. Countries and markets with more appropriate, effective, and complete retail transaction laws tend to realize more gains from trades that require the seller's professional advice and buyer-seller interaction to decide on buyer-good matching case by case. The law plays important roles in disciplining participants in retail transactions, improving the effectiveness of seller-buyer interaction, and

injecting mutual confidence and trust into the retail market. The society's expertise and knowledge get efficiently and righteously used under well-designed regulation, which eventually fuels economic efficiency and growth. We should expect positive relationships among the effectiveness of retail market law, the market for specialized products, quality sorting and advising services, the reliability of seller and buyer behavior, the productivity of their interaction, knowledge and expertise on matching goods to buyers, confidence and trust, welfare from trading, and economic efficiency and growth.

1.8 Conclusions

In this paper we have introduced and analyzed an important new friction that the buyer's costly effort, just like the seller's, helps the seller sort the buyer in their interaction. Our model is one of dual moral hazard with asymmetric information. The seller, relying on his expertise on product-buyer matching, can exert effort and interact with the buyer to privately learn about the suitability of a good or service for the buyer. The buyer's effort improves the quality of the seller's private learning since he knows more about his own situation. The pair take respective costly hidden actions that jointly raise the likelihood that the unfit buyer gets identified.

The unique law that restores the first-best stipulates that if the buyer experiences a loss due to mismatch to the product then the seller fully refunds the buyer and pays the regulator a punitive penalty that equals the buyer's loss. Furthermore, we have successfully used contingent penalty for only the seller to always restore the first-best in this team moral hazard problem with a balanced budget on the equilibrium path, where the social planner uses minimal ex post binary information after a loss occurs on whether the joint productivity of the buyer and the seller met the first-best standard.

The law is more stringent for both the seller and the buyer in the case of a

universal seller who can cater to each type of buyer, because a loss is more likely due to shirking and the resultant interaction failure, though the social planner's actual action on the equilibrium path is more sensitive to whether it eventually knows whether the first-best joint efforts were exerted; if yes then the social planner need not act on the equilibrium path.

We analyze three different constrained optimal regulations: when the first-best is not attainable due to the budget surplus, when the budget is forced to be balanced, and when the budget surplus is costly. A cost to the budget surplus can make all market participants better off and the social optimum more implementable.

If efforts are costless or moral hazard is unilateral or there are no effort choices, then no punitive penalty is needed to always restore the first-best with a balanced budget, and a continuum of laws restore efficiency and regulation becomes simple and flexible, including self-regulation. Costly bilateral moral hazard thus rationalizes punitive penalty, third-party regulation and the determinacy of law, yet complicates regulation.

Our model has diverse applications to goods, services, and non-commercial sectors for countries, organizations, and individuals. We expect positive relationships among apt retail market law, effective buyer-seller interaction, responsible traders' behavior, advising quality, product specialization, knowledge on buyer-good matching, confidence and trust, gains from trade, and economic efficiency and growth.

Disclosure by Contingent Non-Disclosure

2.1 Introduction

No news is news, and no information conveys information. Unfortunately our society seems intolerant of non-disclosure. The Watergate scandal led to the only resignation of a U.S. president to date and the indictment of dozens of top officials. More recently, Edward Snowden's leaks (Mazzetti and Schmidt 2013) put the legitimacy of the nation's secret surveillance programs under debate, and New York had to terminate its surveillance program (Apuzzo and Goldstein 2014). Indeed, transparency and disclosure requirements are normally considered welfare-improving for citizens and stakeholders. However, we show that contingent non-disclosure can be socially optimal, if only for the ultimate purpose of more efficient disclosure.

Our paper concerns how the benevolent sender can convey useful information to receivers to help them with a decision. Receivers' choice issue could be whether to buy an asset or not, whether to take a certain option or not, or whether to behave in a certain way or not, etc. We identify and explore a new benefit of non-disclosure when the sender's unobservable status complicates the situation, so

that with full disclosure efficiency calls for the transmission of more information than can be conveyed directly. The sender optimally mobilizes an extra channel of communication by choosing carefully when not to disclose. Contingent non-disclosure is the only way to attain efficiency when direct communication is congested.

The sender may not always be active and this is unobservable to receivers. Empirically, we seldom see governments or other entities constantly making non-news announcements like “we haven’t heard anything but we are active”, even if they are sometimes inactive. In practice this can be due to communication congestion or constraints. For example, announcements are scarce resources and not everything can make public news, and citizens may prefer not to be swamped or distracted by less important announcements if the sender can make do without them.

When the sender is active, he publicly announces his own official signal if any, and decides whether receivers hear the unofficial signal or rumor if any. Both signals warn against a particular choice. Official warnings are natural in a nation or market where the authority, an organization or a leader intervenes when there is a problem.

The sender can manipulate the rumor by contingently shutting it down. In the optimal equilibrium, he purges the rumor when he disagrees with it. It is surprising that the sender optimally withholds useful information though his welfare is aligned with receivers’. By doing so he breaks a pooled information set containing events that call for different optimal decisions, which enables receivers to distinguish them and optimize in each event. Therefore contrary to the equilibrium behavior of a biased sender who would add noise to the original state variable (Crawford and Sobel 1982), an unbiased sender has the opposite incentive to reduce noise.

The events then get repooled with the same optimal decision applicable to all events in each information set, so that the loss of information due to communication congestion does not harm welfare. In a way, the sender maximizes receivers’ welfare subject to communication constraints, and optimally trades off more welfare-relevant

information for less welfare-relevant information to communicate to receivers.

Specifically, without the sender's intervention the receivers always makes the wrong decision in some event because they are not sure if the sender is active. The sender purges the rumor contingently so the receivers are better informed of the sender's status. However, doing so brings the receivers the new uncertainty of whether the rumor has been suppressed or if it never existed. Fortunately, the lack of this knowledge has no bearing on decision-making or welfare. Hence the sender is willing to introduce new uncertainty in order to resolve the key uncertainty that negatively impacts welfare.

Our theory has wide applications in law, public policy, and political economy. They relate to the roles of governments and other entities such as non-for-profit organizations and labor unions in caring for a group of citizens or members who face a decision involving risk. If citizens receive no warning, they can not tell whether the authority is inactive or is active but does not detect any problem with the situation, such as whether a situation is uninspected or inspected to have met safety standards. For instance, they do not know whether the government approves financial innovation like adjustable-rate mortgages for their purchase, or whether they should stay or leave with possible rumor of terrorism, epidemics, nuclear leaks, or natural disasters, use medicine and surgery that may be said to have side effects, support a health care reform or a possibly corrupt government, join in exciting but potentially dangerous sports or activities, buy weapons and ammunition or support their regulation, eat possibly unhealthy food, drink alcohol, and watch R movies, use products without information on expiration date, accept innovation and progress involving clashes of ideology such as homosexual marriage and experiments with human or animal subjects, etc.

Our paper differs from the existing economics, finance, and accounting literatures on information and disclosure. Grossman (1981), Milgrom (1981), and Grossman and

Hart (1980) show the famous unraveling result where firms fully disclose voluntarily to separate from worse types. Yet there are known situations in the literature where more information and disclosure hurts. Our paper is very different because more information is beneficial in the sense that knowing the sender's status is helpful and two information sources are better than one for receivers. One of our novel contributions is that even though more information always helps receivers, it is still optimal to contingently block some information from them.

Knowing more in Hirshleifer (1971) defeats the purpose of risk management for risk-averse agents, but our agents can be risk-neutral, and inference rather than risk is central to their welfare. Dye (1985a, 1985b, 1986, 1990, 1998) offers multiple reasons for firms' non-disclosure involving investors' knowledge of managerial endowment of information, the mixture of proprietary and nonproprietary information, the conflict between shareholders and managers, the impact of proprietary information, the substitutability of mandatory and voluntary disclosures, strategic accounting choices, externalities, and investor sophistication. Verrecchia (1983, 1990) shows that firms do not disclose as their proprietary disclosure cost extends traders' rational interpretation of non-disclosure. Dye and Sridhar (1995) and Acharya et al (2011) show how firms may delay disclosure due to other firms. Our paper is very different from these because our sender is the social planner who maximizes receivers' welfare.

Bikhchandani et al (1992), Burguet and Vives (2000), Amador and Weill (2012), Carlin, Gervais and Manso (2013) show that public disclosure may hinder the production of information when agents can learn from one another. In our paper all agents receive the same information so there is no mutual learning. Morris and Shin (2002) and Angeletos and Pavan (2007) show how public disclosure may crowd out private information and cause excessive coordination. Ours is not a paper on coordination, and our receivers of public information have no private information of their own. Edmans et al (2014) show that less disclosure creates value by alleviating

firms' distorted incentive to overwork on improving hard information but ignore soft information. Gervais and Strobl (2014) offer a signaling theory of why talented fund managers may optimally choose opaque funds.

Our paper relates to Teoh (1997) where the team leader in a public-good game optimally avoids disclosure when good news does not benefit as much as how bad news hurts in coordination failures. Multiple agents are necessary for Teoh (1997) but our paper works for any number of agents. Non-disclosure is ex ante efficient in Teoh (1997) before information arrives, while our social planner interferes after getting available information. Unlike Teoh (1997), we do not need costly contribution so each agent in our model operates independently. Ours is not on coordination failure as Teoh (1997) but on disclosure efficiency when sender status is unobservable and there are multiple information sources.

The rest of the paper proceeds as follows. The next section presents our model setup. Section 3 derives the agents' optimal decision rule and the first-best. Section 4 shows the inefficiency of truth-telling and transparency. Section 5 proves the efficiency of contingent non-disclosure by purging dissident rumor. Section 6 gives the predictions and policy implications of our theory. Section 7 shows the robustness of our model. The final section concludes.

2.2 The Model

A group of homogeneous agents each decides whether to take an option \tilde{x} , which may be good or bad for them. Let $\tilde{x} = g$ if \tilde{x} is good and $\tilde{x} = b$ if \tilde{x} is bad. The value of \tilde{x} is unobservable. The prior for \tilde{x} is $Pr(\tilde{x} = g) = Pr(\tilde{x} = b) = \frac{1}{2}$. We normalize to zero each agent's payoff from not having \tilde{x} . An agent gains 1 from having a good \tilde{x} but loses $c > 0$ from having a bad \tilde{x} . Each agent maximizes his expected payoff by either choosing \tilde{x} or avoiding it. Examples of \tilde{x} abound in practice. For instance, in business, \tilde{x} could be a financial innovation such as a new type of investment or

mortgage. In political economy, \tilde{x} could be a new health policy.

The social planner who maximizes the welfare of these agents is very busy taking care of many things. Independent of all other randomness, with probability $q \in (0, 1)$ the social planner becomes aware of and hence active in the agents' present choice issue concerning \tilde{x} , but with probability $1 - q$ it does not. We write a if the social planner is active and i if it is inactive. q may be interpreted as attentiveness, skill, time efficiency, judgment, responsibility, accountability, etc.

Two sources of information may aid an agent in his choice. First, if the social planner is active, there may be an official warning w that \tilde{x} is bad. w appears according to $Pr(w | \tilde{x} = b) = p > \frac{1}{2}$ and $Pr(w | \tilde{x} = g) = 1 - p > 0$, where $p \in (\frac{1}{2}, 1)$ can be interpreted as the quality of information. The social planner must report w to the agents whenever w exists¹. Second, there may be a rumor r that \tilde{x} is bad. For example, r may say that a new product is defective. r appears according to $Pr(r | \tilde{x} = b) = p$ and $Pr(r | \tilde{x} = g) = 1 - p$. The existence of w and r are conditionally independent given \tilde{x} . If r exists, the social planner is the first to hear it, and maximizes agents' welfare by choosing to either squash r so that no agent will hear it or let r flow freely to all agents. To save notation we write $Pr(r | \tilde{x} = b)$ as $Pr(r | b)$, and likewise elsewhere.

The sequence of events is as follows. First, Nature determines \tilde{x} . Second, the social planner is either active or not, which is unobservable to agents. If the social planner is inactive, then it is out of the game from this point onward, so there will be no official warning w , and a rumor r reaches all agents if it exists. If the social planner is active, it issues w to agents whenever w exists, and when it hears r it decides whether to purge it so that agents do not hear r . Finally, agents decide whether to accept or reject \tilde{x} , and their payoffs are realized.

