Université de Montréal

ESSAYS IN ECONOMICS OF INFORMATION

par
Catherine Gendron-Saulnier

Département de Sciences Économiques
Faculté des Arts et Sciences

Thèse présentée en vue de l'obtention du grade
Philosophiæ Doctor (Ph.D.) en Sciences Économiques

Avril, 2015

Université de Montréal
Faculté des Études Supérieures et Postdoctorales

Cette thèse intitulée:

Essays in Economics of Information

présentée par:

Catherine Gendron-Saulnier

a été évaluée par un jury composé des personnes suivantes:

Lars Ehlers	Président-rapporteur
Sidartha Gordon	Directeur de recherche
Nicolas Klein	Co-directeur de recherche
Marc Santugini	Co-directeur de recherche
Sean Horan	Membre du jury
Joel Sobel	Examinateur externe
Vincent Larivière	Représentant du doyen de la FAS

Thèse acceptée le 20 août 2015

Résumé

Cette thèse est une collection de trois articles en économie de l'information. Le premier chapitre sert d'introduction et les Chapitres 2 à 4 constituent le coeur de l'ouvrage.

Le Chapitre 2 porte sur l'acquisition d'information sur l'Internet par le biais d'avis de consommateurs. En particulier, je détermine si les avis laissés par les acheteurs peuvent tout de même transmettre de l'information à d'autres consommateurs, lorsqu'il est connu que les vendeurs peuvent publier de faux avis à propos de leurs produits. Afin de comprendre si cette manipulation des avis est problématique, je démontre que la plateforme sur laquelle les avis sont publiés (e.g. TripAdvisor, Yelp) est un tiers important à considérer, autant que les vendeurs tentant de falsifier les avis. En effet, le design adopté par la plateforme a un effet indirect sur le niveau de manipulation des vendeurs. En particulier, je démontre que la plateforme, en cachant une partie du contenu qu'elle détient sur les avis, peut parfois améliorer la qualité de l'information obtenue par les consommateurs. Finalement, le design qui est choisi par la plateforme peut être lié à la façon dont elle génère ses revenus. Je montre qu'une plateforme générant des revenus par le biais de commissions sur les ventes peut être plus tolérante à la manipulation qu'une plateforme qui génère des revenus par le biais de publicité.

Le Chapitre 3 est écrit en collaboration avec Marc Santugini. Dans ce chapitre, nous étudions les effets de la discrimination par les prix au troisième degré en présence de consommateurs non informés qui apprennent sur la qualité d'un produit par le biais de son prix. Dans un environnement stochastique avec deux segments de marché, nous démontrons que la discrimination par les prix peut nuire à la firme et être bénéfique pour les consommateurs. D'un côté, la discrimination par les prix diminue l'incertitude à laquelle font face les consommateurs, c.-à-d., la variance des croyances postérieures est plus faible avec discrimination qu'avec un prix uniforme. En effet, le fait d'observer deux prix (avec discrimination) procure plus d'information aux consommateurs, et ce, même si individuellement chacun de ces prix est moins informatif que le prix uniforme. De l'autre côté, il n'est pas toujours optimal pour la firme de faire de la discrimination par les prix puisque la présence de consommateurs non informés lui donne une incitation à s'engager dans du signaling. Si l'avantage procuré par la flexibilité de fixer deux prix différents est

contrebalancé par le coût du signaling avec deux prix différents, alors il est optimal pour la firme de fixer un prix uniforme sur le marché.

Finalement, le Chapitre 4 est écrit en collaboration avec Sidartha Gordon. Dans ce chapitre, nous étudions une classe de jeux où les joueurs sont contraints dans le nombre de sources d'information qu'ils peuvent choisir pour apprendre sur un paramètre du jeu, mais où ils ont une certaine liberté quant au degré de dépendance de leurs signaux, avant de prendre une action. En introduisant un nouvel ordre de dépendance entre signaux, nous démontrons qu'un joueur préfère de l'information qui est la plus dépendante possible de l'information obtenue par les joueurs pour qui les actions sont soit, compléments stratégiques et isotoniques, soit substituts stratégiques et anti-toniques, avec la sienne. De même, un joueur préfère de l'information qui est la moins dépendante possible de l'information obtenue par les joueurs pour qui les actions sont soit, substituts stratégiques et isotoniques, soit compléments stratégiques et anti-toniques, avec la sienne. Nous établissons également des conditions suffisantes pour qu'une structure d'information donnée, information publique ou privée par exemple, soit possible à l'équilibre.

Mots-Clés: Apprentissage, Acquisition d'Information, Structure d'Information Endogène, Design d'Information, Jeux de Signaling Stochastique, Discrimination par les Prix au Troisième Degré, Complémentarités Stratégiques, Commerce Électronique, Plateforme Internet

Abstract

This thesis is a collection of three essays in economics of information. Chapter 1 is a general introduction and Chapters 2 to 4 form the core of the thesis.

Chapter 2 analyzes information dissemination on the Internet. Online platforms such as Amazon, TripAdvisor or Yelp are now key sources of information for modern consumers. The proportion of consumers consulting online reviews prior to purchasing a good or a service has grown persistently. Yet, sellers have been accused of hiring shills to post fake reviews about their products. This raises the question: Does the presence of shills make reviews less informative? I show that the answers to this question depend on the way the platform presents and summarizes reviews on its website. In particular, I find that withholding information by garbling the reviews benefits information dissemination by inducing the seller to destroy less information with manipulation. Next, I show that the platform's choice regarding how to present reviews hinges on its revenue source. Indeed, a platform that receives sales commissions optimally commits to publishing information differently from a platform that receives revenues from advertisements or from subscription fees. Incidentally, such platforms have contrasting impacts on the amount of information that is transmitted by reviews.

Chapter 3 is co-authored with Marc Santugini. In this chapter, we study the impact of third-degree price discrimination in the presence of uninformed buyers who extract noisy information from observing prices. In a noisy learning environment, it is shown that price discrimination can be detrimental to the firm and beneficial to the consumers. On the one hand, discriminatory pricing reduces consumers' uncertainty, i.e., the variance of posterior beliefs upon observing prices is reduced. Specifically, observing two prices under discriminatory pricing provides more information than one price under uniform pricing even when discriminatory pricing reduces the amount of information contained in each price. On the other hand, it is not always optimal for the firm to use discriminatory pricing since the presence of uninformed buyers provides the firm with the incentive to engage in noisy price signaling. Indeed, if the benefit from price flexibility (through discriminatory pricing) is offset by the cost of signaling quality through two distinct prices, then it is optimal to integrate markets and to use uniform pricing.

Finally, Chapter 4 is co-authored with Sidartha Gordon. In this chapter, we study a class of games where players face restrictions on how much information they can obtain on a common payoff relevant state, but have some leeway in covertly choosing the dependence between their signals, before simultaneously choosing actions. Using a new stochastic dependence ordering between signals, we show that each player chooses information that is more dependent on the information of other players whose actions are either isotonic and complements with his actions or antitonic and substitutes with his actions. Similarly, each player chooses information that is less dependent on the information of other players whose actions are antitonic and complements with his actions or isotonic and substitutes with his actions. We then provide sufficient conditions for information structures such as public or private information to arise in equilibrium.

Keywords: Learning, Information Acquisition, Endogenous Information Structure, Information Design, Noisy-signaling Game, Third-degree Price Discrimination, Complementarities, Online Commerce, Internet Platform

Table of Contents

List of Figures

List of Tables

Remerciements

Mes remerciements vont d'abord à mes directeurs: Sidartha Gordon, Nicolas Klein et Marc Santugini. C'est grâce à leurs précieux conseils, leur patience et leur immense dévouement que j'ai pu progresser tout au long des 6 années qu'aura duré mon doctorat. Je les remercie pour les heures qu'ils ont passées à discuter de ma recherche, à lire mes articles et les efforts qu'ils ont mis pour m'appuyer lors de ma recherche d'emploi.

Je tiens aussi à remercier l'ensemble des professeurs du département de sciences économiques de l'Université de Montréal, particulièrement Marine Carasco, Rui Castro, Deniz Dizdar, Lars Ehlers, Sylvia Gonçalves, Sean Horan, Michel Poitevin, et Yves Richelle. Mes remerciements vont également au personnel administratif, avec une mention spéciale pour Mélanie Chartrand qui a toujours su avoir le mot juste (et pas nécessairement celui que je voulais entendre) pour me motiver.

Je remercie ma collègue et fidèle amie Anabelle pour toutes les discussions de nature académique, mais surtout pour ces discussions anodines sur les couleurs tendance de vernis à ongles. Je dois aussi remercier mes autres collègues Juste, Max, Vincent De Paul, Guy, Louis-Philippe, Hector, Idrissa, Mehrab, Vincent et Jean-François pour leur soutien.

Pour ce qui est du soutien financier, je dois remercier le Fonds Québécois de Recherche Société et Culture (FQRSC), le Département de sciences économiques de l'Université de Montréal, le Centre Inter-universitaire de Recherche en économie Quantitative (CIREQ), le Département d'économie appliquée du HEC Montréal, ainsi que la Faculté des études supérieures et postdoctorales de l'Université de Montréal.

Je remercie également les professeurs rencontrés lors de conférences, de séminaires, ou de cours et qui m'ont prodigué de précieux conseils, notamment Frederic Koessler, Johannes Hörner, Christopher Cotton, Nicolas Sahuguet, Artyom Shneyerov, Ming Li, et Rohan Dutta. Merci à Joel Sobel d'avoir accepté de siéger à titre d'examinateur externe sur mon comité. Merci aussi à la Paris School of Economics et Sciences-Po Paris de m'avoir permis d'effectuer un séjour de recherche chez eux.

Enfin, je tiens à remercier ma famille et mes amis pour leur indulgence et leur soutien indéfectible. Je remercie spécialement mes parents pour m'avoir offert écoute et réconfort et surtout pour m'avoir encouragée à toujours dépasser mes limites. Finalement, mes derniers remerciements vont à mon conjoint, Jean-Simon, qui après toutes ces années, malgré les hauts et les bas (hebdomadaires) qui accompagnent le travail d'un chercheur, ne cesse de m'appuyer.

Chapter 1

Introduction

Our time can definitely be characterized as an information era. The large diffusion of broadband access to the Internet, the increasing popularity of social networks and the proliferation of smartphone users are, among others, some of the phenomena that promote this tremendous production of information. As anecdotal proof, "There will be more words on Twitter in the next two years than contained in all books ever printed."[1]

Means of digital communication and the growing number of information sources should enable economic agents to make better choices by giving them more (and better) opportunities to learn about the variables relevant to the decisions they need to make. Yet, new problems arise and old ones persist. According to Carl Hausman, an expert in media, society and ethics, some of the new problems created by the Digital Age are that the increased speed of communication makes it more likely for errors to occur, that plagiarism is difficult to define and to control, and that all the personal data scattered on the Internet make us vulnerable to security breaches and privacy invasion.[2]

From an economic viewpoint, there are two issues related to the acquisition of information and learning that are not new, but that our inter-connected world makes more acute. The first issue concerns the asymmetric distribution of information among economic agents and its impact on markets' outcomes and welfare. It is true that information asymmetries have always existed. But, diffusion of new technologies and the rise of the Big Data Era help to feed existing inequalities and to create new ones. The second issue concerns the necessity of choice in this vast pool of information resources. The question is not trivial as there are generally more information sources available than what the economic agents, constrained by time-money-ability, can actually learn from.

This thesis is a collection of three essays in economics of information that target these two topics. In particular, the aim of these essays is to analyze how economic contexts and different incentives shape the way individuals acquire, produce and use their information.

Chapter 2 and Chapter 3 form the first part of the thesis. In these chapters, I study the first issue just outlined which is the asymmetric distribution of information among economic agents. Information asymmetries are central to many economic situations and it is important to understand how market aggregates, such as prices or market shares,

[1]Rudder (2014), *Dataclysm: Who We Are (When We Think No One's Looking)*, p.59.
[2]Hausman, C. (2014) Retrieved on April 13, 2015, from: `http://carlhausman.com/2014/02/24/ethical-problems-in-the-digital-era-problems-we-didnt-even-know-we-had/`.

are affected by this phenomenon. Whether or not uninformed individuals can learn by observing market aggregates, what the quality of the information they retrieve is and how their well-being varies with the depth of asymmetries are some of the fundamental questions to address.

In this thesis, one relevant asymmetry that I study is the one that exists between consumers and firms. Indeed, consumers and firms are unequally informed and this gap keeps getting larger. As proof, this Big Data Era in which we live makes it easy for firms to accumulate tons of data on their customers and to carry out complex and extensive analysis of this tremendous amount of information. Chapter 2 and Chapter 3, while analyzing the functioning of markets with asymmetrically informed agents, focus particularly on theses asymmetries among firms and consumers. The aim is to understand how informative some observable economic variables are for consumers when some other protagonists of that same market can influence these variables. I study two particular instances of the phenomenon: in the first instance, I study situations where consumers learn about products from consumer-generated content, in the second, I study situations where consumers learn about products from their price.

More specifically, Chapter 2 is entitled *Can You Trust What You Read on the Internet? Designing Platforms to Deal with Shill Reviews* and examines learning from consumer-generated content. Prior to purchasing a good or a service, it is now common for consumers to consult user-reviews and ratings on online platforms such as Amazon, TripAdvisor or Yelp. Whether feedbacks on the experience of anonymous buyers actually inform consumers is a relevant question. Indeed, the anonymity that prevails on the Internet makes it easier to counterfeit one's identify and to hide one's conflict of interest. For instance, sellers can pay individuals to post fake positive reviews about their products on the platforms with the purpose of manipulating consumers' beliefs. Can consumers rely on these reviews then? In other words, does the presence of fake reviews make the reviews less informative? And if it is the case, can we do something about it? In Chapter 2, I look into theses questions by analyzing the informational design of these Internet platforms taking in consideration that sellers can manipulate reviews.

In particular, I show that the informativeness of reviews does not depend only on the presence of fake reviews, but also on the way the online platform presents reviews, that is how it publishes and summarizes them on its website. Unsurprisingly, I find that review manipulation can have a negative impact on the information that is disseminated to consumers. More interesting, however, is the solution that platforms can implement to deal with the destruction of information by fake reviews. I establish that publishing all the reviews' content is not systematically the design that conveys the more information to consumers. A platform by voluntary withholding information on the reviews' content may improve the overall informativeness of reviews. In the end, the way the platform chooses to present information and how it addresses the issue of review manipulation can be linked to its source of revenues (i.e., sales commissions, advertisements or subscription fees). Hence, my analysis shows that at least two factors, how platforms conduct business and how sellers manipulate reviews, have a determinant impact on the consumers' ability to rely on reviews.

Next, the third chapter of this thesis entitled *Noisy Learning and Price Discrimination: Implications for Information Dissemination and Profits* is co-authored with Marc Santugini and analyzes a situation where consumers learn from prices. In this chapter, we investigate the common commercial practice of third-degree price discrimination in the presence of informational asymmetries among consumers. Third-degree price discrimination is a practice that has gained in popularity in the Digital era with the shift from brick-and-mortar stores to the online marketplace: it is easier for firms to accumulate information on consumers and thus, easier to charge different prices in different consumer segments. Market segmentation seems to advantage firms and to disadvantage consumers in the first place. But, this is not necessarily the case. In the paper, we show that if some of the consumers are uninformed about the characteristics of the good (which are known by the firm), then price discrimination can be detrimental to the firm and beneficial to the consumers – a result somewhat opposite to that known for complete information. On the one hand, we find that it is not always more profitable for the firm to use discriminatory pricing since the presence of uninformed buyers provides the firm with the incentive to engage in noisy price signaling. Indeed, if the benefit from price flexibility (through discriminatory pricing) is offset by the cost of signaling quality through two distinct prices, then it is optimal not to segment the markets and to set the same price. On the other hand, a firm facing two demand segments can use third price discrimination to transmit more information to consumers. Indeed, since there are two prices for the good, uninformed consumers can retrieve more information about the characteristic of the good. This remains true even when discriminatory pricing reduces the amount of information contained in each price in comparison to the amount of information contained in the uniform price.

Then, the second part of the thesis consists of Chapter 4 which is entitled *Information Choice and Diversity: The Role of Strategic Complementarities* and is written in collaboration with Sidartha Gordon. There, we analyze the process of information acquisition, that is, how economic agents make their choice of information and what the resulting information structure is. As mentioned before, in face of the vast pool of information sources available, the issue of information choice arises unlike never before. Indeed, where should one obtain his information? If individuals choose the same sources of information, then they learn (and know) similar things, and thus, are more prone to take similar actions. If they don't choose the same sources of information, then coordination through decentralized markets – whether it is desirable or not – is likely to be more difficult. This goes far beyond than consumers buying decisions, say individuals learning about what the best car to buy is. It extends to countries, central banks, major investors, corporations and so on. Hence, information choices are relevant from the perspective of the agent himself, but also from the perspective of the aggregate economy, and part of the explanation of many economics phenomena relies on information decisions. Most models in economics assume a stylized information structure that abstracts from the process by which individuals come to possess some piece of information. In contrast, we study a model where the information structure is the result of choice and not simply a given of the economy. Doing so furthers our understanding of individual information choice when agents are interacting strategically with each other.

When the players' source choices are different, we could say that information in the economy is diverse, and when the source choices are the same, then we could say that information is rather homogenous. The objective of Chapter 4 is then to characterize the extent of information diversity and to understand what its determinants are. This is crucial as economies with diverse and homogenous information may differ dramatically in terms of outcomes. More specifically, in this chapter, we study a class of games where players face restrictions on how much information they can obtain on a common payoff relevant state, but have some leeway in covertly choosing the dependence between their signals, before simultaneously choosing actions. We obtain a characterization of the extent of information diversity by providing conditions under which the player obtain information from similar sources. In particular, we show that each player chooses information that is more dependent on the information of other players whose actions are either isotonic and complements with his actions or antitonic and substitutes with his actions. Similarly, each player chooses information that is less dependent on the information of other players whose actions are antitonic and complements with his actions or isotonic and substitutes with his actions.

In the end, the general objective of this thesis is to further our understanding of the impact of strategic interactions on the type of information that economic agents choose to acquire and of the impact of information asymmetries on the outcomes of different markets. Studying the process of information acquisition and the resulting quality of information is crucial as information is the keystone of every decisions agents make.

Chapter 2

Can You Trust What You Read on the Internet? Designing Platforms to Deal with Shill Reviews

2.1 Introduction

User-generated content, such as ratings and product reviews, plays an important role in the decision of modern consumers.[1] Not only do reviews guide consumers' choices in electronic marketplaces, they also provide information about products, such as restaurants and movies, purchased outside the online world. Although some reviews are posted on personal blogs and online newspapers, the bulk is posted on dedicated informational platforms. Some of the most popular platforms are Amazon for books and electronics, Yelp for restaurants and other local services, Expedia for hotels and TripAdvisor for hotels and restaurants.

A concern for consumers is that the credibility of reviews can be undermined by the manipulation of reviews by sellers. Because the reviewing process is practically anonymous, consumers have very limited information on individuals who write reviews. Consequently, the sellers can easily hire *shills*[2] to post fake positive feedbacks for their products. Anecdotal evidence suggests that this practice, although illegal, is common.[3] Sometimes, shill reviews are easy to detect, but it is the exception rather than the rule. Most of the time, it is difficult for consumers to distinguish a real review from a fake one.[4] For this reason, it is commonly believed that manipulation makes it harder to extract information on products.

This raises a series of questions: Is manipulation necessarily harmful to consumers? Could reviews be more informative when they are manipulated? If fake reviews impair information transmission, can we do something about it? When answering these questions, I show that the platform hosting the reviews plays a crucial role in the analysis.

First, the platform has a role to play because it chooses how to present the reviews and how to summarize their content on its website. For instance, a platform can display the

[1] Reviews have been shown to be influential in the consumers' decision process (Chevalier and Mayzlin, 2006; Jiang and Wang, 2007; Kwark *et al.*, 2014; Luca, 2011).

[2] A shill is an accomplice of a seller who acts as an enthusiastic customer to entice or encourage others.

[3] For instance, some years ago, Amazon revealed the email addresses from book reviewers by mistake and it was apparent that most of them originate from authors and publishers of the books. Fiverr.com, a site on which individuals sold services for 5$, also contains a large number of ads offering to write positive/negative reviews for a product of your choice. At last, 16 percent of restaurant reviews on the platform Yelp are tagged as potentially fraudulent (Luca and Zervas, 2013).

[4] Li and Hitt (2008) even show that it is erroneous to deduce that a review is false because the opinion it conveys is extreme.

entire set of consumer reviews or only a selection. It can also publish summary statistics like the average review and the reviews' distribution. The decision of what to publish is a design choice for the platform.[5] The design is important because consumers, who are seeking to buy a given product and who are imperfectly informed on that product, heavily rely on the information disclosed by the platform. In this paper, I show that the rule used to summarize content indirectly affects the informativeness of reviews by affecting the extent of seller manipulation.

Then, the second reason why the platform is an important player is because its source of revenues might influence its decision regarding which information design to adopt. Indeed, what (and how much) the platform allows the consumers to learn is intricately related to how it conducts business. In general, the platform can collect revenues in different ways from consumers, sellers, or third parties such as advertisers. In the paper, I concentrate on three business models that are common in the industry: revenues may come from commissions that sellers pay after an item listed on the platform is sold, from advertisers who buy advertising space from the platform, and finally from consumers who pay a fee to subscribe to the platform. In this paper, I show that the platform's business model can be linked to its design choice and thus, to the informative content of reviews.

My analysis builds on the following model. A continuum of consumers considers buying a product for which the quality is either high or low, but is private information of the seller. To guide prospective consumers, it is assumed that previous buyers have left reviews of the product on an Internet platform. More specifically, it is assumed that these reviews can be summarized by some statistic that I call the review statistic. In order to try to increase sales, the seller can exert effort to manipulate reviews, effort which results in an inflation of the review statistic. It is assumed that the extent of the seller's manipulation cannot be verified by consumers nor by the platform.

At the heart of the analysis lies the platform's design for presenting reviews, that is, the signal the platform chooses to publish about the review statistic. In the current paper, I concentrate on two design possibilities. The first possibility consists of the platform publishing the review statistic directly. The second possibility consists of the platform publishing only a binary label saying either that the product is "recommended" or that it is

[5]Throughout the paper, the *design* of the platform refers to functional design rather than visual design such as aesthetic factors or layout considerations.

"not recommended". The product is recommended if the review statistic exceeds a certain threshold, which is predetermined by the platform. In this case, the signal published by the platform is coarser and is not a sufficient metric of the consumer reviews. To put it simply, the platform has the choice to disclose all the information it holds to consumers, or to use a coarse binary label that only partially discloses its information.

In this setting, I investigate two questions:

1. First, taking the possibility for review manipulation into account, which design maximizes information dissemination to consumers?

2. Second, taking the possibility for review manipulation into account, which design generates the most profits for the platform?

To address question 1., one needs to determine what the seller's optimal level of manipulation effort is for each design. Section 2.6 accomplishes this task while focusing on pure manipulation strategies for the seller. In particular, it is established that if the platform is using the design that shows the review statistic, then uniquely two outcomes can occur in equilibrium: either the consumers completely learn the quality, or they learn nothing more than what they already knew. Instead, if the platform is using the coarse design with the binary label, then partial learning of the quality may also be possible in equilibrium. In itself, this eventuality of partial learning with the binary label design does not mean much. But, by comparing the outcome in terms of learning for the design that shows the review statistic to the outcome for the design with the binary label, one can observe the following: for some parameters, the design showing the review statistic is associated to no learning, whereas the design with the binary label is associated to partial learning. This means that the design where the platform discloses all that it knows does not systematically transfer more information to consumers than the binary label design. Paradoxically, this implies that the platform by voluntary withholding information (the review statistic) may improve the overall informativeness of the reviews.

In the context of sender-receiver games, it has already been noted that more noise in a signal (coarser signal) may lead to more information being revealed to the receiver (Blume *et al.*, 2007; Goltsman *et al.*, 2009; Gordon and Nödelke, 2013; Armin, 2014). In these papers, the mechanism at play is that noise has a strategic effect which is to induce the sender to reveal more information than what he otherwise would. In the current paper, noise has a slightly different strategic effect. To see this, one need to observe first that, contrarily to a standard signaling game, here the fake reviews (the sender's message) are combined with the genuine reviews (an exogenous message). This combination of messages implies that a sender can not only produce information, but can also destroy information. Then, in the model, it may be possible that the platform by using the coarse design actually changes the equilibrium of the manipulation game such that the seller destroys less information than what he otherwise would if the platform were to use the design that directly shows the review statistic. In other words, by using the design with a binary label instead, the platform weakens the low-quality seller's incentives to manipulate reviews, which is at the source of information destruction. The main reason behind the result is that the marginal return to manipulation is smaller with the binary design than with the design that reveals the review statistic. Another reason is that manipulation might be more costly with the binary label design: It is necessary that a seller manipulates the reviews to a minimum in order to reach the threshold at which the platform switches from the "not recommended" label to the "recommended" one.

Next, the second question of the paper is related to the design that is preferred by the platform. In particular, in Section 2.7 the design becomes an endogenous choice for the platform. More precisely, I analyze the equilibrium of the game where the platform commits first to a design and then the seller decides on his level of manipulation effort. My analysis shows that a platform receiving advertising revenues can be considered as similar to a platform receiving subscription fees in the sense that both seek to provide information with maximal value for the consumers. A platform that receives commissions, however, is concerned with the maximization of the number of sales, an objective that does not a priori coincide with the maximization of information value. Thus, the analysis splits in two in treating the case of a *transactional* platforms separately from a *non-transactional* platform. A platform is said to be transactional when it receives revenues from sales commissions, and non-transactional when it receives revenues from advertising or from subscription fees.[6]

[6]For example, platforms that could be considered as transactional are Airbnb, Amazon, Etsy, and Expedia, and platforms that could be considered as non-transactional are Angie's List, City Search, Rotten Tomatoes, TripAdvisor and Yelp.

The question of the optimal design choice for a non-transactional platform can be directly addressed using the results of the first part of the paper. In seeking to maximize the information value, the platform may prefer to use the design that reveals the review statistic, may prefer to use the one with the binary label, or may be indifferent between the two. And so, because of the seller's manipulation incentives, the platform may prefer to deprive the consumers of information even though it cares about information quality.

The case of a transactional platform is not as straightforward, in part because one needs to assess how the sales maximization objective relates to the value of information. In other words, does more information translate into larger sales, at least in expectation? Surprisingly, under some conditions the answer is no. This is because more information could mean that there is better evidence pointing to a high quality, in which cases sales will be high, but also that there is better evidence pointing to a low quality, in which cases sales will be low. A transactional platform would ideally want to (but can't) avoid revealing information that would lead consumers to learn that product quality is low. Given the model's specification, it turns out that the platform prefers reviews to be completely uninformative. Note that this last result can be seen as a special case of Kamenica and Gentzkow (2011). That is, the platform (the sender) never benefit from transferring information to consumers (the receiver) since its payoff function is concave in the consumers prior beliefs. In terms of design, this implies that it is always optimal for the platform to commit to using the binary label design with a very low threshold for recommendation so that the label "recommended" is always published for sure.

Overall, my analysis shows that manipulation is one factor among others that can affect how informative reviews are. Ultimately, how platforms conduct business, or more generally their own incentives when choosing how to design information on their website, has also a determinant impact on consumers' ability to rely on consumer reviews.

The remainder of the paper is structured as follows. In Section 2.2, I discuss the related literature and the contribution of the paper. The model is presented in Section 2.3, which is followed by the definition of some important concepts in Section 2.4. In Section 2.5, as a benchmark, I analyze the informational properties of the different designs in the absence of manipulation. In Section 2.6, after characterizing the equilibrium of the manipulation game, I compare the informational properties of the different designs when there is manipulation.

In Section 2.7 follows the analysis of the optimal design choice of the platform in relation to its business model. Finally, Section 2.8 concludes.

2.2 Related Literature

The idea that a seller can interfere with – or at least influence – the learning process of consumers on the Internet is relatively new, but takes root in the well established literature on social learning. The literature has explored various ways in which sellers can affect the information conveyed through online reviews. When consumers arrive sequentially or can strategically decide of the date of their purchase, the sellers can influence learning with an appropriate choice of prices (Bose *et al.*, 2008, 2006; Bhalla, 2012; Papanastasiou and Sava, 2014; Ifrach *et al.*, 2013; Debo *et al.*, 2013). Other papers also explore how to optimally determine the product launching sequence (Liu and Schiraldi, 2012; Bhalla, 2008) or the supply decision (Bar-Isaac, 2003; Sgroi, 2002; Papanastasiou *et al.*, 2014) to channel how consumers learn from reviews. These papers focus on important aspects for sellers, but omit the two-sided platform that is inherent to most electronic transactions. By contrast, I consider the platform as a strategic agent and I abstract from the pricing and supply decision of the seller to concentrate on his manipulation decision.

Two related issues are analyzed throughout the paper: (a) the impact of manipulation on information dissemination and how it depends on the platform's design, and (b) the kind of design a platform wants to use depending on its business model for generating revenues.

The first issue is related to the literature on costly noisy signaling (e.g., Matthews and Mirman, 1983; Kyle, 1985; Harrington, 1986; Carlsson and Dasgupta, 1997). Indeed, since consumers infer the quality of a seller's product from observing reviews, manipulation is like a noisy signaling device. In the model, signaling is noisy for two reasons.

First, the fake reviews (the seller's message) are combined with exogenous real reviews that are noisy because of an exogenous shock. Hence, consumers being uninformed on quality and the realized shock on genuine reviews cannot, in general, perfectly infer quality from the reviews they observe. That is, it is possible that, even in a separating equilibrium, the final reviews observed by consumers convey only partial information. This echoes to the literature starting with Kyle (1985) or Matthews and Mirman (1983). It is worth mentioning that the seller is also uninformed on the realized shock. Consequently, he has imperfect control over the reviews actually observed by consumers.[7] Although the seller knows how the platform transfers information to consumers, he knows only the distribution of genuine reviews. Therefore, the seller can only conjecture the distribution of final reviews when deciding how to manipulate.

The second reason why signaling is noisy in the model is because the platform can choose to garble the reviews. In using the binary design, the platform adds (non-additive) noise and sends a coarser signal of the reviews. Thus, the platform, in choosing a design, actually makes a choice on the amount of noise that it finds optimal acknowledging the seller's manipulation and the consumers' learning process. This question of the optimal level of noise has been exploited in different contexts in sender-receiver games (e.g. Blume *et al.*, 2007; Goltsman *et al.*, 2009; Armin, 2014) or in the design of search platform (Eliaz and Speigler, 2015). Harbaugh and Rasmussen (2013) also explore the optimal level of noise in a certification game and show that a certifying agency by using a coarse grading scheme will maximize information. The mechanism at play is that it induces more firms to be certified. Recently, another example is Gordon and Nödelke (2013). They analyze a signaling model with lying cost where the sender's message gets combined with noise and show that noise can improve information transmission. My work differs from theirs, not only because the type of noise is different, but also because in their case, the sender's message is not combined with an exogenous signal. Indeed, here, the noisy genuine reviews (to which the sender's message is added) act as an exogenous signal and are essential to my results: without genuine reviews, a low-quality seller cannot mimic the high genuine reviews and the interaction between the seller and consumers boils down to a standard costly signaling game.

This paper is also related to the signal-jamming literature (e.g. Riordan, 1985; Fudenberg and Tirole, 1986; Mirman *et al.* 1994). This stream of literature generally assumes that firms, in order to mislead each other, manipulate prices or output by distorting them from some myopic optimal counterparts. In the current paper, the seller manipulates reviews to influence consumers, not his rivals.

[7]The fact that the seller does not perfectly control the signal is in contrast with the standard noisy signaling literature (e.g., Heinsalu (2014) for a survey).

To my knowledge, manipulation of consumer reviews has not been the center of many theoretical papers. Nicollier and Ottaviani (2014), in a work in progress, propose a dynamic game to investigate how prices, reviews and manipulation affect each other, but do not consider the platform. Dellarocas (2006) is closest to my work and has studied manipulation from a number of points of view, except that he does not introduce the platform and its design choice. Dellarocas work is particularly insightful as it applies to situations with multiple sellers. Under specific distributional assumptions for the reviews (that are not always satisfied in my model), he shows that if sellers' equilibrium manipulation strategies are monotonically increasing (decreasing) in quality, then manipulation increases (decreases) the information value of reviews to consumers. Instead of focusing on the comparison of the information value of reviews with and without manipulation, when considering multiple design, what becomes interesting is to compare the effect of manipulation on the information value across the different designs. In that respect, it is in order to reiterate one of the important results of this paper, which is to show that, when there is manipulation, the a priori less informative design may disseminate more information in the end than the a priori more informative design.

The second issue the paper is addressing relates to the platform's choice of information structure, which is generally known as information design. On that account, the Bayesian persuasion literature is particularly relevant. Indeed, one can see the platform and consumers' interaction as being akin to a game of Bayesian persuasion of the like of Kamenica and Gentzkow (2011). In the absence of review manipulation, the model can be seen as a particular case of their Bayesian persuasion game. The platform pre-commits to a signal rule on the reviews' content that is passed on to consumers. Yet, my model goes a step further by adding the seller as a third player and giving him the opportunity to distort the signal received by consumers. An information structure with an exogenously given information value could become less (or more) informative because of the seller's manipulation. Thus, this means that in order to control the information structure, the platform needs to anticipate the direction and the extent of the seller's distortion when choosing the design.

Despite their practical importance, it appears that platform informational design and review manipulation have seldom been analyzed together. Two relative exceptions are Mayzlin *et al.* (2014) and Che and Hörner (2014). Mayzlin *et al.* (2014) conduct an empirical analysis of manipulation by hotels and take advantage of a disparity in platform design between Expedia and TripAdvisor regarding who can post a review. They find that the distribution of reviews for independent hotels and their competitors tend to differ on the two platforms and interpret this difference as a sign that TripAdvisors' reviews are more likely to be fraudulent. They exploit the link between the platforms' organization and the incentives to manipulate the reviews, but do not endogenize the design. As for Che and Hörner (2014), they analyze how a platform making product recommendations can control information revelation to agents to incentivize their experimentation with a product. Thus, they address the design of a recommendation platform, but leave out of their analysis the seller and the possibility for review manipulation. Their work also differs from mine with respect to what the platform is allowed to do. In their case, the platform can manipulate the information published on the product, whereas in my model, the platform commits to transfer information truthfully.

Finally, this paper is also related to a number of streams of literature in marketing and computer science. For instance, Dai *et al.* (2012) consider the optimal rule for the computation of the average reviews on Yelp to address the fact that not all reviews are equally informative. Related to the design of online information systems are also Dinerstein *et al.* (2014), Fradkin (2014), Ghose *et al.* (2013), Hajaj and Sarne (2013, 2014). Horton (2014) analyzes the design of search algorithms, and Dellarocas (2005) and Li and Hitt (2010) analyze the design of reputation systems.

2.3 Model

There are three categories of players: a seller of a product, consumers buying the product, and an online platform (which is independent from the seller) where consumer reviews are posted.

The game lasts for three periods, $t \in \{0, 1, 2\}$. At date 0, the platform chooses how to display the contents of the consumer reviews on its website. I call this choice the *design* of the platform. At date 1, a generation of early consumers buys the product and, after purchasing, post honest reviews about their experience with the product. The seller has the opportunity to engage in review manipulation by hiring shills to post fake positive

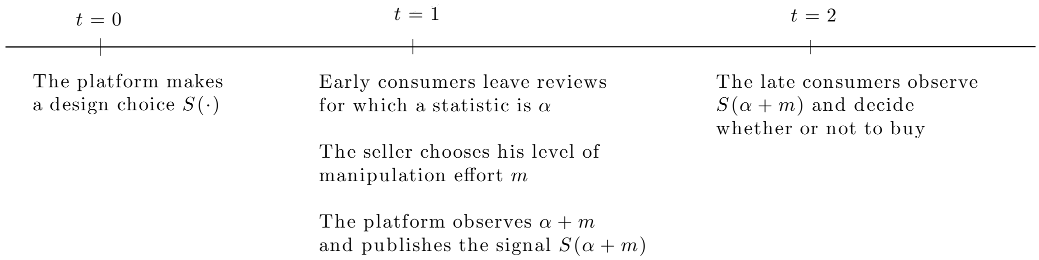


Figure 2.1: Timeline of the Game

reviews. The platform then collects real and fake reviews indistinctly and publishes a signal about the content of the reviews according to the design chosen in $t = 0$. At date 2, a generation of late consumers observes the signal posted by the platform, updates beliefs about product quality and then decides whether or not to buy the product. Figure 2.1 summarizes the timeline.

In the following, I provide the complete game description and, alongside, more details about the seller, the consumers and the consumer reviews. The platform and its design choice are discussed formally at the end of the section.

2.3.1 Game Description

The seller's product has an exogenous quality of θ, which is the realization of the binary random variable $\tilde{\theta} \in \{\theta_H, \theta_L\}$ with $0 < \theta_L < \theta_H < 1$ and $\mathbb{P}(\tilde{\theta} = \theta_H) = q \in (0, 1)$.[8] The seller knows the realization of $\tilde{\theta}$. All consumers and the platform know q, but not the realization of $\tilde{\theta}$.

2.3.1.1 At date 0

The platform chooses the design:

At date 0, the platform commits to a design: for any set of reviews (to be collected at date 1) the platform is committed to publish some signal of this set. In other words,

[8] I adopt the following notational rules: (a) a tilde is used to distinguish a random variable from its realization, i.e., $\tilde{x}$ denotes a random variable, and x its realization, and (b) the notation $\mathbb{P}(y)$ is used to denote the probability of the event y.

the platform decides on a rule to summarize reviews on its website and cannot change it afterward. It is assumed that this choice is common knowledge among all players.

In general, a design for the platform is a signal S that maps the consumer reviews to a set of messages $\mathcal{M}$. The platform chooses the rule S as well as the messages space $\mathcal{M}$ itself. Hereafter, I use $(\mathcal{M}, S)$ to designate a general design, and use s to denote the realization of a particular signal. Section 2.3.2 specifies the designs that are considered in the paper.

2.3.1.2 At date 1

Date-1 Consumers post reviews:

For the purpose of this paper, I assume that the generating process of genuine consumer reviews is exogenously given. I also assume that there exists some statistic $\alpha(\cdot)$ of the consumer reviews that is informative of the product quality.[9] More specifically, it is assumed that $\alpha(\cdot)$ depends on the quality of the product $\theta \in \{\theta_L, \theta_H\}$, a constant b, and a shock λ, as follows

$$\alpha(\theta) \quad \equiv \quad \theta + b + \lambda. \tag{2.1}$$

The higher the quality is, the higher is α. Moreover, the statistic α fluctuates randomly around θ due to the presence of a shock.[10] This shock is the realization of the random variable $\tilde{\lambda}$ which is uniformly distributed on $[-b, b]$, where $(\theta_H - \theta_L)/2 < b \leq 1 - \theta_H$.[11]

[9]Dellarocas (2006) and Mayzlin *et al.* (2014) also assume that the review process is exogenous. The statistic α could be the average review, the number of reviews, and so on. There are two reasons why there is no need to exactly specify what α is. First, because the reviews process is assumed exogenous and consequently, it does not really matter for the manipulation and design decisions. Second, because there is some disagreement in empirical papers regarding which aspects of reviews are influential. Some studies shows that the level of online reviews have a significant effect on sales, while other studies show that it is the volume of reviews that significantly influence sales (See e.g. Chintagunta *et al.*, 2010).

[10]This shock to the quality θ might be the result of a fluctuation in the production process. For instance, in the category of manufactured goods, fluctuation in the quality of inputs used has an effect on the quality of the output. In the case of service goods, variation in the staff might impact the quality of the customer service. Ifrach and al. (2013) also suggest that this time-dependent random fluctuation could be the result of a change in the popularity of the product across time.

[11]The constraint $(\theta_H - \theta_L)/2 < b$ is imposed in order to have a situation with adverse selection. See footnote 12. The constraint $b \leq 1 - \theta_H$ is imposed such that for date-2 consumers $\mathbb{P}(\tilde{u}_{i,2} = 1|\theta, v_i) \leq 1$.

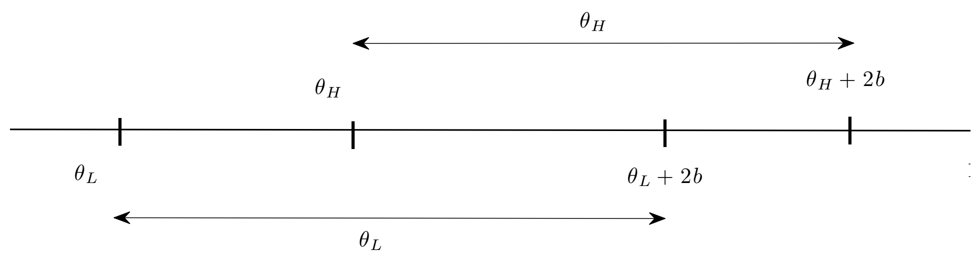


Figure 2.2: The support of $\alpha(\theta_H)$ and $\alpha(\theta_L)$

Thus, conditional on θ, $\tilde{\alpha}(\theta)$ is a random variable that is uniformly distributed on $[\theta, \theta + 2b]$. Figure 2.2 depicts the support of the distribution of $\alpha(\cdot)$ for each possible quality.[12]

<u>The seller manipulates reviews:</u>
At date 1, simultaneously with date-1 consumers posting reviews, the seller decides on a level of manipulation effort in order to improve the reviews. I assume that the marginal cost of production is independent of quality and normalized to 0 and that the product's price is exogenously fixed (cf. Remark 1 in Section 2.3.3). Hence the only decision that the seller takes is how much effort to exert to manipulate reviews.

Specifically, the seller's manipulation strategy is denoted by $m(\cdot)$ and is a function of the product quality θ. Manipulation is assumed to have a positive linear impact on the review statistic α.[13] That is, instead of the statistic $\alpha(\cdot)$ of genuine reviews, it is only the statistic $\alpha^m(\cdot) \equiv \alpha(\cdot) + m(\cdot)$ that can be observed.

In the reality, manipulation is costly because the seller must pay individuals to write false reviews, but also because it could be revealed that the seller has hired shills which would result in some reputation costs for the seller.[14] In accordance with Dellarocas (2006) and Mayzlin *et al.* (2014), I assume that the cost of manipulation is quadratic in the amount of manipulation. This implies that more false reviews are more costly. For $x \in \{H, L\}$, I assume that the cost of the manipulation strategy $m(\theta_x)$ is $c_x \cdot m(\theta_x)^2$ with $c_x \in (0, 1]$.

[12] The assumption that $b > (\theta_H - \theta_L)/2$ implies that the distributions of $\tilde{\alpha}(\theta_H)$ and $\tilde{\alpha}(\theta_L)$ overlap on a portion of their support so that observing a realization α is not always sufficient to recover the value of θ.

[13] Mayzlin *et al.* (2014) and Dellarocas (2006) also analyze manipulation through its effect on some statistic of genuine reviews and they also make the assumption that manipulation has a linear impact on this statistic.

[14] Mayzlin at al. (2013) provide a detailed discussion on the cost of manipulation.

The platform publishes a signal:

Then, it is at date-1 that the platform publishes a signal about the content of the reviews according to the design chosen at date 0. In the model, I assume that this boils down to publishing a signal about α^m. Given $(\mathcal{M}, S)$ is the design chosen at date 0 and $\alpha^m(\cdot)$ summarizes the set of reviews it has received, the platform publishes the signal $S(\alpha^m) = s \in \mathcal{M}$. Date-2 consumers will use the platform signal on α^m to learn about product quality.

2.3.1.3 At date 2

Date-2 consumers make their buying decision:

At date 2, the seller faces a new generation of consumers. The total mass of consumers is normalized to one, and the consumers are assumed to be heterogeneous in their taste for the product. Specifically, the taste v_i of consumer i is the realization of a random variable $\tilde{v}_i$ which is uniformly distributed on $[0, 1]$.

Date-2 consumers make a once-and-for-all decision to buy the product or not. If consumer i chooses not to buy the product, he obtains an outside option that yields zero utility. If consumer i purchases the product, he obtains a net utility, i.e., gross utility net of the price, of $\tilde{u}_{i,2} \in \{-1, 1\}$ with $\mathbb{P}(\tilde{u}_{i,2} = 1|\theta, v_i) = \theta v_i + b$. The probability of a high utility, i.e., $u_{i,2} = 1$, depends, on the quality of the product θ, but also on the consumer's type v_i.

Date-2 consumers are rational Bayesian decision makers, and buy or forgo the product on the basis of their individual taste, the signal published by the platform and their conjectures on the seller's manipulation.[15] As the game ends at date 2, there is no need for the late consumers to post reviews and thus no need for the seller to manipulate reviews at date 2.

[15] For simplicity, it is assumed that consumers rely solely on the signal published by the platform to update their beliefs. This is imposed to capture the following feature of the industry. Platforms provide consumers with convenient access to a large collection of reviews, and consequently to a tremendous amount of information. For that matter, it is not unusual for consumers to be exposed to more reviews than they can (want to) process. Because they are competing with each others for consumers' attention, the platforms seek to reduce the time consumers spend browsing the reviews. To do so, the platforms publish simple summary information about the reviews' content. Dai *et al.* (2012) also argue that because of consumer inattention, the method the platform chooses to aggregate information is especially important.

2.3.2 Two Simple Designs

It is now time to be more explicit on the different designs that I consider in this paper. Given α^m is the observed review statistic, the platform needs to decide which signal to display about α^m.

The class of designs $(\mathcal{M}, S)$ is large as there is an infinite number of possibilities for $\mathcal{M}$ and S. Therefore, I impose restrictions on the design space. Specifically, the set of feasible designs is constrained to the following two transformations of $\alpha^m(\cdot)$:

i. The first possibility for the platform is to report α^m directly. In this case, the signal follows the rule

$$S_c : \alpha^m \to \alpha^m. \tag{2.2}$$

ii. The second possibility is to publish that the product is recommended by posting a "thumbs up" 👍 or that the product is not recommended by posting a "thumbs down" 👎 according to the rule

$$S_b(\alpha^m;\ T) = \begin{cases} \text{👍} & \text{if } \ \alpha^m \geq T \\ \text{👎} & \text{otherwise,} \end{cases} \tag{2.3}$$

where $T \geq 0$ is chosen by the platform. This design can be interpreted as whether or not the platform recommends the product with T being the threshold for recommendation.

To simplify the terminology, I refer to the design $(\text{supp}(\alpha^m), S_c)$, where $\text{supp}(\alpha^m)$ is the support for α^m, as defined in *i.* as the **continuous design**. I refer to the design $(\{$👎, 👍$\}, S_b(;T))$ defined in *ii.* as the **binary design**. At last, since implementing a design boils down to writing a computer code only, I assume that both designs are equally costly for the platform.

2.3.3 Remarks

Before proceeding with the characterization and analysis of the equilibrium, a few comments about the assumptions of the model are in order.

Remark 1. The product's price is exogenously fixed. There are two reasons for this assumption. First, if the seller were to set its price, since he knows product quality, it would be possible for him to signal quality through the price he chooses. This adds an unnecessary layer of complexity to the model. Second, control over the price implies that the seller can control the mass of consumers who buy in date-2. Thus pricing and manipulation decisions can be considered as substitutes. I shut down the price channel so as to concentrate on the seller's manipulation decision.

Remark 2. In practice, review manipulation does not only consist of adding fake positive reviews. A seller might also bribe an unsatisfied buyer to change his negative review into a positive one, he might also try to preempt a negative review with a discount conditional on a good review being written. Another possibility is to have some of the negative reviews removed. Some platforms allow the sellers to dispute the negative reviews (ultimately, it is in the platform's discretion to remove them or not). Although these different types of manipulation may have different costs, I do not make the distinction in the model.

Remark 3. At date 1, when choosing his manipulation strategy $m(\cdot)$, the seller knows the deterministic rule used by the platform, but he does not know the realization of the shock λ. That is, the seller only knows $\mathbb{E}[\alpha(\theta)|\theta]$. Since he does not know $\alpha(\theta)$, he is uncertain about which signal realization will be published after having manipulated the reviews. It can be said that the seller imperfectly controls the platform signal. Nevertheless, for any realization of the shock $\tilde{\lambda}$, the seller is inflating the review statistic by the same amount. In other words, the seller shifts the support of the distribution of $\tilde{\alpha}(\theta)$ to the right. This new random variable is denoted by $\tilde{\alpha}^m(\theta)$.

2.4 Preliminaries

2.4.1 Consumers' Threshold

In order to simplify the exposition of a number of equations and results afterwards, I start with the derivation of date-2 consumers' optimal strategy.

Since all date-2 consumers have the same information (the platform's signal) on product quality, it is without loss of generality to restrict the consumers' strategy to the class of threshold strategies. That is, if a consumer with individual preferences v chooses to buy the product, then all consumers with individual preferences $v' > v$ must also find it profitable to buy the product. In this sense, there exists $\hat{v} \in [0,1]$ such that $\hat{v}$ is the type of the consumers for which the expected utility of buying the product is equal to 0. This optimal threshold will be a function of the signal realization, and thus, depends on the platform design choice and the seller's manipulation effort.

