# Essays on Information Economics in Games

by

Rosina Rodríguez Olivera

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2021

Doctoral Committee:

Associate Professor David Miller, Co-Chair
Assistant Professor Anne-Katrin Roesler, Co-Chair
Professor Tilman Borgers
Associate Professor Vijay Subramanian

Rosina Rodríguez Olivera

rorodrig@umich.edu

ORCID ID: 0000-0001-7232-2082

# ACKNOWLEDGEMENTS

First and foremost I am extremely grateful to my two main supervisors, Anne-Katrin Roesler and David Miller, for the constant support, encouragement and invaluable feedback and guidance. I am also deeply grateful to Tilman Borgers and Vijay Subramanian for their guidance, insights and advice.

I am also thankful for the unending love and support of my family and friends, especially my father Raúl, my mother Adriana, my sister Camila and my best friend Paola. Lastly, I would like to thank my fiancé Louis-Pierre for his constant encouragement, loving support and enlightening discussions.

# TABLE OF CONTENTS

# LIST OF FIGURES

**Figure**

# LIST OF TABLES

**Table**

# LIST OF APPENDICES

**Appendix**

# ABSTRACT

My dissertation focuses on information economics in games. In Chapter II, "*Optimal Disclosure of Private Information to Competitors*", I consider a duopoly model with differentiated substitutes, price competition, and uncertain demand, in which one firm has an information advantage over a competitor. I study the incentives of the informed firm to share its private information with its competitor and the incentives of a regulator to control disclosure in order to benefit consumers. I show that full disclosure of information is optimal for the informed firm, because it increases price correlation and surplus extraction from consumers. I also show that the regulator can increase expected consumer surplus and welfare by restricting disclosure, but that, surprisingly, consumers can benefit from the regulator privately disclosing some information to the competitor. My findings highlight the consequences of an unequal distribution of consumer data between firms on welfare allocation. They also inform an ongoing policy debate about regulating what data online platforms release to other firms who offer goods on its platform.

In Chapter III, "*Strategic Incentives and the Optimal Sale of Information*", I study the optimal sale of information by a monopolist data-seller to multiple privately informed data-buyers who play a two-stage game of incomplete information. In the information stage, data-buyers can simultaneously acquire supplemental information to reduce their uncertainty about the state. In the action stage, data-buyers simultaneously select an action to maximize their expected payoffs. The data-seller offers a menu of Blackwell experiments and prices to screen between two types of data-buyers. I show that the nature of data-buyer's preferences for information allows

the data-seller to extract all surplus from data-buyers, distorting the information provided to the low type such that the high type is indifferent between both experiments. I also show that the features of the optimal menu are determined by the interaction between data-buyers' strategic incentives in the action stage and the correlation of their private information. This interaction can expand the data-seller's ability to serve all segments of the market, increasing expected profits.

In Chapter IV, "*Product reviews - Information Source or Persuasion Device?*", which is joint work with Anne-Katrin Roesler, we study the optimal design of review systems by a platform that has the best interest of consumers in mind. We consider a setting in which a seller offers a good of ex-ante unknown quality through a platform with a review system to sequentially arriving short-lived heterogeneous buyers. Reviews from previous buyers provide consumers with information about the good's quality. Based on the review system, the seller chooses an optimal pricing scheme. Buyers make their purchasing decision based on the information available through reviews, their type, and the price. In a two period setting, we approximate the optimal review system by characterizing the optimal K-piecewise linear distribution over posterior quality estimates.

# CHAPTER I

# Introduction

My dissertation focuses on information economics in games, specifically how information affects the decisions of agents facing uncertainty. Information design studies how an "information designer" can influence agents' individually optimal behavior, by providing information about a payoff-relevant state. In games of incomplete information, the information designer commits to an information structure to affect the behavior of strategically interacting agents. By revealing information, the designer affects the players' beliefs about the payoff-relevant state and their beliefs about the information observed by others, which determine players' optimal choices.

This dissertation studies how an information designer should provide information to multiple privately informed agents who play three different games of incomplete information. I characterize the optimal disclosure of private information between firms to protect consumers, the optimal sale of information by a data-seller to interacting data-buyers, and the design of a review system that allows consumers to share information through product reviews. In each chapter, before observing the payoff-relevant state, a designer commits to an information structure which maps from the state to a joint distribution over signals. Given this information structure, each player observes a signal realization and updates her belief about the state and the information observed by others. Each player then chooses an action to maximize her expected payoff. In this context, the designer chooses the information structure such that players play the Bayes Nash equilibrium that maximizes her own expected payoff.

In Chapter II, "*Optimal Disclosure of Private Information to Competitors*", I consider a duopoly model with differentiated substitutes, price competition, and uncertain demand, in which one firm has an information advantage over a competitor. I study the incentives of the informed firm to share its private information with its competitor and the incentives of a regulator to control disclosure in order to benefit consumers.

I show that full disclosure of information is optimal for the informed firm, because it increases price correlation and surplus extraction from consumers. I also show that the regulator can increase expected consumer surplus and welfare by restricting disclosure, but that, surprisingly, consumers can benefit from the regulator privately disclosing some information to the competitor. Disclosure increases the ability of firms to extract surplus from consumers by pricing to better match the level of demand. But, private disclosure can create a pricing coordination failure between firms by introducing uncertainty in their choices, which increases price volatility and opportunities for consumers to arbitrage prices. The benefit from private disclosure depends on the differentiation between goods, because it determines consumers' willingness to substitute between goods and therefore the extent to which disclosure affects relative demand across firms. Thus, I show that private partial disclosure is optimal for consumers when firms offer sufficiently close substitutes and, otherwise, no disclosure is optimal. When partial disclosure is optimal, I also fully characterize the consumer optimal disclosure policy.

My findings highlight the consequences of an unequal distribution of consumer data between firms on welfare allocation. They also inform an ongoing policy debate about regulating what data online platforms release to other firms who offer goods on its platform.

In Chapter III, "*Strategic Incentives and the Optimal Sale of Information*", I study the optimal sale of information by a monopolist data-seller to multiple privately informed data-buyers who play a two-stage game of incomplete information. In the information stage, data-buyers can simultaneously acquire supplemental information to reduce their uncertainty about the state. In the action stage, data-buyers simultaneously select an action to maximize their expected payoffs. The data-seller offers a menu of experiments and prices to screen between two types of data-buyers. In contrast to standard screening problems, buyer's preferences for information cannot generally be ordered across types since they value information if and only if it affects their choices.

I show that the nature of data-buyer's preferences for information allows the data-seller to extract all surplus from data-buyers, distorting the information provided to the low type such that the high type is indifferent between both experiments offered in the menu. I also show that the features of the optimal menu are determined by the interaction between data-buyers' strategic incentives in the action stage and the correlation of their private information. This interaction can expand the data-seller's ability to serve all segments of the market, increasing expected profits. In particular, the seller's optimal menu offers perfect information to the buyer with highest willingness to pay (the high type) and partial or no information to the other type (the low type). Partial information is offered to the low type whenever the mentioned interaction increases demand for information. That is, i) if they play a coordination game and their private information is negatively correlated; or ii) if they play anti-coordination game and their private information is positively correlated.

In Chapter IV, "*Product reviews - Information Source or Persuasion Device?*", which is joint work with Anne-Katrin Roesler, we study the optimal design of review systems by a platform that has the best interest of consumers in mind. We consider a setting in which a seller offers a good of ex-ante unknown quality through a platform with a review system to sequentially arriving short-lived heterogeneous buyers. Reviews from previous buyers provide consumers with information about the good's quality. Based on the review system, the seller chooses an optimal pricing scheme. Buyers make their purchasing decision based on the information available through reviews, their type, and the price. When buyers have homogeneous preferences, we fully characterize the buyer-optimal review system. We also show that the presence of heterogeneous buyers meaningfully affects the characterization of the buyer-optimal distribution over posterior quality estimates. In a two period model, we characterize the buyer-optimal distribution over posterior quality estimates within the class of $K$-piecewise linear distributions with $K \leq 2$.

# CHAPTER II

# Optimal Disclosure of Private Information to Competitors

## 2.1 Introduction

Some firms can gather more information than their competitors about market features like demand, given their size or incumbency status. For instance, online platforms like Amazon engage in massive data collection and analysis by gathering and processing information generated through trade and consumer searches that other sellers on the platform can't replicate. As a seller themselves, they can use this information to guide their own pricing and control the information observed by other sellers. In settings of information asymmetry, information disclosure between firms affects firm behavior and therefore also impacts consumers and welfare. The use of private information as a competitive advantage by online platforms has attracted the attention of regulatory entities in the US and Europe.[1] When there is an uneven distribution of consumer data between firms, regulatory interventions to control information disclosure can potentially redistribute surplus between firms and consumers or increase welfare.

In this paper, I study the role of information disclosure as a pricing persuasion device, through which a firm with an information advantage or a regulator can influence the pricing of a competing firm. I examine the informed firm's incentives to commit to share its private information with its competitor and the role of a regulator who commits to control information disclosure between firms to benefit consumers.

[1]See for example media coverage in Fung (2020), Green (2018), Lardieri (2019).

Specifically, I analyze a stylized duopoly model with information asymmetry about demand and a binary state. I consider an informed firm which privately learns the level of the demand and an uninformed firm that has no private information. Demand is linear and firms face uncertainty about its level, which can be either low or high. Firms offer differentiated goods, such that consumer willingness to pay for a good depends on its substitutability with the competitor's. Firms compete by simultaneously and non-cooperatively setting prices to maximize their expected profits. In this context, I address the following questions: What is the informed firm's optimal disclosure policy as a competitor in the market? How can a regulator constrain or enforce information disclosure to benefit consumers?

I characterize the optimal disclosure for firms and consumers. I show that the welfare implications of disclosure are determined by the degree of differentiation between goods, because it determines the extent to which disclosure affects firm pricing and relative demand across markets. Regarding optimal disclosure for firms, with substitutes, firm choices are strategic complements and the informed firm thus benefits from sharing its private information with the uninformed firm, because it increases price correlation. As a result, full disclosure is optimal for the informed firm. Full disclosure also maximizes producer surplus, because the uninformed firm also benefits from price correlation as well as from learning about the state. This result highlights that an informed firm may have incentives to share information even when it has no information to gain in return, because it can influence the pricing of its competitor. Furthermore, I generalize this result by showing that the informed firm's optimal disclosure doesn't rely on the linearity of demand: when the informed firm's expected equilibrium profit is supermodular in the state and the choice of the uninformed firm, firms' choices are strategic complements and, accordingly, full disclosure is optimal. I also show that no disclosure is optimal when the informed firm's expected equilibrium profit is submodular.[2]

Regarding optimal disclosure for consumers, I find that a regulator should restrict information disclosure, at least partially.[3] However, some information disclosure is

[2]When the informed firm's profits are supermodular in the state and the choice of the uninformed firm, an increase in the uninformed firm's price has a increasing effect on the informed firm's profits as the state increases. Focusing instead on decision problems, Kolotilin and Wolitzky (2020) shows that supermodularity of a sender's objective function with respect to the state and the receiver's action is a sufficient condition for the optimality of full disclosure.

[3]Luco (2019) presents empirical evidence that full disclosure can be detrimental for consumers

not necessarily detrimental to consumers. First, I show that the optimal disclosure is private, such that the informed firm doesn't observe the signal realization of the uninformed firm. Second, I show that optimal disclosure is determined by the degree of differentiation between goods. In particular, partial disclosure is optimal if firms offer sufficiently close substitutes and no disclosure is optimal otherwise. Information disclosure creates a trade-off for consumers. On the one hand, disclosure reduces the uninformed firm's uncertainty about the state, improving the ability of firms to extract surplus from consumers by increasing price correlation. On the other hand, private partial disclosure introduces uncertainty about the information observed by a firm's competitor. This expands the range of prices in each state, since firms price according to the expected price of its competitor and its expected level of demand. Namely, the uninformed firm may observe a signal which conflicts with the realized state, but the informed firm doesn't observe the signal and therefore cannot adjust, creating a coordination failure. Consumers can benefit from this price heterogeneity by choosing from which firm to buy after observing prices. Overall, the regulator trades-off the opportunity to create this coordination failure in prices with allowing firms to better extract surplus from consumers. The net effect depends on the differentiation between goods, because it determines consumers' willingness to substitute between goods and therefore the extent to which disclosure affects relative demand across firms.

Combining results for firms and consumers, to maximize expected welfare, the regulator trades off the effect of disclosure on consumers and firms. When firms offer sufficiently differentiated goods, no disclosure maximizes expected welfare, since the expected loss of some disclosure for consumers exceeds the expected gain for firms. Conversely, when firms offer sufficiently close substitutes, full disclosure maximizes welfare. For intermediate levels of differentiation, partial disclosure maximizes welfare.

When partial disclosure is optimal, I also fully characterize the consumer and welfare optimal disclosure policies, which share the same qualitative properties. Signals act as equilibrium price recommendations, recommending a price to each firm conditional on the state subject to obedience constraints. First, I show that the regulator recommends at most two prices. Second, I show that one of the prices is only rec-

in the gasoline market in Chile.

ommended when the state is low, revealing the state to the uninformed firm. The other price is recommended in both states, obfuscating the level of demand. The optimality of partial disclosure contrasts with work in the literature focusing on firm incentives, highlighting that optimal disclosure can be more nuanced when considering implications for consumers and welfare.

My analysis emphasizes the wide scope for intervention by a regulator, based on product differentiation and their objective function. My results are of particular interest given current policy debates on the use of private information by firms who act as both a trading platform and a competitor in the market. As I show, it can be optimal for a regulator to intervene by completely preventing or forcing information disclosure, or by designing disclosure policies to partially inform the uninformed firm. Therefore, it is crucial to consider the strategic environment to understand the consequences of information sharing between firms, and whether it is beneficial for both firms and consumers.

**Related literature.** This paper contributes to the literature on strategic information sharing in oligopolies with commitment and the literature on information design in games.[4] Incentives for information sharing about demand among competing firms with symmetric private information and normally distributed linear demand were first studied in Novshek and Sonnenschein (1982), Clarke (1983) as well as Vives (1984), and later generalized in Raith (1996).[5] In these papers, firms commit to share their private information with an intermediary, which then discloses a common signal to all firms to maximize industry-wide profits. These papers focus on the producer surplus optimal public disclosure and on the regulation of industry-wide information sharing by trading organizations. They show the optimality of full disclosure for firms when they compete by choosing prices and offer imperfect substitutes. Instead, I study the incentives of an individual firm to share information to influence its competitor's behavior in a setting of informational advantage, in which signals are privately observed and the distribution of the uninformed firm's signal

[4]Papers like Benoit and Dubra (2006) show that agents' ex-ante and ex-post incentives for information sharing can be disaligned, such that commitment is key.

[5]Other papers in this literature include Gal-Or (1985), Li (1985), Kirby (1988) and Vives (1990). Information sharing about costs are studied in papers like Fried (1984), Gal-Or (1986), Sakai (1986) and Shapiro (1986), in which incentives to share information are reversed.

is unrestricted.[6] My results show that it can be optimal for a firm to unilaterally disclose information about demand to a competitor even without receiving any information in return, because disclosure can influence the behavior of competitors and act as a pricing persuasion device.[7] Further, full disclosure is not only optimal for the informed firm, but also for producer surplus. The intuition for this result relates to Angeletos and Pavan (2007), who study the social value of information with normally distributed signals and find that producer surplus increases with the precision of both public and private signals.

In contrast with this literature, I also analyze the effects of information disclosure on consumers. Vives (1984) and Calzolari and Pavan (2006) show that information disclosure is not necessarily harmful to consumers. Vives (1984) illustrates this by comparing the utility of a representative consumer across full and no disclosure when firms share symmetric normally distributed private information. Calzolari and Pavan (2006) study a sequential setting in which the Stackelberg leader must provide incentives to consumers to reveal their private information to be able to share it with its follower. They focus on the leader's optimal disclosure policy, whereas I focus on the optimal disclosure for consumers. Regarding welfare, Vives (1984) also shows that full disclosure dominates no disclosure if and only if firms offer sufficiently close substitutes, yet I show that restricting to full and no disclosure is with loss of generality since partial disclosure can be consumer and welfare optimal. My results regarding welfare relate to Ui and Yoshizawa (2015), who study the social value of information restricted to symmetric normally distributed signals and symmetric equilibria. They show that welfare decreases in the precision of private information and increases in the precision of public information if goods are close substitutes, intuitively related to the optimality of either full or private partial disclosure as I fully characterize in this paper.

More broadly, the paper contributes to the literature on information design in games

[6] Bergemann and Morris (2013), Bergemann, Heumann, and Morris (2015) and Eliaz and Forges (2015) analyze producer optimal disclosure in Cournot settings with perfect substitutes and information symmetry. They show that it is with loss of generality to restrict attention to a common and, hence, perfectly correlated disclosure.

[7] In sequential settings, the role of current choices as a costly persuasion device to influence the precision of future information has been studied in Mailath (1989), Mirman, Samuelson, and Urbano (1993), Mirman, Samuelson, and Schlee (1994), Harrington (1995), Keller and Rady (2003), Taylor (2004), Bernhardt and Taub (2015), Bonatti, Cisternas, and Toikka (2017).

as studied in papers like Taneva (2019) and Mathevet, Perego, and Taneva (2020). I characterize the optimal recommendation mechanism in a Bertrand setting with product differentiation and information asymmetry. It is most closely related to the literature on consumer optimal information design, which analyzes the effect of information about buyers' valuation on pricing and welfare allocation. This literature has focused on buyer optimal learning, consumer optimal market segmentation and on the incentives of consumers to disclose their preferences directly to firms. Within the buyer optimal learning literature, Roesler and Szentes (2017) analyzes the effect of a buyer's information on monopoly pricing and characterizes optimal buyer learning. In a duopoly setting, Armstrong and Zhou (2019) studies competition between firms when consumers observe a private signal about their valuation and characterizes consumer optimal learning. Within the consumer optimal segmentation literature, Bergemann, Brooks, and Morris (2015) analyzes the welfare consequences of a monopolist having access to additional information about consumer preferences and characterize the feasible welfare outcomes achieved by segmentation. Li (2020) extends the insights from Bergemann, Brooks, and Morris (2015) to an oligopoly setting and characterizes the consumer-optimal market segmentation in competitive markets. Elliott, Galeotti, and Koh (2020) studies how information about consumer preferences should be distributed across firms which compete by offering personalized discounts to consumers and provides necessary and sufficient conditions under which perfect segmentation can be achieved. Lastly, Ichihashi (2020) studies the welfare effects of consumers disclosing information about their valuation with a monopolist, whereas Ali, Lewis, and Vasserman (2020) analyzes the consumer optimal disclosure of information about their preferences in monopolistic and competitive markets. In contrast, I focus on the welfare consequences of an unequal distribution of consumer data across firms and the effect of information disclosure between firms. In particular, I characterize the optimal disclosure policy between firms for consumers, which affects consumers indirectly by affecting prices.

The remainder of the paper is organized as follows: Section 1 presents an example, Section 2 outlines the model and preliminary results, Section 3 derives the informed firm optimal disclosure, Section 4 derives the consumer optimal disclosure, Section 5 derives the producer and welfare optimal disclosures, Section 6 discusses extensions, and Section 7 concludes.

## 2.2 Motivating example

Two firms compete by simultaneously setting prices and offer differentiated substitutes. Firm profits depend on the realization of a binary state, $\theta \in \Theta$, which represents the level of demand. For this example, assume that $\Theta = \{1, 2.8\}$, where the low state occurs with probability $\mu_L = \frac{3}{4}$ and the high state with probability $\mu_H = \frac{1}{4}$. Firm $i$'s demand is given by

$$q_i((p_i, p_{-i}); \theta) = \max\{0, \theta - 2.1p_i + 2p_{-i}\}.$$

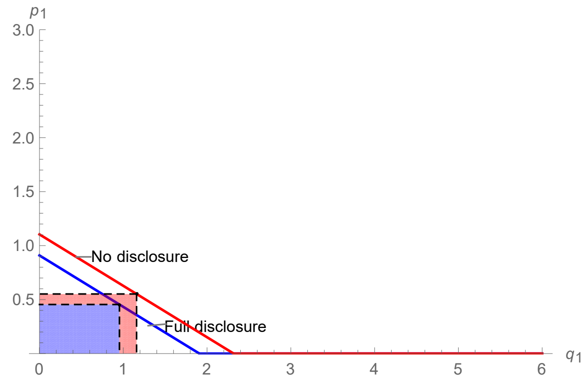
The effect of $i$'s pricing decision on its quantity demanded exceeds the effect of $j$'s pricing decision, representing the differentiation between goods. Firms' costs are zero.

Firm 1 (the informed firm) observes the state, whereas firm 2 (the uninformed firm) initially has no information beyond the prior. Before the realization of the state, the informed firm commits to an information structure which discloses none, some, or all of its private information to the uninformed firm. For simplicity, in this example, I restrict attention to information structures with binary support: conditional on the state, the uninformed firm privately observes either a low signal $s_L$ or a high signal $s_H$. In the low state $\theta_L = 1$, the low signal is observed with probability $x_L \in [0, 1]$, whereas the high signal is observed with the complementary probability $1 - x_L$. In the high state $\theta_H = 2.8$, the high signal is observed with probability $x_H \in [0, 1]$ and the low signal with probability $1 - x_H$. Signal $s_L$ is more likely to occur than $s_H$ when the state is low than when state is high, implying that $x_L + x_H \geq 1$. Denote the information structure by $(x_L, x_H)$, where full (no) disclosure is captured by $x_L = x_H = 1$ ($x_L = x_H = \frac{1}{2}$).

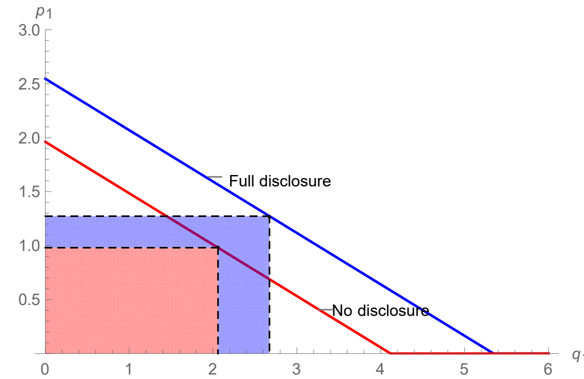
Given the information structure, firms choose a pricing strategy, which consists of a price conditional on their private information to maximize expected profits. Denote by $p_1^*(\theta; (x_L, x_H))$ the informed firm's equilibrium price when the state is $\theta$ and by $p_2^*(s; (x_L, x_H))$ the uninformed firm's equilibrium price after observing signal $s$.

The informed firm's expected equilibrium profits are strictly increasing in the precision of the information structure.[8] Hence, it is optimal for it to commit to share

[8]The informed firm's expected equilibrium profits, $\mathbb{E}[\Pi_1^*(x_L, x_H)] = 2.1\mathbb{E}\left[p_1^*(\theta; (x_L, x_H))^2\right]$ , are




(a) Market of firm 1: $\theta = 1$




(b) Market of firm 1: $\theta = 2.8$

Figure 2.1: Informed firm's expected demand and profits in each state with full disclosure (blue line and blue shaded region) vs no disclosure (red line and red shaded region).

all of its private information with the uninformed firm. In Figure 2.1, I illustrate the optimality of full disclosure, showing its effect on prices, demand and profits in the informed firm's market. Changes in the information observed by the uninformed firm impact equilibrium pricing and, as a result, the expected demand faced by the informed firm. From the informed firm's perspective, a more precise signal increases the uninformed firm's expected equilibrium price in the high state and lowers it in the low state.[9] As a result, with substitutes, a more precise information disclosure increases (decreases) the informed firm's expected demand in the high (low) state, shifting the demand outward (inward) in Figure 2.1. In the high state, the informed firm increases its profits by raising the price on inframarginal consumers who were already buying its product and gaining marginal consumers from the uninformed firm's market. In the low state, the informed firm's profits are lower since it charges a lower price and faces a lower level of demand. Nevertheless, the bigger size of the

strictly increasing in both $x_L$ and $x_H$, because

$$\frac{\partial \mathbb{E}[\Pi_1^*(x_L, x_H)]}{\partial x_k} = 2.1 \sum_{\theta \in \Theta} \mu_\theta p_1^*(\theta; (x_L, x_H)) \frac{\partial p_1^*(\theta; (x_L, x_H))}{\partial x_k}, \ \mu_H \frac{\partial p_1^*(\theta; (x_L, x_H))}{\partial x_k} = -\mu_L \frac{\partial p_1^*(\theta; (x_L, x_H))}{\partial x_k} > 0$$

and the informed firm sets a higher price in the high state.

[9]The uninformed firm's optimal price is linear in its expectation of the state. Then, changes in its pricing across states due to disclosure depends on its prior. After observing the low signal $s_L$, it lowers its price by a relatively small amount if it already believed that the low state was more likely. In contrast, after observing the high signal $s_H$, its price increase is bigger, reflecting its change in beliefs about the state.

market implies that the stakes are higher in the high state, such that the gains from a price bonus there outweigh the losses from a price penalty in the low state. As a result, it is optimal for the informed firm to disclose its private information.

However, full disclosure is detrimental for consumers. Consumers are better off with no disclosure between firms, as illustrated in Table 2.1. Still, while completely restricting disclosure may seem optimal, private partial disclosure can benefit consumers. For instance, private partial disclosure characterized by $(x_L, x_H) = \left(\frac{1}{2}, 1\right)$ yields a higher expected consumer surplus than no disclosure, as also illustrated in Table 2.1.[10]

| | No: $(x_L, x_H) = \left(\frac{1}{2}, \frac{1}{2}\right)$ | Partial: $(x_L, x_H) = \left(\frac{1}{2}, 1\right)$ | Full: $(x_L, x_H) = (1, 1)$ |
|---|---|---|---|
| $(p_1^*(\theta_L), p_1^*(\theta_H))$ | $(0.55, 0.98)$ | $(0.54, 1)$ | $(0.45, 1.27)$ |
| $(p_2^*(s_L), p_2^*(s_H))$ | $(0.66, 0.66)$ | $(0.49, 0.76)$ | $(0.4, 1.27)$ |
| $CS$ | 1.26 | 1.27 | 1.18 |

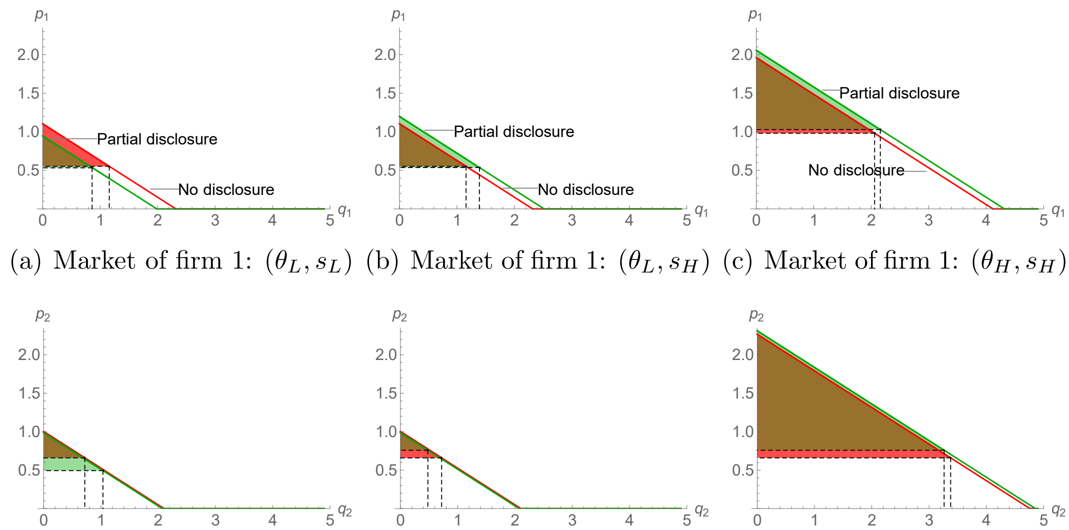
Table 2.1: Prices and Consumer surplus comparison with no, partial and full disclosure.

Intuitively, private partial disclosure impacts consumer surplus in two ways. First, it gives the uninformed firm information about the state, improving the ability of firms to extract surplus from consumers. Second, it introduces uncertainty in the pricing of firms, because each firm is uncertain about the information observed by its competitor. This increases the range of prices in each state, such that consumers can benefit from price arbitrage by choosing from which firm to buy after observing prices. In fact, it creates a pricing coordination failure between firms with positive probability, given that they choose prices without observing its competitor's information.

Figure 2.2 illustrates ex-post consumer surplus with no disclosure and partial disclosure $(x_L, x_H) = \left(\frac{1}{2}, 1\right)$ in each market. Each panel illustrates consumer surplus when the realized state is $\theta$ and the uninformed firm observes signal $s_2$. With this partial disclosure, the uninformed firm observes signal $s_H$ with probability 1 when the state is high, so only three panels per market are displayed.

In the uninformed firm's market, expected consumer surplus decreases with partial disclosure due to changes in its pricing. There is little change in its demand, since

[10] Consumers benefit from partial disclosure when it is private, in which case the informed firm doesn't observe the signal realization of the uninformed firm. If disclosure is public, no disclosure is optimal.



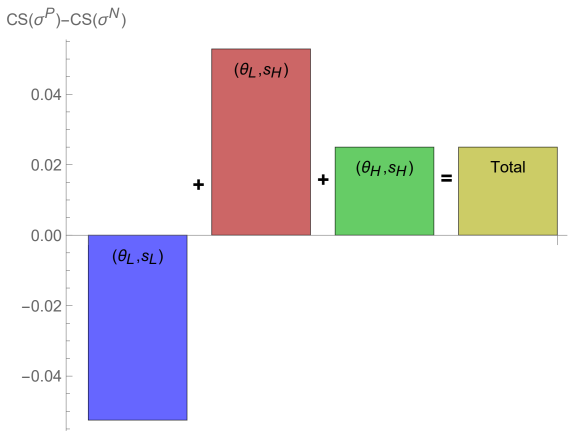

(a) Market of firm 1: $(\theta_L, s_L)$ (b) Market of firm 1: $(\theta_L, s_H)$ (c) Market of firm 1: $(\theta_H, s_H)$

(d) Market of firm 2: $(\theta_L, s_L)$ (e) Market of firm 2: $(\theta_L, s_H)$ (f) Market of firm 2: $(\theta_H, s_H)$

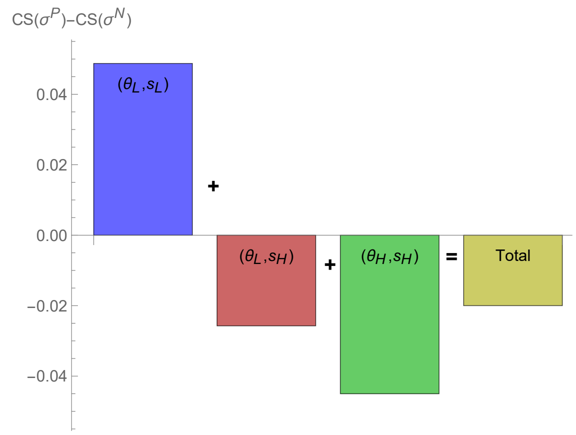
Figure 2.2: Ex-post consumer surplus with no disclosure (shaded red area) vs partial disclosure $(x_L, x_H) = \left(\frac{1}{2}, 1\right)$ (shaded green area).

the informed firm knows the state and therefore does not substantially change its pricing based on the disclosure policy. As shown in panel (d) (panels (e) and (f)), when the uninformed firm observes the low (high) signal, it decreases (increases) its price, increasing (decreasing) quantity sold and consumer surplus. The net effect on expected consumer surplus is negative due to the larger size of the market in the high state, as illustrated in panel (b) of Figure 2.3.

In the informed firm's market, expected consumer surplus increases with partial disclosure. In contrast to market 2, the pricing of the informed firm changes little across disclosure policies, such that changes in consumer surplus are driven by changes in the level of demand from the uninformed firm's pricing. In particular, the informed firm's demand shifts inward when the uninformed firm observes the low signal (panel (a)) and outward when it observes the high signal (panels (b) and (c)). Accordingly, most of the gain in consumer surplus occurs when the uninformed firm observes the high signal, especially when the state is low. In that case, there is a mismatch between the signal realization and the state, implying that the uninformed firm charges a high price. However, since the informed firm doesn't observe the uninformed firm's signal, it can't take advantage of the uninformed firm's high price. Through price arbitrage, consumers avoid the high price in the uninformed firm's market by buying




(a) Market of firm 1




(b) Market of firm 2

Figure 2.3: Weighted change in consumer surplus from partial disclosure vs. no disclosure

from the informed firm at the lower price. Overall, as shown in panel (a) of Figure 2.3, the expected change in consumer surplus for market 1 is positive and exceeds the loss in market 2, showing that partial disclosure is optimal for consumers.

## 2.3 The model

Two symmetric firms offer horizontally differentiated substitutes and compete by simultaneously setting prices. Firm profits depend on the realization of a binary payoff-relevant state, $\theta \in \Theta = \{\theta_L, \theta_H\}$ with $\theta_H > \theta_L > 0$. Firms share a common prior about the state where the probability of $\theta \in \Theta$ is denoted by $\mu_\theta \in (0, 1)$. Firm $i$'s demand, $q_i((p_i, p_{-i}); \theta)$, is given by

$$q_i((p_i, p_{-i}); \theta) = \max\{0, \theta - ap_i + bp_{-i}\} \tag{2.1}$$

where $a$ and $b$ are known parameters with $a > b > 0$.[11] As can be seen from (2.1), the state represents the level of demand and, since firms offer substitutes, both the state and the price of the competitor are positive demand shifters which increase quantity demanded at every price. Define $\delta$ as the ratio of $b$ and $a$, which measures the degree of differentiation. As $\delta$ converges to 0, goods are more differentiated and

[11]The relationship between $a$ and $b$ implies that demand is more sensitive to a firm's own price than the price of its competitor, ensuring that equilibrium prices are finite.

as $\delta$ converges to 1, goods are closer substitutes. Also, assume that

$$\theta_H < \frac{4a^2 - b^2}{2a^2 - b^2}\theta_L,$$

ensuring that equilibrium prices and quantities are non-negative. Firms' costs are zero.[12] Hence, firm $i$'s ex-post profits, $\Pi_i : \mathbb{R}_+^2 \times \Theta \to \mathbb{R}$, correspond to

$$\Pi_i((p_i, p_{-i}); \theta) = p_i \cdot q_i((p_i, p_{-i}); \theta).$$

Given the state and prices $(p_i, p_{-i})$, ex-post consumer surplus in the market of firm $i$ is the difference between consumers' ex-post willingness to pay for the good and the equilibrium price. The ex-post willingness to pay of consumers in the market of firm $i$ is characterized by the ex-post inverse demand, $p_i(q_i; p_{-i}, \theta)$, given by

$$p_i(q_i; p_{-i}, \theta) = \max\left\{0, \frac{\theta + bp_{-i} - q_i}{a}\right\}$$

where demand is generated by a continuum of heterogeneous consumers making discrete choices (Armstrong and Vickers, 2015).[13] Then, ex-post consumer surplus

[12]Including linear or quadratic costs has no impact on the results.

[13]Assume a continuum of consumers with heterogeneous preferences. Consumer $\ell$ has valuation $v_{\ell,i}$ for one unit of the good offered by firm $i$, where $v = (v_{\ell,1}, v_{\ell,2})$ is drawn from a joint cumulative distribution $G(v)$. Consumer $\ell$ attaches no value to more than one unit of either good and wishes to buy either a single unit of one good or to not buy any of them. Then, consumer $\ell$ buys from firm $i$ if $v_{\ell,i} - p_i \geq \max_{j \neq i}\{0, v_{\ell,j} - p_j\}$, where the outside option is normalized to zero. The demand for product $i$, $q_i(p)$, is then the measure of consumers $\ell$ who satisfy $v_{\ell,i} - p_i \geq \max_{j \neq i}\{0, v_{\ell,j} - p_j\}$. Armstrong and Vickers (2015) shows that the linear demand model defined by (2.1) can be micro-founded by this discrete choice model, since

$$\frac{\partial q_i((p_i, p_{-i}); \theta)}{\partial p_{-i}} = \frac{\partial q_{-i}((p_i, p_{-i}); \theta)}{\partial p_i} \text{ for all } i \text{ and } \frac{\partial^2 \sum_{i \in \{1,2\}} q_i((p_i, p_{-i}); \theta)}{\partial p_i \partial p_{-i}} \leq 0.$$

In this context, consumer $\ell$ who buys from firm $i$ receives surplus $v_{\ell,i} - p_i$ and the consumer surplus in market $i$ is simple the sum of the surpluses of consumers $\ell$ who purchase good $i$. Given that there is a continuum of consumers, this coincides with (2.2). Furthermore, since consumers buy at most one product, total ex-post consumer surplus is simply the sum across markets. See Choné and Linnemer (2020) for a survey of micro-foundations of a linear demand system.

in market of firm $i$, $CS_i : \mathbb{R}_+^2 \times \Theta \to \mathbb{R}$, corresponds to

$$CS_i((p_i, p_{-i}); \theta) = \frac{1}{2}\left[\frac{\theta + bp_{-i}}{a} - \frac{\theta + bp_{-i} - q_i((p_i, p_{-i}); \theta)}{a}\right] q_i((p_i, p_{-i}); \theta)$$
$$= \frac{1}{2a} q_i((p_i, p_{-i}); \theta)^2, \tag{2.2}$$

where the term in square brackets corresponds to the difference between the demand intercept and the equilibrium price.

**Information environment.** Firm 1 (the informed firm) knows the state, whereas firm 2 (the uninformed firm) initially has no information beyond the common prior. Assume that a designer can restrict (or require) information sharing between firms by choosing the information observed by the uninformed firm. The designer selects and commits to an information structure before the realization of the state, which discloses none, some, or all of the informed firm's private information to the uninformed firm. Let $S_2$ be the set of signal realizations observed by firm 2. Signal realizations are private.[14] An information structure consists of a set of signal realizations $S_2$ and a family of conditional distributions $\psi_2 : \Theta \to \Delta(S_2)$.

**Pricing game.** Given the information structure $(S_2, \psi_2)$, firms play a pricing game in which they condition their pricing choices on their information by selecting mappings

$$\hat{\beta}_1 : \Theta \to \Delta(\mathbb{R}_+) \text{ and } \hat{\beta}_2 : S_2 \to \Delta(\mathbb{R}_+)$$

to maximize their expected profits. Specifically, the timing is as follows: (i) the designer selects and commits to an information structure $(S_2, \psi_2)$; (ii) the state $\theta$ is realized; (iii) the signal is realized and privately observed by the uninformed firm according to $(S_2, \psi_2)$; (iv) firms update their beliefs according to Bayes' rule and simultaneously choose prices; (v) payoffs are realized.

The solution concept is Bayes Nash equilibrium (BNE). A strategy profile $(\hat{\beta}_1, \hat{\beta}_2)$ is

[14] Allowing for public signals has no effect on the optimal disclosure for firms. However, consumers are better off when signals are private for any disclosure policy. Then, a designer concerned about consumers would optimally commit to private disclosure. See Lemma II.8.

a BNE if, for all $p_i \in \mathbb{R}_+$ selected with positive probability,

$$\int_{S_2}\int_{\mathbb{R}_+} \Pi_1((p_1,p_2);\theta)\mathrm{d}\hat{\beta}_2(p_2|s_2)\mathrm{d}\psi_2(s_2|\theta) \geq \int_{S_2}\int_{\mathbb{R}_+} \Pi_1((p_1',p_2);\theta)\mathrm{d}\hat{\beta}_2(p_2|s_2)\mathrm{d}\psi_2(s_2|\theta) \tag{2.3}$$

for all $p_1' \in \mathbb{R}_+$ and $\theta \in \Theta$ and

$$\sum_{\theta\in\Theta}\mu_\theta \int_{\mathbb{R}_+} \Pi_2((p_2,p_1);\theta)\mathrm{d}\hat{\beta}_1(p_1|\theta) \geq \sum_{\theta\in\Theta}\mu_\theta \int_{\mathbb{R}_+} \Pi_2((p_2',p_1);\theta)\mathrm{d}\hat{\beta}_1(p_1|\theta) \tag{2.4}$$

for all $p_2' \in \mathbb{R}_+$ and $s_2 \in S_2$. Denote by $\hat{\mathcal{E}}(S_2,\psi_2)$ the set of BNE in the pricing game.

For any information structure $(S_2,\psi_2)$, the existence and uniqueness of the BNE is guaranteed by Ui (2016), which provides sufficient conditions for the existence and uniqueness of the BNE in Bayesian games with concave and continuously differentiable payoff functions. This result is formalized in Lemma II.1. The proofs of this result and all subsequent others are in Appendix A.2.

**Lemma II.1** *For all information structures $(S_2,\psi_2)$, the set of BNE in the pricing game $\hat{\mathcal{E}}(S_2,\psi_2)$ is a singleton.*

**Information disclosure.** The choice of information structure $(S_2,\psi_2)$ determines the equilibrium in the pricing game. The designer chooses an information structure to maximize its ex-ante expected payoff such that $(\hat{\beta}_1^*(p_1|\theta), \hat{\beta}_2^*(p_2|s_2))$ is the BNE of the pricing game. I consider four objective functions for the designer:

1. Informed firm expected profits:

$$\mathbb{E}\left[\Pi_1((p_1,p_2);\theta)\right] = \sum_{\theta\in\Theta}\mu_\theta \int_{S_2}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \Pi_1((p_1,p_2);\theta)\mathrm{d}\hat{\beta}_2^*(p_2|s_2)\mathrm{d}\hat{\beta}_1^*(p_1|\theta)\mathrm{d}\psi_2(s_2|\theta)$$

2. Expected Consumer surplus:

$$\mathbb{E}\left[CS((p_1,p_2);\theta)\right] = \frac{1}{2a}\sum_{i\in\{1,2\},\theta\in\Theta}\mu_\theta \int_{S_2}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} q_i((p_i,p_{-i});\theta)^2\mathrm{d}\hat{\beta}_1^*(p_1|\theta)\mathrm{d}\hat{\beta}_2^*(p_2|s_2)\mathrm{d}\psi_2(s_2|\theta)$$

3. Expected Producer surplus:

$$\sum_{i=1,2} \mathbb{E}\left[\Pi_i((p_1,p_2);\theta)\right] = \sum_{i=1,2,\theta\in\Theta} \mu_\theta \int_{S_2}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \Pi_i((p_i,p_{-i});\theta)\mathrm{d}\hat{\beta}_1^*(p_1|\theta)\mathrm{d}\hat{\beta}_2^*(p_2|s_2)\mathrm{d}\psi_2(s_2|\theta)$$

4. Expected Welfare:

$$\mathbb{E}\left[W((p_1,p_2);\theta)\right] := \mathbb{E}\left[CS((p_1,p_2);\theta)\right] + \sum_{i=1,2} \mathbb{E}\left[\Pi_i((p_1,p_2);\theta)\right].$$

The interpretation of the role of the designer varies depending on their objective function. If the designer's objective is to maximize the informed firm's expected profits, then it is as if the informed firm is choosing how much information to disclose to its competitor. If the designer's objective is to maximize expected producer surplus, it is as if there is a collusive agreement between firms to determine optimal disclosure of information among them. Lastly, if the designer's objective is to maximize expected consumer surplus or welfare, the interpretation of the designer is as a regulator.