The underlying state in our model describes the resolution of four uncertain-

¹ Endogenizing the social planner's decision to report w does not affect our results.

ties: the value of \tilde{x} , the social planner's status, and the existence of w and r . To denote a particular state, we stack together our notations for the four uncertainties in above order. For example, $gawr$ means \tilde{x} is good, the social planner is active, and both w and r exist. bi means \tilde{x} is bad, the social planner is inactive (so there is no w), and there is no r . Thus the state space of our model is $\Omega \equiv \{gawr, gaw, gar, ga, gir, gi, bawr, baw, bar, ba, bir, bi\}$.

Before payoffs are realized, the social planner and the agents can not differentiate the states that differ only by the value of \tilde{x} . For example, they can not tell gaw from baw , or bir from gir . To save on notation, we combine each pair of such states into an event: $awr \equiv \{gawr, bawr\}$, $aw \equiv \{gaw, baw\}$, $ar \equiv \{gar, bar\}$, $a \equiv \{ga, ba\}$, $ir \equiv \{gir, bir\}$, $i \equiv \{gi, bi\}$. One and only one of these six events will occur. $\Lambda \equiv \{\{awr\}, \{aw\}, \{ar\}, \{a\}, \{ir\}, \{i\}\}$ will be a useful partition in our analysis.

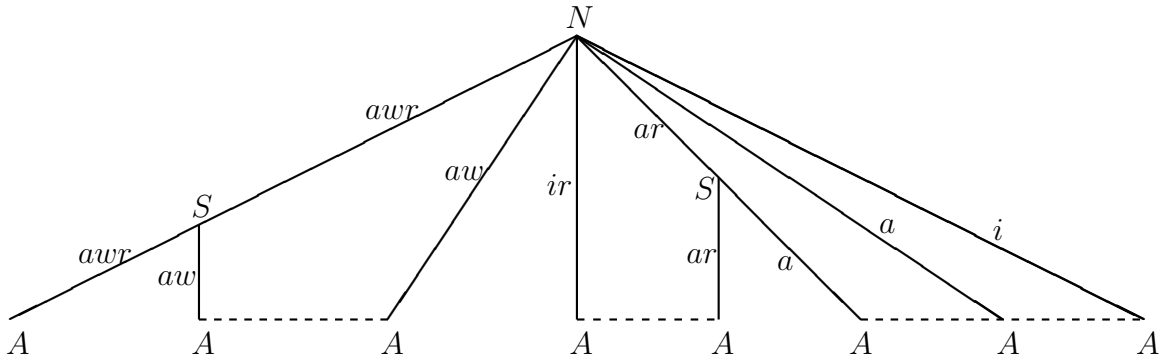


FIGURE 2.1: A Simplified Game Tree

Figure 2.1 shows a simplified game tree. N denotes Nature, who controls the four uncertainties in our model. S denotes the social planner, who when active may quell r . There are two such nodes in events awr and ar where S may act. A denotes agents, who can not tell whether S is active or not unless w exists. If agents hear no r , they can not tell ceteris paribus whether r has been suppressed or it never existed. Due to such information asymmetry between the social planner and the agents, the agents have non-singleton information sets in the game tree denoted by the broken

segments. Each agent makes his decision on \tilde{x} in the end.

2.3 Optimal Decision and the First-Best

Given any information set \mathcal{I} , an agent's expected payoff from having \tilde{x} is

$$[1 - Pr(b|\mathcal{I})] \times 1 - Pr(b|\mathcal{I}) \times c = 1 - (1 + c)Pr(b|\mathcal{I}) \quad (2.1)$$

which is positive and hence better than the status quo of not having \tilde{x} if

$$Pr(b|\mathcal{I}) < p^* \equiv \frac{1}{1 + c} \quad (2.2)$$

So agents optimally choose \tilde{x} if their belief that \tilde{x} is bad is below the critical level p^* ; otherwise they avoid \tilde{x} .

We take \mathcal{I} as a subset of Λ . The smaller the set \mathcal{I} , the more refined the conditioning information. If the prior belief is all the information there is, then $\mathcal{I} = \Lambda$, so $Pr(b|\Lambda) = Pr(b) = \frac{1}{2}$. If $p^* < \frac{1}{2}$, then agents reject \tilde{x} based on only the prior. Since there can only be bad news about \tilde{x} down the road, namely w and r , the optimality of such rejection is never compromised, so agents can decisively and permanently reject \tilde{x} right from the start. We thus assume the more interesting case of $p^* > \frac{1}{2}$ so that it is optimal for agents to accept \tilde{x} if they rely only on the prior belief. This is when future bad news may affect agents' decision.

If new information is not accurate enough to have a sufficient impact on the belief and thus welfare, then agents' optimal decision based only on the prior will be little affected by the arrival of w or r . For our model to be interesting, the quality of information p has to be sufficiently high: $p > p^*$. In sum, the maintained assumption in this paper is

$$p^* \in \left(\frac{1}{2}, p \right) \quad (2.3)$$

or equivalently $c < 1 < (1 + c)p$ in terms of c .

We now find the first-best decision that maximizes the agents' welfare given each of the six elemental events in Λ . Table 2.1 lists some useful information: the six events under header E , their probabilities of occurring $Pr(E)$, the conditional probabilities $Pr(b|E)$ that \tilde{x} is bad given each event E in Λ , their comparison with the critical belief p^* , and whether the welfare from having \tilde{x} is positive given E . The derivation of these results are relegated to the appendix. The agents' welfare from \tilde{x} is negative in events awr and ir but positive in every other event.

Table 2.1: The Agents' First-Best Decision in Each Event in Λ

E	$Pr(E)$	$Pr(b E)$	vs p^*	Welfare
awr	$\frac{1}{2} [p^2 + (1-p)^2] q$	$\frac{p^2}{p^2+(1-p)^2}$	$> p^*$	-
aw	$(1-p)pq$	$\frac{1}{2}$	$< p^*$	+
ir	$\frac{1}{2}(1-q)$	p	$> p^*$	-
ar	$(1-p)pq$	$\frac{1}{2}$	$< p^*$	+
i	$\frac{1}{2}(1-q)$	$1-p$	$< p^*$	+
a	$\frac{1}{2} [p^2 + (1-p)^2] q$	$\frac{(1-p)^2}{p^2+(1-p)^2}$	$< p^*$	+

Proposition 8. *The first-best is for agents to avoid \tilde{x} in events awr and ir , but to choose \tilde{x} otherwise.*

However, agents can not tell ir from ar because in both events they only hear the rumor r and are not sure if the social planner is active. Since the optimal decisions for these two events differ, a potential inefficiency due to the inference problem looms.

The social planner is only aware of the four events in which it is active, so its information structure is $\Lambda_S = \{\{awr\}, \{aw\}, \{ar\}, \{a\}\}$. When the social planner forms a belief, it knows it is active in addition to knowing the status of w and r . Hence in these four events the social planner's information set and belief equal the ones under Λ . We thus have

Lemma 9. (The Social Planner's Belief) *Let the social planner's information set be \mathcal{I}_S . \mathcal{I}_S is defined and the social planner has a belief $Pr(b|\mathcal{I}_S)$ if and only if it*

is active. The social planner can perfectly distinguish the four events awr , aw , ar , a in which it is active. If $E \in \{awr, aw, ar, a\}$ then $\mathcal{I}_S = E$ and $Pr(b | \mathcal{I}_S) = Pr(b | E)$.

2.4 Beliefs Under Full Disclosure

Let the agents' information set be \mathcal{I}_A when an event in Λ occurs. Suppose the social planner acts honestly and never interferes with the rumor. Then agents hear r whenever it exists. Below we compute agents' beliefs.

Lemma 10. *If the social planner always discloses r , then:*

(a). *Agents can perfectly distinguish awr and aw . If $E \in \{awr, aw\}$ then $\mathcal{I}_A = E$ and $Pr(b | \mathcal{I}_A) = Pr(b | E)$.*

(b). *If agents hear r but see no w , then $\mathcal{I}_A = \{ir, ar\}$. Their belief updates to*

$$Pr(b | \{ir, ar\}) = \frac{(1-pq)p}{1-q+2(1-p)pq} \in \left(\frac{1}{2}, p\right). \quad (2.4)$$

$Pr(b | \{ir, ar\})$ is decreasing in q with supremum $Pr(b | ir) = p$ attained as $q \rightarrow 0$ and infimum $Pr(b | ar) = \frac{1}{2}$ attained as $q \rightarrow 1$. $Pr(b | \{ir, ar\})$ is increasing in p with supremum $Pr(b | b) = 1$ attained as $p \rightarrow 1$ and infimum $Pr(b | \Lambda) = \frac{1}{2}$ attained as $p \rightarrow \frac{1}{2}$.

(c). *If agents see nothing, then $\mathcal{I}_A = \{i, a\}$. Their belief updates to*

$$Pr(b | \{i, a\}) = \frac{(1-pq)(1-p)}{1-2(1-p)pq} \in \left(\frac{(1-p)^2}{p^2 + (1-p)^2}, 1-p\right). \quad (2.5)$$

$Pr(b | \{i, a\})$ is decreasing in q with supremum $Pr(b | i) = 1-p$ attained as $q \rightarrow 0$ and infimum $Pr(b | a)$ attained as $q \rightarrow 1$. $Pr(b | \{i, a\})$ is decreasing in p with supremum $Pr(b | \Lambda) = \frac{1}{2}$ attained as $p \rightarrow \frac{1}{2}$ and infimum $Pr(b | g) = 0$ attained as $p \rightarrow 1$.

Therefore agents' information structure under full disclosure is $\Lambda_A = \{\{awr\}, \{aw\}, \{ir, ar\}, \{i, a\}\}$. Λ_A is cruder than Λ since agents do not know whether the social planner is active when they do not see w .

The comparative statics are intuitive. Agents' belief that \tilde{x} is bad given no w is lower if the social planner is more likely active, because the absence of w more likely means no w exists and hence a good \tilde{x} rather than the social planner's inactivity. Agents' belief given r but no w rises with the quality of information p , as the negative information r is more potent if information is more accurate. Agents' belief given nothing is lower if information is of higher quality, because then no bad news is better news. Interpretations of the limiting results are provided in the appendix.

Table 2.2: Information Sets and Beliefs Under Full Disclosure

E	\mathcal{I}_S	to A	\mathcal{I}_A	$Pr(b E)$	$Pr(b \mathcal{I}_S)$	$Pr(b \mathcal{I}_A)$
awr	awr	wr	awr	$\frac{p^2}{p^2+(1-p)^2}$	$\frac{p^2}{p^2+(1-p)^2}$	$\frac{p^2}{p^2+(1-p)^2}$
aw	aw	w	aw	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
ir	n/a	r	$\{ir, ar\}$	p	n/a	$\in (\frac{1}{2}, p)$
ar	ar	r	$\{ir, ar\}$	$\frac{1}{2}$	$\frac{1}{2}$	$\in (\frac{1}{2}, p)$
i	n/a		$\{i, a\}$	$1-p$	n/a	$< 1-p$
a	a		$\{i, a\}$	$\frac{(1-p)^2}{p^2+(1-p)^2}$	$\frac{(1-p)^2}{p^2+(1-p)^2}$	$< 1-p$

Table 2.2 shows the social planner's and the agents' information sets and beliefs under full disclosure where the rumor reaches the agents freely. Column 3 labeled "to A " describes what agents see. Here, wr means agents see w and r ; w means agents see w but not r ; r means agents hear r but not w ; otherwise agents see nothing. We mark the rows that correspond to the non-singleton information sets $\{ir, ar\}$ and $\{i, a\}$ red and blue respectively.