Lemma 2.1. *Assume that the platform design is $(\mathcal{M}, S)$ and the consumers' conjecture on the seller's manipulation strategy is the function $m^e(\cdot)$. Assume further that upon observing signal $s \in \mathcal{M}$, $q(s, m^e; (\mathcal{M}, S))$ is the posterior beliefs updated through Bayes' rule. Then, it is optimal for a date-2 consumer with type v_i to buy the product if and only if*

$$v_i \geq \hat{v}(s, m^e; (\mathcal{M}, S)) \equiv \frac{1-2b}{2\Big(\theta_L + (\theta_H - \theta_L) q(s, m^e; (\mathcal{M}, S))\Big)}. \tag{2.4}$$

Proof. All proofs are relegated to Appendix A.3. □

I refer to $\hat{v}(s, m^e; (\mathcal{M}, S))$ as the consumers' threshold given the platform's design is $(\mathcal{M}, S)$, the signal s has been observed and the manipulation strategy m^e is conjectured.

2.4.2 The Value of Information

A recurrent theme in the paper is to determine which design between the continuous and binary designs disseminates more information to consumers. The criterion used for the comparison of the different designs is the value of information for date-2 consumers. The value of information measures the extend to which consumers' decisions are improved by extracting information from reviews from the situation where no such reviews are available.

Given the consumers' threshold is $\hat{v}(s, m^e; (\mathcal{M}, S))$ as given in Equation (2.4), and the seller's manipulation strategy is the function $m(\cdot)$, the value of information at date 2 for a consumer with type v_i is denoted $IV_i(v_i, \hat{v}, m; (\mathcal{M}, S))$ and is given by

$$IV_i(v_i, \hat{v}, m; (\mathcal{M}, S)) \equiv \mathbb{E}[\tilde{u}_i \mathbb{1}_{\{v_i \geq \hat{v}(\tilde{s}(m), m^e; (\mathcal{M}, S))\}} | v_i; (\mathcal{M}, S)] - \mathbb{E}[\tilde{u}_i \mathbb{1}_{\{v_i \geq \hat{v}(q)\}} | v_i], \tag{2.5}$$

where $\hat{v}(q)$ is the threshold as defined in (2.4) when posterior beliefs equal prior beliefs. This value is the difference between consumer i's ex ante expected utility when beliefs are updated using the platform signal and his expected utility given his prior belief q. For some consumers, the value of information will always be zero. This is the case when a consumer makes the same decision with or without reviews, i.e., he never buys or he always buys.

Then, the aggregate value of information for date-2 consumers is denoted by $IV(\hat{v}, m; (\mathcal{M}, S))$ and is

$$\begin{aligned} IV(\hat{v}, m; (\mathcal{M}, S)) &= \int_0^1 IV_i(v_i, \hat{v}, m; (\mathcal{M}, S))\mathrm{d}v_i && (2.6) \\ &= \int_0^1 \mathbb{E}[\tilde{u}_i \mathbb{1}_{\{v_i \geq \hat{v}(\tilde{s}(m), m^e; (\mathcal{M}, S))\}} | v_i; (\mathcal{M}, S)]\mathrm{d}v_i - \int_0^1 \mathbb{E}[\tilde{u}_i \mathbb{1}_{\{v_i \geq \hat{v}(q)\}} | v_i]\mathrm{d}v_i. && (2.7) \end{aligned}$$

In Equation (2.7), the first term can be considered as the average ex ante expected utility of consumers. The second term can be considered as the average expected utility of consumers given the prior belief. Notice also that this second term is independent of the platform's design choice.

The aggregate value of information depends on the consumers' posterior beliefs, through $\hat{v}(s(m), m^e; (\mathcal{M}, S))$, which in turn are influenced by the design the platform is using and the seller's manipulation strategy. Hereafter, when I refer to the value of information, it is as defined in (2.7).

2.5 A Benchmark

This section provides an analysis of the informational properties of the continuous and binary designs assuming that there is no manipulation of reviews. In addition to providing a benchmark to which compare the designs' informational properties when accounting for manipulation, the case of no manipulation allows me to illustrate the updating of consumers' beliefs in a simple context. Sections 2.5.1.1 and 2.5.1.2 present the posterior beliefs for the continuous and binary designs, respectively. Section 2.5.2 compares the value of information between these two designs.

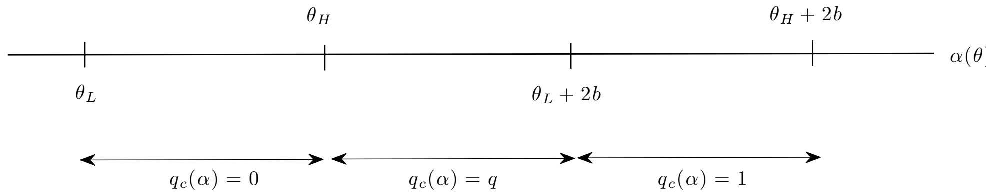


Figure 2.3: Posterior beliefs with the continuous signal

2.5.1 Informational Properties

2.5.1.1 Continuous Design: Posterior Beliefs

When the platform uses the continuous design (assuming no manipulation), the signal observed by the consumers is the realization of $\tilde{\alpha}(\cdot)$ which takes a continuum of values between θ_L and $\theta_H + 2b$. A very high realization allows the consumers to learn that product quality is θ_H, a very low realization allows them to learn that it is θ_L, but a moderate realization does not convey any information.

Let $q_c(\alpha)$ denote consumers' posterior beliefs upon seeing a signal α. Given the binary specification for product quality, $q_c(\alpha)$ is just the bayesian updated probability that $\theta = \theta_H$. Then, Lemma 2.2 specifies the possible values for $q_c(\alpha)$ and Figure 2.3 shows specifically what the posterior beliefs are as a function of $\alpha(\cdot)$.

Lemma 2.2. *For all $\alpha \in [\theta_L, \theta_H + 2b]$, when the platform uses the continuous signal, the posterior beliefs are*

$$q_c(\alpha) = \begin{cases} 0, & \textit{if } \alpha \in [\theta_L, \theta_H) \\ q, & \textit{if } \alpha \in [\theta_H, \theta_L + 2b] \\ 1, & \textit{if } \alpha \in (\theta_L + 2b, \theta_H + 2b]. \end{cases} \tag{2.8}$$

Thus, before α is realized, the posterior beliefs are the random variable $\tilde{q}_c(\alpha)$ with support $\{0, q, 1\}$. To compute the value of information, one needs only to determine the probability of the events $q_c(\alpha) = y$ for $y \in \{0, q, 1\}$.

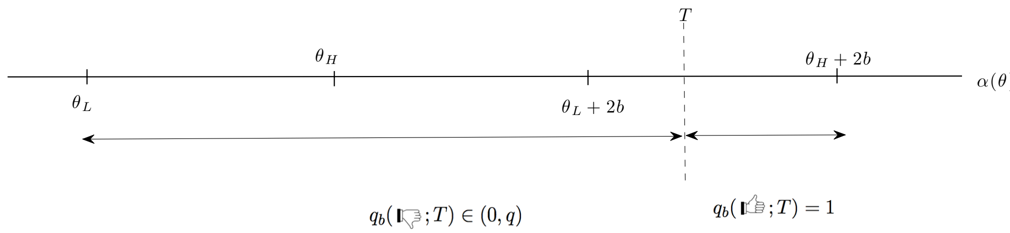


Figure 2.4: Posterior beliefs with a binary signal

2.5.1.2 Binary Design: Posterior Beliefs

Turning next to the the binary design, it is important to notice that the consumers' posterior beliefs depend on the signal realization, but also on the level of the threshold T. For any $T \geq 0$, the signal consumers can observe takes a value in the set {👍, 👎}. Upon observing the realization 👍and the realization 👎, assume that $q_b($👍$;T)$ and $q_b($👎$;T)$ denote the posteriors beliefs, respectively.

Lemma 2.3. *Assume that the platform is using the binary signal. For all threshold $T \geq 0$, if $\mathbb{P}($👎$)\cdot\mathbb{P}($👍$) > 0$, then the posterior beliefs are $q_b($👎$;T) \in [0, q]$ and $q_b($👍$;T) \in [q, 1]$.*

If $\mathbb{P}(s) = 0$ for some $s \in$ {👎, 👍} and $\mathbb{P}(s') = 1$ for $s' \neq s$, then the posterior beliefs are $q_b(s;T) = \circ$ and $q_b(s';T) = q$, where $\circ$ denotes that out-of-equilibrium beliefs need to be defined.

Figure 2.4 illustrates an example where the threshold T is moderately high. In this case, consumers expect that a high quality can be fully revealed with positive probability. That is, by seeing the realization 👍, it is revealed that $\theta = \theta_H$, i.e., $q_b($👍$;T) = 1$. They cannot, however, expect to know that $\theta = \theta_L$ for sure upon seeing the realization 👎, i.e., $q_b($👎$;T) \in (0, q)$.

Once again, before α is realized, the posterior beliefs are the random variable $\tilde{q}_b(s;T)$ with a support in $[0, 1]$. To compute the value of information, one needs to obtain the probability of the events $s =$ 👎 and $s =$ 👍. The former occurs with probability $qG_H(T) + (1-q)G_L(T)$ and the latter with probability $q(1 - G_H(T)) + (1-q)(1 - G_L(T))$

where $G_x(T)$ is given by

$$G_x(T) = \mathbb{P}(s = 👎 | \theta_x) = \min\left\{\max\left\{\frac{T-\theta_x}{2b}, 0\right\}, 1\right\}, \quad x \in \{L, H\}. \tag{2.9}$$

2.5.2 Comparison of the Two Designs

In this section, I compare the properties of the posterior beliefs for the continuous and the binary design. From Lemma 2.2 and Lemma 2.3 follows Proposition 2.1.

Proposition 2.1. *Let $\tilde{q}_c(\cdot)$ and $\tilde{q}_b(\cdot; T)$ be the random posterior beliefs when the platform is using the continuous and the binary design with a threshold $T \geq 0$, respectively. Assume further that $\mathbb{P}(👍)\mathbb{P}(👎) > 0$, then*

a) for all α such that $q_c(\alpha) = q$, $q_c(\alpha) < q_b(👍; T)$ and $q_c(\alpha) > q_b(👎; T)$;

b) $\mathbb{P}(\tilde{q}_c(\alpha) \in \{0, 1\}) > \mathbb{P}(\tilde{q}_b(s(\alpha); T) \in \{0, 1\})$.

Proposition 2.1 reveals that there is an asymmetry between the two types of designs in terms of information transmission. With the continuous design, either the consumers learn everything or they learn nothing more than what they already knew.[16] Proposition 2.1*a*) establishes that when consumers learn nothing with the continuous design, i.e., $q_c(\alpha) = q$, then the platform, by using the binary design instead, would allow consumers to extract some information by observing the signal. Besides, Proposition 2.1*b*) shows that the continuous design is more likely to fully reveal product quality than the binary design. Hence, both types of design have their advantage: With the binary design, information is transmitted in more instances. However, with the continuous design, full revelation of product quality is more likely.

Proposition 2.2, however, establishes that the binary design's signal is a garbling of the continuous design's signal in the sense of Blackwell. In other words, the signal conveyed by a binary design is coarser (noisier) than the signal conveyed by the continuous design.

[16] This hinges on the use of the uniform distribution and the identical length of the supports of $\alpha(\theta_H)$ and $\alpha(\theta_L)$.

Proposition 2.2. *For all threshold $T \geq 0$, the binary design's signal is a garbling of the continuous design's signal.*

Instead of computing explicitly the value of information for the binary and continuous design, and then compare which one is greater, one can use Proposition 2.2 to determine which design is more informative for consumers. Indeed, by Blackwell's theorem (e.g., Blackwell, 1953; Crémer, 1982), Proposition 2.2 implies that the continuous design's signal is more informative than the binary design's signal for all concave utility functions. Hence it is the case for the particular utility function of the model.

2.6 The Manipulation Game

Let me now proceed with the analysis of the manipulation game. Throughout this section, I assume that the platform has (non-strategically) committed to the design $(\mathcal{M}, S)$ and that the seller needs to decide on how much effort to exert to manipulate reviews. More specifically, while date-1 consumers are leaving a set of reviews summarized by the random variable $\tilde{\alpha}$, the seller with quality θ_x is choosing a manipulation effort level $m(\theta_x)$ for $x \in \{L, H\}$. In this case, I determine the manipulation effort exerted by the seller in a perfect Bayesian Nash equilibrium (PBE) of the game.

The section begins with the equilibrium definition in Section 2.6.1. Section 2.6.2 outlines some results on the effect of manipulation on posterior beliefs. Section 2.6.3 and 2.6.4 establish the seller's equilibrium manipulation effort when the design is continuous and binary, respectively. In doing so, a series of questions is addressed. Namely, do both types of seller necessarily manipulate reviews in equilibrium? Does a low quality seller exert more manipulation effort? Is it possible that consumers be better off with than without manipulation? Finally, in Section 2.6.5, I determine which design conveys more information to consumers given the seller's optimal manipulation effort.

2.6.1 Equilibrium Definition

The game is a Bayesian game in which the state of the world is either θ_H or θ_L, i.e., the possible product qualities. The design is fixed to $(\mathcal{M}, S)$ before the game starts. The consumers form beliefs on product quality given the signal realization and the strategy played by the seller. A PBE specifies the consumers' buying threshold $\hat{v}(s) \in [0,1]$ for all signal $s \in \mathcal{M}$ and the seller's manipulation strategy $(m(\theta_L), m(\theta_H)) \in \mathbb{R}^2_+$. In equilibrium, conditional on the design used by the platform, it is required that consumers maximize expected utility and the seller maximizes expected profits. Definition 2.1 gives the requirements a tuple $\{(\hat{v}(s))_{s\in\mathcal{M}}, (m(\theta_L), m(\theta_H))\}$ must meet to form a PBE.

Definition 2.1. *Let $(\mathcal{M}, S)$ be the design the platform is using. Then, the profile of pure strategies $\left\{(\hat{v}^*(s))_{s\in\mathcal{M}}, (m^*(\theta_L), m^*(\theta_H))\right\}$ is a PBE if and only if*

1. *At date 2, given $m^* \equiv (m^*(\theta_L), m^*(\theta_H))$*

$$\hat{v}^*(s) \in \inf\{v \in [0,1] : \mathbb{E}[u_i(v, m^*; (\mathcal{M}, S))|s] \geq 0\} \quad \forall s \in \mathcal{M}; \tag{2.10}$$

2. *At date 1, for θ_x with $x \in \{H, L\}$, given $(\hat{v}^*(s))_{s\in\mathcal{M}}$ such that demand is $1 - \hat{v}^*(s)$,*

$$\mathbb{E}\Big[1 - \hat{v}^*(\tilde{s}(m^*))\Big] - c_x m^*(\theta_x)^2 \geq \mathbb{E}\Big[1 - \hat{v}^*(\tilde{s}(m))\Big] - c_x m^2 \tag{2.11}$$

   for all $m \geq 0$, where $\tilde{s}(m^*)$ is the signal sent by the platform given $m^*(\theta_x)$ and $\tilde{s}(m)$ is the signal sent by the platform given m.

3. *Date-2 consumers' posterior beliefs are computed using Bayes' rule, whenever possible.*

Definition 2.1 specifies that Bayes' rule must be used to compute the consumers' posterior beliefs whenever possible, but there are no requirements on out-of-equilibrium beliefs. Hence, it is necessary to specify posterior beliefs when Baye's rule is vacuous. Specific assumptions on the beliefs off the equilibrium path are made in conjunction with the equilibrium derivation.

2.6.2 The Effects of Manipulation on Information Dissemination

The seller with product quality θ_x must choose how much effort to exert to shift the distribution of $\alpha(\theta_x)$. With a slight abuse of notation, let m_x denote the manipulation effort level for the seller with product quality θ_x, for $x \in \{H, L\}$. Specifically, suppose that the consumers' conjecture on the seller's manipulation efforts is $m^e \equiv (m^e_L, m^e_H)$, such that consumers' posterior beliefs upon seeing the signal s are $q(s, m^e)$ and they adopt the threshold strategy $\hat{v}(s, m^e)$ as given in (2.4). Then, the seller with product quality θ_x has to solve the following maximization problem

$$\max_{m \geq 0} \mathbb{E}[1 - \hat{v}(\tilde{s}(m), m^e) | \theta_x, (\mathcal{M}, S)] - c_x m^2, \tag{2.12}$$

taking into consideration the platform design $(\mathcal{M}, S)$, how m affects the signal $\tilde{s}(m)$, and the consumers buying threshold $\hat{v}(\tilde{s}(m), m^e)$.

The seller is choosing m before knowing the realization of the shock λ. Thus, he expects that $\alpha(\theta_x)$ takes a value between θ_x and $\theta_x + 2b$. By exerting effort to improve reviews, the seller shifts the support of $\alpha(\theta_x)$ by m_x, the level of manipulation. Specifically, the distribution for $\alpha^m(\theta_x)$ is

$$\alpha^m(\theta_x) \sim U\left[\theta_x + m_x,\ \theta_x + 2b + m_x\right]. \tag{2.13}$$

As the level of manipulation effort is restricted to positive levels of m, the distribution of $\alpha^m(\theta_x)$ first-order stochastically dominates the distribution of $\alpha(\theta_x)$. In other words, higher realizations of the review statistic are more likely when reviews are manipulated, and this is why the seller expects to gain from manipulation. In fact, the reason for the gain from manipulation depends on the seller's type. On the one hand, a seller with a high quality product wishes to manipulate since it allows him to pull the distribution of $\alpha^m(\theta_H)$ farther away from $\alpha^m(\theta_L)$ and to separate himself from a low-quality seller. On the other hand, a seller with a low quality product wants to manipulate because it allows him to move the distribution of $\alpha^m(\theta_L)$ closer to the one of $\alpha^m(\theta_H)$, and so, to mimic a high-quality seller.

Without manipulation, higher signals are associated to a higher demand for the product. Whether the same relation holds with manipulation, however, depends on whether the consumers believe that a higher signal realization is associated with a higher quality. This is the issue I now turn to by addressing the impact of manipulation on beliefs.

As sophisticated rational Bayesian agents, date-2 consumers are aware of the seller's incentives to manipulate reviews and they understand that a seller with a high quality product manipulates reviews for a different reason than a seller with a low quality product. More

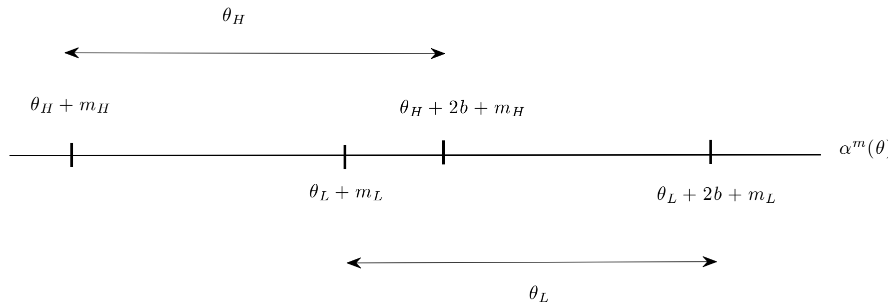


Figure 2.5: A case ruled out in equilibrium

specifically, date-2 consumers form conjectures $m^e = (m^e_L, m^e_H)$ about the manipulation efforts of the seller in order to extract information on product quality contained in the platform's signal. In equilibrium these conjectures are correct, i.e., $m^e \equiv (m^*_L, m^*_H)$.

A relevant issue is then to understand how manipulation affects information dissemination. To address this question, a first useful step is to present a feature of every pure strategy equilibrium. Lemma 2.4 states that, even though the seller can manipulate reviews, the nature of the information provided by the platform's signal cannot be distorted in equilibrium: Higher signals are always associated with a higher expected product quality.

Lemma 2.4. *Suppose that (m^*_L, m^*_H) is an equilibrium in pure strategies of the manipulation game for a given design. Then, the posterior beliefs that $\theta = \theta_H$ are increasing in the signal realization. That is,*

a) with the continuous design, for any realizations $s, s' \in [0,1]$ such that $s < s'$, if $\mathbb{P}(s|m^*_L, m^*_H) > 0$ and $\mathbb{P}(s'|m^*_L, m^*_H) > 0$, we have $q_c(s, m^*) \leq q_c(s', m^*)$;

b) with the binary design, if $\mathbb{P}(\text{👍}|m^*_L, m^*_H) > 0$ and $\mathbb{P}(\text{👎}|m^*_L, m^*_H) > 0$, we have $q_b(\text{👎}, m^*; T) \leq q_b(\text{👍}, m^*; T)$.

Lemma 2.4 does not imply that the equilibrium level of manipulation has to be monotonic in the seller's quality type. It implies only that the support for $\tilde{\alpha}^m(\theta_L)$ cannot be farther to the right than the support for $\tilde{\alpha}^m(\theta_H)$. To understand why this situation is ruled out, suppose that m_H and m_L are such that the support for $\tilde{\alpha}^m(\theta_H)$ and $\tilde{\alpha}^m(\theta_L)$ are as given in Figure 2.5 and that the platform is using the continuous design. It is necessary that $m_L > 0$ for this to happen. In this scenario, consumers associate low signal realizations to high quality and high signal realizations to low quality. Then, there exists a profitable deviation for the low-quality seller. By decreasing m_L, the low type economizes on manipulation costs. Additionally, the deviation increases the low-quality seller's expected revenues since a part of the support of $\tilde{\alpha}^m(\theta_L)$ is now in the region where beliefs are that quality is high, i.e., $q_c(\alpha^m) = 1$.

Lemma 2.4 turns out to be useful when discussing the impact of manipulation on information dissemination.

Effect of m on the distributions' overlap $\mathcal{O}$

To analyze whether manipulation generally enhances or lessens the information conveyed by reviews, it is useful to introduce $\mathcal{O}$ as the *overlap of the distributions* $\tilde{\alpha}(\theta_L)$ *and* $\tilde{\alpha}(\theta_H)$ with $\mathcal{O} = 2b - (\theta_H - \theta_L)$.

To put it simply, $\mathcal{O}$ is the interval of values for which it is impossible to distinguish θ_H from θ_L. If $\mathcal{O} = 0$, then it is always possible to distinguish the quality type. The greater is $\mathcal{O}$, the more difficult it is to distinguish θ_H from θ_L.

With $\mathcal{O}$ as the overlap without manipulation, let $\mathcal{O}^m$ be the overlap when the seller exerts manipulation effort, with[17]

$$\mathcal{O}^m = \begin{cases} 2b - (\theta_H - \theta_L) - (m_H - m_L) & \text{if } \theta_H + m_H < \theta_L + 2b + m_L < \theta_H + 2b + m_H \\ 0 & \text{otherwise} \end{cases} \tag{2.14}$$

The level of the difference $m_H - m_L$ then determines whether $\mathcal{O}^m$ is greater or smaller than $\mathcal{O}$. By comparing m_H to m_L, one can assess the impact of manipulation on information dissemination. If $m_L > m_H$, this means that $\tilde{\alpha}(\theta_H)$ and $\tilde{\alpha}(\theta_L)$ have more possible realizations in common. Therefore, we can conclude that manipulation impairs information dissemination since the two distributions are more difficult to distinguish. If $m_L < m_H$, this means that $\tilde{\alpha}(\theta_H)$ and $\tilde{\alpha}(\theta_L)$ have less possible realizations in common. And so, we can conclude that manipulation improves information dissemination since the two distributions are easier to distinguish.

[17] When $\theta_L + m_L < \theta_H + 2b + m_H < \theta_L + 2b + m_L$, the overlap is $\mathcal{O}^m = 2b + (\theta_H - \theta_L) + (m_H - m_L)$. Lemma 2.4, however, implies that this can never occur in equilibrium.

Effect of m on the distribution of the posterior beliefs

Another way to analyze the impact of manipulation on the information conveyed by reviews would be to examine how the mean and the variance of the posterior beliefs' distributions are affected by the seller's manipulation effort.[18]

Consider first the mean of the posterior beliefs. With Bayesian consumers, it is a well-know property that the posterior beliefs have the martingale property.[19] That is, the mean of the posterior beliefs is equal to the prior beliefs: $\mathbb{E}[\tilde{q}_c(\alpha^m, m)] = \mathbb{E}[\tilde{q}_b(s, m; T)] = q$. This is true independently of the manipulation effort (m_L, m_H). When turning to the variance of $\tilde{q}_c(\alpha^m, m)$ and $\tilde{q}_b(s, m; T)$, the level of m_H and m_L are, however, important. Let $\mathbb{V}(\tilde{y})$ denote the variance of a random variable $\tilde{y}$, then Proposition 2.3 characterizes the impact of manipulation on the variance of the posterior beliefs.

Proposition 2.3. *For* $\tilde{y} \in \{\tilde{q}_c(\alpha^m, m), \tilde{q}_b(s(\alpha^m, m); T)\}$ *and a pair of manipulation efforts* (m_L, m_H) *such that Lemma 2.4 is satisfied,*

a) if $\theta_L + m_L < \theta_H + m_H < \theta_L + 2b + m_L$, *then*

$$\frac{\partial \mathbb{V}(\tilde{y})}{\partial m_H} \geq 0 \quad \text{and} \quad \frac{\partial \mathbb{V}(\tilde{y})}{\partial m_L} \leq 0, \tag{2.15}$$

b) if $\theta_H + m_H = \theta_L + m_L$, *then*

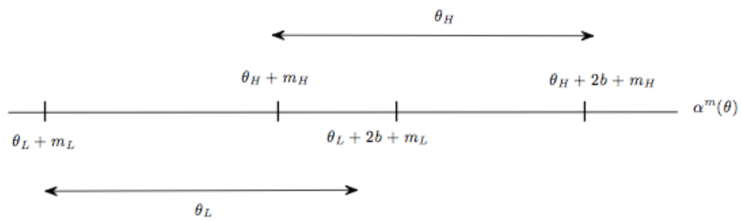
$$\frac{\partial \mathbb{V}(\tilde{y})}{\partial m_H} \geq 0 \quad \text{and} \quad \frac{\partial \mathbb{V}(\tilde{y})}{\partial m_L} \geq 0, \tag{2.16}$$

c) otherwise, a marginal increase in m_H *or in* m_L*, has no impact on the variance.*

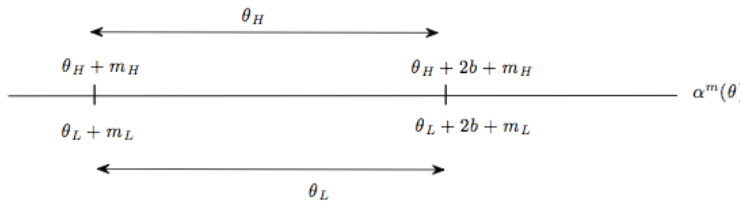
Figure 2.6 illustrates all three cases. Case a) of Proposition 2.3 applies to the situations where the support of $\tilde{\alpha}^m(\theta_H)$ and $\tilde{\alpha}^m(\theta_L)$ overlap, but where the support of $\tilde{\alpha}^m(\theta_H)$ is farther to the right. Case b) illustrates situations where the supports completely overlap. Finally, case c) concerns the situations where the support of $\tilde{\alpha}^m(\theta_H)$ and $\tilde{\alpha}^m(\theta_L)$ do not overlap at all.

In all cases, given that the mean of the posterior beliefs is the prior q, when the variance of the posterior beliefs increases, so does the informativeness of the platform signal. Indeed, suppose the variance is 0, then this means that posterior beliefs stays at q for

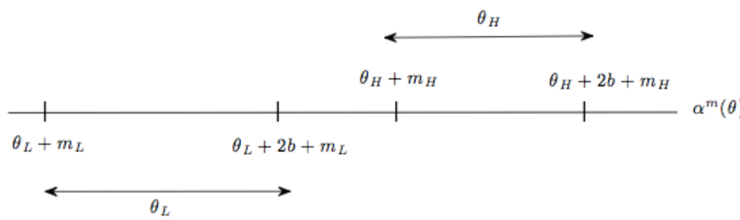
[18]See Appendix A.2 for the effect of manipulation on posterior beliefs.
[19]See for instance Vives (2008), p.389.



(a) An illustration of Proposition 2.3a)



(b) An illustration of Proposition 2.3b)



(c) An illustration of Proposition 2.3c)

Figure 2.6: Illustrations of Proposition 2.3

every realization s occurring with positive probability. Hence, there is no learning at all. Therefore, a higher variability in the beliefs is associated to higher chances to learn product quality and so, to more information.

Proposition 2.3 implies that manipulation by a high-quality seller increases the informativeness of reviews. It becomes easier to distinguish between the distributions of manipulated reviews than to distinguish between the distributions of original reviews. By contrast, manipulation by a low-quality seller decreases the information conveyed by reviews. It becomes more difficult to distinguish between the distributions of manipulated reviews than to distinguish between the distributions of original reviews. The only exception is when the two distributions overlap completely, where, in this case, an increase in manipulation effort from any type increases information.

So far, I have been analyzing the impact of manipulation for any given pair of efforts (m_L, m_H) and not specifically the pairs that can be sustained in an equilibrium. To completely determine how manipulation affects the information conveyed by reviews, one could compare the posterior beliefs' distribution at $(m_L, m_H) = (0, 0)$ to the posterior beliefs' distribution at an equilibrium profile (m_L^*, m_H^*). With this in mind, I solve for the equilibrium of the manipulation game for each design, that is, for the continuous design in Section 2.6.3 and for the binary design in Section 2.6.4.

2.6.3 Continuous Design: Equilibrium of the Manipulation Game

For this section, assume that the platform commits to using the continuous design, that is, it collects α^m and publishes it without further modification. Then, the objective is to characterize the pair of manipulation effort levels (m_L, m_H) that can occur in a PBE of the game.

Before presenting the equilibrium, let me discuss the specification of out-of-equilibrium beliefs. For any pair of manipulation effort levels (m_L, m_H), consumers can observe a realization $\alpha \in \Delta(m_L, m_H) \equiv [\theta_L + m_L, \theta_L + m_L + 2b] \cup [\theta_H + m_H, \theta_H + m_H + 2b]$, and so it is required that out-of-equilibrium beliefs be specified for all $s \in [0, \infty) \setminus \Delta(m_L, m_H)$. By Lemma 2.4, it is the case that in equilibrium $\theta_L + m_L \leq \theta_H + m_H$, which implies $\max\{\alpha : \alpha \in \Delta(m_L, m_H)\} = \theta_H + m_H + 2b$ and $\min\{\alpha : \alpha \in \Delta(m_L, m_H)\} = \theta_L + m_L$. To simplify the analysis, I make the following refinement on admissible out-of-equilibrium beliefs: for all $s > \theta_H + m_H + 2b$, then $\mathbb{P}(\theta = \theta_H | s) = \beta_1 \in [0, 1]$, and for all $s < \theta_L + m_L$, then $\mathbb{P}(\theta = \theta_H | s) = \beta_2 \in [0, 1]$.[20]

Proposition 2.4 characterizes the pair of manipulation effort levels (m_L, m_H) that can occur in a PBE of the game.[21]

[20]Lemma 2.4 does not specify the relation between $\theta_L + m_L + 2b$ and $\theta_H + m_H$. In the case where $\theta_L + m_L + 2b < \theta_H + m_H$, I assume that for all $s \in (\theta_L + m_L + 2b, \theta_H + m_H)$, $\mathbb{P}(\theta = \theta_H | s) = \beta_3 \in [0, 1]$.

[21]I show in Proposition 2.7 that a PBE in mixed strategy exists in the manipulation game with the continuous design. Here, I concentrate on PBE in pure strategies.

Proposition 2.4. *Assume that the platform commits to using the continuous signal. Let $\hat{v}(q_c(\alpha^m, m))$ be the consumers' threshold and $D(q_c(\alpha^m, m)) = 1 - \hat{v}(q_c(\alpha^m, m))$ be the demand for the product when the posterior beliefs are $q_c(\alpha^m, m)$. Then, a pair of manipulation effort levels (m_L^*, m_H^*) is an equilibrium in pure strategy only if either a), b) or c) holds, with*

a) $m_L^* > 0$, $m_H^* > 0$, *such that*

$$m_L^* = \theta_H - \theta_L + m_H^* \quad \text{and} \quad m_x^* \in \left[\frac{D(\beta_1) - D(q)}{4bc_x}, \frac{D(q) - D(\beta_2)}{4bc_x}\right], \tag{2.17}$$

for $x \in \{H, L\}$, *where* β_1, β_2 *are the out-of-equilibrium belief specifications with* $\beta_1 \in [0, 1]$ *and* $0 \leq \beta_2 < q$.

b) $m_L^* > 0$, $m_H^* = 0$, *such that*

$$m_L^* = \theta_H - \theta_L \quad \text{and} \quad m_L^* \in \left[\frac{D(\beta_1) - D(q)}{4bc_L}, \frac{D(q) - D(\beta_2)}{4bc_L}\right], \tag{2.18}$$

where β_1, β_2 *are the out-of-equilibrium belief specifications with* $0 \leq \beta_1 \leq q$, *and* $0 \leq \beta_2 < q$.

c) $m_L^* = 0$, $m_H^* > 0$, *such that*

$$m_H^* \in \left(2b - (\theta_H - \theta_L), \frac{D(1) - D(0)}{4bc_H}\right]. \tag{2.19}$$

Proof. See Appendix A.4. □

The proposition states that, given the continuous design, in an equilibrium manipulation efforts (m_L^*, m_H^*) necessarily take form *a*), *b*) or *c*). Note, however, that multiple equilibria may exist.

The focus of this section is on information dissemination. Thus, an important implication of Proposition 2.4 concerns what the consumers' posterior beliefs are after observation of the platform signal given the manipulation efforts of the seller.

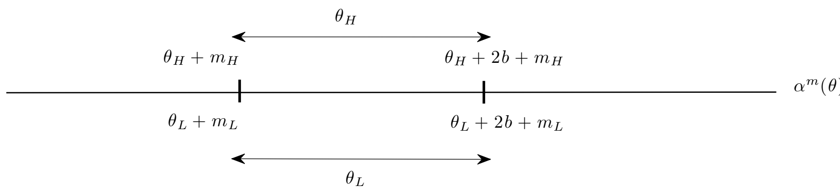


Figure 2.7: Continuous Design: Consumers learn nothing about product quality

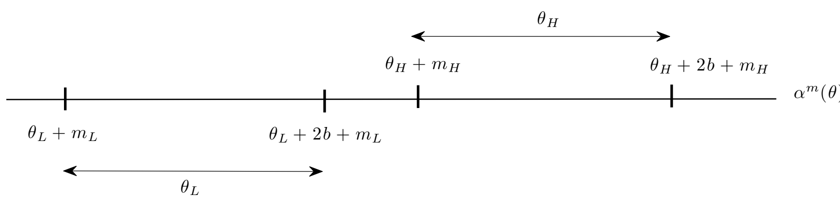


Figure 2.8: Continuous Design: Consumers fully learn product quality

Corollary 2.1. *Assume that the platform is using the continuous design. If* (m_L^*, m_H^*) *is an equilibrium in pure strategies of the manipulation game and*

a) (m_L^*, m_H^*) *is as described in Proposition 2.4a) or 2.4b), then consumers learn nothing more than what they already knew, i.e.,* $q_c(\alpha^m, m) = q$ *for all* $\alpha^m(\cdot)$ *on the equilibrium path;*

b) (m_L^*, m_H^*) *is as described in Proposition 2.4c), then consumers fully learn product quality, i.e.,* $q_c(\alpha^m, m) \in \{0, 1\}$ *for all* $\alpha^m(\cdot)$ *on the equilibrium path.*

More specifically, when consumers learn nothing, it is because the signal supports are the same under both product qualities (see Figure 2.7). When they fully learn the quality it is because the signal supports do not overlap at all (see Figure 2.8). In other words, either manipulation allows the low type to obfuscate the signal completely, or manipulation allows the high type to separate itself perfectly.

The set of necessary and sufficient conditions on parameters for which (m_L^*, m_H^*) as described in Proposition 2.4 are equilibria of the manipulation game are given in the proof of Proposition 2.4 in Appendix A.4. A determinant parameter is, of course, the marginal cost of manipulation for each type of seller. Perfect revelation of quality can occur only if the high type seller separates itself completely from the low type seller. And so, it is necessary that the high-quality seller manipulates the reviews to such an extent that the

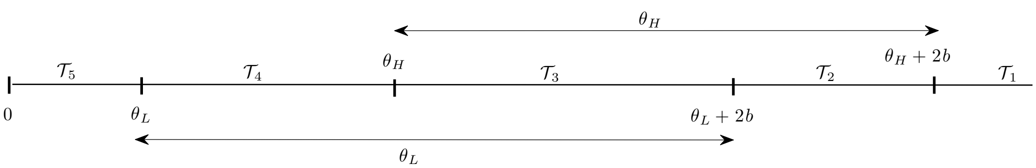


Figure 2.9: Binary Design: Five Zones for T

low-quality seller finds it too costly to imitate him, even partially. In the other case, in order to imitate perfectly the high type and to obfuscate the signal, the seller with a low quality needs to manipulate a lot given that the initial support for $\tilde{\alpha}(\theta_L)$ is farther to the left than the initial support for $\tilde{\alpha}(\theta_H)$. In this case, the high-quality seller continues to manipulate because failure to do so would penalize him even more.

At last, one could remark that Proposition 2.4 rules out as an equilibrium any situations where (m_L, m_H) is such that the supports for $\tilde{\alpha}^m(\theta_L)$ and $\tilde{\alpha}^m(\theta_H)$ overlap partially. The reason is that the low quality seller always has a profitable deviation in this case (c.f. Appendix A.4).

2.6.4 Binary Design: Equilibrium of the Manipulation Game

Consider now the case where the platform commits to using the binary signal and fixes the threshold to some level $T \geq 0$. The level of T jointly determines the probability that the signals 👎 and 👍 are published, and thus the probability that the demand for the product is high or low, respectively.

In general $T \geq 0$, but it is useful to divide the positive real line into different subintervals in order to organize the discussion. More specifically, I identify five zones denoted $\mathcal{T}_1$ to $\mathcal{T}_5$ where $\mathcal{T}_5 = [0, \theta_L)$, $\mathcal{T}_4 = [\theta_L, \theta_H)$, $\mathcal{T}_3 = [\theta_H, \theta_L + 2b)$, $\mathcal{T}_2 = [\theta_L + 2b, \theta_H + 2b)$ and $\mathcal{T}_1 = [\theta_H + 2b, \infty)$. Figure 2.9 depicts these different zones.

There exists a significant difference between the continuous and the binary designs: small levels of manipulation do not have the same impact on the platform signal. When the platform chooses the continuous design, the seller, by adding fake reviews, shifts the support for $\tilde{\alpha}^m(\theta_x)$ to the right. No matter how small the level of manipulation is, the seller knows that it is sufficient to affect the signal realization. In other words, the distribution of $\tilde{\alpha}^m(\theta_x)$ first-order stochastically dominates the distribution of $\tilde{\alpha}(\theta_x)$ for any $m(\cdot)$ and any θ_x.

In comparison, when the platform is using the binary design, it is still true that $\tilde{\alpha}^m(\theta_x)$ first-order stochastically dominates $\tilde{\alpha}(\theta_x)$. Yet, it is possible that small levels of m_x makes no difference to the binary signal realization. For instance, suppose the platform's threshold T is such that $\theta_x + 2b < T$. This implies that without manipulation, i.e., for $m_x = 0$, the platform never publishes the signal 👍 when product quality is θ_x. Suppose next that the seller chooses $m_x = \varepsilon$ instead of 0. Then, if ε is so small such that $\theta_x + 2b + \varepsilon < T$, then manipulation changes nothing to the platform's signal. The probability with which the signal realization is 👍 is still 0.[22]

Consequently, the set of manipulation effort levels where the seller has an impact on the platform signal has a greatest lower bound (which is not a minimum). For $x \in \{H, L\}$, this **greatest lower bound** is denoted $\underline{m}_x(T) \equiv \max\{T - 2b - \theta_x, 0\}$. Specifically, all $m_x \in (0, \underline{m}_x]$ yield exactly the same platform signal distribution as $m_x = 0$.

Similarly, it is possible to define a **least upper bound** of manipulation effort (which is a maximum), $\overline{m}_x(T) \equiv \max\{T - \theta_x, 0\}$. Because the platform publishes the signal 👍 with probability 1 if $m_x \geq \overline{m}_x(T)$, there is no point in choosing an effort level above $\overline{m}_x(T)$. Note that $\underline{m}_H(T) \leq \underline{m}_L(T)$ and $\overline{m}_H(T) \leq \overline{m}_L(T)$ for all T.

All levels of manipulation effort m_x such that $m_x \leq \underline{m}_x(T)$ or such that $m_x > \overline{m}_x(T)$ are never best responses and can be eliminated from the choice set of the seller. And so, after elimination of the strictly dominated strategies, the set of manipulation effort for a seller with quality θ_x that remains is $\{0\} \cup (\underline{m}_x(T), \overline{m}_x(T)]$ for $x \in \{H, L\}$. The value of $\underline{m}_x(T)$ and $\overline{m}_x(T)$ are determined by the threshold T. Hence the value of T plays a crucial role in the equilibrium analysis.

Suppose that the consumers' conjecture on the seller's manipulation efforts is $m^e \equiv (m^e_L, m^e_H)$, such that consumers' posterior beliefs upon seeing the signal $s \in \{$👎, 👍$\}$ are $q_b(s, m^e; T)$

[22] Another possibility where manipulation does not change the platform's signal distribution is when the threshold T is such that $T \leq \theta_x$. In this case, there is no need to manipulate as the signal 👍 is published with certainty for the seller with type x.

and they adopt the threshold strategy $\hat{v}(s, m^e)$ as given in (2.4). Then, the aggregate demand is $D(q_b(s, m^e; T)) = 1 - \hat{v}(s, m^e)$. And so, the expected profits of a seller with quality θ_x are given by

$$\begin{aligned}
&-c_x m_x^2 + D(q_b(\text{👎}, m^e; T)) \\
&\quad + \min\left\{\max\left\{\frac{\theta_x + 2b + m_x - T}{2b}, 0\right\}, 1\right\}\Big(D(q_b(\text{👍}, m^e; T)) - D(q_b(\text{👎}, m^e; T))\Big). \quad (2.20)
\end{aligned}$$

For $m_x \in (\underline{m}_x(T), \overline{m}_x(T))$, the profits are concave in m_x with a local maximum at

$$\hat{m}_x = \frac{D(q_b(\text{👍}, m^e; T)) - D(q_b(\text{👎}, m^e; T))}{4bc_x}. \quad (2.21)$$

Therefore, for a seller with quality θ_x, the candidates for a best-response are $\{0, \hat{m}_x, \overline{m}_x(T)\}$ where $\hat{m}_x$ is given by (2.21). Hence, there are potentially up to nine pairs (m_L, m_H) of manipulation levels that are candidates for a pure-strategy equilibrium. But, before going further, a remark on $\hat{m}_x$ is warranted. In accordance with Definition 2.1, a profile $\left\{(\hat{v}^*(\text{👎}), \hat{v}^*(\text{👍})), (m_L^*, m_H^*)\right\}$ where some type x chooses $m_x^* = \hat{m}_x(\cdot)$ is an equilibrium only if $\hat{m}_x$ is the solution to

$$\hat{m}_x = \frac{D(q_b(\text{👍}, \hat{m}_x, m_{-x}; T)) - D(q_b(\text{👎}, \hat{m}_x, m_{-x}; T))}{4bc_x}, \quad (2.22)$$

that is, Equation (2.21) where m_x^e is replace by the actual $\hat{m}_x$. Hence the value for $\hat{m}_x$ is a fixed point of Equation (2.22).

Proposition 2.5 determines for which threshold T each pair $(m_L, m_H) \in \{0, \hat{m}_L(\cdot), \overline{m}_L(T)\} \times \{0, \hat{m}_H(\cdot), \overline{m}_H(T)\}$ can be an equilibrium of the manipulation game.[23] The conditions provided in the proposition are necessary, but they are not sufficient, i.e., if a pair is an equilibrium, then it satisfies the conditions. The complete set of sufficient conditions, that is the ones on the parameters $(q, \theta_L, \theta_H, c_L, c_H, b)$, are very cumbersome and are only given in Appendix A.5.2.[24]

[23] I show in Proposition 2.7 that a PBE in mixed strategy exists in the manipulation game for any T with the binary design. Here, I concentrate on PBE in pure strategies.

[24] Notice that for a given a set of parameters $(q, \theta_H, \theta_L, c_H, c_L, b, T)$, it is possible for more than one pair (m_L, m_H) of manipulation effort levels to be an equilibrium.

To read the proposition, let $\hat{m}_x(q_b(s,m,T))$ be the solution to (2.22). Consider further that, for the seller with quality θ_x, $\overline{m}_x(T)$ denotes the maximal level of manipulation, and that $\hat{m}_x(q_b(s,m;T))$ denotes a moderate level of manipulation, that is, a level in the interior of $(\underline{m}_x(T), \overline{m}_x(T)]$. Refer also to Figure 2.9 for the definition of $\mathcal{T}_1$ to $\mathcal{T}_5$.

Proposition 2.5. *Assume that the platform is using the binary design with a threshold T. Then, a pair $(m_L, m_H) \in \{0, \hat{m}_L(\cdot), \overline{m}_L(T)\} \times \{0, \hat{m}_H(\cdot), \overline{m}_H(T)\}$ of manipulation efforts is an equilibrium only if the threshold is as specified in the following table:*

(m_L, m_H)	0	$\hat{m}_H(q_b(s,m,T))$	$\overline{m}_H(T)$
0	$\{\mathcal{T}_1, \mathcal{T}_5\}$	$\{\mathcal{T}_1, \mathcal{T}_2\}$	$\{\mathcal{T}_1, \mathcal{T}_2\}$
$\hat{m}_L(q_b(s,m,T))$	$\mathcal{T}_4$	$\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$	$\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$
$\overline{m}_L(T)$	$\mathcal{T}_4$	$\emptyset$	$\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$

The table reads as follows. Take a pair (m_L, m_H), say $(0, \overline{m}_H(T))$. Then, the pair $(0, \overline{m}_H(T))$ is an equilibrium strategy only if the threshold $T \in \{\mathcal{T}_1, \mathcal{T}_2\}$. Note again, however, that $T \in \{\mathcal{T}_1, \mathcal{T}_2\}$ is not sufficient for $(0, \overline{m}_H(T))$ to be an equilibrium.

Let me now highlight some important features of the proposition. First, as soon as the threshold is low enough for a seller with quality θ_x to get the platform to publish the signal 👍 with probability between 0 and 1 without manipulation, then it is optimal for this seller to always manipulate the reviews to some extent. This is because, in this case, the greatest lower bound on the level of manipulation $\underline{m}_x$ reduces to 0 and so even a small amount of manipulation shifts the support of the platform's signal to the right.[25] This is the case for the seller with quality θ_H for $T \in \{\mathcal{T}_2, \mathcal{T}_3\}$ and for the seller with quality θ_L for $T \in \{\mathcal{T}_3, \mathcal{T}_4\}$.

Next, when T is so low such that the platform publishes the signal 👍 with probability 1 even when type θ_x does not manipulate, then $m_x = 0$ is the unique optimal level for the seller. This is the case for the seller with quality θ_H when $T \in \{\mathcal{T}_4, \mathcal{T}_5\}$ and for the seller with quality θ_L when $T \in \mathcal{T}_5$.

Intuitively, it is optimal for a seller to choose the maximal level of manipulation, $\overline{m}_x(T)$, only if c_x, the marginal cost of manipulation, is small enough. Otherwise, a seller chooses a

[25] A shift of the distribution to the right is beneficial for the high-quality seller as it allows him to separate himself from the low-quality seller. A shift of the distribution to the right is beneficial for the low-quality seller as it allows him to mimic the high-quality seller.

moderate level of manipulation, or no manipulation if the benefits are relatively low when compared to the costs. I refer to cost as being sufficiently low/high without being explicit on what the exact requirements are. Once again, refer to Appendix A.5.2 for the conditions for the existence of a pure-strategy equilibrium for each type of manipulation pairs.

A common perception is that a low-quality seller has more incentives to manipulate the reviews than a high-quality seller since it is the low-quality seller that wants to mimic the reviews of the high-quality seller. The current paper shows that it is possible that the low-quality seller manipulates more than the high-quality seller, but it is also possible that the high-quality seller actually manipulates more than the low-quality seller. The fact that the equilibrium manipulation levels are not necessarily monotone in the seller's quality type is also a feature of the continuous design.