The main effects of information disclosure on welfare are captured by the trade offs arising from the informed firm and the consumer optimal disclosures. The insights obtained by analyzing them extend to the producer surplus and welfare optimal disclosures.

### 2.3.1 Preliminary results

#### 2.3.1.1 Equivalence to recommendation mechanisms

This section simplifies the information design problem by constraining the set of information structures. Taneva (2019) shows that it is without loss of generality to restrict attention to information structures where signals are equilibrium recommendations conditional on the state. I present an extension to compact action spaces and bounded, continuous real-valued payoff functions, restricting attention to $p_i \in \left[0, \frac{\theta_H}{a-b}\right]$ for all $i \in \{1,2\}$.[15] In a recommendation mechanism, the pricing rule $\sigma : \Theta \to \Delta\left(\left[0, \frac{\theta_H}{a-b}\right]^2\right)$ recommends a price for each firm such that the obedience

[15] This restriction is without loss of generality, since any price above $\frac{\theta_H}{a-b}$ induces no trade and profits of zero for firm $i$.

constraints are satisfied, ensuring that firms are willing to follow the price recommendation. Any pricing rule which satisfies the obedience constraints is a Bayes Correlated Equilibrium (BCE), as introduced by Bergemann and Morris (2013). A pricing rule $\sigma : \Theta \to \Delta\left(\left[0, \frac{\theta_H}{a-b}\right]^2\right)$ is a BCE if

$$\begin{aligned}
&\sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i, p_{-i}), \theta)\mathrm{d}\sigma((p_i, p_{-i})|\theta) \\
\geq &\sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}), \theta)\mathrm{d}\sigma((p_i, p_{-i})|\theta)
\end{aligned} \tag{2.5}$$

for all $p_i \in \text{supp } \sigma$, $p_i' \in \left[0, \frac{\theta_H}{a-b}\right]$, and $i \in \{1, 2\}$.

Consider an analogous information environment in which both firms observe a private signal about the state such that the informed firm's signal is perfectly informative. In what follows, I show that it is without loss of generality to interpret signals $(s_1, s_2)$ as equilibrium recommendations in which each signal recommends a price to each firm. Define the information structure as the joint distribution of signals. Let $S_i$ be the set of private signal realizations for firm $i$. An information structure consists of a set of signal realizations and a family of conditional distributions $\psi : \Theta \to \Delta(S)$, where $S = S_1 \times S_2 = \{s_L, s_H\} \times S_2$. Let $\psi_i : \Theta \to \Delta(S_i)$ be the marginal distribution of signal $s_i \in S_i$ given the information structure $(S, \pi)$. The marginal distribution $\psi_1$ is fully informative about the state, which implies that the probability of observing signal $s_k$ conditional on state $\theta_k$ is 1.

Given the information structure $(S, \psi)$, firms play a pricing game in which they condition their pricing choices on their signal realization by selecting a mapping $\beta_i : S_i \to \Delta(\left[0, \frac{\theta_H}{a-b}\right])$ to maximize their expected profits. A strategy profile $(\beta_1, \beta_2)$ is a BNE if, for all $p_i \in \left[0, \frac{\theta_H}{a-b}\right]$ with $\beta_i(p_i|s_i) > 0$ for all $i$, we have

$$\begin{aligned}
&\sum_{\theta\in\Theta} \mu_\theta \int_{S_{-i}} \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i, p_{-i}); \theta)\mathrm{d}\beta_{-i}(p_{-i}|s_{-i})\mathrm{d}\psi((s_i, s_{-i})|\theta) \\
\geq &\sum_{\theta\in\Theta} \mu_\theta \int_{S_{-i}} \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i', p_{-i}); \theta)\mathrm{d}\beta_{-i}(p_{-i}|s_{-i})\mathrm{d}\psi((s_i, s_{-i})|\theta)
\end{aligned} \tag{2.6}$$

for all $p_i' \in \left[0, \frac{\theta_H}{a-b}\right]$, $s \in S$ and $i \in \{1, 2\}$. Denote by $\mathcal{E}(S, \psi)$ the set of BNE in the pricing game.

First, Lemma II.2 is an equivalence result stating that every possible BCE distribution can be replicated as a BNE by appropriately choosing the information structure. Intuitively, any correlation between obedient pricing choices can be generated as a BCE. In a BNE, all the correlation between pricing choices is generated through the information structure $(S, \psi)$.

**Lemma II.2** *The set of BCE coincides with* $\cup_{(S,\psi)}\mathcal{E}(S,\psi)$.

Second, Lemma II.3 implies that it is without loss of generality to restrict attention to recommendation mechanisms. Formally, an information structure $(S, \psi)$ is a recommendation mechanism if $S = \left[0, \frac{\theta_H}{a-b}\right]^2$. In a recommendation mechanism, signals act as pricing recommendations which firms are willing to follow as long as their competitor does as well.

**Lemma II.3** *For every* $\sigma \in \cup_{(S,\psi)}\mathcal{E}(S,\psi)$, *there exists a recommendation mechanism* $\left(\left[0, \frac{\theta_H}{a-b}\right]^2, \sigma\right)$ *such that* $\sigma \in \mathcal{E}\left(\left[0, \frac{\theta_H}{a-b}\right]^2, \sigma\right)$.

#### 2.3.1.2 Existence of optimal recommendation mechanism

The existence of the optimal recommendation mechanism stated in Lemma II.4 is guaranteed by the Weierstrass extreme value theorem. First, the existence of correlated equilibria for games in which players receive private signals and simultaneously choose actions from compact sets is established in Stinchcombe (2011). Second, the set of BCE is compact in the weak* topology, since it is the set of all probability measures on a compact set.[16] Then, the designer's problem is to maximize a continuous function of $\sigma$ over a non-empty compact set.

**Lemma II.4** *The optimal recommendation mechanism exists.*

[16] With full disclosure, equilibrium prices are

$$p^F(\theta) = \frac{\theta}{2a-b}.$$

It follows that firms have no incentives to set prices above $p^F(\theta_H)$ or below $p^F(\theta_L)$, because such prices would never be part of a BNE of the pricing game. Hence, the support of any obedient recommendation mechanism must be a subset of $[p^F(\theta_L), p^F(\theta_H)]^2$. See Appendix A.2 for a formal argument.

## 2.4 Informed firm optimal disclosure

In this section, I consider the case in which the informed firm directly determines its optimal information disclosure. That is, assume that the designer's objective is to maximize the informed firm's expected profits,

$$\mathbb{E}_{(\mu,\sigma)}[\Pi_1((p_1,p_2);\theta)] = \sum_{\theta\in\Theta}\mu_\theta \int \Pi_1((p_1,p_2);\theta)\mathrm{d}\sigma((p_1,p_2)|\theta).$$

The informed firm chooses a feasible obedient recommendation mechanism $\sigma$ to maximize its expected equilibrium profits in the pricing game. Proposition II.1 states that it is optimal for the informed firm to share its information.

**Proposition II.1 (Informed firm optimal disclosure)** *It is optimal for the informed firm to fully reveal its private information to the uninformed firm.*

The optimal disclosure policy is determined by the fact that pricing choices are strategic complements, which determines the effect of changes in the precision of the uninformed firm's signal on the informed firm's expected profits. In particular, the informed firm's expected equilibrium profits, $\mathbb{E}_{(\mu,\sigma)}[\Pi_1^*((p_1,p_2);\theta)]$, can be expressed as

$$\mathbb{E}_{(\mu,\sigma)}[\Pi_1^*((p_1,p_2);\theta)] = a\mathbb{E}_\mu\left[\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)^2\right].$$

Then, maximizing the informed firm's expected equilibrium profits is equivalent to maximizing the distance between the expected equilibrium prices set by the uninformed firm across states, $\mathbb{E}_\sigma[p_2|\theta_L]$ and $\mathbb{E}_\sigma[p_2|\theta_H]$, because of the convexity of the informed firm's expected equilibrium profits with respect to the conditional expectation of the uninformed firm's price. With no disclosure, the uninformed firm sets one price in both states. Increasing the precision of the signal observed by the uninformed firm increases the correlation between its expected price and the state and, therefore, variation in its expected price. As a result, full disclosure maximizes the informed firm's expected profits.

Intuitively, increasing the precision of the signal observed by the uninformed firm increases its certainty about the state, increasing (decreasing) expected demand when its posterior beliefs suggest that the high (low) state is more likely. Accordingly, the uninformed firm increases its expected equilibrium price in the high state and decreases it in the low state. The higher competitor price translates to a higher demand for the informed firm, allowing it to increase its own price in the high state, since firms offer substitutes. The informed firm then sells a higher quantity at a higher price, increasing profits. The opposite is true in the low state since it charges a lower price and faces lower demand, but the expected profit gain in the high state exceeds the expected loss in the low state given the larger size of the market in the high state. Hence, the informed firm benefits from price correlation and its expected equilibrium profits increase in the precision of the uninformed firm's signal. Since this precision is maximized by full disclosure, it is optimal for the informed firm to fully disclose its private information.

The optimality of full disclosure doesn't rely on the linearity of demand. As formalized in Proposition A.1 in Appendix A.4, full disclosure is optimal if the informed firm's expected equilibrium profits are supermodular in the state and the uninformed firm's price. I also show that no disclosure is optimal if the informed firm's expected equilibrium profits are submodular in the state and the uninformed firm's price. This implies that it is optimal for the informed firm to reveal no information to its competitor when they compete by setting prices and offer differentiated complement goods. Kolotilin and Wolitzky (2020) obtain a related result in a setting in which the sender and the receiver do no interact. They show that supermodularity of the sender's objective function with respect to the state and the receiver's action is a sufficient condition for the optimality of full disclosure in decision problems. My results strengthen findings from previous work (Vives (1984), Vives (1990) and Raith (1996)), showing the optimality of either full or no information disclosure in a setting of information asymmetry where the distribution of the uninformed firm's signal and the correlation with the informed firm's signal are unrestricted.[17] One takeaway from my results is that it can be optimal for a firm to disclose information to a competitor even when it has no information to gain in return, because the firm can use disclosure to influence competitor prices.

[17] They also strengthen results from Novshek and Sonnenschein (1982), Clarke (1983) and Gal-Or (1985), given that Cournot with substitutes (complements) is equivalent to Bertrand with complements (substitutes) from the point of view of firms, as discussed in Raith (1996).

## 2.5 Consumer optimal disclosure

In this section, I interpret the designer as a regulator whose objective is to determine the consumer optimal disclosure. In particular, assume that the designer's objective is to choose an obedient price recommendation mechanism $\sigma$ that maximizes expected consumer surplus, given by

$$\mathbb{E}_{(\mu,\sigma)}[CS((p_1,p_2);\theta)] = \frac{1}{2a}\sum_{i\in\{1,2\}}\sum_{\theta\in\Theta}\mu_\theta\int q_i((p_i,p_{-i});\theta)^2 \mathrm{d}\sigma((p_i,p_{-i})|\theta).$$

The optimal disclosure is formalized in Proposition II.2. Partial disclosure is optimal for consumers when firms offer sufficiently close substitutes. Otherwise, no disclosure is optimal. Disclosure allows the uninformed firm to better tailor its price to the state, which in turns increases the ability of firms to extract surplus from consumers. But, private disclosure allows consumers to arbitrage prices by creating a potential coordination failure in firm pricing.

**Proposition II.2 (Consumer optimal disclosure)** *If the designer's objective is to maximize expected consumer surplus, there exists an $\hat{\alpha}\in(0,\hat{c}]$ such that partial disclosure is optimal if $\delta\in(\hat{\alpha},1)$ and no disclosure is optimal otherwise, where $\hat{c}$ is the cutoff for binary information structures.*

Intuitively, the impact of disclosure on consumer surplus is determined through two channels. On the one hand, disclosure provides the uninformed firm with information about the state, which increases the correlation between its pricing and the state. Indirectly, this also increases pricing correlation across firms, since the informed firm knows the state. Accordingly, in expectation, firms more accurately tailor their prices to the demand they face, allowing firms to better extract surplus from consumers. On the other hand, it creates uncertainty in firms' pricing decisions, because both firms now have private information. Even if disclosure increases expected price correlation between firms, uncertainty about the signal realization observed by their competitor generates a pricing coordination failure with positive probability. That is, the uninformed firm may observe a signal realization that mismatches with the state, setting a price tailored to the incorrect state. In contrast, the informed firm sets a price tailored to the realized state. When the mismatch occurs and firms set different prices, consumers benefit from arbitraging prices by selecting from which firm to purchase after observing prices. Private disclosure can thus benefit consumers.

The relative impact of these effects on consumer surplus is determined by the degree of differentiation between goods. When goods are close substitutes, a price differential between firms caused by partial private disclosure induces a large segment of the market to buy from the firm with a comparatively low price, creating large gains in consumer surplus with positive probability. In contrast, when goods are not close substitutes, the pricing coordination failure has little impact on the demand that firms face, yielding negligible benefits from price arbitrage. Accordingly, when goods are sufficiently close substitutes, private partial disclosure creates a large enough expected benefit from a potential price coordination failure to provide incentives for the regulator to impose partial disclosure. Otherwise, no disclosure is optimal.

The sketch of the proof is as follows. First, I verify that no disclosure yields a higher expected consumer surplus than full disclosure for all demand parameters and prior distributions of the state. Second, I show that the difference between the expected consumer surplus with partial disclosure $\sigma$ and no disclosure $\sigma^N$, denoted by $\Delta\mathbb{E}[CS](\sigma)$, is a continuous and strictly increasing function of $\delta$ for all $\sigma$. Third, I show that the optimal disclosure is determined by the degree of substitution $\delta$. As a first step, I restrict attention to binary information structures and show that there exists a cutoff $\hat{c}$ in the degree of differentiation above which partial disclosure is optimal.[18] As a second step, I show that no disclosure $\sigma^N$ yields a higher expected consumer surplus than any recommendation mechanism $\sigma$ when $\delta \to 0$. Hence, the intermediate value theorem implies that there exists a cutoff in the degree of differentiation, $\hat{\alpha} \in (0, \hat{c}]$, above which partial disclosure is optimal and in particular better than no disclosure.

More specifically, in the unique BNE of the pricing game, the informed firm's optimal pricing strategy is determined by the price recommendation made to the uninformed firm. Define $\sigma(p_2|\theta)$ as the price recommendation to the uninformed firm given the equilibrium price recommendations $\sigma((p_1, p_2); \theta)$. The informed firm's recommended

[18]Proposition II.3 shows that the consumer optimal disclosure in fact has binary support.

prices satisfy

$$p_1^\sigma(\theta) = \frac{\theta + b \int p_2 \mathrm{d}\sigma(p_2|\theta)}{2a}.$$

This implies that, for any obedient recommendation mechanism $\sigma((p_1, p_2); \theta)$, it is sufficient to pin down $\sigma(p_2|\theta)$ since it determines both firms' equilibrium pricing decisions. Then, relying on properties of expectations and variances, the difference between expected consumer surplus with partial disclosure $\mathbb{E}_{(\mu,\sigma)}[CS((p_1, p_2); \theta)]$ and no disclosure $\mathbb{E}_{(\mu,\sigma^N)}[CS((p_1, p_2); \theta)]$ is a linear combination of three moments of $\sigma(p_2|\theta)$ that can be expressed as

$$\Delta\mathbb{E}[CS](\sigma) = C_1(\delta, a)\mathbb{V}_{(\mu,\sigma)}[p_2] - C_2(\delta)Cov_{(\mu,\sigma)}(\theta, p_2) - C_3(\delta, b)\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]],$$

where $C_k(\cdot)$ is strictly positive for all $k \in \{1, 2, 3\}$, $\mathbb{V}_{(\mu,\sigma)}[p_2]$ represents the variance of the uninformed firm's price, $Cov_{(\mu,\sigma)}(\theta, p_2)$ represents the covariance between the uninformed firm's price and the state and $\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$ denotes the variance of the conditional expectation of the uninformed firm's price conditional on the state.

These three moments determine the impact of information disclosure on consumer surplus through relative changes in the level of demand and pricing in each market.

1. The expected gain in consumer surplus increases in the variance of the uninformed firm's price, $\mathbb{V}_{(\mu,\sigma)}[p_2]$, because it increases the opportunity for consumers to arbitrage price differences between firms and substitute between them.

2. The expected gain decreases in the covariance between the uninformed firm's price and the state, $Cov_{(\mu,\sigma)}[p_2, \theta]$, since it captures surplus extraction from consumers through the uninformed firm's better pricing decision.

3. The expected gain decreases in the variance of the expectation of $p_2$ conditional on the state, $\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$, because it captures the effect of disclosure on the informed firm's pricing. In particular, an increase of $\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$, reduces the informed firm's uncertainty about the uninformed firm's pricing and increases price correlation between firms.

Therefore, $Cov_{(\mu,\sigma)}[p_2, \theta]$ and $\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$ measure the loss in expected consumer surplus from disclosure due to the increased ability of firms to extract surplus from consumers. Increasing information disclosure increases all three moments, but their relative magnitude is determined by the degree of differentiation which pins down firms' optimal pricing and how willing consumers are to substitute between goods. In particular, the benefit for consumers increases as firms offer closer substitutes, whereas the ability of firms to extract surplus from consumers decreases since this decreases their market power. As a result, partial disclosure is optimal for consumers when firms offer sufficiently close substitutes.

Next, when partial disclosure is optimal, I characterize the optimal partially informative recommendation mechanism in Proposition II.3. The optimal price recommendation mechanism recommends at most two prices: a low price only recommended in the low state and a high price recommended in both states. The recommended prices maximize the uninformed firm's expected profits given its beliefs about the state. Then, the optimal price recommendation mechanism is characterized by the probability of recommending the low price in the low state, denoted by $\lambda^*$, where $\lambda^*$ determines the recommended prices $\hat{p}_L$ and $\hat{p}_H$ and is chosen to maximize expected consumer surplus subject to firm optimal pricing.[19]

**Proposition II.3 (Consumer optimal recommendation mechanism)** *Any consumer optimal recommendation mechanism recommends at most two prices. If an optimal mechanism discloses information, then there exists an optimal mechanisms that recommends one price $\hat{p}_L$ only when the state is low and another price $\hat{p}_H$ in both states where*

$$\hat{p}_L = \frac{4a^2[1-\mu_L\lambda^*]\theta_L + b^2\mu_H[(1-\lambda^*)\theta_H - \theta_L]}{(2a-b)\left[4a^2(1-\mu_L\lambda^*) - b^2\mu_H\lambda^*\right]},$$
$$\hat{p}_H = \frac{4a^2\left[\mu_H\theta_H + \mu_L(1-\lambda^*)\theta_L\right] - b^2\mu_H\lambda^*\theta_H}{(2a-b)\left[4a^2(1-\mu_L\lambda^*) - b^2\mu_H\lambda^*\right]},$$

*and $\lambda^* := \sigma(\hat{p}_L|\theta_L) \in (0,1)$ is*

$$\lambda^* = \frac{4\left[\delta(1-3\delta^2) + 6(1-\delta^2)\right]}{\mu_H\delta^5 + 2\mu_H\delta^4 - (12-\mu_H)\delta^3 - 6(4-\mu_H)\delta^2 + 4(1-\mu_H)\delta + 24(1-\mu_H)}.$$

[19] In this context, the price coordination failure occurs when the low state is realized and the uninformed firm is recommended to price high.

Intuitively, consumers gain from disclosure when there are differences in firms' pricing. In particular, when the state is high and the informed firm sets a corresponding high price, it would be good for consumers for the uninformed firm to observe a low price recommendation, because this allows a segment of consumers to substitute goods and purchase from the uninformed firm at a low price. Similarly, it would be best for consumers for the uninformed firm to observe a high price recommendation in the low state, since the informed firm is already setting a low price.

Moreover, in the high state, recommending an intermediate price to the uninformed firm rather than a low price would provide less benefit to consumers, implying that it is optimal for consumers to recommend at most two prices in the that state. Given that no intermediate price would be recommended in the high state, an intermediate price recommendation would reveal to the uninformed firm that the state is low, but the uninformed firm would only be willing to set the low price in that case. Hence, an optimal price recommendation mechanism recommends at most two prices.

Lastly, with linear demand, recommending a unique price in the low state or the high state is equivalent, since both options yield the same expected consumer surplus given that consumers benefit from price differentials induced by uncertainty among firms. This implies that there exists an optimal price recommendation mechanism, characterized in Proposition II.3, which recommends at most two prices with a unique price recommended in the high state. This price recommendation mechanism minimizes the level of prices set by firms.

The sketch of the proof of Proposition II.3 is as follows. First, fixing an arbitraty $\theta$, I show that, for any partially informative obedient recommendation mechanism $\sigma$, at most two prices are recommended in state $\theta$ if only one price is recommended when the state is $\theta' \neq \theta$. If a unique price $\hat{p}$ is recommended when the state is $\theta'$, observing any other recommendation $p_2 \neq \hat{p}$ reveals to the uninformed firm that the state is $\theta \neq \theta'$. When the uninformed firm knows that the level of demand is $\theta$, the obedience constraint implies that there is a unique price that it is willing to set. As a result, it is not possible to recommend more than two obedient prices across states. Second, I show that it is optimal for the regulator to recommend a unique price when the state is $\theta'$. These results imply that the optimal information structure sends at most two price recommendations. Further, it is fully characterized by the probability of recommending one of the optimal prices when the state is $\theta$. Since the benefit from partial disclosure is a consequence of the induced uncertainty between firms, the choice of the state $\theta$ in which two prices are recommended is inconsequential for consumers because of the linearity of demand. Without loss of generality, I focus on the case in which two prices are recommended in the low state.[20]

Proposition II.3 also implies that the cutoff for the optimality of partial disclosure coincides with the cutoff for binary structures ($\hat{\alpha} = \hat{c}$). Then, Figure 2.4 illustrates the optimal disclosure policy. Partial disclosure maximizes expected consumer surplus if the degree of differentiation $\delta$ belongs to the blue shaded region, while no disclosure is optimal for consumers if $\delta$ belongs to the red shaded region. Depending on the demand parameters and distribution of the state, the optimal partial disclosure can increase expected consumer surplus by up to 2% with respect to no disclosure and 10% with respect to full disclosure.[21]

[20] With linear demand and binary price recommendations, characterized by $x_\ell = \mathbb{P}(p_2 = p_\ell | \theta = \theta_\ell)$ with $\ell \in \{L, H\}$, the first order conditions of the regulator's maximization problem are collinear. As a result, it is possible to set either $x_L$ or $x_H$ to 1 since the optimality conditions define a relationship between them. In contrast, preliminary results suggest that with a quadratic demand given by $q_i(p_i, p_{-i}; \theta) = \max\{0, \theta + bp_{-i} - ap_i - cp_i^2\}$ where $c$ is positive and sufficiently small, it is optimal to recommend two prices in the low state and one price in the high state. Similarly, when $c$ is negative and sufficiently small, it is optimal to recommend two prices in the high state and one price in the low state.

[21] Given demand parameters, $a$ and $b$ and the distribution of the state, determined by $\theta_L$, $\theta_H$ and $\mu_H$, define $\eta(a, b, \theta_L, \theta_H, \mu_H)$ as the maximum expected consumer surplus given by

$$\eta(a, b, \theta_L, \theta_H, \mu_H) := \max_{\lambda \in [0,1]} \mathbb{E}[CS((p_1, p_2); \theta)].$$

Then, the maximum increases in consumer surplus compared to no and full disclosure are obtained by maximizing the following functions

$$\max_{(a,b,\theta_L,\theta_H,\mu_H)} \frac{\eta(a, b, \theta_L, \theta_H, \mu_H) - \mathbb{E}_{(\mu,\sigma^N)}[CS((p_1, p_2); \theta)]}{\mathbb{E}_{(\mu,\sigma^N)}[CS((p_1, p_2); \theta)]} \text{ or}$$
$$\max_{(a,b,\theta_L,\theta_H,\mu_H)} \frac{\eta(a, b, \theta_L, \theta_H, \mu_H) - \mathbb{E}_{(\mu,\sigma^F)}[CS((p_1, p_2); \theta)]}{\mathbb{E}_{(\mu,\sigma^F)}[CS((p_1, p_2); \theta)]},$$

with respect to feasible demand parameters and parameters governing the distribution of the state.

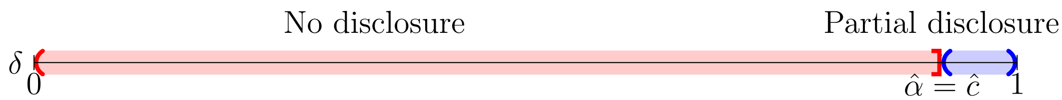



Figure 2.4: Consumer optimal disclosure

## 2.6 Producer surplus and Welfare optimal disclosure

Information disclosure impacts surplus allocation between firms and between firms and consumers, with potential implications for total welfare. In this section, I first characterize the disclosure policy that maximizes expected producer surplus and, combining this result with the consumer optimal disclosure, I derive the expected welfare maximizing disclosure policy.

### 2.6.1 Producer Surplus optimal disclosure

In this section, I interpret the designer as a collusive agreement between firms whose objective is to choose a disclosure policy to maximize expected producer surplus given by

$$\sum_{i=1,2} \mathbb{E}_{(\mu,\sigma)}[\Pi_i((p_i, p_{-i}); \theta)] = \sum_{i\in\{1,2\}} \sum_{\theta\in\Theta} \mu_\theta \int \Pi_i((p_i, p_{-i}); \theta) \mathrm{d}\sigma((p_i, p_{-i})|\theta).$$

First, it is optimal for the uninformed firm to learn the state, as stated in Lemma II.5, because it increases the correlation between its pricing decisions and the state.

**Lemma II.5** *The expected profits of the uninformed firm are maximized by full disclosure.*

Proposition II.1 and Lemma II.5 indicate that full disclosure is optimal for both firms, because it maximizes both the informed and uninformed firm's expected profits. Thus, full disclosure maximizes expected producer surplus.

### 2.6.2 Welfare optimal disclosure

Assume that the designer, interpreted as a regulator, wants to maximize expected welfare, defined as the sum of expected consumer and producer surplus. To maximize expected welfare, the regulator trades off the effect of information disclosure on firms and consumers, given their conflicting preferences over disclosure policies. In particular, firms' expected profits are maximized by full disclosure, whereas expected consumer surplus is maximized by no or partial disclosure. However, the benefits from disclosure for both firms and consumers increase as firms offer closer substitutes. As a result, the optimal disclosure is again determined by the degree of differentiation, as stated in Proposition II.4.

**Proposition II.4 (Welfare optimal disclosure)** *If the designer's objective is to maximize expected welfare, there exists* $\tilde{\alpha}_1 \in (0,1)$ *and* $\tilde{\alpha}_2 \in (0,1)$ *such that* $\tilde{\alpha}_1 \leq \tilde{\alpha}_2$ *and*

1. *no disclosure is optimal when* $\delta \in (0, \tilde{\alpha}_1]$.
2. *partial disclosure is optimal when* $\delta \in (\tilde{\alpha}_1, \tilde{\alpha}_2)$.
3. *full disclosure is optimal when* $\delta \in [\tilde{\alpha}_2, 1)$.

When firms offer sufficiently close substitutes, full disclosure is optimal since it is optimal for firms and their expected gains exceed expected losses for consumers. When firms offer sufficiently differentiated substitutes, no disclosure maximizes expected welfare since it is optimal for consumers and the expected profit gains for firms from disclosure are small. For intermediate levels of differentiation, partial disclosure is optimal.[22]

Ui and Yoshizawa (2015) reach a similar conclusion, restricting attention to symmetric normally distributed private and public signals. When firms offer substitutes, they show that welfare decreases in the precision of private information and increases in the precision of public information, related to the optimality of either partial or full disclosure.

[22] Suppose instead that the regulator maximizes the weighted sum of producer and consumer surplus, where $\omega \in [0,1]$ represents the weight assigned to consumers. The cutoffs $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ increase with $\omega$: for sufficiently high $\omega$, only no or partial disclosure can be optimal; for sufficiently low values of $\omega$, full disclosure is always optimal.

Proposition II.5 characterizes the welfare optimal partially informative disclosure policy. Incentives for partial disclosure are driven by the effect of disclosure on consumer surplus. Hence, the qualitative features of the disclosure policy are shared with the consumer optimal one stated in Proposition II.3. That is, any optimal partially informative recommendation mechanism has binary support, recommends one price only when the state is low, and another price in both states.

**Proposition II.5 (Welfare optimal recommendation mechanism)** *Any welfare optimal recommendation mechanism recommends at most two prices. If the optimal mechanism discloses information, then it recommends one price $\hat{p}_L$ only when the state is low and another price $\hat{p}_H$ in both states where*

$$\hat{p}_L = \frac{4a^2[1-\mu_L\lambda]\theta_L + b^2\mu_H[(1-\lambda)\theta_H - \theta_L]}{(2a-b)\left[4a^2(1-\mu_L\lambda) - b^2\mu_H\lambda\right]},$$
$$\hat{p}_H = \frac{4a^2\left[\mu_H\theta_H + \mu_L(1-\lambda)\theta_L\right] - b^2\mu_H\lambda\theta_H}{(2a-b)\left[4a^2(1-\mu_L\lambda) - b^2\mu_H\lambda\right]},$$

*and $\lambda^* := \sigma(\hat{p}_L|\theta_L) \in (0,1)$ maximizes expected welfare.*

The welfare optimal disclosure is illustrated in Figure 2.4. Full disclosure is optimal if $\delta$ is in the yellow shaded region, partial disclosure in the blue shaded region, while no disclosure is optimal in the red shaded region. The optimal partial disclosure can increase welfare by up to 1% with respect no and full disclosure. These results suggest that a regulator whose objective is to maximize welfare faces a trade off between consumer and producer surplus, and must take into account the relationship between markets. It highlights that the task of a regulator can be more nuanced than simply banning or releasing information: the exact design of information matters.

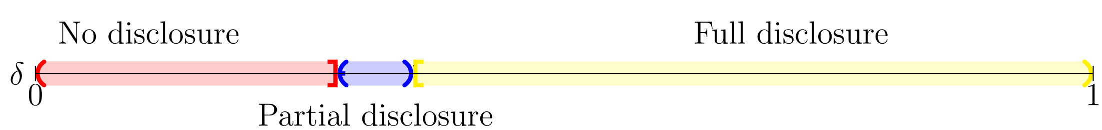



Figure 2.5: Welfare optimal disclosure

## 2.7 Extensions

**Firm optimal disclosure with complements.** Assume $b \in (-a, 0)$ and define the degree of complementary between goods as $\delta = |\frac{b}{a}|$. When goods are complements, disclosure increases the uninformed firm's profits, as stated in Lemma II.5, but reduces the informed firm's profits, as formalized in Proposition A.1 in Appendix A.4. The optimal disclosure policy is determined by comparing the gains of the uninformed firm to the losses of the informed firm, which are determined by the degree of complementarity between goods. In particular, if goods are sufficiently complementary ($\delta \geq \hat{\gamma}$), competitor prices have a significant impact on demand. Then, the negative effect of increased pricing correlation on the informed firm's profits exceeds the positive effect of learning about the state on the uninformed firm's profits. As a result, no disclosure is optimal. Otherwise, the informed firm's expected loss from information disclosure is smaller than the uninformed firm's expected gain, implying that full disclosure is optimal. These results are stated in Lemma II.6.

**Lemma II.6 (Producer surplus optimal disclosure)** *If the designer's objective is to maximize expected producer surplus, full disclosure is optimal if $\delta \in (0, \hat{\gamma})$. Otherwise, no disclosure is optimal.*

Eliaz and Forges (2015) study the producer surplus optimal disclosure policy in a Cournot duopoly with no private information and unknown demand. They show that it is optimal to fully inform one of the duopolists and disclose no information to the other when firms offer perfect substitutes. They also show that the producer surplus optimal disclosure consists of fully informing both firms when they offer perfect complements. Given the correspondence between Cournot and Bertrand discussed in Raith (1996), my results for complements and substitutes nest theirs, while allowing for more general patterns of complementarity and substitution between goods. Relatedly, Angeletos and Pavan (2007) show that producer surplus increases with the precision of public and private normally distributed signals when firms offer substitutes, related to the optimality of full disclosure. When firms offer complements, they show that producer surplus increases in the precision of private information, but can decrease in the precision of public information. In my context, this is exemplified by the designer who may have incentives to force information disclosure between firms when they offer complementary goods, decreasing the informational advantage of the informed firm at the benefit of its competitor.

**Public signals.** Assume that the designer commits to an information structure $(S_2, \pi_2)$ with public signal realizations. Given the information structure, firms play a pricing game in which they condition their choices on their information by selecting a mapping

$$p_1 : \Theta \times S_2 \to \Delta(\mathbb{R}_+) \text{ and } p_2 : S_2 \to \Delta(\mathbb{R}_+)$$

to maximize their expected profits.

**Lemma II.7** *For any disclosure policy $\sigma$, the informed firm's profits are higher with public disclosure than private disclosure.*

Public signals reinforce the informed firm's incentives to disclose information when firms offer substitutes, implying that full disclosure is optimal for the informed firm. When firms offer complements, it is optimal for the informed firm to disclose no information, by the same reasoning as with private signals. In contrast, consumers are better off with private disclosure, since they can benefit from information asymmetry as stated in Lemma II.8.

**Lemma II.8** *For any disclosure policy $\sigma$, expected consumer surplus is higher with private disclosure than public disclosure.*

When signals are public, the gain from partial disclosure disappears and, as a result, no disclosure is optimal for consumers as stated in Lemma II.9.

**Lemma II.9** *With public disclosure, no disclosure is optimal for consumers.*

**Informed firm as the owner of an online platform.** Consider the case in which trade occurs on an online platform run by the informed firm. The informed firm charges the uninformed firm a percentage of its sales for the use of the platform. Given the disclosure policy $\sigma$, the informed firm's expected equilibrium payoff is

$$\mathbb{E}_{(\mu,\sigma)}[\Pi_1^*((p_1, p_2); \theta)] = a\mathbb{E}_\mu \left[ \left( \frac{\theta + b\mathbb{E}_\sigma[p_2|\theta]}{2a} \right)^2 \right] + \alpha\Pi_2^*(\sigma),$$

where $\alpha \in [0, 1]$ is the percentage of sales charged to the uninformed firm.

The informed firm's incentives for information sharing are minimized by setting $\alpha = 0$, since the uninformed firm always benefits from observing information. When firms offer substitutes, the informed firm optimal disclosure doesn't change with $\alpha > 0$. In this case, the informed firm discloses all of its private information for any $\alpha \in [0, 1]$.

When firms offer complements, the informed firm optimal disclosure shares the same qualitative properties as the producer surplus maximizing disclosure with $\alpha = 0$. Full disclosure is optimal if the degree of complementarity is below a certain cutoff, no disclosure is optimal above the cutoff, and the cutoff is an increasing function of $\alpha$. Furthermore, the producer surplus optimal disclosure remains unchanged, since $\alpha$ represents a transfer between firms. Lastly, the consumer and welfare optimal disclosure also remain unchanged, since they are not affected by transfers between firms.

**N symmetric firms with constrained disclosure policies.** Consider a setting with $N \geq 3$ firms who compete by choosing prices. The level of demand depends on the state $\theta \in \{\theta_L, \theta_H\}$ with $\theta_H > \theta_L > 0$ such that firms are active in the market in both states. Firms share a common prior about the state, where the probability of $\theta$ is denoted by $\mu_\theta \in (0, 1)$. Firm $i$'s demand is given by

$$q_i(\mathbf{p}) = \theta - ap_i + \frac{b}{N-1} \sum_{j \neq i} p_j$$

where $a$ and $b$ are known parameters with $a > b > 0$. Firms' costs are zero.

The designer commits to an information structure with private signals, denoted by $\hat{\psi}_k$, to share all of the informed firm's private information with $k$ firms and no information with $N - 1 - k$ firms, where $k \in \{0, 1, 2, ..., N-1\}$. Firms who observe a perfectly informative signal condition their pricing choices on the state and select a mapping $p^F : \Theta \to \mathbb{R}_+$ to maximize their expected profits, whereas firms who observe no information select a price $p^N \in \mathbb{R}_+$ to maximize their expected profits. The optimal information disclosure is stated in Lemma II.10

**Lemma II.10** *If the designer's objective is to maximize the informed firm's expected equilibrium profits or to maximize expected producer surplus, it is optimal to share the informed firm's private information with all other firms. In contrast, if the designer's objective is to maximize expected consumer surplus, it is optimal to share the informed firm's private information with* $k^*(N, \delta)$ *firms where* $\frac{k^*(N,\delta)}{N} \leq \frac{2}{3}$.

First, the informed firm's expected equilibrium profits are maximized by sharing its private information with all other firms because it benefits from price correlation.

Similarly, when the designer's objective is to maximize expected producer surplus, it is optimal to share information with all firms, eliminating information asymmetry between firms, allowing them to better extract surplus from consumers.

Second, if the designer's objective is to maximize expected consumer surplus, information disclosure between firms is at least partially restricted. The optimal information structure, characterized by $k^*(N, \delta)$, is determined by the degree of substitution and the number of firms in the market. In particular, it is optimal to not disclose information to any other firm when $\delta \leq \frac{3}{4}$. When $\delta > \frac{3}{4}$, optimal disclosure is determined by the number of firms in the market and $\delta$, as illustrated in Figure 2.6. In particular, the optimal $k^*(N, \delta)$ increases in both $\delta$ and $N$, and $\frac{k^*(N,\delta)}{N} \leq \frac{2}{3}$. This means that it is optimal to share information with more firms as the number of firms increase in the market and as firms offer closer substitutes, but that it is optimal to leave at least a third of firms uninformed.

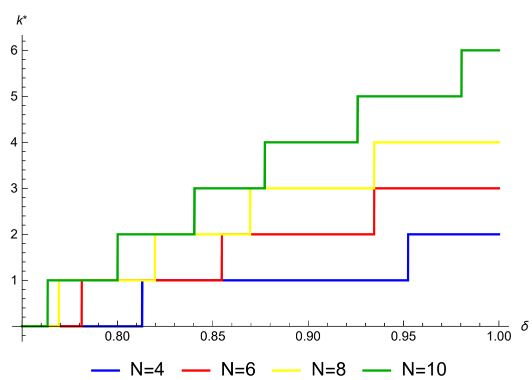



Figure 2.6: Consumer-optimal $k$ as a function of $\delta$ for different market sizes.

## 2.8 Conclusion

This paper studies information disclosure in a setting where two competing firms face ex-ante information asymmetry about the level of demand. I examine the incentives of an informed firm to share its private information with a competitor in a market with product differentiation and price competition. I show that the informed firm can have incentives to fully disclose its private information even without receiving information in return, because it allows them to influence competitor pricing. When firms offer substitutes, they benefit from price correlation, which implies that it is optimal for the informed firm to fully reveal its private information to the uninformed firm. When firms offer complements, it is optimal for the informed firm to not share any private information, which reduces the expected profits of its competitor. Accordingly, it can be optimal for a designer with the objective of maximizing producer surplus or maintaining competition to intervene and force information disclosure.

Further, information disclosure also impacts consumers. Even though complete information disclosure can help firms, it hurts consumers. I find that a regulator with the objective of protecting consumers would either completely restrict information disclosure between firms or only allow private partial disclosure, determined by the degree of differentiation between products. If goods are sufficiently close substitutes, partial disclosure is optimal, because it increases arbitrage opportunities for consumers. The consumer optimal partial disclosure reveals low levels of demand and obfuscates high levels to the uninformed firm.

Moreover, the preferences for information disclosure between firms and consumers are not aligned. When firms offer substitutes, the optimal disclosure depends on the degree of substitution, which determines the effect of disclosure on consumers and firms. If firms offer sufficiently differentiated goods, no disclosure maximizes expected welfare. If firms offer sufficiently close substitutes, full disclosure is optimal. For intermediate levels of differentiation, partial disclosure is optimal. Since incentives for partial disclosure derive from consumers, the optimal partial disclosure also reveals low levels and obfuscates high levels to the uninformed firm. My results highlight the wide scope for potential intervention by regulators, depending on their objective function and product differentiation.

An important aspect not considered in this paper is the effect of information disclosure on firm entry and exit decisions. In particular, the informed firm may reduce its information disclosure, reducing its current profits, to increase its market share and profits in the future by inducing uninformed firms to exit the market. In this context, a regulator may have incentives to force information disclosure between firms in order to maintain the level of competition in the market, which indirectly may also benefit consumers.

# CHAPTER III

# Strategic Incentives and the Optimal Sale of Information

## 3.1 Introduction

Agents often make decisions under uncertainty. The presence of firms who collect, aggregate and sell information allows agents to potentially supplement their private information and improve their decision making. Often, agents who buy information, data-buyers, interact with others in markets. For instance, retailers acquire information about demand conditions to improve their pricing decisions and compete more effectively. Similarly, investors purchase information about the profitability of their investment choices and may have incentives to coordinate with others. Since individual information decisions can affect equilibrium outcomes, demand for information and the optimal information offering of a data-seller depend on the strategic incentives between data-buyers.

In this paper, I analyze the direct sale of supplemental information in a stylized game of incomplete information. A data-seller owns a database containing information about a binary payoff-relevant state. The data-seller offers information to two privately informed data-buyers who play a two-stage game of incomplete information. In the information stage, data-buyers can simultaneously acquire supplemental information to reduce their uncertainty about the state. Data-buyers' information acquisition decisions are unobservable to other data-buyers and signal realizations are private as well as conditionally independent. In the action stage, each data-buyer simultaneously selects an action from a binary set to maximize her expected payoffs. The existence of private information implies that the data-seller is uncertain about demand for information and that data-buyers make inferences about the private information observed by others. In this context, I answer the following questions. What is the data-seller's optimal menu of information offerings with multiple privately informed data-buyers? How does it depend on the strategic incentives of data-buyers and the correlation between their private information?

Data-buyers' willingness to purchase supplemental information from the data-seller is determined by the precision and correlation of their private information, as well as strategic incentives in the action stage. The precision of their private information determines their overall demand for supplemental information. In particular, less informed data-buyers attach a higher value to supplemental information. The correlation between their private information impacts their beliefs regarding the information observed by others, which can affect willingness to pay for information based on strategic incentives in the action stage. Data-buyers have *coordination incentives* (*anti-coordination incentives*) if the expected gain of choosing an action increases (decreases) in the probability that the other data-buyer chooses the same action. Accordingly, with coordination incentives (anti-coordination incentives), the willingness to pay of a data-buyer for information increases (decreases) in the precision of the information observed by others.