Figure 2.2 shows the equilibrium path if the social planner never withholds the rumor r at its two decision points in events awr and ar . The figure also illustrates agents' information sets that confirm their information structure Λ_A . Again, any-

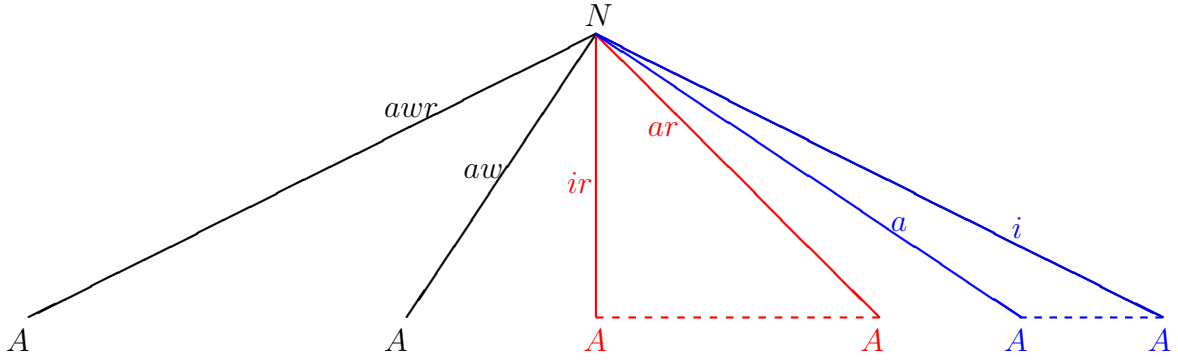


FIGURE 2.2: The Equilibrium and Agents' Information Sets Under Full Disclosure

thing associated with $\{ir, ar\}$ and $\{i, a\}$ are colored red and blue respectively.

2.5 The Inefficiency of Full Disclosure

Agents can perfectly tell events awr and aw from all others, so they hold belief $Pr(b|E)$ for $E \in \{awr, aw\}$ and make the first-best choice in these events. Though agents can not tell event i from a , they can still make the first-best choice of accepting \tilde{x} because this is the first-best in both events.

Inefficiency occurs when agents hear r but observe no w . Agents can not distinguish events ir and ar , which pooled to form their information set $\mathcal{I}_A = \{ir, ar\}$. Recall the agents' belief $Pr(b|\{ir, ar\}) \in (\frac{1}{2}, p)$ and the critical belief $p^* \in (\frac{1}{2}, p)$. Therefore it is possible for agents' belief to be either above or below p^* . We discuss these two parametrization possibilities and show that either way, there exists an inefficiency.

Table 2.3 lists our findings so far on the agents' optimal decision when the social planner never suppresses r . The comparison of agents' belief $Pr(b|\{ir, ar\})$ with p^* and the associated welfare from having \tilde{x} are for the moment unclear. We now fill in the question marks.

Proposition 11. (The Inefficiency of Full Disclosure)

Table 2.3: The Agents' Optimal Decision Under Full Disclosure

E	to A	\mathcal{I}_A	$Pr(b \mathcal{I}_A)$	vs p^*	Welfare
awr	wr	awr	$\frac{p^2}{p^2+(1-p)^2}$	$> p^*$	–
aw	w	aw	$\frac{1}{2}$	$< p^*$	+
ir	r	$\{ir, ar\}$	$\in (\frac{1}{2}, p)$?	?
ar	r	$\{ir, ar\}$	$\in (\frac{1}{2}, p)$?	?
i		$\{i, a\}$	$< 1-p$	$< p^*$	+
a		$\{i, a\}$	$< 1-p$	$< p^*$	+

There is always an inefficiency if the social planner never suppresses r .

If $Pr(b|\{ir, ar\}) < p^*$, then agents optimally accept \tilde{x} when they hear r but see no w . However, in event ir they are better off rejecting \tilde{x} instead. Their welfare loss compared to the first-best is

$$L_{<} = \frac{(1-q)[(c+1)p-1]}{2} > 0 \quad (2.6)$$

which is increasing in p and decreasing in q .

If $Pr(b|\{ir, ar\}) > p^*$ then agents optimally avoid \tilde{x} when they hear r but see no w . However, in event ar they are better off accepting \tilde{x} instead. Their welfare loss compared to the first-best is

$$L_{>} = \frac{(1-c)(1-p)pq}{2} > 0 \quad (2.7)$$

which is decreasing in p and increasing in q .

The higher the quality of information, the worse the agents fare without information, and the more severe the welfare loss $L_{<}$. The less likely the social planner is active, the more likely that event ir occurs in which the wrong decision lead to $L_{<}$. The welfare loss $L_{>}$ is greater if the likelihood of event ar is higher in which the wrong choice is made. The lower the quality of information, the more likely that the two information sources give contradicting signals, making ar more likely. Higher likelihood that the social planner is active also enhances the chance of ar .

We next show how the social planner can avoid such welfare losses by choosing the right occasion to preemptively eliminate the rumor.

2.6 The Efficiency of Contingent Non-Disclosure

In this section we analyze how the social planner can potentially suppress the rumor to improve efficiency. The social planner has two decision points. First, in event awr where the social planner issues the official warning w and hears the rumor r . The social planner's information agrees with the rumor in this case. Second, in event ar where the social planner hears r but has no w . The social planner disagrees with the rumor in this case.

The social planner's strategy we investigate is one where it only quells r in event ar but not in event awr . Because the social planner effectively purges the rumor if and only if its own information source that generates the official warning disagrees with the rumor, we call such contingent non-disclosure policy "purging dissident rumor". We will see how quelling r in event ar improves efficiency by helping agents distinguish events ir and ar . As a first step, we compute agents' beliefs when the social planner intervenes this way.

Lemma 12. *Suppose the social planner quiets r only in event ar .*

- (a). *Agents can perfectly distinguish events ir , awr , and aw . If $E \in \{ir, awr, aw\}$ then $\mathcal{I}_A = E$ and $Pr(b | \mathcal{I}_A) = Pr(b | E)$.*
- (b). *If agents see nothing, then $\mathcal{I}_A = \{ar, i, a\}$ so they believe $Pr(b | \{ar, i, a\}) = 1 - p < \frac{1}{2}$, which is decreasing in p with its supremum $\frac{1}{2}$ attained as $p \rightarrow \frac{1}{2}$ and its infimum 0 attained as $p \rightarrow 1$.*

Therefore agents' information structure when the social planner purges dissident rumor is $\Lambda'_A = \{\{awr\}, \{aw\}, \{ir\}, \{ar, i, a\}\}$, which is still cruder than Λ . When agents see nothing, they do not know whether r has been suppressed by the social

planner or it never existed in the first place, neither do they know whether the social planner is active.

Table 2.4: The Agents' Optimal Decision Under Contingent Non-Disclosure

E	\mathcal{I}_S	to A	\mathcal{I}_A	$Pr(b E)$	$Pr(b \mathcal{I}_A)$	vs p^*	Welfare
awr	awr	wr	awr	$\frac{p^2}{p^2+(1-p)^2}$	$\frac{p^2}{p^2+(1-p)^2}$	$> p^*$	-
aw	aw	w	aw	$\frac{1}{2}$	$\frac{1}{2}$	$< p^*$	+
ir	n/a	r	ir	p	p	$> p^*$	-
ar	ar		$\{ar, i, a\}$	$\frac{1}{2}$	$1 - p$	$< p^*$	+
i	n/a		$\{ar, i, a\}$	$1 - p$	$1 - p$	$< p^*$	+
a	a		$\{ar, i, a\}$	$\frac{(1-p)^2}{p^2+(1-p)^2}$	$1 - p$	$< p^*$	+

Table 2.4 summarizes our findings on agents' optimal decision if the social planner purges dissident rumor. We color the three rows associated with agents' only non-singleton information set $\{ar, i, a\}$ blue. The agents now make the first-best decision in each event. This leads to our main proposition below.

Proposition 13. *The social planner can implement the first-best by the contingent non-disclosure policy of purging dissident rumor.*

It is a good place to pause and ponder why the magic worked. The cause of the inefficiency was that when the social planner acted honestly and never suppressed the rumor r , agents could not distinguish events ir and ar from each other. The problem is that agents' first-best decisions are different in events ir and ar since the two are heterogeneous enough in nature. Hence when ir or ar occurs, agents always make the wrong decision in one event irrespective of what they do. Therefore we have an inefficiency due to agents' inference problem.

The social planner, by purging the dissident rumor, can help agents tell event ir from ar . Now agents know that whenever they hear r but observe no w , ir must have occurred. This helps conquer the loss of information that previously lead to confusion and inefficiency. Therefore the social planner, by being untruthful to

agents, successfully breaks their information set $\{ir, ar\}$ into two singletons ir and ar . One of the two, ar , then gets re-pooled to a different information set $\{i, a\}$, because agents see nothing in each of the three events ar , i , and a . What is nice about the newly formed information set $\{ar, i, a\}$ is that the three events are sufficiently similar or homogeneous to call for the same first-best decision. Therefore the loss of information for agents in this new pooling does not cause any inefficiency.

The pooling of the three events ar , i , and a illustrates an interesting and realistic phenomenon where agents can not tell whether the rumor r has been suppressed (in event ar) or there exists no rumor in the first place (in events i and a). Our result shows that honesty is not the best policy, which justifies the many lies in this world told by governments and organizations out of the best intention that probably have indeed helped.

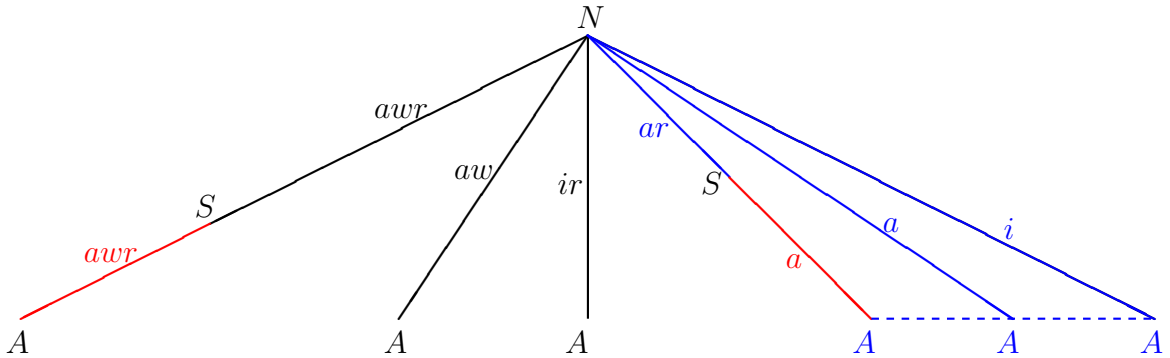


FIGURE 2.3: The Equilibrium and Agents' Information Sets Under Contingent Non-Disclosure

Figure 2.3 shows the equilibrium if the social planner purges dissident rumor r . On event awr , r is let go freely to reach all agents; on event ar , r is purged and agents never get to hear it, so that to them event ar becomes indistinguishable from events a and i . The figure also displays agents' information sets which confirm their information structure Λ'_A . We dye blue anything associated with $\{ar, i, a\}$. The segments following the social planner's decision points are dyed red for emphasis.

In practice only a limited amount of information that are important enough can get a chance to be directly publicized or announced. In the age of information explosion, there are myriads of diverse information sources to learn from, which complicate information and cost enormously to convey them all. Also, there may be a million matters for the social planner to attend to, yet at each given time it can only randomly become aware of and actively engage in some of them. Then it may be prohibitively costly for the busy social planner to constantly notify its status on each duty to the public, or the public may have insufficient attention to receive all the continual notifications or may face exorbitant costs to digest them. Multiple information sources and unobservable sender's status represent complex asymmetric information in practice whose direct communication is congested or constrained, which confuses the agents. Facing communication constraints the social planner optimally uses contingent non-disclosure to give up transmitting welfare-irrelevant information in order that welfare-relevant information gets conveyed to the public. Conversely, if the social planner can find a way to convey all the welfare-relevant information to agents for free by contingent non-disclosure anyway, why would it worry about huge direct communication costs or tight communication constraints?