Another (related) feature shared by the two designs is that different amounts of information are revealed for different pairs of manipulation levels sustainable in equilibrium. Let $K_x(m_x;T)$ be the probability that the platform publishes the signal 👎 when the seller's quality is θ_x and its manipulation level is m_x. More specifically,

$$K_x(m_x,T) \equiv \mathbb{P}r(s = \text{👎} \,|\theta_x, m_x, T) = \min\left\{\max\left\{\frac{T-\theta_x-m_x}{2b}, 0\right\}, 1\right\}. \tag{2.23}$$

The fact that conjectures are correct in equilibrium together with Bayes' rule imply that the posterior beliefs can be written using $K_x(m_x,T)$. More specifically, for $0 < K_L(m_L,T) < 1$ and $0 < K_H(m_H,T) < 1$,

$$q_b(\text{👎}, m, T) \equiv \frac{q \cdot K_H(m_H,T)}{q \cdot K_H(m_H,T) + (1-q)\cdot K_L(m_L,T)} \tag{2.24}$$

$$q_b(\text{👍}, m, T) \equiv \frac{q \cdot (1-K_H(m_H,T))}{q \cdot (1-K_H(m_H,T)) + (1-q)\cdot (1-K_L(m_L,T))}. \tag{2.25}$$

In particular, when $K_L(m_L,T)=1$, these reduce to

$$q_b'(\text{👎}, m, T) \equiv \frac{q \cdot K_H(m_H,T)}{q \cdot K_H(m_H,T) + (1-q)} \quad \text{and} \quad q_b'(\text{👍}, m, T) \equiv 1 \tag{2.26}$$

and when $K_H(m_H,T)=1$, to

$$q_b''(\text{👍}, m, T) \equiv \frac{q}{q + (1-q)(1-K_L(m_L,T))} \quad \text{and} \quad q_b''(\text{👎}, m, T) \equiv 0. \tag{2.27}$$

Corollary 2.2 uses (2.24) to (2.27) to link the posterior beliefs to the manipulation effort levels. Again, refer to Figure 2.9 for the definition of $\mathcal{T}_1$ to $\mathcal{T}_5$.

Corollary 2.2. *Assume that the platform is using the binary design with a threshold T. If (m_L^*, m_H^*) is the seller's pure-strategy in an equilibrium of the manipulation game, then the posterior beliefs $(q_b(\text{👎}, m^*, T), q_b(\text{👍}, m^*, T))$ are given in the following table:*

(m_L^*, m_H^*)	0	$\hat{m}_H(\cdot)$	$\overline{m}_H(T)$
0	$(q, \circ)$ *if* $T \in \mathcal{T}_1$ $(\circ, q)$ *if* $T \in \mathcal{T}_5$	$(q_b'(\text{👎}, (0, \hat{m}_H), T), q)$	$(0, 1)$
$\hat{m}_L(\cdot)$	$(0, q_b''(\text{👍}, (\hat{m}_L, 0), T))$	$(q_b(\text{👎}, (\hat{m}_L, \hat{m}_H), T), q_b(\text{👍}, (\hat{m}_L, \hat{m}_H), T))$	$(0, q_b''(\text{👍}, (\hat{m}_L, \overline{m}_H), T))$
$\overline{m}_L(T)$	$(\circ, q)$	$\emptyset$	$(\circ, q)$

The symbol $\circ$ is used to denote beliefs that are off the equilibrium path.

Two properties of the consumers' learning trajectory need to be emphasized. First, when both types of seller choose not to manipulate the reviews, consumers learn nothing more than what they already knew. This happens in two types of situations: when the threshold is so high that it is too costly to engage even in the minimal level of manipulation, and when the threshold is so low that there is simply no need to manipulate the reviews. There is also no learning in equilibrium as soon as the low type manipulates to its maximal effort level $\overline{m}_L(T)$.

Second, for every other possible equilibrium strategy (m_L, m_H), consumers revise their beliefs about product quality. In particular, it can be the case that consumers learn product quality perfectly after observing the platform signal. That is, full separation of the quality types is possible and it occurs when the pair $(0, \overline{m}_H(\cdot))$ is an equilibrium strategy for the seller. This particular case is possible when the threshold is high enough such that the low-quality seller prefers not to manipulate and the high-quality seller chooses its highest level of manipulation effort.

2.6.5 The Design's Impact on Information Dissemination

In order to analyze the design's influence on the overall dissemination of information, it is necessary to compare the consumers' learning trajectory with a continuous design to the case of a binary design. Before tackling this comparison, I briefly discuss whether manipulation jeopardizes the transmission of information to consumers.

2.6.5.1 With Manipulation versus No Manipulation

It is often advocated that a platform hosting reviews should try to eliminate manipulation because it deteriorates information quality and is detrimental to consumers. This is a path taken by computer scientists, among others, who try to develop algorithms to filter out suspicious reviews (e.g., Yang *et al.*, 2007; Ott *et al.*, 2011). The results of Sections 2.6.3 and 2.6.4 establishes that manipulation can be eliminated in equilibrium with the binary design only for very specific threshold levels. Indeed, unless the platform is using the binary design with a very high or very low threshold, there is always a type of seller that finds it optimal to exert a positive level of manipulation effort. In other words, the only way for the platform to eliminate manipulation in my model is not to transmit any information whatsoever.

Corollary 2.3. *In equilibrium, review manipulation is completely eliminated only if the platform uses the binary design and fixes the threshold to* $T \in \{\mathcal{T}_1, \mathcal{T}_5\}$.

If $T \in \mathcal{T}_1$ ($\mathcal{T}_5$) and no type of seller manipulates the reviews, then the platform publishes the signal 👎 (👍) with probability 1. In this case, the posterior beliefs are equal to the prior so that the consumers learn nothing from the reviews and the value of information provided by the platform is zero for all consumers. Hence, a platform that eliminates manipulation must accept that this also comes at the price of eliminating the informative content of reviews.

For this reason, one can argue that manipulation is not necessarily bad for information dissemination. After all, Proposition 2.3 does state that it is the manipulation efforts of a low quality seller that deteriorate information, not those of a high-quality seller.[26] As it turns out, it is possible that consumers learn more when the seller manipulates reviews

[26] Generally, manipulation of reviews is a device that can increase or decrease the amount of information that is transmitted to consumers. The fact that informativeness can vary in both direction is due to the presence of the genuine review which act as an exogenous signal. Without genuine reviews, information can only be improved.

than when he cannot. One particular instance of this result is when perfect separation of quality types occurs in equilibrium.

Corollary 2.4. *Assume that the platform is using the continuous design. Then, there exists an open set of parameters* $(\theta_L, \theta_H, q, c_L, c_H, b)$ *such that a pair* (m_L^*, m_H^*) *of manipulation efforts allowing for perfect revelation of quality is a perfect Bayesian equilibrium of the manipulation game.*

Assume that the platform is using the binary design. Then, there exists an open set of parameters $(\theta_L, \theta_H, q, c_L, c_H, b, T)$ *with* $T \in \{\mathcal{T}_1, \mathcal{T}_2\}$ *such that a pair* (m_L^{**}, m_H^{**}) *of manipulation efforts allowing for perfect revelation of quality is a perfect Bayesian equilibrium of the manipulation game.*

This corollary means that consumers do not necessarily make poor choices on the basis of fake reviews. Indeed, they may end up with perfect information, and this allows them to make the best possible decision.

2.6.5.2 Continuous Design versus Binary Design

In Section 2.5.2, the benchmark case of no manipulation is analyzed and it is noted that the continuous design is necessarily more informative than the binary design. In this section, I perform the same kind of analysis and compare the informational properties of the binary and continuous design, but this time accounting for manipulation.

As noted at the end of the last section, manipulation is not necessarily bad for consumers as it may be possible that they fully learn the product quality. The next lemma relates this positive outcome of manipulation to the type of design that the platform is using. In particular, it establishes that the binary design is *less likely* to be associated to an equilibrium with complete learning.

Lemma 2.5. *If complete learning cannot occur in equilibrium with the continuous design, then it cannot occur with the binary design either.*

From Lemma 2.5 alone, it would be tempting to conclude that the continuous design has an advantage over the binary design in terms of information dissemination. Yet, it remains to determine what is the information flow associated to each design when complete learning is not possible. Proposition 2.6 shows how the continuous design compares to the binary design with a threshold T for all events, including events where complete learning is impossible.

Proposition 2.6. *Let $\mathcal{E}_C$ and $\mathcal{E}_{B,T}$ be the set of equilibria in pure strategies of the manipulation game with a continuous and binary design with threshold T, respectively. Moreover, let $(m_L^c, m_H^c) \in \mathcal{E}_C$ and $(m_L^b, m_H^b) \in \mathcal{E}_{B,T}$ be the elements that are associated with the highest information value in $\mathcal{E}_C$ and $\mathcal{E}_{B,T}$, respectively. Then, the following situations are possible:*

a) (m_L^c, m_H^c) is associated to the <u>same</u> information value than (m_L^b, m_H^b)

b) (m_L^c, m_H^c) is associated to a <u>greater</u> information value than (m_L^b, m_H^b);

c) (m_L^c, m_H^c) is associated to a <u>lower</u> information value than (m_L^b, m_H^b).

Cases *c*) is of particular interest because it is in direct contradiction with the result established in Section 2.5.2. It says that, because of the seller's manipulation efforts, the platform can transmit more information to consumers by using the coarser signal of reviews. To understand how this is possible, there are two effects to consider when moving from the continuous design to the binary design: a direct effect and an indirect effect.

<u>The Direct Effect</u>

On the one hand, when garbling the reviews with the binary design, the platform has a direct negative effect on the information content of the platform signal. That is, it is possible that the realization consumers would observe with the continuous signal can reveal perfectly product quality. But, by using a binary signal instead, the platform turns out to only transfer partial information on product quality. Figure 2.10 depicts an instance where the realization s would reveal that $\theta = \theta_H$ with the continuous design, but where a binary design with a threshold above s only transfers partial information.

<u>The Indirect Effect</u>

On the other hand, by garbling the reviews, the platform has an indirect effect on information dissemination. This indirect effect operates through the seller's strategic choice of manipulation efforts (an effect that is inexistent in the benchmark case of no manipulation). Proposition 2.6-c) is true when this indirect effect is positive and overcompensates for the direct negative effect.

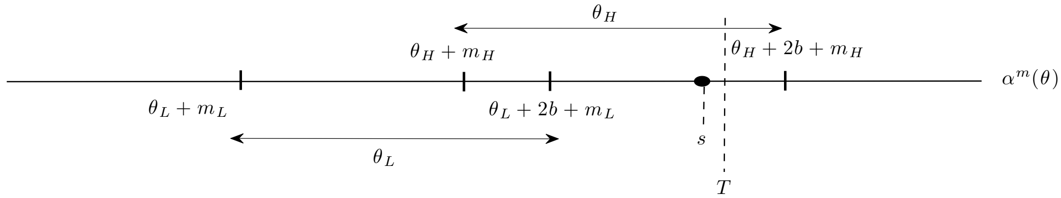


Figure 2.10: The realization s would reveal that $\theta = \theta_H$

The indirect effect has a positive impact on information dissemination if the seller is deterred from obfuscating the reviews as much as he would otherwise if the platform were to use the continuous signal. To see this, fix the parameters $(\theta_L, \theta_H, q, c_L, c_H, b, T)$ and suppose that (m_L^c, m_H^c) and (m_L^b, m_H^b) are equilibria in pure strategies of the game with a continuous and binary design with threshold T, respectively. Then, by comparing $m_H^c - m_L^c$ to $m_H^b - m_L^b$ one can assess which design is most detrimental to information dissemination.

With the continuous design, when complete learning is impossible, the only other equilibrium outcome (with pure strategies) is that nothing is learned (cf. Corollary 2.1). That is, the low-quality seller destroys all the information and obfuscates the signal completely. Specifically, the equilibrium manipulation efforts (m_L^c, m_H^c) have the feature that $m_H^c - m_L^c < 0$.

With the binary design, depending on the threshold T, even though complete learning is impossible, there may exist an equilibrium in pure strategies associated to partial learning. The form of this equilibrium strategy is either $(0, \hat{m}_H(\cdot))$, $(\hat{m}_L(\cdot), 0)$, $(\hat{m}_L(\cdot), \hat{m}_H(\cdot))$ or $(\hat{m}_L(\cdot), \overline{m}_H(T))$. Each of these equilibrium candidates has the feature that $m_H^b - m_L^b > m_H^c - m_L^c$. Moreover, not only the difference $m_H^b - m_L^b$ can be less negative than $m_H^c - m_L^c$, it can even become positive. When $m_H^b - m_L^b$ is negative but less negative than $m_H^c - m_L^c$, the seller is deterred to some extent from clouding the reviews. When, instead, $m_H^b - m_L^b$ becomes positive, the seller is now even incited to increase the information transmitted by reviews.

The reason why the low type is restrained from exerting too much manipulation effort with the binary design is that the marginal return to manipulation is smaller than with the continuous design. Indeed, with the continuous design, consumers see the review statistic

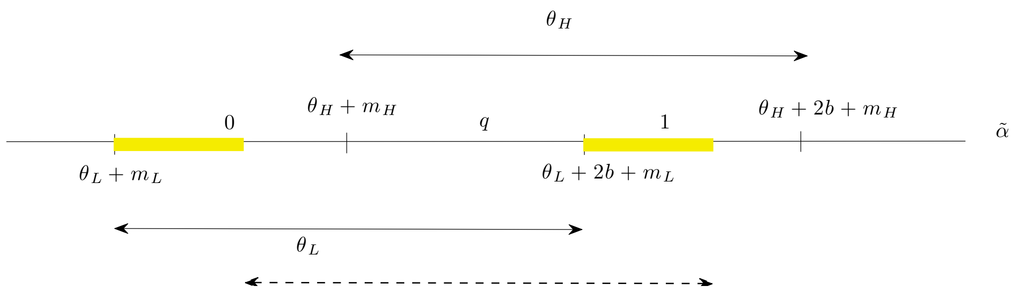


Figure 2.11: Marginal return with the continuous design

α^m, which varies continuously and ranges from very low level to very high level. By exerting effort to manipulate reviews, a low-quality seller is trading off chances that $\alpha^m(\theta_L)$ be very low, which is associated with low posteriors, for chances that $\alpha^m(\theta_L)$ be very high which is associated to high posteriors. This procures the seller the largest possible marginal gain. Figure 2.11 illustrates the gain for the low-quality seller when he deviates from m_L to $m'_L > m_L$. The yellow area at the left of the figure (where posteriors are 0) is the portion of the support for $\alpha^m(\theta_L)$ that is replaced by the yellow area (where posteriors are 1) at the right of the figure following the deviation.

With the coarser binary design, consumers see only whether the product is recommended or not, which can be seen only as a moderately good or moderately bad signal. By exerting effort to manipulate reviews, a low-quality seller is now trading off chances to get the signal 👎, which is associated to moderately low posteriors, for chances to get the signal 👍, which is associated to moderately high posteriors. The marginal return to manipulation is thus smaller than with the continuous design. As a result, the low-quality seller is deterred from clouding the reviews with the binary design. Figure 2.12 illustrates the gain for the low-quality seller when he deviates from m_L to $m'_L > m_L$. The yellow area at the left of the figure (where posteriors are $q_b($👎$, m; T))$ is the portion of the support for $\alpha^m(\theta_L)$ that is replaced by the yellow area (where posteriors are $q_b($👍$, m; T))$ at the right of the figure following the deviation.

Another aspect explaining the result is that, with the binary design, there may be a minimum level of manipulation effort the seller needs to engage in to get over the threshold T with some positive probability. This minimal level of manipulation implies that there are minimal costs to incur if a seller wants to manipulate. In other words, there is a fixed cost to pay and the higher the threshold is, the higher the fixed cost is. Of course, such

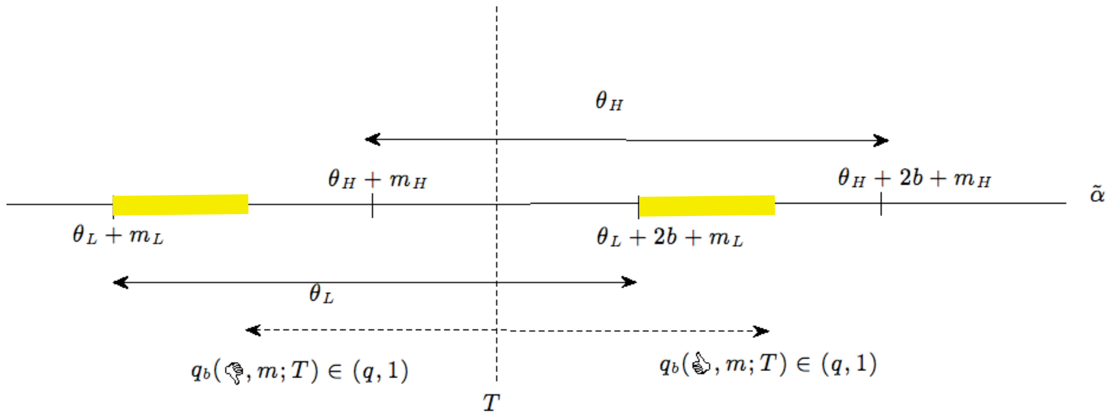


Figure 2.12: Marginal return with the binary design

a fixed cost tempers the incentives to exert manipulation efforts for both type of sellers. The low-quality seller, however, is penalized more severely as, even if c_L is the same as c_H, the highest-end of his support for $\alpha^m(\theta_L)$ is farther away from T. Therefore, a low-quality seller that would want to destroy information would have to incur a much higher cost with the binary design than with the continuous design. It can be suboptimal for him to do so in the former case while it is not in the later.

In a nutshell, the impact of the platform design on information dissemination is the following. Without manipulation, the continuous design always maximizes the value of information. When manipulation is taken into account, the design that maximizes information dissemination can differ dramatically from the one when there is no manipulation. Indeed, if the seller can engage in manipulation, better information may be transmitted through the binary design as it channels manipulation in such a way that less information is destroyed.

2.7 The Platform's Design Choice

2.7.1 The Role of the Business Model

The last section makes the important point that the design used by the platform has an impact on the seller's manipulation efforts and, therefore, on the quality of information that is conveyed to consumers. A fundamental question is then related to what the optimal design is from the platform's perspective. After all, a platform is not a benevolent agent and chooses a design in order to maximize its own profits.

In this section, I link the way a platform's revenues are generated, i.e., its business model, to its choice of design. In particular, I focus on three business models that are common in the industry: a platform that obtains revenues from selling advertising space on its website, a platform that generates revenues from user-paid subscriptions, and a platform that receives commissions from sellers.[27,28,29] I classify platforms in two categories depending on their business model. A platform that generates revenues through sales commissions is called a *transactional platform* and one that generates revenues though advertising or subscription fees is called a *non-transactional platform*.

Transactional and non-transactional platforms, because they differ in their way of generating revenues, may seek for a design with different characteristics. Consider first the case of a non-transactional platform. One can claim that, in this case, the platform cares about providing information with a maximal value for date-2 consumers. Indeed, the higher the value (quality) of information for consumers, the higher the number of consumers who visit the site and the longer they browse on the website. Thus, the platform's ability to attract advertisers or subscribers is positively correlated with the quality of information.[30] Therefore, it can be assumed that a non-transactional platform seeks to maximize the value of information for consumers.

Next, consider the case of a transactional platform. If the platform collects a commission fee on each item sold by the seller, then commissions are proportional to the number of

[27] For example, the biggest revenue source for Yelp comes from local advertisers. According to Forbes, in 2014, the local ads business accounts for over 75% of Yelp's stock value. Retrieved from `http://www.forbes.com/sites/greatspeculations/2014/02/07/yelp-earnings-revenue-growth-keeps-up/`.

[28] Angie's List and Consumers' Checkbook are examples of platforms that require the consumers to pay to subscribe to the site to read reviews.

[29] The Amazon Market Place operates under the commission scheme. The sellers must pay Amazon referral fees and selling fees. The online-marketplace Etsy, which specializes in crafts and other artistic items by artisans, charges 3.5% of the value of the sale when an item sells. Finally, Expedia charges a fee to the hotels for each booking.

[30] Strictly speaking, the consumers are not making the choice to visit the platform or not in the model. But it could be argued that if the consumers expect to obtain valuable information, then they have incentives to visit the platform's site. If they expect, however, to learn invaluable information, it makes sense to assume that they will turn to other learning mechanisms.

sales. Hence it can be assumed, at least on a short-term horizon, that a transactional platform seeks to maximize the number of sales.[31]

In Section 2.7.2, I show that the information maximization objective of a non-transactional platform and the sales maximization objective of a transactional platform drive the platform toward completely different designs. In other words, transactional and non-transactional platforms value the informativeness of reviews differently.

The results in Section 2.7.2 should, however, be interpreted as observations on the platform's incentives on a short-term horizon. Indeed, the analysis of a platform's design choice from a long-term perspective is complicated in at least two ways. First, one needs to account for how the design choice in the first period influences traffic in subsequent periods. Second, one also needs to consider that review manipulation in the first period has an impact on the mass of consumers who buy the product, which has an effect on reviews that are left in subsequent periods. To keep the model simple, I focus on the short-horizon case alone and do not consider what the optimal design is when agents, including the platform, are forward-looking.

2.7.2 A Partial Characterization of the Optimal Design

Once the design is endogenous, the platform becomes the first player to move in the game (c.f. timeline in Figure 2.1). Its design decision is taken at the ex ante stage, that is before $\alpha^m(\cdot)$ is realized. In order to choose the design it prefers, a platform needs to answer the following questions: Which binary design is appropriate? In other words, what is the best level for the threshold T? Then, is the binary design better than the continuous design?

By moving first in the game, the platform acts as a Stackelberg leader and its design choice influences the manipulation in which the seller engages in. Note that how the seller manipulates reviews is independent of the platform's business model. In particular, by choosing the continuous design, any type of platform induces the seller to manipulate in

[31] Notice that maximizing the number of sales is not necessarily the same as maximizing the seller's profits since the latter incurs a cost for manipulating reviews. The platform, by contrast, only gets the benefits of manipulation (if any), but does not have to incur the cost. Consequently, the seller's and a transactional platform's interests are perfectly aligned only if it is impossible to manipulate the reviews.

such a way that only two outcomes are possible: Either consumers learn product quality, or they learn nothing (c.f. Corollary 2.1). By choosing the binary design, any type of platform can induce the seller, through its choice of T, to adopt different levels of manipulation that generate an array of consumer posterior beliefs, from complete learning to no learning (c.f. Corollary 2.2).

One comment is now warranted about the platform's optimization problem. The platform can decide on what its optimal design is if and only if there exists a profile $\{(\hat{v}(s))_{s \in \mathcal{M}}, (m_L, m_H)\}$ (or a mixed strategy profile) that is a PBE according to Definition 2.1 when the design is continuous, as well as when it is binary for any T. The existence of an equilibrium, however, is not straightforward because (a) the seller's payoff function is not always quasi-concave, and (b) the consumers' payoff function is discontinuous for some of the seller's strategy profile. Indeed, a discontinuity in consumers' payoff exists every time a pair (m_L, m_H) of manipulation efforts implies that a signal s occurs with zero probability, a case that requires the specification of out-of-equilibrium beliefs. Despite such discontinuities, using a result due to Reny (1999), it is possible to show that an equilibrium exists for each design.

Proposition 2.7. *For any design choice* $(\mathcal{M}, S) \in \Big\{(supp(\alpha^m), S_c), (\{$ 👎 , 👍 $\}, S_b(;T))\Big\}$, *there exists a Perfect Bayesian Equilibrium (in mixed strategies) of the manipulation game.*

Yet, even with the existence issue for the manipulation game being settled such that a mixed-strategy equilibrium for the continuous and binary design are guaranteed to exist, a complete characterization of the platform's optimal design remains beyond the scope of this paper for two reasons. The first reason is that in order to compare the informational properties of a mixed-strategy equilibrium of the continuous design to the ones of a mixed-strategy equilibrium of the binary design, it is required to know the exact mixed strategies which is not a straightforward task. The second reason is that multiple equilibria of the manipulation game exist and so, to characterize the optimal design, it is necessary to understand how the parameters affect which equilibria may arise. This issue is also very complex and could, at best, be approached by performing extensive numerical simulations.

In case of equilibrium multiplicity in the manipulation game for a given design, I adopt the best-case approach of Kamenica and Gentzkow (2011) and Taneva (2014) and select the equilibrium which yields the highest expected payoffs to the platform. That is, the one that maximizes the value of information in the case of a non-transactional platform and the one that maximizes the expected number of sales in the case of a transactional platform. The platform's problem can then be summarized by the following steps:

1) Take the continuous design. Then, characterize the set of all $(\hat{v}(\cdot), (m_L, m_H))$ that can be sustained in a PBE of the manipulation game. In case of equilibrium multiplicity, select (the) one that maximizes the platform's expected payoff.

2) Fix $T \in [0, \ddot{T}]$ and take the binary design.[32] Then, characterize the set of all $(\hat{v}(\cdot), (m_L, m_H))$ that can be sustained in a PBE of the manipulation game. In case of equilibrium multiplicity, select (the) one that maximizes the platform's expected payoff.

3) Repeat Step 2) for all $T \in [0, \ddot{T}]$.[33]

4) Among all $(\hat{v}(\cdot), (m_L, m_H))$ identified in steps 1), 2) and 3), select (the) one design that is associated to a pair $(\hat{v}(\cdot), (m_L, m_H))$ for which the expected value of the platform's objective function is maximal.

The complete definition of an equilibrium can be found in Appendix A.1.

Although I do not provide a complete characterization of the platform's optimal design, some interesting features of this choice can still be presented, particularly by contrasting the choice of a transactional platform to the choice of a non-transactional platform.

2.7.2.1 Transactional Platform

In Section 2.7.1, it is argued that a transactional platform's objective is the maximization of sales. In contrast to the case of a non-transactional platform, a transactional platform's objective is not stated in terms of preference in information provision. Thus, this raises

[32] For all $T \geq \ddot{T}$, the only equilibrium of the manipulation game is $(m_L, m_H) = (0, 0)$. More specifically, $\ddot{T} = \max\{\ddot{T}_L, \ddot{T}_H\}$ where $\ddot{T}_x = \max\{\sqrt{1/c_L} + \theta_L, 1/8bc_L + \theta_L + 2b\}$ for $x \in \{L, H\}$.

[33] There is no need to check for $T > \ddot{T}$ since, in this case, the only equilibrium of the manipulation game is $(m_L, m_H) = (0, 0)$. This implies that the outcome for a threshold $T > \ddot{T}$ can be reproduce with the binary design with a threshold $\ddot{T}$.

the question of whether or not such a platform benefits from providing information to consumers.

In order to determine what a transactional platform's preferences regarding information provision are, I use one of the results of Kamenica and Gentzkow (2011) about Bayesian persuasion. Indeed, the question of how much information the platform wants to transfer to consumers is akin to the question Kamenica and Gentzkow ask about conditions under which a Sender benefits from persuading a Receiver. Remark 1 in their paper is specifically relevant. Indeed, reinterpreted in the context of my model, their remark implies that a platform never benefits from transferring information to consumers if its expected revenues are concave in the consumers prior beliefs. It turns out that given the model's specific assumptions, expected revenues for a transactional platform are concave in the prior q.

Proposition 2.8. *A transactional platform does not benefit from transferring information to consumers. That is, its expected revenues are at their highest when the posterior beliefs are equal to the prior beliefs and at their lowest when consumers obtain complete information.*

The fact that consumers have the opportunity to fully learn product quality when there is manipulation does not benefit the platform. In fact, perfect revelation of quality is the worst possible scenario for a transactional platform. That is, any other posterior beliefs that are possible given the seller's optimal strategy, according to Corollary 2.1 or to Corollary 2.2, would generate more revenues for the platform.

There are three specific designs that produce the no-learning equilibrium outcome in the manipulation subgame: (a) by using the binary design and setting a threshold $T \leq \theta_L$, (b) by using the binary design and setting a threshold $T \geq \theta_H + 2b$ and (c) by using the continuous design so that a low-quality seller exerts manipulation effort $\theta_H - \theta_L + m_H$, where m_H is the high-quality seller manipulation effort.[34] Note that the designs in (a), (b) and (c), although they have the same posterior beliefs structure, are different in terms of levels of manipulation they induce. Still, the platform is concerned by manipulation only

[34] Although the design described in (a) is always associated with an equilibrium in the manipulation game such that there is no-learning, it is not always the case however that an equilibrium with no-learning of the kind described in (b) or (c) exists. For the binary design, even though the threshold is very high, if the manipulation cost is too low for some type, then this seller will choose to exert positive manipulation effort. The conditions guaranteeing that an equilibrium with $(m_H, m_L) = (0, 0)$ when $T \geq \theta_H + 2b$ exists are given in Appendix A.5.2.1. For the continuous design, see Appendix A.4.2.

because it affects the posterior beliefs of the consumers. Therefore, it is without loss for the platform to choose a binary signal with a very low threshold $T \leq \theta_L$.

Corollary 2.5. *Assume that the platform is transactional. Then, it is always optimal for the platform to commit to using the binary design with $T \leq \theta_L$.*

Corollary 2.5 implies that in equilibrium, when the platform is transactional, consumers obtain no information from reviews. Corollary 2.6 formalizes this claim.

Corollary 2.6. *If $\{(\mathcal{M}^*, S^*), \hat{v}^*(\cdot), m^*(\cdot)\}$ is a PBE, then consumers' posterior beliefs are $q_{S^*}(s; m^*) = q$ for all $s \in \mathcal{M}^*$ on the equilibrium path.*

2.7.2.2 Non-Transactional Platform

A non-transactional platform seeks to maximize the value of information for date-2 consumers. In contrast to the case of a transactional platform, the possibility that information be complete as a result of the seller's manipulation efforts now benefits the platform.

Proposition 2.9. *In the absence of manipulation, it is optimal for a non-transactional platform to use the continuous design. With manipulation, it is not necessarily optimal for the platform to commit to using the continuous design.*

In the case of a non-transactional platform, when manipulation is taken into account, the optimal design can differ dramatically from the one when there is no manipulation. Without manipulation, the continuous design is strictly more informative and thus, it is always optimal for the platform. If the seller can exert effort to manipulate reviews, the platform might be better off by using the binary design instead. The reason is that, as Proposition 2.6 has established, the binary design with its coarser message space channels manipulation in such a way that less information is destroyed.

Assuming the cost of manipulation is the same for both quality types, one can find a sufficient condition under which the continuous design continues to be optimal for a non-transactional platform even when there is manipulation.

Lemma 2.6. *Assume that $c_L = c_H = c$ and that demand is $D(y)$ when posterior beliefs are y. Then, if $\theta_H - \theta_L - 2b + \frac{D(1)-D(0)}{8bc} \geq 0$, it is optimal in the game with manipulation for the platform to commit to using the continuous design.*

Note that the condition is only sufficient and the continuous design might continue to be optimal for a larger set of parameters.

In short, when the platform conducts business by receiving advertising revenues or subscription fees, none of the possible designs are dominant nor dominated in general. Providing comprehensive information on the reviews by using the continuous design or only partial information by using the binary design can both be optimal for different parameter sets.

2.8 Conclusion

As consumers increasingly rely on reviews to guide their decisions, it becomes important to examine whether such content can be trusted and what factors can impact the information content of reviews. Sellers trying to game the system by posting fake reviews for themselves is a first-order concern. Nonetheless, my work suggests that the economic incentives of the platform hosting the reviews also play an important role. Indeed, to understand how much of a problem manipulation is, one need to consider the strategic interaction between platforms and sellers, as the choice of how to present information has an influence on the extent of manipulation.

Two questions are at the core of this paper. First, which design maximizes information transmission given reviews can be manipulated? Second, what is the relation between the platform's business model and its design choice? My analysis establishes that the answer to the first question is not necessarily the design that provides the maximal details on reviews. In some situations, because of review manipulation, the platform induces more information to be revealed by voluntary disclosing less on the reviews' content. The binary design (which wastes information) turns out to be beneficial as it induces the seller to destroy less information with manipulation in the end. As for the second question, this work shows that a platform collecting revenues from advertising or subscription fees and a platform collecting revenues from sales commissions differ in their incentives for choosing a particular design, as they have different preferences regarding information dissemination.

More generally, the current work also establishes that online platforms should not be considered as neutral third parties. The way they conduct business has a crucial importance for the information obtained by consumers. On that account, one could say that the adverse selection problem faced by consumers can become more acute with certain types of platforms, i.e., transactional ones. This observation is particular relevant to the regulation of Internet markets. Competition agencies, which are concerned with sellers hiring shills to boost their reviews, should perhaps be also interested in monitoring transactional platforms more closely than non-transactional ones.[35] Indeed, even though it could be argued that competition across transactional platforms will put upward pressure on information dissemination, one can be worried that in response, platforms will try to collude to weaken these pressures.

Finally, I see several extensions of the model that would further our understanding of information transmission on the Internet. First, it would be interesting to explore which design maximizes social welfare. The answer to that question is not clear since (a) consumers and the platform have aligned preferences only if the platform is non-transactional, and (b) it is not clear how the seller's profits are related to the amount of manipulation. Indeed, Dellarocas (2006) suggests that sellers are trapped in a rat-race and that they are the ones that have the more to gain from the limitation of review manipulation.

Second, the result that a transactional platform prefers to reveal as little information as possible is largely driven by the specificities of the model. It would be interesting to see what happens for a more general model. One might also wonder what would happen if the platform is allowed to use more complex signal rule, say rules with more than one threshold. Third, it would be worthwhile to extend the model to include several platforms and several sellers where both platforms and sellers compete for consumers. Multi-homing, i.e., consultation of reviews on multiple platforms, is a phenomenon that could appear to influence platforms' incentives to provide information and seller's incentives to hire shills. In addition, there is also the issue that with competition, a seller can boost its own reviews but can also denigrate a competitor's product. It would be interesting to address this possibility in future research.

At last, as I have already mentioned, this paper focuses on the short-term horizon. When the long-term comes into play, an important consideration is that review manipulation, by affecting who buys today, will affect who will post review in the next period. A seller

[35]See the announcement of Canada's Competition Bureau at `http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/03782.html`

needs to acknowledge this additional effect. On top of that, the analysis of a dynamic model would also make it possible to explore more complex, yet important, design issues like the need for a differential treatment of newer and older reviews.

Chapter 3

Noisy Learning and Price Discrimination: Implications for Information Dissemination and Profits

3.1 Introduction[1]

As long as there is some easily observable characteristic (e.g., age, income, or geographic location) by which a firm can group buyers and arbitrage can be prevented, it is possible for the firm to segment markets and engage in (third-degree) price discrimination. An important question is whether market segmentation is beneficial for society. The welfare analysis on market segmentation has generally been undertaken under the assumption of complete information on the part of consumers.[2] In this case, the welfare effects are ambiguous. One has to weight the losses of consumers in low-elasticity markets against the gains of those in high-elasticity markets and the gains of the firm. Moreover, one has to consider that discriminatory pricing may lead to the opening of new markets.

In the case of incomplete information, little is known about the effect of market segmentation on welfare and especially on consumers' well-being. This is relevant since the differences among the segmented groups might concern not only tastes, but also information regarding the quality of the good. For instance, with the spread of online commerce, it becomes easier for a firm to introduce a product in a new market. Consumers in the new market might have tastes for the product that differ from consumers' tastes in the original market and they might have less information because of the novelty of the brand. Another example is the case of a prescription drug readily available in the US which is introduced in a developing country. In addition of being able to pay less for the drug, consumers in the developing country might be less informed about the effectiveness of the prescription drug.[3]

The introduction of asymmetric information among buyers leads naturally to the issue of the informative role of prices. Indeed, prices have been shown to be instrumental in disseminating information to market participants (Grossman, 1989).[4] One of the purpose

[1]This chapter is co-authored with Marc Santugini. We thank Masaki Aoyagi, Sidartha Gordon, and Josel Santugini for helpful comments.

[2]See Armstrong (2006) for a survey on price discrimination. See Schmalensee (1981) and Tirole (1988) for a detailed discussion on third-degree price discrimination.

[3]The implementation of drug information centers is a primary concern in many developing countries (Flores Vidotti, 2004). Proper sources of information on drugs are not easily accessible in developing countries. There are several reasons for the absence of information: inadequate translation in local languages, prohibitive cost to acquire information, and even customers' unawareness on how to obtain information.

[4]Several studies have provided conditions under which privately-held information by firms becomes public through prices, beginning with perfectly competitive markets (Kihlstrom and Mirman, 1975; Grossman,

of this paper is to study the effect of market segmentation on the informational content of prices. Specifically, does discriminatory pricing provide less or more information to the uninformed buyers? If it were not for the endogeneity of prices, it could be argued that an increase in the number of price-signals due to market segmentation yields more information to the uninformed buyers (i.e., more precise posterior beliefs). However, since the firm sets prices, the distribution of the price-signals (i.e., the informational content) does depend on whether the firm uses discriminatory pricing. There is thus a trade-off. Price discrimination generates more price-signals, but each of these signals might be less precise.

To study the effect of discriminatory pricing on the dissemination of information via market prices, we consider the simplest model of third-degree price discrimination of a monopoly selling a homogeneous good to two separate markets. In one of the markets, some buyers do not know the quality of the good. Yet, the presence of informed buyers makes it possible for prices to disseminate information. Under noisy demand, we show that market segmentation alters the informational content of price-signals received by the uninformed buyers. Specifically, discriminatory pricing have informational benefits over uniform pricing, i.e., the posterior beliefs of the uninformed buyers have a smaller bias and a lower variance.

The introduction of incomplete information also raises questions on the profitability of third-degree price discrimination from the firm's perspective. It is the second purpose of this paper to address this issue. In an environment of complete information, it is always profitable for a monopoly to segment markets with different demands and to engage in third-degree price discrimination. The reason is that setting different prices – a lower price in the market segments with greater price elasticity and a higher price in those with lower price elasticity – allows the firm to capture more of the consumer surplus. However, it is not known whether market segmentation is systematically profitable for the firm under incomplete information.

In this paper, by endogenizing the firm's decision to segment or integrate the market, we show that, when confronted with uninformed buyers, market segmentation is not necessarily

1976, 1978; Grossman and Stiglitz, 1980) and continuing with imperfectly competitive markets (Wolinsky, 1983; Riordan, 1986; Bagwell and Riordan, 1991; Judd and Riordan, 1994; Daughety and Reinganum, 1995, 2005, 2007, 2008a,b; Janssen and Roy, 2010; Daher et al., 2012).

the more profitable pricing strategy. This is because the firm faces a trade-off when choosing to segment or integrate the markets. On the one hand, market segmentation yields more flexibility and the ability to capture a greater share of the consumer surplus. On the other hand, market segmentation implies that the firm signals quality with two prices instead of one. Hence, two prices are distorted from their complete information counterpart, whereas only one price is when markets are integrated. If the signaling cost (due to the distortion in the prices) is higher under market segmentation than under market integration, then it is possible that the loss due to signaling outweighs the benefit from price flexibility. We find that the higher the number of informed buyers, the more similar the market segments have to be for market integration to be the more profitable option. We also find that it is more likely that market integration be optimal when uninformed buyers are numerous and originate from the market segment with the higher willingness to pay.

The remainder of the chapter is organized as follows. Section 3.2 surveys the literature. Section 3.3 presents the informational benefit of discriminatory pricing for the uninformed buyers. Section 3.4 studies the profitability of discriminatory pricing. Finally, Section 3.5 concludes.

3.2 Literature

We contribute to a large literature on third degree price discrimination, starting with the classic work of Pigou (1920). The reminiscent question with market segmentation is related to the conditions under which price discrimination raises welfare. For instance, Schmalensee (1981) and Varian (1985) identify when an increase in output is necessary for an increase in welfare, whereas Nahata et al. (1990), instead on focusing on output, concentrate on the price effects of discrimination. By contrast, Aguirre et al. (2010) and Cowan (2013) identify sufficient conditions for price discrimination to increase welfare. More recently, Bergemann et al. (2015) show that there is always a way to segment the market such that the combination of consumer and producer surplus satisfy the following: (i) consumer surplus is nonnegative, (ii) producer surplus is at least as high as profits under the uniform monopoly price, and (iii) total surplus does not exceed the efficient surplus.

We depart from this literature in two ways. First, by introducing price discrimination in an environment where some buyers have incomplete information. Second, by concentrating on the signaling aspect of prices instead of on the welfare effects of price discrimination.

We are the first, to our knowledge, to analyze the issue of market segmentation in a stochastic environment with learning. There is, however, a small but growing literature on learning in a stochastic setting. Matthews and Mirman (1983) study a limit pricing environment. Judd and Riordan (1994) and Mirman et al. (2014b) study noisy learning in the monopoly case whereas Mirman et al. (2014a, 2015) study the informational role of prices in competitive markets. **?** study how the communication strategy in a sender-receiver game affects the amount of information that is transmitted given that communication is inherently noisy. It is interesting to note that the use of noisy signaling models is growing, maybe partly because recent experimental work suggests that the stochastic environment in signaling maps better into the behavior of experimental subjects (de Haan et al., 2011; Jeitschko and Norman, 2012).

In this paper, we contribute to the literature on noisy signaling by studying the informational role of prices in the presence of market segmentation. The noisy environment enables us to study thoroughly the effect that market segmentation has on the informational content of prices. In a noiseless environment, the firm reacts to the informational externality, but has limited control over the flow of information. In other words, in equilibrium, either the unknown quality is not revealed and uninformed buyers revert to their prior beliefs, or it is fully revealed. Hence, under noiseless demand, whether the firm uses discriminatory pricing has no particularly meaningful effect on the posterior beliefs.

The second issue we tackle in this paper relates to the profitability of market segmentation for the firm. Whether market integration is optimal is closely related to the question of whether uniform pricing for differentiated goods is optimal. In both problems, the benefits of the increased price flexibility need to be compared to the costs of charging different prices. Some recent papers (McMillan, 2007; Orbach and Einav, 2007; Chen, 2009; Chen and Cui, 2013; Richardson and Stähler, 2013) study such question in the context of differentiated goods. The present paper contributes partially to this strand of the literature by identifying a cost to charging different prices in a signaling context. Hence, we provide a glimpse to what incomplete information can yield when goods are differentiated.

At last, our work is related to several papers in international economics investigating the non-optimality of charging different prices for different markets (Friberg, 2001; Asplund and Friberg, 2000). However, these papers do not study the optimality of market segmentation (or integration) in a noisy signaling environment. Friberg (2001) studies whether a firm selling in regions with different currencies should segment the markets with an emphasis on the impact of the exchange rate, whereas Asplund and Friberg (2000) focus on the transportation cost from one region to another.[5] We do not explore these issues here, but rather provide an information-based reason for the profitability of market integration.

3.3 Information Dissemination

In this section, we study the effect of discriminatory pricing on information dissemination, i.e., how much buyers learn about quality from observing price(s). To that end, we present a model in which a firm sells a good to two segmented markets. We first solve for the equilibrium under discriminatory pricing as well as the benchmark equilibrium of uniform pricing. We then study the effect of discriminatory pricing on the dissemination of information.[6]

3.3.1 Set Up

Consider a firm selling a good of quality $\mu > 0$ in two markets: market A and market B. The firm chooses at which prices she sells the good in each market. That is she chooses P_A and P_B. We assume arbitrage can be prevented such that it is possible to segment the two markets using third-degree price discrimination. In this section, we assume that the decision to segment the markets or not is exogenous. In other words, we abstract from the question of whether market segmentation is profitable.

[5]Other papers such as Friberg (2003), Friberg and Martensen (2001) and Gallo (2010) study the profitability of market segmentation in the context of a duopoly.

[6]In this section, market segmentation is set exogenously. The profitability of market segmentation on the firm is discussed in Section 3.4 when market segmentation is endogenized, i.e., the firm decides whether to segment or to integrate the markets.

In market A, all buyers are informed, i.e., they know μ. Aggregate demand in market A is given by

$$Q_A(P_A, \mu, \eta_A) = \mu - P_A + \eta_A \tag{3.1}$$

where η_A is a demand shock that is unobserved by the buyers. The difference in demand between markets A and B is two-fold. A first difference concerns information. Unlike market A, market B is composed of both informed and uninformed buyers. Specifically, a fraction $\lambda \in [0,1]$ of the buyers knows μ and thus a fraction $1-\lambda$ does not know μ. Although the uninformed buyers have prior beliefs about μ, they also extract partial information about quality from observing prices, i.e., prices are noisy signals. That is, upon observing prices, the uninformed buyers' posterior mean for quality is $\int x\hat{\xi}(x|P_A, P_B)\mathrm{d}x$ where $\hat{\xi}(\cdot|P_A, P_B)$ is the posterior p.d.f. of $\tilde{\mu}$ given P_A and P_B.[7] A second difference is that conditional on μ, the buyers in market B have a reservation price $\gamma\mu$ where $\gamma > 0$ reflects the disparity in demand between the two markets (unless $\gamma = 1$).[8] Aggregate demand in market B is thus given by

$$Q_B(P_B, \mu, \hat{\xi}(\cdot|P_A, P_B), \eta_B) = \lambda(\gamma\mu - P_B) + (1-\lambda)\left(\gamma \int x\hat{\xi}(x|P_A, P_B)\mathrm{d}x - P_B\right) + \eta_B \tag{3.2}$$

where η_B is a demand shock that is unobserved by the buyers and $\int x\hat{\xi}(x|P_A, P_B)\mathrm{d}x$ is the posterior mean of μ. The updating of beliefs reflects the learning activity of the uninformed buyers.

Next, we describe the firm's maximization problem. For simplicity, the firm's marginal cost is assumed to be zero. In addition of knowing the quality μ, the firm has complete information about demand, i.e., both η_A and η_B are known to the firm. This informational asymmetry between the buyers and the firm conveys the idea that the firm knows more about demand than the buyers do.

We consider two cases. First, under discriminatory pricing, using (3.1) and (3.2), the firm's maximization problem is

$$\max_{P_A, P_B} \left\{ P_A \cdot Q_A(P_A, \mu, \eta_A) + P_B \cdot Q_B(P_B, \mu, \hat{\xi}(\cdot|P_A, P_B), \eta_B) \right\}. \tag{3.3}$$

[7]Note that $\hat{\xi}(\cdot|P_A, P_B)$ is the general expression for posterior beliefs upon observing two signals. If there is no market segmentation, then the uninformed buyers receive two *identical* signals, i.e., $P \equiv P_A = P_B$. In that case, posterior beliefs can be simplified to $\hat{\xi}(\cdot|P)$.

[8]The assumption that market A has only informed buyers is without loss of generality. All results continue to hold if we assume a fraction λ_A of buyers is uninformed in market A and a fraction λ_B in market B.

Second, under uniform pricing, the firm sets one price, i.e., $P \equiv P_A = P_B$, and the uninformed buyers receive only one signal. Using (3.1) and (3.2), the firm's maximization problem is

$$\max_{P} \left\{ P \cdot \Big(Q_A(P, \mu, \eta_A) + Q_B(P, \mu, \hat{\xi}(\cdot|P), \eta_B) \Big) \right\}. \tag{3.4}$$

Before proceeding with the definition and characterization of the equilibrium, we discuss the distributional assumption for prior beliefs and the random demand shocks.

Assumption 3.1. *Prior beliefs are $\tilde{\mu} \sim N(\rho, \sigma_\mu^2)$, with $\rho > 0$. Distributions of demand shocks are $\tilde{\eta}_A \sim N(0, \sigma_\eta^2), \tilde{\eta}_B \sim N(0, \sigma_\eta^2)$ such that $\mathbb{E}[\tilde{\eta}_A \tilde{\eta}_B] = 0$.*

The demand shocks are known to the firm, but unobserved by the buyers, which implies that the prices cannot fully reveal quality since they also depend on unobserved demand shocks. We rely on the fact that the family of normal distributions with an unknown mean is a conjugate family for samples from a normal distribution.[9] With the normality assumption, we obtain a unique linear equilibrium, i.e., an equilibrium in which the uninformed buyers' updating rule is linear in the price-signals. Although negative demand shocks can yield a negative price or a negative posterior mean, the values of the parameters of the model can be restricted to ensure that the probability of such events be arbitrarily close to zero. Moreover, it turns out that, for any parameters, equilibrium values for mean prices are always positive.

3.3.2 Equilibrium

First, consider the situation in which the firm uses discriminatory pricing ($\mathcal{D}$). The equilibrium consists of the firm's price strategies, $P^*_{\mathcal{D},A}(\mu, \eta_A, \eta_B)$ and $P^*_{\mathcal{D},B}(\mu, \eta_B, \eta_A)$; the distribution of the price-signals conditional on any quality x, $\phi^*_{\mathcal{D}}(P_A, P_B|x)$; and the uninformed buyers' posterior beliefs about the quality upon observing any prices $\{P_A, P_B\}$, $\hat{\xi}^*_{\mathcal{D}}(x|P_A, P_B)$.[10] In equilibrium, the posterior beliefs are consistent with Bayes' rule and the equilibrium distribution of prices.

[9] Normal assumption combined with linear demand yields closed-form equilibrium values and makes the analysis tractable by focusing on the mean and variance of price and posterior beliefs. See Grossman and Stiglitz (1980), **?**, Judd and Riordan (1994), Mirman et al. (2014a,b, 2015) for the use of normal distributions to study the informational role of prices in problems without market segmentation.

[10] The variable μ refers to the true quality whereas x is used as a dummy variable for quality.