Data-buyers' private information induces two possible interim beliefs, interpreted as their type. When the state is also binary, types are one-dimensional and characterized by the probability that they assign to a given state. The "high type" is defined as the one that attaches a higher value to the fully informative experiment. The data-seller designs a menu of Blackwell experiments and prices to screen data-buyer types, distorting the information provided to the low type in order to charge higher prices to the high type. The interaction between strategic incentives and correlated private information of data-buyers expands the opportunity of the data-seller to serve both segments of the market, increasing profits.

The optimal menu satisfies two standard properties of the screening literature: "no distortion at the top" and "no rent at the bottom". However, the data-seller can also extract all the surplus from the high type.The full surplus extraction result arises from the nature of data-buyers' preferences for information. Since information is valuable to a data-buyer if and only if it affects their choice, data-buyer preferences for information depend on its precision (quality) and what the information is about (position). In fact, different types value experiments differently and may disagree on the ranking of partially informative experiments. As a result, willingness to pay for information cannot be ordered across types, which is a common feature of multidimensional screening models (Rochet and Stole, 2003). The data-seller captures all the surplus by distorting the information provided to the low type and selecting the position of information such that the high type is indifferent between both experiments.

The main features of the optimal menu are as follows. The optimal menu contains the perfectly informative experiment offered to the high type and a concentrated experiment designed for the low type. The complete characterization of the information offered to the low type is determined by the distribution of data-buyer types and strategic incentives in the action stage. Consistent with previous work (Bergemann, Bonatti, and Smolin, 2018), if all data-buyer types would take different actions without supplemental information (non-congruent beliefs), their preferences over partially informative experiments are not aligned. This disagreement implies that there are partially informative experiments which are valuable to the low type but not to the high type, because they improve the low type's decision making without impacting the choices of the high type. Accordingly, the data-seller offers partial information to the low type without attracting the high type. In contrast with previous work, if data-buyer types would choose the same action without supplemental information (congruent beliefs), I show that the data-seller still offers partial information to the low type when: (i) data-buyers have coordination incentives in the action stage and their private information is negatively correlated, or (ii) data-buyers have anti-coordination incentives and their private information is positively correlated. Further, for both congruent and non-congruent beliefs, I show that the quantitative properties of the optimal menu are determined by the strategic incentives and the correlation of private information.

In the special case in which private information is conditionally independent, the qualitative features of the optimal menu are independent of the strategic incentives. If instead private information is conditionally dependent, the data-seller offers partial information to the low type when beliefs are congruent whenever the interaction between strategic incentives and the correlation of private information increases the demand for information. Intuitively, when private information is negatively correlated, data-buyers assign a higher probability to observing different private information. Hence, demand for supplemental information increases when data-buyers have coordination incentives, since it increases the correlation between their action choices. Similarly, when their private information is positively correlated and they have anti-coordination incentives, acquiring conditionally independent information is valuable, because it allows them to reduce the correlation between their actions.

These results highlight that the interaction between strategic incentives and the correlation of private information determines the features of the optimal menu. This interaction can relax the incentive compatibility constraints, allowing the data-seller to increase profits by not excluding the low type from the market. This emphasizes the importance of considering strategic interactions between data-buyers when designing information offerings, given that data-buyers generally interact with others in a market. The results also extend to a setting with $N$ data-buyers in which payoffs depend on whether they match the state and the choice of the majority.

This paper contributes the the literature on information design in games with privately informed players. In contrast to papers in which players have common priors (Taneva (2019) and Mathevet, Perego, and Taneva (2020)), I consider the role of private information and its correlation in determining the seller's optimal information offering. Data-buyers can have heterogeneous previous experiences which provide private information about the state, affecting their demand for information and incentives for the data-seller to provide information. The paper also contributes to the literature on selling information to imperfectly informed decision makers. Within the mechanism design approach to selling information, it is most closely related to Bergemann, Bonatti, and Smolin (2018), which studies the design and ex-ante pricing of Blackwell experiments for a single privately informed receiver. I extend their analysis to a setting with multiple data-buyers and explicitly consider how the optimal information offering depends on strategic interactions among data-buyers and the correlation of private information. Eső and Szentes (2007) and Li and Shi (2017) also consider multi-player settings, but in which the data-seller engages in ex-post pricing and data-buyers' actions are contractible. In contrast, I restrict prices to be contingent only on the information itself. In recent work, Krähmer (2020) studies a mechanism design problem with quasi-linear utility where the principal designs a selling mechanism and can also design and disclose supplemental information that affects agents' valuations. The main difference is that the seller in their setting not only offers information, but also sells a good whereas I consider a model in which only transfers are contractible and the agent takes a non-contractible action after information is revealed. Previous work (Admati and Pfleiderer (1986), Admati and Pfleiderer (1990) and Kastl, Pagnozzi, and Piccolo (2018)) also studies the monopolistic sale of information to multiple receivers with strategic interactions but no private information. In contrast, I consider a setting in which data-buyers are privately informed and study the the interaction between correlated private information and strategic incentives in determining the optimal information offering. Lastly, the non-exclusion result is consistent with the multidimensional screening literature (Rochet and Stole, 2003), in which the data-seller offers distorted partial information to the low type to ensure that the high type is indifferent between the information offerings.[1]

The remainder of the paper is organized as follows: Section 1 outlines the model, Section 2 derives preliminary results, Section 3 characterizes the optimal menu and Section 4 extends results to $N$ data-buyers, and Section 5 concludes.

## 3.2 Model

Consider a setting with two data-buyers and one data-seller. Data-buyers play a game of incomplete information, where $i$ indexes a generic data-buyer and $j$ denotes the other. The payoff-relevant state $\omega$ is drawn from a binary set $\Omega = \{\omega_1, \omega_2\}$. Each data-buyers is privately informed about the state and attaches probability $\theta \in \{\theta_L, \theta_H\}$ to state $\omega_1$, where the correlation between the data-buyers' private information is characterized by $\rho$ and $\nu$. Formally, the joint distribution of data-buyers' private information is defined in Table 3.1, where $\rho \in (0, 1)$ and $\nu \in \left(0, \frac{1-\rho}{2}\right)$.[2]

[1] More specifically, I consider a screening problem in which the data-seller's chooses the quality and position of information. Preferences over these dimensions are determined by their single-dimensional type.

[2] This can be interpreted as data-buyers sharing a common prior and privately observing either good or bad news about the likelihood of state $\omega_1$. Let $\mu_0 = \mathbb{P}(\omega = \omega_1)$ be their common prior and assume they observe a conditionally independent signal $s_0 \in \{s_0^1, s_0^2\}$ where $\mathbb{P}(s = s_0^1 | \omega = \omega_k) = \mu_k$ with $k \in \{1, 2\}$. Then,

$$\rho = \mu_0(1-\mu_1)^2 + (1-\mu_0)(1-\mu_2)^2 \text{ and } \nu = \mu_0(1-\mu_1)\mu_1 + (1-\mu_0)(1-\mu_2)\mu_2.$$

| | | Data-buyer $j$'s type | |
|---|---|---|---|
| | | $\theta_L$ | $\theta_H$ |
| Data-buyer $i$'s type | $\theta_L$ | $1-2\nu-\rho$ | $\nu$ |
| | $\theta_H$ | $\nu$ | $\rho$ |

Table 3.1: Joint distribution of private information.

The game has two stages: the information stage and the action stage. In the information stage, the data-seller offers a menu of Blackwell experiments and prices to the privately informed data-buyers. Data-buyers simultaneously decide whether or not to purchase information from the menu. If a data-buyer purchases information, she observes a private signal realization and updates her beliefs accordingly. Data-buyers don't observe each others' choices from the menu. In the action stage, each data-buyer simultaneously selects her action from the binary set $A = \{a_1, a_2\}$ to maximize her expected payoffs conditional on her signal realization. The payoffs $u : A \times \Omega \to \mathbb{R}$, defined in Table 3.2, are symmetric and characterized by $c > 0$.[3]

| $\omega = \omega_1$ | $a_1$ | $a_2$ |
|---|---|---|
| $a_1$ | $1,1$ | $c,0$ |
| $a_2$ | $0,c$ | $0,0$ |

| $\omega = \omega_2$ | $a_1$ | $a_2$ |
|---|---|---|
| $a_1$ | $0,0$ | $0,c$ |
| $a_2$ | $c,0$ | $1,1$ |

Table 3.2: Action stage payoffs.

Under these assumptions, it is an ex-post dominant strategy for each data-buyer to match the state $\omega$. Data-buyers are said to have coordination (anti-coordination) incentives if the expected gain of choosing an action increases (decreases) in the probability that the other data-buyer chooses the same action. That is, data-buyers have coordination incentives if $c < 1$ and have anti-coordination incentives when $c > 1$.[4]

[3] Any $2 \times 2$ symmetric game in which players prefer to match the state can be normalized in this manner.

[4] Let $I_i$ be data-buyer $i$'s information set. Define $\sigma_k$ as the probability that $i$ assigns to $j$ selecting $a_1$ conditional on state $\omega_k$ and $I_i$. Data-buyer $i$'s expected gain of choosing action $a_1$ instead of $a_2$

**Experiments.** An individual experiment $E^m = (S^m, \{\pi^m(\cdot|\omega)\}_{\omega\in\Omega})$ consists of a finite set of signal realizations $s_\ell^m \in S^m$ and a family of conditional distributions $\pi^m$ where

$$\pi_{\ell,k}^m := \mathbb{P}(s_\ell^m|\omega_k),\ \pi_{\ell,k}^m \geq 0 \text{ and } \sum_{\ell=1}^{L^m} \pi_{\ell,k}^m = 1$$

with $L^m = |S^m|$. Denote by $\mathcal{E}$ the set of feasible experiments. The seller's cost of providing information is zero.

An experiment $E^m$ can represented by a stochastic matrix in which each column represents a state and each row a signal realization, as in Table 3.3.

| | $\omega_1$ | $\omega_2$ |
|---|---|---|
| $s_1$ | $\pi_{1,1}^m$ | $\pi_{1,2}^m$ |
| $s_2$ | $\pi_{2,1}^m$ | $\pi_{2,2}^m$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $s_{L^m}$ | $\pi_{L^m,1}^m$ | $\pi_{L^m,2}^m$ |

Table 3.3: Matrix representation of experiment $E^m$.

Assume that the realizations of data-buyers' private information and the realization of the signal $s \in S^m$ from any experiment $E^m$ are independent conditional on the state $\omega$. Moreover, assume that signal realizations between data-buyers are conditionally independent. The first assumption implies that the value of an experiment is determined by a data-buyer's private information, its correlation with the private information observed by others, and the nature of the strategic incentives. It also implies that the value of an experiment can be derived independently of its price. The second assumption rules out that signals can be used as a coordination device, except through their correlation with the state. As such, data-buyers attach no value to the uninformative experiment.

conditional on her information set $I_i$, $\Delta U_i$, is given by

$$\Delta U_i := \mathbb{P}(\omega = \omega_1|I_i)[\sigma_1 + (1-\sigma_1)c] - (1 - \mathbb{P}(\omega = \omega_1|I_i))\,[\sigma_2 c + (1-\sigma_2)]$$

where $\frac{\partial \Delta U_i}{\partial \sigma_k} \geq 0$ if and only if $c \leq 1$ for all $k$. That is, $i$'s expected gain from selecting action $a_1$ instead of $a_2$ increases in the probability of $j$ choosing action $a_1$ if and only if $c \leq 1$. Analogously, $i$'s gain of choosing $a_2$ instead of $a_1$ increases in the probability that $j$ chooses $a_2$ if and only if $c \leq 1$.

**Data-seller's strategy space.** The data-seller offers a symmetric menu of individual Blackwell experiments and prices with arbitrarily informative signals. Let $\mathcal{M} = (E^m, t^m)_{m=1}^M$ denote the menu of experiments offered by the data-seller, where experiment $E^m$ is offered at price $t^m$ and $M$ is the number of experiments included in the menu with $m \in \{1, 2.., M\}$. Assume that only the experiment itself is contractible, not its realization, the realized state, or the data-buyers' actions. Formally, a strategy for the data-seller is a menu $\mathcal{M} = (E^m, t^m)_{m=1}^M$ where $t^m \in \mathbb{R}$ and $E^m \in \mathcal{E}$.

**Data-buyer's strategy space.** Each data-buyer $i$ of type $\theta$ decides whether to supplement its private information. Let $\iota_{i\theta} \in \{0, 1, ..., M\}$ denote data-buyer $i$'s information acquisition decision, where $\iota_{i\theta} = 0$ represents the case in which $i$ doesn't acquire supplemental information and $\iota_{i\theta} = m$ denotes the case in which $i$ acquires experiment $E^m$. Conditional on all their information, data-buyer $i$ chooses an action from the set $\{a_1, a_2\}$. Formally, a pure strategy for data-buyer $i$ of type $\theta$ consists of a pair $(\iota_{i\theta}, \alpha_{i\theta})$ where $\iota_{i\theta} \in \{0, 1, ..., M\}$ and $\alpha_{i\theta} = (\alpha_{i,\iota_{i\theta}} : S^{\iota_{i\theta}} \to \{a_1, a_2\})_{\iota_{i\theta}=0}^M$.

**Timing.** Before the state $\omega$ is realized, the data-seller offers a menu of experiments and prices. After the state $\omega$ is realized, each data-buyer observes her private information, simultaneously decides whether or not to purchase a Blackwell experiment, and if so which one to acquire. If a data-buyer acquires an experiment, she observes a private signal realization and updates her belief accordingly. Data-buyers don't observe each others' choices from the menu.[5] Lastly, data-buyers simultaneously choose actions from the binary set $\{a_1, a_2\}$ to maximize their expected payoff conditional on their information choices and signal realizations.

**Solution concept.** The solution concept is the data-seller's preferred perfect extended Bayesian equilibrium.[6] An equilibrium is an extended assessment satisfying consistency of beliefs, Bayesian updating, and sequential rationality in each infor-

[5]Data-buyer $i$'s deviations in information choices are unobservable, so that there are no strategic effect on the choices of others. That is, action and information choices are strategically simultaneous.

[6]This definition is equivalent to weak Perfect Bayesian Equilibrium with the additional assumption that data-buyers do not update their beliefs about the state after observing a deviant menu. This holds since the strategic independence assumption only requires this additional constraint given that the data-seller chooses a menu before the state is realized. See Battigalli (1996) or Watson (2016) for details.

mation set. That is, conditional on the offered menu and information choices, each data-buyer $i$'s action choice maximizes her expected payoff. Given the optimal menu, data-buyer $i$'s information choice maximizes the difference between her expected payoff in the action state and the price of information. Lastly, the optimal menu for the data-seller is the one that maximizes their expected profits, anticipating data-buyers' equilibrium choices.

**Definition III.1** *A strategy profile $(\iota^*, \alpha^*)$, a menu $\mathcal{M}^*$ and a belief system $\mu$ form an equilibrium if:*

1. *$(\iota^*, \alpha^*)$ and $\mathcal{M}^*$ satisfy sequential rationality. That is:*

   (a) *Given $\mathcal{M}^*$,*

$$\mathbb{E}[U_{i\theta}(\iota^*, \alpha^*)] \geq \mathbb{E}[U_{i\theta}(\iota^*, (\alpha_i', \alpha_{-i}^*))] \tag{3.1}$$

   *for all $\alpha_i'$ conditional on $\iota_{i\theta}^*$, for all $i \in \{1, 2\}$ and*

$$\iota_{i\theta}^* \in \underset{\iota_{i\theta} \in \{0,\ldots,M\}}{\arg\max}\ \mathbb{E}[U_{i\theta}((\iota_{i\theta}, \iota_{-i}^*), \alpha^*)] - t^{\iota_{i\theta}} \tag{3.2}$$

   *where the expectations are taken over the state $\omega$, the private information of the other data-buyer, and her choices.*

   (b) *A menu $\mathcal{M}^*$ is optimal if*

$$\mathcal{M}^* \in \underset{\mathcal{M}}{\arg\max} \sum_{\theta \in \Theta} \sum_{m=1}^{M} \mathbb{P}(\theta_i = \theta)\mathbb{P}(\iota_{i\theta}^* = m) \cdot t^m.$$

2. *$\mu$ satisfies extended Bayesian updating.*

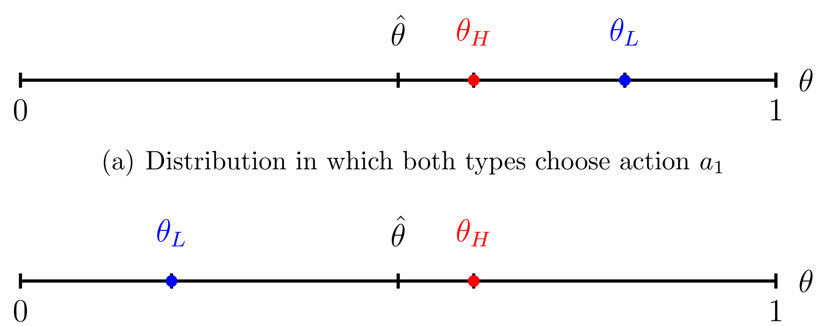
3. *$\mu$ satisfies strategic independence: data-buyers don't infer anything about the state if the data-seller offers a deviant menu.*

**The value of information.** The expected value of experiment $E^m$ is defined as the marginal value of information, which corresponds to the difference in expected equilibrium payoffs with and without observing experiment $E^m$. Denote by $V_{\mathcal{M}}(E^m; \theta)$ data-buyer $i$'s expected value of experiment $E^m$ when her interim belief is $\theta$ and the data-seller offers menu $\mathcal{M}$. Formally,

$$V_{\mathcal{M}}(E^m, \theta) = \max\{0, \mathbb{E}[U_{i\theta}((m, \iota^*_{-i}), \alpha^*)] - \mathbb{E}[U_{i\theta}((0, \iota^*_{-i}), \alpha^*)]\}.$$

The value of experiment $E^m$ for an individual data-buyer depends on her private information, her belief about the private information of others, the strategic incentives in the action stage, and the menu offered. It is determined by the probability of matching the state, the probability of matching the actions of other data-buyers, and the payoff structure. The chance of matching the state depends on $i$'s private information, but is independent of the information acquisition decision of other data-buyers. In contrast, the likelihood of matching data-buyer $j$'s action depends on $j$'s private information. Lastly, the strategic environment in the action stage determines the preferences of data-buyers over equilibrium outcomes.

Let $\overline{E}$ denote the perfectly informative experiment. The high type $\theta_H$ is defined as the type that assigns higher value to $\overline{E}$. That is, $V_{\mathcal{M}}(\overline{E}, \theta_H) \geq V_{\mathcal{M}}(\overline{E}, \theta_L)$. For example, if $c = 1$, data-buyers prefer action $a_k$ if they assign a higher probability to state $\omega_k$ for $k \in \{1, 2\}$ and they are indifferent between actions $a_1$ and $a_2$ if they assign equal probability to each state. In this case, the high type is the type closest to the cutoff $\hat{\theta} = \frac{1}{2}$, as illustrated in Figure 3.1.




(a) Distribution in which both types choose action $a_1$



(b) Distribution in which types choose different actions

Figure 3.1: Example of definition of high (red) and low (blue) types

## 3.3 Preliminary results

### 3.3.1 Simplifications

The data-seller's problem can be simplified along two dimensions. First, the revelation principle of mechanism design implies that it is without loss of generality to focus on direct mechanisms in which the data-seller assigns one experiment to each data-buyer type $\theta$. Second, the revelation principle of games of communication implies that it is without loss of generality to focus on experiments in which signals act as action recommendations. These two results imply that I can restrict attention to menus with at most two elements ($M \leq 2$) and experiments with two possible signal realizations ($S^m = \{s_1, s_2\}$ for all $m$.).

It is not trivial that the revelation principle for mechanism design holds in this setting. Unlike in standard mechanism design, where the information environment $\mathcal{S} = ((S_i)_{i=1}^2, \pi)$ is fixed and the designer chooses $\mathcal{G} = ((\{a_1, a_2\}, u_i)_{i=1}^2, \mu_0)$ to induce a desirable outcome, the data-seller takes the game $\mathcal{G}$ as given and controls $\mathcal{S}$. As such, each menu of experiments induces a different game of incomplete information. Lemma III.1 shows that the outcome of every menu can be attained by a direct menu, which includes at most two elements.

**Lemma III.1** *The data-seller offers a menu which includes at most two experiments.*

Given any direct menu $\mathcal{M}$, an experiment $E^m$ with private signals is *responsive* if every signal $s \in S$ leads to a different action choice for data-buyer $i$ of type $\theta$. A direct menu $\mathcal{M}$ is responsive if every experiment $E^m \in \mathcal{M}$ is responsive. Lemma III.2 shows that it is without loss of generality to focus on menus in which the cardinality of the signal space equals to the cardinality of the action space. I refer to these menus as responsive.

**Lemma III.2** *The outcome of every direct menu can be attained by a responsive menu.*

Lemma III.2 generalizes Proposition 1 from Bergemann, Bonatti, and Smolin (2018) to a setting with multiple data-buyers. This result relies on two main assumptions: signals are private and information acquisition decisions are unobservable. This ensures that a change in the set of signals observed by one data-buyer has no effect on the other data-buyer's action choices. Hence, it is without loss of generality to consider $S^m = \{s_1, s_2\}$ and $\pi^m : \Omega \to [0,1]^2$ for all $m \in \{1, ..., M\}$. Then, an experiment $E^m$ can be represented by the following stochastic matrix:

| $s/\omega$ | $\omega_1$ | $\omega_2$ |
|---|---|---|
| $s_1$ | $\pi_1^m$ | $1 - \pi_2^m$ |
| $s_2$ | $1 - \pi_1^m$ | $\pi_2^m$ |

where $(\pi_1^m, \pi_2^m) \in [0,1]^2$ for all $m \in \{1, ..., M\}$. Given that signals act as action recommendations, after observing signal $s_1$, data-buyer $i$ must be willing to choose action $a_1$ when $j$ follows her action recommendation if she chooses to acquire supplemental information. Similarly, after observing signal $s_2$, data-buyer $i$ must be willing to choose action $a_2$ when $j$ follows her action recommendation if she chooses to acquire supplemental information. That is, if $i$ acquires experiment $n$ and $j$ acquires experiment $m$, then

$$\theta \pi_1^n [\pi_1^m + (1 - \pi_1^m) c] \geq (1 - \theta)(1 - \pi_2^n)[(1 - \pi_2^m) c + \pi_2^m] \text{ and}$$
$$(1 - \theta) \pi_2^n [(1 - \pi_2^m) c + \pi_2^m] \geq \theta (1 - \pi_1^n)[\pi_1^m + (1 - \pi_1^m) c]$$

Also, assume that the likelihood of observing signal $s_1$ is higher conditional on state $\omega_1$ than on $\omega_2$ compared to $s_2$.[7] That is:

$$\frac{\mathbb{P}(s = s_1 | \omega_1)}{\mathbb{P}(s = s_1 | \omega_2)} \geq \frac{\mathbb{P}(s = s_2 | \omega_1)}{\mathbb{P}(s = s_2 | \omega_2)} \Leftrightarrow \frac{\pi_1^m}{1 - \pi_2^m} \geq \frac{1 - \pi_1^m}{\pi_2^m} \Leftrightarrow \pi_1^m + \pi_2^m \geq 1.$$

### 3.3.2 The data-seller's problem

Assume that the data-seller designs experiment $E^L$ for data-buyer type $\theta_L$ and $E^H$ for $\theta_H$. The presence of private information implies that the data-seller is uncertain about the demand for experiments and must screen data-buyer types. The data-seller's problem is then to design a menu of experiments to maximize expected transfers subject to data-buyers' incentive-compatibility and participation constraints.[8]

[7]This condition is equivalent to the monotone likelihood ratio property.

[8]Payments are conditional only on the information product itself and not on the types of other data-buyers. This assumption rules out the use of the Cremer-McLean condition (Crémer and McLean, 1988).

That is:

$$\max_{(E^m,t^m)_{m\in\{L,H\}}} (1-2\nu-\rho)2t^L + 2\nu(t^L+t^H) + \rho 2t^H$$

subject to the participation constraints

$$IR_L : V_{\mathcal{M}}(E^L,\theta_L) - t^L \geq 0,\ IR_H : V_{\mathcal{M}}(E^H,\theta_H) - t^H \geq 0$$

and the incentive-compatibility constraints

$$\begin{aligned} IC_L &: V_{\mathcal{M}}(E^L,\theta_L) - t^L \geq V_{\mathcal{M}}(E^H,\theta_L) - t^H \\ IC_H &: V_{\mathcal{M}}(E^H,\theta_H) - t^H \geq V_{\mathcal{M}}(E^L,\theta_H) - t^L. \end{aligned}$$

**Optimality of a menu with two distinct experiments.** The data-seller can either offer a menu with two distinct items or offer only the perfectly informative experiment at a fixed price, $p$, equal to the low type's willingness to pay. It is optimal for the seller to offer a menu with two distinct items if and only if:

$$(1-\nu-\rho)t^L + (\nu+\rho)t^H \geq p \Leftrightarrow \nu+\rho \geq \frac{p-t^L}{t^H-t^L}$$

where $\nu+\rho$ is the probability that a data-buyer is the high-type.

### 3.3.3 The value of experiments

In this section, I derive a closed form expression for the value of experiments. Assume that data-buyer $j$ type $\theta_j$ follows her equilibrium strategy, and denote by $m$ her experiment choice. Since signals act as action recommendations, data-buyer $j$ type $\theta_j$ conditions her action choice on the realized signal and selects action $a_k$ after observing signal $s_k$.[9] Define $v_k(E^n,\theta_i;m)$ as data-buyer $i$'s expected gain of acquiring experiment $E^n$ if, without information, she would choose action $a_k$ while $j$ plays her equilibrium strategy. Data-buyer $i$'s expected gain of acquiring information when

[9] Hence, consistency of beliefs implies that

$$\mathbb{P}(\iota_j = m|\theta_j) = 1 \text{ and } \mathbb{P}(a_j = a_k|\theta_j, \iota_j = m, s^j = s_k) = 1 \text{ for all } k \in \{1,2\}.$$

choosing action $a_2$ and $a_1$ without supplemental information followed by action $a_k$ after observing signal $s_k$, respectively, are given by:

$$v_2(E^n, \theta_i; m) = \underbrace{\theta_i \pi_1^n \left[\pi_1^m + (1-\pi_1^m)\, c\right]}_{\mathbb{P}(\omega=\omega_1)\mathbb{P}(s^i=s_1|\omega=\omega_1)[\mathbb{P}(s^j=s_1|\omega=\omega_1)+\mathbb{P}(s^j=s_2|\omega=\omega_1)c]} - \overbrace{(1-\theta_i)(1-\pi_2^n)\left[(1-\pi_2^m)\, c + \pi_2^m\right]}^{\mathbb{P}(\omega=\omega_2)\mathbb{P}(s^i=s_1|\omega=\omega_2)[\mathbb{P}(s^j=s_1|\omega=\omega_2)c+\mathbb{P}(s^j=s_2|\omega=\omega_2)]} \text{ and}$$

$$v_1(E^n, \theta_i; m) = (1-\theta_i)\pi_2^n \left[(1-\pi_2^m)\, c + \pi_2^m\right] - \theta_i (1-\pi_1^n)\left[\pi_1^m + (1-\pi_1^m)\, c\right].$$

That is, $v_k(E^n, \theta_i; m)$ is the difference between the data-buyer's expected gain in state $\omega_k'$ and her expected loss in state $\omega_k$ after observing signal $s_k'$. Then, data-buyer $i$'s value of experiment $E^n$ when the data-seller offers the menu $\mathcal{M}$ is given by:

$$V_{\mathcal{M}}(E^n, \theta_i) = \max\left\{0, \sum_{m\in\{L,H\}} \mathbb{P}(\theta_j = \theta_m|\theta_i) v_k(E^n, \theta_i; m) \text{ s.t. } k \text{ solves } \alpha_{i,\iota_{i\theta}=0} = a_k\right\}.$$

Assume that the value of experiment $E^n$ is weakly increasing in its precision, i.e. $c \in (\frac{1}{2}, 2)$.[10]

## 3.4 Optimal menu of experiments

### 3.4.1 General properties

The optimal menu shares some of the structural properties of the one data-buyer setting established in Bergemann, Bonatti, and Smolin (2018). Proposition III.1 generalizes these results to a two data-buyer setting and identifies which constraints are binding as well as the information provided to the high type in any optimal menu.

**Proposition III.1** *In an optimal menu:*

1. *Both participation constraints bind.*

2. *The incentive-compatibility constraint of the high type binds.*

[10] Lemma B.1 in Appendix B.1 shows that $c \in (\frac{1}{2}, 2)$ is a sufficient condition for $V_{\mathcal{M}}(E^n, \theta)$ to be increasing in the precision of experiment $E^n$.

*3.* $E^H$ *is perfectly informative. That is,* $\pi_1^H = \pi_2^H = 1$*.*

*4.* $E^L$ *is concentrated, i.e,* $\pi_k^L = \mathbb{P}_{E^L}(s = s_k | \omega = \omega_k) = 1$ *for some* $k \in \{1, 2\}$*.*

The optimal menu satisfies two standard properties of the screening literature: "no distortion at the top" and "no rent at the bottom". However, the data-seller can also extract all the surplus from the high type, because information is valuable only when it affects data-buyers' decision making. As such, information has two relevant dimensions: its precision and its position.[11] All data-buyer types prefer experiments with higher precision. However, different data-buyer types may disagree on their preferences for the position of information, given that they may select different actions if they don't supplement their private information. Hence, different types value experiments differently and disagree on the ranking of partially informative experiments, a common feature in multidimensional screening models. The data-seller captures all the surplus by selecting the position of information such that the high type is indifferent between the experiments offered.

Additionally, the perfectly informative experiment is part of any optimal menu. It is the most valued by any buyer type, since it allows them to perfectly match the state. Hence, if this experiment is not part of a menu, the data-seller can replace the currently most informative experiment by the perfectly informative one, weakly increasing profits by charging a higher price for this experiment while ensuring that the incentive-compatibility constraints are satisfied. Furthermore, it is optimal for the data-seller to offer this experiment to the high type, since their willingness to pay is higher.

The low type is offered a concentrated experiment in which the distribution of signals conditional on one state $\omega_k$ is degenerated, eliminating uncertainty for one state. Given that data-buyers have incentives to match the state, the data-seller can shift the probability mass from $1 - \pi_k^m$ to $\pi_k^m$ with $k \in \{1, 2\}$ until one of them reaches 1. In particular, if the low type would choose action $a_1$ ($a_2$) without supplemental information, it is optimal to set $\pi_1^L = 1$ ($\pi_2^L = 1$). Intuitively, the data-seller offers the low type an experiment that reveals without noise the state that matches the action that they would have selected without supplemental information.

[11]For example, define $\pi_1^m + \pi_2^m$ as the precision of experiment $E^m$ and $\pi_1^m - \pi_2^m$ as its position.

### 3.4.2 Complete characterization.

The complete characterization of the information provided to the low type depends on the:

1. *Payoff environment*: The payoff environment determines the presence of coordination or anti-coordination incentives, which pins down the effect of information observed by others on their willingness to pay for an experiment.

2. *Correlation between data-buyers' private information:* The correlation between data-buyers' private information affects their beliefs about the information observed by others. In particular, data-buyers' private information is positively (negatively) correlated if $\nu < \sqrt{\rho} - \rho$ ($\nu > \sqrt{\rho} - \rho$).[12]

3. *Support of the distribution of data-buyer types*: The support of the distribution of data-buyer types determines whether or not they choose the same action if they don't supplement their information, which impacts a data-buyer' ranking of partially informative experiments.

Data-buyers' interim beliefs are strictly congruent (non-congruent) if both types choose the same action (different actions) without supplemental information, for any menu $\mathcal{M}$. Assume that the high type $\theta_H$ chooses action $a_1$ without supplemental information. In this case, beliefs are strictly congruent if

$$i)\ \theta_L > \theta_H \geq \frac{1}{2c} \text{ when } c \in \left(\frac{1}{2}, 1\right) \text{ and}$$
$$ii)\ \theta_L > \theta_H \geq \frac{c}{2} \text{ when } c \in (1,2)$$

and strictly non-congruent if

$$i)\ \theta_H \geq \frac{1}{2c},\ \theta_L \leq \frac{c}{2} \text{ and } \theta_L < 1-\theta_H \text{ when } c \in \left(\frac{1}{2}, 1\right)$$
$$ii)\ \theta_H \geq \frac{c}{2},\ \theta_L \leq \frac{1}{2c} \text{ and } \theta_L < 1-\theta_H \text{ when } c \in (1,2).$$

[12] The Pearson's correlation coefficient is positive if and only if $\nu < \sqrt{\rho} - \rho$.

#### 3.4.2.1 Optimal menu with non-congruent beliefs

When beliefs are strictly non-congruent, the optimal experiment offered to the low type is partially informative, as stated in Proposition III.2. Proposition III.2 implies that qualitative results from Bergemann, Bonatti, and Smolin (2018) for one data-buyer setting extend to multiple data-buyers. That is, the high type is offered the perfectly informative experiment and the low type is offered a partially informative experiment.

**Proposition III.2** *Suppose that beliefs are strictly non-congruent. In an optimal menu, the data-seller offers perfect information to the high type and partial information to the low type.*

With strictly non-congruent beliefs, different data-buyer types may disagree on the ranking of partially informative experiments, as illustrated in Figure 3.2. For instance, if $(\theta_L, \theta_H) = (0.2, 0.6)$, the low type prefers experiment $(1/2, 1)$ to experiment $(1, 1/2)$, whereas the high type prefers experiment $(1, 1/2)$ to $(1/2, 1)$. As a result, the data-seller can offer the low type information that has no value to the high type, by making it sufficiently imprecise ($\pi_1^L$ sufficiently low). In an optimal menu, the data-seller selects $\pi_1^L$ such that the high type is indifferent between acquiring experiment $E^L$ or $E^H$. That is,

$$t^H - t^L = \theta_H(1-\pi_1^L)\left[\frac{\nu}{\nu+\rho}(\pi_1^L + (1-\pi_1^L)c) + \frac{\rho}{\nu+\rho}\right].$$

The right-hand side is the product of the probability of state $\omega_1$, the additional precision, and the gain of choosing action $a_1$ over action $a_2$ when the state is $\omega_1$. The left-hand side is the price differential. Hence, the information offered to the low type is such that the price differential equals the expected gain in state $\omega_1$.

Even though the qualitative properties of the optimal menu are independent of the strategic incentives and the correlation of private information, its quantitative properties are determined by their interaction, as stated in Lemma III.3.

**Lemma III.3** *The precision of the optimal $E^L$ decreases as coordination incentives increase. Moreover, the effect of increasing the correlation of private information depends on coordination incentives:*

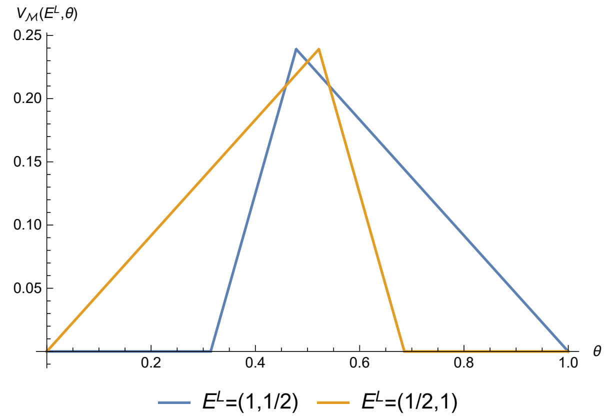



Figure 3.2: Value of $E^L$ when private information is conditionally independent $((\nu, \rho) = \left(\frac{1}{4}, \frac{1}{4}\right))$, data-buyers have coordination incentives ($c = \frac{2}{3}$) and $E^H$ is fully informative.

1. *If data-buyers have coordination incentives, the precision of $E^L$ decreases in the correlation of private information.*

2. *If data-buyers have anti-coordination incentives, the precision of $E^L$ increases in the correlation of private information.*

The precision of the optimal experiment $E^L$ decreases as coordination incentives increase, because the value of $E^L$ increases for the high type and decreases for the low type. Then, an increase in coordination incentives implies that the high type has higher incentives to deviate, reducing the data-seller's scope to provide information to the low type. As a result, the precision of the optimal $E^L$, characterized by $V_{\mathcal{M}}(E^L, \theta_H) = V_{\mathcal{M}}(E^L, \theta_L)$, decreases. Moreover, with coordination (anti-coordination) incentives, the precision of $E^L$ decreases (increases) in the correlation of private information. When data-buyers have coordination (anti-coordination) incentives, an increase in the correlation of their private information decreases (increases) the willingness to pay of both types and decreases (increases) the precision of the optimal $E^L$.

#### 3.4.2.2 Optimal menu with congruent beliefs

When data-buyers' beliefs are strictly congruent, the information offered to the low type is determined by the interaction between strategic incentives in the action stage and the correlation of private information, as stated in Proposition III.3.

**Proposition III.3** *Assume data-buyers' beliefs are strictly congruent. In an optimal menu, the data-seller offers perfect information to the high type and offers the low type:*

1. *partial information, if data-buyers' private information is negatively correlated and $\theta_L < \hat{\theta}$ if they have coordination incentives.*

2. *partial information, if data-buyers' private information is positively correlated and $\theta_L < \hat{\theta}$ if they have anti-coordination incentives.*

3. *no information, otherwise.*

Coordination incentives play no role in determining the features of an optimal menu if and only if private information is conditionally independent ($\nu = \sqrt{\rho} - \rho$) or data-buyers' payoffs are independent of each others' choices ($c = 1$). In both cases, the qualitative properties of a one data-buyer menu generalize to a two data-buyer setting. That is, in any optimal menu, the high type learns the state and the low type is offered no information. Otherwise, the optimal menu is determined by the interaction between strategic interactions and the correlation of private information.

When there are coordination incentives ($c < 1$), data-buyers face no trade-off between matching the state and each others' actions. Hence, the value of an experiment increases in the precision of the experiment observed by others, because it increases the correlation between the state and their action choices, allowing data-buyer $i$ to better predict $j$'s action choice. Information acts a coordination device and is valuable for two reasons: it reduces uncertainty about the state and about the choices of other data-buyers. When predicting the action choice of other data-buyers, each data-buyer makes inferences about what information has been gathered by others, which depends on the correlation between their private information. If their private information is positively (negatively) correlated, data-buyers assign a higher (lower) probability to observing the same private information and acquiring the same experiment. Thus, demand for information is higher when private information is negatively correlated, since it reduces further the probability of mis-coordination. The increase in demand creates a scope for the data-seller to offer partial information to the low type, as long as the low type is sufficiently unsure about the state. When private information is positively correlated, demand for information is reduced in comparison to the conditionally independent case in which the low type is offered no information. As such, the low type is also offered no information.

When data-buyers have anti-coordination incentives ($c > 1$), the trade-off between matching the state and each others' actions implies that the value of experiment $E^n$ decreases in the precision of the information observed by others, since it increases the correlation between their action and the state. Data-buyers want to be as informed as possible about the state, but their choices to be as uncorrelated with each other as possible. Information is still valuable because it allows data-buyers to learn about the state, but they value information more when their private information is positively correlated, because it reduces the correlation between their action choices. If the low type is sufficiently uncertain about the state, this increase allows the data-seller to offer partial information to the low type even when beliefs are strictly congruent without incurring any cost in terms of surplus extraction from the high type. In contrast, acquiring supplemental information when private information is negatively correlated increases the correlation between action choices through increasing the correlation with the state, decreasing data-buyers' willingness to pay for supplemental information with respect to the conditionally independent case. As such, the data-seller also offers no information to the low type.

In both cases, the partially informative experiment offered to the low type is such that the price differential equals the expected gain in state $\omega_2$. That is:

$$t^H - t^L = (1-\theta_H)(1-\pi_2^L)\left[1 + \pi_2^L\left(\frac{\nu}{\nu+\rho}\right)(1-c)\right].$$

The left-hand side is the price differential. The right-hand side is the product of the probability of state $\omega_2$, the probability of observing signal $s_1$ in state $\omega_2$ and the expected payoff gain. The expected payoff gain depends on the probability of observing different private information, the presence of coordination incentives and the precision of experiment $E^1$ in state $\omega_2$.

The quantitative properties of the optimal menu also depend on strategic incentives and the correlation of private information, as stated in Lemma III.4. This result and its intuition are analogous to Lemma III.3.

**Lemma III.4** *Assume that the optimal $E^L$ is partially informative. The precision of the optimal $E^L$ decreases as coordination incentives increase and:*

1. *decreases in the correlation of private information if data-buyers have coordinate incentives.*

2. *increases in the correlation of private information if data-buyers have anti-coordination incentives.*

#### 3.4.2.3 Example

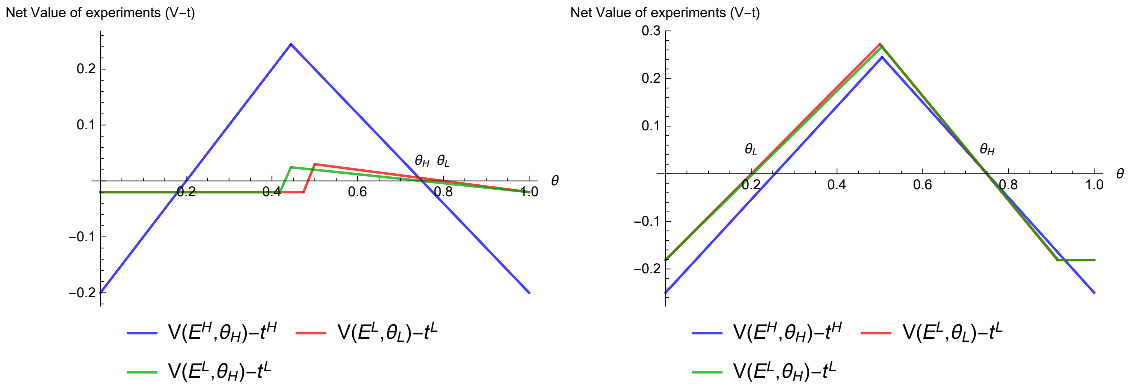
This example derives the optimal menu for two data-buyer type distributions, one with strictly congruent beliefs and one with strictly non-congruent beliefs, assuming the same strategic environment and correlation of private information. It illustrates how the interaction between strategic incentives and correlated private information creates scope for the data-seller to offer partial information to the low type, even when data-buyers have strictly congruent types.

Let $c = \frac{2}{3}$ and $(\nu, \rho) = (0.4, 0.2)$ so data-buyers have coordination incentives and negatively correlated private information and consider two type distributions:

$$(\theta_L, \theta_H) \in \left\{ \left(\frac{4}{5}, \frac{3}{4}\right), \left(\frac{1}{5}, \frac{3}{4}\right) \right\}.$$

If $(\theta_L, \theta_H) = \left(\frac{4}{5}, \frac{3}{4}\right)$, data-buyers have strictly congruent beliefs, whereas if $(\theta_L, \theta_H) = \left(\frac{1}{5}, \frac{3}{4}\right)$, data-buyers have strictly non-congruent beliefs. In the first case, the optimal menu offers the perfectly informative experiment to the high type at a price of $t^H = 0.2$ and partial information to the low type, characterized by $\pi_1^L = 1$ and $\pi_2^L = 0.1$, at a price of $t^L = 0.05$. In the second case, the high type is offered the perfectly informative experiment at a price of $t^H = 0.25$ and partial information to the low type, characterized by $\pi_1^L = 0.9$ and $\pi_2^L = 1$, at a price of $t^L = 0.18$.