2.7 Predictions

Our theory has an interesting policy implication in the realms of political economy, public policy, and law. It helps for the social planner to have its agents under surveillance so that it can know as much as possible, such as the rumor, and do so as early as possible. In this way it has the potential to restore the efficient outcome by appropriately blocking information from the public. The government should also be endowed the power to manipulate information such as purging dissident rumor in order to attain the first-best. This is related to surveillance programs at the national (prism and Edward Snowden, see Mazetti and Schmidt 2013), federal (the Watergate

scandal), and local (New York government, see Apuzzo and Goldstein 2014) levels. What we have essentially shown is that it can be optimal for a government, even a liberal and democratic one, to use such surveillance programs, refrain from truth-telling or disclosure, and suppress dissident voices or opinions in order to restore efficient communication. Surveillance programs such as the prism can thus serve efficiency purposes beyond national security. We have therefore rationalized autocracy, dishonesty, non-democracy, unfairness, tyranny oppression, and persecution, etc.

Implication 14. *It can be optimal for the government to use surveillance and withhold information from the public such as by suppressing dissident rumor.*

In practice, some social planners are unable to withhold whatever information whenever they want to. This could be due to disclosure regulations or the difficulties in implementation. For example, democratic, transparent, and truthful governments tend to have less leeway to rig information due to the nature of democracy and related customs, laws and constitutions that require immediately and full disclosure. In many countries the citizens' right to know is considered one of their basic rights. Moreover, larger countries with larger populations, decentralized organizations, and better communication channels may find it harder to block rumors before any citizens know or spread rumors. However, citizens of governments that are better able to manipulate information benefit more from the type of indirect communication that we have identified. Autocratic, opaque, and dishonest regimes have more such flexibility and often manage to withhold or delay reporting relevant information. Therefore we have the unconventional prediction below, which also applies to non-for-profit organizations and labor unions whose interests are well-aligned with their members.

Prediction 15. *Citizens of non-democratic, opaque, dishonest, small, and more centralized countries (cross-sectional) or periods (time-series) with lower disclosure*

requirements for the government benefit from the government's effective communication by contingent non-disclosure like the purging of dissident rumors. Citizens of democratic, transparent, truthful, large, and less centralized countries tend to be confused by complex information and make mistakes. Democracy, transparency, honesty and disclosure laws impair informational efficiency and diminish welfare. The contrast is more apparent in the age or places of high information complexity or explosion.

Our next prediction is derived from the comparative statics with respect to q and p . Although our main proposition on how to attain efficiency is robust to the values of q and p , for firms that are unable to always attain efficiency, how they are likely to err does depend on these probabilities. For example, for purchase decisions of say assets, the direction and magnitude of errors can be observed ex post.

Suppose a social planner is unable to manipulate information to influence agents' choice. Recall that agents' decision is inefficient in events ir and ar , and their belief $Pr(b|\{ir, ar\})$ is decreasing in q and increasing in p . Therefore by their optimal decision rule, agents are ceteris paribus more likely to err against the direction of the social planner's news if the social planner is more likely active or if the news is of lower quality. This is intuitive, because when the social planner is more active then when agents do not hear official news they tend to believe that there is no news rather than that the social planner is inactive and hence err against the direction of its news. When the news is less potent, agents are less worried about making a mistake against the social planner's news in case it is inactive, and are thus more likely to err this way.

Recall that when $Pr(b|\{ir, ar\}) < p^*$ and agents go against the social planner's potential news in event $\{ir, ar\}$, they err in event ir whose probability is $\frac{1-q}{2}$ which is decreasing in q . The welfare loss in event ir is $(1+c)p - 1$, which is increasing

in p . When $Pr(b|\{ir, ar\}) > p^*$ so that agents go in the direction of the social planner's news in event $\{ir, ar\}$, the chance to err is $Pr(ar) = (1 - p)pq$, which is decreasing in $p \in (\frac{1}{2}, 1)$ and increasing in q . We already provided interpretation to these comparative statics earlier when analyzing the inefficiency of full disclosure. They allow us to derive more empirical predictions.

Social planners who are busier are less likely active in their agents' choice issues. Governments who are constantly in war state are less able to care for its citizens. Governments with larger populations like China and India are less likely to have the time to engage in its citizens' choices than less populous nations. More transparent governments and those with good research institutions like the United States are better able to generate high quality information, while opaque governments and those lacking knowledge tend to settle for noisy signals. Better signals also become available in the new age of information when the collection, processing, and analysis of information become easier and faster. There are also heterogeneity in and exogenous shocks to information precision. Information on exchange-traded assets are more precise than information on over-the-counter assets. Stock price has become less discrete from 1/8s to 1/16s to decimals, and finer grids mean finer information (Graham et al 2003).

Prediction 16. *If the social planner is more likely active in agents' choices (less busy or at war, smaller population), or if information is noisier (opaque markets, industries, and regime, lack of research institutes, age of turbulence), agents tend to err against the direction of the social planner's news and this happens when the social planner is inactive. Such error is less likely if the social planner is more likely active, and the magnitude of loss from such error is lower if information is noisier. If the social planner is unlikely active in agents' choices, or if information is accurate, agents tend to err in the direction of the social planner's news and this happens when*

the social planner is active. Such error is less likely if the firm is less likely active or if information is more precise.

2.8 Robustness

First, our model and conclusions are robust to $p \in (\frac{1}{2}, 1)$ and $q \in (0, 1)$. Second, our focus on binary news is without the loss of generality. If the underlying state space for \tilde{x} is continuous, it is often that case that only states above a threshold is considered severe enough to warrant a warning. Then our model still applies. The model can be symmetrically rewritten with a focus on good news, or both good and bad news, but the core message that it can be optimal to block information for better disclosure remains the same.

Finally, agents in our model do not know if the social planner is active. This assumption is nonessential for our result and below we present a simple model without it but nevertheless conveys our core message. The idea and conclusions of our model are robust to this structure and we included it in our main model to fit better into the applications in practice and generate predictions.

A group of homogeneous agents try to know the content of two independent random signals: $\tilde{\alpha}, \tilde{\beta} \in \{0, 1\}$. An agent gets 0 if he knows both signals correctly, -1 if he gets one of them wrong, and -2 if he gets both wrong. The social planner observes $\tilde{\alpha}$ and $\tilde{\beta}$ but owing to communication constraints can not directly convey their values to agents. Agents observe $\tilde{\alpha}$ but only a noisy version of $\tilde{\beta}$, which is $\tilde{\gamma} = \tilde{\delta}\tilde{\beta} + (1 - \tilde{\delta})(1 - \tilde{\beta})$, where $Pr(\tilde{\delta} = 0) = \rho \in (0, \frac{1}{2})$ and $Pr(\tilde{\delta} = 1) = 1 - \rho$, so agents mistake $1 - \tilde{\beta}$ for $\tilde{\beta}$ with probability ρ . $\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}$ are mutually independent. We now state the first-best.

Lemma 17. *The first-best is for agents to always correctly know $\tilde{\alpha}$ and $\tilde{\beta}$. The first-best welfare is zero.*

The social planner decides whether to disclose $\tilde{\alpha}$ and $\tilde{\beta}$ to agents. Since $\tilde{\gamma}$ is fully derived from $\tilde{\beta}$, when the social planner withholds $\tilde{\beta}$, agents do not observe $\tilde{\gamma}$ either. Thus agents know they can observe $\tilde{\gamma}$ if and only if $\tilde{\beta}$ has been disclosed.

Suppose the social planner always discloses the signals. Agents read $\tilde{\alpha}$ correctly so they get 0 by knowing $\tilde{\alpha}$. Agents optimally use $\tilde{\gamma}$ in lieu of $\tilde{\beta}$ because this is their best available information on $\tilde{\beta}$: they read $\tilde{\beta}$ correctly with probability $1 - \rho > \frac{1}{2}$ which is better than random guessing. Their expected utility is $0 + \rho \times (-1) + (1 - \rho) \times 0 = -\rho \in (-\frac{1}{2}, 0)$. Therefore if the social planner always truthfully discloses both signals $\tilde{\alpha}$ and $\tilde{\beta}$, then the most efficient outcome is for agents to read the signals as $\{\tilde{\alpha}, \tilde{\gamma}\}$ and receive an expected payoff of $-\rho$. By wisely suppressing $\tilde{\beta}$ some of the time, the social planner can do much better:

Proposition 18. *The social planner attains the first-best by always disclosing $\tilde{\alpha}$ but disclosing $\tilde{\beta}$ if and only if $\tilde{\beta} = \tilde{\alpha}$.*

Intuitively, it makes sense for the social planner to contingently withhold $\tilde{\beta}$ that agents find harder to interpret, and let them infer it from $\tilde{\alpha}$ instead which they read correctly. Some essence of our theory is thus captured by this simple model.

2.9 Conclusions

The benevolent sender has to withhold some information contingently so as to efficiently communicate information when the sender may not always be active and the receiver does not know whether the sender is active. Our theory has wide applications in fields like political economy, law, and public policy. Efficiency is restored if the government has the authority to monitor and manipulate information like the power to purge dissident rumor. We show transparency hurts and rationalize a myriad of phenomena like surveillance programs, suppression of rumors, non-democracy, opaqueness, and oppression. The modeling per se illustrates a distinct situation

where the sender's incentive is aligned with the receiver's, more information helps the receiver, yet it is optimal for the sender to block some information from the receiver contingently in order to communicate more efficiently to the receiver. We also contribute to the economics of information transmission. Our theory generates interesting predictions and policy implications.

Appendix A

Proofs for Chapter 1

A.1 Proof of Proposition 1 (The First-Best)

The optimal $\lambda_l = 0$ since $l < 0$.

If $(1 - \phi)h + (1 - f)\phi l < 0$ then $\lambda_h = 0$ is optimal. The social planner solves $\max_{a,e \in [0,1]} w(a, e, 0, 0) = \max_{a,e \in [0,1]} -c - k$. The unique optimal solution is $(a, e) = (0, 0)$ and the value function is 0.

If $(1 - \phi)h + (1 - f)\phi l > 0$ then $\lambda_h = 1$ is optimal if the highest resulting social welfare is positive. The social planner solves $\max_{a,e \in [0,1]} w(a, e, 1, 0) = \max_{a,e \in [0,1]} (1 - \phi)h + (1 - f)\phi l - c - k$. Let the optimal solution be (a^r, e^r) and the associated value function be $w(a^r, e^r, 1, 0)$. (a^r, e^r) is unique for concave f and convex c and k .

The higher of the two value functions is the first-best total welfare, and the corresponding $(a, e, \lambda_h, \lambda_l)$ is the first-best.

Suppose $(a, e, \lambda_h, \lambda_l) = (0, 0, 0, 0)$ is the first-best. If there is no law then the market surplus equals the total surplus which is negative unless no transaction occurs, so the market breaks down. With no transaction, no one exerts effort. The market attains the first-best by itself so no law is needed.

A.2 Proof of Lemma 2 (The Need for Law)

The seller's expected profit is $\pi(a, e, \lambda_h, \lambda_l, p, x) = [(1 - \phi f)p - (1 - f)\phi x]\lambda_h + (p - x)\phi f\lambda_l - k$ if the buyer buys. $x < p$ so the optimal $\lambda_l = 1$. $(1 - f)\phi x < (1 - \phi f)x < (1 - \phi f)p$ so the optimal $\lambda_h = 1$. The seller solves $\max_{a \in [0,1]} \pi(a, e, 1, 1, p, x) = \max_{a \in [0,1]} p - \phi x - k$ at $t = 4$. The optimal $a = 0$ with profit function $p - \phi x > p - \phi p \geq 0$. The seller exerts no effort and expects zero profit if the buyer does not buy. Thus the seller prefers that the buyer buys and satisfies the buyer's participation constraint whenever possible. If the buyer participates at $t = 3$ then he buys at $t = 6$ since he learns no news. If so, then at $t = 4$ he solves $\max_{e \in [0,1]} u(0, e, 1, 1, p, y) = \max_{e \in [0,1]} (1 - \phi)h + (y + l)\phi - p - c$; the optimal $e = 0$. If the buyer does not buy he exerts no effort and gets 0. $(a, e, \lambda_h, \lambda_l) = (0, 0, 1, 1)$ disagrees with the first-best so the outcome is inefficient.