Definition 3.1. *For any $\mu > 0$, the tuple $\Big\{P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B), P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A),$ $\phi^*_{\mathcal{D}}(P_A,P_B|\cdot)$, $\hat{\xi}^*_{\mathcal{D}}(\cdot|P_A,P_B)\Big\}$ is a noisy signaling equilibrium with discriminatory pricing if,*

1. *Given $\hat{\xi}^*_{\mathcal{D}}(\cdot|P_A,P_B)$, and for any η_A and η_B, the firm's price strategies are*

$$\Big\{P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B), P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A)\Big\} = \arg\max_{P_A,P_B} \Big\{P_A \cdot Q_A(P_A,\mu,\eta_A) + P_B \cdot Q_B(P_B,\mu,\hat{\xi}^*_{\mathcal{D}}(\cdot|P_A,P_B),\eta_B)\Big\}. \tag{3.5}$$

2. *Given the distribution of $\{\tilde{\eta}_A,\tilde{\eta}_B\}$, $\phi^*_{\mathcal{D}}(P_A,P_B|x)$ is the p.d.f. of the random price-signals $\Big\{P^*_{\mathcal{D},A}(x,\tilde{\eta}_A,\tilde{\eta}_B), P^*_{\mathcal{D},B}(x,\tilde{\eta}_B,\tilde{\eta}_A)\Big\}$ conditional on any quality x.*

3. *Given $\phi^*_{\mathcal{D}}(P_A,P_B|\cdot)$ and prior beliefs $\xi(\cdot)$, the uninformed buyers' posterior beliefs about quality upon observing P_A and P_B is $\tilde{\mu}^*_{\mathcal{D}}|P_A,P_B$ with the p.d.f.*

$$\hat{\xi}^*_{\mathcal{D}}(x|P_A,P_B) = \frac{\xi(x)\phi^*_{\mathcal{D}}(P_A,P_B|x)}{\int_{x'\in\mathbb{R}} \xi(x')\phi^*_{\mathcal{D}}(P_A,P_B|x')\mathrm{d}x'}, \quad \forall\ x \in \mathbb{R}. \tag{3.6}$$

Using Definition 3.1, Proposition 3.1 characterizes the noisy signaling equilibrium when the firm engages in third-degree price discrimination. Specifically, the price strategies and the posterior beliefs (as a function of the price-signals) are provided. The joint distribution of the price-signals is immediate from Assumption 3.1, (3.7), and (3.8).

Proposition 3.1. *Suppose that markets are segmented. For $\lambda \in [0,1)$, there exists a noisy signaling equilibrium with discriminatory pricing.*[11] *In equilibrium,*

1. *Given quality μ and demand shocks $\{\eta_A,\eta_B\}$, the firm sets prices*

$$P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B) = \frac{\delta^*_0\delta^*_1\gamma^2(1-\lambda)^2 + (2-2\delta^*_2\gamma(1-\lambda)+\delta^*_1\gamma^2\lambda(1-\lambda))\mu}{4-\delta^{*2}_1\gamma^2(1-\lambda)^2-4\delta^*_2\gamma(1-\lambda)} + \frac{(2-2\delta^*_2\gamma(1-\lambda))\eta_A + \delta^*_1\gamma(1-\lambda)\eta_B}{4-\delta^{*2}_1\gamma^2(1-\lambda)^2-4\delta^*_2\gamma(1-\lambda)} \tag{3.7}$$

and

$$P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A) = \frac{2\delta^*_0\gamma(1-\lambda) + (\delta^*_1\gamma(1-\lambda)+2\gamma\lambda)\mu + \delta^*_1\gamma(1-\lambda)\eta_A + 2\eta_B}{4-\delta^{*2}_1\gamma^2(1-\lambda)^2-4\delta^*_2\gamma(1-\lambda)}. \tag{3.8}$$

[11] When $\lambda = 1$, all buyers are informed and no updating rule needs to be specify. In this case, there exists an equilibrium with equilibrium prices $P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B) = (\mu+\eta_A)/2$ and $P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A) = (\gamma\mu+\eta_B)/2$, which are equal to (3.7) and (3.8) evaluated at $\lambda = 1$, respectively.

2. *Given any observation* $\{P_A, P_B\}$*, the uninformed buyers' posterior beliefs are*

$$\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B \sim N\left(\delta^*_0 + \delta^*_1 P_A + \delta^*_2 P_B, \frac{\sigma^2_\eta \sigma^2_\mu}{\sigma^2_\eta + (1+\gamma^2\lambda^2)\sigma^2_\mu}\right). \tag{3.9}$$

Here,

$$\delta^*_0 = \frac{\rho\sigma^2_\eta}{\sigma^2_\eta + \sigma^2_\mu(1+\gamma^2\lambda)}, \tag{3.10}$$

$$\delta^*_1 = \frac{2\sigma^2_\mu}{\sigma^2_\eta + \sigma^2_\mu(1+\gamma^2\lambda)}, \tag{3.11}$$

$$\delta^*_2 = \frac{2\gamma(\lambda\sigma^2_\mu(\sigma^2_\eta + 2\sigma^2_\mu) - \sigma^4_\mu(1-\gamma^2\lambda^2))}{(\sigma^2_\eta + \sigma^2_\mu(1+\gamma^2\lambda))(\sigma^2_\eta + \sigma^2_\mu(1+\gamma^2\lambda(2-\lambda)))}. \tag{3.12}$$

Proof. See Appendix B.1. □

Next, we define and characterize the noisy signaling equilibrium in the benchmark model in which the firm uses uniform pricing ($\mathcal{U}$).

Definition 3.2. *For any* $\mu > 0$*, the tuple* $\left\{P^*_{\mathcal{U}}(\mu, \eta_A, \eta_B), \phi^*_{\mathcal{U}}(P|\cdot), \hat{\xi}^*_{\mathcal{U}}(\cdot|P)\right\}$ *is a noisy signaling equilibrium with uniform pricing if,*

1. *Given* $\hat{\xi}^*_{\mathcal{U}}(\cdot|P)$*, and for any* η_A *and* η_B*, the firm's price strategy is*

$$P^*_{\mathcal{U}}(\mu, \eta_A, \eta_B) = \arg\max_P \left\{P \cdot \left(Q_A(P, \mu, \eta_A) + Q_B(P, \mu, \hat{\xi}^*_{\mathcal{U}}(\cdot|P), \eta_B)\right)\right\}. \tag{3.13}$$

2. *Given the distribution of* $\{\tilde{\eta}_A, \tilde{\eta}_B\}$*,* $\phi^*_{\mathcal{U}}(P|x)$ *is the p.d.f. of the random price-signal* $P^*_{\mathcal{U}}(x, \tilde{\eta}_A, \tilde{\eta}_B)$ *conditional on any quality* x*.*

3. *Given* $\phi^*_{\mathcal{U}}(P|\cdot)$ *and prior beliefs* $\xi(\cdot)$*, the uninformed buyers' posterior beliefs upon observing any* P *is* $\tilde{\mu}^*|P$ *with p.d.f.*

$$\hat{\xi}^*_{\mathcal{U}}(x|P) = \frac{\xi(x)\phi^*_{\mathcal{U}}(P|x)}{\int_{x'\in\mathbb{R}} \xi(x')\phi^*_{\mathcal{U}}(P|x')\mathrm{d}x'}, \quad \forall\, x \in \mathbb{R}. \tag{3.14}$$

Proposition 3.2 characterizes the noisy signaling equilibrium when the firm does not price discriminate. The distribution of the price-signal is immediate from Assumption 3.1 and from Equation (3.15).

Proposition 3.2. *Suppose that markets are not segmented. For $\lambda \in [0,1)$, there exists a noisy signaling equilibrium with uniform pricing.*[12] *In equilibrium,*

1. *Given quality μ and demand shocks $\{\eta_A, \eta_B\}$, the firm sets the price*

$$P^*_{\mathcal{U}}(\mu, \eta_A, \eta_B) = \frac{\beta^*_0 \gamma(1-\lambda) + (1+\gamma\lambda)\mu + \eta_A + \eta_B}{4 - 2\beta^*_1 \gamma(1-\lambda)} \tag{3.15}$$

in markets A and B.

2. *Given any observation P, the uninformed buyers' posterior beliefs are*

$$\tilde{\mu}^*_{\mathcal{U}} | P \sim N\left(\beta^*_0 + \beta^*_1 P, \frac{2\sigma^2_\eta \sigma^2_\mu}{2\sigma^2_\eta + (1+\lambda\gamma)^2\sigma^2_\mu}\right). \tag{3.16}$$

Here,

$$\beta^*_0 = \frac{2\rho\sigma^2_\eta}{2\sigma^2_\eta + \sigma^2_\mu(1+\gamma+\gamma\lambda+\gamma^2\lambda)}, \tag{3.17}$$

$$\beta^*_1 = \frac{4(1+\gamma\lambda)\sigma^2_\mu}{2\sigma^2_\eta + \sigma^2_\mu(1+2\gamma+2\gamma^2\lambda-\gamma^2\lambda^2)}. \tag{3.18}$$

Proof. See Appendix B.1. □

From (3.7), (3.8), and (3.15), equilibrium prices are linear functions of the quality μ as well as demand shocks η_A and η_B. Although prices are informative about quality, the presence of unknown demand shocks prevents the buyers from learning the exact value of quality, i.e., price is partially revealing of quality. Hence, noise in demand allows us to study the impact of pricing strategies (discriminatory vs. uniform) on the dissemination of information.[13]

As a first step to study the impact of pricing strategies, one can notice that the amount of information conveyed by the price(s) depends on the pricing strategy adopted by the firm.

---

[12] When $\lambda = 1$, all buyers are informed and no updating rule needs to be specify. In this case, there exists an equilibrium with $P^*_{\mathcal{U}}(\mu, \eta_A, \eta_B) = ((1+\gamma)\mu + \eta_A + \eta_B)/4$ which is equal to (3.15) evaluated at $\lambda = 1$.

[13] In our model, if demand shocks are known to buyers, then quality is perfectly inferred from observing the price(s) under both uniform and discriminatory pricing.

Remark 3.1. *The pricing strategy chosen by the firm alters the amount of information conveyed by the price-signal(s). More specifically, for* $s \in \{A, B\}$

1. $\mathbb{V}ar[P^*_{\mathcal{U}}(\tilde{\mu}, \eta_A, \eta_B)|\, \eta_A, \eta_B] \geq \mathbb{V}ar[P^*_{\mathcal{D},s}(\tilde{\mu}, \eta_A, \eta_B)|\, \eta_A, \eta_B]$,
2. $\mathbb{V}ar[P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|\, \mu] < \mathbb{V}ar[P^*_{\mathcal{D},s}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|\, \mu]$.

Remark 3.1 is important since it means that the pieces of information from which the buyers learn can have a lower quality when the firm uses discriminatory pricing. Indeed, on the one hand, observation 1. establishes that, conditional on demand shocks η_A and η_B, the price-signal under uniform pricing is more sensitive to quality μ than the price-signal in segment s under discriminatory pricing. It is thus easier to distinguish between different qualities under uniform pricing. On the other hand, observation 2. establishes that, conditional on quality μ, the price-signal under uniform pricing is less sensitive to the demand shocks (η_A, η_B) than the price-signal in segment s under discriminatory pricing. In other words, there is less noise in the price-signal under uniform pricing. Taken together, these observations imply that the price-signal with uniform pricing is a better signal than the price-signal from market segment s under discriminatory pricing.

Yet, one cannot already conclude that uniform pricing is better for the dissemination of information to uninformed buyers. The reason is that when the firm segments the market, uninformed buyers obtain two price-signals instead of one. Hence, there is a trade-off: the uniform price-signal is better, but there are two signals when markets are segmented. In the next section, we show that, in spite of this trade-off, discriminatory pricing induces more learning.

3.3.3 Comparison of Pricing Strategies

We now compare discriminatory and uniform pricing for the dissemination of information. We consider two aspects to measure the amount of information conveyed by prices: the bias of the posterior mean and the posterior variance.

3.3.3.1 Bias of the Posterior Mean

The bias of the posterior mean measures the distance between the true quality μ and the uninformed buyers' mean posterior evaluation of this quality. For instance, consider that discriminatory pricing is used by the firm and that μ is the true quality. A posteriori, upon seeing P_A and P_B, uninformed buyers believe that quality has a mean $\mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B]$. Then, on average, the posterior mean is $\int_{P_A,P_B} \mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B]\phi^*_{\mathcal{D}}(P_A, P_B|\mu)\mathrm{d}P_A\mathrm{d}P_B$ and the expected bias denoted by $\mathcal{B}^*_{\mathcal{D}}$ is the average distance between the true quality μ and the uninformed buyers' mean evaluation of this quality.

Using Proposition 3.1, the expected bias under discriminatory pricing is the absolute value of the difference between the unconditional posterior mean for quality and the true quality μ, i.e.,

$$\mathcal{B}^*_{\mathcal{D}} \equiv \left| \int_{P_A,P_B} \mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B]\phi^*_{\mathcal{D}}(P_A, P_B|\mu)\mathrm{d}P_A\mathrm{d}P_B - \mu \right|. \tag{3.19}$$

Consider next that uniform pricing is used by the firm. Then, the expected bias under uniform pricing follows similarly, from Proposition 3.2, and is defined as

$$\mathcal{B}^*_{\mathcal{U}} \equiv \left| \int_{P} \mathbb{E}[\tilde{\mu}^*_{\mathcal{U}}|P]\phi^*_{\mathcal{U}}(P|\mu)\mathrm{d}P - \mu \right|. \tag{3.20}$$

Proposition 3.3 establishes that the expected bias of the posterior mean is always smaller under discriminatory pricing than under uniform pricing. In other words, with third-degree price discrimination, the buyers are, on average, closer to the truth.[14]

Proposition 3.3. *From* (3.19) *and* (3.20)*,* $|\mathcal{B}^*_{\mathcal{D}}| \leq |\mathcal{B}^*_{\mathcal{U}}|$.

Proof. From (3.7), (3.8), and (3.9),

$$\int_{P_A,P_B} \mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B]\phi^*_{\mathcal{D}}(P_A, P_B|\mu)\mathrm{d}P_A\mathrm{d}P_B = \frac{\rho\sigma^2_\eta + \mu(1+\gamma^2\lambda^2)\sigma^2_\mu}{\sigma^2_\eta + (1+\gamma^2\lambda^2)\sigma^2_\mu}. \tag{3.21}$$

From (3.15) and (3.16),

$$\int_{P} \mathbb{E}[\tilde{\mu}^*_{\mathcal{U}}|P]\phi^*_{\mathcal{U}}(P|\mu)\mathrm{d}P = \frac{2\rho\sigma^2_\eta + \mu(1+\gamma\lambda)^2\sigma^2_\mu}{2\sigma^2_\eta + (1+\gamma\lambda)^2\sigma^2_\mu}. \tag{3.22}$$

[14] In the particular case of $\rho = \mu$, the posterior mean of quality is on average unbiased regardless of the pricing regime, i.e., $|\mathcal{B}^*_{\mathcal{D}}| = |\mathcal{B}^*_{\mathcal{U}}| = 0$.

Plugging (3.21) into (3.19) and (3.22) into (3.20) yields the results. □

From Proposition 3.3, we can conclude that observing two signals instead of one reduces the expected bias of the posterior beliefs. However, for specific realization of demand shocks $\{\eta_A, \eta_B\}$, posterior beliefs might be closer to the true quality with uniform pricing than with discriminatory pricing. Thus, in order to compare the quality of information dissemination under the two pricing regimes, we need to discuss the effect of price discrimination on the variance of the posterior beliefs.

3.3.3.2 Posterior Variance

Upon observing the price-signal(s), uninformed buyers believe that quality is normally distributed with a mean and a variance as given in (3.9) or in (3.16). We now study the effect of the firm's pricing strategy on this posterior variance.[15]

The size of the posterior variance is related to the uninformed buyers' learning speed. More specifically, the smaller the posterior variance is, the faster is the learning speed and the more information is conveyed to uninformed buyers. The reason is that the posterior variance tells us how likely it is that values far from the posterior mean be the actual true quality. For instance, if one were to construct a 95% confidence interval for the value of the true quality, the size of this interval would be determined by the posterior variance.

Proposition 3.4 states that the posterior variance for quality is always greatest under uniform pricing. Hence, price discrimination provides more information to the uninformed buyers, i.e., the posterior beliefs for quality are less variable. Equation (3.23) characterizes the variance differential. Note that the presence of both demand uncertainty and prior uncertainty are necessary for market segmentation to have an effect on the informational content of prices. If there is no prior uncertainty (i.e., $\sigma_\mu^2 \to 0$), then there is no reason to learn from prices. Moreover, if there is no unknown demand shock (i.e., $\sigma_\eta^2 \to 0$), then observing more prices does not provide more information to the uninformed buyers since the uninformed buyers can infer exactly the true quality regardless of the pricing strategy.

[15] Note that the posterior variance of the posterior beliefs, the object studied in this section, is different from the posterior beliefs' variance where the posterior beliefs is taken as a random variable.

Proposition 3.4. *For* $\lambda \in [0, 1)$, *from* (3.9) *and* (3.16),

$$\mathbb{V}[\tilde{\mu}^*_{\mathcal{U}}] - \mathbb{V}[\tilde{\mu}^*_{\mathcal{D}}] = \frac{(1-\gamma\lambda)^2\sigma_\eta^2\sigma_\mu^4}{(\sigma_\eta^2 + (1+\gamma^2\lambda^2)\sigma_\mu^2)(2\sigma_\eta^2 + (1+\gamma\lambda)^2\sigma_\mu^2)} \geq 0. \tag{3.23}$$

Before discussing Proposition 3.4, it is worth considering three special cases of (3.23). Consider first the benchmark case of full information with two identical markets, i.e., $\gamma = \lambda = 1$. Hence, for an uninformed *outsider*, market segmentation yields no gain in precision of the posterior beliefs. Next, consider two special cases for which discriminatory prices provide more precise posterior beliefs. If $\lambda = 1$ and $\gamma \in [0, 1)$, then discriminatory prices provide better information to an uninformed *outsider*. The fact that two signals about two fully informed markets be available makes the posterior beliefs more precise. In other words, the market price is more informative to outsiders.[16] Finally, if $\gamma = 1$ and $\lambda \in (0, 1)$, then preferences over the good are the same between the two markets, but some buyers in market B are uninformed. In the presence of uninformed buyers, segmenting a market into two identical markets provides more precise information to the uninformed buyers.[17]

We now discuss Proposition 3.4. This discussion builds in part on Remark 3.1 outlining the fact that the variance of a price-signal changes with the pricing regime. Yet, in spite of changes in the variances of the price-signals, the gain in precision due to discriminatory prices always holds. On the one hand, discriminatory pricing (compared to uniform pricing) implies that the buyers receive two signals instead of one. Hence, holding everything else constant, price discrimination provides more signals and increases the precision of the posterior beliefs. On the other hand, since the firm sets prices, the variance of the price-signals are endogenous. In particular, it is possible for the variance of the price-signals to increase as a consequence of market segmentation. Hence, there is a trade-off. Price discrimination offers more signals but each of these signals might be less precise. Although there is a trade-off, it turns out that third-degree price discrimination always reduces the variance of the posterior mean for quality even when each price-signal becomes less precise.

[16] Given $\lambda = 1$, from (3.9) and (3.16) we have $\mathbb{V}[\tilde{\mu}^*] = \sigma_\eta^2\sigma_\mu^2/(\sigma_\eta^2 + (1+\gamma^2)\sigma_\mu^2)$ and $\mathbb{V}[\tilde{\mu}^*] = 2\sigma_\eta^2\sigma_\mu^2/(2\sigma_\eta^2 + (1+\gamma)^2\sigma_\mu^2)$ respectively.

[17] Given $\gamma = 1$ and $\lambda \in (0, 1)$, from (3.9) and (3.16), we have $\mathbb{V}[\tilde{\mu}^*] = \sigma_\eta^2\sigma_\mu^2/(\sigma_\eta^2 + (1+\lambda^2)\sigma_\mu^2)$ and $\mathbb{V}[\tilde{\mu}^*] = 2\sigma_\eta^2\sigma_\mu^2/(2\sigma_\eta^2 + (1+\lambda)^2\sigma_\mu^2)$, respectively.

We finish this section with a comparative analysis on (3.23). Specifically, we show how the parameters of the model mitigate or reinforce the positive effect of discriminatory pricing on the variance of the posterior beliefs. Remark 3.2 presents the effect of noise on the variance differential.

Remark 3.2. *From* (3.23)*, for* $\lambda \in [0, 1)$*,*

$$\frac{\partial(\mathbb{V}[\tilde{\mu}^*_{\mathcal{U}}] - \mathbb{V}[\tilde{\mu}^*_{\mathcal{D}}])}{\partial \sigma_\mu} \geq 0. \tag{3.24}$$

An increase in the variance of the prior beliefs always increases the variance differential. Specifically, from (3.24), if the prior beliefs are very precise, then the differential in the posterior variance stemming from the observation of two price-signals instead of one is relatively small. Since the uninformed buyers are quite certain that the true quality lies within a constraint interval, the firm's signaling activity does not play a prominent role as the informational reaction to the price-signals is small, i.e., β_1^* and $\{\delta_1^*, \delta_2^*\}$ are small. On the other hand, if the prior beliefs are very diffuse, the firm's signaling activity matters a lot and the differential in information that two price-signals convey instead of one is significantly more important.

Next, consider the effect of the proportion of informed buyers and the differential in demand on (3.23). From (3.23), notice that only the product $\lambda\gamma$ matters for the variance differential. Since $\gamma > 0$ and $\lambda \in [0, 1)$, two cases can occur, that is, $\lambda\gamma < 1$ and $\lambda\gamma \geq 1$. Remark 3.3 concentrates on the case $\lambda\gamma < 1$ and states that the larger the proportion of informed buyers is and the lesser the differential in buyers' valuation is, then the smaller is the variance differential coming from market segmentation.

Remark 3.3. *From* (3.23)*, for* $\lambda \in [0, 1)$ *and* $\gamma \in (0, 1]$*,*

$$\frac{\partial \mathbb{V}[\tilde{\mu}^*_{\mathcal{U}}] - \mathbb{V}[\tilde{\mu}^*_{\mathcal{D}}]}{\partial \lambda} \leq 0 \tag{3.25}$$

and

$$\frac{\partial \mathbb{V}[\tilde{\mu}^*_{\mathcal{U}}] - \mathbb{V}[\tilde{\mu}^*_{\mathcal{D}}]}{\partial \gamma} \leq 0. \tag{3.26}$$

From (3.25), as λ increases, under both uniform pricing and third-price discrimination, the posterior beliefs become less volatile as the price-signals incorporate more information from the mere fact that a larger proportion of buyers knows the true quality μ. However, the posterior variance decreases more rapidly when the uninformed buyers observe a single price-signal rather than two price-signals. This means that the benefit on the flow of information arising from a signal of a better quality is subject to some form of decreasing return. From (3.9) and (3.16), $\beta_1^* > \delta_1^* + \delta_2^*$. Hence, the uninformed buyers are always more sensitive to an improvement in the quality of a price-signal (due to more informed buyers) under uniform pricing than under discriminatory pricing. This higher sensibility translates into a higher decay rate of the posterior variance.

Next, we investigate the effect of the parameter γ on the variance differential in (3.23). From (3.26), an increase in γ reduces the benefit from observing two price-signals rather than one. Recall that γ is an indicator of how elastic is market B relative to market A, i.e., as $\gamma \to 1$, the two market segments are more similar to each other. Hence an increase in γ, by homogenizing the two markets, implies that P_A and P_B incorporate and convey a more similar content to the uninformed buyers.[18] Therefore, as $\gamma \to 1$, the uninformed buyers gain less from observing a second price-signal as it contains little supplementary informative content.[19]

3.4 On the Profitability of Discriminatory Pricing

We now extend the model by allowing the firm to either segment or integrate the two markets.[20] Our model has now two stages. In a nutshell, at the first stage, the firm decides whether or not to split the market into two separate markets. Then, at the second stage, the firm uses uniform pricing if there is market integration and discriminatory price if there is market segmentation. In either case, the firm takes into account the fact that prices can provide partial information about quality to the uninformed buyers.

[18]Using the criterion of mutual information $MI(\tilde{P}_A, \tilde{P}_B) = -log(1-\rho^2)/2$ where ρ is the correlation coefficient between $\tilde{P}_A$ and $\tilde{P}_B$, then we have $\partial MI(\tilde{P}_A, \tilde{P}_B)/\partial\gamma > 0$ such that the mutual information of $\tilde{P}_A$ and $\tilde{P}_B$ increases with γ.

[19]Comparative static in the case $\lambda\gamma \geq 1$ is more complicated. Overall, results of Remark 3.3 continue to hold if $\lambda\gamma$ is sufficiently high.

[20]Since we are studying third-degree price discrimination, arbitrage is assumed to be impossible.

3.4.1 Preliminaries

The demand side is unchanged, that is, demands in market A and B are given by (3.1) and (3.2), respectively. Except for the quality parameter μ and the demand shocks η_A and η_B, all the other parameters of the model (including the uninformed buyers' prior beliefs and the distribution of the demand shocks) are common knowledge.

The timing is as follows. At the first stage, the firm does not know the quality, nor the demand shocks and decides whether or not to segment the markets by comparing the expected profit under discriminatory pricing with the expected profit under uniform pricing, rationally anticipating quality, the demand shocks as well as the learning activity of the uninformed buyers.[21] Formally, let $M \in \{\mathcal{U}, \mathcal{D}\}$ be the firm's decision in the first stage. If $M = \mathcal{U}$, then the markets are integrated and pricing is uniform. If $M = \mathcal{D}$, then the markets are segmented leading to discriminatory pricing. At the second stage, after observing quality and the demand shocks, the firm sets the price(s). The uninformed buyers observe the segmentation decision, but do not know the quality nor the demand shocks.[22] Upon observing the price(s), the buyers update beliefs (if uninformed), and purchase the good. Figure 3.1 summarizes the timeline.

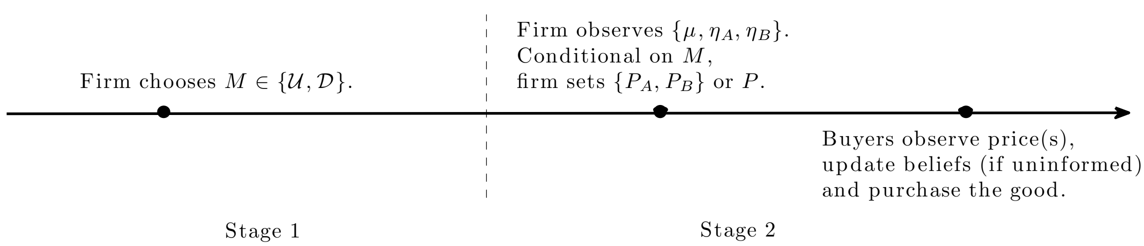


Figure 3.1: Timeline

[21]This reflects the idea that the firm faces some uncertainty in demand before making a decision about market segmentation. If the firm knows μ before making the segmentation decision, then its choice, if not independent of μ, signals information to the uninformed buyers. In order to abstract for the increased complexity of having both prices and segmentation decision acting as signals, we assume that the firm learns μ only after her segmentation decision is made.

[22]The fact that the buyers do not observe the demand shocks conveys the idea that the firm knows more about demand than the buyers do. Moreover, this informational asymmetry enables prices to provide partial (noisy) information about the quality of the good.

We now describe formally the behavior of the firm at each stage.[23] We begin with the second stage. If the markets are not segmented, then the firm sets one price. Using (3.1) and (3.2) evaluated at $P \equiv P_A = P_B$, stage-2 maximization problem (given $M = \mathcal{U}$) is the same as in (3.4). If the markets are segmented, then the firm sets a price in each market. Using (3.1) and (3.2), given that the firm has decided to segment the market at stage 1, stage-2 maximization problem (given $M = \mathcal{D}$) is the same as in (3.3). Note that from (3.3) and (3.4), the firm's expected profits are influenced by the uninformed buyers' posterior mean or equivalently by the p.d.f. $\hat{\xi}_{\mathcal{U}}(\cdot|P)$ and $\hat{\xi}_{\mathcal{D}}(\cdot|P_A, P_B)$.

At the first stage, the firm decides whether to use discriminatory or uniform pricing strategies. Formally, let the tuple $\{\{P_{\mathcal{U}}(\mu, \eta_A, \eta_B), \{P_{\mathcal{D},A}(\mu, \eta_A, \eta_B), P_{\mathcal{D},B}(\mu, \eta_B, \eta_A)\}\}, \{\hat{\xi}_{\mathcal{U}}(\cdot|P_A, P_B)$ $\hat{\xi}_{\mathcal{D}}(\cdot|P)\}\}$ define a profile of strategies at the second stage. Specifically, $P_{\mathcal{U}}(\mu, \eta_A, \eta_B)$ is the firm's price strategy under uniform pricing and $\{P_{\mathcal{D},A}(\mu, \eta_A, \eta_B), P_{\mathcal{D},B}(\mu, \eta_B, \eta_A)\}$ is the firm's price strategy under discriminatory. The terms $\hat{\xi}_{\mathcal{U}}(\cdot|P)$ and $\hat{\xi}_{\mathcal{D}}(\cdot|P_A, P_B)$ are the uninformed buyers' posterior beliefs under uniform pricing and discriminatory pricing, respectively. Given these strategies and posterior beliefs at stage-2, the expected profits of the firm under uniform pricing and discriminatory pricing are

$$\begin{aligned}\mathbb{E}[\Pi_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)] = \mathbb{E}\Big[P_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B) \cdot \Big(&Q_A(P_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B), \tilde{\mu}, \tilde{\eta}_A) \\ &+ Q_B(P_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B), \tilde{\mu}, \hat{\xi}_{\mathcal{U}}(\cdot|P_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)), \tilde{\eta}_B)\Big)\Big] \end{aligned} \tag{3.27}$$

and

$$\begin{aligned}\mathbb{E}[\Pi_{\mathcal{D}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)] = \mathbb{E}[P_{\mathcal{D},A}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B) \cdot Q_A(P_{\mathcal{D},A}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B), \tilde{\mu}, \tilde{\eta}_A)] + \mathbb{E}[P_{\mathcal{D},B}(\tilde{\mu}, \tilde{\eta}_B, \tilde{\eta}_A) \\ \cdot Q_B(P_{\mathcal{D},B}(\tilde{\mu}, \tilde{\eta}_B, \tilde{\eta}_A), \tilde{\mu}, \hat{\xi}_{\mathcal{D}}(\cdot|P_{\mathcal{D},A}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B), P_{\mathcal{D},B}(\tilde{\mu}, \tilde{\eta}_B, \tilde{\eta}_A)), \tilde{\eta}_B)],\end{aligned} \tag{3.28}$$

respectively. Here, $\mathbb{E}[\cdot]$ is the expectation operator over $\{\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B\}$ where a tilde sign is used to distinguish a random variable from its realization.

[23]A definition of the perfect Bayesian equilibrium is provided in Appendix B.2.

Let $\{M^*, \{P^*_{\mathcal{U}}(\mu, \eta_A, \eta_B), \{P^*_{\mathcal{D},A}(\mu, \eta_A, \eta_B), P^*_{\mathcal{D},B}(\mu, \eta_B, \eta_A)\}\}, \{\hat{\xi}^*_{\mathcal{U}}(\cdot|P_A, P_B), \hat{\xi}^*_{\mathcal{D}}(\cdot|P)\}\}$ be a perfect Bayesian equilibrium. Then, using (3.27) and (3.28), at the first stage the firm chooses not to split the two markets (i.e., $M^* = \mathcal{U}$) if and only if

$$\mathbb{E}[\Pi^*_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)] > \mathbb{E}[\Pi^*_{\mathcal{D}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)]. \tag{3.29}$$

3.4.2 Comparisons of Profits

We now provide conditions under which (3.29) holds. Specifically, we show that the presence of uninformed buyers (inducing the firm to engage in noisy signaling) makes it possible for the firm to obtain higher expected profits by not segmenting the market.

In order to do so, we need to obtain the equilibrium profits for each possible state in stage-2 (i.e., $M \in \{\mathcal{U}, \mathcal{D}\}, \forall\ (\mu, \eta_A, \eta_B)$) and then take the expectation with respect to $(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)$ to obtain stage-1 expected profits. Proposition 3.5 gives the expected profits for $M = \mathcal{U}$ and $M = \mathcal{D}$.

Proposition 3.5. *Given $M = \mathcal{U}$ and given $M = \mathcal{D}$, there exists an equilibrium in the second stage characterized by Proposition 3.1 and Proposition 3.2, respectively. Then, stage-1 expected profits are*

1. *For $M = \mathcal{U}$,*

$$\mathbb{E}[\Pi^*_{\mathcal{U}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)] = \frac{(1+\gamma)^2(\rho + \sigma^2_\mu) + 2\sigma^2_\eta}{8} - (1-\lambda)^2\Psi^*_{\mathcal{U}}, \tag{3.30}$$

where

$$\Psi^*_{\mathcal{U}} = \frac{\gamma^2\sigma^2_\mu}{8}\left(1 + \frac{(1+\gamma)^2(1+\gamma\lambda)^2\rho^2\sigma^2_\mu}{(2\sigma^2_\eta + (1+\gamma)(1+\gamma\lambda)\sigma^2_\mu)^2}\right) \tag{3.31}$$

2. *For $M = \mathcal{D}$,*

$$\mathbb{E}[\Pi^*_{\mathcal{D}}(\tilde{\mu}, \tilde{\eta}_A, \tilde{\eta}_B)] = \frac{(1+\gamma^2)(\rho^2 + \sigma^2_\mu) + 2\sigma^2_\eta}{4} - (1-\lambda)^2\Psi^*_{\mathcal{D}}, \tag{3.32}$$

where

$$\Psi^*_{\mathcal{D}} = \frac{\gamma^2\sigma^2_\mu(\sigma^4_\eta + 2(1+\gamma^2\lambda)\sigma^2_\eta\sigma^2_\mu + \gamma^2(1+\gamma^2\lambda^2)\sigma^2_\mu(\rho^2+\sigma^2_\mu))}{4\left(\sigma^4_\eta + 2(1+\gamma^2\lambda)\sigma^2_\eta\sigma^2_\mu + (1+\gamma^2)(1+\gamma^2\lambda^2)\sigma^4_\mu\right)}. \tag{3.33}$$

Proof. It follows from $\left\{P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B), P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A), \phi^*_{\mathcal{D}}(P_A,P_B|\cdot), \hat{\xi}^*_{\mathcal{D}}(\cdot|P_A,P_B)\right\}$ in Proposition 3.1 and from $\left\{P^*_{\mathcal{U}}(\mu,\eta_A,\eta_B), \phi^*_{\mathcal{U}}(P|\cdot), \hat{\xi}^*_{\mathcal{U}}(\cdot|P)\right\}$ in Proposition 3.2. □

Proposition 3.5 shows that regardless of the firm's decision to segment or integrate the markets, the first-stage expected profits are the sum of two components. The first component is the *full-information* expected profits, i.e., when all buyers are informed (i.e., $\lambda = 1$). The second component is a distortion that emanates from the firm's need to signal quality via prices. Indeed, in order to signal the quality of the good to the uninformed buyers, the firm alters prices. This distortion in prices translates into a loss in expected profits. Formally, from (3.31) and (3.33), $-\Psi^*_{\mathcal{U}} \leq 0$ and $-\Psi^*_{D} \leq 0$ denote the loss in expected profits (due to signaling) under no market segmentation and market segmentation, respectively.

Using Proposition 3.5, we show that when there is noisy signaling, it is possible that the firm prefers *not* to segment the market. Taking the difference in profits between (3.32) and (3.30) yields

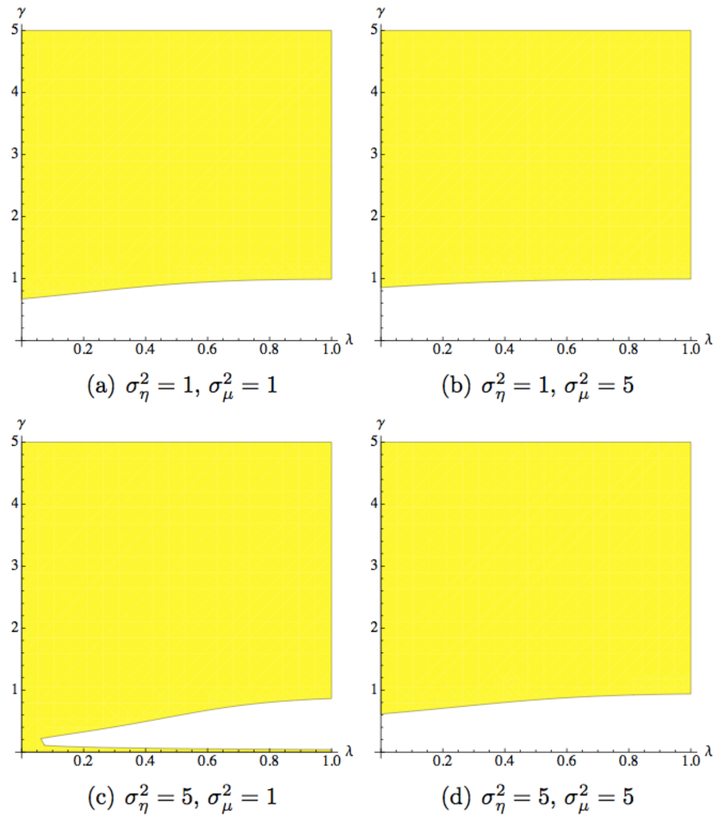
$$\frac{(1-\gamma)^2(\rho^2+\sigma^2_\mu)}{8} + \frac{\sigma^2_\eta}{4} - (1-\lambda)^2(\Psi^*_D - \Psi^*_{\mathcal{U}}). \tag{3.34}$$

When (3.34) is positive (negative), the firm prefers to (not to) segment the markets. Consider first the benchmark case of full information when all buyers are informed, i.e., $\lambda = 1$. If every buyer is informed, then it is always profitable for a firm to segment the market. Indeed, in that case, there is no loss in expected profits due to signaling. Using two prices instead of one always yields higher expected profits for two reasons: a) because using two prices allows the firm to set prices that are more appropriate for each market segments, this is captured by the first term in (3.34), and b) because, by setting two prices instead of one, the firm accounts for the specific demand shocks in each market segment and not only for the average shock, this is captured by the second term in (3.34).

Remark 3.4. *If $\lambda = 1$, then*

$$\mathbb{E}[\Pi^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)] < \mathbb{E}[\Pi^*_{\mathcal{D}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)].$$

Remark 3.4 implies that a necessary condition for the firm to prefer not to segment the market is the presence of uninformed buyers, which is related to the loss (due to signaling) in



(a) $\sigma_\eta^2 = 1$, $\sigma_\mu^2 = 1$

(b) $\sigma_\eta^2 = 1$, $\sigma_\mu^2 = 5$

(c) $\sigma_\eta^2 = 5$, $\sigma_\mu^2 = 1$

(d) $\sigma_\eta^2 = 5$, $\sigma_\mu^2 = 5$

Figure 3.2: Shaded area indicates $\Psi^*_{\mathcal{U}} \leq \Psi^*_{\mathcal{D}}$ with $\rho = 10$.

expected profits. Indeed, in order to offset the benefit from price flexibility (by segmenting the market), it is necessary (but not sufficient) for the loss in expected profits under market segmentation to be greater than the loss in expected profits under no market segmentation. For $\lambda \in [0, 1)$, it is possible that $\Psi^*_{\mathcal{D}} > \Psi^*_{\mathcal{U}}$. Figure 3.2 depicts the region of the parameters space $\{\lambda, \gamma\}$ corresponding to $\Psi^*_{\mathcal{D}} > \Psi^*_{\mathcal{U}}$.

Proposition 3.6 establishes the condition under which the firm chooses not to segment the market. Condition (3.35) compares the gains and losses in expected profits from integrating the markets. Intuitively, the firm faces a trade-off. On the one hand, market segmentation yields more flexibility and the ability to capture more of the consumer surplus. On the other hand, the firm also has to incur a signaling cost, i.e., the distortion needed to signal quality via prices depends on whether the market is integrated or segmented. Specifically, the firm does not segment the market if there is a reduction in cost due to signaling which is greater than the loss from price flexibility. While there is always a loss from giving up price flexibility, whether there is a reduction in cost due to signaling depends on the parameter values.

Proposition 3.6. *If $\lambda \in [0,1)$, then, at the first stage, it is optimal for the firm not to segment the market (i.e., $M^* = \mathcal{U}$) if and only if*

$$\Psi_{\mathcal{D}}^* - \Psi_{\mathcal{U}}^* \geq \frac{(1-\gamma)^2(\rho^2+\sigma_\mu^2)+2\sigma_\eta^2}{8(1-\lambda)^2} \tag{3.35}$$

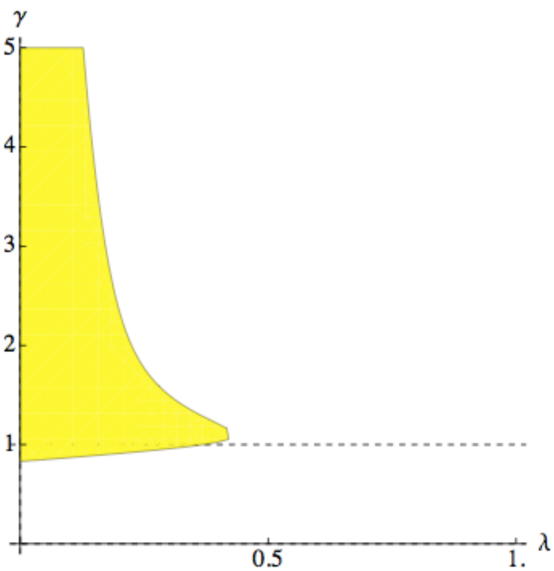
where $\Psi_{\mathcal{U}}^*$ and $\Psi_{\mathcal{D}}^*$ are given by (3.31) *and* (3.33), *respectively.*

It is convenient to depict the condition stated in Proposition 3.6. Figure 3.3 illustrates Proposition 3.6 by showing regions of the parameters space $\{\lambda, \gamma\}$ corresponding to $M^* = \mathcal{U}$ for some σ_η^2 and σ_μ^2. The firm chooses not to segment the markets when the fraction of informed buyers is low enough and the reservation price on market B is either almost similar to the one of market A, or higher. In terms of the parameters, this implies that λ is low and γ is not too low. This is consistent with the decomposition of expected profits provided in Proposition 3.5. Indeed, as noted, the firm faces a trade-off between a benefit from price flexibility and a cost from having to signal quality from prices.

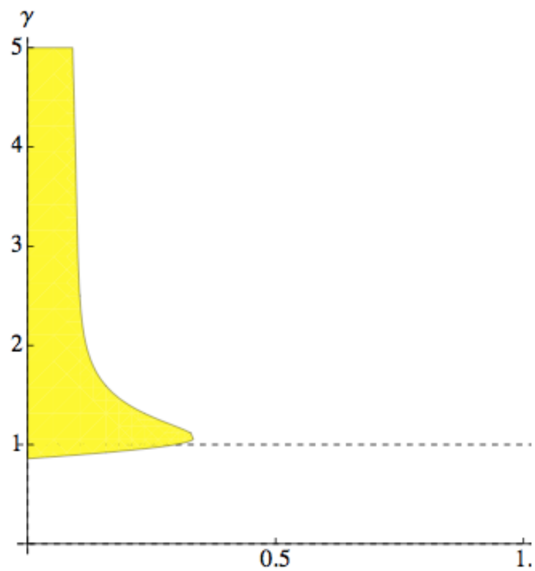
The farther γ is from 1, the greater the gain from price flexibility and splitting the market since markets A and B are very different. That is, the first component in (3.34) is increasing in $(1-\gamma)^2$. The firm will prefer not to segment the market and to avoid the double signaling cost when the price-signals are worthy for a large mass of buyers (for low values of λ).

When the two markets are identical, i.e., $\gamma = 1$, market integration can turn out to be the optimal pricing regime for the firm.[24] In fact, this is the situation in which market integration is the most susceptible to be profitable since the benefit from price flexibility is, in this case, at its lowest.

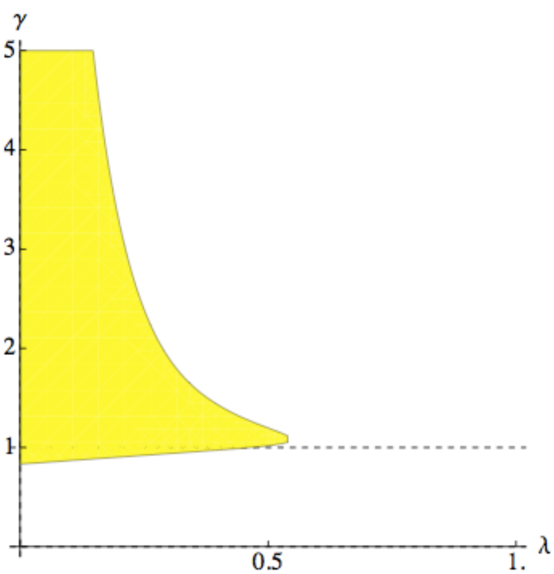
[24] There is still a difference between the two markets because there are some uninformed buyers in market B.



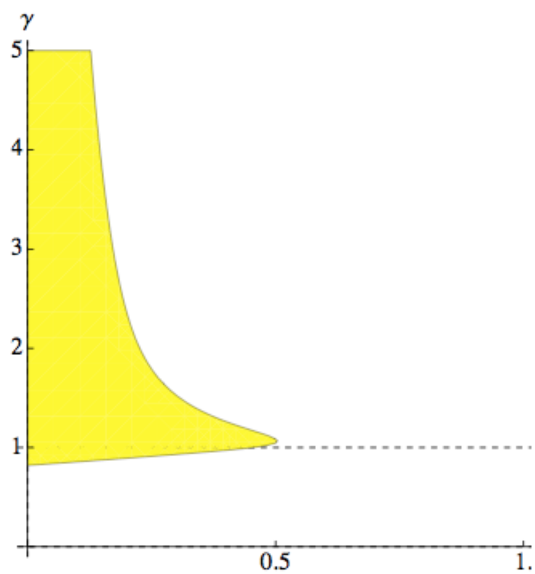
(a) $\sigma_\eta^2 = 1$, $\sigma_\mu^2 = 1$



(b) $\sigma_\eta^2 = 1$, $\sigma_\mu^2 = 0.5$



(c) $\sigma_\eta^2 = 0.5$, $\sigma_\mu^2 = 1$



(d) $\sigma_\eta^2 = 0.5$, $\sigma_\mu^2 = 0.5$

Figure 3.3: The shaded area shows $M^* = \mathcal{U}$ for $\rho = 10$.

3.5 Conclusion

In this paper, we have studied the common commercial practice of third-degree price discrimination in the presence of consumer learning using prices as informative signals of quality. Understanding the implications of price discrimination is particularly important since the practice has gained in popularity with the shift from brick-and-mortar stores to the online marketplace as it is easier for firms to accumulate information on consumers and thus, easier to charge different prices to different consumer segments.

In this particular paper, we study the effect of market segmentation on the informational content of prices and highlight that it can be beneficial for consumers and detrimental for the firm. More specifically, we find that market segmentation improves the informational content of the price-signals, which benefits the uninformed buyers by yielding more precise posterior beliefs. Since the introduction of noise precludes complete learning, the uninformed buyers continue to face uncertainty about the product's quality. In future work, it would be interesting to study the effect of risk-aversion under incomplete learning.

One caveat is however in order. That is, our analysis implicitly assumes that both markets are served whether pricing is discriminatory or not. In general, price discrimination makes it profitable to serve markets that would otherwise not be served with uniform pricing. In order words, discriminatory pricing may lead to the opening of new markets. In the presence of uninformed buyers, uniform pricing might make it more likely to exclude the buyers of one of the markets. The reason is that the informational externality generally leads to an increase in the mean prices. Hence, the benefit of market segmentation (in terms of accessibility of the good) is enhanced by the presence of uninformed buyers. See Appendix B.3.

Our second contribution is to show that market segmentation is not necessarily optimal for the firm. Under complete information on the demand side, a monopoly obtains a higher expected profit by charging different prices for market segments having different price elasticities. We show that this conclusion does not hold because of the firm's need to engage in price signaling. Therefore, we outline an important difference regarding the effect of market segmentation between complete and incomplete information environments.

An extension of the model would be to assume the firm already knows μ when choosing whether to segment the market or not. In this case, her best-response would be to segment the market only when μ is above some function of the parameters. Consequently, her choice together with the price-signal(s) will convey information to uninformed buyers. Instead of following a normal distribution, the uninformed buyers' posterior beliefs will now follow a truncated normal distribution. Whether an equilibrium exists in this case is a question we leave for future research.

Chapter 4

Information Choice and Diversity: The Role of Strategic Complementarities

4.1 Introduction[1]

Economic agents are often exposed to more information than they can process. They are also often surrounded by more information sources than their limited cognitive abilities enable them to pay attention to (Sims 2003, 2005, 2006). The increase of information flows and the proliferation of information sources has accelerated in the Digital Age. As Google's CEO Eric Schmidt stated it in 2010, *every two days now we create as much information as we did from the dawn of civilization up until 2003.*[2]

This plethora of information sources is perceived as enabling individuals to make better choices by giving them the possibility to learn about various variables relevant to their decision making. In the theory of decision making under uncertainty, various orderings and measures have been developed that allow a decision maker to rank information structures, and the general conclusion of this research is that, for a decision maker, more information is always better.

When many agents interact, information choice is complicated in at least two ways. The first complication is that in a strategic context, information choice becomes a strategic decision, and the value of the various information choices depends on other agents' information choices, as it is the case for any strategic decision. The second complication is that more information is not always better, because in some games, ignorance has commitment value.

In this paper, we study the strategic choice of information, together with the strategic choice of actions. Our focus is not on the quality, quantity or precision of the information acquired –those are the main issues in decision making and remain important in a strategic context–, but on the diversity of the information that agents choose to acquire, which is meaningful only in a model with many agents. Information is diverse if agents choose to acquire dissimilar information and it is homogenous if agents choose to acquire similar information. In our framework, diversity is an endogenous outcome that results from economic fundamentals, such as the payoffs of the agents.