Figure 3.3 illustrates the optimal menus for these two type distributions. It depicts the value of two experiments, $E^L$ and $E^H$, net of their prices, as a function of data-buyer type $\theta$. A first feature of the optimal menu is full surplus extraction by the data-seller. This is shown by the intersection of $V_{\mathcal{M}}(E^L, \theta_L) - t^L$ and $V_{\mathcal{M}}(E^H, \theta_H) - t^H$ with the x-axis (net value of zero) at their type. Second, the net value of $E^L$ for the high type $(V_{\mathcal{M}}(E^L, \theta_H) - t^H)$ is also zero at $\theta_L$, implying that the high type is




(a) Strictly congruent beliefs: $\theta_H = \frac{3}{4}$ and $\theta_L = \frac{4}{5}$

(b) Strictly non-congruent beliefs: $\theta_H = \frac{3}{4}$ and $\theta_L = \frac{1}{5}$

Figure 3.3: Optimal menu with coordination incentives ($c = \frac{2}{3}$) and negatively correlated private information ($\rho = 0.2$ and $\nu = 0.4$).

indifferent between acquiring experiments $E^L$ and $E^H$. Compared to previous work in the literature, $V_{\mathcal{M}}(E^L, \theta_H)$ and $V_{\mathcal{M}}(E^L, \theta_L)$ differ due to the correlation between data-buyers' private information.

## 3.5 N data-buyers

In this section, I extend my results to a setting with $N \geq 2$ data-buyers. Data-buyers are privately informed about the state and attach probability $\theta \in \{\theta_L, \theta_H\}$ to state $\omega_1$. Data-buyers' private information is correlated. In particular, assume that data-buyers' types $\theta_1, ...\theta_N$ are exchangeable random variables with correlation given by

$$\eta_L = \mathbb{P}(\theta_j = \theta_L | \theta_i = \theta_L) \in (0, 1] \text{ and } \eta_H = \mathbb{P}(\theta_j = \theta_H | \theta_i = \theta_H) \in (0, 1]$$

for all $j$.[13] Let $k \in \{0, 1, ..., N\}$ be the number of high types among the $N$ data-buyers. Denote by $\rho_k$ the probability of observing $k$ high type data-buyers and $N-k$ low types, where $\rho_k \geq 0$ and $\sum_{k=0}^{N} \rho_k = 1$.

In the action stage, ex-post payoffs depend on whether or not data-buyer $i$ matches

[13] Exchangeability is a property of the joint distribution of random variables. Exchangeable random variables, though correlated, have equal distributions, i.e., the probability of $\theta_i = \theta_H$ is constant across $i$.

the state and on whether or not the majority of data-buyers choose the same action as $i$. In particular, ex-post payoffs are given by:

$$u_i(a,\omega) = \begin{cases} 1 & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 > \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_\ell \\ c & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 \leq \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_\ell \\ 0 & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 \leq \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_{\ell'} \\ 0 & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 > \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_{\ell'} \end{cases}$$

where $\kappa^\ell_{-i}$ is the number of data-buyers $-i$ who choose action $a_\ell \in \{a_1, a_2\}$ and $c > 0$. As in the two data-buyer case, it is an ex-post dominant strategy to select the action that matches the state. Data-buyers are said to have coordination incentives if they prefer to match the majority and anti-coordination incentives otherwise. That is, data-buyers have coordination (anti-coordination) incentives if $c < 1$ ($c > 1$).

The data-seller's problem is to select the optimal menu of experiments to maximize her expected profits subject to the data-buyers' participation and incentive-compatibility constraints. That is:

$$\max_{(E^m,t^m)_{m=1\in\{L,H\}}} \sum_{k=0}^{N} \rho_k \left((N-k)t^L + k \cdot t^H\right)$$

subject to

$$\begin{aligned} V_{\mathcal{M}}(E^L,\theta_L) - t^L &\geq 0 \\ V_{\mathcal{M}}(E^H,\theta_H) - t^H &\geq 0 \\ V_{\mathcal{M}}(E^L,\theta_L) - t^L &\geq V_{\mathcal{M}}(E^H,\theta_L) - t^H \text{ and} \\ V_{\mathcal{M}}(E^H,\theta_H) - t^H &\geq V_{\mathcal{M}}(E^L,\theta_H) - t^L. \end{aligned}$$

**Value of information.** Suppose that all buyers but $i$ purchase the experiment designed for their corresponding type. The number of data-buyers $-i$ who choose action $a_1$ conditional on the state $\omega$ and on the type of data-buyer $i$, $\kappa^1_{-i}|(\omega,\theta_i)$, is distributed according to a Conway-Maxwell-Binomial distribution,[14] with parameters

[14]The Conway-Maxwell-Binomial distribution generalizes the binomial distribution and allows both positive and negative correlation among the exchangeable Bernoulli trials. See Kadane et al.

$N-1$, $\nu$ and

$$p_{\omega,\theta_i} = \mathbb{P}(\theta_j = \theta_H|\theta_i)\mathbb{P}(s_j = s_1|\omega, \theta_j) + \mathbb{P}(\theta_j = \theta_L|\theta_i)\mathbb{P}(s_j = s_1|\omega, \theta_j).$$

The parameter $\nu$ characterizes the underlying correlation among Bernoulli trials, which captures the correlation among data-buyers' private information. In particular, if $\nu > 1$, the Bernoulli random variables are negatively correlated. Conversely, when $\nu < 1$, the Bernoulli random variables are positively correlated. Lastly, if $\nu = 1$, the Conway-Maxwell-Binomial distribution simplifies to a Binomial distribution in which Bernoulli trials are independent.

Define $\Lambda_k^\theta$ as the expected gain of choosing the action that matches state $\omega_k$ with $k \in \{1, 2\}$ conditional on data-buyer $i$ being type $\theta$. That is:[15]

$$\Lambda_1^\theta = \mathbb{P}\left(\kappa_{-i}^1 + 1 \leq \left\lceil \frac{N}{2} \right\rceil |\omega = \omega_1\right) c + \mathbb{P}\left(\kappa_{-i}^1 + 1 > \left\lceil \frac{N}{2} \right\rceil |\omega = \omega_1\right)$$

and

$$\Lambda_2^\theta = \mathbb{P}\left(\kappa_{-i}^1 \geq N - \left\lceil \frac{N}{2} \right\rceil |\omega = \omega_2\right) c + \mathbb{P}\left(\kappa_{-i}^1 < N - \left\lceil \frac{N}{2} \right\rceil |\omega = \omega_2\right).$$

Data-buyer $i$'s expected gain of acquiring information when she would choose $a_2$ and $a_1$ without observing supplemental information are respectively given by:

$$V_2(E^n, \theta|\kappa_{-i}^1) = \theta\pi_1^n\Lambda_1^\theta - (1-\theta)(1-\pi_2^n)\Lambda_2^\theta \text{ and}$$
$$V_1(E^n, \theta|\kappa_{-i}^1) = (1-\theta)\pi_2^n\Lambda_2^\theta - \theta(1-\pi_1^n)\Lambda_1^\theta.$$

Then, data-buyer $i$'s willingness to pay for experiment $E^n$ is

$$V_{\mathcal{M}}(E^n, \theta) = \begin{cases} \max\{0, V_2(E^n, \theta|\kappa_{-i}^1)\} & if\ \alpha_{i,\iota_{i\theta}=0} = a_2 \\ \max\{0, V_1(E^n, \theta|\kappa_{-i}^1)\} & if\ \alpha_{i,\iota_{i\theta}=0} = a_1. \end{cases}$$

(2016) and Daly and Gaunt (2015) for details. Note that the probability of a data-buyer $j$ choosing action $a_1$ conditional on the state and $\theta_i$ is constant across $j$.

[15]Note that $\kappa_{-i}^1 + \kappa_{-i}^2 = N - 1$. Then, $\kappa_{-i}^2 + 1 \leq \lceil \frac{N}{2} \rceil$ is equivalent to $\kappa_{-i}^1 \geq \lceil \frac{N}{2} \rceil$.

Assume that the value of experiment $E^n$ is increasing in its precision.[16] That is:

$$c < 1 \text{ and } c\left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq \left\lceil \frac{N}{2} \right\rceil \tag{3.3}$$

or

$$c > 1 \text{ and } \left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq c\left\lceil \frac{N}{2} \right\rceil. \tag{3.4}$$

**Optimal menu.** Proposition III.1 extends to the case with $N$ data-buyers if willingness to pay for an experiment goes up as its precision increases. Hence, it is optimal for the data-seller to offer the perfectly informative experiment to the high type, whereas the information provided to the low type depends on the coordination incentives and the distribution of data-buyers types. Specifically, it depends on whether or not interim beliefs are strictly congruent.

The information provided to the low type is stated in Proposition III.4 and Proposition III.5, which extend the previous results to the $N$ data-buyer case. The interpretation is analogous as for the two data-buyer case.

**Proposition III.4** *Assume that data-buyers' beliefs are strictly non-congruent. In an optimal menu, the high type observes perfect information and the low type observes partial information.*

**Proposition III.5** *Assume that data-buyers' beliefs are strictly congruent. In an optimal menu, the high type observes perfect information and the low type observes*

1. *partial information when data-buyers' private information is negatively correlated and $\theta_L < \tilde{\theta}$ if data-buyers have coordination incentives.*

2. *partial information when data-buyers' private information is positively correlated and $\theta_L < \tilde{\theta}$ if data-buyers have anti-coordination incentives.*

3. *no information otherwise.*

[16] These restrictions on the payoff structure are sufficient but not necessary conditions for the value of experiment $E^n$ to be increasing in its precision. See appendix B.2 for details.

## 3.6 Conclusion

This paper considers a setting in which a monopolist data-seller offers supplemental information to privately informed data-buyers. Consistent with previous work, data-buyers' demand for information depends on the precision of their private information. However, it also depends on the correlation of the data-buyers' private information and their strategic interactions. The correlation between private information and coordination incentives interact in a meaningful way. In particular, positive correlation increases (decreases) data-buyer's demand for information with respect to the conditionally independent case if they have anti-coordination (coordination) incentives. Similarly, negative correlation increases (decreases) data-buyers demand for information with respect to the conditionally independent case if they have coordination (anti-coordination) incentives.

The data-seller offers a menu of experiments to screen the data-buyers' types. The interaction between coordination incentives and the correlation of private information is the main determinant of the features of the optimal menu. Whenever this interaction increases demand for information, the data-seller is able to offer partial information to the low type even when beliefs are strictly congruent. In any optimal menu, the data-seller reveals the state to the data-buyer type with the highest willingness to pay for such information. If private information leads data-buyers to choose different actions in the absence of supplemental information, the data-seller can exploit the position of information to provide partial information to the data-buyer with the lowest willingness to pay, without conceding rents to the high type. Indeed, the data-seller can provide partial information to the low type if data-buyers have coordination incentives and their private information is negatively correlated or if data-buyers have anti-coordination incentives and their private information is positively correlated.

These results highlight that the interaction of strategic incentives and correlated private information can relax the incentive compatibility constraints, allowing the data-seller to increase profits by not excluding the low type segment from the market. Considering strategic interactions between data-buyers when designing information offerings is of central importance, both qualitatively and quantitatively, given that data-buyers often interact with others in markets. This non-exclusion result is consistent with results from the multidimensional screening, in which the data-seller offers distorted partial information to the low type designed to ensure that the high type is indifferent between the information offerings.

## CHAPTER IV

# Product Reviews - Information Source or Persuasion Device?

**joint work with Anne-Katrin Roesler**

## 4.1 Introduction

Online platforms with reviews are widely used by consumers. Reviews allow consumers to share information about a good's quality, influencing others' purchasing decisions and their willingness to pay for a good. When buyers have heterogeneous preferences, the seller's pricing choices can influence the set of consumers who trade and, therefore, the information shared with future consumers through reviews. As such, reviews are not only an information source for consumers, but can also act as a persuasion device if the seller appropriately chooses prices to persuade consumers to trade in the future.

We analyze a setting in which a seller offers a good of ex-ante unknown quality through a platform with a review system to sequentially arriving short-lived heterogeneous buyers. Reviews by previous buyers provide consumers with information about the quality of the good. Based on the review system, the seller chooses an optimal pricing scheme. Buyers make their purchasing decision based on the information available through reviews, their type, and the price. In this context, we plan to address the following questions: How should a platform design the optimal review system, having the best interest of consumers in mind? What pricing scheme will a seller adopt on such a platform? Is the review system only an information transmission source or can it be used as an indirect persuasion device by the seller?

When buyers have homogeneous preferences, we fully characterize the buyer-optimal review system. We show that when there is only one buyer-type and two periods, the designer's problem is equivalent to Roesler and Szentes (2017), who characterize buyer-optimal learning in a bilateral-trade model in which the buyer only observes a (noisy) signal about her valuation prior to facing a take-it-or-leave-it offer by the seller. This is because in the first period, the buyer has no information about the quality of the good prior to making her purchasing decision and, therefore, she cannot base her decision on any information about the quality of the good except her prior. Accordingly, from an ex-ante point of view, the distribution of reports is equal to the prior distribution of quality. It follows that the set of feasible distributions over reviews is given by the set of all distributions $G$ for which the prior distribution of the quality is a mean-preserving spread. The optimal review system implements buyer-optimal learning as defined in Theorem 1 from Roesler and Szentes (2017). Given the review system, it is straightforward to identify the optimal pricing strategy for the seller and the optimal purchasing decision for the buyer. We also use this insight to solve the $T$-period model with homogeneous buyers. Since the distribution of reports is equal to the prior distribution of quality in every period and there is no inter-temporal effect of prices on the distribution of reports, maximizing expected buyer-surplus is equivalent to maximizing per-period buyer-surplus.

We then show that the presence of heterogeneous buyers meaningfully affects the characterization of the buyer-optimal distribution over posterior quality estimates. In a two period model, we show that the buyer-optimal distribution with homogeneous buyers is no longer optimal. For the special case of $K$-piecewise linear distributions with $K \leq 2$, we characterize the buyer-optimal distribution over posterior quality estimates, which approximates the buyer-optimal distribution. Given the distribution of types, we show that the buyer-optimal distribution induces the seller to set the lowest price in the set of undominated prices, benefiting both the low and high type buyers. Trade occurs with a probability strictly less than 1 and is typically inefficient since a subset of buyers with strictly positive expected valuation don't purchase the good.

**Related Literature.** This paper contributes to the large literature on information design surveyed in Kamenica (2019) and Bergemann and Morris (2019). Specifically, it contributes to the literature on information design in monopoly pricing models, focusing on buyer-optimal learning. Roesler and Szentes (2017) considers a one-period bilateral-trade model in which a single buyer observes a signal about her willigness to pay and the seller makes a take-it-or-leave-it offer. We show that their setting provides a benchmark for the simplest version of our model with two periods and homogeneous consumers, and we extend this benchmark to consider a dynamic bilateral trade model with heterogenous buyer preferences. In a related paper, Ravid, Roesler, and Szentes (2019) consider a static bilateral trade model in which buyer-learning is unobservable but costly. They argue that there is a qualitative difference between information being extremely cheap and truly free, by showing that even an infinitesimal cost to the buyer can lead to discontinuously worse outcomes.

In dynamic settings, this paper relates to Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2019) and Che and Hörner (2018), which study buyer-learning in review systems. Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2019) analyze learning dynamics and characterizes the conditions for asymptotic learning of two classes of review systems: one in which buyers observe all previous reviews and on in which buyers observe only summary statistics. In their setting, the set of buyers who trade in a given period also influences the set of buyers who trade in future periods through the information transmitted by reviews. In our model, the seller has the ability to influence the set of buyers who purchase the good by selecting a pricing scheme and, therefore, can persuade future consumers to purchase by determining the information transmitted through reviews. Che and Hörner (2018) analyzes how a recommendation system provides incentives to users to learn about a product in a bandit setting.

The project also ties into the dynamic pricing literature. Related papers include Libgober and Mu (2018), Bergemann and Ozmen (2006) and Bergemann and Välimäki (2006). The first paper considers a setting with short-lived buyers, while the others focus on heterogeneous long-lived buyers. Libgober and Mu (2018) studies a dynamic pricing model where buyers learn about their value for a good over time. The authors characterize the optimal pricing scheme robust to information arrival. In contrast, we characterize the optimal review system, conditional on a seller's optimal pricing.

Bergemann and Ozmen (2006) and Bergemann and Välimäki (2006) study optimal pricing by a seller who offers a good of unknown quality. Over time, buyers can learn their valuation by purchasing the good. In contrast with our setting, these papers take the information transmission process as given and assume that buyers learn from their own experiences, rather than the experiences of others.

The remainder of the paper is organized as follows: Section 1 presents the model, Section 2 derives results in a two period setting, Section 3 discusses future plans, and Section 4 concludes.

## 4.2 The model

Consider a long-lived seller who offers a good through a platform operated by a third party, the designer. Time is discrete and finite, with $t \in \{0, 1, ..., T\}$. In each period, one short-lived buyer arrives.

Let $q$ denote the good's quality, distributed according to the continuous CDF $H$ on $[\underline{q}, \overline{q}]$ where $\mu = \int_{\underline{q}}^{\overline{q}} q \mathrm{d}H(q)$. The good's quality is constant over time. Initially, the prior $H$ is common knowledge, but the realization $q$ is unknown to both players. A buyer's private type is either high or low, $\alpha_t \in \{\alpha_\ell, \alpha_h\}$ with $\alpha_\ell < \alpha_h$ and $\mathbb{P}(\alpha_t = \alpha_i) = f_i$, and $\mathbb{E}[\alpha_t] = 0$ for all $t \in \{0, 1, ..., T\}$. Buyer types are independent across periods. A buyer's valuation, $v_t$, depends on the good's quality and her private type, $v_t := q + \alpha_t$. If trade occurs in period $t$ at price $p_t$, the seller's in-period payoff is $p_t$ and the buyer's payoff is $v_t - p_t$. Otherwise, both get a payoff of zero. The seller and the designer discount future payoffs at rate $\delta \in (0, 1)$. The seller, the designer, and buyers are risk neutral.

The timing is as follows. At time $t = 0$, prior to observing the value of the good and buyer types, the designer chooses and implements a *review system* to maximize expected consumer surplus. A *review system* $\pi = (S, \{G(\cdot|r)\}_{r\in R})$ is a set of signals $S \subseteq \mathbb{R}$ and conditional distributions $\{G(\cdot|r)\}_{r\in R}$ that determines how feasible reports $r \in R \subseteq \mathbb{R}$ are mapped into signal realizations (reviews) $s \in S$ observed by future consumers. Based on the implemented review system, but prior to observing the quality of the good, the seller commits to a pricing scheme.[1] A buyer arrives each

[1]From a technical perspective, we abstract away from the signaling component of prices - we do

period, observes her private type, the price, and information about the good's quality provided through previous reviews. Based on this information, the buyer chooses whether or not to buy the good. If the buyer purchases the good, she learns her valuation and reports the quality of the good to the review system, turning it into a signal realization for future buyers. That is, only buyers who purchase the good write reviews. Since buyers have no stakes in the decisions of future consumers, we assume that buyers report truthfully.

Each buyer's objective is to maximize her expected payoff given the price, her type, and the information provided by the review system. The seller chooses a pricing scheme in order to maximize her expected profits given the implemented review system. The designer chooses and implements the review system that maximizes expected consumer surplus.

**Benchmark: Two-periods with homogeneous consumers.** As a benchmark, consider the case in which there is only one buyer-type and two periods, $T = 1$. For simplicity, assume $\alpha = 0$. The designer's problem reduces to selecting what information about the good's quality is revealed to future customers through the review system, in order to maximize expected buyer-surplus. As it turns out, from the designer's perspective, the two period benchmark essentially reduces to a static setting. Accordingly, the results from Roesler and Szentes (2017), who characterize buyer-optimal learning in a bilateral-trade model in which the buyer only observes a (noisy) signal about her valuation prior to facing a take-it-or-leave-it offer by the seller, naturally generalize to this setting.

To see the connection to the Roesler-Szentes setting, note that in the first period, the buyer has no information about the quality of the good prior to making her purchasing decision. Hence, the buyer cannot base her decision on any information about the quality of the good except her prior. The price in the first period is such that the buyer either always or never buys. In the former case, the buyer reports the good's quality. Since reports in the first period are not pre-selected, from an ex-ante point of view, the distribution of reports is equal to the prior distribution of quality. It follows that the set of feasible distributions over reviews is given by the set of all distributions $G$ for which the prior distribution of $q$ is a mean-preserving

not allow the seller to use prices as signals about quality.

spread. The optimal review system implements buyer-optimal learning as defined in Theorem 1 from Roesler and Szentes (2017). In period $t = 1$, this review system generates a unit-elastic demand, the seller sets the lowest price on its support, and trade occurs with probability 1. In period $t = 0$, demand only depends on the price and the expectation of the quality $\mu$, implying that it is optimal for the seller to set a price $p_0 = \mu$ which induces trade with probability 1.

This insight can be used to solve the $T$-period model with homogeneous buyers. With homogeneous buyers, the distribution of reports is equal to the prior distribution of quality in every period. Since there is no inter-temporal effect of prices on the distribution of reports, maximizing expected buyer-surplus is equivalent to maximizing per-period buyer-surplus. Hence, the static solution from Roesler and Szentes (2017) is optimal. In period $t = 0$, the seller sets a price $p_0 = \mu$ and trade occurs with certainty, whereas for all $t \geq 1$, the seller sets the lowest price on the support of the buyer-optimal distribution and trade also occurs with probability 1.

Our model builds on the analysis of Roesler and Szentes (2017) by introducing two features: buyers with heterogeneous preferences and a dynamic model in which sequentially arriving buyers learn from and post reviews.

## 4.3 Heterogeneous buyers with two-periods

We now consider the case with heterogeneous buyers and two periods $t \in \{0, 1\}$. We assume that each buyer's type $\alpha_t$ is either high, $\alpha_h$, or low, $\alpha_\ell$. As previously discussed, the two-period case essentially reduces to a static setting from the point of view of the designer. In this section, we show that the presence of heterogeneous buyers meaningfully affects the characterization of the buyer-optimal distribution over posterior quality estimates. In fact, we show that the buyer-optimal distribution with homogeneous buyers is no longer optimal.

**Set of feasible information structures.** Recall that the review system is an information structure which determines how reports are translated into reviews. Since buyers and the seller are risk-neutral, only the posterior quality estimate after observing a signal matters for the buyer's purchasing decision (besides her type), and we can thus restrict attention to unbiased signals $s = \mathbb{E}[q|s]$ without loss of generality. It is by now well-known in the literature that the set of feasible CDFs of posterior estimates is the set of mean-preserving contractions of the prior:

$$\mathcal{G}_H = \{G \in \Delta([\underline{q}, \overline{q}]) : \int_{\underline{q}}^{x} H(q)\mathrm{d}q \geq \int_{\underline{q}}^{x} G(s)\mathrm{d}s \text{ for all } x \in [\underline{q}, \overline{q}] \text{ and } \int_{\underline{q}}^{\overline{q}} s\mathrm{d}G(s) = \mu\}.$$

**The designer's problem.** In period $t = 1$, the buyer's trading decision depends on $\mathbb{E}[q|s]$ and her type. A buyer type $\alpha_1$ purchases the good if $\alpha_1 + \mathbb{E}[q|s] \geq p_1$. If the CDF induced by the review system is $G$, then the buyer's payoff is

$$U(\alpha_1, G, p_1) = \int_{p_1-\alpha_1}^{\overline{q}} (\alpha_1 + s - p_1)\mathrm{d}G(s).$$

The buyer's type is unknown to the seller. From the seller's perspective, expected demand at price $p_1$ is given by the mixture distribution

$$\overline{G} = f_\ell \cdot \widetilde{G}_\ell + f_h \cdot \widetilde{G}_h,$$

where $\widetilde{G}_i(s) := G(s - \alpha_i)$ with $i \in \{\ell, h\}$. The seller's optimal price in period $t = 1$, $p_1$, is thus given as a solution to[2] $\max_s s[1 - \overline{G}(s-)]$.

Similarly, from the designer's perspective, the buyer's expected payoff in period $t = 1$ is

$$f_\ell \int_{p_1-\alpha_\ell}^{\overline{q}} (s + \alpha_\ell - p_1)\mathrm{d}G(s) + f_h \int_{p_1-\alpha_h}^{\overline{q}} (s + \alpha_h - p_1)\mathrm{d}G(s) = \int_{p_1}^{\overline{q}+\alpha_h} (s - p_1)\mathrm{d}\overline{G}(s).$$

Putting this together, the designer's problem is

$$\max_{G \in \mathcal{G}_H} \int_{p_1}^{\overline{q}+\alpha_h} (s - p_1)\mathrm{d}\overline{G}(s)$$
$$\text{s.t. } p_1 \in \arg\max_s s[1 - \overline{G}(s-)]$$

where $\overline{G}(s) = f_\ell G(s - \alpha_\ell) + f_h G(s - \alpha_h)$.

[2]Here $G(s-)$ denotes the left limit.

**Benchmarks: Full and no information.** Consider first as a benchmark the case in which the implemented review system reveals no information about the good's quality. Then, the induced distribution over posterior value estimates, denoted by $\overline{G}_{No}$, is given by

$$\overline{G}_{No}(s) = \begin{cases} 0 & \text{if } s < \mu + \alpha_\ell \\ f_\ell & \text{if } s \in [\mu + \alpha_\ell, \mu + \alpha_h) \\ 1 & \text{if } s \geq \mu + \alpha_h \end{cases}.$$

Facing this distribution of posterior value estimates, the seller's optimal price is

$$p_1^* = \mu + \alpha_\ell \text{ if } f_h \leq \frac{\mu + \alpha_\ell}{\mu + \alpha_h} \text{ and } p_1^* = \mu + \alpha_h \text{ otherwise.}$$

Hence, trade occurs with probability one when the probability of the high type $\alpha_h$ is sufficiently low, yielding profits of $\mu + \alpha_\ell$ for the seller and an expected buyer surplus of $f_h(\alpha_h - \alpha_\ell)$. Otherwise, only the high type purchases the good and pays her willingness to pay for it, implying that the seller's profits are $f_h(\mu + \alpha_h)$ and the expected buyer surplus is 0.

As a second benchmark, if the implemented review system reveals all information about the good's quality, the induced distribution over posterior value estimates, $\overline{G}_{Full}$ corresponds to

$$\overline{H}(q) = f_\ell H(q - \alpha_\ell) + f_h H(q - \alpha_h).$$

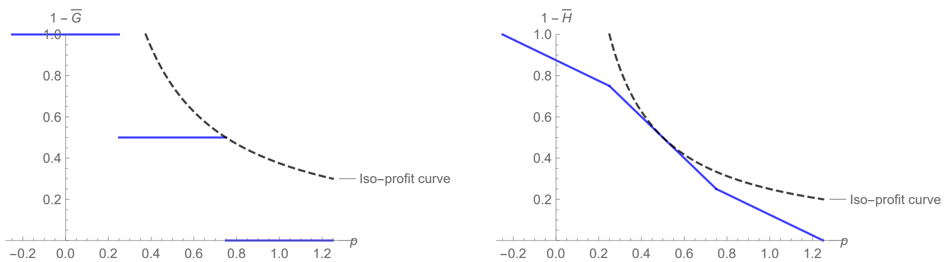
The seller's optimal price $p_1$ maximizes his expected profits,

$$p_1 \in \arg\max_s s[1 - \overline{H}(s-)],$$

and the expected buyer surplus is $\int_{p_1}^{\overline{q}+\alpha_h} (s - p_1) \mathrm{d}\overline{H}(s)$.

Another natural benchmark to consider is the buyer-optimal signal distribution with homogeneous buyers, to investigate whether buyers are better off under this distribution than under no or full information. The following example illustrates and compares these three benchmarks for the case in which the good's quality is uniformly distributed on $[0, 1]$. It shows that buyers can be better off when the distribution of posterior quality estimates is given by the buyer-optimal distribution with homogeneous types than than when they learn the good's quality.

**Example 1** *Assume that the good's quality is distributed uniformly on* $[0,1]$ *and that the distribution of buyer types is given by* $\alpha_h = -\alpha_\ell = \frac{1}{4}$ *and* $f_\ell = f_h = \frac{1}{2}$. *When the implemented review system reveals no information, the induced distribution over posterior value estimates and the seller's problem are illustrated in Figure 4.1(a). The solid blue line represents the seller's expected demand and the dashed black line, the highest attainable iso-profit curve. It is optimal for the seller to set a price of* $p^* = \frac{3}{4}$, *yielding expected profits of* $\frac{3}{8}$ *and a buyer expected surplus of* $0$.




(a) No information (b) Full information

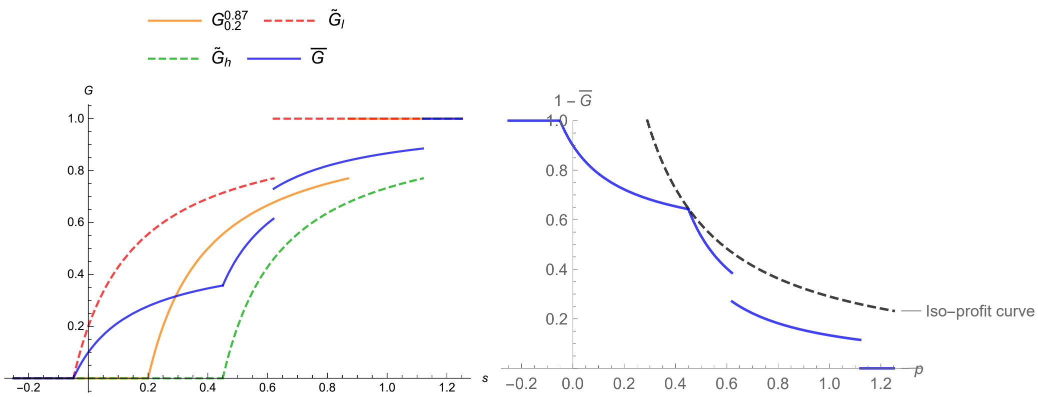
Figure 4.1: Seller's optimal price

*In contrast, if the implemented review system reveals all information, the induced distribution over posterior value estimates is* $\overline{G}_{Full}(s) = \overline{H}(s)$. *With full information revelation, the seller's problem is illustrated in Figure 4.1(b). It is optimal for the seller to set a price of* $p^* = \frac{1}{2}$, *yielding expected profits of* $\frac{1}{4}$ *and a buyer expected payoff of* $\frac{5}{32}$.

*As mentioned before, it is natural to analyze how the buyer-optimal signal with homogeneous consumers performs in comparison to these two benchmarks. Accordingly, suppose now that the implemented review system is the buyer-optimal signal for the case in which buyers have homogeneous preferences. That is, the induced distribution over posterior quality estimates is*

$$G_{0.2}^{0.87}(s) = \begin{cases} 0 & \text{if } s < 0.2 \\ 1 - \frac{0.2}{s} & \text{if } s \in [0.2, 0.87) \\ 1 & \text{if } s \geq 0.87. \end{cases}$$

*Then, the induced mixture distribution, which corresponds to the distribution of posterior value estimates, is* $\overline{G}(s) = f_\ell G_{0.2}^{0.87}(s-\alpha_\ell)+f_h G_{0.2}^{0.87}(s-\alpha_h)$. *Figure 4.2(a) illustrates the construction of this mixture distribution, which characterizes the expected demand faced by the seller in the presence of buyer heterogeneity. The solid orange line represents the distribution of posterior quality estimates. The dashed lines represent the distribution of posterior value estimates for the high and low types. Lastly, the solid blue line represents the expected demand faced by the seller.*




(a) Construction of the induced demand.

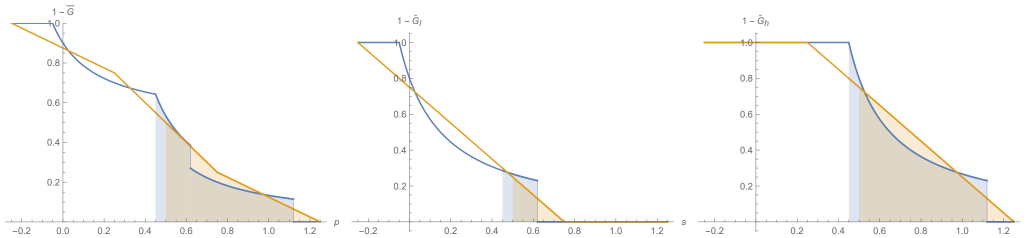
(b) Seller's optimal price.

Figure 4.2: The seller's problem.

*The seller's problem is illustrated in Figure 4.2(b). As before, the solid blue line represents the seller's expected demand and the dashed black line, the highest attainable iso-profit curve. It is optimal for the seller to set a price of* $p^* = 0.45$, *yielding expected profits of* 0.29 *and a buyer expected surplus of* 0.17. *Hence, the seller is better off when the review system reveals no information about the good's quality, whereas the buyer is better off when the distribution over posterior quality estimates is given by the buyer-optimal distribution for the case of homogeneous preferences.*

*Figure 4.3(a) compares the buyer's expected surplus for the truncated Pareto and fully disclosing distributions, illustrating that the buyer is better off under the truncated Pareto distribution than fully learning the quality of the good. In fact, both buyer-types are better off when the distribution over posterior quality is given by the buyer-optimal signal for the case in which buyers have homogeneous preferences, as illustrated in Figure 4.3(b) and (c). In particular, the low type's expected surplus under the truncated Pareto distribution is* 0.04 *whereas it is* $\frac{1}{32}$ *with full disclosure.*

*Similarly, the high type's expected surplus under the truncated Pareto distribution is* 0.29 *whereas it is* $\frac{9}{32}$ *with full disclosure.*



(a) Expected demand (b) Low type (c) High type

Figure 4.3: Buyer-surplus comparison under $\overline{G}$ (Blue) and $\overline{H}$ (Orange).

However, as shown in Figure 4.2(a), buyer heterogeneity implies that the distributions of their valuations are given by shifts of the quality distribution. As a result, the demand curve that the seller faces is a mixture distribution. Therefore, the presence of heterogeneous preferences creates a kink (and a jump) in the demand-curve faced by the seller when applying the results from the homogeneous buyer case, because it yields a truncated Pareto distribution over the posterior quality. This suggests that one can improve buyer surplus by smoothing out the demand curve in a way that reduces the seller's profit while not decreasing the gains from trade. This is illustrated in Figure 4.4, which compares the expected buyer surplus when the distribution of posterior quality estimates is given by a truncated Pareto distribution, represented as the blue line, and the one resulting from a distribution of posterior quality estimates given by

$$\hat{G}(s) = \begin{cases} 0 & if \ s < 0.13 \\ 0.26 & if \ s \in [0.13, 0.26) \\ s & if \ s \in [0.26, 1] \end{cases},$$

represented as the blue line. As shown in this figure, when buyers have heterogeneous preferences, the buyer-optimal solution is no longer a truncated Pareto distribution.

More generally, any review system that yields a truncated Pareto distribution as the CDF of posterior quality estimates has a kink and a jump in the resulting mixture distribution $\overline{G}$ over posterior value estimates. Hence, a similar construction to the

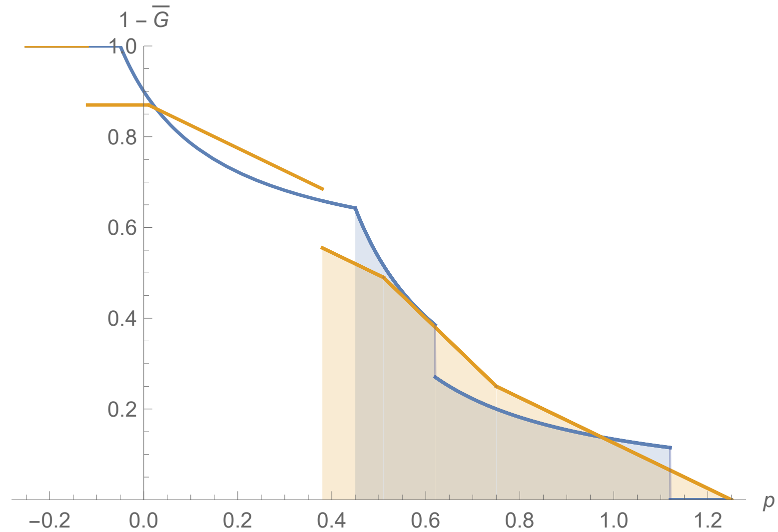



Figure 4.4: The result from the homogeneous buyer case does not directly apply.

one just illustrated can be used to obtain a signal structure that makes the buyer better off.

**Remark:** Initially, one may consider taking the distribution of buyer-valuations as the prior, i.e. the mixture distribution

$$\overline{H}(q) = f_\ell H(q - \alpha_\ell) + f_h H(q - \alpha_h),$$

and then apply the construction from Roesler and Szentes (2017) to this. However, with heterogeneous buyers and a review system that uses qualities as inputs, the resulting demand faced by the seller is given by $1 - \overline{G}$, where $\overline{G}$ is the mixture distribution over posterior value estimates induced by the signal. That is, any CDF in the set of feasible posterior value estimate CDFs induced by signals is a mixture distribution. However, a truncated Pareto distribution, and hence the solution from Roesler and Szentes (2017) applied to $\overline{H}$, cannot be represented as a mixture distribution.[3]

[3] To see this, notice that a truncated Pareto distribution has an atom at the top, but is smooth everywhere else. Such a CDF cannot be represented as a mixture distribution.

### 4.3.1 Simple review systems

As a first step to characterize the buyer optimal review system when buyers have heterogeneous preferences, we restrict attention to a simple class of review systems which induce piecewise linear distributions over posterior quality estimates. A $K$-piecewise linear distribution is defined as

$$G_K(s) = \begin{cases} 0 & \text{if } s < a_0 \\ m_1 s + n_1 & \text{if } s \in [a_0, a_1) \\ ... & \\ m_K s + n_K & \text{if } s \in [a_{K-1}, 1) \\ 1 & \text{if } s \geq 1 \end{cases}$$

where $\underline{q} \leq a_0 \leq ... \leq a_K \leq \overline{q}$, $m_k \geq 0$ for all $k \in \{1, ..., K\}$, $m_k a_k + n_k = m_{k+1} a_k + n_{k+1}$ for all $k \in \{1, ..., K-1\}$ and $m_K + n_K = 1$. Note that it is possible to approximate the buyer optimal review system as we increase $K$. In the future, we plan to use the insights and intuition gained from these simple distributions to derive a general result, yet the simplicity of these review systems makes them particularly interesting from an application perspective.

In what follows, we assume that the good's quality is distributed uniformly on $[0, 1]$ and that the distribution of buyer types is characterized by $\alpha_h = -\alpha_\ell = \alpha \in \left(0, \frac{1}{2}\right)$, where each buyer type is equally likely.

#### 4.3.1.1 K=1

We start by considering $K = 1$ and characterize the optimal 1-piecewise linear distribution, given by

$$G_1(s) = \begin{cases} 0 & \text{if } s < a \\ m \cdot s + n & \text{if } s \in [a, 1) \\ 1 & \text{if } s \geq 1 \end{cases}$$

where $0 \leq m \cdot a + n \leq m + n = 1$ and $m \geq 0$. $G_1(s)$ is a feasible distribution if and only if

$$a \in \left[0, \frac{1}{2}\right], \; m = \frac{1}{(1-a^2)} \text{ and } n = -\frac{a^2}{1-a^2}.$$

Hence, the distribution over posterior quality estimates is characterized by $a \in \left[0, \frac{1}{2}\right]$.

Lemma IV.1 characterizes the seller's optimal pricing strategy as a function of the distribution of buyer types and the one over posterior quality estimates, which determine the seller's expected demand.

**Lemma IV.1** *Given the distribution over posterior quality estimates $G_1(s)$ characterized by $a \in \left[0, \frac{1}{2}\right]$ and the distribution of buyer types characterized by $\alpha \in \left(0, \frac{1}{2}\right)$, the seller's optimal price, $p^*(\alpha, a)$, satisfies*

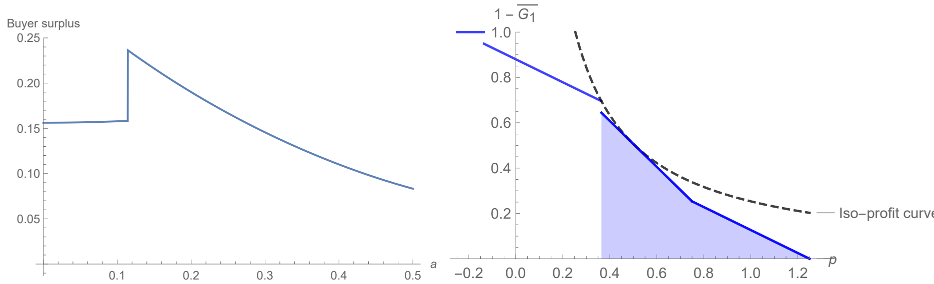
$$p^*(\alpha, a) \in P^*(\alpha, a) := \begin{cases} \{a - \alpha, a + \alpha, 1/2\} & if \; (\alpha, a) \in \left(0, \frac{1}{4}\right] \times \left[0, \frac{1}{2}\right] \\ \{a + \alpha, 1/2, (1+\alpha)/2\} & if \; (\alpha, a) \in \left(\frac{1}{4}, \frac{1}{2}\right] \times [0, 1-2\alpha) \\ \{a + \alpha, (1+\alpha)/2\} & if \; (\alpha, a) \in \left(\frac{1}{4}, \frac{1}{2}\right] \times \left[1-2\alpha, \frac{1}{2}\right] \end{cases}.$$

The following example illustrates the seller's optimal pricing strategy and the characterization of the buyer optimal distribution over posterior quality estimates within this class of 1-piecewise linear distributions.