A.3 Proof of Proposition 3 (Implementing the First-Best)

Suppose the buyer participates at $t = 3$. The seller's expected profit under x^r is $\pi(a, e, \lambda_h, \lambda_l, p, x^r) = [(1 - \phi f)p - (1 - f)(p - l)\phi]\lambda_h + l\phi f\lambda_l - k$ if the buyer buys. The optimal $\lambda_l = 0$ since $l < 0$. If $\lambda_h = 0$, the buyer and the seller expect 0. If $\lambda_h = 1$, then when the seller sells, the buyer at $t = 6$ updates his belief to $Pr(\tilde{\tau} = l | \tilde{s} = h)$ and expects $[u(a, e, 1, 0, p, y) + c]/(1 - \phi f) - c \geq u(a, e, 1, 0, p, y) \geq 0$ if he buys. So if the buyer participates then he buys when the seller sells.

The seller at $t = 4$ and the social planner effectively solve the same problem $\max_{a \in [0,1]} (1 - f)l\phi - k$ when $\lambda_h = 1$. So the seller chooses a^r if the buyer chooses e^r . Under y^r the buyer at $t = 4$ and the social planner solve the same problem $\max_{e \in [0,1]} (1 - f)\phi l - c$. So the buyer chooses e^r if the seller chooses $\lambda_h = 1$ and a^r . The social planner expects to keep $(x^r - y^r)q_l = (f^r - 1)\phi l > 0$. The seller extracts the residual surplus and makes profit $w(a^r, e^r, 1, 0) + (1 - f^r)\phi l$. If this is positive,

the seller sets $\lambda_h = 1$; otherwise No-sale is chosen instead.

For uniqueness, assume differentiability. The first-order conditions are $(x - p)\phi f'_a = k'$ for the seller and $(p - l - y)\phi f'_e = c'$ for the buyer. The social planner's first-order conditions are $-\phi l f'_a = k'$ and $-\phi l f'_e = c'$. The unique law that matches coefficients and implements (a^r, e^r) is $(x^r, y^r) = (p - l, p)$.

A.4 Proof of Proposition 4 (Balance the Budget)

Similar arguments as the proof for implementing the first-best are not repeated here. Suppose the buyer participates at $t = 3$. If the seller faces penalty p then he expects $q_h p - k = (1 - \phi)\lambda_h p - k$ if the buyer buys, so the optimal $\lambda_h = 1$. If the seller faces x^r then $\lambda_l = 0$ by earlier analysis. So when the seller sells, the buyer's belief on his type is no worse than at $t = 3$ so he buys. If the seller chooses $(\lambda_h, \lambda_l) = (1, 0)$ and a^r and satisfies the buyer's participation constraint then the buyer exerts e^r and buys under y^r by earlier analysis. Suppose the buyer exerts e^r and buys. If the seller exerts at least a^r then he faces penalty p and expects $(1 - \phi)p - k$; the optimal $a = a^r$. The buyer expects $u(a, e, 1, \lambda_l, p, y^r) = [(1 - \phi)h + (1 - f)\phi l - c] + \lambda_l \phi f l - (1 - \phi)p$. For any (a, e) , the buyer's willingness to pay and thus the seller's profit is the highest if $\lambda_l = 0$, so the seller optimally sets $\lambda_l = 0$. The buyer faces the same incentive as the social planner and exerts e^r . The budget is balanced and the seller expects $w(a^r, e^r, 1, 0) \geq 0$. He optimally motivates the buyer to participate. The first-best is always attained. If the seller exerts at most a^r then he faces x^r and makes less than $w(a^r, e^r, 1, 0)$ due to the budget surplus so he prefers exerting at least a^r .

A.5 Proof of Lemma 5 (The First-Best for One-Stop Seller)

First, a few preliminaries. Let q_s , q_h , q_l be the probabilities that the buyer gets the good, gets h , and gets l . We have

$$\begin{aligned}
q_s &= \frac{1}{2}(\lambda_u + \lambda_{lh} + \lambda_{hl} + \lambda_{hh}), \\
q_h &= \lambda_u Pr(\tilde{\tau} = \tilde{s} = l) + \lambda_{hh} Pr(\tilde{\tau} = \tilde{s} = h) + \\
&\quad \lambda_{lh} Pr(\tilde{\tau} = l, \tilde{s} = h) + \lambda_{hl} Pr(\tilde{\tau} = h, \tilde{s} = l) \\
&= \frac{1}{4}[(1+f)(\lambda_u + \lambda_{hh}) + (1-f)(\lambda_{lh} + \lambda_{hl})], \\
q_l &= q_s - q_h = \frac{1}{4}[(1-f)(\lambda_u + \lambda_{hh}) + (1+f)(\lambda_{lh} + \lambda_{hl})]
\end{aligned} \tag{A.1}$$

The seller's expected profit is

$$\begin{aligned}
\pi(a, e, \lambda_{hh}, \lambda_{hl}, \lambda_{lh}, \lambda_u, p, x) &= q_s p - q_l x - k \\
&= \left[\frac{p}{2} - \frac{(1-f)x}{4} \right] (\lambda_u + \lambda_{hh}) + \left[\frac{p}{2} - \frac{(1+f)x}{4} \right] (\lambda_{lh} + \lambda_{hl}) - k
\end{aligned} \tag{A.2}$$

The buyer's expected utility is

$$\begin{aligned}
u(a, e, \lambda_{hh}, \lambda_{hl}, \lambda_{lh}, \lambda_u, p, y) &= q_h h + (y + l)q_l - q_s p - c \\
&= \left[\frac{1+f}{4}h + \frac{1}{4}(1-f)(y+l) - \frac{p}{2} \right] (\lambda_u + \lambda_{hh}) + \\
&\quad \left[\frac{1-f}{4}h + \frac{1}{4}(1+f)(y+l) - \frac{p}{2} \right] (\lambda_{lh} + \lambda_{hl}) - c
\end{aligned} \tag{A.3}$$

The social welfare is

$$\begin{aligned}
w(a, e, \lambda_{hh}, \lambda_{hl}, \lambda_{lh}, \lambda_u) &= q_h h + q_l l - c - k \\
&= \frac{(1+f)h + (1-f)l}{4} (\lambda_u + \lambda_{hh}) + \frac{(1-f)h + (1+f)l}{4} (\lambda_{lh} + \lambda_{hl}) - c - k
\end{aligned} \tag{A.4}$$

Since $(1+f)h + (1-f)l > (1-f)h + (1+f)l$, the optimal $\lambda_{lh} = \lambda_{hl} = 0$.

If $(1 + f)h + (1 - f)l < 0$ then the optimal $\lambda_{hh} = \lambda_{ll} = 0$. Efforts are useless but costly, so the unique socially optimal $a = e = 0$. The associated social welfare is 0.

If $(1 + f)h + (1 - f)l > 0$ then the optimal $\lambda_{hh} = \lambda_{ll} = 1$ if the resulting optimal social welfare is positive. The total welfare is $w(a, e, 1, 0, 0, 1) = [(1 + f)h + (1 - f)l]/2 - c - k$. Let (a_2^r, e_2^r) maximize $w(a, e, 1, 0, 0, 1)$. The social planner chooses $(a_2^r, e_2^r, 1, 0, 0, 1)$ over zero effort and no sale if $w(a_2^r, e_2^r, 1, 0, 0, 1) > 0$. When f is concave and c, k are convex, this optimum is unique.

A.6 Proof of Proposition 6 (Implementing the First-Best for One-Stop Seller)

Similar arguments as the proofs for the standard case are not repeated here. Suppose the buyer participates at $t = 3$. The seller's expected profit under x_2^r is $(\lambda_{ll} + \lambda_{hh})[2p - (1 - f)(h - l)]/4 + (\lambda_{lh} + \lambda_{hl})[2p - (1 + f)(h - l)]/4 - k$ if the buyer buys. $(1 - f)(h - l) < (1 + f)(h - l)$ so the optimal $\lambda_{lh} = \lambda_{hl} = 0$. If $\lambda_{hh} = \lambda_{ll} = 0$, the buyer and the seller expects 0. If $\lambda_{hh} = \lambda_{ll} = 1$ then the seller always sells and the buyer expects $u(a, e, 1, 0, 0, 1, p, y)$ if he buys. If the buyer participates then he buys when the seller sells.

The seller at $t = 4$ and the social planner solve the same problem $\max_{a \in [0,1]} (h - l)f/2 - k$ when $\lambda_{hh} = \lambda_{ll} = 1$. So the seller chooses a_2^r if the buyer chooses e_2^r . The buyer at $t = 4$ and the social planner solve the same problem $\max_{e \in [0,1]} (h - l)f/2 - c$. So the buyer chooses e_2^r if the seller chooses $\lambda_{hh} = \lambda_{ll} = 1$ and a_2^r . The social planner expects to keep $x_2^r q_l = (1 - f_2^r)(h - l)/2 > 0$. The seller extracts the residual surplus and makes profit $w(a_2^r, e_2^r, 1, 0, 0, 1) - (1 - f_2^r)(h - l)/2 = f_2^r h + (1 - f_2^r)l - c_2^r - k_2^r$. If this is positive, the seller sets $\lambda_{hh} = \lambda_{ll} = 1$; otherwise No-sale is chosen instead.

For uniqueness, assume differentiability. The first-order conditions are $x f'_a = 2k'$ for the seller and $(h - y - l)f'_e = 2c'$ for the buyer. The social planner's first-order conditions are $(h - l)f'_a = 2k'$ and $(h - l)f'_e = 2c'$. The unique law that matches

coefficients and implements (a_2^r, e_2^r) is $(x_2^r, y_2^r) = (h - l, 0)$. Under y^r , the max price that motivates the buyer to participate is $h + (q_l l - c)/q_h < h$, so $x^r < x_2^r$.

Suppose the social planner can infer $1\{\tilde{\tau} = l, f < f_2^r\}$ ex post. Suppose the buyer participates at $t = 3$. If the seller faces no penalty he expects $q_s p - k$ if the buyer buys, so he always sells. If the seller faces x_2^r then $\lambda_{hl} = \lambda_{lh} = 0$ by earlier analysis. So when the seller sells, the buyer's belief on matching is no worse than at $t = 3$ so he buys. If the seller chooses $\lambda_{hh} = \lambda_{ll} = 1$ and a_2^r and satisfies the buyer's participation constraint then the buyer exerts e_2^r and buys by earlier analysis. Suppose the buyer exerts e_2^r and buys. If the seller exerts at least a_2^r then he expects $p - k$; the optimal $a = a_2^r$. The buyer expects $[(1 - f)h + (1 + f)l]/2 + (\lambda_{ll} + \lambda_{hh})(h - l)f/2 - p - c$. For any (a, e) , the buyer's willingness to pay and thus the seller's profit is the highest if $\lambda_{hh} = \lambda_{ll} = 1$, so the seller optimally chooses so. The buyer faces the same incentive as the social planner and exerts e_2^r . The budget is balanced and the seller expects $w(a_2^r, e_2^r, 1, 0, 0, 1) \geq 0$. He optimally motivates the buyer to participate. The first-best is always attained. If the seller exerts at most a_2^r then he faces x_2^r and makes less than $w(a_2^r, e_2^r, 1, 0, 0, 1)$ due to the budget surplus so he prefers exerting at least a_2^r . If loss occurs the social planner does not act on the equilibrium path since $f = f_2^r$.

A.7 Proof of Proposition 7 (The Importance of Dual Moral Hazard)

If $c = k \equiv 0$, the total welfare is $w(a, e, \lambda_h, \lambda_l) = [(1 - \phi)h + (1 - f)\phi l]\lambda_h + \phi f l \lambda_l$. $\lambda_l = 0$ is optimal. If $(1 - \phi)h + [1 - f(1, 1)]\phi l > 0$ then $(a, e, \lambda_h, \lambda_l) = (1, 1, 1, 0)$ is the first-best; otherwise any $(a, e, 0, 0)$ is the first-best with $w(a, e, 0, 0) = 0$ and the market breaks down. We take the more interesting case of $(1, 1, 1, 0)$ as the first-best since no law is needed in the other case.