[1]This chapter is co-authored with Sidartha Gordon.

[2]Siegler, MG. (2010, August 4). Eric Schmidt: Every 2 Days We Create As Much Information As We Did Up To 2003. Retrieved from http://techcrunch.com/2010/08/04/schmidt-data/

Information asymmetries have played an important role in many economic models in many different fields, including industrial organization, the economics of organizations, political economy, macroeconomics and financial economics. As a first step to understand the influence exerted by information on action choices, the models assume an exogenous information structure. But soon, questions arise on where the information structure comes from, who chooses it and how and why they choose it. In an influential paper, Morris and Shin (2002) initiate a literature on the preferences of the central bank (or the planner) over the macroeconomic information structure. In the context of an auction, Bergemann and Pesendorfer (2007) consider the joint design problem of a seller choosing the rules of an auction and the precision of the information of the bidders, in order to maximize his expected profits. Similar questions have been asked in other contexts and a literature on information design has emerged that seeks, more generally, to describe for a given game or for a game to be designed, all possible outcomes a planner could achieve by choosing the players' information structure (Gentzkow and Kamenica, 2011; Bergemman and Morris, 2013; Taneva 2014).

Another approach, the one we adopt, is to model the information structure as the result of the agents' decentralized information choices. Until recently, the literature that followed this path restricted attention to the acquisition of private information and was exclusively concerned with the strategic decision of how much private information (in terms of information quantity, quality or precision) agents choose to acquire when interacting with others. Hellwig and Veldkamp (2009) and Myatt and Wallace (2010) depart from this tradition by allowing agents to acquire potentially public or correlated information in a model of a beauty contest with a continuum of symmetric agents. These authors note that in a strategic context, information about an unknown payoff relevant parameter is at the same time information about what other players know. In these model, actions are either strategic complements or strategic substitutes and agents play actions that are increasing in their signal. The authors make the informal observation that when actions are complements, agents would like to know what the others know and when actions are substitute, they would prefer not to know what the others know, but to have independent information. Thus they make a claim about the *players' preferences over information dependence.* At the same time, in both papers, the authors establish a *complementarity inheritance* result: the sign of the complementarity in actions is passed on to the complementarity in information precisions. If actions are complements, precisions are also complements. If actions are substitutes, precisions are also substitutes. Last, the authors claim that players' preferences over information dependence reflect the players' preferences over action dependence.

Although complementarity inheritance has been noted in several models, all these models are very similar in that they rely on very similar functional forms and distributions. It has been shown that the result is not robust to even slight deviations from the models in which it holds. For example, Jimenez-Martinez (2013) consider the same functional forms as Hellwig and Veldkamp (2009), but assumes two players instead of a continuum, and obtains that complementarity inheritance only holds in a subset of the parameter space. Szkup and Trevino (2014) consider a model that departs from Hellwig and Veldkamp (2009)'s only in that they assume binary actions instead of a continuum. They find that while actions are strategic complements in their model, information precisions need not be complements.

In our view, although complementarity inheritance is related in some way with the players' preferences over information dependence, these are two rather different phenomena. First, complementarity inheritance is a result on preferences over information precision. Preferences over information precision are tricky for two reasons. First, because they mix two different considerations: (i) whether the player wants or not to know more on the uncertain variable; (ii) whether the player wants or not to know what the other players know. Second, because the uncertainty about the other players' knowledge is not held fixed when varying the other players' precisions: it increases as the other players choose a greater precision. Thus, a positive complementarity in precisions, i.e. the fact that following an increase in the information precision by the other players, the remaining player also prefers to increase his precision in reaction, could be due to various reasons. First, it could be that the remaining player is now more willing to know the state more precisely; Second, it could be that he is now more willing to know what the other agents know more precisely; Finally, it could be that he is now more uncertain about what the others know and as a result, is willing to compensate the uncertainty by increasing his own precision. Of course, it could also be some complicated combination of the three reasons we just listed.

This confusion arises even if the players acquire signals that are independent conditional on state realizations. Indeed, when a player acquires private information on the state, he cannot avoid acquiring information on what the other players know, because what they know is correlated with the state. This implies that the issue of precision choice cannot be fully isolated from the players' preferences over information dependence.

It is however possible to study the players' preference over dependence, and the consequences of these preferences on choice and equilibrium, in isolation from the choice of precision. In this paper, we disentangle the two issues and concentrate on the issue of information diversity. We build a model where the only information choice each player has is one between signals that are all equally informative on the state, but are diverse in the sense that they are not perfectly correlated with each other. The amount of information a player can acquire is fixed, perhaps endogenously determined at an earlier stage, but exogenously given from the perspective of the game.[3] Therefore, the only motive driving the players' information choices is whether they want or not to observe the same signal as this or the other player.[4]

We show that there exists a general link between complementarities in payoffs and the players' preferences over information, and that the equilibrium structure can be linked to strategic complementarities in a large class of models, without relying on particular functional forms, distributional assumptions, nor a continuum of agents. Our result is robust in these dimensions, but it requires the amount of information that each player acquires to be held fixed.

We distinguish two components in the choice made by a player. The first one is which signal he chooses to observe. The second one is his action strategy, namely the function that maps the signal he obtains to his actions. One feature of the equilibrium that plays a crucial role are the monotonicity properties of equilibrium action strategies. In general, equilibrium action strategies need not posses any monotonicity property, but in some

[3]This assumption is relatively reasonable in many applications. Firms usually set budgets to information gathering activities, individuals subscribe to newspapers and magazines on a year-term basis, etc. Van Nieuwerburgh and Veldkamp (2009) show how this assumption is not restrictive using a duality argument.

[4]In our model, the players' information choice is akin to a location choice in some abstract information space in which positive dependence can be thought of as an incomplete ranking of distance. By choosing their signals, agents determine the information diversity, in the same way as firms determine the level of brand diversity or geographical dispersion in a market by choosing their brand or their location, as captured in the Hotelling and Salop firm location models.

games they do. Even when this is the case, these properties depend on the information structure.[5]

The problem we study is a complex one: equilibrium action strategies depend on the information structure the players believe in, but the equilibrium information structure depends on the action strategies they expect. We connect player's preferences over information, their equilibrium information choices and the equilibrium information diversity to two types of payoff complementarities: positive or negative complementarities between own action and others' actions, and positive or negative complementarities between own action and the unknown state.

It is useful to decompose our analysis in two parts. First, we need to understand how the information structure determines action strategy monotonicity properties. Second, we need to understand how the interplay between these monotonicity properties and the payoff complementarities in actions shapes the players' preferences over signals and information structures.

For the first part, we show that for any exogenous information structure, if complementarities between own action and state are strong and complementarities between actions are weak, there exists a Nash Bayesian equilibrium in action strategies where action strategies are monotonic in a way that agrees with the state complementarity between own action and state: if this complementarity is positive, the action strategy is increasing and if it is negative, the action strategy is decreasing. Although action complementarities may work against these monotonicities, if they are weak, they do not prevent the existence of such an equilibrium. When the dominance of state complementarities over action complementarities is sufficiently strong, an equilibrium where action monotonicity agrees with state complementarities exists, regardless of what the information structure is. In some special cases, studied in particular by Van Zandt and Vives (2007), action complementarities work in the same direction as state complementarities. This is the case for example if all complementarities, action and state, are positive. In this case, a monotone equilibrium that agrees with state complementarities can be obtained under weaker assumptions.

[5] A large literature studies, for games with exogenous information structures, which fundamentals (in particular, which information structures) guarantee the existence of equilibria where all players' action strategies are increasing in their type (Athey, 2001; McAdams, 2003; Van Zandt and Vives, 2007, Reny, 2010).

For the second part, we study the players' preferences over signals. These preferences are hedonic: a player does not prefer one signal over another per se. Instead the preference depends on what signals the other players choose, on how they react to their signal, and on the dependence properties between the different signals. We show that these hedonic preferences depend on the interplay between action strategy monotonicities and complementarities in actions. More precisely, we show that the preference of a player between two signals depends on which of the two signals is more dependent on (or more similar to) the signals of a certain group of players the agent would like to be informationally close to, and less dependent (or less similar to) the signals of another group of players the agent would like to be informationally far away from. The action strategies of two players are isotonic if they move in the same direction with their signals and antitonic if they move in opposite directions. We show that each player seeks to be informationally close to players whose actions are either isotonic and complement with his own, or antitonic and substitute with his own, and informationally far away from players whose actions are either antitonic and complement with his own, or isotonic and substitute with his own.

We obtain our results for a general class of payoff functions and distributions. With the functional forms and distributions that are usually studied in the literature (e.g. linear and quadratic payoff functions, Gaussian distributions), dependence boils down to the conditional correlation between the signals. To tackle the general case, we define a new notion of statistical dependence between signals. In the case of two players, our dependence ordering between signals coincides with familiar orderings (supermodular, concordance, positive orthant dependence orderings), but in the case of three agents or more, our dependence ordering is novel and of independent interest.

Assembling the two parts of the analysis, we provide sufficient conditions for certain monotonicity patterns and information structures to arise in equilibrium, as well as sufficient conditions for these structures to be the most plausible ones in equilibrium. In particular, we show that if all complementarities (state and action) are positive and public information is feasible, there exists an equilibrium where information is public. In this case, knowing what the other knows allows a player to know a lot on the action of that other player. It is perhaps not surprising that the players choose to obtain the highest level of information by having perfectly correlated information.

However, if all state complementarities are positive, but all action complementarities are negative, and state complementarities sufficiently dominate action complementarities, then there exists an equilibrium where information is as private as possible. It may seem surprising that not knowing what the other players know is better. The reason is that this allows a player to rely more on his signal, without incurring the cost of playing an action that covariates positively with the other players' actions.

Along the way, we show by an example, that equilibrium information structures can be inefficient in an ex ante sense. Interestingly, efficiency can sometimes be restored if the players observe the others' signal choice (but not their realizations). This is because, in this case, the agents internalize some payoff relevant decisions that are ignored when information choice are not observed. In particular, a deviation in information has an effect on actions that is absent when information choice are not observed. These reactions in the action stage may sometimes serve to discipline the players from choosing a suboptimal information structure in the information acquisition stage.

The paper is structured as follows. In Section 2, we present the model. In Section 3, we illustrate all of our results with a simple example. In Sections 4, we introduce definitions that are needed in the analysis of the general case, in particular our new ordering of signal conditional dependence. Section 5 presents the core results of the paper. In Section 6, we show how our model can be applied to different situations. In Section 7, we conclude with a more precise discussion of the literature.

4.2 The Model

In this section, we define a Bayesian game with information choice. Let $I = \{1, ..., N\}$ be a finite set of players. In the game, each player i chooses an action $a_i \in A_i \subseteq \mathbb{R}$. An action profile is denoted $a = (a_1, ..., a_N)$. The players' payoffs depend on a, but also on some unknown state of the world $\theta \in \mathbb{R}$. Each player has a von Neumann-Morgenstern utility function $u_i(a, \theta)$. Actions are chosen simultaneously, but prior to choosing an action, each player chooses a piece of information about θ and observes this information. The information structure is therefore endogenous. We now describe how players choose information.

From the players' point of view, before they acquire information, the state of the world θ is the unknown realization of a random variable Θ, whose support is $T \subseteq \mathbb{R}$. A **signal** is a finite support random variable X_s, which is correlated with the state, and therefore may carry payoff-relevant information, but does not itself directly enter the players' payoffs. Each player i has a set $\mathbb{X}_i$ of signals he can potentially observe, but he can only choose exactly one signal $X_i \in \mathbb{X}_i$, of which he observes the realization $x_i \in \mathbb{R}$. For simplicity, we assume that all the signals that a player can potentially observe have the same finite support $\mathcal{X}_i$ and in addition, we assume without loss of generality that this support is symmetric around zero, i.e. $-\mathcal{X}_i = \mathcal{X}_i$.[6] Let $\mathbb{X} = \bigcup_i \mathbb{X}_i$ be the set of all available signals for all players, $X = (X_1, ..., X_N)$ be a profile of signal choices and $x = (x_1, ..., x_N)$ be a profile of signal realizations. Let F be the joint cdf of the random vector $(\Theta, (X_s : X_s \in \mathbb{X}))$. The tuple $(\mathbb{X}_1, ..., \mathbb{X}_N, F)$ is the **signal structure** of the game. A Bayesian game with information choice is defined by the tuple $\Gamma = (I, (A_1, ..., A_N), (u_1, ..., u_N), (\mathbb{X}_1, ..., \mathbb{X}_N, F))$.

The game unfolds as follows in two stages. Initially all players start with the common prior F, which can be thought of as Nature's mixed strategy. In the first stage, each player i simultaneously chooses a signal $X_i \in \mathbb{X}_i$. In the second stage, each player i privately observes his own signal realization x_i, and then simultaneously chooses an action a_i, without having observed the other players' first stage choices, nor the other players' signal realizations.

In the normal form of this game, a pure strategy for player i is a pair (X_i, α_i) in $\mathbb{X}_i \times A_i^{\mathcal{X}_i \times \mathbb{X}_i}$. The first component in the pair is the signal X_i chosen by player i in stage 1. The second component in the pair is the player's **action strategy** α_i, a mapping that determines player i's action choice, given the realization x_i he observed and the source X_i he chose. In most of the paper, we will restrict attention to pure strategies.[7] For simplicity and without loss of generality, given the focus on pure strategies, we restrict attention to action strategies in $A_i^{\mathcal{X}_i}$, such that the action chosen by each player only depends on his signal

[6] The finite support assumption is made to simplify the exposition, and to avoid uninteresting technical complications. We conjecture that our results extend to the case where $\mathcal{X}$ is infinite. The symmetry assumption is without loss of generality, because the problem is invariant, under any increasing transformation of the set $\mathcal{X}_i$ and any finite set $\mathcal{X}_i$ can be transformed into a symmetric set. The assumption simplifies notations later on.

[7] We consider mixed strategies in Appendix C.5.

observation.[8] A strategy profile $(X, \alpha) = (X_i, \alpha_i)_{i \in I}$ is a **full-fledged Nash-Bayesian equilibrium** if for all i and all $(X_i', \alpha_i') \neq (X_i, \alpha_i)$, we have

$$\mathbb{E}_{\Theta,X}\left(u_i\left(\alpha_i\left(X_i\right), \alpha_{-i}\left(X_{-i}\right), \Theta\right)\right) \geq \mathbb{E}_{\Theta,X_i',X_{-i}}\left(u_i\left(\alpha_i'\left(X_i'\right), \alpha_{-i}\left(X_{-i}\right), \Theta\right)\right).$$

Our goal is to understand what type of information structure can arise in a Nash-Bayesian equilibrium of a Bayesian game with information choice Γ, and how the equilibrium information structure relates to the equilibrium action strategies. Of course, both questions are very broad, and could be analyzed from a variety of angles. Important considerations in a player's choice of a signal could be how informative on θ the different available signals are, how costly they are, or which aspects of θ they reveal.[9] We do not consider this type of choice here. Instead, we focus on the choice of signal conditional dependence: does a player want his signal to depend, conditionally on θ, on the other players' signals or not? In order to eliminate the other motives, and to concentrate on the conditional dependence motive, we assume that a player has access to signals that are all equally informative on the state in the sense of Blackwell. Namely, for each i, not only the support of all signals in $\mathbb{X}_i$ is the same, but in addition the joint marginal distribution of Θ and X_s is the same for all signals X_s in $\mathbb{X}_i$.[10]

Given this restriction, the only remaining degree of freedom the players have when choosing their information is the conditional dependence of their signals with each other. The goal of the paper is to study which dependence patterns between players' signals can arise in a Nash Equilibrium. We interpret these dependence patterns in terms of informational diversity. We also relate the dependence patterns with the monotonicity properties of action strategies and the payoff complementarities in actions. For some information structures, we provide sufficient conditions on the primitives of the model, which ensure that this information structure is chosen by the players in some equilibrium of the game.

[8] With pure strategies, it is without loss of generality to restrict attention to action strategies α_i that do not depend on X_i. Indeed, holding a strategy profile for the other players (X_{-i}, α_{-i}) fixed, any joint distribution over $T \times A^N$ induced by some profile (X, α) such that α_i depends on X_i, can also be induced by some other profile $(X_i, \alpha_i', \alpha_{-i})$ such that α_i' does not depend on X_i.

[9] For example, one signal could reveal θ's sign, whereas another could reveal θ's absolute value.

[10] In particular, if all players have access to the same signals, i.e. all the sets $\mathbb{X}_i$ are equal to $\mathbb{X}$, then our assumption is that all signals in $\mathbb{X}$ are equally informative on θ in the sense of Blackwell.

In the rest of the paper, we refer to Γ as the game with endogenous information structure, that is the game where the players choose which signal to observe. We also refer to Γ_X as the game with an exogenous information structure such that the profile of signal observed by the players is X.

In Section 4.3, we first examine the questions in a simple example and provide complete answers in this context. We then show in Section 4.5 that several of the insights gained from studying the example can be generalized to a large class of games, and do not rely on specific payoffs nor on a particular information structure.

4.3 An Illustrative Example

To fix ideas, we start with a simple example.[11] Suppose that $N = 2$, $\mathbb{X}_1 = \mathbb{X}_2 = \{X_I, X_{II}\}$ and player i chooses an action $a_i \in \mathbb{R}$. The payoffs of the game are

$$u_i(a, \theta) = -a_i^2 + 2b_{ia}a_ia_j + 2b_{i\theta}a_i\theta + K(a_j, \theta) \tag{4.1}$$

where b_{ia} and $b_{i\theta}$ are real numbers for $i \in \{1, 2\}$ and $K(\cdot, \cdot)$ is a function that does not affect the set of Nash-Bayesian equilibria, but may have an effect on welfare. The parameter b_{ia} captures the level of strategic interaction between player i's and player j's actions and the parameter $b_{i\theta}$, the strategic interaction between player i's action and the state. Positivity implies action complementarity and negativity, action substitutability.

The information structure is as follows. The random vector (θ, X_I, X_{II}) is distributed in $\{-1, 1\}^3$ according to a probability distribution function such that the vectors (Θ, X_I) and (Θ, X_{II}) have the same joint marginal distribution given by

	$X_\ell = -1$	$X_\ell = 1$
$\Theta = -1$	$\frac{1-\varepsilon}{2}$	$\frac{\varepsilon}{2}$
$\Theta = 1$	$\frac{\varepsilon}{2}$	$\frac{1-\varepsilon}{2}$

for $\ell \in \{I, II\}$ and where $\varepsilon \in (0, 1/2)$.

[11] One could also consider the normal quadratic payoff setting to illustrate our results. Such an example, however, is not strictly speaking a special case of our model, because the support of the signals is infinite. Note, that the finiteness assumption is made to keep the exposition simple, not for more fundamental reasons.

Moreover, we assume that $\mathbb{P}(\theta = -1) = \mathbb{P}(\theta = 1) = 1/2$, and that the joint distribution of two signals, conditional on $\Theta = \theta \in \{-1, 1\}$ is given by the following matrix:

	$X_{II} = \theta$	$X_{II} \neq \theta$
$X_I = \theta$	$(1-\varepsilon)^2$	$\varepsilon(1-\varepsilon)$
$X_I \neq \theta$	$\varepsilon(1-\varepsilon)$	ε^2

Fixing signal choices $X_i \in \{X_I, X_{II}\}$ for $i = 1, 2$, the ex ante expected payoff of player i given the profile of signal choice X is

$$\begin{aligned}
&\mathbb{E}_{\Theta,X}(u_i(\alpha(X),\Theta)) = \\
&\mathbb{P}(X_i = -1) \cdot \Big(\mathbb{P}(X_j = 1|X_i = -1)\mathbb{E}(u_i(\Theta, \alpha_i(-1), \alpha_j(1))|X_i = -1) \\
&\qquad\qquad +\mathbb{P}(X_j = -1|X_i = -1)\mathbb{E}(u_i(\Theta, \alpha_i(-1), \alpha_j(-1)))|X_i = -1)\Big) \\
&+\mathbb{P}(X_i = 1) \cdot \Big(\mathbb{P}(X_j = -1|X_i = 1)\mathbb{E}(u_i(\Theta, \alpha_i(1), \alpha_j(-1)))|X_i = 1) \\
&\qquad\qquad +\mathbb{P}(X_j = 1|X_i = 1)\mathbb{E}(u_i(\Theta, \alpha_i(1), \alpha_j(1)))|X_i = 1)\Big),
\end{aligned} \tag{4.2}$$

where $\mathbb{P}(X_j = x|X_i = x) = 1$ if $X_i = X_j$ and $\mathbb{P}(X_j = x|X_i = x) = 1-2\varepsilon(1-\varepsilon)$ if $X_i \neq X_j$. By taking the first-order condition to (4.2) with respect to $\alpha_i(1)$ and $\alpha_i(-1)$ for $i = 1, 2$ and then solving for $(\alpha_1(-1), \alpha_1(1), \alpha_2(-1), \alpha_2(1))$, we can compute the equilibrium in the second-stage, that is, once the information structure is fixed. In particular, we obtain $\alpha_i(-1) = -\alpha_i(1)$ and

$$\alpha_i(1) = \frac{(b_{i\theta} + b_{ia}b_{j\theta}[2\mathbb{P}(X_j = x|X_i = x) - 1])}{1 - b_{ia}b_{ja}[2\mathbb{P}(X_j = x|X_i = x) - 1]^2}(1-2\varepsilon). \tag{4.3}$$

Because $\alpha_i(-1) = -\alpha_i(1)$, the number $\alpha_i(1)$ equals the slope of the action strategy of player i, and its sign indicates whether this strategy is increasing or decreasing in his signal.

To avoid non generic trivial cases, we will assume that for all $i, j \in \{1, 2\}$, such that $i \neq j$, we have $b_{i\theta} + b_{ia}b_{j\theta} \neq 0$, $b_{ia}b_{ja} \neq 1$, $b_{i\theta} + b_{ia}b_{j\theta}(1-2\varepsilon)^2 \neq 0$ and $b_{ia}b_{ja}(1-2\varepsilon)^4 \neq 1$. These conditions ensure that (a) for any profile of pure signal strategies (X_1, X_2), a unique pure Nash-Bayesian equilibrium exists in the action game with exogenous information structure (X_1, X_2), and (b) that in this Nash-Equilibrium, each player's action strategy is strictly monotonic: either it is strictly increasing, or it is strictly decreasing.

A look at Equation (4.3) shows that whether player i's action strategy is strictly increasing or strictly decreasing in his signal depends on b_{ia}, the level of strategic complementarity, $b_{i\theta}$, the level of state complementarity and $\mathbb{P}(X_j = x|X_i = x)$, the information structure.

The profile of action strategies (α_1, α_2) is said to be strictly isotonic if both players's actions are either strictly increasing $(\alpha_i(1) > 0)$ or strictly decreasing $(\alpha_i(1) < 0)$. This occurs when $b_{1\theta} + b_{1a}b_{2\theta}(2\mathbb{P}(X_2 = x|X_1 = x) - 1)$ and $b_{2\theta} + b_{2a}b_{1\theta}(2\mathbb{P}(X_1 = x|X_2 = x) - 1)$ have the same sign.

4.3.1 Information Choices

We turn now to the information choice stage of the game. The main question that motivates our work is to understand which assumptions about the payoffs are necessary for information diversity to emerge as a result of the players' individual choice. The binary example allows us to illustrate very clearly the main contribution of the paper.

Fixing the action strategies to (α_1, α_2), where the α_i are the odd functions given by (4.3), the expected payoff for player i can be written as

$$\mathbb{E}_{\Theta,X}(u_i(\alpha(X), \Theta)) = 2b_{ia}(2\mathbb{P}(X_j = x|X_i = x) - 1)\,\alpha_i(1)\,\alpha_j(1) + \text{Constant}. \tag{4.4}$$

In the rest of this section, we perform a partial analysis where (α_1, α_2) is fixed to the expression given in (4.3). Once (α_1, α_2) is fixed, player i's information choice determines $\mathbb{P}(X_j = x|X_i = x)$ and which one is optimal depends on the monotonicity of actions strategies (the sign of $\alpha_i(1)\,\alpha_j(1)$) and on the strategic motive in actions (the sign of b_{ia}). In our example, since the players have access to exactly the same signal, the information structure is either public, if both players observe the same signal, or private, if the players observe different signals.

4.3.1.1 Conflict on the Information Structure

A first observation is that whenever the players have conflicting preferences over the information structure, which is this context means that one of the two players would prefer the information to be public, while the other would prefer it to be private, there cannot be an equilibrium of the endogenous information game that is in pure strategies. This is because the players are playing a game akin to Matching Pennies in the first stage. The players have conflicting preferences over the information structure when $b_{1a}b_{2a} < 0$.

Proposition 4.1. *In the binary quadratic example, if* $b_{1a}b_{2a} < 0$, *players have conflicting preferences over the information structure. In this case, the game with endogenous information acquisition does not have an equilibrium in pure strategies.*

However, it can be shown in a more general context, that this game always has an equilibrium in mixed strategies.[12] In the rest of the analysis of the binary quadratic example, we will restrict attention to the case where players agree on the information structure they like best, that is, we assume that $b_{1a}b_{2a} > 0$.

4.3.1.2 Agreement on the Information Structure

Proposition 4.2 characterizes the equilibrium information structure when (α_1, α_2) is fixed. With isotonic action strategies, a low value for $\mathbb{P}(X_j = x|X_i = x)$ is desirable for player i only if $b_{ia} < 0$, i.e., when actions are substitutes. With antitonic action strategies, a low value for $\mathbb{P}(X_j = x|X_i = x)$ is desirable for player i only if $b_{ia} > 0$, i.e., when actions are complements.

Proposition 4.2. *In the binary quadratic example, let* $b_{1a}b_{2a} > 0$ *and let* $(X, (\alpha_1, \alpha_2))$ *be a pure Nash-Bayesian equilibrium of the game. Then,*

1. $X_1 = X_2$ *only if* $\alpha_1(1)\,\alpha_2(1)\,b_{ia} > 0$ *for* $i = 1, 2$.
2. $X_1 \neq X_2$ *only if* $\alpha_1(1)\,\alpha_2(1)\,b_{ia} < 0$ *for* $i = 1, 2$.

The result in Proposition 4.2 leads us to a more general phenomenon, which we will analyze in greater generality in Theorem 4.4.

A general feature of games with endogenous information structure is that multiple equilibria can exist. For instance, it can be the case that the players choose to acquire the same signal, so that they hold public information, but that the actual signal they observe can be either one contained in the set $\mathbb{X}$. This type of multiplicity is trivial since the dependence pattern among the players' signals is the same for all equilibria. More interesting is the fact

[12] See Appendix C.5.

that non-trivial multiplicity can also occur with endogenous information choice. One such example would be a game where two types of equilibria can be sustained, an equilibrium where the players choose the same signal and another one where the players choose different signals.

Theorem 4.1. *In a game with endogenous information choice, non-trivial multiple equilibria can exist.*

Basically, Theorem 4.1 establishes that the dependence pattern in information choice is not always uniquely pin down by action complementarities. We prove Theorem 4.1 using our binary quadratic example. More specifically, we construct an example with strategic substitutability in actions and complementarities in a player's action and the state and show that both public and private information can be sustained in some equilibrium of the game.

4.3.2 Ex-ante Constrained Inefficiency of the Equilibrium Information Structure

Next, we use the binary quadratic example to show that, under certain conditions, the players' equilibrium signal choices do not result in the information structure a planner would design. We compare the equilibrium of the endogenous information game (in cases covered by Proposition 4.2) with an auxiliary game in which the planner chooses the information structure. In particular, we assume the planner chooses between either (X_I, X_I) or (X_I, X_{II}) , then this information structure becomes common knowledge, and the players simultaneously choose actions in a noncooperative manner. Of course, it is not necessarily obvious what the preferences of the planner should be. In order to avoid this difficulty, we focus on the case where the two players have symmetric payoffs, given by

$$u_i(\theta, a) = -a_i^2 + 2b_a a_i a_j + 2b_\theta a_i \theta + 2b_{\theta a} a_j \theta + 2b_{aa} a_j^2. \tag{4.5}$$

This is a special case of Equation (4.1) considered before, when the players have symmetric payoffs and with $K(a_j, \theta) = 2b_{\theta a} a_j \theta + 2b_{aa} a_j^2$. The terms in $K(a_j, \theta)$ capture an externality that does not affect the Nash-Bayesian equilibrium in the game with endogenous information choice, but contributes to determine which information structures are constrained efficient.

Since the players' payoffs are symmetric, the unique equilibrium is also symmetric. We may then safely assume that the planner maximizes the expected payoff of player 1. Given a profile of signal choices (X_1, X_2) and action strategies α as in (4.3) , the ex ante expected utility of player 1 is written as

$$\begin{aligned}\mathbb{E}_{\Theta,X}(u_1(\alpha(X),\Theta)) &= 2b_a\left(2\mathbb{P}(X_2=x|X_1=x)-1\right)\alpha_1(1)\,\alpha_2(1)-(\alpha_1(1))^2 \\ &\quad +2b_{aa}\left(\alpha_2(1)\right)^2+2b_\theta\alpha_1(1)\,(1-2\varepsilon) \\ &\quad +4b_{\theta a}\alpha_2(1)\,(1-2\varepsilon) \end{aligned} \quad (4.6)$$

Player 1 would optimize his signal choice by considering that he has a direct impact on the first term through a change in $\mathbb{P}(X_2 = x|X_1 = x)$. On the other hand, the planner, when pondering over which information structure to impose, considers that player 1's utility, also depends indirectly on $\mathbb{P}(X_2 = x|X_1 = x)$ as this term enters $\alpha_1(1)$ and $\alpha_2(1)$.

Therefore, the social planner, since he knows the signal choices, uses a different expected payoff function when maximizing welfare, and thus, would not necessarily choose the Nash-Bayesian equilibrium for the signal choice structure.

Theorem 4.2. *A pure Nash-Bayesian equilibrium (X, α) of the game with endogenous information acquisition need not be constrained ex-ante Pareto efficient.*

Note that Theorem 4.2 applies to every game that fits the description of our model in Section 4.2 and not just the particular binary quadratic example. Essentially, the reason for the inefficiency is that the planner will take into consideration the impact of the signal choices on the actions when making a choice on the information structure, an effect that the players do not individually consider. This result suggests that policy intervention is sometimes beneficial in markets for information. In a decentralized system, players may choose either too similar or too dissimilar information, and policy intervention can help to mitigate this type of inefficiency.

In the main model, we make the assumption that signal choices of the first stage are not observed by the players. This is important, since it implies that a deviation from equilibrium play does not affect the other player's action choices in the second stage: the choice of signal and actions are strategically simultaneous. One can imagine situations where signal choices are observable. For example, a company may sign a contract with a market research firms and this may be observable by all other companies.

Interestingly, this difference can have important effects. To see this, consider again the case of two symmetric players. In this case, both players in stage 1 face the problem of the planner, which we analyzed earlier. As we showed, the planner's solution may be disjoint from the set of Nash equilibria of the game where signal choices are unobservable. We can thus deduce the following result.

Theorem 4.3. *Suppose the players publicly observe the profile of signal choices. Then, any pure Nash-Bayesian equilibrium* (X, α) *of the game is constrained ex ante Pareto efficient.*

In this alternative model where the players publicly observe the profile of signal choices, the actions in the second stage are functions of the profile of signal choices. Therefore, a shift of signal by a player has an impact on the other players' actions, which is in turn acknowledged by the deviating player. So it turns out that allowing for the public observation of information choices induces the players to internalize the impact of their signal choice and to behave as the planner would want them to.

Theorem 4.3 suggests that an intervention that mandates players to publicly disclose their sources of information may sometimes be desirable, in that it could help to mitigate excessive information similarity or dissimilarity that may result from a decentralized market for information.

4.4 General Case: Preliminary Definitions

In this Section, we introduce the concepts that are needed in order to generalize some of the insights obtained in the example studied in Section 4.3. We first introduce monotonicity properties, then strategic complementarities in actions. Last, we introduce a new partial order on a set of information structures, the "dependence ordering," which compares the positive dependence between a single player's signal and all the other players' signals across information structures.

4.4.1 Monotonicity Properties of Action Strategies

For any action strategy $\alpha_i : \mathcal{X} \to \mathbb{R}$, we say that α_i is **increasing** if for all $x_i, x_i' \in \mathcal{X}$, we have $x_i \leq x_i' \Longrightarrow \alpha_i(x_i) \leq \alpha_i(x_i')$, and that α_i is **strictly increasing** if for all $x_i, x_i' \in \mathcal{X}$, we have $x_i < x_i' \Longrightarrow \alpha_i(x_i) < \alpha_i(x_i')$. We say that α_i is **(strictly) decreasing** if $-\alpha_i$ is (strictly) increasing.

A profile of action strategies α is **(strictly) monotonic** if for all i, the action strategy α_i is either (strictly) increasing or (strictly) decreasing. It is (strictly) **isotonic** if either, for all i, the action strategy α_i is (strictly) increasing, or for all i, α_i is (strictly) decreasing.

We also want to encompass the cases where the action profiles are (strictly) monotonic, but not necessarily (strictly) isotonic. For any vector $m \in \{1, -1\}^I$, we say that the profile of action strategies α is **(strictly)** m**-monotonic** if for all i, the function $m_i\alpha_i$ is (strictly) increasing. In particular, for any vector m, a (strictly) m-monotonic profile of action strategies α is (strictly) isotonic if, for all i, the m_i have the same sign. Fixing a player i, a profile of action strategies α is (**strictly) antitonic for** i if it is (strictly) m-monotonic, and m satisfies $m_i = -m_j$ for all $j \neq i$.

4.4.2 Strategic Complementarities in Actions

Let $a_{-i,j} \in \mathbb{R}^{I\setminus\{i,j\}}$ be the action strategies of players $I \setminus \{i, j\}$. We say that **player** i **has (strict) positive complementarities in actions with player** $j \neq i$, if for all $a_i' < a_i''$, and all $a_{-i,j} \in \mathbb{R}^{I\setminus\{i,j\}}$ the difference $u_i(a_i'', a_j, a_{-i,j}) - u_i(a_i', a_j, a_{-i,j})$ is (strictly) increasing in a_j. We say that player i has **(strict) negative complementarities in actions with player** $j \neq i$, if for all $a_i' < a_i''$, and all $a_{-i,j} \in \mathbb{R}^{I\setminus\{i,j\}}$, the difference $u_i(a_i'', a_j, a_{-i,j}) - u_i(a_i', a_j, a_{-i,j})$ is (strictly) decreasing in a_j. We say that player i has (strict) positive complementarities in actions if he has (strict) positive complementarities with all the other players. We say that he has (strict) negative complementarities in actions if he has strict (negative) complementarities with all the other players.[13]

[13] In a complete information game in which the best response function of player i is well defined, if u_i has positive (negative) complementarities in actions with player j, his best response function is increasing (decreasing) in a_j (Topkis, 1998; Milgrom and Roberts, 1994).

Although these definitions can be used to describe many situations, we are interested in a richer class of payoff functions where each player may have a (strict) positive complementarity in actions with some players and a (strict) negative complementarity with some other players. The complementarity properties of a payoff function u_i are encoded by a complementarity vector $c^i = \left(c^i_j\right)_{j\in I} \in \{-1,1\}^I$ with $c^i_i = 1$. We say that player i has **(strict) c^i-complementarities in actions** if he has (strict) positive complementarities in actions with all players $j \neq i$ such that $c^i_j = 1$ and (strict) negative complementarities in actions with all players $j \neq i$ such that $c^i_j = -1$. In particular, the case where $c^i = (1, ..., 1)$ corresponds to the case of a player that has a (strict) positive complementarity in actions. Similarly, the case where $c^i_{-i} = (-1, ..., -1)$ corresponds to the case of a player that has a (strict) negative complementarity in actions.[14]

4.4.3 Conditional Dependence Orderings

We now introduce a family of weak partial ordering on a set $\mathbb{X}_i$ of signals accessible to player i, the "conditional dependence orderings." Each such ordering is indexed by some fixed profile of the other player's signals X_{-i} and compares across the signals X_i accessible to player i the positive dependence between X_i and X_{-i}. These orderings play a central role in all the results in Section 4.5.

This notion requires the following definition. For any $k \geq 1$, a subset $L \subseteq \mathcal{X}^k$ is an **increasing subset** of $\mathcal{X}^k$ if for all $x, x' \in \mathcal{X}^k$, such that $x \leq x'$, $x \in L \Rightarrow x' \in L$. Equivalently, L is an increasing subset if its indicator function $\mathbf{1}_L(x)$ is increasing.

Definition 4.1 (weakly greater conditional dependence)**.** *Let $i \in I$ and let X_{-i} be a profile of signals for all players different from i. For all X'_i and X''_i in $\mathbb{X}_i$, we say that X'_i* ***depends at least as much as X''_i on X_{-i} conditionally on Θ****, if for all (θ, x), and all increasing set $L \subseteq \mathcal{X}^{I\setminus\{i\}}$, we have*

$$\mathbb{P}\left(X'_i \geq x \mid X_{-i} \in L,\ \Theta = \theta\right) \geq \mathbb{P}\left(X''_i \geq x \mid X_{-i} \in L,\ \Theta = \theta\right).^{15}$$

[14] If the payoff u_i is twice continuously differentiable, then player i has c^i-complementarities in actions if and only if $c^i_j \frac{\partial^2 u_i}{\partial a_i \partial a_j}(a) \geq 0$ for all $j \neq i$ and all a, and he has strict complementarities in actions if this inequalities hold strictly, almost everywhere. But we do not assume that payoffs have this regularity property.

[15] We provide an equivalent definition, based on the notion of multivariate first order stochastic dominance in Appendix C.1.

For each profile X_{-i} of signals chosen by the other players, this defines a weak partial order over the signals accessible to player i.

Similarly, we define a larger class of weak partial orders over $\mathbb{X}_i$. It enables us to compare, for two signals X_i' and X_i'', whether one signal depends more on some other player signal, but less on another player's signal than the other.

Definition 4.2 (weakly greater conditional d^i-dependence)**.** *Let $i \in I$ and $d^i \in \{-1, 1\}^I$. For all X_i' and X_i'' in $\mathbb{X}_i$, we say that X_i' d^i-**depends at least as much as** X_i'' **on** X_{-i} **conditionally on** Θ, if for all (θ, x), $d_i^i X_i'$ depends at least as much as $d_i^i X_i''$ on $\left(d_j^i X_j\right)_{j \neq i}$.*

For each profile X_{-i} of signals chosen by the other players, and for each dependence vector d^i, this defines a weak partial order over the signals accessible to player i.[16]

Our interpretation of weakly greater conditional d^i-dependence is that player i's signal X_i' depends at least as much on the signals of the players j such that $d_i^i = d_j^i$ and at most as little on the signals of the players j such that $d_i^i = -d_j^i$, as player i's signal X_i'', conditional on θ. Note that weakly greater conditional dependence is precisely the special case of weakly greater conditional d^i-dependence, when $d_1^i = ... = d_N^i \in \{-1, 1\}$.

For the two weak partial orders defined in this subsection, a strict partial order is defined as follows: we say that X_i' **depends more than** X_i'' **on** X_{-i} if X_i' depends as much as X_i'' on X_{-i}, and X_i'' does not depend as much as X_i' on X_{-i}.

4.5 Equilibrium Information Structures

In this section, which is the core of the paper, we determine which information structures can be part of a full-fledged Nash-Bayesian equilibrium of the game with information choice.

The analysis is in two steps, each subdivided in two sub-steps. In the first step (Sections 4.5.1 and 4.5.2), we assume that action strategies are m-monotonic, for some exogenously

[16] Similarly to Definition 4.1 (respectively to Definition 4.2), one can define a weakly greater unconditional dependence (respectively d^i-dependence) partial ordering, which is weaker than the orderings in this definition. Only the conditional versions play a role in the paper, because we are interested in information structures, not in random variables.

fixed vector m. The action strategies themselves may not be fixed, but their monotonicity is. We show in Section 4.5.1 that together with m, the complementarities in actions determine preferences over conditional dependence between own and others' signals. For each pair (i, j) of players, we determine whether player i wants his signal to be as conditionally dependent as possible on player j's signal, whichever this signal is, or as independent as possible of this signal, whichever this signal is. The answer to this question depends on the action complementarities between i and j for player i and on the monotonicities m_i and m_j. More precisely, it only depends on the sign of the product $m_i m_j c^i_j$.

We then characterize in Section 4.5.2 the set of signal profiles which are "compatible" with the maximization of the preferences over conditional dependence described in Section 4.5.1, while still holding the monotonicity vector m of the action strategies fixed. Throughout this first step, the problem we study is akin to the study of a "location game," where a finite number of players choose a *location* from a set of possible locations, and have preferences over locating *close t*o of *far from* each of the other players, except that the "locations" are in fact the signals and the distance is replaced by our notion of conditional dependence.

In the second step (Sections 4.5.3 and 4.5.4), we proceed to endogenize the monotonicity vector m, so as to obtain a full-fledged Nash-Bayesian equilibrium of the Bayesian game with information choice, where both the information structure and the action strategies are jointly determined. In Section 4.5.3, we provide sufficient conditions for a signal profile X to be part of an equilibrium. The way this works is that if the Bayesian game Γ_X with exogenous information equilibrium X admits an equilibrium α in m-monotonic strategies, such that in addition, X is compatible with α in the sense of the characterization of Section 4.5.2, then an equilibrium (X, α) turns out to be a full-fledged equilibrium, and then, it follows from this that X is the signal profile of some equilibrium. In Section 4.5.4, we provide conditions under which the equilibrium information structure is essentially unique, in the sense that all (possibly multiple) equilibria have the same information structure.

4.5.1 Preferences for Conditional (In)Dependence

We now show how the monotonicity m of action strategies and the complementarities in actions for a given player jointly determine this player's preferences over the conditional dependence between his own and other players' signals. Throughout Section 4.5.1, we suppose that a monotonicity vector $m \in \{-1,1\}^I$ is fixed and that players are restricted to play second stage action strategies that are m-monotonic. The restriction to m-monotonic action strategies for some given m is a step in the analysis, but in some cases, the restriction may follow from the primitives. For example, the restriction could result from an external constraint, or from iterative elimination of never best-response action strategies.

4.5.1.1 Preferences for Conditional Dependence

We are now ready to present our characterization of the preferences for conditional dependence.

Theorem 4.4. *Let $i \in I$ and let c^i be a complementarity vector for i. Fix a monotonicity profile $m \in \{-1,1\}^I$ and a profile of signals X_{-i}.*

(i) Suppose that u_i has c^i-complementarity in actions. Suppose that X_i' and X_i'' are two signals in $\mathbb{X}_i$ such that X_i' d^i-depends at least as much on X_{-i} as X_i'' does, where d^i is the conditional dependence vector such that

$$d_j^i = m_i m_j c_j^i \tag{4.7}$$

for all $j \in I \setminus \{i\}$ and $d_i^i = 1$. Then for any profile of pure m-monotonic action strategies α, player i finds signal X_i' at least as good as signal X_i'' :

$$\mathbb{E}_{\Theta,X_i',X_{-i}}\left(u_i\left(\alpha_i\left(X_i'\right),\alpha_{-i}\left(X_{-i}\right),\Theta\right)\right) \geq \mathbb{E}_{\Theta,X_i'',X_{-i}}\left(u_i\left(\alpha_i\left(X_i''\right),\alpha_{-i}\left(X_{-i}\right),\Theta\right)\right). \tag{4.8}$$

(ii) If, in addition, u_i has strictly c^i-complementarity in actions, and X_i' d^i-depends more on X_{-i} than X_i'' does, then for any profile of pure strictly m-monotonic action strategies α, player i strictly prefers signal X_i' to signal X_i'' :

$$\mathbb{E}_{\Theta,X_i',X_{-i}}\left(u_i\left(\alpha_i\left(X_i'\right),\alpha_{-i}\left(X_{-i}\right),\Theta\right)\right) > \mathbb{E}_{\Theta,X_i'',X_{-i}}\left(u_i\left(\alpha_i\left(X_i''\right),\alpha_{-i}\left(X_{-i}\right),\Theta\right)\right). \tag{4.9}$$

Part (i) in Theorem 4.4 is tight in the sense that if X_i' and X_i'' do *not* satisfy X_i' d^i-depends at least as much on X_{-i} as X_i'', where d^i is player i's most preferred dependence vector under m-monotonic action strategies and c^i-complementarity in actions, then a payoff function u_i with c^i-complementarity in actions and an m-monotonic profile of action strategies can be found, such that the inequality (4.8) does not hold. Part *(ii)* is also tight, in a similar sense. In other words, the d^i-dependence ordering over signals is the weakest ordering for which the inequalities (4.8) and (4.9) hold.

In a nutshell, Theorem 4.4 states that player i prefers a signal that is

- as conditionally dependent as possible on the signals of players who belong to one of two groups: first, the players whose actions are complement to his own and whose monotonic strategy varies in the same direction as his own; and second, the players whose actions are substitute to his own and whose monotonic strategy varies in the direction opposite to his own;

- as conditionally independent as possible of the signals of players who belong to one of two groups: first, the players whose actions are complement to his own and whose monotonic strategy varies in a direction opposite to his own; and second, the players whose actions are substitute to his own and whose monotonic strategy varies in the same direction as his own;

Moreover, part *(i)* of Theorem 4.4 simplifies in the four following cases:

a. If player i has a positive complementarity in actions, i.e. $c^i = (1, ..., 1)$, and α is *isotonic,* i.e. $m = (1, ..., 1)$ or $m = (-1, ..., -1)$, then player i prefers a signal that is as conditionally dependent as possible on X_{-i}.

b. If player i has a negative complementarity in actions, i.e. $c_{-i}^i = (-1, ..., -1)$, and α is *isotonic,* i.e. $m = (1, ..., 1)$ or $m = (-1, ..., -1)$, then player i prefers a signal that is as conditionally independent as possible of X_{-i}.

c. If player i has a positive complementarity in actions, and α is *antitonic for* i, i.e. $m_j = -m_i$ for all $j \neq i$, then player i prefers a signal that is as conditionally independent as possible of X_{-i}.

d. If player i has a negative complementarity in actions, and α is *strictly antitonic for* i, then player i prefers a signal that is as conditionally dependent as possible on X_{-i}

Part *(ii)* of Theorem 4.4 also simplifies is a similar way in the analogous four cases. Theorem 4.4 motivates and justifies the following definition.

Definition 4.3. *For any $i \in I$, any monotonicity vector m and any complementarity vector c^i, let* ***player i's most preferred dependence vector under m-monotonic action strategies and c^i-complementarity in actions*** *be the vector d^i such that for all $j \in I$, the equation (4.7) holds.*

4.5.1.2 Link with the Literature on Dependence Orderings

Before we proceed with the rest of the analysis, we shall now pause and discuss how Theorem 4.4 relates to the literature in applied probability, which studies dependence orderings and the logical relations between them.

Part (i) in Theorem 4.4 can be viewed as a generalization of a classic result in this literature, due to Tchen (1980). This scholar compares, for N-variate random vectors with fixed marginals, two dependence orderings: the *Positive Quadrant Dependence ordering* (PQD) and the *Supermodular Dependence ordering* (SPM). While it is well known that SPM dependence implies PQD dependence, Tchen shows that in the case $N = 2$, PQD dependence also implies SPM dependence, i.e. the two are equivalent.[17] In contrast, for $N \geq 3$, PQD dependence no longer implies SPM dependence. Müller and Scarsini (2003) provide a counterexample in the case $N = 3$.[18]

One difference between the two cases, which in our view is crucial for the difference in results, is that while in two dimensions, dependence only involves a single pair of components $(1, 2)$, in three dimensions, it involves three pairs of components: $(1, 2)$, $(2, 3)$ and $(1, 3)$. In other words, it becomes a measure of multilateral dependence (or interdependence).

The paradigm in the applied probability literature is to conceive dependence orderings as multilateral dependence between multiple univariate components. Our result departs from

[17] For definitions of SPM and PQD, see the Appendix C.4. See Müller and Stoyan (2002, Theorem 3.8.2) for related results.

[18] One way implications between various interdependence orderings and some equivalences have been obtained for the case $N \geq 3$. They are reviewed by Strulovici and Meyer (2012), who also establish new implications (see also Christofides and Vaggelatou, 2004; Müller and Stoyan, 2002; and Hu, Müller and Scarsini, 2004).

this paradigm by defining dependence between two components of a multivariate random vector, one of which is itself multivariate.

In applied probability literature, different concepts of dependence relate to our definition of dependence between an univariate and a multivariate components.[19] These are called concepts of *setwise dependence* (e.g. Chhetry et al., 1989) and can be seen as generalizations of the positive upper (lower) orthant dependence concept.[20]

The various concepts of setwise dependence describe the dependence between random vectors, while disregarding the dependence between the univariate components within each of these vectors. In our particular problem, we only need to study the dependence between the signal choice of a given player and the choices of other players, not the dependence patterns among these other players' signals. Although such concepts of setwise dependence between vectors have been studied, we are not aware of any work studying setwise dependence *orderings,* such as the one we define.

Our Theorem 4.4 can be viewed as an extension of Tchen's result to the more general setwise case. Indeed, our concept of dependence is an appropriate generalization of PQD dependence in a setwise setting, and the inequality (4.8) is also an appropriate generalization of SPM dependence.[21] More specifically, in the special case where the state Θ is deterministic, $N = 2$, $i = 1$, α_1 and α_2 are the identity functions (so that $m_1 = m_2 = 1$) and $c^1 = (1, 1)$, we obtain the following result.