**Example 2** *Assume $\alpha = \frac{1}{4}$ as in Example 1. The seller's optimal pricing strategy is*

$$p^*\left(\frac{1}{4}, a\right) = \begin{cases} \frac{1}{2} & if \; a < 0.11 \\ a + \frac{1}{4} & if \; a \geq 0.11 \end{cases}.$$

*Figure 4.5(a) illustrates the buyer's expected surplus as a function of $a$. It shows that the buyer's expected surplus is maximized by choosing the smallest $a$ such that the seller is willing to set a price of $a + \frac{1}{4}$. That is, $a^* = 0.11$. Figure 4.5(b) shows the demand induced by the buyer-optimal distribution $G_1^*$. Under this distribution, the seller charges a price $p^* = 0.11 + \frac{1}{4}$, obtaining expected profits of $0.253$ and the buyer, an expected surplus of $0.236$.*




(a) Buyer's expected surplus (b) Demand induced by the buyer-optimal $G_1^*$

Figure 4.5: Buyer-optimal distribution within this class and induced demand

| | Seller's profit | $\mathbb{E}_\alpha[U(\alpha, G, p^*)]$ | $U(\alpha_\ell, G, p^*)$ | $U(\alpha_h, G, p^*)$ |
|---|---|---|---|---|
| $G_{No}$ | $\frac{3}{8}$ | 0 | 0 | 0 |
| $G_{Full}$ | $\frac{1}{4}$ | $\frac{5}{32}$ | $\frac{1}{32}$ | $\frac{9}{32}$ |
| $G_{0.2}^{0.87}$ | 0.29 | 0.17 | 0.04 | 0.29 |
| $G_1^*$ | 0.253 | 0.236 | 0.075 | 0.3975 |

Table 4.1: Seller's profits and Buyer surplus

*Table 1 compares the buyer's expected surplus for each type for the benchmarks of full and no information, the buyer optimal distribution $G_1^*$ and the buyer optimal distribution with homogeneous types. Under the buyer optimal distribution $G_1^*$, both types achieve higher expected surplus.*

In the previous example, the main intuition behind the characterization of the buyer optimal $G_1$ is to choose $a \in \left[0, \frac{1}{2}\right]$ such that the seller is willing to set the lowest feasible price given the distribution of buyer types. Formally, define

$$\underline{p}^*(a|\alpha) := \min\{p : p \in P^*(\alpha, a) \text{ for a fixed } \alpha\}$$

as the minimum optimal price given the distribution of buyer types characterized by $\alpha$. Proposition IV.1 formalizes the buyer-optimal distribution $G_1^*$.

**Proposition IV.1** *Given the distribution of buyer types characterized by $\alpha \in \left(0, \frac{1}{2}\right)$, the buyer-optimal distribution $G_1^*$ is characterized by the smallest $a$ such that the seller is willing to set a price equal to $\underline{p}^*(a|\alpha)$.*

**Remark IV.1** *For all $\alpha \geq 0.003$, the buyer is better off when the distribution of posterior quality estimates is given by $G_1^*$ instead of the buyer-optimal distribution* *for homogeneous types. Moreover, the buyer is always better off when the distribution over posterior estimates is $G_1^*$ than full information.*

#### 4.3.1.2 K=2

Consider now the class of 2-piecewise linear distributions under the assumption that $G_2(s)$ is continuous at $a_0$, given by

$$G_2(s) = \begin{cases} 0 & \text{if } s < a_0 \\ m_1 s - m_1 a_0 & \text{if } s \in [a_0, a_1) \\ m_2 s - m_2 a_1 + m_1(a_1 - a_0) & \text{if } s \in [a_1, 1) \\ 1 & \text{if } s \geq 1 \end{cases}$$

where $m_1 \geq m_2 \geq 0$, $0 \leq a_0 \leq a_1 \leq 1$ and $m_2(1 - a_1) + m_1(a_1 - a_0) \leq 1$. This distribution is feasible if and only if

$$i)\ m_1^*(a_0, a_1) = \frac{a_1}{(1 - a_0)(a_1 - a_0)},\ m_2^*(a_0, a_1) = \frac{1 - a_0 - a_1}{(1 - a_0)(1 - a_1)} \text{ and}$$
$$ii)\ a_1 \in \left[0, \frac{1}{2}\right] \text{ and } a_0 \in [0, a_1] \text{ or } a_1 \in \left(\frac{1}{2}, 1\right) \text{ and } a_0 \in [0, 1 - a_1].$$

The following example illustrates the construction of the buyer optimal 2-piecewise linear distribution, which follows the same intuition as in the $K = 1$ case.

**Example 3** *Assume again $\alpha = \frac{1}{4}$. Consider the case in which $a_1 \in \left[0, \frac{1}{2}\right]$ and $a_0 \in [0, a_1]$.*[4] *It is easy to verify that any price $p < a_0 + \frac{1}{4}$ yields strictly lower profits for the seller than $p = a_0 + \frac{1}{4}$, implying that the seller would never find it optimal to set such prices. Similarly, any price $p > \frac{3}{4}$ is also never optimal for the seller since $p = a_0 + \frac{1}{4}$ yields also strictly higher profits. In fact, the seller's optimal price satisfies*

$$p^*(a_0, a_1) \in \left\{ a_0 + \frac{1}{4}, \frac{(1 - a_0)a_0(4a_1 - 7) + 8(1 - a_1)a_1}{8(2a_1(1 - a_1) - a_0(1 - a_0))}, \frac{1}{2} \right\},$$

*as illustrated in Figure 4.6.*

[4]The procedure to construct the buyer optimal $G_2$ when $a_1 \in \left(\frac{1}{2}, 1\right)$ and $a_0 \in [0, 1 - a_1]$ is analogous, but yields a strictly lower buyer surplus.

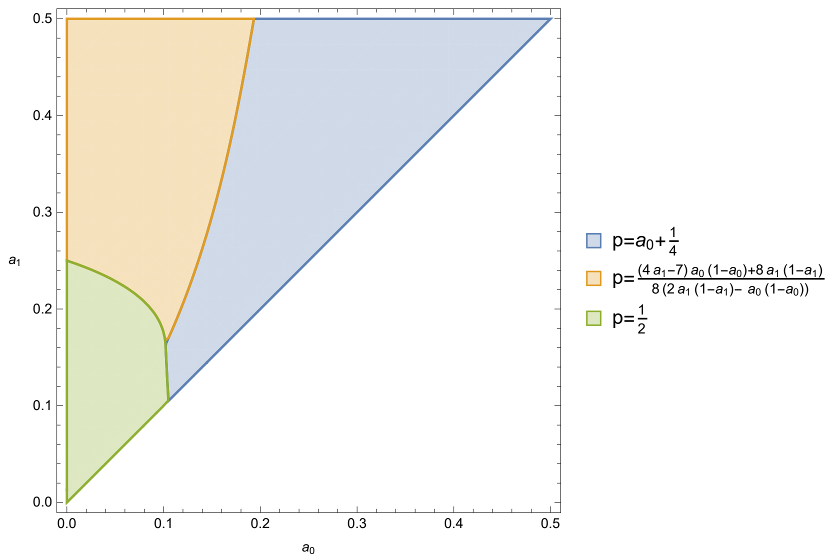
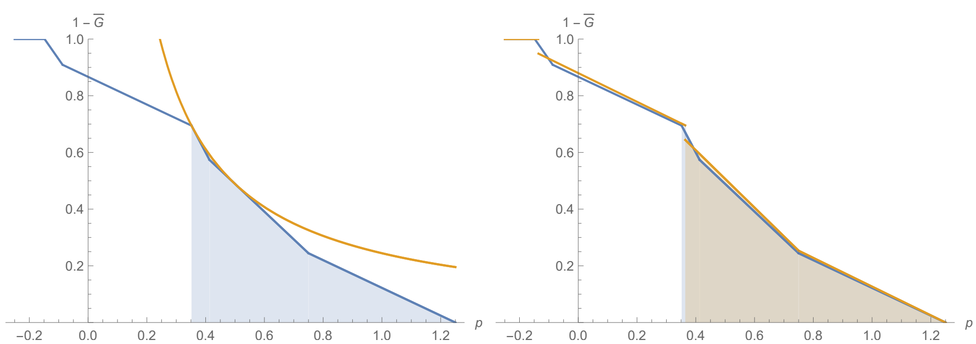



Figure 4.6: Seller's optimal pricing.

*The buyer's expected surplus is maximized by choosing the smallest* $a_0$ *such that the seller is willing to charge a price* $p = a_0 + \frac{1}{4}$*. In fact, the optimal* $(a_0^*, a_1^*)$ *are such that the seller is indifferent between the three possible optimal prices. That is, the optimal distribution over posterior quality estimates is characterized by*

$$(a_0^*, a_1^*) = (0.102, 0.163).$$

*Under this distribution, the seller charges a price* $p^* = a_0^* + \frac{1}{4}$*, the seller's profits are* 0.244 *and the buyer's expected surplus is* 0.238*, as illustrated in Figure 4.7.*




(a) Induced demand by $G_2^*$ (b) Comparison to induced demand by $G_1^*$

Figure 4.7: Induced demand by distribution $G_2^*$ (Blue) and $G_1^*$ (Orange).

| | Seller's profit | $\mathbb{E}_\alpha[U(\alpha, G, p^*)]$ | $U(\alpha_\ell, G, p^*)$ | $U(\alpha_h, G, p^*)$ |
|---|---|---|---|---|
| $G_{No}$ | $\frac{3}{8}$ | 0 | 0 | 0 |
| $G_{Full}$ | $\frac{1}{4}$ | $\frac{5}{32}$ | $\frac{1}{32}$ | $\frac{9}{32}$ |
| $G_{0.2}^{0.87}$ | 0.29 | 0.17 | 0.04 | 0.29 |
| $G_1^*$ | 0.253 | 0.236 | 0.075 | 0.3975 |
| $G_2^*$ | 0.244 | 0.238 | 0.0775 | 0.398 |

Table 4.2: Seller's profits and Buyer surplus

*Table 2 compares the buyer's expected surplus for each type for the benchmarks of full and no information, the buyer optimal distributions $G_1^*$, $G_2^*$ and the buyer optimal distribution with homogeneous types. Under the buyer optimal distribution $G_2^*$, both types achieve higher expected surplus.*

Formally, given the distribution over buyer-types characterized by $\alpha \in \left(0, \frac{1}{2}\right)$, define $\mathcal{P}((a_0, a_1)|\alpha)$ as the set of undominated prices for the seller. That is, a price $p \in \mathcal{P}((a_0, a_1)|\alpha)$ if and only if $p \in \arg\max_s s[1 - \overline{G}_2(s-)]$ for some feasible $(a_0, a_1)$ given $\alpha$. Define

$$\underline{p}^*((a_0, a_1)|\alpha) := \min\{p : p \in \mathcal{P}((a_0, a_1)|\alpha) \text{ for a fixed } \alpha\}$$

as the minimum optimal price given the distribution of buyer types characterized by $\alpha$. Proposition IV.2 formalizes the buyer-optimal distribution $G_2^*$.

**Proposition IV.2** *Given the distribution of buyer types characterized by $\alpha \in \left(0, \frac{1}{2}\right)$, the buyer-optimal distribution $G_2^*$ is characterized by the smallest $(a_0, a_1)$ such that the seller is willing to set a price equal to $\underline{p}^*((a_0, a_1)|\alpha)$.*

## 4.4 Discussion and conclusion

We use an approximation approach to characterize the buyer-optimal distribution over posterior quality estimates within a simple class of $K$-piecewise linear distributions. We show that these distributions yield higher buyer surplus in comparison to the buyer-optimal distribution with homogeneous buyers. The simplicity of these review systems is attractive from an application perspective and helps to understand the trade-off between the simplicity of a review system and its effects on the expected buyer surplus. In the future, we plan to use the insights and intuition gained from the $K$-piecewise linear distributions to derive a general result for the two period setting.

We also plan to extend the analysis to a dynamic setting. When consumers have heterogeneous preferences, the seller's pricing influences the set of consumers who trade and the information shared with future consumers through reviews. Accordingly, we conjecture that reviews are not only an information source for consumers, but can also act as a persuasion device, through which the seller induces selection bias in the set of reviews, persuading consumers to trade in the future.

# CHAPTER V

# Conclusion

This dissertation focuses on information economics in games. In Chapter II, I analyze the welfare effects of information disclosure in a duopoly model with differentiated substitutes, price competition, and uncertain demand, in which one firm has an information advantage over a competitor. My main result shows that a regulator can increase expected consumer surplus and welfare by restricting disclosure between firms, but that, surprisingly, consumers can benefit from the regulator privately disclosing some information to the competitor. Disclosure increases the ability of firms to extract surplus from consumers by pricing to better match the level of demand. But, private disclosure can create a pricing coordination failure between firms by introducing uncertainty in their choices, which increases price volatility and opportunities for consumers to arbitrage prices. The benefit from private disclosure depends on the differentiation between goods, because it determines consumers' willingness to substitute between goods and therefore the extent to which disclosure affects relative demand across firms. Thus, I show that private partial disclosure is optimal for consumers when firms offer sufficiently close substitutes and, otherwise, no disclosure is optimal.

In Chapter III, I study the optimal sale of information by a monopolist data-seller to multiple privately informed data-buyers who play a two-stage game of incomplete information. In the information stage, buyers can simultaneously acquire supplemental information to reduce their uncertainty about the state. In the action stage, buyers simultaneously select an action between two options to maximize their expected payoffs. The seller's optimal menu screens between the two types of buyers. The interaction between coordination incentives and the correlation of private information is the main determinant of the features of the optimal menu. In particular, the seller offers perfect information to the buyer with highest willingness to pay (the high type) and partial or no information to the other type (the low type). Partial information is offered to the low type whenever the mentioned interaction increases demand for information. That is, i) if they play a coordination game and their private information is negatively correlated; or ii) if they play anti-coordination game and their private information is positively correlated.

In Chapter IV, which is joint work with Anne-Katrin Roesler, we study the optimal design of review systems by a platform that has the best interest of consumers in mind. We analyze a setting in which a seller offers a good of ex-ante unknown quality through a platform with a review system to sequentially arriving short-lived heterogeneous buyers. Reviews by previous buyers provide consumers with information about the quality of the good. Based on the review system, the seller chooses an optimal pricing scheme. Buyers make their purchasing decision based on the information available through reviews, their type, and the price. We focus on how a platform should design the optimal review system, having the best interest of consumers in mind. When buyers have homogeneous preferences, we fully characterize the buyer-optimal review system. We also show that the presence of heterogeneous buyers meaningfully affects the characterization of the buyer-optimal distribution over posterior quality estimates. In a two period model, we characterize the buyer-optimal distribution over posterior quality estimates within the class of $K$-piecewise linear distributions with $K \leq 2$.

# APPENDICES

# APPENDIX A

# Appendix for Chapter II

## A.1 Useful results for the binary signal benchmark

In this section, I derive the optimal information structures for the benchmark case in which signals are restricted to be binary. Assume that the set of signals is binary and given by $S = \{s_L, s_H\}^2$. The information structure $(S, x)$ with conditional distributions $x : \Theta \to \Delta(S)$ can be represented in matrix form where rows represent firm 1's signal realization and columns firm 2's signal realization. Define $\pi$ as follows:

| $\theta = \theta_L$ | $s_L$ | $s_H$ |
|---|---|---|
| $s_L$ | $x_L$ | $1 - x_L$ |
| $s_H$ | 0 | 0 |

| $\theta = \theta_H$ | $s_L$ | $s_H$ |
|---|---|---|
| $s_L$ | 0 | 0 |
| $s_H$ | $1 - x_H$ | $x_H$ |

The set of feasible information structure, denoted by $\mathcal{D}$, is

$$\mathcal{D} := \{(x_L, x_H) \in [0, 1]^2 : x_L + x_H \geq 1\}.$$

### A.1.1 Optimal pricing and equilibrium outcomes

Given the information structure $(S, x)$ and conditional on the realization of signal

$s_i$, firm $i$ chooses $p_i(s_i) \geq 0$ to maximize her expected profits,

$$\max_{p_i(s_i)\geq 0} \Pi_i(p_i(s_i), p_{-i}(s_{-i})) = p_i(s_i) \left[\mathbb{E}[\theta|s_i] + b\mathbb{E}\left[p_{-i}\left(s_{-i}\right)|s_i\right] - ap_i(s_i)\right].$$

Equilibrium prices $(p_1^*(s_1), p_2^*(s_2))$ are the unique solution to:

$$\mathbb{E}[\theta|s_i] + b\mathbb{E}\left[p_j(s_j)|s_i\right] - 2ap_i(s_i) = 0$$

for all $s_i \in S_i$, $i \in \{1,2\}$ and $j \neq i$ where

$$\mathbb{E}[\theta|s_i = s_\ell] = \mathbb{P}(\theta = \theta_L|s_i = s_\ell)\theta_L + \mathbb{P}(\theta = \theta_H|s_i = s_\ell)\theta_H \text{ and}$$
$$\mathbb{E}\left[p_{-i}(s_{-i})|s_i = s_\ell\right] = \mathbb{P}(s_{-i} = s_L|s_i = s_\ell)p_{-i}(s_L) + \mathbb{P}(s_{-i} = s_H|s_i = s_\ell)p_{-i}(s_H).$$

### A.1.2 Consumer and welfare optimal disclosure

**Consumer optimal disclosure.** Assume that the designer's objective is to maximize expected consumer surplus, given by

$$\mathbb{E}[(CS(p_1,p_2);\theta)] = \sum_{i\in\{1,2\},\theta\in\Theta,(k,n)\in\{L,H\}^2} \mu_\theta \left[\mathbb{P}(s_i = s_k \cap s_j = s_n|\theta)\frac{\left(\theta + bp_i^*(s_k) - ap_j^*(s_n)\right)^2}{2a}\right]$$

The optimal disclosure is determined by the relationship between goods, as stated in Lemma A.1.

**Lemma A.1** *If the designer's objective is to maximize expected consumer surplus, partial disclosure is optimal if $\delta \in (\hat{c}, 1)$ and no disclosure is CS-optimal, otherwise.*

**Proof. Lemma A.1.** First, full disclosure is never CS-optimal since expected consumer surplus is higher with no information disclosure than with full disclosure since

$$CS\left(x_L, 1-x_L\right) - CS\left(1,1\right) \geq \frac{\mu_L\mu_H\left(a^4+b^4\right)\left(\theta_H-\theta_L\right)^2}{8a^3(2a-b)^2} \geq 0,$$

implying that either no or partial disclosure maximizes consumer surplus.

Second, I show that there exists $\hat{c} \in (0,1)$ such that partial disclosure is optimal if $\delta \geq \hat{c}$ and no disclosure is optimal otherwise. Define $\Delta\mathbb{E}[CS](x)$ as the difference between the expected consumer surplus with no disclosure $\pi^N = (x_L, 1 - x_L)$ and the expected consumer surplus with disclosure $(x_L, x_H)$. The sign of $\Delta\mathbb{E}[CS](x)$ is determined by

$$\Phi(a,b,x) = f_1(a,b)\mathbb{V}[s_2] + f_2(a,b)\mu_L\mu_H(x_L + x_H - 1)^2 - f_3(a,b)\mathbb{E}[\mathbb{V}[s_2|\theta]]$$

where

$$\begin{aligned} f_1(a,b) &= a^2\left(4a^2(6a+b) - b^2(18a+7b)\right) \\ f_2(a,b) &= b^4(2a+b) \\ f_3(a,b) &= a^2b^2(6a+5b) \end{aligned}$$

and $f_k(a,b) > 0$ for all $k \in \{1,2,3\}$, $\min\{f_1(a,b), f_3(a,b)\} > f_2(a,b)$ for all $a > b > 0$, $f_1(a,b) > f_3(a,b)$ if and only if $\delta < \hat{c} \approx 0.9$. This implies that

$$f_1(a,b)\mathbb{V}[s_2] > f_3(a,b)\mathbb{E}[\mathbb{V}[s_2|\theta]] \text{ if } \delta < \hat{c}$$

since $f_1(a,b) > f_3(a,b)$ and $\mathbb{V}[s_2] > \mathbb{E}[\mathbb{V}[s_2|\theta]]$. Thus, no disclosure maximizes the expected consumer surplus if $\delta \leq \hat{c}$. Otherwise, partial disclosure is CS-optimal since for all $\delta > \hat{c}$, there exists $x \in \mathcal{D}$ such that $\Phi(a,b,x) < 0$. ■

**Welfare optimal disclosure.** Assume that the designer's objective is to maximize expected welfare, defined as the sum of expected consumer surplus and expected firm profits.

**Lemma A.2** *Assume that the designer's objective is to maximize expected welfare. If*

1. $\delta \in (0, \tilde{c}_1]$, *no disclosure is optimal*

2. $\delta \in (\tilde{c}_1, \tilde{c}_2)$, *partial disclosure is optimal.*

3. $\delta \in [\tilde{c}_2, 1)$, *full disclosure is optimal.*

**Proof. Lemma A.2.** First, define $\Delta\mathbb{E}[TS_1](x)$ as the difference in expected welfare with full disclosure $x^F$ and a partial disclosure characterized by $x$. The sign of this difference is determined by

$$\rho_1(a,b,\mu,x) = f_4(a,b)\mathbb{V}[s_2] + f_5(a,b)\mathbb{E}[\mathbb{V}[s_2|\theta]] + \mu_L\mu_H f_6(a,b)(x_L + x_H - 1)^2$$

where

$$f_4(a,b) = 16a^5(3b-a),\ \ f_5(a,b) = 4a^2b^2\left(5a^2-b^2\right) \text{ and}$$
$$f_6(a,b) = b^2(16a^4 - 12a^3b + a^2b^2 - b^4)$$

Note that $\rho_1(a,b,\mu,x) > 0$ for all $x \in \mathcal{D}$ if $\delta \geq \tilde{c}_2 \approx 0.31$ and there exists $x \in \mathcal{D}$ such that $\rho_1(a,b,\mu,x) < 0$ if $\delta < \tilde{c}_2$. Thus, full disclosure is optimal if $\delta \geq \tilde{c}_2$ and either partial or no disclosure is optimal otherwise.

Second, define $\Delta\mathbb{E}[TS_2](x)$ as the difference of expected welfare with no disclosure $x^N$ and with partial disclosure $x$. The sign of $\Delta\mathbb{E}[TS_2](x)$ is determined by the sign of

$$\rho_2(a,b,\mu,x) = f_7(a,b)\mathbb{V}[s_2] - f_8(a,b)\mathbb{E}[\mathbb{V}[s_2|\theta]] - f_9(a,b)\mu_L\mu_H(x_L + x_H - 1)^2$$

where

$$f_7(a,b) = 4a^4(2a-5b),\ \ f_8(a,b) = 12a^2b^2(2a+b) \text{ and } f_9(a,b) = b^2\left(22a^3 + 5a^2b - b^2(2a+b)\right)$$

and $f_k(a,b) > 0$ for all $k \in \{7,8,9\}$ since $a > b > 0$. Note that $\rho_2(a,b,\mu,x) \geq 0$ for all $x \in \mathcal{D}$ if $\delta \leq \tilde{c}_1 \approx 0.29$ and for $\delta > \tilde{c}_1$, there exists $x \in \mathcal{D}$ such that $\rho_2(a,b,\mu,x) < 0$. Thus, no disclosure is optimal if $\delta \leq \tilde{c}_1$ and either full or partial disclosure is optimal otherwise. In summary, no disclosure is optimal if $\delta \leq \tilde{c}_1$, partial disclosure is optimal if $\delta \in (\tilde{c}_1, \tilde{c}_2]$ and full disclosure is optimal if $\delta > \tilde{c}_2$. ■

## A.2 Proofs

### A.2.1 Preliminary results: proofs

**Proof. Lemma II.1.** The pricing game is a smooth concave game since $\Pi_i((\cdot, p_{-i});\theta) : \mathbb{R}_+ \to \mathbb{R}$ is concave and continuously differentiable for each $p_{-i} \in \mathbb{R}_+$ since the demand is linear in $p_{-i}$. Define the payoff gradient as

$$\nabla\Pi(\mathbf{p},\theta) := \left(\frac{\partial\Pi_i((p_i,p_{-i});\theta)}{\partial p_i}\right)_{i\in\{1,2\}},$$

where firm $i$'s ex-post payoff function is given by $\Pi_i((p_i,p_{-i});\theta) = p_i(\theta - ap_i + bp_{-i})$. Then, the payoff gradient, given by

$$\nabla\Pi(\mathbf{p},\theta) = (\theta + bp_{-i} - 2ap_i)_{i\in\{1,2\}},$$

is continuously differentible. The Jacobian matrix of the payoff gradient, given by

$$F_{\nabla\Pi}(\mathbf{p},\theta) := \begin{pmatrix} \frac{\partial^2\Pi_1((p_1,p_2);\theta)}{\partial p_1^2} & \frac{\partial^2\Pi_1((p_1,p_2);\theta)}{\partial p_1\partial p_2} \\ \frac{\partial^2\Pi_2((p_2,p_1);\theta)}{\partial p_1\partial p_2} & \frac{\partial^2\Pi_2((p_2,p_1);\theta)}{\partial p_2^2}\end{pmatrix} = \begin{pmatrix} -2a & b \\ b & -2a\end{pmatrix},$$

is negative definite because $-2a < 0$ and $4a^2 - b^2 > 0$ since $a > |b|$. This implies that the payoff gradient $\nabla\Pi(\mathbf{p},\theta)$ is strictly monotone by Lemma 4 from Ui (2016). Furthermore, since for all $\mathbf{p}$, there exists $c > 0$ such that

$$\mathbf{p}^T F_{\nabla\Pi}(\mathbf{p},\theta)\mathbf{p} < -c\mathbf{p}^T\mathbf{p},$$

the payoff gradient is also strongly monotone by the same lemma. Then, the uniqueness of the Bayesian Nash equilibrium of the pricing game follows from Proposition 1 from Ui (2016), which states that if the payoff gradient is strictly monotone, the Bayesian game as at most one Bayesian Nash equilibrium. The existence of a unique Bayesian Nash equilibrium follows from Proposition 2 from Ui (2016). ■

**Proof. Lemma II.2.** First, I show that the set of BCE is a subset of $\cup_{(S,\psi)}\mathcal{E}(S,\psi)$. Assume $\sigma \in BCE$. Then, $\sigma$ satisfies

$$\sum_{\theta\in\Theta}\mu_\theta\int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]}\Pi_i((p_i,p_{-i}),\theta)\mathrm{d}\sigma((p_i,p_{-i})|\theta) \geq \sum_{\theta\in\Theta}\mu_\theta\int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]}\Pi_i((p_i',p_{-i}),\theta)\mathrm{d}\sigma((p_i,p_{-i})|\theta) \tag{A.1}$$

for all $p_i \in \operatorname{supp}\sigma$, $p_i' \in \left[0,\frac{\theta_H}{a-b}\right]$ and $i \in \{1,2\}$.

Consider an information structure $\left(\left[0,\frac{\theta_H}{a-b}\right]^2,\psi^*\right)$ where $\left[0,\frac{\theta_H}{a-b}\right]^2$ is the set of signal realizations and $\psi^*: \Theta \to \Delta(\left[0, \frac{\theta_H}{a-b}\right]^2)$ coincides with $\sigma$, i.e. $\sigma = \psi^*$. Let

$$\beta_i^*(p_i|p_i') = \begin{cases} 1 & if\ p_i = p_i' \\ 0 & otherwise \end{cases}$$

be the obedient strategy. Then, the right-hand side of (A.1) can be written as

$$\begin{aligned} &\sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}), \theta)\mathrm{d}\sigma((p_i, p_{-i})|\theta) \\ &= \sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}), \theta)\mathrm{d}\psi^*((p_i, p_{-i})|\theta) \\ &= \sum_{\theta\in\Theta} \mu_\theta \int_{s_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}); \theta)\mathrm{d}\beta_{-i}^*(p_{-i}|s_{-i})\mathrm{d}\psi^*((s_i, s_{-i})|\theta) \end{aligned}$$

The first equality holds by definition of $\psi^*$. The second equality holds by definition of the obedient strategy and Fubini's theorem since, fixing $\theta$, $\Pi_i((p_i, p_{-i}); \theta)$ is $\sigma$-integrable because $\Pi_i|\theta : \left[0, \frac{\theta_H}{a-b}\right]^2 \to \mathbb{R}_+$ is a bounded and continuous real-valued function on a compact set.[1] Hence, the BNE incentive-compatibility constraints are implied by the BCE obedience constraints. This, in turn, implies that if $\sigma \in BCE$, then $\sigma$ is also a BNE of the game. Thus, the set of BCE is a subset of the set of BNE of the game.

Second, I show that $\cup_{(S,\psi)}\mathcal{E}(S, \psi)$ is a subset of BCE. Consider a BNE composed by an information structure $(\hat{S}, \hat{\psi})$ with $\hat{\psi} : \Theta \to \Delta(S)$ and measurable behavioral strategies $(\hat{\beta}_i, \hat{\beta}_{-i})$.[2] Given the behavioral strategies $(\hat{\beta}_i, \hat{\beta}_{-i})$, define $\hat{\beta} : S \to \Delta\left(\left[0, \frac{\theta_H}{a-b}\right]^2\right)$ as the joint measure. Let $\hat{\sigma} : \Theta \to \Delta\left(\left[0, \frac{\theta_H}{a-b}\right]^2\right)$ be the composition of $\hat{\psi}$ and $\hat{\beta}$, defined as $\hat{\sigma} = \hat{\beta} \circ \hat{\psi}$. Then, by definition $\hat{\sigma} \in \cup_{(S,\psi)}\mathcal{E}(S, \psi)$. The definition of BNE implies that $(\hat{S}, \hat{\psi})$ and $\hat{\beta}$ satisfy:

$$\begin{aligned} &\sum_{\theta\in\Theta} \mu_\theta \int_{\hat{S}_{-i}} \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i, p_{-i}); \theta)\mathrm{d}\hat{\beta}_{-i}(p_{-i}|s_{-i})\mathrm{d}\hat{\psi}((s_i, s_{-i})|\theta) \\ &\geq \sum_{\theta\in\Theta} \mu_\theta \int_{\hat{S}_{-i}} \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}); \theta)\mathrm{d}\hat{\beta}_{-i}(p_{-i}|s_{-i})\mathrm{d}\hat{\psi}((s_i, s_{-i})|\theta) \qquad \text{(A.2)} \end{aligned}$$

[1] See theorem 11.27 from Aliprantis and Border (2013) where the condition of theorem are satisfied by Proposition 3.3 and Theorem 4.4 from from Royden (1968)

[2] Behavioral strategies $\beta_i : S_i \to \Delta\left(\left[0, \frac{\theta_H}{a-b}\right]\right)$ for all $i \in \{1, 2\}$ are defined as a regular conditional probabilities as defined in Appendix C from Bass (2011).

for all $p_i' \in \left[0, \frac{\theta_H}{a-b}\right]$, $s \in S$ and $i \in \{1,2\}$. Integrating both sides of the BNE incentive-compatibility constraint, we have

$$\sum_{\theta\in\Theta} \mu_\theta \int_{\hat{S}_i} \int_{\hat{S}_{-i}} \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i, p_{-i}); \theta) \mathrm{d}\hat{\beta}_i(p_i|s_i) \mathrm{d}\hat{\beta}_{-i}(p_{-i}|s_{-i}) \mathrm{d}\hat{\psi}(s|\theta)$$
$$\geq \sum_{\theta\in\Theta} \mu_\theta \int_{\hat{S}_i} \int_{\hat{S}_{-i}} \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}); \theta) \mathrm{d}\hat{\beta}_i(p_i|s_i) \mathrm{d}\hat{\beta}_{-i}(p_{-i}|s_{-i}) \mathrm{d}\hat{\psi}(s|\theta)$$

Then, (A.2) implies that

$$\sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i, p_{-i}), \theta) \mathrm{d}\sigma((p_i, p_{-i})|\theta)$$
$$\geq \sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\left[0,\frac{\theta_H}{a-b}\right]} \Pi_i((p_i', p_{-i}), \theta) \mathrm{d}\sigma((p_i, p_{-i})|\theta)$$

■

**Proof. Lemma II.3.** Consider a distribution $\sigma \in \cup_{(S,\psi)}\mathcal{E}(S,\psi)$. Lemma II.2 implies that $\sigma \in BCE$. Consider the recommendation mechanism $(\left[0, \frac{\theta_H}{a-b}\right]^2, \psi_\sigma)$ where $\psi_\sigma = \sigma$ for all $(p_1, p_2) \in \left[0, \frac{\theta_H}{a-b}\right]^2$ and $\theta \in \Theta$ and the obedient behavioral strategy

$$\beta_i^*(p_i|p_i') = \begin{cases} 1 & if\ p_i = p_i' \\ 0 & otherwise \end{cases}.$$

The interim expected payoff of firm $i$ when firm $-i$ follows $\beta_{-i}^*$ is

$$\sum_{\theta\in\Theta} \mu_\theta \int_{S_{-i}} \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i', p_{-i}); \theta) \mathrm{d}\beta_{-i}^*(p_{-i}|p_{-i}') \mathrm{d}\psi_\sigma((p_i, p_{-i}')|\theta)$$
$$= \sum_{\theta\in\Theta} \mu_\theta \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i', p_{-i}); \theta) \mathrm{d}\psi_\sigma((p_i, p_{-i})|\theta)$$
$$= \sum_{\theta\in\Theta} \mu_\theta \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i', p_{-i}); \theta) \mathrm{d}\sigma((p_i, p_{-i})|\theta) \tag{A.3}$$

for all $i$. Hence, the definition of BCE and (A.3) imply

$$\sum_{\theta\in\Theta} \mu_\theta \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i, p_{-i}); \theta) \mathrm{d}\psi_\sigma((p_i, p_{-i})|\theta) \geq \sum_{\theta\in\Theta} \mu_\theta \int_0^{\frac{\theta_H}{a-b}} \Pi_i((p_i', p_{-i}); \theta) \mathrm{d}\psi_\sigma((p_i, p_{-i})|\theta)$$

for all $p_i' \in \left[0, \frac{\theta_H}{a-b}\right]$ and $i$. The distribution of prices conditional on the state $\theta$ under $\beta^*$ and $(\left[0, \frac{\theta_H}{a-b}\right]^2, \sigma)$ is $\psi_\sigma = \sigma$. Thus, $\sigma \in \mathcal{E}\left(\left[0, \frac{\theta_H}{a-b}\right]^2, \sigma\right)$. ■

**Lemma A.3** *The support of the distribution $\sigma(\mathbf{p}|\theta)$ is a subset of $[p^F(\theta_L), p^F(\theta_H)]^2$ for all $\theta \in \Theta$, where $p^F(\theta)$ is the equilibrium price with full disclosure when the state $\theta$ is realized.*

**Proof. Lemma A.3.** The minimum and maximum price in any equilibrium is charged when both firms know that the state is low and that the state is high, respectively. That is, the highest and lowest equilibrium prices occur with full disclosure. Under full disclosure $\sigma^F$, both firms learn the state. Let $p^F(\theta)$ be the equilibrium price under full disclosure when the state is $\theta$, where

$$p^F(\theta_L) = \frac{\theta_L}{(2a-b)} \text{ and } p^F(\theta_H) = \frac{\theta_H}{(2a-b)}$$

Hence, any obedient recommendation mechanism must recommend prices in the set of feasible equilibrium prices denoted by $[p^F(\theta_L), p^F(\theta_H)]^2$. ■

**Proof. Lemma II.4.** The set of BCE is the collection of distributions $\sigma : \Theta \to \Delta([p^F(\theta_L), p^F(\theta_H)]^2)$ such that

*i)* $\sigma((p_1, p_2)|\theta) \geq 0$ for all $(p_1, p_2) \in [p^F(\theta_L), p^F(\theta_H)]^2$ and $\theta \in \Theta$,

*ii)* $\displaystyle\int d\sigma((p_1, p_2)|\theta) = 1$ for all $\theta \in \Theta$ and

*iii)* $\displaystyle\sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\mathbb{R}_+} \Pi_i((p_i, p_{-i}), \theta) d\sigma((p_i, p_{-i})|\theta) \geq \sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\mathbb{R}_+} \Pi_i((p_i', p_{-i}), \theta) d\sigma((p_i, p_{-i})|\theta)$

for all $p_i \in \text{supp } \sigma$, $p_i' \in \mathbb{R}_+$ and $i \in \{1, 2\}$.

First, Theorem A from Stinchcombe (2011) establishes the existence of Correlated equilibrium in games in which players receive private signals and then simultaneously choose actions from compact sets. Formally, consider a game in which the set of players $I$ is finite and for each $i$, the type $\omega_i$ belongs to the measure space $(\Omega_i, \mathcal{F}_i)$. Each player $i$ simultaneously chooses an action from a compact set $A_i$ and denote by $\Delta_i$ the set of countably additive Borel probabilities in $A_i$, with the weak* topology. Let $\mathbb{B}_i(\mathcal{F}_i)$ be the set of $i$'s behavioral strategies, defined as the $\mathcal{F}_i$-measurable functions from $\Omega_i$ to $\Delta_i$. Given a vector $b \in \mathbb{B} := \times_i \mathbb{B}_i(\mathcal{F}_i)$, player $i$'s expected utility if $b$ is played is defined by

$$u_i^P(b) = \int_\Omega \langle u_i(\omega), \times_i b_i(\omega)\rangle P(\mathrm{d}\omega)$$

where $\langle f, \nu\rangle := \int_A f(a)\nu(\mathrm{d}a)$ for $f : A \to \mathbb{R}$ and Borel probabilities $\nu$, and $\times_i b_i$ is the product probability on $A$ having $b_i$ as the marginal. $(\mathbb{B}_i(\mathcal{F}_i), u_i^P)_{i\in I})$ denotes the normal form game. Then, Theorem A shows that all games $(\mathbb{B}_i(\mathcal{F}_i), u_i^P)_{i\in I})$ have correlated equilibria.

In the pricing game, two firms simultaneously choose a price to maximize their expected equilibrium profits,

$$\sum_{\theta\in\Theta} \mu_\theta \int_{p_{-i}\in\mathbb{R}_+} \Pi_i((p_i, p_{-i}), \theta)\mathrm{d}\sigma((p_i, p_{-i})|\theta).$$

Note that $p_i \in \left[0, \frac{\theta_H}{a-b}\right]$ for all $i \in \{1, 2\}$. Hence, firms simultaneously choose prices from compact sets. Thus, this result implies that the set of BCE is non-empty.

Second, the set of BCE is the collection of distributions

$$\sigma : \Theta \to \Delta([p^F(\theta_L), p^F(\theta_H)]^2),$$

which corresponds to the set of all probability measures on $[p^F(\theta_L), p^F(\theta_H)]^2$ for each $\theta \in \Theta$ where $\Theta$ is finite. Then, the set of BCE is compact since $[p^F(\theta_L), p^F(\theta_H)]^2$ is compact in the weak* topology, by Theorem 15.11 from Aliprantis and Border (2013).

The designer's objectives are

$$i)\ \text{Informed firm optimal}: \sum_{\theta\in\Theta} \mu_\theta \int \Pi_1((p_1, p_2), \theta)\mathrm{d}\sigma((p_1, p_2)|\theta)$$

$$ii)\ \text{PS optimal}: \sum_{i\in\{1,2\}} \sum_{\theta\in\Theta} \mu_\theta \int \Pi_i((p_i, p_{-i}), \theta)\mathrm{d}\sigma((p_i, p_{-i})|\theta)$$

$$iii)\ \text{CS-optimal}: \frac{1}{2a} \sum_{i\in\{1,2\}} \sum_{\theta\in\Theta} \mu_\theta \int q_i((p_i, p_{-i}); \theta)^2\mathrm{d}\sigma((p_i, p_{-i})|\theta)$$

and

$$iv) \text{ Welfare-optimal}: \sum_{i\in\{1,2\}}\sum_{\theta\in\Theta}\mu_\theta\int\Pi_i((p_i,p_{-i}),\theta)\mathrm{d}\sigma((p_i,p_{-i})|\theta)$$
$$+\frac{1}{2a}\sum_{i\in\{1,2\}}\sum_{\theta\in\Theta}\mu_\theta\int q_i((p_i,p_{-i});\theta)^2\mathrm{d}\sigma((p_i,p_{-i})|\theta)$$

Third, the continuity of all objective functions in the weak* topology follows from Corollary 15.7 from Aliprantis and Border since because both $\Pi_i((p_i,p_{-i}),\theta)$ and $q_i((p_i,p_{-i}),\theta)$ are continuous and bounded functions. Hence, the integral

$$\int\Pi_i\mathrm{d}\sigma((p_i,p_{-i})|\theta) \text{ and } \int q_i^2\mathrm{d}\sigma((p_i,p_{-i})|\theta)$$

is continuous in $\sigma$. Thus, the designer's problem is to maximize a continuous objective function in a compact set. The existence of a solution is guaranteed by the Weierstrass extreme value theorem. ■

### A.2.2 Informed firm optimal disclosure: proofs

**Proof. Proposition II.1.** The fully disclosing information structure recommends prices $(p^F(\theta),p^F(\theta))$ with probability 1 for all $\theta\in\Theta$.