Suppose the buyer participates at $t = 3$. The seller's expected profit under $x > p$ is $[(1 - \phi f)p - (1 - f)\phi x]\lambda_h + (p - x)\phi f \lambda_l$ if the buyer buys. The optimal $\lambda_l = 0$. If $\lambda_h = 0$, the buyer and the seller expect 0. If $\lambda_h = 1$, then when

the seller sells, the buyer at $t = 6$ updates his belief to $Pr(\tilde{\tau} = l | \tilde{s} = h)$ and expects $[u(a, e, 1, 0, p, y)/(1 - \phi f)] \geq u(a, e, 1, 0, p, y) \geq 0$ if he buys. So if the buyer participates then he buys when the seller sells.

Suppose $\lambda_h = 1$. The seller's expected profit $(1 - \phi f)p - (1 - f)\phi x = (p - \phi x) + (x - p)\phi f$ is increasing in f , so $a = 1$ is optimal if the profit is positive. Under $y = p + \epsilon \in (p, p - l)$ the buyer's expected utility $(1 - \phi)h + (1 - f)(\epsilon + l)\phi - (1 - \phi)p$ is increasing in f so $e = 1$ is optimal and the buyer can be motivated to participate and buy if his willingness to pay is positive. When $(a, e) = (1, 1)$ the seller's profit $[1 - \phi f(1, 1)]p - [1 - f(1, 1)]\phi x$ is positive for $x < \frac{1 - \phi f(1, 1)}{[1 - f(1, 1)]\phi}p$. The buyer's expected payoff gross of payment is $(1 - \phi)h + [1 - f(1, 1)](\epsilon + l)\phi > (1 - \phi)h + [1 - f(1, 1)]\phi l = w(1, 1, 1, 0) > 0$, so the price is positive. Therefore any $x \in \left(p, \frac{1 - \phi f(1, 1)}{[1 - f(1, 1)]\phi}p\right)$ and $y \in (p, p - l)$ restore the first-best. To see that $p + \epsilon < \frac{1 - \phi f(1, 1)}{[1 - f(1, 1)]\phi}p$, note this reduces to $[1 - f(1, 1)]\phi \epsilon < (1 - \phi)p$. The price satisfies the buyer's participation constraint so $(1 - \phi)p = (1 - \phi)h + [1 - f(1, 1)](\epsilon + l)\phi$. The inequality further reduces to $(1 - \phi)h + [1 - f(1, 1)]\phi l > 0$ which holds since the left-hand side is $w(1, 1, 1, 0)$. Therefore any $x = y = p + \epsilon$ where $\epsilon \in (0, -l)$ always restores the first-best and the budget is always balanced.

If there is no buyer's moral hazard, then the social welfare is $w(a, \lambda_h, \lambda_l) = [(1 - \phi)h + (1 - f)\phi l]\lambda_h + \phi f l \lambda_l - k$. The optimal $\lambda_l = 0$. Let $\max_{a \in [0, 1]} w(a, 1, 0) > 0$ since otherwise no law is needed. The first-best implementation is very similar to the standard model. x^r and any $y \leq x^r$ restores the first-best; in particular, $y = x^r$ runs a balanced budget and always restores efficiency.

If there are no effort choices by the buyer and the seller, then the social welfare is $w(\lambda_h, \lambda_l) = [(1 - \phi)h + (1 - f)\phi l]\lambda_h + \phi f l \lambda_l$. The optimal $\lambda_l = 0$. Let $w(1, 0) = (1 - \phi)h + (1 - f)\phi l > 0$ so that $(1, 0)$ is the first-best, since otherwise no law is needed. The derivation is very similar to the standard case. $x \in \left(p, \frac{1 - \phi f}{(1 - f)\phi}p\right)$

and $y \leq x$ restore the first-best and a wide range of $y = x \in \left(p, \frac{1-\phi f}{(1-f)\phi} p\right)$ does so unconditionally with a balanced budget.

A.8 Motivation for Ex Post Inference

The social planner may learn $1\{\tilde{\tau} = l, f < f^r\}$ ex post in various ways. The loss can be different for $f < f^r$ and $f \geq f^r$. This is reasonable because effective seller-buyer interaction before and during the purchase can often help the buyer better use the good or service. Therefore high joint efforts reduce the loss to the type l buyer. More generally, the loss can be stochastic. For instance, if $f < f^r$ then the buyer's loss distribution is worse than and does not overlap with his loss distribution if $f \geq f^r$. Overlapping distributions can be accommodated if we model a unit mass of buyers. Alternatively, if the loss $l = l(a, e)$ directly depends on, is monotone in and thus informative of joint efforts, then the social planner can infer joint efforts from $l(a, e)$. Our analysis on balancing the budget still goes through if the social planner can tell $l \geq l(a^r, e^r)$ from $l < l(a^r, e^r)$, except the algebra is messier. The social planner may also learn $1\{\tilde{\tau} = l, f < f^r\}$ directly. After the good fails for the unfit buyer the underlying reason for the failure usually surfaces. The social planner can learn more about f given the reason for the failure. If some easy questions and answers could have revealed the mismatch, then the social planner can infer that f had been low. For example, if young teenagers are sold alcohol or firearms. On the other hand, if the cause for the failure is difficult to anticipate, then even if both parties exerted high efforts the mismatch may still have gone unnoticed. Either way, more can be deduced about f the quality of their interaction due to their joint efforts. Even more information can be revealed if the communication between the seller and the buyer was recorded via formal files, emails, meeting minutes, video or audio recording, transcripts of online chats, etc. Such records are routinely kept during customer

services and sales, interviews, (financial) consulting, interaction over the counter, video conferences and phone calls, legal communication and interrogation, etc. In reviewing records of communication, the social planner may be able to infer whether the logic flow of the conversation reflects reasonable thinking and deliberation, why certain questions were avoided and answers concealed, whether the questions were carefully formulated and answers informative and well-articulated, and how likely the mismatch could have been identified given the reason of the failure and the communication. Note however we are assuming that whether f had been high or low can be inferred only ex post after the buyer incurs a loss. Without the fundamental cause for the failure or the realized amount of loss, no one can learn $1\{f < f^r\}$. Therefore f is not deducible before payoffs are realized or for the buyer who does not experience a bad outcome. The social planner may also infer f via the sales volume q_s and the overall fraction or long-run frequency of losses or lawsuits q_l . Such inferred f represents the average productivity of the seller with a large number of buyers, and can be a good indicator of joint productivity in typical disputes.

Appendix B

Proofs for Chapter 2

B.1 Derivations for Results in the First Table

\tilde{x} is good with probability $\frac{1}{2}$; given this, r exists with probability $1-p$; independently with probability q the social planner is active; given this and the good \tilde{x} , w exists with probability $1-p$. Hence $Pr(gawr) = \frac{1}{2}(1-p)^2q$. Similarly, $Pr(bawr) = \frac{1}{2}p^2q$. Therefore $Pr(awr) = Pr(bawr) + Pr(gawr) = \frac{1}{2}[p^2 + (1-p)^2]q$. Likewise, $Pr(aw) = Pr(gaw) + Pr(baw) = \frac{1}{2}(1-p)pq + \frac{1}{2}(1-p)pq = (1-p)pq$. $Pr(ir) = Pr(gir) + Pr(bir) = \frac{1}{2}(1-q)(1-p) + \frac{1}{2}(1-q)p = \frac{1-q}{2}$. $Pr(ar) = Pr(gar) + Pr(bar) = \frac{1}{2}(1-p)pq + \frac{1}{2}(1-p)pq = (1-p)pq$. $Pr(i) = Pr(gi) + Pr(bi) = \frac{1}{2}(1-q)p + \frac{1}{2}(1-q)(1-p) = \frac{1-q}{2}$. $Pr(a) = Pr(ga) + Pr(ba) = \frac{1}{2}p^2q + \frac{1}{2}(1-p)^2q = \frac{1}{2}[p^2 + (1-p)^2]q$. By the law of conditional probability,

$$Pr(b|awr) = \frac{Pr(bawr)}{Pr(awr)} = \frac{p^2}{p^2 + (1-p)^2} > p \quad (\text{B.1})$$

To see the inequality, note since $p \in (\frac{1}{2}, 1)$, we have $(p-1)(2p-1) < 0$, or $p^2 + (1-p)^2 < p$, or $Pr(b|awr) > p$. Hence $Pr(b|awr) > p^*$. Likewise,

$$Pr(b|a) = \frac{Pr(ba)}{Pr(a)} = \frac{(1-p)^2}{p^2 + (1-p)^2} < \frac{(1-p)^2}{(1-p)^2 + (1-p)^2} = \frac{1}{2} < p^* \quad (\text{B.2})$$

Similarly,

$$Pr(b|aw) = \frac{Pr(baw)}{Pr(aw)} = \frac{1}{2} < p^*$$

The same is true of $Pr(b|ar)$. We also have

$$Pr(b|ir) = \frac{Pr(bir)}{Pr(ir)} = p > p^*, \quad Pr(b|i) = \frac{Pr(bi)}{Pr(i)} = 1 - p < \frac{1}{2} < p^*$$

B.2 Proof of Lemma 9 (The Social Planner's Belief)

If event ir or i occurs then the social planner is inactive and has no belief. Among the six events in Λ , the unique event where the social planner sees both w and r is event awr ; the unique event where it sees w but not r is event aw ; the unique event where it has no w but hears r is event ar ; the unique event where it is active but sees nothing is event a . Therefore the social planner can perfectly recognize each of the four events awr, aw, ar, a . When one of them happens, the social planner precisely knows which. So if $E \in \{awr, aw, ar, a\}$, $\mathcal{I}_S = E$ and the social planner's belief updates to $Pr(b|\mathcal{I}_S) = Pr(b|E)$.

B.3 Proof of Lemma 10

(a). Among the six events in Λ , the unique event where agents see w and hear r is awr ; the unique event where they see w but hear no r is aw . Therefore agents can perfectly distinguish each of awr and aw . When one of them occurs agents know precisely which. So if $E \in \{awr, aw\}$, $\mathcal{I}_A = E$ and agents' belief updates to

$$Pr(b | \mathcal{I}_A) = Pr(b | E).$$

(b). There are two events among the six in Λ where agents hear r but see no w : ir and ar . Agents can not tell ir from ar because they do not know if the social planner is active. So $\mathcal{I}_A = \{ir, ar\}$ if either occurs. Agents' belief updates to

$$Pr(b | \{ir, ar\}) = \frac{Pr(bir) + Pr(bar)}{Pr(ir) + Pr(ar)} = \frac{(1-pq)p}{1-q+2(1-p)pq} \in \left(\frac{1}{2}, p\right) \quad (\text{B.3})$$

$Pr(b | \{ir, ar\}) > \frac{1}{2}$ reduces to $2p > 1 - q + 2pq$, or $(1-q)(2p-1) > 0$, which holds by assumption. $Pr(b | \{ir, ar\}) < p$ reduces to $(1-p)q < 2(1-p)pq$, or $1 < 2p$.

$Pr(b | \{ir, ar\})$ is strictly decreasing in q because

$$\frac{\partial}{\partial q} \left[\frac{(1-pq)p}{1-q+2(1-p)pq} \right] = - \frac{(2p-1)(1-p)p}{[1-q+2(1-p)pq]^2} < 0 \quad (\text{B.4})$$

So for a given p , its sup $Pr(b | \{ir, ar\}) = \lim_{q \rightarrow 0} Pr(b | \{ir, ar\}) = p$, inf $Pr(b | \{ir, ar\}) = \lim_{q \rightarrow 1} Pr(b | \{ir, ar\}) = \frac{1}{2}$. When the social planner is almost never active, agents take whether an official warning exists unknowable because there is almost no chance to learn it. Thus agents only condition on whether there is an rumor r . This is when agents' posterior belief rises to the highest level p , which is the posterior belief after only one negative signal r . When the social planner is almost always active, then agents' not observing an official warning w means no w exists in the first place. Agents condition on there being r but no w . The existence of r and the non-existence of w offset each other, so agents' posterior is unchanged from the prior, reaching its lowest level.