Corollary 4.1 (Tchen, 1980)**.** *Suppose that u_1 has $(1, 1)$-complementarity in actions. Fix a signal $X_2 \in \mathbb{X}_2$. Suppose that X_1 and X_1' are two signals in $\mathbb{X}_1$ such that X_1' depends at least as much as X_1 on X_2. Then player 1 finds signal X_1' at least as good as signal X_1*

$$\mathbb{E}_{X_1',X_2}\left(u_1\left(X_1', X_2\right)\right) \geq \mathbb{E}_{X_1,X_2}\left(u_1\left(X_1, X_2\right)\right).$$

[19] We thank Marco Scarsini for pointing us to this literature.

[20] One such concept in Chhetry et al. (1989) is setwise positive upper (lower) orthant dependence, SPUOD (SPLOD). The set $(X_1, ..., X_k)$ with X_t a $p_t \times 1$ vector in $\mathbb{R}_t^p$ is said to be setwise positively upper (lower) orthant dependent if for all $x_t \in \mathbb{R}_t^p,\ t = 1, .., k$,

$$\mathbb{P}\left[\cap_{t=1}^k \{X_t > (\leq) x_t\}\right] \geq \Pi_{t=1}^k \mathbb{P}\left[X_t > (\leq) x_t\right].$$

[21] An appropriate name for this generalization of the SPM dependence ordering would be *Increasing Differences dependence ordering.*

This Corollary is a reformulation of Tchen's result, because $(1, 1)$-complementarity in actions coincides with supermodularity for a function of two variables, and the assumption that X_1' depends at least as much on X_2 as X_1 is equivalent to the assumption that (X_1', X_2) is at least as PQD dependent as (X_1, X_2) in the bivariate case.

Another way in which Theorem 4.4 extends Tchen's result is not mathematical, but purely conceptual. Our comparison dependence orderings are conditional on the state Θ. While this does not raise any mathematical difficulty, it allows us to interpret our dependence ordering as similarity between information sets, rather than between the components of a random vector. Similarly, inequality (4.8) indicates a preference for an information set over another. This generalization enables us to interpret Theorem 4.4 as telling us, between two pieces of information, which one a player prefers to have, depending on his preferences over actions, when one piece is more (or less) similar than the other to the other players' information.

We believe that both our new bilateral (conditional) dependence ordering for multivariate distributions and the extension of Tchen's result in Theorem 4.4 are of independent interest, and that they are likely to have applications in economics, in addition to the particular one we study in this paper. We now return to the analysis of Bayesian games with information choice.

4.5.2 Equilibria of the Information Choice Game

Every Bayesian game with information choice Γ and every fixed profile of action strategies α induce an information choice game Γ_α, which is the normal form in which each players $i \in I$ chooses a signal $X_i \in \mathbb{X}_i$ and receives the payoffs $\mathbb{E}_{\Theta,X}\left(u_i\left(\alpha_i\left(X_i\right), \alpha_{-i}\left(X_{-i}\right), \Theta\right)\right),$ where α is the fixed profile of action strategies.

We now use part (ii) in Theorem 4.4 to obtain a characterization of the equilibria of the information choice game Γ_α when α is a strictly m-monotonic action strategy profile, for some monotonicity vector m.

Corollary 4.2. *Let* $\left(c^1, ..., c^N\right)$ *be a profile of complementarity vectors. Suppose that for each* i, *player* i *has strict* c^i*-complementarities in actions. Suppose that* (X, α) *is a full-fledged Nash-Bayesian equilibrium profile in pure strategies. Suppose that* α *is strictly* m*-monotonic, for some monotonicity vector* m. *Then for all* i, *there exists no signal* X_i' *that* d^i*-depends more on* X_{-i} *than* X_i, *where* d^i *is player* i*'s most preferred dependence vector under* m*-monotonic action strategies and* c^i*-complementarity in actions.*

In some cases, by this result, Nash-Bayesian equilibrium conditions imply that m-monotonicity of action strategies together with c^i-complementarity in actions pin down an essentially unique information structure. Whether or not this is the case depends on the geometric structure of the signal structure. We now study this question in more detail.

4.5.2.1 Most d^i-dependent Signals

The first set of situations where monotonicity pins down the information structure is when the dependence preference of each player are to some extent independent on the information choices of the other players. For example, if there are two players and both have access to two different signals X_I and X_{II}, none of these signals is intrinsically more public than the other. For each player, the more public signal is the signal the other player chooses. Yet, in many contexts, the different available signals do not possess that kind of symmetry. Some signals are unambiguously more public than others, some are unambiguously more private than others. We push this idea even further and introduce the concept of a most d^i-dependent signal, independently of the signals chosen by others.

For any signal structure $(\mathbb{X}_1, ..., \mathbb{X}_N)$, any distribution F and any dependence vector d^i, and for all X_i in $\mathbb{X}_i$, we say that X_i^* is the **most d^i-dependent signal in** $\mathbb{X}_i$ if X_i^* has the property that, for all signal profiles X_{-i}, the signal X_i^* is a greatest element of the "as d^i-dependent on X_{-i} as" weak partial order on $\mathbb{X}_i$. In plain words, the signal X_i^* provides player i with information that is more d^i-dependent on the other player's signals, than any other signal player i could choose to observe, regardless of what signals the other players choose to observe. In particular, for $d^i = (1, ..., 1)$, we call this signal player i**'s most**

public signal and for $d^i_{-i} = (-1, ..., -1)$, for $d^i_{-i} = (d^i_j)_{j \neq i}$ we call this signal player i**'s most private signal.**[22]

We obtain the following direct implication of Corollary 4.2 (a direct implication of Theorem 4.4).

Corollary 4.3. *Let $\left(c^1, ..., c^N\right)$ be a profile of complementarity vectors. Suppose that for each i, player i has strict c^i-complementarities in actions. Suppose that (X, α) is a full-fledged Nash-Bayesian equilibrium profile in pure strategies. Suppose that α is strictly m-monotonic, for some monotonicity vector m. Suppose that for all $i \in I$, player i has a most d^i-dependent signal in $\mathbb{X}_i$, where d^i is player i's most preferred dependence vector under m-monotonic action strategies and c^i-complementarity in actions. Then X_i must be a most d^i-dependent signal in $\mathbb{X}_i$.*

To illustrate the usefulness of this result, it is helpful to consider the following three special cases:

1. If every player has a most public signal, and $m_i m_j c^i_j = 1$ for all i, j such that $i \neq j$, then in any Nash-Bayesian equilibrium in which actions are strictly m-monotonic, every player must be choosing his most public signal.

2. More specifically, if every player has a most public signal and strict positive complementarities in actions, then in any Nash-Bayesian equilibrium in which actions are strictly isotonic, every player must be choosing his most public signal.

3. If every player has a most private signal, and $m_i m_j c^i_j = -1$ for all i, j such that $i \neq j$, then in any Nash-Bayesian equilibrium in which actions are strictly m-monotonic, every player must be choosing his most private signal.

Of course, whether or not Corollary 4.3 has bite hinges upon whether there exists or not a most d^i-dependent signal for each player i. The answer to this question depends on the geometry of the signal structure $\mathbb{X}$.

[22] Most d^i-dependent signals in $\mathbb{X}_i$ need not be unique, but they are payoff-equivalent under m-monotonic action strategies. See the last paragraph of Section 4.5.4. If two different players i and j both have a most public signal, then it must be that the intersection $\mathbb{X}_i \cap \mathbb{X}_j$ has at most (essentially) one element. Moreover, if this intersection is indeed nonempty, its (essentially) unique element is both i's and j's most public signal.

To gain intuition, it is helpful to visualize the information choice game (where action monotonicities are held fixed) as a location game in a spatial setting. For each player i, the set $\mathbb{X}_i$ is the analog of a set of admissible locations where this player can locate. Each player's preferences over locations only depends on where the other players locate. He wants to be close to some of them and far away from some of them. The preferences of player i are summarized by the vector d^i. For each $j \neq i$, if $d^i_j = 1$, call j a friend of i, and if $d^i_j = -1$, call j an enemy of i. Note that j may be a friend of i, while i is an enemy of j. Then each player i wants to locate as close as possible to all his friends and as far as possible from all his enemies. Corollary 4.3 says that if each player has a location X_i that minimizes distance with all his friends and maximizes distance with all his enemies, regardless of where all of them locate, then player i must choose this location in any Nash equilibrium in pure strategies.

In Appendix C.2 we show how to construct examples of signal structures $\mathbb{X}$ that admit a most public signal, a most private signal and a most d^i-dependent signal.

One case of interest where Corollary 4.3 does not apply is when the signal structure is symmetric, i.e. when $\mathbb{X}_1 = ... = \mathbb{X}_N$. In this case, except in degenerate cases, players do not have most d^i-dependent signals. But, in the symmetric case, we can still obtain sharp predictions in the case where $N = 2$.

If $N = 2$ and the two players have strict positive (negative) complementarities in actions, in any pure Nash-Bayesian equilibrium, whose actions are strictly isotonic (antitonic for both players) they choose to acquire essentially the same information. We say that (X_1, X_2) is **public information** if the event $X_1 = X_2$ has probability one.

Corollary 4.4. *Suppose that $N = 2$, that $\mathbb{X}_1 = \mathbb{X}_2$ and suppose that the payoff functions u_i have strict positive (negative) complementarities in actions, for $i = 1, 2$. Let $(X_1, X_2, \alpha_1, \alpha_2)$ be a pure full-fledged Nash-Bayesian equilibrium of the game. If α is strictly isotonic (antitonic for both players), the equilibrium information structure must be public information.*[23]

Note that, under the assumptions of Corollary 4.4, the equilibrium is never unique: if $(X_I, X_I, \alpha_1, \alpha_2)$ is a full-fledged Nash-Bayesian equilibrium, then for any $X_{II} \in \mathbb{X}$, the

[23]It is worth noting that Corollary 4.4 does not generalize to the case of three players or more.

profile $(X_{II}, X_{II}, \alpha_1, \alpha_2)$ is also a Nash-Bayesian equilibrium. But all of these equilibria are payoff equivalent. The players higher order beliefs are also the same across all the equilibria. In the example of Section 4.3 we referred to this as trivial multiplicity.

Corollary 4.4 describes a case where the two players "agree" on the information structure they want. In the case of a symmetric signal structure with two players, this agreement is necessary. Indeed, in the case where they do not agree, it is easy to see that no equilibrium in pure strategies can exist, even when, for any information structure X, the game with exogenous information structure X admits a Nash-Bayesian equilibrium in pure strategies. This was illustrated in the example studied in Section 4.3, but it is a more general phenomenon.

Corollary 4.5. *Suppose that $N = 2$, that $\mathbb{X}_1 = \mathbb{X}_2$ and contain at least two signals whose realizations are not equal with probability one, and suppose that both players have strict complementarities in actions, but of opposing signs. There is no pure full-fledged Nash Bayesian equilibrium $(X_1, X_2, \alpha_1, \alpha_2)$ such that α is strictly monotonic.*

To see why Corollary 4.5 is true, suppose for example that player 1 has a strict positive complementarity in actions, while player 2 has a strict negative complementarity in actions. Suppose further that action strategies are set to be strictly isotonic in the second stage. Then, there cannot be an equilibrium in pure strategies in the first stage. This is because player 1 wants to observe the same signal as player 2, in order to increase dependence, whereas player 2 wants to observe a signal different from player 1, in order to decrease dependence. Similarly, if action strategies are set to be strictly antitonic in the second stage, there cannot be an equilibrium in pure strategies in the first stage either. This is because player 2 now wants to observe the same signal as player 1, in order to increase dependence, whereas player 1 wants to observe a signal different from player 2, in order to decrease dependence. In both cases, the signal choice in the first stage of the game has a structure à la matching pennies. No equilibrium in pure strategies exists, although (as shown in Appendix C.5) a mixed equilibrium can always be constructed.

Note, however, that Corollary 4.5 need not hold when the signal structure is not symmetric. For example, as shown in Corollary 4.3, and with the payoff configuration of the previous paragraph, if player 1 has access to a most-dependent signal and player 2 has access to a less dependent signal, then the information structure defined by these two signals is a (unique) candidate for an equilibrium where the strategies are strictly isotonic. Similarly, if player 1 has access to a least-dependent signal and player 2 has access to a most dependent signal, then the information structure defined by these two signals is a (unique) candidate for an equilibrium where the strategies are strictly isotonic.

4.5.3 Full-fledged Equilibrium Information Structures

In general, the information structure and the actions strategies (and their monotonicity properties) are jointly determined in equilibrium. The conditional dependence properties between the signals chosen by the different players contribute to determine incentives to choose actions strategies that are either increasing or decreasing in signal realizations. Conversely, in Sections 4.5.1 and 4.5.2, we showed how the monotonicity properties of the action strategies chosen by the players contribute to determine their incentives to choose more or less conditionally dependent signals. In this section, we propose conditions that guarantee that an equilibrium exists, with certain pre-specified monotonicity characteristics and with a certain pre-specified information structure.

The key condition that guarantees existence of an equilibrium is a form of compatibility between the monotonicity properties of the candidate action strategies and the candidate information structure. But other conditions are required as well. In the equilibrium we construct, the m-monotonicity of the equilibrium action strategies agrees with the state complementarity in the players' payoffs. For example, if for player i, the state and his action are positive complements, his equilibrium action strategy will be increasing in the signal. If they are negative complements instead, his equilibrium strategy will be decreasing in the signal. Existence of such an equilibrium is established in two cases. First, when the action complementarities are aligned with the state complementarities, and in effect reinforce them, by giving players additional incentives to play m-monotonic action strategies. Second, when the sign of the monotonicity of equilibrium action strategies is predictable.

4.5.3.1 When Action Complementarities Reinforce State Complementarities

We first provide sufficient conditions that ensure that an equilibrium exists where players have public information (or most public information). We show that a very simple condition ensuring this is that all players have positive complementarities, both in actions and in state. But we also show that it is the case for a weaker condition. It only requires that the state complementarities be *aligned* with the action complementarities in the sense that if the sign of the players' state complementarities are given by the vector m and their action complementarities are given by the vectors $(c^i)_{i \in I}$, then for all $i \neq j$, we have $c^i_j = m_i m_j$. Clearly, the situation where all (action and state) complementarities are positive is the special case where $m = (1, ..., 1)$ and $c^i = (1, ..., 1)$ for all i.

First, when state complementarities are aligned with action complementarities, we establish, using a result due to Van Zandt and Vives (2007), that fixing the information structure X, the game Γ_X admits an m-monotonic Nash-Bayesian equilibrium α^X. A partial intuition for why this is true is that the monotonicity of each player i's equilibrium strategy m_i is dictated by his state complementarity, also m_i, but is further reinforced by the action complementarities.

For example, consider the situation where all players have positive action and state complementarity. Suppose further that each signal depends positively on the state so that a high realization is evidence that the state is likely to be high, and that conditionally on the state, all signals are positively dependent among each other. Thus, when a player observes a high (low) realization, he believes that the state is high and that other players' realizations are also high. A first-order effect is that he wants to play a high (low) action, so that his action will be aligned with the state, which he believes to be high (low). But there is a second-order effect, which is that he also believes that, conditional on the state, the other players' realizations are high (low), so that they are likely to be playing high (low) actions. Because the action complementarities are positive, this gives an additional reason to play a high (low) action. Since all players can realize this, there is then a third order effect which further increases the incentive to play a high (low) action. The process then goes on ad infinitum.

Second, we show that when state complementarities are aligned with action complementarities, and players choose either a public or a most public information structure X, and play an m-monotonic Nash-Bayesian equilibrium α^X of the game Γ_X, no player can strictly gain by deviating to another strategy $(X'_i, \alpha'_i) \neq \left(X_i, \alpha_i^X\right)$. To establish this, we first argue that assuming by contradiction that the deviation (X'_i, α'_i) is a profitable deviation, there exists another deviation (X'_i, α''_i) that is even better, where α''_i is m_i-monotonic. But in this case, we can show that the deviation (X_i, α''_i) is even better, so it must be profitable, which contradicts the fact that α^X was an equilibrium of the game Γ_X in the first place.

The precise conditions that guarantee the existence of an equilibrium where dependence among signals is maximized (most public signals are chosen) are the following.

Theorem 4.5. *Let $N \geq 2$. Let m be a monotonicity vector and let $(c^i)_{i \in I}$ be the profile of complementarity vectors such that for all i, j, $c^i_j = m_i m_j$. Suppose that*

i. For each $i \in I$, u_i has c^i-complementarities in actions.

ii. For each $i \in I$, u_i has m_i-complementarities in a_i and θ.

iii. For all $x_i < x'_i$, the distribution of θ conditional on $X_i = x'_i$ first order stochastically dominates the distribution of θ conditional on $X_i = x_i$.

iv. For every profile X, all i and all $x_i < x'_i$, the distribution of X_{-i} conditional on $X_i = x'_i$ first order stochastically dominates the distribution of X_{-i} conditional on $X_i = x_i$.

Then for any profile X of signal choices such that for each i, the signal X_i is most dependent on X_{-i} in $\mathbb{X}_i$, there exists an m-monotonic action strategy profile α such that (X, α) is a full-fledged Nash-Bayesian equilibrium for the game.

The following result is a direct implication of Theorem 4.5.

Corollary 4.6. *Let $N \geq 2$. Suppose that conditions (i) to (iv) of Theorem 4.5 hold. In addition, suppose that a public (most public) signal profile exists. Then for any public (most public) information signal profile X, there exists an m-monotonic action strategy profile α such that (X, α) is a full-fledged Nash-Bayesian equilibrium for the game Γ.*

4.5.3.2 When Monotonicity of Action Strategies Is Predictable

We now move away from the case where action complementarities reinforce state complementarities and consider cases where action complementarities may create incentives for the players to play actions that vary in the direction opposite to the one which agrees with the state complementarity. In that case, it is not possible to predict in general whether the action strategies of any full-fledged Nash-Bayesian equilibrium will be m-monotonic for any particular monotonicity vector m.

Nevertheless, in some context, we may have enough information to know that in equilibrium, action strategies are m-monotonic for some pre-specified vector m. For example, it could be that all profiles of action strategies α that are not m-monotonic are strictly dominated, or do not survive iterated elimination of strictly dominated strategies.

One natural reason why m may be predictable is that each player i could have an m_i-complementarity in state which is strong enough that it dominates any potential higher order effect and that it single-handedly determines the monotonicity of equilibrium action strategies. For example, in an symmetric Cournot duopoly, where firms 1 and 2 should produce a larger quantity when the state is high and a lower quantity when the state is low, the negative action complementarity creates a contrarian incentive. But if this second order effect and all other higher order effects are negligible compared to the first order effect, it could be predictable that equilibrium action strategies are increasing in signal realizations.

Alternatively, the monotonicity of action strategies could be predictable for other reasons. For example, in the same setting, it could be that firm 1's complementarity in state dominates firm 1's (negative) action complementarities, so that this firm always plays an increasing action strategy in any equilibrium. In contrast, firm 2 could have a (negative) action complementarity that is much stronger than its positive complementarity in state. This and the fact that firm 1 plays an increasing action strategy in any equilibrium could imply that firm 2 plays a decreasing action strategy in any equilibrium. As a result, in any equilibrium, the action strategy profile is $(1, -1)$-monotonic.

When the monotonicity of action strategies is predictable in this sense, and under the assumption that each player i has c^i-complementarities in actions, in light of Theorem 4.4, a natural candidate X may emerge for an equilibrium information structure: one that has the property that for each i, X_i is most d^i-dependent on X_{-i} than any other signal in $\mathbb{X}_i$, where d^i is agent i's most preferred dependence vector under m-monotonic action strategies and c^i-complementarities in actions (defined in Equation (4.7)). Our main result in this section provides a simple sufficient condition for this candidate to form a full-fledged Nash-Bayesian equilibrium, together with some m-monotonic action strategy α. The condition imposes that for any player i, and for any signal deviation X_i', in the game $\Gamma_{X_i',X_{-i}}$, player i has at least one best response to α_{-i} that is m_i-monotonic.

The logic at play here is the same as in Theorem 4.5. There, because action complementarities reinforce state complementarities, monotonic equilibrium action strategies are known to exist. We thus obtain $d_j^i = m_i m_j c_i^j = 1 \cdot 1 \cdot 1$ for all i and j and the natural candidate that emerges is any information structure where positive dependence is maximized for all players. The same condition on deviations to other signals X_i' as the one stated in the previous paragraph can then be obtained from primitives and is sufficient to establish that this candidate is indeed an equilibrium.

The difference now is that, both existence and the condition on deviation to signals X_i' are assumed rather than derived from primitives, and therefore need to be verified directly. But the result shows that this approach can be adapted beyond the case where complementarities reinforce each other.

For any strategy profile (X, α), we say that the action strategy α_i is a best response for player i in game Γ_X, if for all x_i,

$$\alpha_i(x_i) \in \arg\max_{a_i \in A_i} \mathbb{E}_{\Theta,X} \left(u_i\left(a_i, \alpha_{-i}(X_{-i}), \Theta\right) \mid X_i = x_i \right).$$

Theorem 4.6. *Let $N \geq 2$. Let m be a monotonicity vector. For each i, let c^i be a complementarity vector, and let d^i be agent i's most preferred dependence vector under m-monotonic action strategies and c^i-complementarities in actions. Let (X, α) be a strategy profile such that:*

i. For each i, the payoff of player i has c^i-complementarities in actions.

ii. For each i, X_i is most d^i-dependent on X_{-i} in $\mathbb{X}_i$.

iii. The profile α is an m-monotonic Nash-Bayesian equilibrium of the game Γ_X.

iv. For each i and each X_i', in the game $\Gamma_{X_i',X_{-i}}$, player i has a best response α_i' to the action profile α_{-i} that is m_i-monotonic.

Then, the profile (X, α) is a full-fledged Nash-Bayesian equilibrium of the game Γ.

Note that assumptions (*iii*) and (*iv*) are not on primitives, but they could be derived from assumptions on primitives: for example, one could assume that only m-monotonic profiles survive iterated elimination of strictly dominated strategies (and existence of a Nash-Bayesian equilibrium of game Γ_X could be established using Kakutani's fixed point theorem). However, we feel that in practice, both assumptions (*iii*) and (*iv*) are much less restrictive than those assumptions on primitives, and can often be easily checked in most applications. The scope of our result is thus larger than if we imposed those assumptions.

4.5.4 Conditions for A Unique Full-fledged Equilibrium Information Structure

We would now like to provide conditions under which the information structure is essentially the same in all full-fledged Nash-Bayesian equilibria. In fact, this question is already answered in Corollary 4.3, which only needs to be reformulated and reinterpreted.

In the reformulation, the key assumptions we make are (i) that for all information structures X, all Nash-Bayesian equilibria are strictly m-monotonic for some monotonicity vector m such that there is a profile of dependence vectors $\left(d^1, ..., d^N\right)$ where each d^i is agent i's most preferred dependence vector under m-monotonic action strategies and his actual c^i-complementarities in actions and (ii) that each agent has a most d^i-dependent signal in $\mathbb{X}_i$. When this holds, then all full-fledged Nash-Bayesian equilibria of the game must have an information structure where each agent chooses a most d^i-dependent signal in $\mathbb{X}_i$.

Again, as in Section 4.5.3.2, the predetermined monotonicity m in assumption (i) could be the resulting balance of a number of contrarian forces.

Corollary 4.7. *Let $N \geq 2$. Suppose that for each i, player i has c^i-complementarities in actions. Suppose for any signal profile X, all Nash-Bayesian equilibria of the game Γ_X are strictly m-monotonic, for some m such that there exist vectors $\left(d^1, ..., d^N\right)$ such that for all i, d^i is player i's most preferred dependence vector under m-monotonic action strategies and c^i-complementarities in actions. Suppose that each player i has a most d^i-dependent signal in $\mathbb{X}_i$. Then in any full-fledged Nash-Bayesian equilibrium (X, α) of the game Γ, for each i, the signal X_i is a most d^i-dependent signal in $\mathbb{X}_i$.*

Although the formulation and the interpretation is different, this result is formally equivalent to Corollary 4.3. Strictly speaking, Corollary 4.7 does not pin down a unique information structure, because each player i may have multiple most d^i-dependent signals. But most d^i-dependent signals are interchangeable when actions are played according to a strictly m-monotonic profile α, in the sense that if X and X' are two signal profiles, each made up of (possibly distinct) most d^i-dependent signals X_i and X_i', where for each i, d^i is player i's most preferred dependence vector under m-monotonic action strategies and his c^i-complementarities in actions, then (X', α) is also a full-fledged Nash Bayesian equilibrium of the game and all players obtain the same expected payoff in both equilibria.

4.6 Applications

In this section, we now illustrate how the model can be applied to different contexts.

4.6.1 Currency Speculation

This example is adapted from the model of currency speculation of Morris and Shin (1998), to which we add information choice. The game is played between N agents. Each agent i decides whether $(a_i = 1)$ or not $(a_i = 0)$ she speculates against a currency. The bank then observes the realization of the number n of agents who speculate and the realization $\theta \in \{\theta_1, ..., \theta_p\}$, with $\theta_1 < ... < \theta_p$, of an uncertain but relevant fundamental state Θ. It defends the currency if and only if $n \leq \theta$. If this condition holds, the attack is "unsuccessful" and it is "successful" otherwise. The payoff of an agent who chooses to speculate is $\pi - b\theta - K$, with $\pi > 0$, $b > 0$ and $K > 0$ if the attack is successful and $-K$ if it is unsuccessful. The payoff of not speculating is 0.

Before deciding whether or not to speculate, each of the agents i chooses a single signal X_i from a common set of accessible signals $\mathbb{X}_i = \mathbb{X}$. Suppose that assumptions (*iii*) and (*iv*) of Theorem 4.5 hold, so that high realizations of any of the accessible signals are associated with high realizations both of the state and of the other accessible signals.

In this game, all agents have negative complementarities in state and positive complementarities in the other agents' actions: if other agents are more likely to attack, the agent is more willing to attack.

All assumptions of Theorem 4.5 are satisfied. Therefore we know that at least one full-fledged Nash-Bayesian equilibrium exists, where all agents choose to observe the same (i.e. public) signal. Doing so enables them to perfectly coordinate: whenever the realization of the public signal is lower than some threshold, all agents attack.[24]

An implication is that unanimous attacks followed by devaluation sometimes occur even for high realizations of the fundamental.

These unanimous attacks are a poor signal of the fundamental, since they only reflect part of the information of one signal, and ignore the information contained in all the other signals that the agents choose not to observe.

In this equilibrium, the movements of the currency are essentially driven by random realizations of a signal used by speculators for coordination purposes.

While such an equilibrium may be detrimental to society, it may be good for speculators. But in a version of the model, this "herding on the same signal" equilibrium may be bad for the speculators themselves.

Following Goldstein, Ozdenoren and Yuan (2011), consider instead a setting where the central bank does not observe the realization of Θ at all, nor does it observe a signal

[24] Under our assumption that the set of accessible signals $\mathbb{X}_i$ is the same for all speculators, we cannot rule out other information structures in equilibria. Under the alternative assumption that each speculator i has a unique most public signal in $\mathbb{X}_i$, an argument in the spirit of Corollary 4.7 can be used to establish that for an open set of parameters, a unique information structure arise in all full-fledged Nash-Bayesian equilibria, such that all agents choose their most public signal (possibly the same one).

of it, and where n does not directly enter its decision on whether or not to defend the currency. Suppose instead that the central bank learns about the state from the occurrence and potentially also from the size of an attack, and that it defends the currency peg if and only if $\mathbb{E}(\Theta \mid n) > 0$. Taking again the bank's decision rule as given, the agents play a coordination game with information choice, the payoffs of which are endogenous, since they depend on which signal action strategies the bank expects them to play. For the same reasons as in the previous model, there exists an equilibrium where all agents choose to observe the same signal. This implies that a unanimous attack is a weak signal that the realization of the fundamental is low. Consider the case where $\mathbb{E}(\Theta) > 0$ and $\mathbb{E}(\Theta \mid X_i \leq x^*) > 0$, where x^* is the threshold realization of the common signal below which the agents choose to attack. Then, when observing an attack, the bank does not find evidence in favor of abandoning the peg convincing enough, because the unanimous attack only reflects the information of one signal.

Because of the excessive similarity in speculators' information, the bank chooses to always defend the currency. Consequently, attacks are never successful and therefore they never occur in equilibrium.

What happens in this case is that, while individually, each speculator has an incentive to observe the same signal as the others, their collective interest is that the bank expects them to acquire diversified information. The equilibrium is however determined by their individual interest. From the speculators' point of view, informational diversity is a public good that they under-provide in equilibrium.

The above analysis can also be applied to technology adoption in the presence of positive network externalities, or the problem of collective action in a revolutionary movement. In both cases, complementarities imply that players might observe the same signal so that the aggregate action is not a good aggregator of all available information.

4.6.2 Other Applications

In each of the following examples, θ is an uncertain parameter with support $\{-1, 1\}$, with $\mathbb{P}(\Theta = -1) = \mathbb{P}(\Theta = 1) = 1/2$. Moreover, available signals are $\mathbb{X} = (X_1, ..., X_L)$ with $L \geq N$, each with support in $\{-1, 1\}$ such that the random vector (θ, X_ℓ) is distributed in

$\{-1,1\}^2$ according to a joint marginal distribution given by

$$\begin{array}{c|c|c} & X_\ell = -1 & X_\ell = 1 \\ \hline \Theta = -1 & \frac{1-\varepsilon}{2} & \frac{\varepsilon}{2} \\ \hline \Theta = 1 & \frac{\varepsilon}{2} & \frac{1-\varepsilon}{2} \end{array} \tag{4.10}$$

for all X_ℓ and $\varepsilon \in (0, 1/2)$.

The signals X_ℓ are independent, conditional on any realization of θ. Conditional on $\theta \in \{-1,1\}$, two signals (X_s, X_ℓ) have the joint distribution

$$\begin{array}{c|c|c} & X_\ell = \theta & X_\ell \neq \theta \\ \hline X_s = \theta & (1-\varepsilon)^2 & (1-\varepsilon)\varepsilon \\ \hline X_s \neq \theta & (1-\varepsilon)\varepsilon & \varepsilon^2 \end{array} \tag{4.11}$$

Given the information structure, when two players observe the same signal, their signal's realization is the same with probability one. When two players do not observe the same signal, their signal's realization is the same with probability $1 - 2\varepsilon(1-\varepsilon)$.

4.6.2.1 Supply Chains

Suppose there is a supply chain with 1 manufacturer (denoted M) and 2 retailers (denoted R_1 and R_2). The manufacturer chooses the wholesale price w, and retailer R_i, the markup p_i over the wholesale price. Final prices are $w + p_1$ and $w + p_2$. The demand for retailer R_i is

$$Q_i(p_i, p_j, w; \theta) = A + b_i\theta + \lambda_i(w + p_j) - (w + p_i), \tag{4.12}$$

where $b_i > 0$, $A > b_i$ and $0 < \lambda_i < 1$ for $i = 1, 2$. All three players are uncertain on the intercept of the retailers' demand function θ and have access to the signals $(X_1, ..., X_L)$. Retailer R_i's profits function is given by

$$\Pi_{R_i}(p_i, p_j, w; \theta) = p_i \cdot (A + b_i\theta + \lambda_i(w + p_j) - (w + p_i)), \tag{4.13}$$

for $i = 1, 2$, and the manufacturer's profits function is given by

$$\Pi_M(w, p_i, p_j; \theta) = w \cdot (2A + (b_1 + b_2)\theta - (1 - \lambda_1)p_2 - (1 - \lambda_2)p_1 + w(\lambda_1 + \lambda_2 - 2)). \tag{4.14}$$

The vector encoding action complementarities is $c^{R_i} \equiv (c^{R_i}_{R_j}, c^{R_i}_{M}) = (1, -1)$ for retailer R_i, and $c^{M} \equiv (c^{M}_{R_1}, c^{M}_{R_2}) = (-1, -1)$ for the manufacturer.

It is natural to look for an equilibrium where (p_1, p_2, w) are strictly increasing in θ, so that the action strategies are strictly m-monotonic with $m \equiv (m_{R_1}, m_{R_2}, m_M) = (1, 1, 1)$. Then, assuming that m describes the monotonicity of equilibrium action strategies, Theorem 4.4 implies that the players' dependence preferences are obtained by combining the monotonicity and the complementarity vectors, such that $d^{R_i} \equiv (d^{R_i}_{R_j}, d^{R_i}_{M}) = (1, -1)$ for retailer R_i and $d^{M} \equiv (d^{M}_{R_1}, d^{M}_{R_2}) = (-1, -1)$ for the manufacturer. Hence, the retailers prefer to observe the same signal and the manufacturer a signal different from the one observed by the retailers.

4.6.2.2 Beauty Contests

Suppose a set I of players interacts in a beauty contest game. Two versions of the beauty contest model can be found in the literature. In both versions, each player $i \in I$ chooses an action and his payoff depends on the others' average action $\bar{a} = \frac{1}{N}\left(\sum_{j \in I \setminus \{i\}} a_j\right)$. Version 1 (as in Myatt and Wallace (2011)) assumes the payoff function

$$u_i(a_i, \bar{a}; \theta) = -(1-r)(a_i - \theta)^2 - r(a_i - \bar{a})^2,$$

while version 2 (as in Hellwig and Veldkamp (2009)) assumes the payoff function

$$u_i(a_i, \bar{a}; \theta) = -(a_i - (1-r)\theta - r\bar{a})^2,$$

where $r \in (-1, 1)$. In both versions, all players have positive complementarities in actions if $r > 0$ and negative complementarities in actions if $r < 0$, and they all have positive complementarities in state.

For both versions, given the information structure is fixed, a player's best response in action is $\alpha_i(x_i) = (1-r)\mathbb{E}(\Theta|x_i, X)) + r\mathbb{E}(\bar{a}|x_i, X)$.

If r is sufficiently small (such that $1 - r$ is big enough), then the best response will be increasing in the signal's realization for all players. Then, by Theorem 4.6, the situations where all players play action strategies that are strictly increasing in state and i) all players observe the same signal if $r > 0$, and ii) all players observe different signals if $r < 0$ are equilibria of the game with an endogenous information structure.

4.6.2.3 Technological Spillovers

Suppose a set I of players have the possibility to exert effort in developing a new technology. A player chooses his level of effort a_i to the development of this technology. The cost of effort a_i is ca_i^2, with $c > 0$, while the benefit for each individual is $\left((Ka_i + b\theta)\sum_{j\in I} a_j\right)$, with $b > 0$ so that the payoff of player i is

$$u_i(a;\theta) = \left((Ka_i + b\theta)\sum_{j\in I} a_j\right) - ca_i^2.$$

The effort exerted by the other players has a positive impact on player i's payoff. Hence, all players have positive complementarities in actions and positive complementarities in state. When all players observe the same signals, there exists an equilibrium where all players choose an effort level that is increasing in their signal if b is large enough. Then, Theorem 4.6 ensures that the situations where all players choose a strictly increasing action strategies and all choose to observe the same signals is still an equilibrium of the game with an endogenous information structure. Therefore, information acquisition on the state is suboptimal from a social viewpoint since acquired information will be homogenous.

4.6.2.4 Policy Choice in Federations

This example is adapted from Loeper (2011). Assume there is a set I of jurisdictions and that each of them needs to decide on a policy. A jurisdiction's policy choice is a_i, and its payoff is

$$u_i(a;\theta) = -(a_i - \beta_{i\theta}\theta)^2 - \sum_{j\neq i}\beta_{ij}(a_i - a_j)^2,$$

where $\beta_{i\theta}$ is jurisdiction i's alignment preference with the state θ and $\beta_{ij} \in (-1, 1)$ the coordination externality that jurisdiction j imposes on i. Jurisdiction i has positive (negative) complementarities in state when $\beta_{i\theta} > (<)\ 0$ and positive (negative) complementarity in actions with jurisdiction j when $\beta_{ij} > (<)\ 0$.

A player's action best response is

$$\alpha_i(x_i) = \frac{\beta_{i\theta}\mathbb{E}[\theta|x_i] + \sum_{j\neq i}\beta_{ij}\mathbb{E}[a_j|x_i]}{(1+\sum_{j\neq i}\beta_{ij})}.$$

When all jurisdictions have positive (negative) complementarity in state and positive complementarity in actions, such that the action complementarities are aligned with the state complementarities, Theorem 4.5 ensures that there exists an equilibrium where policy choices are increasing (decreasing) in the signal and all jurisdictions observe the same signal. In this type of equilibrium, the vector of policy choices would not be very informative on the fundamental θ, since all policies will be based on the same signal.

Consider next the case where all jurisdictions have negative (positive) complementarity in state and negative complementarity in actions. Suppose further that when it is exogenously determined that all jurisdictions observe different signals then, in this case, there exists an equilibrium where policy choices are decreasing (increasing) in the signal. Then, Theorem 4.6 implies that observing different signals is still an equilibrium in Γ, the game with endogenous information choice, when all jurisdictions have sufficiently negative (positive) complementarity in state.

4.6.2.5 Imperfect Competition

These examples are adapted from Jiménez-Martinez (2013).

Cournot: Consider a set I of firms interacting in an oligopoly and competing in Cournot. Let the aggregate demand function be given by $P(q_1,...,q_I;\theta) = A + b\theta - \delta(q_1 + ... + q_I)$, with $b > 0$, $\delta > 0$.

Each of the firms sets a quantity q_i, $i = 1,...,I$ assuming the marginal cost is constant and equals to $c_i > 0$ for each firm. Firm i's payoff is

$$u_i(q_i, q_{-i};\theta) = \Big(A + b\theta - \delta(q_1 + ... + q_I) - c_i\Big)q_i.$$

In this context, the firms have negative complementarities in actions and positive complementarities in state.

Thus Theorem 4.6 implies that when condition iv) holds, which is the case if b is large enough (and ε is small enough relative to the other parameters), there exists an equilibrium where all firms choose different signals and play a quantity strategy increasing in their signal. In particular, the price P is very informative on the state since it aggregates the information of N independent signals.

Bertrand: Consider next an oligopoly where two firms produce a differentiated product and compete in Bertrand. Each of the two firms $i = 1, 2$ sets a price p_i for its product and faces the linear demand $Q_i = b_0 + b_1\theta - \gamma(p_i - p_j)$ with $b_0 > 0$, $b_1 > 0$ and $\gamma > 0$. Each firm i has a constant marginal cost c_i.

Firm i's profits function is then

$$u_i(p_i, p_j; \theta) = \Big(b_0 + b_1\theta - \gamma(p_i - p_j)\Big)\Big(p_i - c_i\Big).$$

The firms have positive complementarities in actions and positive complementarities in states. Thus, since state and action complementarities reinforce each other, Theorem 4.5 implies that there exists an equilibrium where both firms choose the same (public) signal and choose a price strategy that is increasing in the signal. In this case, the quantity $Q_1 + Q_2$ only reflects the information contained in one signal and does not aggregate all the information potentially available.

4.7 Related Literature

We contribute to the applied probability literature on dependence orderings. We explain this contribution in section 4.5.1.2.

Regarding decentralized information acquisition, the most commonly studied framework is a two stage game where players start with a common prior on some unknown common value state that affects all players' payoffs.[25] In the first stage, each player makes an

[25] See Li, McKelvey and Page (1987), Vives (1988), Hellwig and Veldkamp (2009), Myatt and Wallace (2011), Szkup and Trevino (2014), Yang (2014), and many others. Veldkamp's monograph (2011) and Hellwig, Kohls and Veldkamp (2013) provide excellent surveys on the widely studied special case of the beauty contest games with a continuum of actions and players, quadratic payoffs and a Gaussian information structure, and their applications to macroeconomics and finance. Our paper covers a larger class of models, since we do not rely on specific functional forms and allow for a finite number of players.

information choice (for example, the precision of the signal he receives) that determines the information on the state that he has when entering the second stage. In the second stage, players simultaneously choose an action. Two different extensive forms have been considered, depending on whether the choices made at the first stage are observed or not. In some models, the acquisition is publicly observed. The game is then an extensive form game where each profile of information acquisition choices defines a subgame, and in each subgame, the information structure is common knowledge: in these games, acquisition is *overt.* In other models, the choices of the players in the first period are not observed before actions are taken. Acquisition is then *covert.* A game where information acquisition is covert is essentially static, as it is equivalent to one where all players simultaneously choose both their information and a commitment to an action strategy that maps the signal they will observe to the action they choose. The difference between overt and covert information acquisition is in the way a deviation in the first stage is treated: Under overt acquisition, a deviation on information choice is commonly observed, and the information structure is common knowledge in the second stage subgame following the deviation; Under covert acquisition, players form a belief of what the information structure is in the second stage, and this belief is correct in equilibrium. But whenever a player deviates, all other players' belief on the information structure is incorrect. It should be noted that in games with a continuum of players (Hellwig and Veldkamp, 2009; Myatt and Wallace, 2011; Szkup and Trevino, 2014), where players' payoffs only depend on the statistical distribution of the other players' actions, the two forms of acquisition are equivalent. Thus, there is no need to make a distinction in this case. The distinction matters only for games with finitely many players. In this paper, we derive results that apply to games with covert acquisition and finitely many players, and to games with covert or overt acquisition and a continuum of players. The case where acquisition is overt and the number of players is finite is also considered but only in section 4.7.3.

4.7.1 The Motive Inheritance Result

The main focus in the literature has been on the player's choice of amount of information (their signal's precision for signals that are independent conditional on the state), and on the acquisition of private information. A central question in this context is whether the players' amount of information acquisition are complements, substitutes, or neither complements nor substitutes. With finitely many players, the question is meaningful only when acquisition is overt. With a continuum of players, the question is meaningful for both overt and covert acquisition, since the two are in this case equivalent. Li, McKelvey and Page (1987) study a Cournot market with finitely many firms and overt acquisition. The unknown common value state is the demand intercept and the information structure satisfies certain conditions. Actions are substitutes and they find that the precision levels of the private information acquired in the first stage are substitutes as well.[26] Vives (1988) obtains a similar result in the case of a continuum of players. Assuming as well a continuum of players, Hellwig and Veldkamp (2009) obtain a similar result in the context of a beauty contest game, where actions can be either substitutes or complements. They find that when actions are substitutes, acquisition levels are substitutes and when actions are complements, acquisition levels are complements: the strategic motive in actions is *inherited* by the acquisition game. All of these papers assume an unbounded continuum of actions (the real line), quadratic payoffs, and a Gaussian information structure.

In spite of the large number of contexts where the inheritance result is confirmed, it does not generalize to the larger class of all games with strategic complementarities or substitutabilities. In particular, the unbounded continuum of actions, the continuum of players and a Gaussian information structure seem to be crucial for the result. Even with unbounded actions, quadratic payoffs and a Gaussian information structure, but only two players (as in the case of a differentiated Bertrand game, which the author uses as an example), Jimenez-Martinez (2013) shows that it only holds for some parameters: when the complementarity in actions is strong, levels of acquired precision may be substitutes. And even with a continuum of players, quadratic payoffs, a Gaussian information structure, but binary actions, i.e. a global game, Szkup and Trevino (2014) present a model where the actions are complements but the acquired precision levels are not.

In contrast with this literature, we do not allow players to choose how much information they acquire. We hold the amount of information fixed. Instead, we let them choose whether the information they acquire is private or public. More generally, we allow players

[26] Hwang (1993) exploits this result in a duopoly to derive various comparative statics results. Hwang (1995) studies a similar model but focuses on payoff comparisons between different market structures and different ways in which the levels of information precision of the firms is set. Bergemann, Shi and Välimäki (2009) obtain conditions under which information acquisition levels are substitutes or complements, in a VCG auction with interdependent valuations. Their setting differs from the common value models listed here in several ways.

to choose the level of conditional dependence between their signals. We show that another type of inheritance result holds: complementarity in actions implies a preference for positive informational dependence, and substitutability in actions implies a preference for informational independence. But unlike the motive inheritance result on precision, our dependence inheritance results hold for all games where actions are strategic complements or substitutes and do not rely on specific functional forms, provided that the second stage strategies are monotonic (which is a possibility in some games, and an implication of Nash-Bayesian equilibrium in a subset of these games).

4.7.2 Public and Private Information

The issue of the role of public and private information is a central one in the entire literature on endogenous information structures. Morris and Shin (2002) for example, show that in a beauty contest game with a continuum of players, when the planner (the central bank) increases the precision of public information, it can be detrimental to welfare, because players rely less on their private information.

In the context of information acquisition, Hellwig and Veldkamp (2009) and Myatt and Wallace (2011) let players choose whether the information is private and potentially public. They provide models where public information is an equilibrium outcome of players choosing to observe the same potentially public signals. Thus in their models, unlike Morris and Shin (2002), public information is not provided by an external third party, but is the result of the market forces themselves. We follow up on this idea, and go one step further. While their model has private and potentially public signals (public signals are the potentially public ones that all players chose to observe), in a version of our model, all signals are potentially public: a private signal is one that only one player chose to observe, while a public signal is one that all players chose to observe.

Like Morris and Shin (2002), Hellwig and Veldkamp (2009) are interested in the marginal value of additional public information compared to an initial situation. But because no information is intrinsically public or private, what they really look at is the marginally value of additional potentially public information. They make the important observation that marginal value of acquiring *more* potentially public information is kinked at some profiles that they call symmetric. At a symmetric profile, defined as one where all players observe the same potentially public signals, if a player deviates and observes one more potentially public signal, he obtains additional information that in effect is private, since nobody else observes it. If he instead drops one of his potential signals, he decreases his own access to public information. This asymmetry and discontinuity causes multiple equilibria that differ in the level of public information. In Myatt and Wallace (2011), public information obtains when all players pay a substantial amount of attention to the same signal. An implication of this assumption is that players who hold public information are necessarily well informed players. In both of these two papers, the problem of the division of information between private and public (and everything in between) is intrinsically intertwined with the more widely studied issue of the amount of information that the players acquire. In Hellwig and Veldkamp (2009), it is because of the question they choose to ask, and in Myatt and Wallace (2011), it is in the way they define public information.

In contrast, we choose to completely disentangle the two issues. At the risk of making the model seem less realistic (because in practice, economic agents often face the choice of how much information to acquire), we hold the amount of information fixed by assuming that all signals an agent can choose to observe are equally informative of the unknown state: they all have the same joint marginal distribution with the state. By doing so, we isolate the issue of the partition of the information structure between public, private and neither private nor public information, from the issue of the amount of information. Doing so enables us to identify a robust force and to obtain general results that hold for a large class of games, not only the Gaussian-quadratic model with a continuum of actions and players. As we argued earlier, no such result holds when the issue of the amount of information is not excluded, even when only private information can be acquired.

Our assumption that players are restricted in the amount of information they acquire (formally, the joint marginal distribution between their own signal and the state is fixed, no matter what signal they choose) can be thought of as a form of rational inattention. Players are limited in how much information they can acquire, (Sims, 2003, 2005, 2006), and thus face a choice of what to observe.

4.7.3 Inefficiency of Equilibrium Under Hidden Information Acquisition

A number of papers are dedicated to the analysis of inefficiencies in the collection of information and in the use of that information when the information structure is exogenously given. Angeletos and Pavan (2007) for instance, study a model with a continuum of players, quadratic payoff and a Gaussian information structure, where each player observe a private and a public signals. By comparing the equilibrium use of information to an efficiency benchmark (the best society could achieve keeping information decentralized), they show that information use can be inefficient when the incentives to coordinate actions and the social value for coordination are different. The welfare impact depends on the degree of strategic interaction and on its nature (complementarity or substitutability).

Angeletos and Pavan (2007)'s finding is recurrent in the literature. Morris and Shin (2002) among others also show that an increase in the amount of public information can impair welfare. This, however, does not hold necessarily if information is a choice for the players. Chahrour (2012) proposes a model of endogenous information acquisition where public information can still have a detrimental effect. In the model, a central authority chooses both how many signals to divulge and their precisions. He finds that the authority always chooses the highest possible precision and releases a positive but finite number of signals. An important result is that too many signals can cause the players to decrease the amount of information they acquire which in turn decreases welfare.

Colombo and Femminis (2008, 2011), on the other hand, are examples where endogenizing the information structure makes additional public information beneficial for welfare. By allowing the players to choose the precision of their private signals once the central authority has announced the precision of the public signal, they show that the precisions of private and public signals are strategic substitutes. Moreover, if the cost of public information is lower than the cost of private information, then increasing the precision of the public information increases welfare. While Colombo and Femminis (2008, 2011) investigate the welfare implications of public information provision on incentives to acquire private information, Llosa Gonzalo and Venkateswaran (2012), by considering models different from the beauty-contest type, study how different links and externalities among players affect the acquisition process of private information.

Existing work allows the players to choose the level of information precision. In this paper, our approach was different. Indeed, we take the analysis in an other direction by keeping the amount of information fixed and focusing instead on information dependence. We show that covert information acquisition sometimes leads to inefficiencies when there are payoff externalities that are not reflected in the players' equilibrium choice. Interestingly, we show that these inefficiencies can sometimes be eliminated when information acquisition is overt.

Bibliography

Aguirre, I., Cowan, S., and Vickers, J. (2010). Monopoly Price Discrimination and Demand Curvature. *Amer. Econ. Rev.*, 100(4):1601–1615.

Amador, M. (2010). Learning from Prices: Public Communication and Welfare. *Journal of Political Economy*, 118:866–907.