Full disclosure is optimal for the informed firm if her expected equilibrium payoffs with full disclosure exceed her expected equilibrium payoffs induced by any other obedient recommendation mechanism. That is,

$$\sum_{\theta\in\Theta}\mu_\theta\Pi_1((p^F(\theta),p^F(\theta));\theta)\geq\sum_{\theta\in\Theta}\mu_\theta\int\Pi_1((p_1,p_2);\theta)\mathrm{d}\sigma((p_1,p_2)|\theta) \tag{A.4}$$

for all $\sigma:\Theta\to\Delta([p^F(\theta),p^F(\theta)]^2)$ that satisfy the obedience constraints and $p_1$. The obedience constraints requires that given $p_2$, $p_1$ must be a best response for firm 1. Then, for all recommendation mechanism $\sigma$ that satisfy the obedience constraints, the RHS of (A.4) is given by

$$\sum_{\theta\in\Theta}\mu_\theta\int\Pi_1((p_1,p_2);\theta)\mathrm{d}\sigma((p_1,p_2)|\theta)=\sum_{\theta\in\Theta}\mu_\theta\int\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\left[\theta+bp_2-\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2}\right]\mathrm{d}\sigma(p_1,p_2|\theta)$$
$$=a\mathbb{E}_\mu\left[\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)^2\right]=a\mathbb{E}_\mu[p_1^\sigma(\theta)^2]$$

The first equality holds since $\Pi_1((p_1,p_2);\theta)=p_1(\theta+bp_2-ap_1)$ and since firm 1's best response, denoted by $p_1^\theta(\theta)$, is

$$p_1^\sigma(\theta)=\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}.$$

When firms offer substitutes, firm 1's expected equilibrium profit is an increasing and convex function of the expected equilibrium price $p_2$. Then, Jensen's inequality implies that maximizing expected equilibrium profits is equivalent to maximizing the distance between the expected equilibrium prices set by firm 2, $\mathbb{E}_\sigma[p_2|\theta]$ (or, equivantly, by maximizing the distance between $p_1^\sigma(\theta)$). When firms offer substitutes, Lemma A.3 shows that $\sup\sigma((p_1,p_2)|\theta)\in[p^F(\theta_L),p^F(\theta_H)]^2$ for all $\theta\in\Theta$. Hence, recommending $(p^F(\theta_L),p^F(\theta_L)))$ in the low state and $(p^F(\theta_H),p^F(\theta_H)))$ in the high state maximizes expected equilibrium profit which implies that full disclosure is optimal for the informed firm.[3] ■

### A.2.3 Consumer optimal disclosure: proofs

**Proof. Proposition II.2.** Lemma A.1 shows that full disclosure is never optimal for consumers. Consider instead any partial disclosure policy $\sigma$ and define $\sigma(s_2|\theta)$ the distribution of price recommendation $p_2$ conditional on the state $\theta$. Expected consumer surplus, denoted by $\mathbb{E}_{(\mu,\sigma)}[CS((p_1,p_2);\theta)]$, is

$$\mathbb{E}_{(\mu,\sigma)}[CS((p_1,p_2);\theta)]=\frac{1}{2a}\sum_{\theta\in\Theta}\mu_\theta\left[\int(\theta+bp_2-ap_1)^2\mathrm{d}\sigma((p_1,p_2)|\theta)\right]$$
$$+\frac{1}{2a}\sum_{\theta\in\Theta}\mu_\theta\left[\int(\theta+bp_1-ap_2)^2\mathrm{d}\sigma((p_1,p_2)|\theta)\right]\tag{A.5}$$

[3]That is, to maximize the expectation of a quadratic function in an interval, it is necessary to put all mass on the extremes of such interval.

where, in the unique BNE, $p_1$ satisfies

$$p_1 = \frac{1}{2a}\left[\theta + b\int p_2 \mathrm{d}\sigma(p_2|\theta)\right].$$

Substituting this expression in (A.5), expected consumer surplus can be written as

$$\begin{aligned}\mathbb{E}_{(\mu,\sigma)}[CS((p_1,p_2);\theta)] = &\frac{1}{2a}\mathbb{E}_\mu\left[\mathbb{E}_\sigma\left[\left(\frac{\theta}{2} + b\left(p_2 - \frac{1}{2}\mathbb{E}_\sigma[p_2|\theta]\right)\right)^2\Bigg|\theta\right]\right]\\ &+ \frac{1}{2a}\mathbb{E}_\mu\left[\mathbb{E}_\sigma\left[\left[\theta\left(1+\frac{b}{2a}\right) + \frac{b^2}{2a}\mathbb{E}_\sigma[p_2|\theta] - ap_2\right]^2\Bigg|\theta\right]\right]\end{aligned}$$

Define $\Delta\mathbb{E}[CS](\sigma)$ as the difference in expected consumer surplus with partial and no disclosure. This difference is given by

$$\begin{aligned}\Delta\mathbb{E}[CS](\sigma) = &\frac{a}{2}\left(\delta^2+1\right)\mathbb{V}_{(\mu,\sigma)}[p_2] - \left[\left(1-\frac{\delta^2}{2}\right)\left(\frac{\delta}{2}+1\right) - \frac{\delta}{4}\right]Cov_{(\mu,\sigma)}(\theta,p_2)\\ &- \frac{b\delta}{8}\left(7-\delta^2\right)\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]],\end{aligned}$$

where the equality holds by the law of iterated expectations, the definition of variance, conditional variance and covariance and the law of total variance. Hence, the difference in expected consumer surplus, $\Delta\mathbb{E}[CS](\sigma)$, is a continuous function of $\delta$. This difference is also a strictly increasing function of $\delta$. In particular, Lemma A.1 shows that $\Delta\mathbb{E}[CS](\sigma)$ converges to a positive number as $\delta \to 1$ which, in turn, implies that

$$b > \frac{2Cov_{(\mu,\sigma)}[\theta,p_2]}{4\mathbb{V}_{(\mu,\sigma)}[p_2] - 3\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]},$$

and this condition ensures that $\Delta\mathbb{E}[CS](\sigma)$ is a strictly increasing function of $\delta$.

First, if $\delta \to 0$, the expected consumer surplus with partial and no disclosure converge to

$$\Delta\mathbb{E}[CS](\sigma) \underset{\delta\to 0}{\to} \frac{a}{2}V_{(\mu,\sigma)}[p_2] - Cov_{(\mu,\sigma)}[\theta,\mathbb{E}[p_2|\theta]] = -Cov_{(\mu,\sigma)}\left[\theta - \frac{a}{2}p_2, p_2\right].$$

The equality holds by properties of covariance and since the covariance between $\theta$ and $\mathbb{E}_\sigma[p_2|\theta]$ equals the covariance between $\theta$ and $p_2$. The price $p_2$ is an increasing function of $\theta$ since the state is a positive demand shifter and

$$\frac{\partial p_2}{\partial \theta} \leq \frac{1}{2a-b} \leq \frac{2}{a}$$

since $a > b > 0$. Then, the covariance between $\theta - \frac{a}{2}p_2$ and $p_2$ is the covariance between two increasing functions of $\theta$. Hence, this covariance is positive, which implies that $\Delta\mathbb{E}[CS](\sigma)$ converges to a negative number when $\delta \to 0$.

Second, Lemma A.1 shows that partial disclosure yields higher expected consumer surplus than no disclosure if $\delta \in (\hat{c}, 1)$. That is, $\Delta\mathbb{E}[CS](\sigma) > 0$ for all $\delta \in (\hat{c}, 1)$. Hence, the Intermediate Value theorem implies that there exists $\hat{\alpha} \in (0, \hat{c}]$ such that $\Delta\mathbb{E}[CS](\sigma) = 0$ when $\delta = \hat{\alpha}$. Moreover, since $\Delta\mathbb{E}[CS](\sigma)$ is strictly increasing in $\delta$, partial disclosure is optimal for all $\delta \in (\hat{\alpha}, 1)$ where $\hat{\alpha} \in (0, \hat{c}]$ and no disclosure is optimal otherwise. ■

**Lemma A.4** *Assume that $\sigma$ is partially informative and $\sigma(p_2|\theta)$ is degenerated, placing all mass on $\hat{p} \in [p_L^F, p_H^F]$. For any obedient $\sigma$, supp $\sigma(p_2|\theta') = \{\hat{p}, \hat{p}'\}$ for all $\theta \neq \theta'$.*

**Proof. Lemma A.4.** The recommendation mechanism $\sigma$ is not fully informative. First, I show that $\hat{p} \in$ supp $\sigma(p_2|\theta')$. Suppose not. Then, supp $\sigma|\theta \cap$ supp $\sigma|\theta' = \emptyset$ which implies that price recommendations fully reveal the state. However, this contradicts the assumption that $\sigma$ is partially informative. Hence, $\hat{p} \in$ supp $\sigma(p_2|\theta')$.

Second, I show that the support of $\sigma(p_2|\theta')$ is binary. Firm $i$'s obedience constraint is

$$\sum_{\theta \in \Theta} \mu_\theta \int_{p_{-i} \in \mathbb{R}_+} \Pi_i((p_i, p_{-i}), \theta) \mathrm{d}\sigma((p_i, p_{-i})|\theta) \geq \sum_{\theta \in \Theta} \mu_\theta \int_{p_{-i} \in \mathbb{R}_+} \Pi_i((p_i', p_{-i}), \theta) \mathrm{d}\sigma((p_i, p_{-i})|\theta)$$

for all $i$, $p_i \in$ supp $\sigma$ and $p_i' \in [p_L^F, p_H^F]$. The left-hand side of the uninformed firm obedience constraint can be simplified as follows:

$$\begin{aligned}
&\sum_{\theta\in\Theta}\mu_\theta\int_{p_1\in[p_L^F,p_H^F]} p_2\left(\theta+bp_1-ap_2\right)\mathrm{d}\sigma((p_1,p_2)|\theta)\\
&=\sum_{\theta\in\Theta}\mu_\theta\int_{p_1\in[p_L^F,p_H^F]} p_2\left[\theta+b\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)-ap_2\right]\mathrm{d}\sigma((p_1,p_2)|\theta)\\
&=\sum_{\theta\in\Theta}\mu_\theta p_2\left[\theta+b\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)-ap_2\right]\int_{p_1\in[p_L^F,p_H^F]}\mathrm{d}\sigma((p_1,p_2)|\theta)\\
&=\sum_{\theta\in\Theta}\mu_\theta p_2\left[\theta+b\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)-ap_2\right]\sigma(p_2|\theta)
\end{aligned}$$

The first equality holds by the best response function of firm 1. The last equality holds since $\int_{p_1\in[p_L^F,p_H^F]}\mathrm{d}\sigma((p_1,p_2)|\theta)=\sigma(p_2|\theta)$. Hence, the uninformed firm obedience constraint is

$$\sum_{\theta\in\Theta}\mu_\theta p_2\left[\theta+b\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)-ap_2\right]\sigma(p_2|\theta)\geq\sum_{\theta\in\Theta}\mu_\theta p_2'\left[\theta+b\left(\frac{\theta+b\mathbb{E}_\sigma[p_2|\theta]}{2a}\right)-ap_2'\right]\sigma(p_2|\theta)$$

for all $p_2\in$ supp $\sigma(p_2|\theta)$ and $p_2'\in[p_L^F,p_H^F]$. The obedience constraint for $p_2=\hat{p}$ is

$$\begin{aligned}
&\mu_{\theta'}\sigma(\hat{p}|\theta')\hat{p}\left[\theta'+b\left(\frac{\theta'+b\mathbb{E}_\sigma[p_2|\theta']}{2a}\right)-a\hat{p}\right]+\mu_\theta\hat{p}\left[\theta+b\left(\frac{\theta+b\hat{p}}{2a}\right)-a\hat{p}\right]\\
&\geq\mu_{\theta'}\sigma(\hat{p}|\theta')p_2'\left[\theta'+b\left(\frac{\theta'+b\mathbb{E}_\sigma[p_2|\theta']}{2a}\right)-ap_2'\right]+\mu_\theta p_2'\left[\theta+b\left(\frac{\theta+b\hat{p}}{2a}\right)-ap_2'\right]
\end{aligned}$$

for all $p_2'\in[p_L^F,p_H^F]$. Similarly, the obedience constraint of $\sigma$ for $p_2\neq\hat{p}$ is

$$p_2\left[\theta'+b\left(\frac{\theta'+b\mathbb{E}_\sigma[p_2|\theta']}{2a}\right)-ap_2\right]\geq p_2'\left[\theta'+b\left(\frac{\theta'+b\mathbb{E}_\sigma[p_2|\theta']}{2a}\right)-ap_2'\right]\tag{A.6}$$

for all $p_2'\in[p_L^F,p_H^F]$. The uninformed firm's profits are strictly concave in $p_2$ which implies there exists a unique $\hat{p}'\in[p_L^F,p_H^F]$ that satisfies (A.6) and $\hat{p}'\neq\hat{p}$. Hence, the support of $\hat{\sigma}|\theta'$ is binary and given by $\{\hat{p},\hat{p}'\}$. ■

**Lemma A.5** *Assume that $\sigma$ is partially informative and $\sigma(p_2|\theta_H)$ is degenerated, placing all mass on $\hat{p}_H\in[p_L^F,p_H^F]$. For any obedient $\sigma$, supp $\sigma(p_2|\theta_L)=\{\hat{p}_L,\hat{p}_H\}$*

*where* $\lambda = \sigma(\hat{p}_L|\theta_L)$,

$$\hat{p}_L = \frac{4a^2[1-\mu_L\lambda]\theta_L + b^2 mu_H[(1-\lambda)\theta_H - \theta_L]}{(2a-b)\left[4a^2(1-\mu_L\lambda) - b^2\mu_H\lambda\right]} \quad \textit{and}$$
$$\hat{p}_H = \frac{4a^2\left[\mu_H\theta_H + \mu_L(1-\lambda)\theta_L\right] - b^2\mu_H\lambda\theta_H}{(2a-b)\left[4a^2(1-\mu_L\lambda) - b^2\mu_H\lambda\right]}$$

**Proof. Lemma A.5.** Lemma A.4 implies that the support of $\sigma(p_2|\theta_L)$ is binary and given by $\{\hat{p}_L, \hat{p}_H\}$ if the support of $\sigma(p_2|\theta_H)$ is degenerated and given by $\hat{p}_H$. Define $\sigma(\hat{p}_L|\theta_L) = 1 - \sigma(\hat{p}_H|\theta_L) := \lambda \in (0,1)$. By definition,

$$\mathbb{E}_\sigma[p_2|\theta_L] = \lambda\hat{p}_L + (1-\lambda)\hat{p}_H \text{ and } \mathbb{E}_\sigma[p_2|\theta_H] = \hat{p}_H.$$

Then, taking $\mathbb{E}_\sigma[p_2|\theta_L]$ and $\mathbb{E}_\sigma[p_2|\theta_H]$ as given, $\hat{p}_L$ and $\hat{p}_H$ are characterized by

$$\hat{p}_L = \arg\max_{p_2} p_2\left[\theta_L + b\left(\frac{\theta_L + b\mathbb{E}_\sigma[p_2|\theta_L]}{2a}\right) - ap_2\right]$$
$$\hat{p}_H = \arg\max_{p_2} \mu_L(1-\lambda)p_2\left[\theta_L + b\left(\frac{\theta_L + b\mathbb{E}_\sigma[p_2|\theta_L]}{2a}\right) - ap_2\right] + \mu_H p_2\left[\theta_H + b\left(\frac{\theta_H + b\mathbb{E}_\sigma[p_2|\theta_H]}{2a}\right) - ap_2\right]$$

The first order conditions of the previous maximization problems are

$$\hat{p}_L = \frac{1}{2a}\left[\theta_L + \frac{b}{2a}\left(\theta_L + b\mathbb{E}_\sigma[p_2|\theta_L]\right)\right]$$
$$\hat{p}_H = \frac{\mu_L(1-\lambda)\left[\theta_L + \frac{b}{2a}\left(\theta_L + b\mathbb{E}_\sigma[p_2|\theta_L]\right)\right] + \mu_H\left[\theta_H + \frac{b}{2a}\left(\theta_H + b\mathbb{E}_\sigma[p_2|\theta_H]\right)\right]}{2a\left(\mu_L(1-\lambda) + \mu_H\right)}$$

Using the definition of $\mathbb{E}_\sigma[p_2|\theta_L]$ and $\mathbb{E}_\sigma[p_2|\theta_H]$, we have that $\hat{p}_L$ and $\hat{p}_H$ are given by

$$\hat{p}_L = \frac{4a^2[1-\mu_L\lambda]\theta_L + b^2\mu_H[(1-\lambda)\theta_H - \theta_L]}{(2a-b)\left[4a^2(1-\mu_L\lambda) - b^2\mu_H\lambda\right]} \quad \text{and}$$
$$\hat{p}_H = \frac{4a^2\left[\mu_H\theta_H + \mu_L(1-\lambda)\theta_L\right] - b^2\mu_H\lambda\theta_H}{(2a-b)\left[4a^2(1-\mu_L\lambda) - b^2\mu_H\lambda\right]}$$

where $\lambda$ fully characterizes $\sigma$. ■

**Proof. Proposition II.3.** This proof applies to a more general result which states that is optimal for the designer to select $\sigma(p_2|\theta)$ to be degenerated for any $\theta$. Here I present the proof for $\sigma(p_2|\theta_H)$ but the proof for the other case is analogous.

Suppose not. Assume that the optimal recommendation mechanism $\sigma^* = \{\sigma^*(p_2|\theta)\}_{\theta\in\Theta}$

is partially informative where both $\sigma^*(p_2|\theta)$ are not degenerated. Consider an alternative partially informative recommendation mechanism $\hat{\sigma}$ in which $\hat{\sigma}(p_2|\theta_H)$ is degenerated and places all its mass on one point $\hat{p}_H \in [p^F(\theta_L), p^F(\theta_H)] = [p_L^F, p_H^F]$ where $\hat{p}_H \in \text{supp } \hat{\sigma}(p_2|\theta_L)$. By Lemma A.5, for any obedient $\hat{\sigma}$, the support of $\hat{\sigma}|\theta_L$ is $\{\hat{p}_L, \hat{p}_H\}$ where $\hat{p}_L$ and $\hat{p}_H$ are defined in Lemma A.4 and $\lambda = \hat{\sigma}(\hat{p}_L|\theta_L)$ fully characterizes $\hat{\sigma}$. Next, I show that there exists $\lambda \in (0,1)$ such that $\Delta\mathbb{E}[CS](\hat{\sigma}) \geq \Delta\mathbb{E}[CS](\sigma^*)$. Given that $\mathbb{E}_\sigma[p_2] = \mathbb{E}_{\sigma'}[p_2]$ for all feasible $\sigma, \sigma'$,[4] the difference between $\Delta\mathbb{E}[CS](\hat{\sigma})$ and $\Delta\mathbb{E}[CS](\sigma^*)$, denoted as $\Delta\mathbb{E}[CS]_{\hat{\sigma}-\sigma^*}$, is

$$\begin{aligned}\Delta\mathbb{E}[CS]_{\hat{\sigma}-\sigma^*} = &\frac{a}{2}\left(1+\delta^2\right)\left(\mathbb{E}_{\hat{\sigma}}[p_2^2] - \mathbb{E}_{\sigma^*}[p_2^2]\right) - \left[\left(1-\frac{\delta^2}{2}\right)\left(1+\frac{\delta}{2}\right) - \frac{\delta}{4}\right]\left(\mathbb{E}_{\hat{\sigma}}[\theta \cdot p_2] - \mathbb{E}_{\sigma^*}[\theta \cdot p_2]\right) \\ &- \frac{b\delta}{8}(7-\delta^2)\left(\mathbb{E}_{\hat{\sigma}}[\mathbb{E}[p_2|\theta]^2] - \mathbb{E}_{\sigma^*}[\mathbb{E}[p_2|\theta]^2]\right)\end{aligned}$$

For any feasible $\sigma^*$, the expectation $\mathbb{E}_{\sigma^*}[p_2|\theta_L]$ satisfies

$$\mathbb{E}_{\sigma^*}[p_2|\theta_L] \in \left(\frac{\theta_L}{2a-b}, \frac{\mathbb{E}_\mu[\theta]}{2a-b}\right).$$

Moreover, by definition, $\mathbb{E}_{\hat{\sigma}}[p_2|\theta_L] = \lambda\hat{p}_L + (1-\lambda)\hat{p}_H$, and

$$\mathbb{E}_{\hat{\sigma}}[p_2|\theta_L] = \frac{\theta_L}{2a-b} \text{ if } \lambda = 1 \text{ and } \mathbb{E}_{\hat{\sigma}}[p_2|\theta_L] = \frac{\mathbb{E}[\theta]}{2a-b} \text{ if } \lambda = 0.$$

The intermediate value theorem implies that there exists $\tilde{\lambda} \in (0,1)$ such that $\mathbb{E}_{\hat{\sigma}}[p_2|\theta_L]$ equals $\mathbb{E}_{\sigma^*}[p_2|\theta_L]$ since $\mathbb{E}_{\hat{\sigma}}[p_2|\theta_L]$ is a continuous function of $\lambda$. Since $\mathbb{E}_\sigma[p_2] = \mathbb{E}_{\sigma'}[p_2]$ for all feasible $\sigma$ and $\sigma'$, $\tilde{\lambda}$ also satisfies $\mathbb{E}_{\hat{\sigma}}[p_2|\theta_H] = \mathbb{E}_{\sigma^*}[p_2|\theta_H]$. Then, the difference between $\Delta\mathbb{E}[CS](\hat{\sigma})$ and $\Delta\mathbb{E}[CS](\sigma^*)$ for $\hat{\sigma}$ characterized by $\tilde{\lambda}$ is

[4] Note that $\mathbb{E}_\pi[p_2] = \mathbb{E}_{\pi'}[p_2]$ for all feasible $\pi, \pi'$ since

$$\mathbb{E}_\pi[p_2] = \frac{1}{2a}\left[\mathbb{E}[\theta]\left(1+\frac{b}{2a}\right) + \frac{b^2}{2a}\mathbb{E}_\pi[p_2]\right] \Leftrightarrow \mathbb{E}_\pi[p_2] = \frac{\mathbb{E}_\mu[\theta]}{2a-b}.$$

The equality holds by the uninformed firm's and informed firm's best response functions and by the law of iterated expectations. Then, $\mathbb{E}_\pi[p_2]$ doesn't depend on $\pi$. Given the equivalence between $\pi_2$ and $\sigma$, it also follows that $\mathbb{E}_\sigma[p_2] = \mathbb{E}_{\sigma'}[p_2]$ for all feasible $\sigma, \sigma'$.

$$\Delta\mathbb{E}[CS]_{\hat{\sigma}-\sigma^*} = \frac{a}{2}\left(1+\delta^2\right)\left[\mathbb{E}_{\hat{\sigma}}[p_2^2] - \mathbb{E}_{\sigma^*}[p_2^2]\right]$$
$$= \frac{a}{2}\left(1+\delta^2\right)\left[\mu_L\left(\tilde{\lambda}\hat{p}_L^2 + (1-\tilde{\lambda})\hat{p}_H^2 - \mathbb{E}_{\sigma^*}[p_2^2|\theta_L]\right) + \mu_H\left(\hat{p}_H^2 - \mathbb{E}_{\sigma^*}[p_2^2|\theta_H]\right)\right]$$

Hence, $\mathbb{E}_{\hat{\sigma}}[p_2^2] \geq \mathbb{E}_{\sigma^*}[p_2^2]$ by Jensen's inequality. Then, for all demand parameters and $\sigma^*$ such that $\delta \leq \hat{c}$, there exists $\lambda \in (0,1)$ such that $\Delta\mathbb{E}[CS]_{\hat{\sigma}-\sigma^*} \geq 0$. This contradicts the optimality of $\sigma^*$. Thus, the optimal partially informative recommendation mechanism is such that supp $\sigma|\theta_H = \{\hat{p}_H\}$ and supp $\sigma|\theta_L = \{\hat{p}_L, \hat{p}_H\}$.

Lastly, the optimal recommendation mechanism is characterized by

$$\lambda^* \in \arg\max_{\lambda\in[0,1]} \Delta\mathbb{E}[CS](\lambda)$$

where

$$\begin{aligned}\Delta\mathbb{E}[CS](\lambda) = {} & \frac{a}{2}\left(\delta^2+1\right)\mu_L\lambda\left[\mu_L(1-\lambda)+\mu_H\right]\left(\hat{p}_H-\hat{p}_L\right)^2 \\ & - \left[\left(1-\frac{\delta^2}{2}\right)\left(\frac{\delta}{2}+1\right) - \frac{\delta}{4}\right]\left[\mu_L\theta_L\left[\lambda\hat{p}_L + (1-\lambda)\hat{p}_H\right] + \mu_H\theta_H\hat{p}_H - \frac{\left(\mu_L\theta_L+\mu_H\theta_H\right)^2}{2a-b}\right] \\ & - \frac{b\delta}{8}\left(7-\delta^2\right)\left[\mu_L\left[\lambda\hat{p}_L+(1-\lambda)\hat{p}_H\right]^2 + \mu_H\hat{p}_H^2 - \left[\mu_L\lambda\hat{p}_L + (1-\mu_L\lambda)\hat{p}_H\right]^2\right]\end{aligned}$$

and $\hat{p}_L$ and $\hat{p}_H$ are functions of $\lambda$ defined in Lemma A.4. The optimal $\lambda^* \in (0,1)$ is characterized by the first order condition of $\Delta\mathbb{E}[CS](\lambda)$ and it is given by

$$\lambda^* = \frac{4\left[\delta(1-3\delta^2)+6(1-\delta^2)\right]}{\mu_H\delta^5 + 2\mu_H\delta^4 - (12-\mu_H)\delta^3 - 6(4-\mu_H)\delta^2 + 4(1-\mu_H)\delta + 24(1-\mu_H)}.$$

■

### A.2.4 Producer surplus optimal disclosure: proofs

**Proof. Lemma II.5.** For the informed firm, the difference in expected profits with full disclosure $\sigma^F$ and any disclosure $\sigma$ is

$$\begin{aligned}\Pi_2(\sigma^F) - \Pi_2(\sigma) = {} & a\mathbb{E}_\mu\left[\left(\frac{\theta}{2a-b}\right)^2 - 2\frac{\theta}{2a-b}\mathbb{E}_\sigma[p_2|\theta] + \mathbb{E}_\sigma[p_2^2|\theta]\right] \\ & + \frac{b^2}{2a}\mathbb{E}_\mu\left[\frac{\theta}{2a-b}\mathbb{E}_\sigma[p_2|\theta] - \mathbb{E}_\sigma[p_2|\theta]^2\right]\end{aligned}$$

Then, this difference is positive since

$$\Pi_2(\sigma^F) - \Pi_2(\sigma) \geq a\mathbb{E}_\mu\left[\left(\frac{\theta}{2a-b} - \mathbb{E}_\sigma[p_2|\theta]\right)^2\right] + \frac{b^2}{2a}\sum_{\theta\in\theta}\mu_\theta\left[\mathbb{E}_\sigma[p_2|\theta]\left(\frac{\theta}{2a-b} - \mathbb{E}_\sigma[p_2|\theta]\right)\right]$$
$$\geq a\mathbb{E}_\mu\left[\left(\frac{\theta}{2a-b} - \mathbb{E}_\sigma[p_2|\theta]\right)^2\right] \geq 0$$

The first inequality holds by Jensen's inequality. The second since $a > |b| > 0$,

$$\frac{\theta_L}{2a-b} \leq \mathbb{E}_\sigma[p_2|\theta_L] \leq \mathbb{E}_\sigma[p_2|\theta_H] \leq \frac{\theta_H}{2a-b} \text{ and } \sum_{\theta\in\Theta}\mu_\theta\mathbb{E}_\sigma[p_2|\theta] = \frac{\mathbb{E}_\mu[\theta]}{2a-b}$$

for all feasible $\sigma$. Hence, $\Pi_2(\sigma^F) \geq \Pi_2(\sigma)$ which implies that full disclosure is optimal for the uninformed firm. ■

### A.2.5 Welfare optimal disclosure: proofs

**Proof. Proposition II.4.** First, I show that full disclosure yields a higher expected welfare than no disclosure if $\delta \geq \tilde{\alpha}$. First, full disclosure $\sigma^F$ yields a higher expected welfare than no disclosure $\sigma^N$ if and only if $\delta \geq \tilde{\alpha}$. Hence, if $\delta \geq \tilde{\alpha}$, either full or partial disclosure is optimal whereas if $\delta < \tilde{\alpha}$, either no or partial disclosure is optimal.

Second, consider $\delta < \tilde{\alpha}$. The difference between the total expected surplus with partial disclosure $\sigma$ and no disclosure $\sigma^N$ is given by

$$\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] - \mathbb{E}_{(\mu,\sigma^N)}\left[W((p_1,p_2);\theta)\right]$$
$$= \frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right]Cov_{(\mu,\sigma)}[\theta,p_2] - \frac{a}{2}\left(1-\delta^2\right)\mathbb{V}_{(\mu,\sigma)}[p_2] - \frac{b}{8}\delta\left(1-\delta^2\right)\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$$

This difference is a continuous and strictly increasing function of $\delta$. Moreover, as $\delta$ converges to 0, the difference in expected consumer surplus converges to

$$\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] - \mathbb{E}_{(\mu,\sigma^N)}\left[W((p_1,p_2);\theta)\right] \underset{\delta\to 0}{\to} -\frac{a}{2}\mathbb{V}_{(\mu,\sigma)}[p_2] < 0$$

and Lemma A.2 shows that $\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] > \mathbb{E}_{(\mu,\sigma^N)}\left[W((p_1,p_2);\theta)\right]$ for all $\delta > \tilde{c}_1$. Then, the intermediate value theorem implies that there exists a $\tilde{\alpha}_1 \in (0, \tilde{c}_1]$

such that

$$\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] = \mathbb{E}_{(\mu,\sigma^N)}\left[W((p_1,p_2);\theta)\right].$$

Also, since this difference is strictly increasing in $\delta$, this also implies that

$$\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] > \mathbb{E}_{(\mu,\sigma^N)}\left[W((p_1,p_2);\theta)\right] \text{ for all } \delta > \tilde{\alpha}_1 \text{ and}$$

$$\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] < \mathbb{E}_{(\mu,\sigma^N)}\left[W((p_1,p_2);\theta)\right] \text{ for all } \delta < \tilde{\alpha}_1.$$

That is, partial disclosure is welfare if $\delta \in [\tilde{\alpha}_1, \tilde{\alpha})$ and no disclosure is welfare optimal if $\delta < \tilde{\alpha}_1$.

Now, consider $\delta \geq \tilde{\alpha}$. The difference between expected welfare with full disclosure $\mathbb{E}_{(\mu,\sigma^F)}\left[W((p_1,p_2);\theta)\right]$ and with partial disclosure $\mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right]$ is given by

$$\begin{aligned}
&\mathbb{E}_{(\mu,\sigma^F)}\left[W((p_1,p_2);\theta)\right] - \mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] \\
&= \frac{b\delta}{2}\left(\delta^2 - \frac{3}{2}\right)\mathbb{E}_\mu\left[\left(\frac{\theta}{2a-b}\right)^2 - \mathbb{E}_\sigma[p_2|\theta]^2\right] - \frac{b\delta}{2}\mathbb{E}_\mu\left[\mathbb{E}_\sigma[p_2^2|\theta] - \mathbb{E}_\sigma[p_2|\theta]^2\right] \\
&+ \delta\left(\frac{3}{4} + \frac{\delta}{2} + \frac{\delta^2}{4}\right)\mathbb{E}_\mu\left[\theta\left(\frac{\theta}{2a-b} - \mathbb{E}_\sigma[p_2|\theta]\right)\right] - \frac{a}{2}\mathbb{E}_\mu\left[\left(\frac{\theta}{2a-b}\right)^2 - \mathbb{E}_\sigma[p_2^2|\theta]\right],
\end{aligned}$$

which is a continuous function of $\delta$.

First, as $\delta \to 1$, this difference converges to

$$\mathbb{E}_{(\mu,\sigma^F)}\left[W((p_1,p_2);\theta)\right] - \mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] \underset{\delta\to 1}{\to} \frac{3}{4a}\mathbb{E}_\mu\left[\left(\frac{\theta}{a} - \mathbb{E}_\sigma[p_2|\theta]\right)^2\right] > 0$$

Second, Lemma A.2 shows that partial disclosure yields higher expected welfare than full disclosure if $\delta < \tilde{c}_2$. Then, the intermediate value theorem implies that there exists $\tilde{\alpha}_2 \in [\tilde{c}_2, 1)$ such that

$$\mathbb{E}_{(\mu,\sigma^F)}\left[W((p_1,p_2);\theta)\right] = \mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right].$$

Analogously as before, $\mathbb{E}_{(\mu,\sigma^F)}\left[W((p_1,p_2);\theta)\right] - \mathbb{E}_{(\mu,\sigma)}\left[W((p_1,p_2);\theta)\right] > 0$ for $\delta > \tilde{\alpha}_2$

and negative for $\delta \in [\alpha, \tilde{\alpha}_2)$. Hence, full disclosure is welfare optimal if $\delta > \tilde{\alpha}_2$ and partial disclosure is optimal if $\delta \in [\alpha, \tilde{\alpha}_2)$. In summary, full disclosure is welfare optimal if $\delta \geq \tilde{\alpha}_2$, partial disclosure is welfare optimal if $\delta \in (\tilde{\alpha}_1, \tilde{\alpha}_2)$ and no disclosure is welfare optimal if $\delta \in (0, \tilde{\alpha}_1]$. ■

**Proof. Proposition II.5.** The proof is analogous to the proof of Proposition II.3. Suppose not. Assume that the optimal recommendation mechanism $\sigma^* = \{\sigma^*(p_2|\theta)\}_{\theta\in\Theta}$ is partially informative where both $\sigma^*(p_2|\theta)$ are not degenerated. Consider an alternative partially informative recommendation mechanism $\hat{\sigma}$ in which $\hat{\sigma}(p_2|\theta_H)$ is degenerated and places all its mass on one point $\hat{p}_H \in [p^F(\theta_L), p^F(\theta_H)] = [p^F_L, p^F_H]$ where $\hat{p}_H \in \text{supp } \hat{\sigma}(p_2|\theta_L)$. By Lemma A.5, for any obedient $\hat{\sigma}$, the support of $\hat{\sigma}|\theta_L$ is $\{\hat{p}_L, \hat{p}_H\}$ where $\hat{p}_L$ and $\hat{p}_H$ are defined in Lemma A.4 and $\lambda = \hat{\sigma}(\hat{p}_L|\theta_L)$ fully characterizes $\hat{\sigma}$.

Consider first the case in which $\delta < \widetilde{\alpha}$. Next, I show that there exists $\lambda \in (0, 1)$ such that $\mathbb{E}[W](\hat{\sigma}) - \mathbb{E}[W](\sigma^N) \geq \mathbb{E}[W](\sigma^*) - \mathbb{E}[W](\sigma^N)$ where

$$\begin{aligned}\mathbb{E}[W](\sigma) - \mathbb{E}[W](\sigma^N) = \frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right]Cov_{(\mu,\sigma)}[\theta, p_2] - \frac{a}{2}\left(1-\delta^2\right)\mathbb{V}_{(\mu,\sigma)}[p_2]\\ - \frac{b}{8}\delta\left(1-\delta^2\right)\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]\end{aligned}$$

The difference between $\mathbb{E}[W](\hat{\sigma}) - \mathbb{E}[W](\sigma^N)$ and $\mathbb{E}[W](\sigma^*) - \mathbb{E}[W](\sigma^N)$ is

$$\begin{aligned}\Delta\mathbb{E}[W]^N_{\hat{\sigma}-\sigma^*} \geq \frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right]\left(\mathbb{E}_{\hat{\sigma}}[\theta\cdot p_2] - \mathbb{E}_{\sigma^*}[\theta\cdot p_2]\right)\\ - \left(\frac{a}{2}+\frac{b}{8}\delta\right)\left(1-\delta^2\right)\left(\mathbb{E}_{\hat{\sigma}}[p_2^2] - \mathbb{E}_{\sigma^*}[\mathbb{E}[p_2|\theta]^2]\right)\end{aligned}$$

where the inequality holds since $\mathbb{E}_\sigma[p_2] = \mathbb{E}_{\sigma'}[p_2]$ for all feasible $\sigma, \sigma'$ and $\mathbb{E}_\sigma[p_2^2] \geq \mathbb{E}_\sigma[\mathbb{E}[p_2|\theta]^2]$ for all $\sigma$. Note that $\mathbb{E}_{\hat{\sigma}}[p_2^2]$ is a continuous function of $\lambda$ and for any $\sigma^*$,

$$\mathbb{E}_{\sigma^*}[\mathbb{E}[p_2|\theta]^2] \in \left(\left(\frac{\mathbb{E}[\theta]}{2a-b}\right)^2, \mathbb{E}_{\sigma^*}[p_2^2]\right) \subset \left(\left(\frac{\mathbb{E}[\theta]}{2a-b}\right)^2, \mu_L\left(\frac{\theta_L}{2a-b}\right)^2 + \mu_H\left(\frac{\theta_H}{2a-b}\right)^2\right)$$

$$\mathbb{E}_{\hat{\sigma}}[p_2^2] = \left(\frac{\mathbb{E}[\theta]}{2a-b}\right)^2 \text{ if } \lambda = 0 \text{ and } \mathbb{E}_{\hat{\sigma}}[p_2^2] = \mu_L\left(\frac{\theta_L}{2a-b}\right)^2 + \mu_H\left(\frac{\theta_H}{2a-b}\right)^2 \text{ if } \lambda = 1$$

Hence, the intermediate value theorem implies that there exists $\hat{\lambda} \in (0, 1)$ such that $\mathbb{E}_{\hat{\sigma}}[p_2^2] = \mathbb{E}_{\sigma^*}[\mathbb{E}[p_2|\theta]^2]$. Then, the difference between $\mathbb{E}[W](\hat{\sigma}) - \mathbb{E}[W](\sigma^N)$ and $\mathbb{E}[W](\sigma^*) - \mathbb{E}[W](\sigma^N)$ for $\hat{\lambda}$ satisfies

$$\begin{aligned}\Delta\mathbb{E}[W]^N_{\hat{\sigma}-\sigma^*} &\geq \frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right](\mathbb{E}_{\hat{\sigma}}[\theta \cdot p_2] - \mathbb{E}_{\sigma^*}[\theta \cdot p_2]) \\ &\geq \frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right]\left(\theta_L \mathbf{1}_{(\mathbb{E}_{\hat{\sigma}}[p_2|\theta] \geq \mathbb{E}_{\sigma^*}[p_2|\theta])} + \theta_H \mathbf{1}_{(\mathbb{E}_{\hat{\sigma}}[p_2|\theta] < \mathbb{E}_{\sigma^*}[p_2|\theta])}\right)[\mathbb{E}_{\hat{\sigma}}[p_2] - \mathbb{E}_{\sigma^*}[p_2]]\end{aligned}$$

where $\mathbb{E}_{\hat{\sigma}}[p_2] = \mathbb{E}_{\sigma^*}[p_2]$, contradicting the optimality of $\sigma^*$. Thus, the optimal disclosure has binary support and it is characterized by $\lambda^* \in \arg\max_{\lambda \in [0,1]} \Delta\mathbb{E}[W]^N(\lambda)$ where

$$\begin{aligned}\Delta\mathbb{E}[W]^N(\lambda) = &\frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right]\left[\mu_L\theta_L\left[\lambda\hat{p}_L + (1-\lambda)\hat{p}_H\right] + \mu_H\theta_H\hat{p}_H - \frac{(\mu_L\theta_L + \mu_H\theta_H)^2}{2a-b}\right] \\ &- \frac{a}{2}\left(1-\delta^2\right)\mu_L\lambda\left[\mu_L(1-\lambda)+\mu_H\right](\hat{p}_H - \hat{p}_L)^2 \\ &- \frac{b}{8}\delta\left(1-\delta^2\right)\left[\mu_L\left[\lambda\hat{p}_L + (1-\lambda)\hat{p}_H\right]^2 + \mu_H\hat{p}_H^2 - \left[\mu_L\lambda\hat{p}_L + (1-\mu_L\lambda)\hat{p}_H\right]^2\right]\end{aligned}$$

The optimal $\lambda^* \in (0,1)$ is characterized by the first order condition of $\Delta\mathbb{E}[W]^N(\lambda)$.

Consider now the case in which $\delta \geq \widetilde{\alpha}$. Next, I show that

$$\mathbb{E}[W](\hat{\sigma}) - \mathbb{E}[W](\sigma^F) \geq \mathbb{E}[W](\sigma^*) - \mathbb{E}[W](\sigma^F).$$

The difference between $\mathbb{E}[W](\hat{\sigma}) - \mathbb{E}[W](\sigma^F)$ and $\mathbb{E}[W](\sigma^*) - \mathbb{E}[W](\sigma^F)$, denoted by $\Delta\mathbb{E}[W]^F_{\hat{\sigma}-\sigma^*}$, is

$$\begin{aligned}\Delta\mathbb{E}[W]^F_{\hat{\sigma}-\sigma^*} = &\frac{\delta}{2}\left[\left(\frac{\delta}{2}+1\right)\delta+\frac{3}{2}\right](\mathbb{E}_{\hat{\sigma}}[\theta \cdot p_2] - \mathbb{E}_{\sigma^*}[\theta \cdot p_2]) - \frac{a}{2}\left(1-\delta^2\right)\left(\mathbb{E}_{\hat{\sigma}}[p_2^2] - \mathbb{E}_{\sigma^*}[p_2^2]\right) \\ &- \frac{b\delta}{2}\left(\frac{5}{2}-\delta^2\right)\left(\mathbb{E}_{\hat{\sigma}}[\mathbb{E}[p_2|\theta]^2] - \mathbb{E}_{\sigma^*}[\mathbb{E}[p_2|\theta]^2]\right)\end{aligned}$$

Analogously as before, there exists $\hat{\lambda} \in (0,1)$ such that $\mathbb{E}_{\hat{\sigma}}[p_2^2] = \mathbb{E}_{\sigma^*}[\mathbb{E}[p_2|\theta]^2]$ and the difference between $\mathbb{E}[W](\hat{\sigma}) - \mathbb{E}[W](\sigma^F)$ and $\mathbb{E}[W](\sigma^*) - \mathbb{E}[W](\sigma^F)$ for $\hat{\lambda}$ satisfies

$$\Delta\mathbb{E}[W]^F_{\hat{\sigma}-\sigma^*} \geq \frac{\delta}{2}\left[\left(\frac{\delta}{2}+1\right)\delta+\frac{3}{2}\right](\mathbb{E}_{\hat{\sigma}}[\theta \cdot p_2] - \mathbb{E}_{\sigma^*}[\theta \cdot p_2])$$

and since $\mathbb{E}_{\hat{\sigma}}[\theta \cdot p_2] \geq \mathbb{E}_{\sigma^*}[\theta \cdot p_2]$, this contradicts the optimality of $\sigma^*$. Thus, the optimal partially informative recommendation mechanism is such that supp $\sigma|\theta_H = \{\hat{p}_H\}$ and supp $\sigma|\theta_L = \{\hat{p}_L, \hat{p}_H\}$. The optimal disclosure is characterized by $\lambda^* \in \arg\max_{\lambda \in [0,1]} \Delta\mathbb{E}[W]^F(\lambda)$ where

$$\begin{aligned}\Delta\mathbb{E}[W]^F(\lambda) = \frac{\delta}{2}\left[\delta\left(\frac{\delta}{2}+1\right)+\frac{3}{2}\right]&\left[\mu_L\theta_L\left[\lambda\hat{p}_L+(1-\lambda)\hat{p}_H\right]+\mu_H\theta_H\hat{p}_H-\frac{(\mu_L\theta_L+\mu_H\theta_H)^2}{2a-b}\right]\\ &-\frac{a}{2}\left(1-\delta^2\right)\mu_L\lambda\left[\mu_L(1-\lambda)+\mu_H\right](\hat{p}_H-\hat{p}_L)^2\\ &-\frac{b}{2}\delta\left(\frac{5}{2}-\delta^2\right)\left[\mu_L\left[\lambda\hat{p}_L+(1-\lambda)\hat{p}_H\right]^2+\mu_H\hat{p}_H^2-\left[\mu_L\lambda\hat{p}_L+(1-\mu_L\lambda)\hat{p}_H\right]^2\right]\end{aligned}$$

The optimal $\lambda^* \in (0,1)$ is characterized by the first order condition of $\Delta\mathbb{E}[W]^F(\lambda)$.
■

## A.3 Non-linear demand

Consider the same environment as before but assume firm $i$'s demand, $q(p_i, p_{-i}; \theta)$, is continuous and differentiable and satisfies the following properties:

$$i)\ \frac{\partial q(p_i, p_{-i};\theta)}{\partial p_i} \leq 0,\ \ ii)\quad \frac{\partial q(p_i, p_{-i};\theta)}{\partial \theta} > 0,\ \text{and } iii)\ \left|\frac{\partial q(p_i, p_{-i};\theta)}{\partial p_i}\right| > \left|\frac{\partial q(p_i, p_{-i};\theta)}{\partial p_{-i}}\right|$$

The first condition ensures that quantity demanded decreases as price increases, the second condition implies that the state is a positive demand shifter and, lastly, the third condition implies that goods are differentiated and that a change of its own price has a bigger effect on the demand than a change of the price of a competitor.[5] Assume that firm's ex-post profits are strictly concave in $p_i$. That is,

$$p_i\frac{\partial^2 q(p_i, p_{-i};\theta)}{\partial p_i^2} < -2\frac{\partial q(p_i, p_{-i};\theta)}{\partial p_i} \text{ for all } p_i.$$

Furthermore, assume that

$$\frac{\partial^2 \Pi_i(p_i, p_{-i};\theta)}{\partial p_i^2}\frac{\partial^2 \Pi_{-i}(p_{-i}, p_i;\theta)}{\partial p_{-i}^2} \geq \left(\frac{\partial^2 \Pi_i(p_i, p_{-i};\theta)}{\partial p_i \partial p_{-i}}\right)^2.$$

[5]This ensures that equilibrium prices are finite.