$Pr(b | \{ir, ar\})$ is strictly increasing in p because

$$\frac{\partial}{\partial p} \left[\frac{(1-pq)p}{1-q+2(1-p)pq} \right] = \frac{[1-2(1-p)pq](1-q)}{[1-q+2(1-p)pq]^2} > 0 \quad (\text{B.5})$$

The numerator is positive if $2(1-p)pq < 1$. Since

$$(1-p)p \leq \frac{[(1-p)+p]^2}{4} = \frac{1}{4} \quad (\text{B.6})$$

we have $4(1-p)p \leq 1$, or $2(1-p)pq \leq \frac{q}{2} < 1$. Therefore $Pr(b | \{ir, ar\})$ is strictly increasing in p . So for a given q , $\sup Pr(b | \{ir, ar\}) = \lim_{p \rightarrow 1} Pr(b | \{ir, ar\}) = 1$ and $\inf Pr(b | \{ir, ar\}) = \lim_{p \rightarrow \frac{1}{2}} Pr(b | \{ir, ar\}) = \frac{1}{2}$. With very high quality information, agents' posterior rises all the way to 1 since hearing the rumor r almost surely reflects that the underlying \tilde{x} is bad. With very low quality information, agents' posterior belief equals the prior because the signals become useless noises; hearing r and seeing no w reflects nothing about the underlying \tilde{x} .

(c). There are two events among the six in Λ where agents see nothing: i and a . Agents can not tell i from a because they do not know if the social planner is active. So $\mathcal{I}_A = \{i, a\}$ if either occurs. Agents' belief updates to

$$Pr(b | \{i, a\}) = \frac{Pr(bi) + Pr(ba)}{Pr(i) + Pr(a)} = \frac{(1-pq)(1-p)}{1 - 2(1-p)pq} \in \left(\frac{(1-p)^2}{(1-p)^2 + p^2}, 1-p \right)$$

Recall $2(1-p)pq < 1$ from our proof for part (b), so we can cross-multiply and reduce the lower-bound inequality to $(1-pq)[(1-p)^2 + p^2] > (1-p)[1 - 2(1-p)pq]$, or $2(1-q)p^2 > (1-q)p$, or $2p > 1$ which holds by assumption. The upper-bound inequality $Pr(b | \{i, a\}) < 1-p$ reduce to $1-pq < 1 - 2(1-p)pq$, or $pq > 2(1-p)pq$, or $2p > 1$.

$Pr(b | \{i, a\})$ is strictly decreasing in q because

$$\frac{\partial}{\partial q} \left[\frac{(1-pq)(1-p)}{1 - 2(1-p)pq} \right] = -\frac{(2p-1)(1-p)p}{[1 - 2(1-p)pq]^2} < 0 \quad (\text{B.7})$$

So for a given p , $\sup Pr(b | \{i, a\}) = \lim_{q \rightarrow 0} Pr(b | \{i, a\}) = 1 - p$ and $\inf Pr(b | \{i, a\}) = \lim_{q \rightarrow 1} Pr(b | \{i, a\}) = \frac{(1-p)^2}{(1-p)^2 + p^2}$. When the social planner is almost never active, agents take whether w exists unknowable because there is almost no chance to learn it. Thus agents only condition on there being no r . This is when agents' posterior belief reaches the highest level $1 - p$ because there is only one piece of good news that there is no r . When the social planner is almost always active, agents' not observing w means no w exists in the first place. Agents condition on there being no w or r . The two pieces of good news reinforce each other, so agents' posterior belief is the lowest.

$Pr(b | \{i, a\})$ is strictly decreasing in p because

$$\frac{\partial}{\partial p} \left[\frac{(1-pq)(1-p)}{1-2(1-p)pq} \right] = -\frac{1-q+2pq(1-p)}{(1-2pq+2p^2q)^2} < 0 \quad (\text{B.8})$$

So for a given q , $\sup Pr(b | \{i, a\}) = \lim_{p \rightarrow \frac{1}{2}} Pr(b | \{i, a\}) = \frac{1}{2}$ and $\inf Pr(b | \{i, a\}) = \lim_{p \rightarrow 1} Pr(b | \{i, a\}) = 0$. With very low quality information, the agents' posterior belief equals the prior since the signals become useless noises; observing nothing has no bearing on the belief regarding \tilde{x} . With very high quality information, the agents' posterior belief falls all the way to zero because observing nothing almost surely means that the underlying \tilde{x} is good.

B.4 Proof of Proposition 11 (The Inefficiency of Full Disclosure)

If $Pr(\{ir, ar\}) < p^*$ then agents optimally choose \tilde{x} when they hear r but see no w . Agents make the wrong decision in event ir since the first-best is to avoid \tilde{x} . Their welfare from \tilde{x} in event ir is $[1 - Pr(b | ir)] \times 1 - Pr(b | ir) \times c = 1 - (1 + c)p < 0$. Since $Pr(ir) = \frac{1-q}{2}$, agents' expected welfare loss is

$$L_{<} \equiv \frac{(1-q)[(c+1)p-1]}{2} > 0. \quad (\text{B.9})$$

which is increasing in p and decreasing in q .

If $Pr(\{ir, ar\}) > p^*$ then agents optimally avoid \tilde{x} when they hear r but see no w . Agents make the wrong decision in event ar since the first-best is to choose \tilde{x} . Their welfare from \tilde{x} in event ar is $[1 - Pr(b|ar)] \times 1 - Pr(b|ar) \times c = \frac{1-c}{2} > 0$. Since $Pr(ar) = (1-p)pq$, agents' expected welfare loss is

$$L_{>} \equiv \frac{(1-c)(1-p)pq}{2} > 0 \quad (\text{B.10})$$

which is decreasing in p for $p \in (\frac{1}{2}, 1)$ and increasing in q .

Therefore an inefficiency exists irrespective of how $Pr(\{ir, ar\})$ compares with the critical belief p^* .

B.5 Proof of Lemma 12

(a). Since r is purged in event ar , now there are only two events among the six in Λ where agents can hear r : awr and ir . In event awr , agents see both w and r ; in event ir , agents only hear r but see no w . The only event among the six where agents see w but hear no r is aw . So agents can recognize each of awr , ir , and aw . If $E \in \{sir, awe, aw\}$, $\mathcal{I}_A = E$ and agents' belief updates to $Pr(b|\mathcal{I}_A) = Pr(b|E)$.

(b). This leaves us with three events ar , i , a , which agents can not distinguish because they see nothing in all three events. So if $E \in \{ar, i, a\}$, then $\mathcal{I}_A = \{ar, i, a\}$ and agents' belief updates to

$$Pr(b|\{ar, i, a\}) = \frac{Pr(bar) + Pr(bu) + Pr(bi)}{Pr(ir) + Pr(u) + Pr(i)} = \frac{(1-p)/2}{1/2} = 1-p < \frac{1}{2} \quad (\text{B.11})$$

which is strictly decreasing in p . So $\inf Pr(b|\{ar, i, a\}) = \lim_{p \rightarrow 1} Pr(b|\{ar, i, a\}) = 0$ and $\sup Pr(b|\{ir, ar\}) = \lim_{p \rightarrow \frac{1}{2}} Pr(b|\{ar, i, a\}) = \frac{1}{2}$. With very high quality information the agents' posterior belief falls all the way to 0 because the absence of

w in events ar and a plus the absence of r in events i and a almost surely reflect that the underlying \tilde{x} is good. With very low quality information the agents' posterior belief equals the prior because the signals become useless noises; observing nothing has no bearing on the belief regarding \tilde{x} .

B.6 Proof of Proposition 13

Since agents can perfectly recognize each of the three events awr , aw , and ir , their information set $\mathcal{I}_A = E$ and their belief $Pr(b|\mathcal{I}_A) = Pr(b|E)$ for $E \in \{awr, aw, ir\}$. Therefore agents make the first-best decision in these three events.

In each of the other three events ar, i and a agents observe nothing so their information set is $\mathcal{I}_A = \{ar, i, a\}$ and their belief is $Pr(b|\{ar, i, a\}) = 1 - p < p^*$. Agents optimally accept \tilde{x} . Since the first-best in each of the three events ar, i , and a is to accept \tilde{x} , agents' choice is efficient in these events.

In conclusion, the first-best is attained in every event if the social planner purges dissident rumor by blocking r in event ar .

B.7 Proof of Proposition 18

Since $\tilde{\gamma}$ the noisy version of the signal $\tilde{\beta}$ exists if and only if $\tilde{\beta}$ is disclosed, agents get to observe $\tilde{\gamma}$ if and only if the social planner discloses $\tilde{\beta}$.

Suppose the social planner always discloses $\tilde{\alpha}$ but discloses $\tilde{\beta}$ if and only if $\tilde{\beta} = \tilde{\alpha}$.

Given such equilibrium strategy of the social planner, agents know that whenever they get to observe $\tilde{\gamma}$, it must be that $\tilde{\beta} = \tilde{\alpha}$; and whenever they do not get to observe $\tilde{\gamma}$, it must be that $\tilde{\beta} \neq \tilde{\alpha}$, or $\tilde{\beta} = 1 - \tilde{\alpha}$.

Since agents always read $\tilde{\alpha}$ correctly, they also always infer $\tilde{\beta}$ correctly from their knowledge of $\tilde{\alpha}$. Therefore agents always correctly know $\tilde{\alpha}$ and $\tilde{\beta}$. The first-best is thus attained.

Bibliography

- [1] Acharya, Viral V., Peter DeMarzo and Ilan Kremer. 2011. Endogenous Information Flows and the Clustering of Announcements. *American Economic Review* 101:2955-2979.
- [2] Akerlof, George A. 1970. The Market for Lemons. *Quarterly Journal of Economics* 84(3):488-500.
- [3] Allen, Franklin, and Douglas Gale. 1999. Innovations in Financial Services, Relationships, and Risk Sharing. *Management Science* 45(9):1239-53.
- [4] Amador, Manuel, and Pierre-Olivier Weill. 2012. Learning from Private and Public Observations of Others' Actions. *Journal of Economic Theory* 147:910-40.
- [5] Angeletos, George-Marios, and Alessandro Pavan. 2007. Efficient Use of Information and Social Value of Information. *Econometrica* 75:1103-42.
- [6] Apuzzo, Matt and Joseph Goldstein. 2014. New York Drops Unit That Spied Among Muslims. *New York Times* April 16:A1.
- [7] Arrow, Kenneth J. 1963. Uncertainty and the Welfare Economics of Medical Care. *American Economic Review* 53(5):941-973.
- [8] Barro, Robert J. 1998. *Determinants of Economic Growth: A Cross-Country Empirical Study*. The MIT Press.
- [9] BBC News. 2010. Warren Buffett Defends Credit Rating Agencies. June 2.
- [10] Bhattacharyya, Sugato, and Francine Lafontaine. 1995. Double-Sided Moral Hazard and the Nature of Share Contracts. *RAND Journal of Economics* 26:761-781.
- [11] Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 100:992-1026.

- [12] Bolton, Patrick, Xavier Freixas, and Joel Shapiro. 2007. Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry. *Journal of Financial Economics* 82: 297-330.
- [13] Boot, Arnoud W. A. 2000. Relationship Banking: What Do We Know? *Journal of Financial Intermediation* 9: 7-25.
- [14] Bradsher, Keith. 2011. Two Top Executives Resign From Wal-Mart China. *The New York Times*. October 17.
- [15] Burguet, Roberto, and Xavier Vives. 2000. Social Learning and Costly Information Acquisition. *Economic Theory* 15:185-205.
- [16] Carlin, Bruce Ian, and Simon Gervais. 2009. Work Ethic, Employment Contracts, and Firm Value. *Journal of Finance* 64(2):785-821.
- [17] Carlin, Bruce Ian, and Simon Gervais. 2012. Legal Protection in Retail Financial Markets. *Review of Corporate Finance Studies* 1:68-108.
- [18] Carlin, Bruce Ian, Simon Gervais, and Gustavo Manso. 2013. Libertarian Paternalism, Information Production, and Financial Decision Making. *Review of Financial Studies* 26(9): 2204-2228.
- [19] Carlin, Bruce Ian, Florin Dorobantu, and Sean Viswanathan. 2009. Public Trust, the Law, and Financial Investment. *Journal of Financial Economics*, 92:321-341.
- [20] Crawford, Vincent P., and Joel Sobel. 1982. Strategic Information Transmission. *Econometrica* 50(6):1431-1451.
- [21] Cheung, Steven N.S. 1968. Private Property Rights and Sharecropping. *Journal of Political Economy* 76(6):1107-22.
- [22] Cross, Frank B. 2002. Law and Economic Growth. *Texas Law Review* 80(7):1737-1776.
- [23] Darby, Michael R., and Edi Karni. 1973. Free Competition and the Optimal Amount of Fraud. *Journal of Law and Economics* 16:67-88.
- [24] Demski, Joel S., and David E. M. Sappington. 1987. Delegated Expertise. *Journal of Accounting Research* 25(1):68-89.
- [25] Diamond, Douglas W. 1984. Financial Intermediation and Delegated Monitoring. *The Review of Economic Studies*. 51(3):393-414.