Amador, M. (2012). Learning from Private and Public Observation of Others' Actions. *Journal of Economic Theory*, 147:910–940.

Amir, R. and Lazzati, N. (2014). Endogenous Information Acquisition in Bayesian Games with Strategic Complementarities. Working Paper.

Anderson, S. P. and Renault, R. (2006). Advertising Content. *American Economic Review*, 96(1):93–113.

Angeletos, G. and Pavan, A. (2007). Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142.

Armin, R. (2014). The Benefits of Miscommunication. Working Paper.

Armstrong, M. (2006). Recent Developments in the Economics of Price Discrimination. In Blundell, R., Newey, W., and Persson, W., editors, *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress of the Econometric Society*, volume 2. Cambridge University Press.

Asplund, M. and Friberg, R. (2000). The Law of One Price in Scandinavian Duty-Free Stores. SSE/EFI Working Paper no 351.

Athey, S. (2001). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69(4):861–889.

Athey, S. and Levin, J. (2001). The value of information in monotone decision problems. MIT Working Paper 98-24.

Bagwell, K. and Riordan, M. (1991). High and Declining Prices Signal Product Quality. *Amer. Econ. Rev.*, 81(1):224–239.

Bar-Isaac, H. (2003). Reputation and Survival: Learning in a Dynamic Signalling Model. *Review of Economic Studies*, 70:231–251.

Bergemann, D., Brooks, B., and Morris, S. (2015). The Limits of Price Discrimination. *Amer. Econ. Rev.*, 105(3):921–957.

Bergemann, D. and Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308.

Bergemann, D. and Pesendorfer, W. (2007). Information structure in optimal auctions. *Journal of Economic Theory*, 137:580–609.

Bergemann, D., Shi, X., and Välimäki, J. (2009). Information acquisition in interdependent value auctions. *Journal of the European Economic Association*, 7(1):61–89.

Besbes, O. and Scarsini, M. (2013). On Information Distortions in Online Ratings. Working Paper.

Bhalla, M. (2012). Dynamic Pricing under Social Learning with Strategic Consumers. Working Paper.

Bhalla, M. (2013). Waterfall versus Sprinkler Product Launch Strategy: Influencing the Herd. *The Journal of Industrial Economics*, 61:138–165.

Blackwell, D. (1953). Equivalent Comparisons of Experiments. *Annals of Mathematical Statistics*, 24:265–272.

Blume, A., Board, O., and Kawamura, K. (2007). Noisy Talk. *Theoretical Economics*, 2:395–440.

Bose, S., Orosel, G., Ottaviani, M., and Vesterlund, L. (2006). Dynamic Monopoly Pricing and Herding. *RAND Journal of Economics*, 37:910–928.

Bose, S., Orosel, G., Ottaviani, M., and Vesterlund, L. (2008). Monopoly Pricing in the Binary Herding Model. *Economic Theory*, 37:204–241.

Brown, J., Hossain, T., and Morgan, J. (2010). Shrouded Attributes and Information Suppression: Evidence from the Field. *The Quaterly Journal of Economics*, pages 859–876.

Calvo-Armengol, A. and de Marti Beltran, J. (2007). Communication Networks: Knowledge and Decisions. *The American Economic Review*, 97(2):86–91.

Calvo-Armengol, A. and de Marti Beltran, J. (2009). Information gathering in organizations: Equilibrium, welfare, and optimal network structure. *Journal of the European Economic Association*, 7(1):116–161.

Calvo-Armengol, A., de Marti Beltran, J., and Prat, A. (2009). Endogenous Communication in Complex Organizations. Working Paper.

Carlsson, H. and Dasgupta, S. (1997). Noise-Proof Equilibria in Two-Action Signaling Games. *Journal of Economic Theory*, 77(2):432–460.

Chahrour, R. (2014). Public communication and information acquisition. *American Economic Journal: Macroeconomics*, 6(3):73–101.

Che, Y.-K. and Hörner, J. (2014). Optimal Design for Social Learning. Working Paper.

Chen, C. (2009). A Puzzle or a Choice: Uniform Pricing for Motion Pictures at the Box. *Atl. Econ. J.*, 37(1):73–85.

Chen, Y. and Cui, T. (2013). The Benefit of Uniform Price for Branded Variants. *Market. Sci.*, 32(1):36–50.

Chevalier, J. A. and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43:345–354.

Chhetry, D., Sampson, A., and Kimeldorf, G. (1989). Concepts of Setwise Dependence. *Probability in the Engineering and Informational Sciences*, 3:367–380.

Chintagunta, P., Gopinath, S., and Venkataraman, S. (2010). The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science*, 29(5):944–957.

Christofides, T. and Vaggelatou, E. (2004). A connection between supermodular ordering and positive/negative association. *Journal of Multivariate Analysis*, 88(1):138–151.

Colombo, L. and Femminis, G. (2008). The Social Value of Public Information with Costly Information Acquisition. *Economic Letters*, 100:196–199.

Colombo, L., Femminis, G., and Pavan, A. (2014). Information Acquisition and Welfare. *Forthcoming in Review of Economic Studies.*

Cornand, C. and Heinemann, F. (2008). Optimal Degree of Public Information Dissemination. *The Economic Journal*, 118:718–742.

Cowan, S. (2013). Welfare-Increasing Thirs-Degree Price Discriminaion. University of Oxford, Economics Series Working Paper 652.

Crémer, J. (1982). A Simple Proof of Blackwell's "Comparison of Experiments" Theorem. *Journal of Economic Theory*, 27:239–443.

Daher, W., Mirman, L., and Santugini, M. (2012). Information in Cournot: Signaling with Incomplete Control. *Int. J. Ind. Organ.*, 30(4):361–370.

Dai, W., Jin, G. Z., Lee, J., and Luca, M. (2012). Optimal Aggregation of Consumer Ratings: An Application to Yelp.com. Working Paper.

Daughety, A. and Reinganum, J. (1995). Product Safety: Liability, R&D, and Signaling. *Amer. Econ. Rev.*, 85(5):1187–1206.

Daughety, A. and Reinganum, J. (2005). Secrecy and Safety. *Amer. Econ. Rev.*, 95(4):1074–1091.

Daughety, A. and Reinganum, J. (2007). Competition and Confidentiality: Signaling Quality in a Duopoly when there is Universal Private Information. *Games Econ. Behav.*, 58(1):94–120.

Daughety, A. and Reinganum, J. (2008a). Communicating Quality: A Unified Model of Disclosure and Signalling. *RAND J. Econ.*, 39(4):973–989.

Daughety, A. and Reinganum, J. (2008b). Imperfect Competition and Quality Signalling. *RAND J. Econ.*, 39(1):163–183.

de Haan, T., Offerman, T., and Sloof, R. (2011). Noisy Signaling: Theory and Experiment. *Games Econ. Behav.*, 73(2):402–428.

Dellarocas, C. (2005). Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard. *Information Systems Research*, 16(2):209–230.

Dellarocas, C. (2006). Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms. *Management Science*, 52:1577–1593.

Dellarocas, C. (2010). Designing Reputation Systems for the Social Web. Working Paper.

Dewan, T. and Myatt, D. (2008). The Qualities of Leadership: Direction, Communication and Obfuscation. *American Political Science Review*, 102(3):351–368.

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2014). Consumer Price Search and Platform Design in Internet Commerce. Working Paper.

Eliaz, K. and Spiegler, R. (2015). Search Design and Broad Matching. Working Paper.

Flores Vidotti, C. (2004). Drug information centers in developing countries and the promotion of rational use of drugs: A viewpoint about challenges and perspectives. *Int. Pharm. J.*, 18(1):21–23.

Fradkin, A. (2014). Search Frictions and the Design of Online Marketplaces. Working Paper.

Friberg, R. (2001). Two Monies, Two Markets?: Variability and the Option to Segment. *J. Int. Econ.*, 55(2):317–327.

Friberg, R. (2003). Common Currency, Common Market? *J. Eur. Econ. Assoc.*, 1(2-3):650–661.

Friberg, R. and Martensen, K. (2001). Endogenous Market Segmentation and the Law of One Price. Working Paper.

Fudenberg, D. and Tirole, J. (1986). A Signal-Jamming Theory of Predation. *RAND Journal of Economics*, 17:366–376.

Gallo, F. (2010). To Segment or Not to Segment Markets? A Note on the Profitability of Market Segmentation for an International Oligopoly. Working Paper.

Gendron-Saulnier, C. and Santugini, M. (2013). The Informational Benefit of Price Discrimination. Cahiers de Recherche 13-02, HEC Montréal, Institut d'économie appliquée.

Ghose, A., Ipeirotis, P. G., and Li, B. Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue. *Management Science*, page Forthcoming.

Goldstein, I., Ozdenoren, E., and Yuan, K. (2011). Learning and complementarities in speculative attacks. *The Review of Economic Studies*, 78(1):263–292.

Goltsman, M., Hörner, J., Pavlov, G., and Squintani, F. (2009). Mediation, Arbitration and Negotiation. *Journal of Economic Theory*, 144(4):1397–1420.

Gordon, S. and Nödelke, G. (2013). Figures of Speech and Informative Communication. Working Paper.

Grossman, S. (1976). On the Efficiency of Competitive Stock Markets where Traders Have Different Information. *J. Finance*, 31(2):573–585.

Grossman, S. (1978). Further Results on the Informational Efficiency of Competitive Stock Markets. *J. Econ. Theory*, 18(1):81–101.

Grossman, S. (1989). *The Informational Role of Prices*. MIT Press.

Grossman, S. and Stiglitz, J. (1980). On the Impossibility of Informationally Efficient Markets. *Amer. Econ. Rev.*, 70(3):393–408.

Hajaj, C., Hazon, N., Sarne, D., and Elmalech, A. (2013). Search more, disclose less. In *AAAI*.

Hajaj, C. and Sarne, D. (2014). Strategic information platforms: selective disclosure and the price of free. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 839–856. ACM.

Harbaugh, R. and Rasmussen, E. (2013). Coarse Grades: Informing the Public by Withholding Information. Working Paper.

Harrington Jr., J. E. (1986). Limit Pricing when the Potential Entrant is Uncertain of its Cost Function. *Econometrica*, 54(2):429–437.

Hauk, E. and Hurkens, S. (2001). Secret information acquisition in cournot markets. *Economic Theory*, 18(3):661–681.

Heinsalu, S. (2014). Noisy Signalling Over Time. Working Paper.

Hellwig, C., Kohls, S., and Veldkamp, L. (2012). Information choice technologies. *American Economic Review: Papers and Proceedings*, 102(3):35–40.

Hellwig, C. and Veldkamp, L. (2009). Knowing what Others Know: Coordination Motives in Information Acquisition. *Review of Economic Studies*, 76:223–251.

Horton, J. (2014). Misdirected Search Effort in a Matching Market: Causes, Consequences and a Partial Solution. Working Paper.

Hu, T., A. M. and Scarsini, M. (2004). Some counterexamples in positive dependence. *Journal of Statistical Planning and Inference*, 1(153-158).

Hu, N., Pavlou, P., and Zhang, J. (2009). Overcoming the J-Shaped Distribution of Product Reviews. *Communications of the ACM*, 52(10).

Hwang, H. (1993). Optimal information acquisition for heterogenous duopoly firms. *Journal of Economic Theory*, 59(2):385–402.

Hwang, H. (1995). Information acquisition and relative efficiency of competitive, oligopoly and monopoly markets. *International Economic Review*, 36(2):325–340.

Ifrach, B., Maglaras, C., and Scarsini, M. (2014). Bayesian Social Learning from Consumer Reviews. Working Paper.

James, J. and Lawler, P. (2011). Optimal Policy Intervention and the Social Value of Public Information.

Janssen, M. and Roy, S. (2010). Signaling Quality through Prices in an Oligopoly. *Games Econ. Behav.*, 68(1):192–207.

Jeitschko, T. and Norman, H.-T. (2012). Signaling in Deterministic and Stochastic Settings. *J. Econ. Behav. Organ.*, 82(1):39–55.

Jewitt, I. (2007). Information order in decision and agency problems. Unpublished manuscript.

Jiang, B.-J. and Wang, B. (2008). Impact of Consumer Reviews and Ratings on Sales, Prices, and Profits: Theory and Evidence. *ICIS 2008 Proceedings*.

Jiménez-Martinez, A. (2011). Strategic Interactions in Information Decisions with a Finite Set of PLayers. Working Paper.

Jiménez-Martinez, A. (2013). Information acquisition interactions in two-player quadratic games. *International Journal of Game Theory*.

Jing, B. (2011). Seller Honesty and Product Line Pricing. *Quantitative Marketing and Economics*, 9:403–427.

Judd, K. and Riordan, M. (1994). Price and Quality in a New Product Monopoly. *Rev. Econ. Stud.*, 61(4):773–789.

Kamenica, E. and Gentzkow, M. (2011). Bayesian Persuasion. *American Economic Review*, 101:2590–2615.

Kihlstrom, R. and Mirman, L. (1975). Information and Market Equilibrium. *Bell. J. Econ.*, 6(1):357–376.

Kreps, D. and Porteus, E. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, 46(1):185–200.

Kwark, Y., Chen, J., and Raghunathan, S. Online Product Reviews: Implications for Retailers and Competing Manufacturer. *Information Systems Research*, page Forthcoming.

Kyle, A. (1985). Continuous Auctions and Insider Trading. *Econometrica*, 53(6):1315–1335.

Lehmann, E. (1955). Ordered Families of Distributions. *The Annals of Mathematical Statistics*, 26:399–419.

Levhari, D., Paroush, J., and Peleg, B. (1975). Efficiency Analysis for Multivariate Distributions. *Review of Economic Studies*, 42:87–91.

Li, L., McKelvey, R. D., and Page, T. (1987). Optimal research for cournot oligopolists. *Journal of Economic Theory*, 42(1):140 – 166.

Li, X. and Hitt, L. M. (2008). Self-Selection and Information Role of Online Product Reviews. *Information Systems Research*, 19(4):456–474.

Li, X. and Hitt, L. M. (2010). Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis. *MIS Quaterly*, 34:809–831.

Liu, T. and Schiraldi, P. (2012). New Product Launch: Herd Seeking or Herd Preventing? *Economic Theory*, 51:627–648.

Llosa Gonzalo, L. and Venkateswaran, V. (2012). Efficiency with Endogenous Information Choice. Working Paper.

Loeper, A. (2011). Coordination in heterogeneous federal systems. *Journal of Public Economics*, 95:900–912.

Luca, M. (2011). Reviews, Reputation and Revenue: The Case of Yelp.com. Working Paper.

Luca, M. and Smith, J. (2013). Salience in Quality Disclosure: Evidence from the U.S. News College Rankings. *Journal of Economics and Management Strategy*, 22:58–77.

Luca, M. and Zervas, G. (2013). Fake It Till You Make It: Reputation, Competition and Yelp Review Fraud. Working Paper.

Matthews, S. and Mirman, L. (1983). Equilibrium Limit Pricing: The Effects of Private Information and Stochastic Demand. *Econometrica*, 51(4):981–996.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *American Economic Review*, 104:2241–2255.

McMillan, R. (2007). Different Flavor, Same Price: The Puzzle of Uniform Pricing for Differentiated Products. Working Paper.

Meyer, M. and Strulovici, B. (2012). Increasing interdependence of multivariate distributions. *Journal of Economic Theory*, 147(4):1460–1489.

Meyer, M. and Strulovici, B. (2013). The supermodular stochastic ordering. Unpublished manuscript.

Mirman, L., L., S., and Schlee, E. (1994). Strategic Information Manipulation in Duopolies. *Journal of Economic Theory*, 62:363–384.

Mirman, L., Salgueiro, E., and Santugini, M. (2014a). Learning in a Perfectly Competitive Market. CIRPEE working paper 1423.

Mirman, L., Salgueiro, E., and Santugini, M. (2014b). Noisy Signaling in Monopoly. *Int. Rev. Econ. Finance*, 29(C):504–511.

Mirman, L., Salgueiro, E., and Santugini, M. (2015). Noisy Learning in a Competitive Market with Risk Aversion. CIRPEE working paper 1502.

Morris, S. and Shin, H. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *The American Economic Review*, 88(3):pp. 587–597.

Morris, S. and Shin, H. (2002). Social Value of Public Information. *American Economic Review*, 92(5):1521–1534.

Morris, S. and Shin, H. (2007). Optimal Communication. *Journal of the European Economic Association*, 5(2-3):594–602.

Morris, S., Shin, H., and Tong, H. (2006). Social Value of Public Information: Morris and Shin (2002) Is Actuallt Pro-Transparency, not Con.: Reply. *American Economic Review*, 96(1):453–455.

Myatt, D. and Wallace, C. (2011). Endogenous Information Acquisition in Coordination Games. *Review of Economic Studies*, 78(4):1–38.

Nahata, B., Ostaszewski, K., and Sahoo, P. (1990). Direction of Price Changes in Third-Degree Price Discrimination. *Amer. Econ. Rev.*, 80(5):1254–1258.

Nicollier, L. and Ottaviani, M. (2014). Production and Manipulation of Reviews. Workshop on Rating, Ranking, and Recommendation systems, Paris School of Economics.

Orbach, B. and Einav, L. (2007). Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry. *Int. Rev. Law Econ.*, 27(2):129–153.

Østerdal, L. (2010). The mass transfer approach to multivariate discrete first order stochastic dominance: Direct proof and implications. *Journal of Mathematical Economics*, 46:1222–1228.

Ott, M., Choi, Y., Cardie, C., and Hancock, J. T. (2011). Finding Deceptive Opinion Spam by Any Stretch of the Imagination. 49th Annual Meeting of the Association for Computational Linguistics.

Papanastasiou, Y., Bakshi, N., and Savva, N. (2014). Social Learning from Early Buyer Reviews: Implications for New Product Launch. Working Paper.

Papanastasiou, Y. and Savva, N. (2014). Dynamic Pricing in the Presence of Social Learning and Strategic Consumers. Working Paper.

Pigou, A. (1920). *The Economics of Welfare*. London: Macmillan.

Pope, D. G. (2009). Reacting to Rankings: Evidence from "America's Best Hospitals". *Journal of Health Economics*, 28:1154–1165.

Radner, R. (1962). Team decision problems. *The Annals of Mathematical Statistics*, 33(3):857–881.

Reny, P. (1999). On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Game. *Econometrica*, 67(5):1029–1056.

Reny, P. (2011). On the existence of monotone pure strategy equilibria in bayesian games. *Econometrica*, 79(2):499–553.

Richardson, M. and Stähler, F. (2013). On the "uniform pricing puzzle" in recorded music. ANU Working Papers in Economics and Econometrics no. 612.

Riordan, M. (1985). Imperfect Information and Dynamic Conjectural Variations. *RAND Journal of Economics*, 16:41–50.

Riordan, M. (1986). Monopolistic Competition with Experience Goods. *Quart. J. Econ.*, 101(2):265–279.

Russo, G. and Senatore, L. (2012). A note on contribution game with loss function. *Economic Letters*, 115:211–214.

Schmalensee, R. (1981). Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination. *Amer. Econ. Rev.*, 71(1):242–247.

Schofield, N. and Sened, I. (2002). Local nash equilibrium in multiparty politics. *Annals of Operation Research*, 109:193–211.

Sgroi, D. (2002). Optimizing Information in the Herd: guinea pigs, profits and welfare. *Games and Economic Behavior*, 39:137–166.

Shaked, M. and Shanthikumar, G. (2007). *Stochastic Orders*. Springer Series in Statistics. Springer.

Sims, C. (1998). Stickiness. *Carnegie-Rochester Conference Series on Public Policy*, 49(1):317–356.

Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Sims, C. (2005). Rational inattention: A research agenda. Discussion Paper Series 1: Economic Studies 2005, 34, Deutsche Bundesbank, Research Centre.

Sims, C. (2006). Rational inattention: Beyond the linear-quadratic case. *The American Economic Review*, 96(2):158–163.

Sonderegger, S. (2012). Centralized or Decentralized Information: Which Is Better For Providing Incentives? *Economic Inquiry*, 48(2):290–305.

Svensson, L. (2006). Social Value of Public Information: Comment: Morris and Shin (2002) Is Actuallt Pri-Transparency, not Con. *American Economic Review*, 96(1):448–452.

Szkup, M. and Trevino, I. (2014). Information acquisition and transparency in global games. Unpublished manuscript.

Taneva, I. (2014). Information design. Unpublished manuscript.

Tchen, A. (1980). Inequalities for distributions with given marginals. *The Annals of Probability*, 8(4):814–827.

Tirole, J. (1988). *The Theory of Industrial Organization*. MIT press.

Ui, T. and Yoshizawa, Y. (2013). Characterizing social value of information. Working Paper.

Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information Acquisition and Under-Diversification. Working Paper.

Van Zandt, T. and Vives, X. (2007). Monotone equilibria in bayesian games of strategic complementarities. *Journal of Economic Theory*, 134(1):339–360.

Varian, H. (1985). Price Discrimination and Social Welfare. *Amer. Econ. Rev.*, 75(4):970–875.

Veldkamp, L. (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.

Vives, X. (1988). Aggregation of information in large cournot markets. *Econometrica*, 56(4):851–876.

Vives, X. (2008). *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press.

Vives, X. (2011). Endogenous Public Information and Welfare. Working Paper.

Wolinsky, A. (1983). Prices as Signals of Product Quality. *Rev. Econ. Stud.*, 50(4):647–658.

Yang, M. (2014). Coordination with flexible information acquisition. Unpublished manuscript.

Yang, Y., Sun, Y. L., Ren, J., and Yang, Q. (2007). Building Trust in Online Rating Systems Through Signal Modeling. *ICDCSW '07: Proceedings of the 27th International Conference on Distributed Computing Systems Workshops*.

Yu, M., Debo, L., and Kapuscinski, R. (2013). Strategic Waiting for Consumer-Generated Quality Information: Dynamic Pricing of New Experience Goods. Working Paper.

Appendix A

Appendix to Chapter 2

A.1 Equilibrium Definition for the Game with Endogenous Design

The game with an endogenous design is a Bayesian game in which the state of the world is either θ_H or θ_L, i.e., the possible product qualities. Only the seller knows the product quality. Consumers form beliefs on product quality given the signal realization and the strategies played by the seller and the platform.

A perfect Bayesian equilibrium (PBE) specifies the platform design choice $(\mathcal{M}, S)$, the seller's manipulation strategy $m(\cdot; (\mathcal{M}, S))$ for each possible $(\mathcal{M}, S)$, and consumers' threshold $\hat{v}(s; (\mathcal{M}, S))$ for each $s \in \mathcal{M}$ for each possible design $(\mathcal{M}, S)$. In equilibrium, it is required that the seller maximizes expected profits, consumers maximize their expected utility and the platform maximizes revenues conditional on a business model. Definition A.1 gives the requirements a tuple $\{(\mathcal{M}^*, S^*), m^*(\cdot; (\mathcal{M}, S)), \hat{v}^*(\cdot; (\mathcal{M}, S))\}$ must meet to form a PBE.

Definition A.1. *Let $y \in \{transactional,\ non\text{-}transacational\}$ be the type of the platform as determined by its business model. Then, the tuple $\{(\mathcal{M}^*, S^*), m^*(\cdot; (\mathcal{M}, S)), \hat{v}^*(\cdot; (\mathcal{M}, S))\}$ is a PBE if*

1. *at date 2, given $m^*(\cdot; (\mathcal{M}, S))$ and for all $(\mathcal{M}, S)$*
$$\hat{v}^*(s, m^*) \in \inf\{v \in [0,1] : \mathbb{E}[u_i(v, S, m^*)|s] \geq 0\} \quad \forall s \in \mathcal{M}; \tag{A.1}$$

2. *at date 1, for θ_x with $x \in \{H, L\}$, given $\hat{v}^*(\cdot; (\mathcal{M}, S))$ and for all $(\mathcal{M}, S)$,*
$$\mathbb{E}[1 - \hat{v}^*(\tilde{s}(m^*), m^*)] - c_x m^*(\theta_x; S)^2 \geq \mathbb{E}[1 - \hat{v}^*(\tilde{s}(m), m^*)] - c_x m^2 \tag{A.2}$$
for all $m \geq 0$, where $1 - \hat{v}^*(s, m^*)$ is the demand for the seller's product given signal realization $s \in \mathcal{M}$;

3. *at date 0, given $\hat{v}^*(\cdot; (\mathcal{M}, S))$ and $m^*(\cdot; (\mathcal{M}, S))$*
$$\mathbb{E}\left[Rev^y(\hat{v}^*(\tilde{s}, m^*), \mathcal{M}^*, S^*)\right] \geq \mathbb{E}\left[Rev^y(\hat{v}^*(\tilde{s}, m^*), \mathcal{M}, S)\right] \tag{A.3}$$
for all $(\mathcal{M}, S) \in \Big\{\Big(supp(\alpha^m), S_c\Big), \Big(\{$👎, 👍$\}, S_b(\cdot\ ;\ T), T \geq 0\Big)\Big\}$, where $Rev^y(\cdot)$ is the platform's revenues fonction;

4. *posterior beliefs are computed using Bayes' rule, whenever possible.*

A.2 The Effects of Manipulation on Consumers' Posteriors Beliefs

Effect of m on $\tilde{q}_c(\alpha^m, m)$

When the platform is using the continuous design, the posterior beliefs take a value in $\{0, q, 1\}$ (cf. Lemma 2.2), and this does not change when the seller manipulates the reviews. What changes is the probability of observing a realization of the platform's signal that leads consumers to form each level of posterior beliefs. Therefore, one could say that manipulation impacts the distribution of the posterior beliefs, but not the support of this distribution, i.e., only $\mathbb{P}(q_c(\alpha^m, m) = y)$ for $y \in \{0, q, 1\}$ changes.

Effect of m on $\tilde{q}_b(s, m; T)$

When the platform is rather using the binary design with a threshold T, manipulation affects both the distribution of the posterior beliefs and the support of this distribution. Not only do the respective probabilities that the posterior beliefs are $q_b(\text{👍}, m; T)$ or $q_b(\text{👎}, m; T)$ change, so do the value of $q_b(\text{👍}, m; T)$ and $q_b(\text{👎}, m; T)$. More specifically, the value of $q_b(\text{👍}, m; T)$ and $q_b(\text{👎}, m; T)$ change with m_H and m_L according to Lemma A.1.

Lemma A.1. *For all manipulation pairs* (m_L, m_H)*,*

$$\frac{\partial q_b(\text{👍}, m; T)}{\partial m_H} \geq 0 \quad ; \quad \frac{\partial q_b(\text{👎}, m; T)}{\partial m_H} \leq 0, \tag{A.4}$$

and

$$\frac{\partial q_b(\text{👍}, m; T)}{\partial m_L} \leq 0 \quad ; \quad \frac{\partial q_b(\text{👎}, m; T)}{\partial m_L} \geq 0. \tag{A.5}$$

Proof. Fix the pair of manipulation levels to (m_H, m_L). For $x \in H, L$, let

$$K_x(m_x, T) = \mathbb{P}r(s = \text{👎} | \theta_x, m_x) = \min\left\{\max\left\{\frac{T - \theta_x - m_x}{2b}, 0\right\}, 1\right\}, \tag{A.6}$$

be the probability that the platform signal be 👎 when quality is θ_x and manipulation m_x. Then, $\partial K_x(m_x, T)/\partial m_x \leq 0$.

According to Bayes' rule,

$$q_b(\text{👍}, m; T) = \frac{q}{q + (1-q)\left(\frac{1 - K_L(m_L, T)}{1 - K_H(m_H, T)}\right)} \tag{A.7}$$

and

$$q_b(\text{👎}, m; T) = \frac{q}{q + (1-q)\left(\frac{K_L(m_L,T)}{K_H(m_H,T)}\right)}. \tag{A.8}$$

Therefore,

$$\frac{\partial q_b(\text{👍}, m; T)}{\partial m_H} = \frac{\partial q_b(\text{👍}, m; T)}{\partial K_H(m_H,T)} \frac{\partial K_H(m_H,T)}{\partial m_H} \geq 0 \tag{A.9}$$

$$\frac{\partial q_b(\text{👍}, m; T)}{\partial m_L} = \frac{\partial q_b(\text{👍}, m; T)}{\partial K_L(m_L,T)} \frac{\partial K_L(m_L,T)}{\partial m_L} \leq 0. \tag{A.10}$$

The results for the effect of m_H and m_L on $q_b(\text{👎}, m; T)$ follow similarly. □

All else equal, the higher is the level of manipulation of the high-quality seller, the higher are the beliefs that the product has a high quality upon seeing the realization 👍. The reason is because it is now more likely that $\alpha^m(\theta_H) \geq T$ and so more likely that the signal 👍 is published when product quality is high. Incidentally, it is also the case that it is less likely that $\alpha^m(\theta_H) < T$, which lowers the beliefs that the product has a high quality upon seeing the signal realization 👎. The intuition is the reverse for the impact of the low-quality seller's manipulation.

A.3 Proofs

Lemma 2.1: Assume that the platform design is $(\mathcal{M}, S)$ and the consumers' conjecture on the seller's manipulation strategy is the function $m^e(\cdot)$. Assume further that upon observing signal $s \in \mathcal{M}$, $q(s, m^e; (\mathcal{M}, S))$ is the posterior beliefs updated through Bayes' rule. At date 2, consumer with type v_i decides to buy the good if and only if $\mathbb{E}[\tilde{u}_i|v_i, s] \geq 0$. Given the intrinsic quality's binary structure, a consumer buys if and only if

$$\begin{aligned}
\mathbb{E}[\tilde{u}_i|v_i,s] &= q(s,m^e;(\mathcal{M},S))\left(\mathbb{P}(u_i=1|\theta_H) - \mathbb{P}(u_i=-1|\theta_H)\right) \\
&\quad +(1-q(s,m^e;(\mathcal{M},S)))\left(\mathbb{P}(u_i=1|\theta_L) - \mathbb{P}(u_i=-1|\theta_L)\right) \qquad \text{(A.11)} \\
&= q(s,m^e;(\mathcal{M},S))\left(\theta_H v_i + b - (1-\theta_H v_i - b)\right) \\
&\quad +(1-q(s,m^e;(\mathcal{M},S)))\left(\theta_L v_i + b - (1-\theta_L v_i - b)\right) \qquad \text{(A.12)} \\
&= 2v_i(q(s,m^e;(\mathcal{M},S))\theta_H + (1-q(s,m^e;(\mathcal{M},S)))\theta_L) + 2b - 1 \geq 0 \qquad \text{(A.13)}
\end{aligned}$$

The condition given in (A.13) can be equivalently written as

$$v_i \geq \hat{v}(s, m^e; (\mathcal{M}, S)) \equiv \frac{1-2b}{2(\theta_L + (\theta_H - \theta_L) q(s, m^e; (\mathcal{M}, S)))}. \tag{A.14}$$

I refer to $\hat{v}(s, m^e; (\mathcal{M}, S))$ as the consumers's threshold given they received the signal s. That is, all consumers with a type above $\hat{v}(s, m^e; (\mathcal{M}, S))$ will buy one unit of the good at date 2. This implies that the seller's aggregate demand at date 1 is $D(q(s, m^e; (\mathcal{M}, S))) = \mathbb{P}[v_i \geq \hat{v}(s, m^e; (\mathcal{M}, S))] = 1 - \hat{v}(s, m^e; (\mathcal{M}, S))$.

Note that no other strategy can be optimal for consumers.

Lemma 2.2: The result follows from Bayes' rule. Specifically, upon seeing $\alpha(\cdot) \in [\theta_L, \theta_H)$, the beliefs are

$$\begin{aligned} q_c(\alpha) &= \frac{\mathbb{P}(\alpha\,(\theta_H) \in [\theta_L, \theta_H))\,\mathbb{P}(\theta = \theta_H)}{\mathbb{P}(\alpha(\theta_H) \in [\theta_L, \theta_H))\mathbb{P}(\theta = \theta_H) + \mathbb{P}(\alpha(\theta_L) \in [\theta_L, \theta_H))\mathbb{P}(\theta = \theta_L)} && \text{(A.15)} \\ &= \frac{0 \cdot q}{0 \cdot q + (\theta_H - \theta_L) \cdot (1-q)/2b} && \text{(A.16)} \\ &= 0, && \text{(A.17)} \end{aligned}$$

for $\alpha(\cdot) \in [\theta_H, \theta_L + 2b]$,

$$\begin{aligned} q_c(\alpha) &= \frac{\mathbb{P}(\alpha\,(\theta_H) \in [\theta_H, \theta_L + 2b])\,\mathbb{P}(\theta = \theta_H)}{\mathbb{P}(\alpha(\theta_H) \in [\theta_H, \theta_L + 2b])\mathbb{P}(\theta = \theta_H) + \mathbb{P}(\alpha(\theta_L) \in [\theta_H, \theta_L + 2b])\mathbb{P}(\theta = \theta_L)} && \text{(A.18)} \\ &= \frac{q \cdot (\theta_L + 2b - \theta_H)/2b}{q \cdot (\theta_L + 2b - \theta_H)/2b + (1-q) \cdot (\theta_L + 2b - \theta_H)/2b} && \text{(A.19)} \\ &= q, && \text{(A.20)} \end{aligned}$$

and for $\alpha(\cdot) \in (\theta_L + 2b, \theta_H + 2b]$,

$$\begin{aligned} q_c(\alpha) &= \frac{\mathbb{P}(\alpha\,(\theta_H) \in (\theta_L + 2b, \theta_H + 2b])\,\mathbb{P}(\theta = \theta_H)}{\mathbb{P}(\alpha(\theta_H) \in (\theta_L + 2b, \theta_H + 2b])\mathbb{P}(\theta = \theta_H) + \mathbb{P}(\alpha(\theta_L) \in (\theta_L + 2b, \theta_H + 2b])\mathbb{P}(\theta = \theta_L)} && \text{(A.21)} \\ &= \frac{q \cdot (\theta_H - \theta_L)/2b}{q \cdot (\theta_H - \theta_L)/2b + (1-q) \cdot 0} && \text{(A.22)} \\ &= 1. && \text{(A.23)} \end{aligned}$$

Lemma 2.3: For $x \in \{H, L\}$, let $G_x(T)$ be the probability that the platform publishes the signal 👎 when the seller's quality is θ_x and the threshold is T, that is

$$G_x(T) = \min\left\{\max\left\{\frac{T-\theta_x}{2b}, 0\right\}, 1\right\}. \tag{A.24}$$

Depending on the value of T, $G_x(T)$ reduces to

for $T \in [\cdot, \cdot]$	$[0, \theta_L]$	$[\theta_L, \theta_H]$	$[\theta_H, \theta_L + 2b]$	$[\theta_L + 2b, \theta_H + 2b]$	$[\theta_H + 2b, \infty)$
$G_H(T)$	0	0	$(T-\theta_H)/2b$	$(T-\theta_H)/2b$	1
$G_L(T)$	0	$(T-\theta_L)/2b$	$(T-\theta_L)/2b$	1	1

Then, from Bayes' rule,

$$q_b(👍; T) = \frac{q(1-G_H(T))}{q(1-G_H(T)) + (1-q)(1-G_L(T))}, \tag{A.25}$$

and

$$q_b(👎; T) = \frac{qG_H(T)}{qG_H(T) + (1-q)G_L(T)}. \tag{A.26}$$

Then, we have

for $T \in [\cdot, \cdot]$	$[0, \theta_L]$	$[\theta_L, \theta_H]$	$[\theta_H, \theta_L + 2b]$	$[\theta_L + 2b, \theta_H + 2b]$	$[\theta_H + 2b, \infty)$
$q_b(👍; T)$	q	$\frac{q}{q+(1-q)\left(\frac{\theta_L+2b-T}{\theta_H+2b-T}\right)}$	$\frac{q}{q+(1-q)\left(\frac{\theta_L+2b-T}{\theta_H+2b-T}\right)}$	1	$\circ$
$q_b(👎; T)$	$\circ$	0	$\frac{q}{q+(1-q)\left(\frac{T-\theta_L}{T-\theta_H}\right)}$	$\frac{q}{q+(1-q)\left(\frac{T-\theta_L}{T-\theta_H}\right)}$	q

where $\circ$ denotes that the signal s occurs with probability zero.

Since $G_L(T) \geq G_H(T)$ for all T, then $q \leq q_b(👍; T) \leq 1$ and $0 \leq q_b(👎; T) \leq q$.

Proposition 2.1: The proof follows directly from Lemmas 2.2 and 2.3.

Proposition 2.2: By garbling of information, I mean applying a stochastic map on a signal to create a new one. An information structure σ' is a garbled version of the information structure σ if $\sigma' = z \cdot \sigma$ for some stochastic map z. Blackwell (1953) defined one structure σ to be sufficient to another structure σ', if σ' is a garbled version of σ.

The binary design's signal can be represented by the stochastic matrix $Q(T)$ where

$$Q(T) = \begin{bmatrix} \mathbb{P}(👍|\theta_H, T), & \mathbb{P}(👎|\theta_H, T) \\ \mathbb{P}(👍|\theta_L, T), & \mathbb{P}(👎|\theta_L, T) \end{bmatrix} = \begin{bmatrix} 1 - G_H(T), & G_H(T) \\ 1 - G_L(T), & G_L(T) \end{bmatrix}. \tag{A.27}$$

The continuous design's signal can represented by the stochastic matrix P

$$\begin{aligned} P &= \begin{bmatrix} \mathbb{P}(\alpha(\theta_H) > \theta_L + 2b) & \mathbb{P}(\theta_H \leq \alpha(\theta_H) \leq \theta_L + 2b) & \mathbb{P}(\alpha(\theta_H) < \theta_H) \\ \mathbb{P}(\alpha(\theta_L) > \theta_L + 2b) & \mathbb{P}(\theta_H \leq \alpha(\theta_L) \leq \theta_L + 2b) & \mathbb{P}(\alpha(\theta_L) < \theta_H) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\theta_H - \theta_L}{2b} & \frac{\theta_L + 2b - \theta_H}{2b} & 0 \\ 0 & \frac{\theta_L + 2b - \theta_H}{2b} & \frac{\theta_H - \theta_L}{2b} \end{bmatrix} \end{aligned} \tag{A.28}$$

Then, given the model's specification, there exists a 3x2 non-negative stochastic matrix $W(T)$ such that $Q(T) = P \cdot W(T)$. The matrix W depends on the value of the threshold T in the following way,

$T \in$	$[0, \theta_L]$	$[\theta_L, \theta_H]$	$[\theta_H, \theta_L + 2b]$	$[\theta_L + 2b, \theta_H + 2b]$	$[\theta_H + 2b, \infty)$
W	$\begin{bmatrix} 0 & \frac{2b}{\theta_H - \theta_L} \\ 0 & 0 \\ 0 & \frac{2b}{\theta_H - \theta_L} \end{bmatrix}$	$\begin{bmatrix} \frac{2b - \theta_H - T}{\theta_H - \theta_L} & \frac{T - \theta_H}{\theta_H - \theta_L} \\ 0 & 0 \\ 0 & \frac{2b}{\theta_H - \theta_L} \end{bmatrix}$	$\begin{bmatrix} \frac{2b - \theta_H - T}{\theta_H - \theta_L} & \frac{T - \theta_H}{\theta_H - \theta_L} \\ 0 & 0 \\ \frac{2b - \theta_L - T}{\theta_H - \theta_L} & \frac{T - \theta_L}{\theta_H - \theta_L} \end{bmatrix}$	$\begin{bmatrix} \frac{2b}{\theta_H - \theta_L} & 0 \\ 0 & 0 \\ \frac{2b - \theta_L - T}{\theta_H - \theta_L} & \frac{T - \theta_L}{\theta_H - \theta_L} \end{bmatrix}$	$\begin{bmatrix} \frac{2b}{\theta_H - \theta_L} & 0 \\ 0 & 0 \\ \frac{2b}{\theta_H - \theta_L} & 0 \end{bmatrix}$

(A.29)

Lemma 2.4: First, consider the continuous signal. Without manipulation, it is the case that for $s' > s$, $q_c(s, m^*) \leq q_c(s', m^*)$ (cf. Lemma 2.2). With manipulation, suppose on the contrary that $\exists\ s' > s$ with $\mathbb{P}(s'|m_L^*, m_H^*)\mathbb{P}(s|m_L^*, m_H^*) > 0$ such that $q_c(s) > q_c(s')$. Given Bayes' rule, this is possible only if $m_L > m_H \geq 0$. In this case, there exists $\varepsilon > 0$ such that $s < s' - \varepsilon < s'$ and $m_L - \varepsilon > 0$. In this case, choosing $m_L - \varepsilon$ instead of m_L is a profitable deviation for the low-quality seller.

The case of the binary signal follows from a similar argument. Suppose $q_b(👎, m^*; T) > q_b(👍, m^*; T)$, once again, given Bayes' rule, this is possible only if $m_L^* > m_H^* \geq 0$. Given Lemma 2.2, demand when the signal is 👎 is going to be greater than demand when the signal is 👍. Then, decreasing m_L is a profitable deviation for the low-quality seller because manipulation cost are lower and the expected demand will be higher.

Proposition 2.3: Fix the manipulation levels to (m_H, m_L). When the platform uses the continuous design, the probability mass function of $\tilde{q}_c(\alpha^m, m)$ is $f_c^m(\cdot)$ with

$$f_c^m(q_c(\alpha^m, m)) = \begin{cases} qA(m_H, m_L) & \text{if } q_c(\alpha^m, m) = 1 \\ \\ 1 - A(m_H, m_L) & \text{if } q_c(\alpha^m, m) = q \\ \\ (1-q)A(m_H, m_L) & \text{if } q_c(\alpha^m, m) = 0 \end{cases} \tag{A.30}$$

where

$$A(m_H, m_L) = \begin{cases} \frac{\theta_H + m_H - \theta_L - m_L}{2b} & \text{if } \theta_L + m_L \leq \theta_H + m_H < \theta_L + 2b + m_L \\ \frac{\theta_L + m_L - \theta_H - m_H}{2b} & \text{if } \theta_H + m_H \leq \theta_L + m_L < \theta_H + 2b + m_H \\ 1 & \text{otherwise.} \end{cases} \tag{A.31}$$

Given the posterior beliefs' mean is q, the variance is

$$\begin{aligned} \mathbb{V}(\tilde{q}_c(\alpha^m, m)) &= (1-q)^2 qA(m_H, m_L) + (0-q)^2(1-q)A(m_H, m_L) \\ &= q(1-q)A(m_H, m_L). \end{aligned}$$

The results follow by taking the derivative of $A(m_H, m_L)$ with respect to m_H and m_L.

When the platform uses the binary design, the probability mass function of $\tilde{q}_b(s, m; T)$ is $f_b^m(\cdot)$ with

$$f_b^m(q_b(s, m; T)) = \begin{cases} q(1 - K_H(T)) + (1-q)(1 - K_L(T)) & \text{if } s = \text{👍} \\ \\ qK_H(T) + (1-q)K_L(T) & \text{if } s = \text{👎}, \end{cases} \tag{A.32}$$

where the expression for $K_H(T)$ and $K_L(T)$ are given by Equation (A.6)

The variance is

$$\begin{aligned} \mathbb{V}(\tilde{q}_b(s, m; T)) &= \left(q_b(\text{👍}, m; T) - q\right)^2 \left(q(1 - K_H(T)) + (1-q)(1 - K_L(T))\right) \quad \text{(A.33)} \\ &\quad + \left(q_b(\text{👎}, m; T) - q\right)^2 \left(qK_H(T) + (1-q)K_L(T)\right) \\ &= \frac{(1-q)^2 q^2 (K_L(T) - K_H(T))^2}{((1-q)K_L(T) + qK_H(T))(1 - (1-q)K_L(T) - qK_H(T))}. \quad \text{(A.34)} \end{aligned}$$

The results follow by taking the derivative of $\mathbb{V}(\tilde{q}_b(s,m;T))$ with respect to m_H and m_L taking in consideration that

$$\begin{cases} K_L(T) \geq K_H(T) & \text{if } \theta_L + m_L \leq \theta_H + m_H < \theta_L + 2b + m_L \\ K_L(T) \leq K_H(T) & \text{if } \theta_H + m_H \leq \theta_L + m_L < \theta_H + 2b + m_H. \end{cases} \tag{A.35}$$

Proposition 2.4: See Appendix A.4

Proposition 2.5: See Appendix A.5

Lemma 2.5: In the case of the continuous design, Appendix A.4.1 gives the conditions for full revelation to occur in equilibrium. In the case of the binary design, see Appendix A.5.2.7.

A set of conditions that is sufficient to insure that complete learning cannot occur in an equilibrium of the continuous design is

$$b > (\theta_H - \theta_L)/2 \tag{A.36}$$

$$\theta_H - \theta_L + \frac{D(1) - D(0)}{4bc} - \sqrt{\frac{D(1) - D(0)}{c}} < 0. \tag{A.37}$$

When conditions (A.36) and (A.37) are true, then conditions (A.64), (A.65) and (A.66) given in Appendix A.4.1 cannot hold.

Under conditions (A.36) and (A.37), then it is also impossible that complete learning be an equilibrium outcome when the platform is using the binary design. To see this, first note that the conditions given in Appendix A.5.2.7 are equivalent to the following set of conditions:

$$\text{set}_1 = \begin{cases} T \in \mathcal{T}_1 \\ c \leq \frac{D(1)-D(0)}{16b^2} \\ \sqrt{\frac{D(1)-D(0)}{c}} + \theta_L < T < \frac{D(1)-D(0)}{4bc} + \theta_L \\ T \leq \sqrt{\frac{D(1)-D(0)}{c}} + \theta_H \end{cases} \tag{A.38}$$

$$\text{set}_2 = \begin{cases} T \in \mathcal{T}_1 \\ c \leq \frac{D(1)-D(0)}{16b^2} \\ \frac{D(1)-D(0)}{4bc} + \theta_L \leq T \leq \sqrt{\frac{D(1)-D(0)}{c}} + \theta_H \end{cases} \tag{A.39}$$

$$\text{set}_3 = \begin{cases} T \in \mathcal{T}_1 \\ c > \frac{D(1)-D(0)}{16b^2} \\ \frac{D(1)-D(0)}{8bc} + \theta_L + 2b < T < \frac{D(1)-D(0)}{4bc} + \theta_H \end{cases} \tag{A.40}$$

$$\text{set}_4 = \begin{cases} T \in \mathcal{T}_2 \\ c > \frac{D(1)-D(0)}{16b^2} \\ \frac{D(1)-D(0)}{8bc} + \theta_L + 2b < T < \frac{D(1)-D(0)}{4bc} + \theta_H \end{cases} \tag{A.41}$$

$$\text{set}_5 = \begin{cases} T \in \mathcal{T}_2 \\ c \leq \frac{D(1)-D(0)}{16b^2} \\ \frac{D(1)-D(0)}{4bc} + \theta_L \leq T < \frac{D(1)-D(0)}{4bc} + \theta_H \end{cases} \tag{A.42}$$

$$\text{set}_6 = \begin{cases} T \in \mathcal{T}_2 \\ c < \frac{D(1)-D(0)}{16b^2} \\ \sqrt{\frac{D(1)-D(0)}{c}} + \theta_L \leq T < \frac{D(1)-D(0)}{4bc} + \theta_L \end{cases} \tag{A.43}$$

In other words, each set of conditions is sufficient for $(0, \overline{m}_H(T))$ to be an equilibrium such that complete learning occurs in the game. Together the conditions in set_1 to set_6 are necessary for complete learning to be possible with the binary design.

It is possible to show that conditions (A.36) and (A.37) are incompatible with the conditions in set_1 to set_6. Therefore, there cannot be complete learning with the binary design if complete learning is not an equilibrium outcome of the continuous design.

Proposition 2.6: To prove that cases a), b) and c) of the proposition are possible, it is sufficient to provide a set of parameters where those cases occur.

Note that for all cases, I impose the restriction $c_H = c_L = c$.

Case a: I construct an example where both designs produce an equilibrium outcome with complete learning.

When $(\theta_L, \theta_H, q, c, b) = (0.6,\ 0.65,\ 0.5,\ 0.75,\ 0.05)$, then there is an equilibrium with $(m_L^c, m_H^c) = (0,\ 0.23)$ because condition (A.65) in Appendix A.4.1 holds which insures that complete learning can occur in equilibrium with the continuous design.

For those parameters and fixing $T = 0.88$, there is an equilibrium with $(m_L^b, m_H^b) = (0,\ 0.23)$, because conditions (A.113), (A.114), (A.117) and (A.118) in Appendix A.5.2.7

hold, which means that complete learning also occur in equilibrium with the binary design.

Case b: I construct an example where the continuous design produces an equilibrium outcome with complete learning, whereas the binary design is associated with partial learning only.

When $(\theta_L, \theta_H, q, c, b) = (0.4,\ 0.48,\ 0.85,\ 0.75,\ 0.12)$, then there is an equilibrium with $(m_L^c, m_H^c) = (0,\ 0.38)$ because condition (A.64) in Appendix A.4.1 holds which insures that complete learning can occur in equilibrium with the continuous design.