Firms offer substitutes (complements) if

$$\frac{\partial q(p_i, p_{-i}; \theta)}{\partial p_{-i}} > 0 (< 0).$$

When firms offer substitutes, assume that the elasticity of demand of firm $i$ is a non-increasing function of the other firm' price and that the demand is supermodular in the state $\theta$ and the price of the other firm $p_{-i}$, i.e.,

$$\frac{\partial^2 q(p_i, p_{-i}; \theta)}{\partial p_i \partial p_{-i}} \geq 0 \text{ for all } (p_i, p_{-i}) \text{ and } \frac{\partial^2 q(p_i, p_{-i}; \theta)}{\partial \theta \partial p_{-i}} \geq 0.$$

Similarly, when firms offer complements, assume that the elasticity of demand of firm $i$ is a non-decreasing function of the other firm' price and that the demand is submodular in the state $\theta$ and the price of the other firm. That is,

$$\frac{\partial^2 q(p_i, p_{-i}; \theta)}{\partial p_i \partial p_{-i}} \leq 0 \text{ for all } (p_i, p_{-i}) \text{ and } \frac{\partial^2 q(p_i, p_{-i}; \theta)}{\partial \theta \partial p_{-i}} \leq 0.$$

Note that these assumptions imply that firms choices are strategic complements (substitutes) when they offer substitutes (complements).

**Pricing game equilibrium.** For all information structures $(S_2, \pi_2)$, the existence and uniqueness of the BNE is guaranteed by Ui (2016), which provides sufficient conditions for the existence and uniqueness of BNE in Bayesian games with concave and continuously differentiable payoff functions. This is formalized in Lemma A.6.

**Lemma A.6** *For all information structures $(S_2, \pi_2)$, the set of Bayesian Nash equilibria in the pricing game $\hat{\mathcal{E}}(S_2, \pi_2)$ is a singleton.*

**Simplifications.** The strict concavity of firm's ex-post profits in $p_i$ imply that firms' profits are bounded and continuous functions and that there exists $\overline{p}$ such that it is without loss of generality to restrict attention to the compact action space $p_i \in [0, \overline{p}]$. The equivalence to recommendation mechanism $\sigma$ is established in Lemma II.2 and Lemma II.3. The existence and uniqueness of BNE imply that it is sufficient to restrict attention to the distribution $\sigma(p_2|\theta)$ since for any obedient recommendation mechanism there exists a function $p_1(\theta, \sigma(p_2|\theta))$ which represents firm 1's best response when the state is $\theta$ and the price recommendations are given by $\sigma$ where

$$p_1(\theta, \sigma(p_2|\theta)) = \arg\max_{p_1} \int_{p_2} v_1(p_1, p_2; \theta) \mathrm{d}\sigma(p_2|\theta).$$

By Leibniz rule, $p_1(\theta, \sigma(p_2|\theta))$ is implicitly characterized by

$$\int_{p_2} q(p_1, p_2; \theta) \mathrm{d}\sigma(p_2|\theta) + p_1 \int_{p_2} \frac{\partial q(p_1, p_2; \theta)}{\partial p_1} \mathrm{d}\sigma(p_2|\theta) = 0.$$

Then, firm 1's expected equilibrium profits given information structure $\sigma$ are

$$\begin{aligned}
\mathbb{E}_(\mu, \sigma)[\Pi_1^*(p_1, p_2; \theta)] &= \sum_{\theta \in \Theta} \mu_\theta \mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta)|\theta] \\
&= \sum_{\theta \in \Theta} \mu_\theta \int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta)), p_2; \theta) \mathrm{d}\sigma(p_2|\theta)
\end{aligned}$$

Furthermore, for any information structure $\sigma$, the set of recommended equilibrium prices for firm 2 in the pricing game is a subset of the interval between the equilibrium prices with full disclosure. This is formalized in Lemma A.7.

**Lemma A.7** *The support of any obedient distribution $\sigma(p_2|\theta)$ is a subset of*

$$[p^F(\theta_L), p^F(\theta_H)] \textit{ for all } \theta \in \Theta$$

*where $p^F(\theta)$ is the equilibrium price with full disclosure when the state $\theta$ is realized.*

**Informed firm optimal disclosure.** Assume the designer wants to maximize the informed firm's expected equilibrium payoffs. First, I show that when firms offer substitutes, the informed firm's expected equilibrium payoff conditional on the state is supermodular in the state and the price of the other firm. Similarly, I also show that the informed firm's expected equilibrium payoff conditional on the state is submodular in the state and the price of the other firm when firms offer complements. This is formalized in Lemma A.8.

**Lemma A.8** *When firms offer substitutes (complements), $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta)|\theta]$ is supermodular (submodular) in $\theta$ and $p_2$.*

Second, I show that it is optimal for the informed firm to share all its private information with the uninformed firm when the informed firm expected equilibrium profits are supermodular in the state and the uninformed firm's price. I also show that it is optimal for the informed firm to share none of its private information with the uninformed firm when the informed firm expected equilibrium profits are submodular in the state and the uninformed firm's price. This is formalized in Proposition A.1.

**Proposition A.1** *If $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta)|\theta]$ is supermodular in $p_2$ and $\theta$, full disclosure is optimal for the informed firm. Similarly, if $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta)|\theta]$ is submodular in $p_2$ and $\theta$, no disclosure is optimal for the informed firm.*

These two results imply that full disclosure is optimal for the informed firm when firms offer substitutes and no disclosure is optimal when firms offer complements. These results also extend to Cournot competition using same equivalence arguments as before.

### A.3.1 Proofs

**Proof. Lemma A.6.** The pricing game is a smooth concave game since $\Pi_i(\cdot, p_{-i}; \theta) : \mathbb{R}_+ \to \mathbb{R}$ is concave and continuously differentiable for each $p_{-i} \in \mathbb{R}_+$ since

$$\frac{\partial^2 \Pi_i(p_i, p_{-i}; \theta)}{\partial p_i^2} < 0 \text{ for all } p_{-i} \in \mathbb{R}_+.$$

Define the payoff gradient as

$$\nabla \Pi(\mathbf{p}, \theta) := \left( \frac{\partial \Pi_i((p_i, p_{-i}); \theta)}{\partial p_i} \right)_{i \in \{1,2\}}.$$

The payoff gradient is continuously differentible. The Jacobian matrix of the payoff gradient, given by

$$F_{\nabla \Pi}(\mathbf{p}, \theta) := \begin{pmatrix} \frac{\partial^2 \Pi_1((p_1,p_2);\theta)}{\partial p_1^2} & \frac{\partial^2 \Pi_1((p_1,p_2);\theta)}{\partial p_1 \partial p_2} \\ \frac{\partial^2 \Pi_2((p_2,p_1);\theta)}{\partial p_1 \partial p_2} & \frac{\partial^2 \Pi_2((p_2,p_1);\theta)}{\partial p_2^2} \end{pmatrix},$$

is negative definite because

$$\frac{\partial^2 \Pi_i(p_i, p_{-i};\theta)}{\partial p_i^2} < 0 \text{ and } \frac{\partial^2 \Pi_i(p_i, p_{-i};\theta)}{\partial p_i^2}\frac{\partial^2 \Pi_{-i}(p_{-i}, p_i;\theta)}{\partial p_{-i}^2} \geq \left(\frac{\partial^2 \Pi_i(p_i, p_{-i};\theta)}{\partial p_i \partial p_{-i}}\right)^2.$$

This implies that the payoff gradient $\nabla\Pi(\mathbf{p},\theta)$ is strictly monotone by Lemma 4 from Ui (2016). Furthermore, since for all $\mathbf{p} := (p_i, p_{-i})$, there exists $c > 0$ such that

$$\mathbf{p}^T F_{\nabla\Pi}(\mathbf{p},\theta)\mathbf{p} < -c\mathbf{p}^T\mathbf{p},$$

the payoff gradient is also strongly monotone by the same lemma. Then, the uniqueness of the Bayesian Nash equilibrium of the pricing game follows from Proposition 1 from Ui (2016), which states that if the payoff gradient is strictly monotone, the Bayesian game as at most one Bayesian Nash equilibrium. The existence of a unique Bayesian Nash equilibrium follows from Proposition 2 from Ui (2016). ■

**Proof. Lemma A.7.** With full disclosure, there is no uncertainty about the state. Each firm chooses $p_i : \Theta \to \mathbb{R}_+$ to maximize $\Pi_i(p_i, p_{-i};\theta)$. That is, firm $i$'s best response to $p_{-i}$ is implicitly defined by

$$q(p_i, p_{-i};\theta) + p_i\frac{\partial q(p_i, p_{-i};\theta)}{\partial p_i} = 0$$

In equilibrium, both firms choose the same price, denoted by $p^F(\theta)$. Since $q(p_2, p_1;\theta_L) < q(p_2, p_1, \theta_H)$, the highest (lowest) equilibrium price the uninformed firm is willing to price is when both firms are certain that the state is high (low). Hence, the support of any obedient recommendation $\sigma(p_2|\theta)$ is a subset of $[p^F(\theta_L), p^F(\theta_H)]$. ■

**Proof. Lemma A.8.** By definition, $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2;\theta)|\theta]$ is given by

$$\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2;\theta)|\theta] = \int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta)), p_2;\theta)\mathrm{d}\sigma(p_2|\theta)$$

When firms offer substitutes, for any obedient $\sigma(p_2|\theta)$ we have that

$$\begin{aligned}&\int_{p_2} q(p_1, p_2;\theta)\mathrm{d}\sigma(p_2|\theta_H) \geq \int_{p_2} q(p_1, p_2;\theta)\mathrm{d}\sigma(p_2|\theta_L) \text{ and}\\ &\int_{p_2} \frac{\partial q(p_1, p_2;\theta)}{\partial p_1}\mathrm{d}\sigma(p_2|\theta_H) \geq \int_{p_2} \frac{\partial q(p_1, p_2;\theta)}{\partial p_1}\mathrm{d}\sigma(p_2|\theta_L) \text{ for all } p_1 \text{ and } \theta\end{aligned} \tag{A.7}$$

since $q(p_1, p_2; \theta)$ is strictly increasing in $p_2$, $\int_0^x \mathrm{d}\sigma(p_2|\theta_L) \geq \int_0^x \mathrm{d}\sigma(p_2|\theta_H)$ for all $x$ and $\frac{\partial^2 q(p_1,p_2;\theta)}{\partial p_1 \partial p_2} > 0$. Then, since $p_1(\theta, \sigma(p_2|\theta))$ is implicitly defined by

$$\int_{p_2} q(p_1, p_2; \theta)\mathrm{d}\sigma(p_2|\theta) + p_1 \int_{p_2} \frac{\partial q(p_1, p_2; \theta)}{\partial p_1} \mathrm{d}\sigma(p_2|\theta) = 0,$$

(A.7) implies that $p_1(\theta, \sigma(p_2|\theta_H)) \geq p_1(\theta, \sigma(p_2|\theta_L))$ for all $\theta \in \Theta$. Then, $\frac{\partial^2 q(p_1,p_2,\theta)}{\partial p_1 \partial p_2} \geq 0$ also implies that

$$\int_{p_2} q(p_1(\theta, \sigma(p_2|\theta_H)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta_H) \geq \int_{p_2} q(p_1(\theta, \sigma(p_2|\theta_L)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta_L).$$

Then, when firms offer substitutes,

$$\int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta_H)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta_H) - \int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta_L)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta_L) \geq 0 \tag{A.8}$$

for all $\theta \in \Theta$. By Leibnitz rule,

$$\begin{aligned}
&\frac{\partial \int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta_H)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta_H)}{\partial \theta} \\
&= \int_{p_2} \left.\frac{\partial p_1(t, \sigma(p_2|\theta_H))}{\partial t}\right|_{t=\theta} \left[q(p_1(\theta, \sigma(p_2|\theta_H)), p_2; \theta) + p_1(\theta, \sigma(p_2|\theta_H))\frac{\partial q(p_1, p_2; t)}{\partial p_1}\right] \mathrm{d}\sigma(p_2|\theta_H) \\
&+ \int_{p_2} p_1(\theta, \sigma(p_2|\theta_H)) \left.\frac{\partial q(p_1(\theta, \sigma(p_2|\theta_H)), p_2; t)}{\partial t}\right|_{t=\theta} \mathrm{d}\sigma(p_2|\theta_H) \\
&= \int_{p_2} p_1(\theta, \sigma(p_2|\theta_H)) \left.\frac{\partial q(p_1(\theta, \sigma(p_2|\theta_H)), p_2; t)}{\partial t}\right|_{t=\theta} \mathrm{d}\sigma(p_2|\theta_H)
\end{aligned}$$

where the last inequality holds by the first order condition of the informed firm's pricing decision. Similarly,

$$\frac{\partial \int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta_L)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta_L)}{\partial \theta} = \int_{p_2} p_1(\theta, \sigma(p_2|\theta_L)) \left.\frac{\partial q(p_1(\theta, \sigma(p_2|\theta_L)), p_2; t)}{\partial t}\right|_{t=\theta} \mathrm{d}\sigma(p_2|\theta_L)$$

Then, the left-hand side of (A.8) is non-decreasing in $\theta$ since

$$\begin{aligned}
&\int_{p_2} p_1(\theta, \sigma(p_2|\theta_H)) \left.\frac{\partial q(p_1(\theta, \sigma(p_2|\theta_H)), p_2; t)}{\partial t}\right|_{t=\theta} \mathrm{d}\sigma(p_2|\theta_H) \\
\geq &\int_{p_2} p_1(\theta, \sigma(p_2|\theta_L)) \left.\frac{\partial q(p_1(\theta, \sigma(p_2|\theta_L)), p_2; t)}{\partial t}\right|_{t=\theta} \mathrm{d}\sigma(p_2|\theta_L)
\end{aligned}$$

because $p_1(\theta, \sigma(p_2|\theta_H)) > p_1(\theta, \sigma(p_2|\theta_L))$ and $\frac{\partial^2 q(p_1,p_2;\theta)}{\partial \theta \partial p_2} > 0$. Thus, when firms offer substitutes,

$$\int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_H)), p_2; \theta_H) \mathrm{d}\sigma(p_2|\theta_H) - \int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_L)), p_2; \theta_H) \mathrm{d}\sigma(p_2|\theta_L)$$
$$\geq \int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_H)), p_2; \theta_L) \mathrm{d}\sigma(p_2|\theta_H) - \int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_L)), p_2; \theta_L) \mathrm{d}\sigma(p_2|\theta_L)$$

which implies that $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta)|\theta]$ is supermodular in $\theta$ and $p_2$. The proof for the complement case is analogous. ■

**Proof. Proposition A.1.** Consider first the case in which $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2\theta)|\theta]$ is supermodular in $p_2$ and $\theta$. Next, I show that for all $\sigma$ and $p_2 \in [p^F(\theta_L), p^F(\theta_H)]$,

$$\mathbb{E}_{\sigma^F}[\Pi_1^*(p_1, p_2; \theta_L)|\theta_L] \leq \mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta_L)|\theta_L] \leq \mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta_H)|\theta_H] \leq \mathbb{E}_{\sigma^F}[\Pi_1^*(p_1, p_2; \theta_H)|\theta_H].$$

That is,

$$\Pi_1(p_1(\theta_L, \sigma^F(p_2|\theta_L)), p^F(\theta_L); \theta_L) \leq \int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_L)), p_2; \theta_L) \mathrm{d}\sigma(p_2|\theta_L)$$
$$\leq \int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_H)), p_2; \theta_H) \mathrm{d}\sigma(p_2|\theta_H) \leq \Pi_1(p_1(\theta_H, \sigma^F(p_2|\theta_H)), p^F(\theta_H); \theta_H)$$

First,

$$\int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_L)), p_2; \theta_L) \mathrm{d}\sigma(p_2|\theta_L) \geq \int_{p_2} \Pi_1(p_1(\theta_L, \sigma^F(p_2|\theta_L)), p_2; \theta_L) \mathrm{d}\sigma^F(p_2|\theta_L)$$
$$= \Pi_1(p_1(\theta_L, \sigma^F(p_2|\theta_L)), p^F(\theta_L); \theta_L) \quad \text{(A.9)}$$

since $\sigma^F(p_2|\theta_L)$ recommends $p^F(\theta_L)$ with probability 1, the informed firm's demand is increasing in $p_2$ and $p_1(\theta_L, \sigma(p_2|\theta_L)) \geq p_1(\theta_L, \sigma^F(p_2|\theta_L))$.[6] Similarly,

$$\int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_H)), p_2; \theta_H) \mathrm{d}\sigma(p_2|\theta_H) \leq \int_{p_2} \Pi_1(p_1(\theta_H, \sigma^F(p_2|\theta_H)), p_2; \theta_H) \mathrm{d}\sigma^F(p_2|\theta_H)$$
$$= \Pi_1(p_1(\theta_H, \sigma^F(p_2|\theta_H)), p^F(\theta_H); \theta_H) \quad \text{(A.10)}$$

because $\sigma^F(p_2|\theta_H)$ recommends $p^F(\theta_H)$ with probability 1, the informed firm's demand is increasing in $p_2$ and $p_1(\theta_H, \sigma(p_2|\theta_H)) \leq p_1(\theta_H, \sigma^F(p_2|\theta_H))$.

[6]The proof of $p_1(\theta_L, \sigma(p_2|\theta_L)) \geq p_1(\theta_L, \sigma^F(p_2|\theta_L))$ follows an analogous argument as in Lemma A.8.

Second, supermodularity implies that

$$\int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_L)), p_2; \theta_L)\mathrm{d}\sigma(p_2|\theta_L) \leq \int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_H)), p_2; \theta_H)\mathrm{d}\sigma(p_2|\theta_H) \tag{A.11}$$

since $p_1(\theta_L, \sigma(p_2|\theta_L)) \leq p_1(\theta_H, \sigma(p_2|\theta_H))$, $\frac{\partial^2 q(p_1,p_2;\theta)}{\partial\theta\partial p_2} > 0$ and the state is a positive demand shifter, implying that $\sigma(p_2|\theta_H)$ recommends on average higher prices than $\sigma(p_2|\theta_L)$. Thus, (A.9), (A.10) and (A.11) imply that

$$\begin{aligned}
&\Pi_1(p_1(\theta_L, \sigma^F(p_2|\theta_L)), p^F(\theta_L); \theta_L) \leq \int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_L)), p_2; \theta_L)\mathrm{d}\sigma(p_2|\theta_L) \\
&\leq \int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_H)), p_2; \theta_H)\mathrm{d}\sigma(p_2|\theta_H) \leq \Pi_1(p_1(\theta_H, \sigma^F(p_2|\theta_H)), p^F(\theta_H); \theta_H)
\end{aligned}$$

Then,

$$\begin{aligned}
\mathbb{E}_{\sigma^F,\mu}[\Pi_1^*(p_1, p_2; \theta)] &= \sum_{\theta\in\Theta} \mu_\theta \Pi_1(p_1(\theta, \sigma^F(p_2|\theta)), p^F(\theta); \theta) \\
&\geq \sum_{\theta\in\Theta} \mu_\theta \int_{p_2} \Pi_1(p_1(\theta, \sigma(p_2|\theta)), p_2; \theta)\mathrm{d}\sigma(p_2|\theta) \\
&= \mathbb{E}_{\sigma,\mu}[\Pi_1^*(p_1, p_2; \theta)]
\end{aligned}$$

where the inequality holds by Jensen's inequality.

Consider now the case in which $\mathbb{E}_\sigma[\Pi_1^*(p_1, p_2; \theta)|\theta]$ is submodular in $\theta$ and $\sigma(p_2|\theta)$. Analogously as in the supermodular case, it is possible to show that

$$\begin{aligned}
&\Pi_1(p_1(\theta_L, \sigma^N(p_2|\theta_L)), p^N; \theta_L) \leq \int_{p_2} \Pi_1(p_1(\theta_L, \sigma(p_2|\theta_L)), p_2; \theta_L)\mathrm{d}\sigma(p_2|\theta_L) \\
&\leq \int_{p_2} \Pi_1(p_1(\theta_H, \sigma(p_2|\theta_H)), p_2; \theta_H)\mathrm{d}\sigma(p_2|\theta_H) \leq \Pi_1(p_1(\theta_H, \sigma^N(p_2|\theta_H)), p^N; \theta_H)
\end{aligned}$$

which in turn implies that

$$\begin{aligned}\mathbb{E}_{\sigma^N,\mu}[\Pi_1^*(p_1,p_2;\theta)] &= \sum_{\theta\in\Theta}\mu_\theta\Pi_1(p_1(\theta,p^N),p^N;\theta)\\ &\geq \sum_{\theta\in\Theta}\mu_\theta\int_{p_2}\Pi_1(p_1(\theta,\sigma(p_2|\theta)),p_2;\theta)\mathrm{d}\sigma(p_2|\theta_H)\\ &= \mathbb{E}_{\sigma,\mu}[\Pi_1^*(p_1,p_2;\theta)]\end{aligned}$$

where the inequality holds by Jensen's inequality. ■

## A.4 Extensions proofs

**Proof. Lemma II.6.** Proposition II.1 and Lemma II.5 imply that full disclosure is optimal if firms offer imperfect substitutes. If firms offer complements, the informed firm prefers to not disclose her private information whereas the uninformed firm prefers to learn the state. First, full disclosure yields higher producer surplus than no disclosure if and only if $\delta < \frac{2}{1+\sqrt{2}}$. Next, I show that full disclosure is optimal when firms offer complements if $\delta < \frac{2}{1+\sqrt{2}}$ and no disclosure is optimal otherwise.

Consider first the case in which $\delta < \frac{2}{1+\sqrt{2}}$. The difference in expected producer surplus of full disclosure $\sigma^F$ and disclosure policy $\sigma$ is

$$PS(\sigma^F) - PS(\sigma) \geq \left(a + b - \frac{b^2}{4a}\right)\mathbb{E}_\mu\left[\left(\frac{\theta}{2a-b} - \mathbb{E}_\sigma[p_2|\theta]\right)^2\right] \geq 0$$

The first inequality holds by Jensen's inequality, $a > |b|$ and $b < 0$, whereas the second one holds for all $\delta < \frac{2}{1+\sqrt{2}}$. Hence, full disclosure is optimal if $\delta < \frac{2}{1+\sqrt{2}}$.

Consider now the case in which $\delta \geq \frac{2}{1+\sqrt{2}}$. The difference in expected producer surplus of no disclosure $\sigma^N$ and disclosure policy $\sigma$ is

$$PS(\sigma^N) - PS(\sigma) = a\mathbb{V}_{(\mu,\sigma)}[p_2] - (1-\delta)\,Cov_{(\mu,\sigma)}[p_2,\theta] + \frac{3b}{4}\cdot\delta\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$$

Note that this difference is a strictly increasing function of $\delta$ because

$$\begin{aligned}&\frac{\partial PS(\sigma^N) - PS(\sigma)}{\partial\delta} = Cov_{(\mu,\sigma)}[p_2,\theta] + \frac{3b}{4}\mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]],\\ &Cov_{(\mu,\sigma)}[p_2,\theta] > 2a\mathbb{V}_{(\mu,\sigma)}[p_2],\ \delta < 1 \text{ and } \mathbb{V}_{(\mu,\sigma)}[p_2] \geq \mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]] \geq 0.\end{aligned}$$

This implies that

$$PS(\sigma^N) - PS(\sigma) \geq a\mathbb{V}_{(\mu,\sigma)}[p_2] - \left(1 - \frac{2}{1+\sqrt{2}}\right) Cov_{(\mu,\sigma)}[p_2,\theta] + \frac{3b}{4} \cdot \frac{2}{1+\sqrt{2}} \mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]] \geq 0$$

for all $\delta > \frac{2}{1+\sqrt{2}}$ since $Cov_{(\mu,\sigma)}[p_2,\theta] \in \big[2a\mathbb{V}_{(\mu,\sigma)}[p_2], (2a-b)\mathbb{V}_{(\mu,\sigma)}[p_2]\big]$ and $\mathbb{V}_{(\mu,\sigma)}[p_2] \geq \mathbb{V}_\mu[\mathbb{E}_\sigma[p_2|\theta]]$.[7] Hence, no disclosure is optimal if $\delta > \frac{2}{1+\sqrt{2}}$. ■

**Proof. Lemma II.7.** With public signals, the informed firm's expected equilibrium profits, given by

$$\mathbb{E}_{(\mu,\sigma^{Pub})}[\Pi_1^*((p_1,p_2);\theta)] = a\mathbb{E}_\mu\left[\mathbb{E}_{\sigma^{Pub}}\left[\left(\frac{\theta+bp_2}{2a}\right)^2\middle|\theta\right]\right],$$

are higher than its expected equilibrium profits with private signals. This holds because

$$\begin{aligned}\mathbb{E}_{(\mu,\sigma^{Pub})}[\Pi_1^*((p_1,p_2);\theta)] &= a\mathbb{E}_\mu\left[\mathbb{E}_{\sigma^{Pub}}\left[\left(\frac{\theta+bp_2}{2a}\right)^2\middle|\theta\right]\right] \geq a\mathbb{E}_\mu\left[\left(\mathbb{E}_{\sigma^{Pub}}\left[\frac{\theta+bp_2}{2a}\middle|\theta\right]\right)^2\right]\\ &= a\mathbb{E}_\mu\left[\left(\frac{\theta+b\mathbb{E}_{\sigma^{Pub}}[p_2|\theta]}{2a}\right)^2\right] = a\mathbb{E}_\mu\left[\left(\frac{\theta+b\mathbb{E}_{\sigma^{Priv}}[p_2|\theta]}{2a}\right)^2\right]\\ &= \mathbb{E}_{(\mu,\sigma^{Priv})}[\Pi_1^*((p_1,p_2);\theta)]\end{aligned}$$

where the inequality holds by Jensen's inequality. ■

**Proof. Lemma II.8.** The expected consumer surplus with public disclosure is given by

$$CS(\sigma^{Pub}) = \frac{1}{2a}\mathbb{E}_\mu\left[\mathbb{E}_{\sigma^{Pub}}\left[q_1\left(\frac{\theta+bp_2}{2a},p_2;\theta\right)^2 + q_2\left(p_2,\frac{\theta+bp_2}{2a};\theta\right)^2\middle|\theta\right]\right],$$

whereas expected consumer surplus with private disclosure is

$$CS(\sigma^{Priv}) = \frac{1}{2a}\mathbb{E}_\mu\left[\mathbb{E}_{\sigma^{Priv}}\left[q_1\left(\frac{\theta+b\mathbb{E}_{\sigma^{Priv}}[p_2|\theta]}{2a},p_2;\theta\right)^2 + q_2\left(p_2,\frac{\theta+b\mathbb{E}_{\sigma^{Priv}}[p_2|\theta]}{2a};\theta\right)^2\middle|\theta\right]\right].$$

[7] The highest covariance between prices and the state occurs with full disclosure. In this case, $p_2(\theta) = \frac{\theta}{2a-b}$. Hence,

$$Cov_{(\mu,\sigma)}[p_2,\theta] \leq Cov_{(\mu,\sigma)}[p_2,(2a-b)p_2] = (2a-b)\mathbb{V}_{(\mu,\sigma)}[p_2]$$

The difference between expected consumer surplus with private and public disclosure is

$$CS(\sigma^{Priv}) - CS(\sigma^{Pub}) = \frac{b^2}{8a}\left(7 - \frac{b^2}{a^2}\right)\left(\mathbb{E}_{(\mu,\sigma)}[p_2^2] - \mathbb{E}_{\mu}\left[\mathbb{E}_{\sigma}[p_2|\theta]^2\right]\right)$$

Then, $CS(\sigma^{Priv}) \geq CS(\sigma^{Pub})$ because $a > |b|$ and

$$\mathbb{E}_{(\mu,\sigma)}[p_2^2] - \mathbb{E}_{\mu}\left[\mathbb{E}_{\sigma}[p_2|\theta]^2\right] = \mathbb{E}_{\mu}[\mathbb{E}_{\sigma}[p_2^2|\theta]] - \mathbb{E}_{\mu}\left[\mathbb{E}_{\sigma}[p_2|\theta]^2\right] \geq 0$$

where the equality holds by the law of iterated expectations and the inequality by Jensen's inequality. ■

**Proof. Lemma II.9.** First, no disclosure is optimal when firms offer complements since $CS(\sigma^{Pub}) \leq CS(\sigma^{Priv})$. Similarly, no disclosure is optimal when firms offer substitutes and firms offer sufficiently far substitutes ($\delta < \hat{c}$). Consider then the case in which firms offer substitutes ($b > 0$) and $\delta \geq \hat{c}$. The expected gain of consumer surplus with public disclosure with respect to no disclosure is given by:

$$\begin{aligned} CS(\sigma^{Pub}) - CS(\sigma^{N}) &\leq \frac{1}{2a}\left(\frac{a}{2}V_{(\mu,\sigma^{Pub})}[p_2] - Cov_{(\mu,\sigma^{Pub})}(\theta, p_2)\right) \\ &< 0 \end{aligned}$$

where the first inequality holds by definition of variance, covariance and $\delta$. The second inequality holds because $CS(\sigma^{Pub}) < CS(\sigma^{N})$ for $\delta = 0$. ■

**Proof. Lemma II.10.** The designer commits to an information structure with private signals, denoted by $\hat{\psi}_k$, to share all the informed firm's private information with $k$ firms and share no information with $N-1-k$ firms, where $k \in \{0, 1, 2, ..., N-1\}$. Firms who observe a perfectly informative signal condition their pricing choices on the state and select a mapping $p^F : \Theta \to \mathbb{R}_+$ to maximize their expected profits, whereas firms who observe no information select a price $p^N \in \mathbb{R}_+$ to maximize their expected profits. Equilibrium prices are

$$
\begin{aligned}
p^F(\theta_L) &= \frac{\theta_L(2a(N-1)-bk) + b\mu_H(N-k-1)(\theta_H-\theta_L)}{(2a-b)(2a(N-1)-bk)}, \\
p^F(\theta_H) &= \frac{\theta_H(2a(N-1)-bk) - b\mu_L(N-k-1)(\theta_H-\theta_L)}{(2a-b)(2a(N-1)-bk)}, \text{ and} \\
p^N &= \frac{\mu_L\theta_L + \mu_H\theta_H}{2a-b}.
\end{aligned}
$$

Consider first the case in which the designer's objective is to maximize the informed firm's expected equilibrium profits, given by

$$\mathbb{E}[\Pi_1^*(\hat{\psi}_k)] = a\sum_{\theta\in\Theta}\mu_\theta p^F(\theta)^2.$$

The informed firm's expected equilibrium profits are maximized by sharing its private information with all other firms ($k^* = N-1$). Similarly, when the designer's objective is to maximize expected producer surplus, given by

$$PS(\hat{\psi}_k) = (N-k-1)a(p^N)^2 + (k+1)a\sum_{\theta\in\Theta}\mu_\theta p^F(\theta)^2,$$

it is optimal to share information with all firms ($k^* = N-1$), eliminating all information asymmetry between firms.

In contrast, if the designer's objective is to maximize expected consumer surplus, information disclosure between firms is at least partially restricted. Expected consumer surplus, given by,

$$
\begin{aligned}
CS(\hat{\psi}_k) = & \frac{(k+1)}{2a}\sum_{\theta\in\Theta}\mu_\theta\left[\theta + b\frac{(N-k-1)}{N-1}p^N - \left(a - b\frac{k}{N-1}\right)p^F(\theta)\right]^2 \\
& + \frac{(N-k-1)}{2a}\sum_{\theta\in\Theta}\mu_\theta\left[\theta + b\frac{k+1}{N-1}p^F(\theta) - \left(a - b\frac{N-k-2}{N-1}\right)p^N\right]^2
\end{aligned}
$$

The optimal information structure, characterized by $k^*(N,\delta)$, is determined by the degree of substitution and the number of firms in the market, where

$$k^*(N,\delta) = \begin{cases} 0 & \text{if } \delta \leq \frac{3}{4} \text{ for all } N \geq 3 \\ 0 & \text{if } \delta \in \left(\frac{3}{4}, 0.76\right) \text{ and } N \in \left[3, 1 + \frac{1}{2}\sqrt{\frac{\delta^2}{4\delta-3}} - \frac{\delta}{2}\right] \\ f(N,\delta) & \text{otherwise} \end{cases}$$

and

$$f(N,\delta) = \lceil \frac{2(N-1)\left(\delta^3 + \delta^2(4N-5) + \delta(4N-7)(N-1) - 3(N-1)^2\right)}{\delta\left(\delta + (N-1)\right)\left(\delta + 3(N-1)\right)} \rceil$$

if

$$CS\left[\hat{\pi}_{\lceil \frac{2(N-1)\left(\delta^3+\delta^2(4N-5)+\delta(4N-7)(N-1)-3(N-1)^2\right)}{\delta(\delta+(N-1))(\delta+3(N-1))} \rceil}\right] \geq CS\left[\hat{\pi}_{\lfloor \frac{2(N-1)\left(\delta^3+\delta^2(4N-5)+\delta(4N-7)(N-1)-3(N-1)^2\right)}{\delta(\delta+(N-1))(\delta+3(N-1))} \rfloor}\right]$$

and

$$f(N,\delta) = \lfloor \frac{2(N-1)\left(\delta^3 + \delta^2(4N-5) + \delta(4N-7)(N-1) - 3(N-1)^2\right)}{\delta\left(\delta + (N-1)\right)\left(\delta + 3(N-1)\right)} \rfloor,$$

otherwise.

■

# APPENDIX B

# Appendix for Chapter III

## B.1 Proofs with two data-buyers

**Proof. Lemma III.1**. I show that for any menu $\hat{\mathcal{M}}$ and BNE $\hat{\sigma}$, there exists a direct menu $\mathcal{M}^D$ and $\sigma^D$ such that (i) every data-buyer $i$ of type $\theta \in \{\theta_L, \theta_H\}$ purchases the experiment designed for her type; (ii) for every type vector, the distribution over outcomes under $\hat{\mathcal{M}}$ if $\hat{\sigma}$ is played is the same as the distribution over outcomes that results under $\mathcal{M}^D$ and $\sigma^D$.

Suppose instead that $M \geq 3$. First, suppose that $\hat{\iota}_{i\theta} \in \{0, 1, ..., M\}$ for all $i$ type $\theta$. Hence, only two elements of the menu are traded in equilibrium. Then, it is possible to eliminate the redundant elements of the menu and offer menu $\mathcal{M}^D$ which only includes the ones that are purchased in equilibrium. In this case, it is trivial that the distribution over outcomes remains unchanged. Second, assume that there exist $i$ type $\theta$ such that $\hat{\iota}_i \in \Delta(\{0, 1, .., M\})$. Construct an alternative menu $\mathcal{M}^D$ in which the experiments that are chosen by $i$ type $\theta$ with positive probability are replace by one experiment that randomizes over those experiments such that induces the same distribution over outcomes. That is, define

$$\pi^D((s_i, s_j)|\omega, (\iota_i, \iota_j)) = \sum_{m=0}^{M} \mathbb{P}(\hat{\iota}_i = m|\omega)\hat{\pi}((s_i, s_j)|\omega, (m, \iota_j)).$$

Note that the overall distribution over outcomes remains unchanged. Thus, $\mathcal{M}^D$

implements the same outcome as $\hat{\mathcal{M}}$ and, since $\hat{\sigma}$ is a Bayes Nash equilibrium, $\sigma^D$ is also an equilibrium. Therefore, it is without loss of generality to consider menus with at most two elements. ■

**Proof. Lemma III.2**. Consider $i$ type $\theta$ and an experiment $E^m \in \mathcal{M}$ where the menu $\mathcal{M}$ is a incentive-compatible and individually rational. Let $i$ type $\theta$ choose a single action after each signal. Given the equilibrium strategies, let $S_k^m$ denote the subset of signals in experiment $E^m$ that induces buyer $i$ type $\theta$ to choose action $a_k \in \{a_1, a_2\}$ where $\cup_{k=1}^2 S_k^m = S^m$.

Construct $\widehat{E}^m = (\widehat{S}^m, \widehat{\pi}^m)$ where $\widehat{S}^m = \{s_1, s_2\}$ and, for all $s_\ell \in \{s_1, s_2\}$ and $\omega \in \{\omega_1, \omega_2\}$,

$$\widehat{\pi}^m(s_\ell|\omega) = \int_{S_k^m} \pi^m(s|\omega)\mathrm{d}s.$$

$E^m$ and $\hat{E}^m$ are constructed such that both experiments induce the same outcome for buyer $i$ type $\theta$. Thus, $i$ attaches the same value to both experiments, i.e., $V_{\mathcal{M}}(E^m, \theta) = V_{\hat{\mathcal{M}}}(\widehat{E}^m, \theta)$. Furthermore, $\widehat{E}^m$ is a weakly less informative than $E^m$ and Blackwell's theorem implies that $V_{\mathcal{M}}(E^m, \theta') \leq V_{\hat{\mathcal{M}}}(\widehat{E}^m, \theta')$ for all $\theta'$. This relaxes the incentive constraints of types $\theta' \neq \theta$. Therefore, for any $\mathcal{M}$, it is possible to construct $\widehat{\mathcal{M}}$ that replaces $E^m$ with $\widehat{E}^m$ that is also incentive compatible and individually rational and yields weakly larger profits. ■

**Lemma B.1** *The value of experiment $E^n$ is weakly increasing in its precision if $c \in \left(\frac{1}{2}, 2\right)$.*

**Proof. Lemma B.1.** Data-buyer $i$ type $\theta$'s willingness to pay for an experiment is determined by $v_1(E^n, \theta; m)$ and $v_2(E^n, \theta\ m)$. First, if $i$ purchases experiment $E^n$ with $n \neq m$, both expressions are increasing in $\pi_1^n$ and $\pi_2^n$ since

$$\frac{\partial v_2(E^n, \theta; m)}{\partial \pi_1^n} = \frac{\partial v_1(E^n, \theta; m)}{\partial \pi_1^n} = \theta[\pi_1^m + (1 - \pi_1^m)c] \text{ and}$$
$$\frac{\partial v_2(E^n, \theta; m)}{\partial \pi_2^n} = \frac{\partial v_2(E^n, \theta; m)}{\partial \pi_2^n} = (1 - \theta)[(1 - \pi_2^m)c + \pi_2^m]$$

Second, if $i$ acquires the same experiment as $j$ $(n = m)$, we have that

$$\frac{\partial v_k(E^m, \theta; m)}{\partial \pi_1^m} = \begin{cases} \theta\left[\pi_1^m(2-c) + (1-\pi_1^m)c\right] & if\ k=2 \\ \theta\left[\pi_1^m + (1-\pi_1^m)(2c-1)\right] & if\ k=1 \end{cases}$$

and

$$\frac{\partial v_k(E^m, \theta; m)}{\partial \pi_2^m} = \begin{cases} (1-\theta)\left[(1-\pi_2^m)(2c-1) + \pi_2^m\right] & if\ k=2 \\ (1-\theta)\left[(1-\pi_2^m)c + \pi_2^m(2-c)\right] & if\ k=1 \end{cases}.$$

Hence, $c \in \left(\frac{1}{2}, 2\right)$ ensures that $V_{\mathcal{M}}(E^n, \theta)$ increases in $\pi_1^n$ and $\pi_2^n$ for all experiments $E^n$. ■

**Proof. Proposition III.1**.

1. First, I show that $IR_L$ binds. Suppose not. Then, $IR_H$ must bind.[1] Since the high type observes $\overline{E}$, $t^H = V_{\mathcal{M}}(\overline{E}, \theta_H) > V_{\mathcal{M}}(\overline{E}, \theta_L)$ and $t^L < V_{\mathcal{M}}(E^L, \theta_L)$. Moreover, the incentive compatibility constraints imply

$$IC_L : V_{\mathcal{M}}(E^L, \theta_L) - t^L \geq V_{\mathcal{M}}(\overline{E}, \theta_L) - t^H \text{ and } IC_H :\ 0 \geq V_{\mathcal{M}}(E^L, \theta_H) - t^L.$$

Since $V_{\mathcal{M}}(E^L, \theta_L) - t^L > 0$ and $V_{\mathcal{M}}(\overline{E}, \theta_L) - t^H < 0$, it is possible to increase $t^L$ by a small $\epsilon > 0$ without violating any compatibility constraint, yielding a contradiction.

Second, I show that the participation constraint of the high type also binds. Since the $IR_L$ and $IC_H$ bind, the data-seller's maximization can be written as:

$$\max_{(E^m, t^m)_{m\in\{L,H\}}} (1-\nu-\rho)V_{\mathcal{M}}(E^L, \theta_L) + (\nu+\rho)\left(V_{\mathcal{M}}(E^H, \theta_H) - [V(E^L, \theta_H) - V_{\mathcal{M}}(E^L, \theta_L)]\right)$$

subject to

$$IR_H : V(E^L, \theta_H) - V_{\mathcal{M}}(E^L, \theta_L) \geq 0 \text{ and}$$
$$IC_L : V_{\mathcal{M}}(E^H, \theta_H) - V(E^H, \theta_L) \geq V_{\mathcal{M}}(E^L, \theta_H)] - V(E^L, \theta_L).$$

Hence, $IR_H$ implies $IC_L$. As a result, we can further simplify the data-seller's max-

[1] At least one participation constraint binds. Otherwise, the seller can increase prices and profits.

imization problem to

$$\max_{(E^m,t^m)_{m\in\{L,H\}}} V_{\mathcal{M}}(E^L,\theta_L)+(\nu+\rho)\left(V_{\mathcal{M}}(E^H,\theta_H)-V(E^L,\theta_H)\right)$$

such that $V(E^L,\theta_H)\geq V_{\mathcal{M}}(E^L,\theta_L)$. Note that

$$\Pi_{\mathcal{M}}\leq V_{\mathcal{M}}(E^L,\theta_H)+(\nu+\rho)\left(V_{\mathcal{M}}(E^H,\theta_H)-V(E^L,\theta_H)\right)$$

since $V_{\mathcal{M}}(E^L,\theta_H)\geq V_{\mathcal{M}}(E^L,\theta_L)$. Thus, it is optimal for the data-seller to choose $E^L$ such that $V_{\mathcal{M}}(E^L,\theta_H)=V_{\mathcal{M}}(E^L,\theta_L)$ which implies that $IR_H$ binds.