- [26] Diamond, Douglas W. 1991. Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt. *Journal of Political Economy*. 99(4):689-721.
- [27] Doney, Patricia M., and Joseph P. Cannon. 1997. An Examination of the Nature of Trust in Buyer-Seller Relationships. *Journal of Marketing* 61(2):35-51.
- [28] Drucker, Steven and Manju Puri. 2009. On Loan Sales, Loan Contracting and Lending Relationships. *Review of Financial Studies* 22(7):2835-2872.
- [29] Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh. 1987. Developing Buyer-Seller Relationships. *Journal of Marketing* 51(2):11-27.
- [30] Dye, Ronald A. 1985a. Disclosure of Nonproprietary Information. *Journal of Accounting Research* 23, 123-145.
- [31] Dye, Ronald A. 1985b. Strategic Accounting Choice and the Effects of Alternative Financial Reporting Requirements. *Journal of Accounting Research* 23, 544-574.
- [32] Dye, Ronald A. 1986. Proprietary and Nonproprietary Disclosures. *Journal of Business* 59, 331-336.
- [33] Dye, Ronald A. 1990. Mandatory versus Voluntary Disclosures: the Case of Financial and Real Externalities. *Accounting Review* 65, 1-24.
- [34] Dye, Ronald A. 1998. Investor Sophistication and Voluntary Disclosures. *Review of Accounting Studies* 3, 261-287.
- [35] Dye, Ronald A. and Sri S. Sridhar 1995. Industry-wide Disclosure Dynamics. *Journal of Accounting Research* 33, 157-174.
- [36] Edmans, Alex, Mirko Stanislav Heinle, and Chong Huang. 2014. The Real Costs of Financial Efficiency When Some Information is Soft. Working paper.
- [37] Eswaran, Mukesh, and Ashok Kotwal. 1985. A Theory of Contractual Structure in Agriculture. *American Economic Review* 75:352-367.
- [38] Gervais, Simon, and Günter Strobl. 2014. The Industrial Organization of Money Management. Working paper.

- [39] Graham, John R., Roni Michaely, and Michael R. Roberts. 2003. Do Price Discreteness and Transactions Costs Affect Stock Returns? Comparing Ex-Dividend Pricing before and after Decimalization. *Journal of Finance* 58(6):2611-2636.
- [40] Grossman, Sanford J. 1981. The Informational Role of Warranties and Private Disclosures about Product Quality. *Journal of Law and Economics* 24(3):461-483.
- [41] Grossman, Sanford J., and Oliver D. Hart. 1980. Disclosure Laws and Takeover Bids. *Journal of Finance* 35(2):323-334.
- [42] Groves, Theodore. 1973. Incentives in Teams. *Econometrica* 41(4):617-631.
- [43] Gupta, Srabana, and Richard E. Romano. 1998. Monitoring the Principal with Multiple Agents. *The RAND Journal of Economics* 29(2):427-442.
- [44] Heal, Geoffrey. 1976. Do Bad Products Drive Out Good? *The Quarterly Journal of Economics* 90(3):499-502.
- [45] Heal, Geoffrey. 1977. Guarantees and Risk-Sharing. *Review of Economic Studies* 44(3):549-560.
- [46] Hirshleifer, Jack. 1971. The Private and Social Value of Information and the Reward to Inventive Activity. *American Economic Review* 61(4):561-574.
- [47] Holmström, Bengt. 1982. Moral Hazard in Teams. *Bell Journal of Economics* 13(2):324-340.
- [48] Hurst, James Willard. 1984. *Law and Economic Growth: The Legal History of the Lumber Industry in Wisconsin 1836-1915*. The University of Wisconsin Press.
- [49] Inderst, Roman, and Marco Ottaviani. 2009. Misselling Through Agents. *American Economic Review* 99(3):883-908.
- [50] Inderst, Roman, and Marco Ottaviani. 2012a. Competition through Commissions and Kickbacks. *American Economic Review* 102(2):780-809.
- [51] Inderst, Roman, and Marco Ottaviani. 2012b. How (Not) to Pay for Advice: A Framework for Consumer Financial Protection. *Journal of Financial Economics*. 105(2):393-411.
- [52] Inderst, Roman, and Marco Ottaviani. 2013. Sales Talk, Cancellation Terms, and the Role of Consumer Protection. *Review of Economic Studies* 80(3): 1002-1026.

- [53] Iyer, Rajkamal, and Manju Puri. 2012. Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks. *American Economic Review*, 102(4): 1414-45.
- [54] Kronman, Anthony T. 1978. Mistake, Disclosure, Information, and the Law of Contracts. *Journal of Legal Studies* 7(1):1-34.
- [55] La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2006. What Works in Securities Laws? *Journal of Finance* 61(1):1-32.
- [56] La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2008. The Economic Consequences of Legal Origins. *Journal of Economic Literature* 46(2):285-332.
- [57] La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, and Robert W. Vishny. 1997. Legal Determinants of External Finance. *Journal of Finance* 52(3):1131-1150.
- [58] La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, and Robert W. Vishny. 1998. Law and Finance. *Journal of Political Economy* 106(6):1113-55.
- [59] Levine, Ross. 1998. The Legal Environment, Banks, and Long-Run Economic Growth. *Journal of Money, Credit and Banking* 30(3):596-613.
- [60] Levine, Ross. 1999. Law, Finance, and Economic Growth. *Journal of Financial Intermediation* 8:8-35.
- [61] Mazzetti, Mark, and Michael S. Schmidt. 2013. Ex-C.I.A. Worker Says He Disclosed U.S. Surveillance. *New York Times* June 10:A1(L).
- [62] McCann, Colum. 2011. Occupy Wall Street. *The New York Times*. December 25.
- [63] Milgrom, Paul R. 1981. Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics* 12(2):380-391.
- [64] Morris, Stephen, and Hyun Song Shin. 2002. Social Value of Public Information. *American Economic Review* 92(5):1521-34.
- [65] Nunn, Nathan. 2007. Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade. *Quarterly Journal of Economics* 122:569-600.

- [66] Oi, Walter Y. 1973. The Economics of Product Safety. *Bell Journal of Economics and Management Science* 4(1):3-28.
- [67] Palfrey, Thomas, and Thomas Romer. 1983. Warranties, Performance, and the Resolution of Buyer-Seller Disputes. *Bell Journal of Economics* 14(1):97-117.
- [68] Petersen, Mitchell A., and Raghuram G. Rajan. 1994. The Benefits of Lending Relationships: Evidence from Small Business Data. *Journal of Finance* 49(1): 3-37.
- [69] Prendergast, Canice. 1993. The Role of Promotion in Inducing Specific Human Capital Acquisition. *Quarterly Journal of Economics* 108:523-534.
- [70] Puri, Manju, and Jorg Rocholl. 2008. On the Importance of Retail Banking Relationships. *Journal of Financial Economics* 89(2):253-267.
- [71] Ramakrishnan, Ram T.S., and Anjan V. Thakor. 1984. Information Reliability and a Theory of Financial Intermediation. *Review of Economic Studies* 51(3):415-432.
- [72] Rasmusen, Eric B. 2004. Agency Law and Contract Formation. *American Law and Economics Review* 6(2):369-409.
- [73] Romano, Richard E. 1994. Double Moral Hazard and Resale Price Maintenance. *RAND Journal of Economics*. 25(3):455-466.
- [74] Shavell, Steven. 1994. Acquisition and Disclosure of Information Prior to Sale. *RAND Journal of Economics* 25(1):20-36.
- [75] Simon, Marilyn J. 1981. Imperfect Information, Costly Litigation, and Product Quality. *Bell Journal of Economics* 12(1):171-184.
- [76] Simon, Ruth, and Victoria McGrane. 2011. Wells Draws \$85 Million Penalty in Subprime Case. *The Wall Street Journal*: July 21.
- [77] Spence, Michael. 1977. Consumer Misperceptions, Product Failure and Producer Liability. *Review of Economic Studies* 44(3):561-572.
- [78] Spengler, Joseph J. 1968. The Economics of Safety. *Law and Contemporary Problems* 33(3):619-638.
- [79] Stein, Jeremy C. 2002. Information Production and Capital Allocation: Decentralized vs. Hierarchical Firms. *Journal of Finance* 57(5):1891-1921.

- [80] Stiglitz, Joseph E. 1974. Incentives and Risk Sharing in Sharecropping. *Review of Economic Studies*. 41(2):219-255.
- [81] Stiglitz, Joseph E. 2010. *Freefall: America, Free Markets, and the Sinking of the World Economy*. W.W. Norton and Company, Inc., NY.
- [82] Sykes, Alan O. 1984. The Economics of Vicarious Liability. *Yale Law Journal* 93(7):1231-1280.
- [83] Sykes, Alan O. 1988. The Boundaries of Vicarious Liability: An Economic Analysis of the Scope of Employment Rule and Related Legal Doctrines. *Harvard Law Review* 101(3):563-609.
- [84] Taylor, Curtis R. 1995. The Economics of Breakdowns, Checkups, and Cures. *Journal of Political Economy* 103(1):53-74.
- [85] Teoh, Siew Hong. 1997. Information Disclosure and Voluntary Contributions to Public Goods. *RAND Journal of Economics* 28(3):385-406.
- [86] Verrecchia, Robert E. 1983. Discretionary Disclosure. *Journal of Accounting and Economics* 5, 179-194.
- [87] Verrecchia, Robert E. 1990. Information Quality and Discretionary Disclosure. *Journal of Accounting and Economics* 12(4), 365-380.
- [88] Volcker, Paul. 2010. *The Time We Have Is Growing Short*. New York Review of Books. June 24.
- [89] Wolinsky, Asher. 1993. Competition in a Market for Informed Experts' Services. *RAND Journal of Economics* 24(3):380-398.

Biography

Ying Xue was born on February 19, 1987 in Lanzhou City, Gansu Province, China. He was a freshman (2004~5) at the School of Economics and Management of Tsinghua University in Beijing, China, a sophomore (05~6) and senior (07~8) in the Department of Statistics and Actuarial Science in the Faculty of Science at the University of Hong Kong, and a junior (06~7) in the Faculty of Mathematics at the University of Waterloo in Ontario, Canada. Ying earned a Bachelor of Science in Actuarial Science from the University of Hong Kong in June 2008 and a Master of Business Research in Finance from Stanford University in June 2011, and expects a PhD in Business Administration (Finance) from Duke University in September 2014. Ying will immediately and permanently return to China and serve the Chinese government after graduation.

Thanks to the strong recommendation of his advisor and chair Prof. Simon Gervais, Ying won 1 of 118 travel awards worldwide to the 2014 American Finance Association Annual Meeting. Thanks to the strong recommendation of his PhD committee member Prof. Vish Viswanathan, Ying won 1 of 63 travel awards worldwide to the 2013 Princeton Initiative Conference. Ying was 1 of 100 elected Goldman Sachs Global Leaders worldwide, 1 of 20000 certified actuaries worldwide, 1 of 9 worldwide awarded by the Actuarial Foundation the John Culver Woody Scholarship, and 1 of 108 delegates worldwide to the London International Youth Science Forum. Ying won 1st Prizes in National Math and Physics (provincial champion) Olympiads.