2. Suppose $IC_H$ does not bind. Since $IR_H$ binds, it is possible to increase both the precision of $E^L$ and $t^L$ since payoffs are continuous in $\pi_k^L$, yielding a contradiction.

3. First, I show that if $\overline{E}$ is part of the optimal menu, it is offered to the high type. Suppose instead that only the low type purchases this experiment, implying that its price cannot exceed $V_{\mathcal{M}}(\overline{E},\theta_L)$. If the high type does not purchase this experiment, incentive-compatibility implies that $t^H<V_{\mathcal{M}}(\overline{E},\theta_L)$. Thus, only offering $\overline{E}$ at price $V_{\mathcal{M}}(\overline{E},\theta_L)$ improves the seller's profits. This yields a contradiction.

Second, I show that $\overline{E}$ is part of the optimal menu. Assume without loss of generality that $\theta_H$ chooses action $a_1$ in the absence of supplemental information. Consider first the case in which $\theta_L$ also chooses action $a_1$ under the same circumstances. Proposition 1.4 implies that $\pi_1^H=\pi_1^L=1$. The data-seller's expected profits from a menu $\mathcal{M}$ are

$$\begin{aligned}\Pi_{\mathcal{M}}=&(1-\theta_L)\,\pi_2^L(1-\nu-\rho)\left[\frac{(1-2\nu-\rho)(c(1-\pi_2^L)+\pi_2^L)}{1-\nu-\rho}+\frac{\nu(c(1-\pi_2^H)+\pi_2^H)}{1-\nu-\rho}\right]\\&+(\nu+\rho)\,(1-\theta_H)\,\pi_2^H\left[\frac{\nu(c(1-\pi_2^L)+\pi_2^L)}{\nu+\rho}+\frac{\rho(c(1-\pi_2^H)+\pi_2^H)}{\nu+\rho}\right]\end{aligned}$$

First, if $c<1$, it is trivial that the data-seller's profits are strictly increasing in $\pi_2^H$, implying that it is optimal to set $\pi_2^H=1$. Second, if $c>1$, the same result is true since

$$\frac{\partial\Pi_{\mathcal{M}}}{\partial\pi_2^H}\geq(1-\theta_H)\,[\nu(1-\pi_2^L)+\rho(1-\pi_2^H)],$$

$c \in (1,2)$ and $\theta_L > \theta_H$. Consider now the case in which the low type chooses $a_2$ in the absence of supplemental information. Proposition 1.4 implies that $\pi_1^H = \pi_2^L = 1$ and the expected profits from a menu $\mathcal{M}$ are

$$\Pi_{\mathcal{M}} = \theta_L \pi_1^L \left[\nu + (1-2\nu-\rho)(c(1-\pi_1^L)+\pi_1^L)\right] + \pi_2^H (1-\theta_H) \left[\nu + \rho \left[(c(1-\pi_2^H)+\pi_2^H)\right]\right].$$

Analogously as before, it is possible to verify that the data-seller's profits are strictly increasing in $\pi_2^H$ when $c \in (\frac{1}{2}, 2)$.

4. Assume without loss of generality that $\theta_H$ chooses action $a_1$ in the absence of supplemental information. Consider first the case in which $\theta_L$ also chooses action $a_1$ under the same circumstances. Suppose that in an optimal menu $\widehat{\mathcal{M}}$, $\pi_1^L < 1$. Consider an alternative menu $\overline{\mathcal{M}}$ that replaces $\widehat{E}^L$ characterized by $\pi_1^L < 1$ and $\pi_2^L$ with $\overline{E}^L$ which replaces $\pi_1^L$ with 1. Since $E^H = \overline{E}$, the difference between the data-seller expected profits with $\overline{\mathcal{M}}$, $\Pi_{\overline{\mathcal{M}}}$, and with $\widehat{\mathcal{M}}$, $\Pi_{\widehat{\mathcal{M}}}$, is

$$\frac{\Pi_{\overline{\mathcal{M}}} - \Pi_{\widehat{\mathcal{M}}}}{(1-\nu-\rho)}$$
$$= \min\left\{\pi_2^L (1-\theta_H) \left[\frac{\nu(c(1-\pi_2^L)+\pi_2^L)}{\nu+\rho} + \frac{\rho}{\nu+\rho}\right], (1-\pi_1^L)\theta_H \left[\frac{\nu(c(1-\pi_1^L)+\pi_1^L)}{\nu+\rho} + \frac{\rho}{\nu+\rho}\right]\right\}.$$

This contradicts the optimality of $\widehat{\mathcal{M}}$.

Consider now the case in which $\theta_L$ chooses action $a_2$ in the absence of supplemental information. Suppose that in an optimal menu, $\widehat{\mathcal{M}}$, $\pi_2^L < 1$. Consider an alternative menu $\overline{\mathcal{M}}$ that replaces $\pi_2^L$ with 1. The difference between the profits resulting from these two menus is

$$\Pi_{\overline{\mathcal{M}}} - \Pi_{\widehat{\mathcal{M}}} = (1-\theta_H)[1-\rho-\nu((1-\pi_2^L)c+\pi_2^L)]$$
$$+ \min\left\{\pi_1^L \theta_L[(1-2\nu-\rho)(c(1-\pi_1^L)+\pi_1^L)+\nu], (1-\pi_2^L)(1-\theta_L)[(1-2\nu-\rho)(c(1-\pi_2^L)+\pi_2^L)+\nu]\right\}.$$

Note that $\Pi_{\overline{\mathcal{M}}} - \Pi_{\widehat{\mathcal{M}}} \geq (1-\theta_H)(1-2\nu-\rho) > 0$ since both terms inside the minimum are positive and $c < 2$. ■

**Proof. Proposition III.2.** Proposition III.1 implies that $\pi_2^L = 1$ if beliefs are strictly non-congruent and that $\pi_1^L$ is such that $IC_H$ binds. When $\theta_L$ attaches a positive value to $E^L$,

$$V_{\mathcal{M}}(E^L, \theta_L) = \theta_L \pi_1^L \left[ \left( \frac{1-2\nu-\rho}{1-\nu-\rho} \right) [\pi_1^L + (1-\pi_1^L)c] + \left( \frac{\nu}{1-\nu-\rho} \right) \right] \text{ and}$$
$$V_{\mathcal{M}}(E^L, \theta_H) = \max \left\{ 0, (1-\theta_H) - (1-\pi_1^L)\theta_H \left( \frac{\nu}{\nu+\rho}(\pi_1^L + (1-\pi_1^L)c) + \frac{\rho}{\nu+\rho} \right) \right\}.$$

Define $f(\pi_1^L)$ as follows:

$$f(\pi_1^L) := (1-\theta_H) - (1-\pi_1^L)\theta_H \left( \frac{\nu}{\nu+\rho}(\pi_1^L + (1-\pi_1^L)c) + \frac{\rho}{\nu+\rho} \right).$$

Note that $f(\pi_1^L)$ is a continuous function of $\pi_1^L \in [0,1]$, $f(1) = (1-\theta_H) > 0$ and $f(0) < 0$ since $\theta_H > \frac{1}{2c}$ when $c \in \left(\frac{1}{2}, 1\right)$ and $\theta_H > \frac{c}{2}$ if $c \in (1,2)$. Then, the intermediate value theorem implies that there exists $\underline{\pi_1^L} \in (0,1)$ such that $f(\underline{\pi_1^L}) = 0$. Furthermore, since the graph of $f(\pi_1^L)$ is a convex parabola, $f(\pi_1^L) \geq 0$ for all $\pi_1^L \geq \underline{\pi_1^L}$ and negative otherwise. Hence, $V_{\mathcal{M}}(E^L, \theta_H) = f(\pi_1^L)$ if $\pi_1^L \geq \underline{\pi_1^L}$ and $V(E^L, \theta_H) = 0$ otherwise. Define $S^N(\pi_1^L)$ as the high type surplus from acquiring experiment $E^L$. That is:

$$S^N(\pi_1^L) = V_{\mathcal{M}}(E^L, \theta_H) - V_{\mathcal{M}}(E^L, \theta_L).$$

Note that $S^N(\pi_1^L = \underline{\pi_1^L}) < 0$, $S^N(\pi_1^L = 1) > 0$ because $\theta_L < 1 - \theta_H$ and $S^N(\pi_1^L)$ is continuous on the closed interval $[\underline{\pi_1^L}, 1]$. Then, by the Intermediate value theorem, there exists $\pi_1^L \in (\underline{\pi_1^L}, 1)$ such that $S^N(\pi_1^L) = 0$. This implies that the low type observes partial information whenever beliefs are strictly non-congruent. ■

**Proof. Lemma III.3.** The high type surplus from acquiring $E^L$, $S^N(\pi_1^L)$, is an increasing function of $\pi_1^L$ for all $\pi_1^L$ since

$$\begin{aligned} \frac{\partial S^N(\pi_1^L)}{\partial \pi_1^L} &\geq \theta_H \left[ \frac{\nu}{\nu+\rho}[\pi_1^L + (1-\pi_1^L)(2c-1)] + \frac{\rho}{\nu+\rho} \right] \\ &\quad - (1-\theta_H) \left[ \frac{1-2\nu-\rho}{1-\nu-\rho}[\pi_1^L(2-c) + (1-\pi_1^L)c] + \frac{\nu}{1-\nu-\rho} \right] \geq 0 \end{aligned}$$

where the first inequality holds since $\theta_L < 1 - \theta_H$ by definition of the high type and the second holds for all non-congruent distribution of types, $c \in \left(\frac{1}{2}, 2\right)$ and distribution of private information. The optimal $\pi_1^L$, defined as $\pi_1^L$ such that $S^N(\pi_1^L) = 0$, increases in $c$ since $S^N(\pi_1^L)$ is a decreasing function of $c$. This implies that the precision of $E^L$ decreases as the coordination incentives increase.

The effect of $\nu$ on $S^N(\pi_1^L)$ is determined by the coordination incentives since

$$\frac{\partial S^N(\pi_1^L)}{\partial \nu} = (1-\pi_1^L)(c-1)\left[\frac{1-\rho}{(1-\nu-\rho)^2}\theta_L \pi_1^L - \theta_H(1-\pi_1^L)\frac{\rho}{(\rho+\nu)^2}\right].$$

First, if data-buyers have coordination incentives, $S^N(\pi_1^L)$ is an decreasing function of $\nu$ if

$$\theta_H(1-\pi_1^L)\frac{\rho}{(\rho+\nu)^2} - \frac{1-\rho}{(1-\nu-\rho)^2}\theta_L \pi_1^L \leq 0. \tag{B.1}$$

Note that (B.1) holds for all

$$\pi_1^L \geq \frac{\theta_H \rho(1-\nu-\rho)^2}{\theta_H \rho(1-\nu-\rho)^2 + \theta_L(1-\rho)(\nu+\rho)^2} \text{and } S^N\left(\frac{\theta_H \rho(1-\nu-\rho)^2}{\theta_H \rho(1-\nu-\rho)^2 + \theta_L(1-\rho)(\nu+\rho)^2}\right) < 0. \tag{B.2}$$

Since the optimal $\pi_1^L$ satisfies (B.2), $S^N(\pi_1^L)$ is decreasing function of $\nu$ for all $\pi_1^L$ that satisfies (B.2) with coordination incentives. This also implies that the precision of $E^L$ increases in $\nu$ with anti-coordination incentives. ■

**Lemma B.2** *Assume $c \in \left(\frac{1}{2}, 1\right)$ and that beliefs are strictly congruent. In an optimal menu, the data-seller offers no information to $\theta_L$ if $\nu \leq \sqrt{\rho} - \rho$ and partial information otherwise.*

**Proof. Lemma B.2**. Proposition III.1 implies that in an optimal menu $E^H = \overline{E}$, $\pi_1^L = 1$ and that $\pi_2^L$ is determined such that $IC_H$ binds. Given that both participation constraints bind, $IC_H$ simplifies to $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$ where both expressions are computed by assuming that $j$ does not deviate form her equilibrium choices. The value of experiment $E^L$ for the $\theta_L$ and $\theta_H$ are given by

$$V_{\mathcal{M}}(E^L, \theta_L) = (1-\theta_L)\pi_2^L\left[\left(\frac{1-2\nu-\rho}{1-\nu-\rho}\right)\left[(1-\pi_2^L)c + \pi_2^L\right] + \left(\frac{\nu}{1-\nu-\rho}\right)\right] \text{ and}$$
$$V_{\mathcal{M}}(E^L, \theta_H) = (1-\theta_H)\pi_2^L\left[\left(\frac{\nu}{\nu+\rho}\right)\left[(1-\pi_2^L)c + \pi_2^L\right] + \left(\frac{\rho}{\nu+\rho}\right)\right].$$

It is trivial that $IC_H$ binds if $\pi_2^L = 0$. Suppose now that $\pi_2^L \in (0,1]$. Assume first that $\nu \le \sqrt{\rho} - \rho$. Then:

$$\begin{aligned} V_{\mathcal{M}}(E^L, \theta_H) - V_{\mathcal{M}}(E^L, \theta_L) &> (1-\theta_L)\,\pi_2^L \left( \frac{(1-2\nu-\rho)}{1-\nu-\rho}[\pi_2^L + (1-\pi_2^L)c] + \frac{\nu}{1-\nu-\rho} \right) \\ &\quad - (1-\theta_L)\,\pi_2^L \left( \frac{(1-2\nu-\rho)}{1-\nu-\rho}[\pi_2^L + (1-\pi_2^L)c] + \frac{\nu}{1-\nu-\rho} \right) \\ &= 0 \end{aligned}$$

where the inequality holds since $\theta_L > \theta_H$, $c < 1$ and $\nu \le \sqrt{\rho} - \rho$. Thus, it is not possible for the data-seller to offer partial information to the low type without inducing a deviation from the high type. Assume now that $\nu > \sqrt{\rho} - \rho$. In this case, there exists $\pi_2^L \in (0, 1]$ such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$ if and only if

$$\theta_L \le \frac{(1-\nu-\rho)[c \cdot \nu + \rho]\theta_H + [(\nu+\rho)^2 - \rho](1-c)}{[c(1-2\nu-\rho)+\nu](\nu+\rho)}$$

where

$$\pi_2^L = \frac{(1-\nu-\rho)\,[\nu c + \rho]\,\theta_H + \left[(\nu+\rho)^2 - \rho\right](1-c) - [c(1-2\nu-\rho)+\nu]\,(\nu+\rho)\theta_L}{(1-c)\,[(\nu+\rho)^2 - \rho - \nu(1-\nu-\rho)\theta_H + (\nu+\rho)(1-2\nu-\rho)\theta_L]}. \tag{B.3}$$

Thus, if the low type is sufficiently uncertain about the state, the data-seller is able to provide supplemental information to the low type without attracting the high type. Otherwise, the low type observes no supplemental information. ■

**Lemma B.3** *Assume $c \in (1, 2)$ and that beliefs are strictly congruent. In an optimal menu, the data-seller offers no information to $\theta_L$ if $\nu \ge \sqrt{\rho} - \rho$ and partial information otherwise.*

**Proof. Lemma B.3.** The proof is analogous to the proof of Lemma B.2 ■

**Proof. Proposition III.3** This proof is contained in Lemma B.2 and Lemma B.3. ■

**Proof. Lemma III.4.** In an optimal menu in which the low type observes partial information, (B.3) defines the optimal $E^2$. First, if data-buyers' private information is negatively correlated and they have coordination incentives, $c \in \left(\frac{1}{2}, 1\right)$ and $\nu > \sqrt{\rho} - \rho$. In this case, the sign of $\frac{\partial \pi_2^L}{\partial c}$ depends on the sign of

$$(\nu+\rho)^2 - \rho - \nu(1-\nu-\rho)\theta_H + (1-2\nu-\rho)(\nu+\rho)\theta_L$$

which is positive for all $\theta_L > \theta_H$ and $\nu > \sqrt{\rho} - \rho$. Hence, the precision of $E^L$ increases in $c$. As $c$ increases, data-buyers incentives to coordinate decrease. Thus, the precision of $E^L$ increases as the incentives to coordinate decrease. Similarly, the sign of $\frac{\partial \pi_2^L}{\partial \nu}$ depends on the sign of

$$(\nu + \rho)(\theta_H \rho(2 - \nu - \rho) - \theta_L(1 - \rho)(\nu + \rho) + \nu - \rho) + (1 - \theta_H)\rho$$

which is positive for all $\theta_L > \theta_H$, $\nu > \sqrt{\rho} - \rho$ and $c < 1$. Hence, the precision of $E^L$ increases in $\nu$. An increase in $\nu$ decreases the correlation between their private information. Thus, the precision of $E^L$ decreases in the correlation of private information.

Second, if data-buyers' private information is positively correlation and they have anti-coordination incentives, $c \in (1, 2)$ and $\nu < \sqrt{\rho} - \rho$. Analogously, it is straightforward to show that $\frac{\partial \pi_2^L}{\partial c} > 0$ and $\frac{\partial \pi_2^L}{\partial \nu} < 0$. ■

## B.2 Proofs with N data-buyers

**Lemma B.4** *The value of experiment $E^n$ is increasing in its precision if*

$$c < 1 \text{ and } c\left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq \left\lceil \frac{N}{2} \right\rceil \text{ or } c > 1 \text{ and } \left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq c \left\lceil \frac{N}{2} \right\rceil.$$

**Proof. Lemma B.4.** Consider the case in which data-buyer $i$ acquires experiment $E^L$. $V(E^L, \theta)$ is increasing in $\pi_1^L$ if

$$\Lambda_1^\theta \geq \max\left\{ -\pi_1^L \cdot \frac{\partial \Lambda_1^\theta}{\partial \pi_1^L}, (1 - \pi_1^L)\frac{\partial \Lambda_1^\theta}{\partial \pi_1^L} \right\}. \tag{B.4}$$

A sufficient but not necessary condition for (B.4) is

$$\Lambda_1^\theta \geq \max\left\{ -\frac{\partial \Lambda_1^\theta}{\partial \pi_1^L}, \frac{\partial \Lambda_1^\theta}{\partial \pi_1^L} \right\} = \mid \frac{\partial \Lambda_1^\theta}{\partial \pi_1^L} \mid \tag{B.5}$$

since $\pi_1^L \in [0, 1]$. $\Lambda_1^\theta$ depends on the distribution of the Conway-Maxwell-Binomial random variable, $\kappa_{-i}^1$. The distribution function of a Conway-Maxwell-Binomial

$(n, p, \nu)$ random variable is given by

$$F(k; n, p, \nu) = \frac{\sum_{\ell=0}^{k} p^{\ell}(1-p)^{n-\ell}\binom{n}{\ell}^{\nu}}{S(p, \nu)}.$$

where $S(p, \nu) = \sum_{k=0}^{n} p^{k}(1-p)^{n-k}\binom{n}{k}^{\nu}$ is a normalizing constant. Given that $\pi_1^L$ only affects $p_{\omega_1,\theta}$, using the chain rule, we have:

$$\frac{\partial F(k; n, p_{\omega_1,\theta}, \nu_{\omega_1,\theta})}{\partial \pi_1^L} = \frac{\partial F(k; n, p_{\omega_1,\theta}, \nu_{\omega_1,\theta})}{\partial p_{\omega_1,\theta}} \frac{\partial p_{\omega_1,\theta}}{\partial \pi_1^L}$$

where $\frac{\partial p_{\omega_1,\theta}}{\partial \pi_1^L} = \mathbb{P}(\theta_j = \theta_L | \theta)$ and

$$\frac{\partial F(k; n, p, \nu)}{\partial p_{\omega_1,\theta}}$$
$$= \frac{\sum_{\ell=1}^{k}\binom{n}{\ell}^{\nu} p^{\ell-1}(1-p)^{n-\ell-1}(\ell - np) - \left[\sum_{\ell=1}^{n}\binom{n}{\ell}^{\nu} p^{\ell-1}(1-p)^{n-\ell-1}(\ell - np)\right] F(k; n, p, \nu)}{S(p, \nu)}.$$

Note that

$$\begin{aligned}
\frac{\partial F(k; n, p, \nu)}{\partial p_{\omega_1,\theta}} &\leq (k - np)\,\frac{\sum_{\ell=1}^{k}\binom{n}{\ell}^{\nu} p^{\ell-1}(1-p)^{n-\ell-1}}{S(p, \nu)} + np\frac{\sum_{\ell=1}^{n}\binom{n}{\ell}^{\nu} p^{\ell-1}(1-p)^{n-\ell-1}}{S(p, \nu)} \\
&\leq (k - np) + np = k.
\end{aligned}$$

Then, $\frac{\partial \Lambda_1^{\theta}}{\partial \pi_1^L}$ is given by:

$$\left|\frac{\partial \Lambda_1^{\theta}}{\partial \pi_1^L}\right| = \mathbb{P}(\theta_j = \theta_L|\theta)\left|(1-c)\frac{\partial F(\lceil \frac{N}{2} \rceil - 1; \cdot)}{\partial p_{\omega_1,\theta}}\right| \leq \left|(1-c)\lceil \frac{N}{2} \rceil\right|$$

where the inequality holds since $\mathbb{P}(\theta_j = \theta_L|\theta) \in [0, 1]$. Moreover, $\Lambda_1^{\theta} \geq \min\{1, c\}$. Then, (B.5) holds for all $\pi_1^L \in [0, 1]$ (or $p_{\omega_1,\theta}$) if

$$\begin{aligned}
c \geq (1-c)\lceil \frac{N}{2} \rceil \text{ if } c < 1 &\Leftrightarrow c\left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq \left\lceil \frac{N}{2} \right\rceil \\
1 \geq (c-1)\lceil \frac{N}{2} \rceil \text{ if } c > 1 \ &\Leftrightarrow \left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq c\left\lceil \frac{N}{2} \right\rceil.
\end{aligned}$$

Analogously, I can show that $V_{\mathcal{M}}(E^H, \theta)$ increases in $\pi_1^H$ under the same sufficient conditions and that $V_{\mathcal{M}}(E^n, \theta)$ increases in $\pi_2^n$. ■

**Proposition B.1** *Assume that payoffs satisfy (3.3). The high type observes $\overline{E}$ and*

1. *If $\nu \leq 1$, the low type observes no information if beliefs are strictly congruent and partial information if beliefs are strictly non-congruent.*
2. *If $\nu > 1$, the low type observes partial information if beliefs are strictly congruent and $\theta_L < \tilde{\theta}$ and no information otherwise.*

**Proof. Proposition B.1.** When beliefs are congruent, $\pi_1^L = 1$ and $\pi_2^L$ such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$. First, the value that the low type attaches to experiment $E^L$ is given by:

$$V_{\mathcal{M}}(E^L, \theta_L) = (1 - \theta_L)\pi_2^L \Lambda_2^{\theta_L}$$

where $\kappa_{-i}^1|(\omega_2, \theta_L)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_2,\theta_L} = \eta_L(1-\pi_2^L)$ and $\nu$. Second, the value that the high type attaches to experiment $E^L$ is given by:

$$V_{\mathcal{M}}(E^L, \theta_H) = (1 - \theta_H)\pi_2^L \Lambda_2^{\theta_H}$$

where $\kappa_{-i}^1|(\omega_2, \theta_H)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_2,\theta_H} = (1-\eta_H)(1-\pi_2^L)$ and $\nu$. Denote by $F(k; N-1, p_{\omega,\theta_i}, \nu)$ the distribution function of a CMB distribution with these parameters. Given that the distribution of $\kappa_{-i}^1|(\omega_2, \theta_L)$ and $\kappa_{-i}^1|(\omega_2, \theta_H)$ share two of those parameters, I simplify the notation to $F(k; p_{\omega,\theta_i})$.

It is trivial the incentive compatibility constraint of the high type binds if $\pi_2^L = 0$. Assume now that $\pi_2^L \in (0, 1]$ and consider first the case in which data-buyers' private information is positively correlated or $\nu < 1$. In this case, $\eta_L > 1 - \eta_H$ which implies that $p_{\omega_2,\theta_L} > p_{\omega_2,\theta_H}$ and

$$V_{\mathcal{M}}(E^L, \theta_L) < (1-\theta_H)\pi_2^L \left[c + (1-c)F\left(N - \lceil\frac{N}{2}\rceil - 1; (1-\eta_H)(1-\pi_2^L)\right)\right] = V(E^1, \theta_H)$$

where the inequality holds since $\theta_H < \theta_L$, $\eta_L > 1 - \eta_H$ and $c < 1$. Then, $V_{\mathcal{M}}(E^L, \theta_L) < V_{\mathcal{M}}(E^L, \theta_H)$ which implies that the incentive-compatibility constraint of the high type is violated. Thus, it is not possible for the seller to offer information to the low type. Analogously, we reach the same conclusion for $\nu = 1$.

In contrast, if data-buyer types are negatively correlated ($\nu > 1$) or $\eta_L < 1 - \eta_H$, we have that $(1-\theta_L) < (1-\theta_H)$ but

$$c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; \eta_L(1-\pi_2^L)\right) > c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; (1-\eta_H)(1-\pi_2^L)\right).$$

Define

$$\Delta V_{\mathcal{M}}(\pi_2^L) := \frac{V_{\mathcal{M}}(E^L, \theta_L) - V_{\mathcal{M}}(E^L, \theta_H)}{\pi_2^1}.$$

Note that $\Delta V_{\mathcal{M}}(1) < 0$ by the definition of high type and

$$\lim_{\pi_2^L \to 0} \Delta V_{\mathcal{M}}(\pi_2^L) \geq 0 \text{ iff } \theta_L \leq 1 - \frac{(1-\theta_H)\left(c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; (1-\eta_H)\right)\right)}{c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; \eta_L\right)}. \quad \text{(B.6)}$$

Then, the Intermediate value theorem implies that there exists at least one $\pi_2^L \in (0,1)$ such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$. Thus, the seller is able to offer partial information to the low type if and only if (B.6) holds.

Consider now the case in which beliefs are non-congruent. In this case, $\pi_2^L = 1$ and $\pi_1^L$ is such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$. First, the value that the low type attaches to experiment $E^L$ is given by:

$$V_{\mathcal{M}}(E^L, \theta_L) = \theta_L \pi_1^L \Lambda_1^{\theta_L}$$

where $\kappa_{-i}^1 | (\omega_1, \theta_L)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_1, \theta_L} = \eta_L \pi_1^L + (1-\eta_L)$ and $\nu$. Second, the high type attaches a value to experiment $E^L$ given by

$$V_{\mathcal{M}}(E^L, \theta_H) = \max\{0, (1-\theta_H)\Lambda_2^{\theta_H} - \theta_H(1-\pi_1^L)\Lambda_1^{\theta_H}\}$$

where $\kappa_{-i}^1 | (\omega_1, \theta_H)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_1, \theta_L} = \eta_H + (1-\eta_H)\pi_1^L$ and $\nu$ and $\kappa_{-i}^1 | (\omega_2, \theta_H)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_2, \theta_H} = 0$ and $\nu$.[2] Note

[2]Then, $\mathbb{P}(\kappa_{-i}^1 \leq k | \omega_2) = 1$ for all $k \geq 0$.

that $V_{\mathcal{M}}(E^L, \theta_H)$ is a continuous and increasing function of $\pi_1^L$, $V_{\mathcal{M}}(E^L, \theta_H) < 0$ if $\pi_1^L = 0$ and $V_{\mathcal{M}}(E^L, \theta_H) > 0$. Then, the intermediate value theorem implies that there exists $\hat{\pi}_1^L \in (0, 1)$ such that $V_{\mathcal{M}}(E^L, \theta_H) \geq 0$ for all $\pi_1^L \geq \hat{\pi}_1^L$. Define $\Delta V_{\mathcal{M}}(\pi_1^L) = V_{\mathcal{M}}(E^L, \theta_L) - V_{\mathcal{M}}(E^L, \theta_H)$. Note that $\Delta V_{\mathcal{M}}(\pi_1^L)$ is also a continuous function of $\pi_1^L$, $\Delta V_{\mathcal{M}}(\hat{\pi}_1^L) \geq 0$ and $\Delta V_{\mathcal{M}}(1) = (\theta_L - (1 - \theta_H)) < 0$ by the definition of high type. Thus, the intermediate value theorem implies that there exists $\pi_1^L \in (\hat{\pi}_1^L, 1)$ such that $\Delta V_{\mathcal{M}}(\pi_1^L) = 0$. ■

**Proposition B.2** *Assume that payoffs satisfy (3.4). The high type observes $\overline{E}$ and*

1. *If $\nu \geq 1$, the low type observes no information if beliefs are strictly congruent and partial information if beliefs are strictly non-congruent.*
2. *If $\nu < 1$, the low type observes partial information if beliefs are strictly congruent and $\theta_L \leq \tilde{\theta}$ and no information otherwise.*

**Proof. Proposition B.2.** The proof is analogous to the proof of Proposition B.1. ■

**Proof. Proposition III.4** This proof is contained in Proposition B.1 and Proposition B.2. ■

**Proof. Proposition III.5** This proof is contained in Proposition B.1 and Proposition B.2. ■

# APPENDIX C

# Appendix for Chapter IV

## C.1 Proofs

**Proof. Lemma IV.1.** Consider first the case in which $\alpha \in \left[0, \frac{1}{4}\right]$. The distribution over posterior value estimates induced by the distribution over buyer types and the distribution over posterior quality estimates is given by

$$\overline{G}_1(s) = \begin{cases} 0 & ifs < a - \alpha \\ \frac{1}{2}\frac{s+\alpha-a^2}{1-a^2} & if\ s\,[a-\alpha, a+\alpha) \\ \frac{s-a^2}{1-a^2} & if\ s\,[a+\alpha, 1-\alpha) \\ \frac{1}{2}+\frac{1}{2}\frac{s-\alpha-a^2}{1-a^2} & if\ s\,[1-\alpha, 1+\alpha) \\ 1 & if\ s \geq 1+\alpha \end{cases}$$

Next, I show that the seller's optimal price $p^*(\alpha, a) \in \{a-\alpha, a+\alpha, \frac{1}{2}\}$. Define by $\Pi_1(s)$ the seller's expected profits of setting a price of $s$ given the distribution over posterior value estimates characterized by $\overline{G}_1$, $\Pi_1(s) := s(1-\overline{G}_1(s-))$. First, note that $\Pi_1(a-\alpha) \geq \Pi_1(s)$ for all $s < a-\alpha$, $\Pi_1(a+\alpha) \geq \Pi_1(s)$ for all $s \in [a-\alpha, a+\alpha)$ and $\Pi_1(a+\alpha) \geq \Pi_1(s)$ for all $s \geq 1-\alpha$. Second, for all $s \in [a+\alpha, 1-\alpha)$, $\max\{\Pi_1\left(\frac{1}{2}\right), \Pi_1(a+\alpha)\} \geq \Pi_1(s)$. Hence, all prices except $a-\alpha$, $a+\alpha$ and $\frac{1}{2}$

are dominated for the seller and

$$p^*(\alpha,a) = \begin{cases} a-\alpha & \text{if and only if } \Pi_1(a-\alpha) \geq \max\{\Pi_1(a+\alpha), \Pi_1\left(\frac{1}{2}\right)\} \\ a+\alpha & \text{if and only if } \Pi_1(a+\alpha) > \max\{\Pi_1(a-\alpha) \text{ and } \Pi_1(a+\alpha) \geq \Pi_1\left(\frac{1}{2}\right) \\ \frac{1}{2} & \text{if and only if } \Pi_1\left(\frac{1}{2}\right) > \max\{\Pi_1(a-\alpha), \Pi_1(a+\alpha)\} \end{cases}$$

Analogously, it is possible to show that $p^*(\alpha, a) \in \{a+\alpha, \frac{1}{2}, \frac{1+\alpha}{2}\}$ when $\alpha \in \left(\frac{1}{2}, \frac{1}{2}\right)$ and $a \in [0, 1-2\alpha)$ and $p^*(\alpha, a) \in \{a+\alpha, \frac{1+\alpha}{2}\}$ when $\alpha \in \left(\frac{1}{2}, \frac{1}{2}\right)$ and $a \in [1-2\alpha, \frac{1}{2}]$. ■

**Proof. Proposition IV.1.** From the designer's perspective, the buyer's expected payoff is depends on the distribution of types and the distribution of posterior value estimates characterized by $a \in \left[0, \frac{1}{2}\right]$ and it is given by

$$\mathcal{U}(\alpha,a) = \frac{1}{2}\int_{p^*(\alpha,a)+\alpha}^{1} (s-\alpha-p^*(\alpha,a))\mathrm{d}G_1(s) + \frac{1}{2}\int_{p^*(\alpha,a)-\alpha}^{1} (s+\alpha-p^*(\alpha,a))\mathrm{d}G_1(s).$$

where $G_1(s) = \mathbf{1}_{[a,1]}(s)\frac{s-a^2}{1-a^2} + \mathbf{1}_{[1,\infty)}(s)$. Note that $\mathcal{U}(\alpha,a)$ is a decreasing function of the price $p^*(\alpha,a)$ and that $p^*(\alpha,a)$ is weakly increasing in $a$. Moreover, the effect of $a$ on $\mathcal{U}(\alpha,a)$ is driven by the effect on the price $p^*(\alpha,a)$, implying that it is optimal for consumers to choose $a$ which induces the seller to choose the smallest price.

Consider the lowest price the seller is willing to set, $\underline{p}^*(a|\alpha) := \min\{p : p \in P^*(\alpha,a) \text{ for a fixed } \alpha\}$ given the information structure characterized by $a$ and the distribution of types characterized by $\alpha$. Define $\underline{a}(\alpha,p)$ as the minimum $a \in \left[0, \frac{1}{2}\right]$ such that it is optimal for the seller to set a price of $p$ given $\alpha$. The optimal $a^*(\alpha)$ is given by $\underline{a}(\alpha,p)$ such that $p = \underline{p}(a|\alpha)$. ■

**Proof. Proposition IV.2.** The proof of this result is analogous to the proof of Proposition IV.1. ■

# BIBLIOGRAPHY

# BIBLIOGRAPHY


Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2019): "Learning from reviews: The selection effect and the speed of learning," Discussion paper, working paper.

Admati, A. R., and P. Pfleiderer (1986): "A monopolistic market for information," *Journal of Economic Theory*, 39(2), 400–438.

——— (1990): "Direct and indirect sale of information," *Econometrica: Journal of the Econometric Society*, pp. 901–928.

Ali, S. N., G. Lewis, and S. Vasserman (2020): "Voluntary disclosure and personalized pricing," in *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 537–538.

Aliprantis, C. D., and K. C. Border (2013): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin and Heidelberg GmbH & Company KG.

Angeletos, G.-M., and A. Pavan (2007): "Efficient use of information and social value of information," *Econometrica*, 75(4), 1103–1142.

Armstrong, M., and J. Vickers (2015): "Which demand systems can be generated by discrete choice?," *Journal of Economic Theory*, 158, 293–307.

Armstrong, M., and J. Zhou (2019): "Consumer information and the limits to competition," .

Bass, R. F. (2011): *Regular conditional probabilities*p. 380–381, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Battigalli, P. (1996): "Strategic independence and perfect Bayesian equilibria," *Journal of Economic Theory*, 70(1), 201–234.

Benoit, J.-P., and J. Dubra (2006): "Information revelation in auctions," *Games and Economic Behavior*, 57(2), 181–205.

Bergemann, D., A. Bonatti, and A. Smolin (2018): "The design and price of information," *American Economic Review*, 108(1), 1–48.

Bergemann, D., B. Brooks, and S. Morris (2015): "The limits of price discrimination," *American Economic Review*, 105(3), 921–57.

Bergemann, D., T. Heumann, and S. Morris (2015): "Information and volatility," *Journal of Economic Theory*, 158, 427–465.

Bergemann, D., and S. Morris (2013): "Robust predictions in games with incomplete information," *Econometrica*, 81(4), 1251–1308.

——— (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57(1), 44–95.

Bergemann, D., and D. Ozmen (2006): "Optimal pricing with recommender systems," in *Proceedings of the 7th ACM conference on Electronic commerce*, pp. 43–51.

Bergemann, D., and J. Välimäki (2006): "Dynamic pricing of new experience goods," *Journal of Political Economy*, 114(4), 713–743.

Bernhardt, D., and B. Taub (2015): "Learning about common and private values in oligopoly," *The RAND Journal of Economics*, 46(1), 66–85.

Bonatti, A., G. Cisternas, and J. Toikka (2017): "Dynamic oligopoly with incomplete information," *The Review of Economic Studies*, 84(2), 503–546.

Calzolari, G., and A. Pavan (2006): "On the optimality of privacy in sequential contracting," *Journal of Economic theory*, 130(1), 168–204.

Che, Y.-K., and J. Hörner (2018): "Recommender systems as mechanisms for social learning," *The Quarterly Journal of Economics*, 133(2), 871–925.

Choné, P., and L. Linnemer (2020): "Linear demand systems for differentiated goods: Overview and user's guide," *International Journal of Industrial Organization*, p. 102663.

Clarke, R. N. (1983): "Collusion and the incentives for information sharing," *The Bell Journal of Economics*, pp. 383–394.

Crémer, J., and R. P. McLean (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica: Journal of the Econometric Society*, pp. 1247–1257.

Eliaz, K., and F. Forges (2015): "Information disclosure to Cournot duopolists," *Economics Letters*, 126, 167–170.

Elliott, M., A. Galeotti, and A. Koh (2020): "Market segmentation through information," .

Eső, P., and B. Szentes (2007): "Optimal information disclosure in auctions and the handicap auction," *The Review of Economic Studies*, 74(3), 705–731.

Fried, D. (1984): "Incentives for information production and disclosure in a duopolistic environment," *The Quarterly Journal of Economics*, 99(2), 367–381.

Fung, B. (2020): "Congress grilled the CEOs of Amazon, Apple, Facebook and Google. Here are the big takeaways," *CNN*, Available at https:https://www.cnn.com/2020/07/29/tech/tech-antitrust-hearing-ceos/index.html, Date Published: Jul 29, 2020, Date Accessed: Jul 29, 2020.

Gal-Or, E. (1985): "Information sharing in oligopoly," *Econometrica: Journal of the Econometric Society*, pp. 329–343.

——— (1986): "Information transmission—Cournot and Bertrand equilibria," *The Review of Economic Studies*, 53(1), 85–92.

Green, D. (2018): "One of Amazon's most powerful tools is putting a target on its back, and it could turn into a problem for the company," *Business Insider*, Available at https://www.businessinsider.com/amazon-scrutinized-third-party-seller-data-use-2018-12, Date Published: Dec 1, 2018, Date Accessed: Jul 10, 2020.

Harrington, Jr, J. E. (1995): "Experimentation and learning in a differentiated-products duopoly," *Journal of Economic Theory*, 66(1), 275–288.

Ichihashi, S. (2020): "Online privacy and information disclosure by consumers," *American Economic Review*, 110(2), 569–95.

Kamenica, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

Kastl, J., M. Pagnozzi, and S. Piccolo (2018): "Selling information to competitive firms," *The RAND Journal of Economics*, 49(1), 254–282.

Keller, G., and S. Rady (2003): "Price dispersion and learning in a dynamic differentiated-goods duopoly," *RAND Journal of Economics*, pp. 138–165.

Kirby, A. J. (1988): "Trade associations as information exchange mechanisms," *The RAND Journal of Economics*, pp. 138–146.

Kolotilin, A., and A. Wolitzky (2020): "Assortative information disclosure," .

Krähmer, D. (2020): "Information disclosure and full surplus extraction in mechanism design," *Journal of Economic Theory*, 187, 105020.

Lardieri, A. (2019): "European Union Opens Antitrust Probe Into Amazon," *World Report US News*, Available at https://www.usnews.com/news/world-report/articles/2019-07-17/european-union-opens-antitrust-probe-into-amazon, Date Published: Jul 17, 2019, Date Accessed: Sep 16, 2019.

Li, H., and X. Shi (2017): "Discriminatory information disclosure," *American Economic Review*, 107(11), 3363–85.

Li, L. (1985): "Cournot oligopoly with information sharing," *The RAND Journal of Economics*, pp. 521–536.

Li, W. (2020): "Using Information to Amplify Competition," *arXiv preprint arXiv:2010.05342*.

Libgober, J., and X. Mu (2018): "Informational Robustness in Intertemporal Pricing," *Available at SSRN 2892096*.

Luco, F. (2019): "Who benefits from information disclosure? The case of retail gasoline," *American Economic Journal: Microeconomics*, 11(2), 277–305.

Mailath, G. J. (1989): "Simultaneous signaling in an oligopoly model," *The Quarterly Journal of Economics*, 104(2), 417–427.

Mathevet, L., J. Perego, and I. Taneva (2020): "On information design in games," *Journal of Political Economy*, 128(4), 1370–1404.

Mirman, L. J., L. Samuelson, and E. E. Schlee (1994): "Strategic information manipulation in duopolies," *Journal of Economic Theory*, 62(2), 363–384.

Mirman, L. J., L. Samuelson, and A. Urbano (1993): "Duopoly signal jamming," *Economic Theory*, 3(1), 129–149.

Novshek, W., and H. Sonnenschein (1982): "Fulfilled expectations Cournot duopoly with information acquisition and release," *The Bell Journal of Economics*, pp. 214–218.

Raith, M. (1996): "A general model of information sharing in oligopoly," *Journal of economic theory*, 71(1), 260–288.

Ravid, D., A.-K. Roesler, and B. Szentes (2019): "Learning before trading: on the inefficiency of ignoring free information," *Available at SSRN 3317917*.

ROCHET, J.-C., AND L. A. STOLE (2003): *The Economics of Multidimensional Screening*vol. 1 of *Econometric Society Monographs*, p. 150–197. Cambridge University Press.

ROESLER, A.-K., AND B. SZENTES (2017): "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 107(7), 2072–80.

ROYDEN, H. L. (1968): *Real analysis*. Krishna Prakashan Media.

SAKAI, Y. (1986): "Cournot and Bertrand equilibria under imperfect information," *Journal of Economics*, 46(3), 213–232.

SHAPIRO, C. (1986): "Exchange of cost information in oligopoly," *The review of economic studies*, 53(3), 433–446.

STINCHCOMBE, M. B. (2011): "Correlated equilibrium existence for infinite games with type-dependent strategies," *Journal of Economic Theory*, 146(2), 638–655.

TANEVA, I. (2019): "Information design," *American Economic Journal: Microeconomics*, 11(4), 151–85.

TAYLOR, C. R. (2004): "Consumer privacy and the market for customer information," *RAND Journal of Economics*, pp. 631–650.

UI, T. (2016): "Bayesian Nash equilibrium and variational inequalities," *Journal of Mathematical Economics*, 63, 139–146.

UI, T., AND Y. YOSHIZAWA (2015): "Characterizing social value of information," *Journal of Economic Theory*, 158, 507–535.

VIVES, X. (1984): "Duopoly information equilibrium: Cournot and Bertrand," *Journal of economic theory*, 34(1), 71–94.

——— (1990): "Trade association disclosure rules, incentives to share information, and welfare," *the RAND Journal of Economics*, pp. 409–430.

WATSON, J. (2016): "Perfect bayesian equilibrium: General definitions and illustrations," *Unpublished, University of California at San Diego*, 133.