For those parameters and fixing $T = 0.74 \in \mathcal{T}_1$, the set of conditions set$_1$ to set$_3$ (those only are relevant for $T \in \mathcal{T}_1$) of Lemma 2.5 are all violated, which means that complete learning cannot occur in equilibrium with the binary design. But, there exists an equilibrium that induces partial learning. The pair of manipulation efforts $(m_L^b, m_H^b) = (0,\ 0.068)$ is an equilibrium. This corresponds to a pair of type $(0, \hat{m}_H)$ given in Appendix A.5.2.4. To check that it is the case, it is sufficient to notice that $m_H^b \approx 0.068$ is one of the solution to (A.95) and that conditions (A.97), (A.98), (A.101) and (A.102) in Appendix A.5.2.4 hold.

Case c: I construct an example where the only equilibrium in pure strategies with the continuous design produces an outcome with no learning, whereas the binary design induces an equilibrium outcome with partial learning.

When $(\theta_L, \theta_H, q, c, b) = (0.5,\ 0.56,\ 0.65,\ 0.95,\ 0.1)$, conditions (A.64) to (A.66) in Appendix A.4.1 are violated which means that complete learning cannot occur in equilibrium with the continuous design.

The pair of manipulation efforts $(m_L^c, m_H^c) = (0.1,\ 0.04)$ is an equilibrium that induces no learning because condition (A.71) in Appendix A.4.2 holds with $\beta_1 = 0.8$ and $\beta_2 = 0.2$ as out-of-equilibrium beliefs.

For those parameters, by Lemma 2.5 we know that complete learning cannot occur in equilibrium with the binary design. Fixing $T = 0.8$, there exists an equilibrium with the manipulation efforts $(m_L^b, m_H^b) = (0,\ 0.087)$. This corresponds to a pair of type $(0, \hat{m}_H)$ given in Appendix A.5.2.4. Indeed, $m_H^b \approx 0.087$ is one of the solution to (A.95) and conditions (A.97), (A.98), (A.101) and (A.102) in Appendix A.5.2.4 hold.

Proposition 2.7: The proof is detailed for the binary design, the case for the continuous design follows similarly.

<u>Existence with the Binary Design</u>

Assume that the platform is using the binary design with a threshold $T \geq 0$. That is, given θ_x the platform observes $\alpha^m(\theta_x)$ and publishes a signal $s \in \{\text{👎}, \text{👍}\}$ following the rule

$$s(\alpha^m) = \begin{cases} \text{👍} & \text{if } \alpha^m(\theta_x) \geq T \\ \text{👎} & \text{if } \alpha^m(\theta_x) < T. \end{cases} \tag{A.44}$$

The game is $\Gamma = (Z_i, u_i)_{i \in I}$, where I is the set of players, Z_i the pure strategy set of player i and $u_i : Z \to \mathbb{R}$ with $Z = \times_{i \in I} Z_i$ is player i's payoff function.

More specifically,

- the set of players is $I = \{$seller of type H, seller of type L, a representative consumer$\}$
- the pure strategy set of a seller of type $x \in \{L, H\}$ is $m_x \in [0, T - \theta_x]$ (note that any $m_x > T - \theta_x$ is strictly dominated by $T - \theta_x$ such that it is not necessary to include those in the set of strategies)
- the pure strategy set of the representative consumer is $\Big(\hat{v}(\text{👎}), \hat{v}(\text{👍})\Big) \in [0,1]^2$.
- Payoff of the seller of type θ_x:

  Given $\Big(\hat{v}(s), (m_L, m_H)\Big)$, the payoff of the seller of type x is

$$\mathbb{P}(s = \text{👍} | \theta_x, m_x)(1 - \hat{v}(\text{👍})) + \mathbb{P}(s = \text{👎} | \theta_x, m_x)(1 - \hat{v}(\text{👎})) - c_x m_x^2 \tag{A.45}$$

  with

$$\mathbb{P}(s = \text{👎} | \theta_x, m_x) = \min\left\{\max\left\{\frac{T - \theta_x - m_x}{2b}, 0\right\}, 1\right\}, \tag{A.46}$$

$$\mathbb{P}(s = \text{👍} | \theta_x, m_x) = 1 - \mathbb{P}(s = \text{👎} | \theta_x, m_x). \tag{A.47}$$

The payoff function is continuous in $(m_x, m_{-x}, v(👎), v(👍))$, but, in general, it is not quasi-concave in m_x.

- Payoff of the representative consumer:

  The representative consumer is choosing $\hat{v}(s)$. Note that $1 - \hat{v}(s)$ can be interpreted as the volume of units that the representative consumer is buying. Utility from consuming a unit takes a value in $\{-1, 1\}$ with $\theta_x v_i + b$ the probability that a particular unit is associated to a utility of 1 when quality is θ_x.

  For $s \in \{👎, 👍\}$, given $\Big(\hat{v}(s), (m_L, m_H)\Big)$, the payoff of the representative consumer is

$$\int_{\hat{v}(s)}^{1} q_b(s, m_L, m_H; T)\Big[(1)\cdot(\theta_H v_i + b) + (-1)\cdot(1 - \theta_H v_i - b)\Big]\, \mathrm{d}v_i$$
$$+ \int_{\hat{v}(s)}^{1} (1 - q_b(s, m_L, m_H; T))\Big[(1)\cdot(\theta_L v_i + b) + (-1)\cdot(1 - \theta_L v_i - b)\Big]\, \mathrm{d}v_i \quad \text{(A.48)}$$
$$= \Big(\theta_L + (\theta_H - \theta_L) q_b(s, m_L, m_H; T)\Big)\Big(1 - \hat{v}(s)^2\Big) - (1 - 2b)(1 - \hat{v}(s)) \quad \text{(A.49)}$$

  where $q_b(s, m_L, m_H; T)$ is the probability that $\theta = \theta_H$ given that the signal s is observed

$$q_b(s, m_L, m_H; T) = \frac{q\mathbb{P}(s|\theta_H, m_H)}{q\mathbb{P}(s|\theta_H, m_H) + (1-q)\mathbb{P}(s|\theta_L, m_L)}. \quad \text{(A.50)}$$

  The representative consumer's payoff function is not necessarily continuous. For instance, at the profile $(m_L, m_H) = (T - \theta_L, T - \theta_H)$, we have $\mathbb{P}(👎|\theta_x, m_x) = 0$ for $x \in \{L, H\}$ such that $q_b(👎, m_L, m_H; T)$ is not defined.

  Equation (A.50) is undefined, and so is Equation (A.49), every time a signal $s \in \{👎, 👍\}$ is observed with probability 0 for both type of sellers. These cases require that out-of-equilibrium beliefs be specified.

Despite the discontinuity in the representative consumer's payoff and the non quasi-concavity of the seller's payoff, an equilibrium in mixed strategy can be proven to exist using Reny (1999).

Let Φ_i denotes the set of probability measures on the Borel subsets of Z_i. Then, extend u_i to $\Phi = \times_{i\in I}\Phi_i$ by defining $u_i(\phi) = \int_Z u_i(z)\mathrm{d}\phi$ for all $\phi \in \Phi$ and let $\overline{\Gamma} = (\Phi_i, u_i)_{i\in I}$ denote the mixed extension of Γ.

Let the index C refers to the representative consumer and $-C$ to the sellers. Assume that the sellers are playing the mixed strategy (ϕ_L, ϕ_H), then the representative consumer's payoff from a strategy v is now

$$\Big(\theta_L + (\theta_H - \theta_L)\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]\Big)\left(1 - v^2\right) - (1 - 2b)(1 - v) \tag{A.51}$$

where $\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)] = \int_{Z_{-C}} q_b(s, m_L, m_H; T)\mathrm{d}\phi_L(m_L)\mathrm{d}\phi_H(m_H)$.

The next definitions are needed in order to present Reny (1999)'s existence theorem.

Definition A.2. *(Reny, 1999) Player i can* ***secure a payoff*** *of $\delta \in \mathbb{R}$ at $z \in Z$, if there exists $\bar{z}_i \in Z_i$ such that $u_i(\bar{z}_i, z'_{-i}) \geq \delta$ for all z'_{-i} in some open neighborhood of z_{-i}.*

Thus, a payoff can be secured by i at z if i has a strategy that guarantees at least that payoff even if the other players deviate slightly from z.

Definition A.3. *Let $u(z) = (u_C(z), u_L(z), u_H(z))$ be the vector of the players' payoff functions at strategy profile z. The graph of the vector payoff function is $\{(z, u) \in Z \times \mathbb{R}^I : u = u(z)\}$.*

The closure of the graph of the vector payoff function is $\{(z, u) \in Z \times \mathbb{R}^I : \exists\ z^k \to z \text{ with } u(z^k) \to u\}$.

Definition A.4. *(Reny, 1999) The game Γ is* ***better-reply secure*** *if whenever (z^*, u^*) is in the closure of the graph of its vector payoff function and z^* is not an equilibrium, some player i can secure a payoff strictly above u_i^* at z^*.*

The game $\overline{\Gamma}$ is ***better-reply secure*** *if whenever (ϕ^*, u^*) is in the closure of the graph of its mixed extension vector payoff function and ϕ^* is not an equilibrium, some player i can secure a payoff strictly above u_i^* at ϕ^*.*

Theorem A.1. *(Reny, 1999) Suppose Γ is a compact, Hausdorff game. Then Γ possesses a mixed strategy Nash equilibrium if its mixed extension, $\overline{\Gamma}$, is better-reply secure.*

Hence, it is sufficient to show that $\overline{\Gamma}$ is better-reply secure to establish that an equilibrium in mixed strategy exists.

Given (ϕ_L, ϕ_H), let $\hat{v}(s, \phi_L, \phi_H)$ denotes the best response of the representative consumer for $s \in \{$👎, 👍$\}$

$$\hat{v}(s, \phi_L, \phi_H) = \frac{1 - 2b}{2\Big(\theta_L + (\theta_H - \theta_L)\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]\Big)}. \tag{A.52}$$

Let $\overline{\Sigma}$ denote the closure of the graph of the mixed extension's vector payoff function. Then, $\overline{\Sigma}$ can be defined as the union of two subsets:

$$\Lambda = \{(\phi, u) \in \Sigma : \phi_C(\hat{v}(s, \phi_L, \phi_H)) = 1, \ \ \forall \ s \in \{\text{👎}, \text{👍}\}\}$$

and its complements

$$\Lambda^c = \{(\phi, u) \in \Sigma : \exists \ s \in \{\text{👎}, \text{👍}\} \text{ with } \phi_C(\hat{v}(s, \phi_L, \phi_H)) < 1\}.$$

In other words, the set Λ contains the elements where the representative consumer is playing the pure strategy given by (A.52).

Note the following two observations:

1. for any $(\phi, u) \in \overline{\Sigma}$ and for all $s \in \{$👎, 👍$\}$, $\phi_C(\hat{v}(s, \phi_L, \phi_H)) = 1$ is always the only best response to (ϕ_L, ϕ_H) (c.f. Lemma 2.1). That is, all strategies (pure or mixed) of the representative consumer are dominated by the pure strategy in (A.52),

2. when the sellers deviate from (ϕ_L, ϕ_H), the impact on the representative consumer's payoff is only through $\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]$. If $\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]|_{(\phi_L, \phi_H)} = y$, then for small deviations (ϕ'_L, ϕ'_H), we have $\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]|_{(\phi'_L, \phi'_H)} = y + \varepsilon$.

I now show that the game $\overline{\Gamma}$ is better-reply secure.

<u>Step 1:</u> Take an element $(\phi^*, u^*) \in \Lambda^c$ and assume that $\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]|_{(\phi^*_L, \phi^*_H)} = y$. Deviations from (ϕ^*_L, ϕ^*_H) means that $\mathbb{E}[q_b(s, \tilde{m}_L, \tilde{m}_H; T)]|_{(\phi_L, \phi_H)} = y + \varepsilon$ for some $\varepsilon \in [-y, 1-y]$. Then for the representative consumer, for all y, the strategy $\phi_C\left(\hat{v}(s, \phi^*_L, \phi^*_H)\right) = 1$ secures him a higher payoff than ϕ^*_C as long as ε is small enough. In other words, there is

always a neighborhood of (ϕ_L^*, ϕ_H^*) such that $u_C(\hat{v}(s, \phi_L^*, \phi_H^*), \phi_L', \phi_H') > u_C^*$ for all (ϕ_L', ϕ_H') in that neighborhood.

Step 2: Take an element $(\phi^*, u^*) \in \Lambda$. By assumption, the representative consumer is playing a best-response, which implies that if $(\phi_C^*, \phi_L^*, \phi_H^*)$ is not an equilibrium, then there exists a type of seller $x \in \{L, H\}$ for which there exists ϕ_x' that is a better response than ϕ_x^* to (ϕ_C^*, ϕ_{-x}^*). As the payoff function of seller of type x is continuous in the others' strategies, this means that there is also some neighborhood of (ϕ_C^*, ϕ_{-x}^*) where ϕ_x' secures seller of type x a higher payoff than the payoff at $(\phi_C^*, \phi_L^*, \phi_H^*)$.

Hence, Step 1 and Step 2 prove that $\overline{\Gamma}$ is better-reply secure, and thus, that $\overline{\Gamma}$ admits a Nash equilibrium. Note that the equilibrium is also perfect Bayesian because subsequent play is always optimal and beliefs are derived by Bayes' rule.

Existence with the Continuous Design

The proof that a mixed strategy equilibrium exists when the platform is using the continuous design follows similar steps. Given a profile of pure strategy (v, m_H, m_L), the seller of type $x \in \{L, H\}$ profits are given by

$$\Pi_x(m_x, m_{-x}, v) \quad = \quad -c_x m_x^2 + \int_s (1 - v(s)) f(s|\theta_x, m_x) \mathrm{d}s \tag{A.53}$$

while the representative consumer's payoff for a signal s is

$$u_C(s, v, m_L, m_H) = (\theta_L + (\theta_H - \theta_L) q_c(s, m_L, m_H))(1 - v^2) - (1 - 2b)(1 - v) \tag{A.54}$$

with $q_c(s, m_L, m_H)$ as the Bayes updated probability that $\theta = \theta_H$ given (m_L, m_H) and design is continuous.

As with the binary design, the payoff function of seller x is continuous while the payoff function of the representative consumer can be discontinuous since $q_c(s, m_L, m_H)$ is undefined for s such that $\mathbb{P}(s|\theta_L, m_L) = \mathbb{P}(s|\theta_H, m_H) = 0$. The strategy set of the seller with quality θ_x can be restricted to $[0, \sqrt{1/c_x}]$. Indeed any $m_x > \sqrt{1/c_x}$ yields to the seller with quality θ_x strictly negative profits for any profile $\{m_{-x}, \hat{v}(s)\}$. So, $m_x > \sqrt{1/c_x}$ is strictly dominated by 0 which yields to the seller, in the worst case, profits of 0. Then, showing that Reny (1999)'s result can be used to prove the existence of an equilibrium boils down to show that the game is better-reply secure. The remainder of the proof proceeds exactly as with the binary design and is left to the reader.

Proposition 2.8: Given the consumers threshold function is $\hat{v}(s,m)$, the consumers' demand for the product is $1-\hat{v}(s,m)$. Then the expected number of sales from a design $(\mathcal{M}, S)$ is

$$q\mathbb{E}\left[1-\hat{v}(\tilde{s},m))|\theta_H\right] + (1-q)\mathbb{E}\left[1-\hat{v}(\tilde{s},m))|\theta_L\right]. \tag{A.55}$$

Assume that for each sale, the commission earned by the platform is a fixed percentage of the price. Then, the platform's expected revenues are proportional of the expected number of sales.

At the prior q and with $\hat{v}(q)$ as the consumers' threshold, the platform's expected revenues are proportional to $1-F(\hat{v}(q))$, where $F(\cdot)$ is the uniform CDF on $[0,1]$. This threshold is $\hat{v}(q) = \frac{1-2b}{2(\theta_L+(\theta_H-\theta_L)q)}$, such that $\partial^2\hat{v}(q)/\partial q^2 > 0$. It is then the case that $1-F(\hat{v}(q))$ in concave in q. Therefore, by Kamenica and Gentzkow (2011)'s result, a transactional platform does not benefit from providing information to consumers.

Note that the concavity of the expected revenues hinges on the fact that $\hat{v}(q)$ is convex in q in combination to the fact that $F(\cdot)$ is linear. The distribution of consumers' type is therefore important.[1] Note also that the expected profits are higher in the case where consumers learn nothing than when consumers learn the quality perfectly, i.e., the expected demand is higher at the prior than with full revelation

$$\mathbb{E}[D(q)] = 1-\left(\frac{1-2b}{2}\right)\left(\frac{1}{(1-q)\theta_L+q\theta_H}\right) > 1-\left(\frac{1-2b}{2}\right)\left(\frac{1-q}{\theta_L}+\frac{q}{\theta_H}\right) = qD(1)+(1-q)D(0) \tag{A.56}$$

Proposition 2.9: The statement for the case with no manipulation follows from Proposition 2.2. The statement for the case with manipulation follows from Proposition 2.6.

Lemma 2.6: Appendix A.4.1 gives the condition for complete learning to occur with the continuous design. Assume that $c_H = c_L = c$. Then, when $\theta_H - \theta_L - 2b + \frac{D(1)-D(0)}{8bc} \geq 0$, it implies that there exists $m_H \in (2b-(\theta_H-\theta_L), \frac{D(1)-D(0)}{4bc}]$ such that Equation (A.64) is satisfied. In turn, this means that an equilibrium inducing complete learning exists when the platform is using the continuous design.

[1]One could imagine a situation where $F(\cdot)$ is sufficiently concave such that the result does not hold.

A.4 Proof of Proposition 2.4

Throughout the proof, let $D(y)$ be the consumers' demand given posterior beliefs are y. For an arbitrary pair of manipulation levels (m_L, m_H), there are five cases to consider.

A. $\theta_L + 2b + m_L < \theta_H + m_H$

B. $\theta_H + m_H = \theta_L + m_L$

C. $\theta_L + m_L < \theta_H + m_H < \theta_L + 2b + m_L$

D. $\theta_H + 2b + m_H < \theta_L + m_L$

E. $\theta_H + m_H < \theta_L + m_L < \theta_H + 2b + m_H$

- Cases D and E are possible only with $m_L > 0$ and therefore, by Lemma 2.4, they can be eliminated as potential equilibrium candidates.

- Cases C can also be eliminated: Let $D(1), D(q), D(0)$ be the demand for the product given posterior beliefs are $1, q$, and 0, respectively. Assume that for out-of-equilibrium realizations of α such that $\alpha < \theta_L + m_L$, consumers' posterior beliefs are that quality is high with probability β_2 and that for $\alpha > \theta_H + 2b + m_H$, consumers' posterior beliefs are that quality is high with probability β_1.

  Suppose $m_L > 0$, then it is not profitable for the low-quality seller to increase m_L if $\frac{D(1)-D(0)}{2b} - 2c_L m_L \leq 0$ which is equivalent to $\frac{D(1)-D(0)}{4bc_L} \leq m_L$. It is not profitable for the low-quality seller to decrease m_L if $-\frac{D(q)-D(\beta_2)}{2b} + 2c_L m_L \leq 0$ which is equivalent to $\frac{D(q)-D(\beta_2)}{4bc_L} \geq m_L$. Thus, the absence of a profitable deviation requires that $\frac{D(1)-D(0)}{4bc_L} \leq m_L \leq \frac{D(q)-D(\beta_2)}{4bc_L}$, which is impossible because $D(1)-D(0) > D(q)-D(\beta_2)$ for all β_2.

  Suppose $m_L = 0$, then increasing m_L is a strict profitable deviation for the low type since $\frac{D(1)-D(0)}{2b} - 2c_L 0 > 0$.

- Case A is a equilibrium candidate only if $m_L = 0$. To see this, suppose on the contrary that $m_L > 0$. The support of $\alpha^m(\theta_L)$ and $\alpha^m(\theta_H)$ are disjoint under case A's assumptions. Assume further that for out-of-equilibrium realizations of α such that $\alpha < \theta_L + m_L$, consumers' posterior beliefs are that quality is high with probability

β_2. Then, decreasing m_L is profitable for the low type as $-\frac{D(0)-D(\beta_2)}{2b} + 2c_L m_L > 0$ for all β_2. By deviating, the low type cannot do worst in terms of demand, but he economizes on the manipulation cost.

A.4.1 Case A (Complete learning)

Case A requires that $m_L = 0$ and $m_H > 2b - (\theta_H - \theta_L)$. Fix the consumers' conjectures to $(0, m_H)$. Then for any realizations $\alpha \in [\theta_L, \theta_L + 2b]$, the beliefs are that product quality is low, and for any realizations $\alpha \in [\theta_H + m_H, \theta_H + 2b + m_H]$, that product quality is high.

Assume that for out-of-equilibrium realizations of α such that $\alpha < \theta_L$, consumers' posterior beliefs are that quality is high with probability β_2, for $\theta_L + 2b < \alpha < \theta_H + m_H$, consumers' posterior beliefs are that quality is high with probability β_3, and that for $\alpha > \theta_H + 2b + m_H$ consumers' posterior beliefs are that quality is high with probability β_1.

Local deviations

- For the high-quality seller, it is not profitable to increase m_H if $\frac{D(\beta_1)-D(1)}{2b} - 2c_H m_H \leq 0$, which is true for all β_1. Moreover, it is not profitable to decrease m_H if $-\frac{D(1)-D(\beta_3)}{2b} + 2c_H m_H \leq 0$, which requires that $m_H \leq \frac{D(1)-D(\beta_3)}{4bc_H}$.
- For the low-quality seller, it is not profitable to increase m_L if $\frac{D(\beta_3)-D(0)}{2b} - 2c_L \cdot 0 \leq 0$, which is true only if $\beta_3 = 0$.

Case A remains an equilibrium candidate only if the consumers' posterior beliefs are that quality is high with probability 0 when $\theta_L + 2b + m_L < \alpha < \theta_H + m_H$. I make this assumption for the remainder of the proof.

Global deviations

For the high-quality seller, it must also be the case that $m_H = 0$ is not a profitable deviation. Two situations can arise when $m_H = 0$: either it is possible for the high type to obtain $D(1)$ with a positive probability by not manipulating the reviews, either it is not.

When the conjectured equilibrium level m_H is above or equal to $2b$, the high type by choosing no manipulation cannot obtain the demand $D(1)$ with positive probability. The expected profits are

$$\mathbb{E}[\Pi_H(0);(0,m_H)|m_H \geq 2b] = D(0). \tag{A.57}$$

Otherwise, when the conjectured equilibrium level m_H is below $2b$, the high type by choosing no manipulation can obtain the demand $D(1)$ with positive probability. The expected profits are

$$\mathbb{E}[\Pi(0);(0,m_H)|m_H < 2b] = D(0) + \left(\frac{2b - m_H}{2b}\right)(D(1) - D(0)) \tag{A.58}$$

(A.58) is greater than (A.57) and so it suffices to show that a deviation to $m'_H = 0$ with payoff given by (A.58) is not profitable for the high type:

$$\begin{aligned} D(1) - c_H m_H^2 &\geq D(0) + (D(1) - D(0))\left(\frac{2b - m_H}{2b}\right) \\ (D(1) - D(0))\left(\frac{m_H}{2b}\right) - c_H m_H^2 &\geq 0 \\ m_H\left(\frac{D(1) - D(0)}{2b} - c_H m_H\right) &\geq 0 \end{aligned}$$

which is true since $m_H \leq \frac{D(1)-D(0)}{4bc_H}$. Therefore, the high-quality seller does not have a global deviation.

For the low-quality seller, the deviations to eliminate are levels above 0. Given that $\theta_H + m_H > \theta_L + 2b$, the low type has an impact on its demand by choosing m'_L instead of 0, only if $\theta_L + 2b + m'_L > \theta_H + m_H$. Moreover, there is no need to manipulate more than $\theta_H - \theta_L + m_H$, as this levels insures that $D(1)$ occurs with probability 1. Hence, only deviations in $(\theta_H - \theta_L + m_H - 2b, \theta_H - \theta_L + m_H]$ need to be eliminated.

For $m'_L \in (\theta_H - \theta_L + m_H - 2b, \theta_H - \theta_L + m_H]$, the low type expected profits are

$$\mathbb{E}[\Pi_L(m'_L);(0,m_H)] = D(1) - (D(1) - D(0))\left(\frac{(\theta_H - \theta_L) - m'_L + m_H}{2b}\right) - c_L m'^2_L \tag{A.59}$$

which are concave in m'_L. Taking the first order condition with respect to m'_L, equalizing to 0, and solving for the critical point, I obtain $\hat{m}'_L = \frac{D(1)-D(0)}{4bc_L}$.

There are three scenarios to consider: a) $\hat{m}'_L \in (\theta_H - \theta_L + m_H - 2b, \theta_H - \theta_L + m_H]$; b) $\hat{m}'_L > \theta_H - \theta_L + m_H]$; c) $\hat{m}'_L \leq \theta_H - \theta_L + m_H - 2b$. Scenario a) concerns the situations where the critical point is in the set of "admissible" values for a deviation. In this case, one needs to check that $\hat{m}'_L$ is not a profitable deviation from $m_L = 0$. Scenarios b) and c) concern the situations where the critical point is not in the set of "admissible" values for a deviation. In scenario b), the critical point $\hat{m}'_L$ is too high, and one needs to check that $\theta_H - \theta_L + m_H$ is not a profitable deviation from $m_L = 0$. In scenario c), the critical point $\hat{m}'_L$ is too low, and there is no profitable deviation from $m_L = 0$.

Hence, to show that no profitable deviation exists, it is sufficient to show that the deviations to $m'_L = \frac{D(1)-D(0)}{4bc_L}$ and to $m'_L = \theta_H - \theta_L + m_H$ are not profitable.

For $m'_L = \hat{m}'_L$ the low type profits reduce to

$$\mathbb{E}\left[\Pi_L(\hat{m}'_L);(0,m_H)\right] = D(1) - (D(1) - D(0))\left(\frac{\theta_H - \theta_L + m_H}{2b}\right) + \frac{(D(1)-D(0))^2}{16b^2c_L} \quad \text{(A.60)}$$

and for $m'_L = \theta_H - \theta_L + m_H$, they reduce to

$$\mathbb{E}[\Pi_L(\theta_H - \theta_L + m_H);(m_H,0)] = D(1) - c_L(\theta_H - \theta_L + m_H)^2. \quad \text{(A.61)}$$

Scenario a): if $m'_L = \frac{D(1)-D(0)}{4bc_L}$ is admissible, it is sufficient to check that $0 \gtrsim \frac{D(1)-D(0)}{4bc_L}$, which requires that

$$\begin{aligned} & D(0) - \left(D(1) - (D(1) - D(0))\left(\frac{(\theta_H - \theta_L) + m_H}{2b}\right) + \frac{(D(1)-D(0))^2}{16b^2c_L}\right) \\ = \;& (D(1) - D(0))\left(\frac{(\theta_H - \theta_L) + m_H - 2b}{2b}\right) - \frac{(D(1)-D(0))^2}{16b^2c_L} \\ = \;& (\theta_H - \theta_L) + m_H - 2b - \frac{(D(1)-D(0))}{8bc_L} \geq 0 \end{aligned} \quad \text{(A.62)}$$

Scenario b): $m'_L = \frac{D(1)-D(0)}{4bc_L}$ is not admissible because it is too high, that is, if $\frac{D(1)-D(0)}{4bc_L} > \theta_H - \theta_L + m_H$, it is necessary that $0 \gtrsim \theta_H - \theta_L + m_H$, which requires that

$$\begin{aligned} & D(0) - \left(D(1) - c_L(\theta_H - \theta_L + m_H)^2\right) \\ = \;& (\theta_H - \theta_L) + m_H - \sqrt{\frac{(D(1)-D(0)}{c_L}} \geq 0 \end{aligned} \quad \text{(A.63)}$$

Scenario c): $m'_L = \frac{D(1)-D(0)}{4bc_L}$ is not admissible because it is too low, that is, if $\frac{D(1)-D(0)}{4bc_L} < \theta_H - \theta_L + m_H - 2b$, then no profitable deviation.

SUMMARY: If the out-of-equilibrium beliefs are $\beta_1, \beta_2 \in [0,1]$ and $\beta_3 = 0$ and that

i) $m_H^* \in \left(2b - (\theta_H - \theta_L), \frac{D(1)-D(0)}{4bc_H}\right]$

ii)

$$\begin{cases} \frac{D(1)-D(0)}{4bc_L} \in (\theta_H - \theta_L - 2b + m_H^*, \theta_H - \theta_L + m_H^*] \\ \theta_H - \theta_L + m_H^* - 2b - \frac{D(1)-D(0)}{8bc_L} \geq 0 \end{cases} \tag{A.64}$$

or

$$\begin{cases} \frac{D(1)-D(0)}{4bc_L} > \theta_H - \theta_L + m_H^* \\ \theta_H - \theta_L + m_H^* - \sqrt{\frac{D(1)-D(0)}{c_L}} \geq 0 \end{cases} \tag{A.65}$$

or

$$\frac{D(1)-D(0)}{4bc_L} \leq \theta_H - \theta_L + m_H^* - 2b \tag{A.66}$$

then $(0, m_H^*)$, such that Case A occurs, is an equilibrium of the manipulation game with a Continuous design.[2]

A.4.2 Case B (No learning)

Case B requires that $\theta_H + m_H = \theta_L + m_L$, or equivalently that $m_L = \theta_H - \theta_L + m_H$.

In this case, for $\theta_H + m_H \leq \alpha \leq \theta_H + 2b + m_H$, the consumers' posterior beliefs are equal to the prior. Assume further that out-of-equilibrium beliefs for any realizations $\alpha < \theta_H + m_H = \theta_L + m_L$, the beliefs are that product quality is high with probability β_2, and for any realizations $\alpha > \theta_H + 2b + m_H$, that product quality is high with probability β_1.

[2] Notice that when Equation (A.64) is true, then Equation (A.65) must also be true. Indeed suppose that Equation (A.64) is true, but not Equation (A.65), then $(\theta_H - \theta_L) + m_H \geq 2b + \frac{D(1)-D(0)}{8bc_L}$ and $\theta_H - \theta_L + m_H < \sqrt{\frac{D(1)-D(0)}{c_L}}$ which then implies $2b + \frac{D(1)-D(0)}{8bc_L} < \sqrt{\frac{D(1)-D(0)}{c_L}} \Rightarrow \left(\sqrt{2b} - \sqrt{\frac{D(1)-D(0)}{8bc_L}}\right)^2 < 0$, where the last inequality is false.

Local Deviations

For a seller with quality θ_x with $x \in \{H, L\}$, it is not profitable to increase m_x if $(D(\beta_1) - D(q))/2b - 2c_x m_x \leq 0$ and it is not profitable to decrease m_x if $-(D(q) - D(\beta_2))/2b + 2c_x m_x \leq 0$. Hence if (m_H, m_L) are such that

$$\frac{D(\beta_1) - D(q)}{4bc_x} \leq m_x \leq \frac{D(q) - D(\beta_2)}{4bc_x}, \tag{A.67}$$

for $x \in \{H, L\}$, then no seller has a local profitable deviation.

Note that it is necessary that $\beta_2 < q$, otherwise a type with $m_x > 0$ can for sure profitably deviate in decreasing m_x as $-(D(q) - D(\beta_2))/2b + 2c_x m_x > 0$ if $\beta_2 \geq q$.

Global Deviations

Given $m_x \in \left[\frac{D(\beta_1)-D(q)}{4bc_x}, \frac{D(q)-D(\beta_2)}{4bc_x}\right]$, the expected profits of a seller are

$$\mathbb{E}[\Pi_x(m_x); (m_L, m_H)] = D(q) - c_x m_x^2. \tag{A.68}$$

The only global deviation to be eliminated is the deviation to $m_x = 0$. If $\theta_x + m_x < \theta_x + 2b$, it means that the seller with quality θ_x by choosing not to manipulate the reviews can obtain the demand $D(q)$ with positive probability. The expected profits are

$$\mathbb{E}[\Pi_x(0); (m_L, m_H)|m_x < 2b] = D(\beta_2) + \left(\frac{2b - m_x}{2b}\right)(D(q) - D(\beta_2)). \tag{A.69}$$

If instead $\theta_x + m_x \geq \theta_x + 2b$, it means that the seller with quality θ_x by choosing not to manipulate the reviews can only obtain the demand $D(\beta_2)$. The expected profits are

$$\mathbb{E}[\Pi_x(0); (m_L, m_H)|m_x \geq 2b] = D(\beta_2). \tag{A.70}$$

Because it is required that $\beta_2 < q$, it suffices to show that the profits in (A.69) are smaller than the profits in (A.68) to check that there is no profitable global deviations.

$$D(q) - c_x m_x^2 - \left(D(\beta_2) + \left(\frac{2b - m_x}{2b}\right)(D(q) - D(\beta_2))\right) \geq 0$$

$$D(q) - D(\beta_2) - c_x m_x^2 - \left(\frac{2b - m_x}{2b}\right)(D(q) - D(\beta_2)) \geq 0$$

$$-c_x m_x^2 + \left(\frac{m_x}{2b}\right)(D(q) - D(\beta_2)) \geq 0$$

which is true since $m_x \leq \frac{D(q)-D(\beta_2)}{4bc_x}$.

Now, case B can occur only if $m_L > 0$, but it is not necessarily the case that $m_H > 0$ also. Suppose that $m_H = 0$ and $m_L = \theta_H - \theta_L$ so that we are still in case B. Then, there will be no profitable deviation for the high type only if it is not better for him to increase m_H which requires that $(D(\beta_1) - D(q))/2b - 2c \cdot 0 \leq 0 \Rightarrow D(\beta_1) \leq D(q) \Rightarrow \beta_1 \leq q$.

SUMMARY: If out-of-equilibrium beliefs values are $\beta_2 < q$, $\beta_1 \in [0, 1]$, and (m_L^*, m_H^*) are such that

$$\begin{cases} m_x^* \in \left[\frac{D(\beta_1)-D(q)}{4bc_x}, \frac{D(q)-D(\beta_2)}{4bc_x}\right] & \text{for } x \in \{H, L\}, \\ m_L^* = \theta_H - \theta_L + m_H^* \end{cases} \tag{A.71}$$

then (m_L^*, m_H^*), such that Case B occurs, is an equilibrium of the manipulation game with a Continuous Design.

If out-of-equilibrium beliefs values are $\beta_2 < q$, $\beta_1 \leq q$, and $(m_L^*, 0)$ are such that

$$\begin{cases} m_L^* \in \left[\frac{D(\beta_1)-D(q)}{4bc_L}, \frac{D(q)-D(\beta_2)}{4bc_L}\right] \\ m_L^* = \theta_H - \theta_L \end{cases} \tag{A.72}$$

then $(m_L^*, 0)$, such that Case B occurs, is an equilibrium of the manipulation game with a Continuous Design.

A.5 Proof of Proposition 2.5

Throughout the proof, let $D(y)$ be the consumers' demand given posterior beliefs are y.

A.5.1 Equilibrium Candidates

The threshold T has a critical impact on the seller's payoffs and on the impact of a manipulation effort m. More specifically, for a given T there are two important values for m: $\underline{m}_x(T) = \max\{0, T - \theta_x - 2b\}$ and $\overline{m}_x(T) = \max\{0, T - \theta_x\}$. The level $\underline{m}_x(T)$

represents the minimal level of manipulation that is needed in order for the demand to be different from the one obtained with $m_x = 0$. The level $\overline{m}_x(T)$ represents the maximal level of manipulation that has an impact on the demand.

Assuming consumers posterior beliefs are $q_b(\ThumbsDown; T)$ when they see the signal $s = \ThumbsDown$ and $q_b(\ThumbsUp; T)$ when they see the signal $s = \ThumbsUp$, then the seller's expected profits are

$$\begin{aligned} \mathbb{E}[\Pi_x(m_x, q_b(\cdot; T), T)] &= \min\left\{\max\left\{\frac{\theta_x + 2b + m_x - T}{2b}, 0\right\}, 1\right\}\Big(D(q_b(\ThumbsUp; T)) - D(q_b(\ThumbsDown; T))\Big) \\ &\quad + D(q_b(\ThumbsDown; T)) - c_x m_x^2 \end{aligned} \tag{A.73}$$

A seller with quality θ_x chooses $m_x \geq 0$ given the consumers beliefs are $q_b(\ThumbsUp; T)$ or $q_b(\ThumbsDown; T)$, and demand is $D(q_b(\ThumbsUp; T))$ or $D(q_b(\ThumbsDown; T))$ when they see the signal $\ThumbsUp$ or $\ThumbsDown$.

Then,

- all levels $m_x \in [0, \underline{m}_x(T)]$ are strictly dominated by 0. For $m_x = 0$, expected profits are

$$\min\left\{\max\left\{\frac{\theta_x + 2b - T}{2b}, 0\right\}, 1\right\}\Big(D(q_b(\ThumbsUp; T)) - D(q_b(\ThumbsDown; T))\Big) + D(q_b(\ThumbsDown; T)) \tag{A.74}$$

- all levels $m_x \in (\overline{m}_x(T), \infty)$ are strictly dominated by $\overline{m}_x(T)$. For $m_x = \overline{m}_x(T)$, expected profits are

$$D(q_b(\ThumbsUp; T)) - c_\theta \overline{m}_x(T)^2 \tag{A.75}$$

- For $m_x \in (\underline{m}_x(T), \overline{m}_x(T))$, the expected profits in (A.73) are concave in m_x. Taking the first order condition to (A.73) yields the local extremum point

$$\hat{m}_x = \frac{D(q_b(\ThumbsUp; T)) - D(q_b(\ThumbsDown; T))}{4bc_x}, \tag{A.76}$$

and the expected profits at $\hat{m}_x$ are

$$D(q_b(\text{👎}; T)) + \left(\frac{\theta_x + 2b - T}{2b}\right)\Big(D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T))\Big) + \frac{\Big(D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T))\Big)^2}{16b^2 c_x}. \tag{A.77}$$

For a seller with quality θ_x, candidates for a best-response are $\{0, \hat{m}_x, \overline{m}_x(T)\}$. Hence, there are potentially up to nine pairs of manipulation levels that are candidates for an equilibrium.

The next three lemmas are useful to find the conditions under which each candidate profile is an equilibrium. In the remainder, I use $m \gtrsim m'$ to denote the fact that m is at least as good as m' for the seller.

Lemma A.2. *If $\hat{m}_x$ is in the set of admissible values, i.e., $\hat{m}_x \in (\underline{m}_x(T), \overline{m}_x(T))$, then $\hat{m}_x \gtrsim \overline{m}_x(T)$.*

Proof.

$$\begin{aligned}
&\mathbb{E}[\Pi_x(\hat{m}_x)] - \mathbb{E}[\Pi_x(\overline{m}_x(T))] \\
&= D(q_b(\text{👎}; T)) + \left(\frac{\theta_x + 2b - T}{2b}\right)(D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T))) \\
&\quad + \frac{(D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T)))^2}{16b^2 c_x} - \Big(D(q_b(\text{👍}; T)) - c_x \overline{m}_x(T)^2\Big) \\
&= \left(\frac{\theta_x - T}{2b}\right)(D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T))) + \frac{(D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T)))^2}{16b^2 c_x} + c_x(T - \theta_x)^2 \\
&= \left(\sqrt{c_x}(T - \theta_x) - \frac{D(q_b(\text{👍}; T)) - D(q_b(\text{👎}; T))}{4b\sqrt{c_x}}\right)^2 \geq 0.
\end{aligned} \tag{A.78}$$

□

Lemma A.3. *For $T \geq \theta_x + 2b$,*

i. $\overline{m}_x(T) \gtrsim 0$ *if*

$$\frac{D(q_b(\text{👍};T)) - D(q_b(\text{👎};T))}{(T-\theta_x)^2} \geq c_x, \tag{A.79}$$

ii. $\hat{m}_x \gtrsim 0$ *(provided that $\hat{m}_x$ is admissible) if*

$$\frac{D(q_b(\text{👍};T)) - D(q_b(\text{👎};T))}{8b(T-2b-\theta_x)} \geq c_x. \tag{A.80}$$

Proof. i. From (A.74) and (A.75),

$$\mathbb{E}[\Pi_x(\overline{m}_x(T))] - \mathbb{E}[\Pi_x(0)] = D(q_b(\text{👍};T)) - c_x(T-\theta_x)^2 - D(q_b(\text{👎};T))$$

ii. From (A.74) and (A.77),

$$\begin{aligned}
&\mathbb{E}[\Pi_x(\hat{m}_x);(m_H,m_L)] - \mathbb{E}[\Pi_x(0);(m_H,m_L)] \\
&= D(q_b(\text{👎};T)) + \left(\frac{\theta_x+2b-T}{2b}\right)(D(q_b(\text{👍};T)) - D(q_b(\text{👎};T))) \\
&\quad + \frac{(D(q_b(\text{👍};T)) - D(q_b(\text{👎};T)))^2}{16b^2c_x} - D(q_b(\text{👎};T)) \\
&= \left(\frac{\theta_x+2b-T}{2b}\right)(D(q_b(\text{👍};T)) - D(q_b(\text{👎};T))) + \frac{(D(q_b(\text{👍};T)) - D(q_b(\text{👎};T)))^2}{16b^2c_x}
\end{aligned} \tag{A.81}$$

□

Lemma A.4. *For $\theta_x < T < \theta_x + 2b$,*

i. $\hat{m}_x \gtrsim 0$ *whenever $\hat{m}_x$ is admissible,*

ii. $\overline{m}_x(T) \gtrsim 0$ *if*

$$\frac{D(q_b(\text{👍};T)) - D(q_b(\text{👎};T))}{2b(T-\theta_x)} \geq c_x. \tag{A.82}$$

Proof. i. From (A.74) and (A.77),

$$\begin{aligned}
&\mathbb{E}[\Pi_x(\hat{m}_x)] - \mathbb{E}[\Pi_x(0)] \\
&= D(q_b(\leftthumbsdown; T)) + \left(\frac{\theta_x + 2b - T}{2b}\right)(D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T))) \\
&\quad + \frac{(D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T)))^2}{16b^2 c_x} \\
&\quad - \left(D(q_b(\leftthumbsup; T)) + \left(\frac{\theta_x - T}{2b}\right)(D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T)))\right) \\
&= \frac{(D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T)))^2}{16b^2 c_x} \geq 0
\end{aligned} \tag{A.83}$$

ii. From (A.74) and (A.75),

$$\begin{aligned}
&\mathbb{E}[\Pi_x(\overline{m}_x(T))] - \mathbb{E}[\Pi_x(0)] \\
&= D(q_b(\leftthumbsup; T)) - c_x(T - \theta_x)^2 - \left(D(q_b(\leftthumbsup; T)) + \left(\frac{\theta_x - T}{2b}\right)(D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T)))\right) \\
&= -c_x(T - \theta_x)^2 + \left(\frac{T - \theta_x}{2b}\right)(D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T))) \\
&= -c_x(T - \theta_x) + \left(\frac{D(q_b(\leftthumbsup; T)) - D(q_b(\leftthumbsdown; T))}{2b}\right)
\end{aligned} \tag{A.84}$$

□

The table on the next page summarizes the conditions found in the previous lemmas.

for $T \in [\cdot,\cdot]$	$[0,\theta_L]$	$[\theta_L,\theta_H]$	$[\theta_H,\theta_L+2b]$	$[\theta_L+2b,\theta_H+2b]$	$[\theta_H+2b,\infty)$
$\hat{m}_H \gtrsim 0$	$\emptyset$	$\emptyset$	$c_H > 0$	$c_H > 0$	$\frac{(D(q_b(\text{👍};T))-D(q_b(\text{👎};T)))}{8b(T-2b-\theta_H)} \geq c_H$
$\overline{m}_H(T) \gtrsim 0$	$\emptyset$	$\emptyset$	$\frac{D(q_b(\text{👍};T))-D(q_b(\text{👎};T))}{2b(T-\theta_H)} \geq c_H$	$\frac{D(q_b(\text{👍};T))-D(q_b(\text{👎};T))}{2b(T-\theta_H)} \geq c_H$	$\frac{(D(q_b(\text{👍};T))-D(q_b(\text{👎};T)))}{(T-\theta_H)^2} \geq c_H$
$\hat{m}_L \gtrsim 0$	$\emptyset$	$c_L > 0$	$c_L > 0$	$\frac{(D(q_b(\text{👍};T))-D(q_b(\text{👎};T)))}{8b(T-2b-\theta_L)} \geq c_L$	$\frac{(D(q_b(\text{👍};T))-D(q_b(\text{👎};T)))}{8b(T-2b-\theta_L)} \geq c_L$
$\overline{m}_L(T) \gtrsim 0$	$\emptyset$	$\frac{D(q_b(\text{👍};T))-D(q_b(\text{👎};T))}{2b(T-\theta_L)} \geq c_L$	$\frac{D(q_b(\text{👍};T))-D(q_b(\text{👎};T))}{2b(T-\theta_L)} \geq c_L$	$\frac{(D(q_b(\text{👍};T))-D(q_b(\text{👎};T)))}{(T-\theta_L)^2} \geq c_L$	$\frac{(D(q_b(\text{👍};T))-D(q_b(\text{👎};T)))}{(T-\theta_L)^2} \geq c_L$

Table A.1: Binary design: Summary of the conditions for which $m_x, m'_x \in \{0, \hat{m}_x, \overline{m}_x(T)\}$ are such that $m_x \gtrsim m'_x$

$T \in \mathcal{T}$	$\hat{m}_H$ admissible if
$\mathcal{T}_1$	$\frac{D(q_b(\text{👍},(m_L,\hat{m}_H);T))-D(q_b(\text{👎},(m_L,\hat{m}_H);T))}{4b(T-\theta_H)} \leq c_H \leq \frac{D(q_b(\text{👍},(m_L,\hat{m}_H);T))-D(q_b(\text{👎},(m_L,\hat{m}_H);T))}{4b(T-\theta_H-2b)}$
$\mathcal{T}_2$	$\frac{D(q_b(\text{👍},(m_L,\hat{m}_H);T))-D(q_b(\text{👎},(m_L,\hat{m}_H);T))}{4b(T-\theta_H)} \leq c_H$
$\mathcal{T}_3$	$\frac{D(q_b(\text{👍},(m_L,\hat{m}_H);T))-D(q_b(\text{👎},(m_L,\hat{m}_H);T))}{4b(T-\theta_H)} \leq c_H$
$\mathcal{T}_4$	$\emptyset$
$\mathcal{T}_5$	$\emptyset$

Table A.2: Binary design: Conditions for $\hat{m}_H \in (\underline{m}_H(T), \overline{m}_H(T))$

$T \in \mathcal{T}$	$\hat{m}_L$ admissible if
$\mathcal{T}_1$	$\frac{D(q_b(\text{👍},(\hat{m}_L,m_H);T))-D(q_b(\text{👎},(\hat{m}_L,m_H);T))}{4b(T-\theta_L)} \leq c_L \leq \frac{D(q_b(\text{👍},(\hat{m}_L,m_H);T))-D(q_b(\text{👎},(\hat{m}_L,m_H);T))}{4b(T-\theta_L-2b)}$
$\mathcal{T}_2$	$\frac{D(q_b(\text{👍},(\hat{m}_L,m_H);T))-D(q_b(\text{👎},(\hat{m}_L,m_H);T))}{4b(T-\theta_L)} \leq c_L \leq \frac{D(q_b(\text{👍},(\hat{m}_L,m_H);T))-D(q_b(\text{👎},(\hat{m}_L,m_H);T))}{4b(T-\theta_L-2b)}$
$\mathcal{T}_3$	$\frac{D(q_b(\text{👍},(\hat{m}_L,m_H);T))-D(q_b(\text{👎},(\hat{m}_L,m_H);T))}{4b(T-\theta_L)} \leq c_L$
$\mathcal{T}_4$	$\frac{D(q_b(\text{👍},(\hat{m}_L,m_H);T))-D(q_b(\text{👎},(\hat{m}_L,m_H);T))}{4b(T-\theta_L)} \leq c_L$
$\mathcal{T}_5$	$\emptyset$

Table A.3: Binary design: Conditions for $\hat{m}_L \in (\underline{m}_L(T), \overline{m}_L(T))$

The conditions presented in Table A.1 determine which profile of manipulation can be sustained as best-responses when $D(q_b(\text{👍}; T))$ and $D(q_b(\text{👎}; T))$ are fixed in accordance with the consumer conjectures on m_H and m_L. For an equilibrium, more is required. More specifically, in an equilibrium, $q_b(\text{👍}; T)$ and $q_b(\text{👎}; T)$ are functions of m_H and m_L that are chosen by the seller and not conjectures. Hence, to be exact, $q_b(\text{👍}; T)$ and $q_b(\text{👎}; T)$ should be instead written $q_b(\text{👍}, m_L, m_H; T)$ and $q_b(\text{👎}, m_L, m_H; T)$ where

$$q_b(\text{👍}, m_L, m_H; T) = \frac{1}{1 + \frac{(1-q)}{q}\left(\frac{2b+\theta_L+m_L-T}{2b+\theta_H+m_H-T}\right)} \quad \text{and} \quad q_b(\text{👎}, m_L, m_H; T) = \frac{1}{1 + \frac{(1-q)}{q}\left(\frac{T-\theta_L-m_L}{T-\theta_H-m_H}\right)}.$$

There are nine profiles (m_L, m_H) that are equilibrium candidates:

1. $(0, 0)$
2. $(\hat{m}_L, 0)$ where $\hat{m}_L$ solves

$$m_L = \frac{D(q_b(\text{👍}, m_L, 0; T)) - D(q_b(\text{👎}, m_L, 0; T)}{4bc_L} \tag{A.85}$$

3. $(\overline{m}_L, 0)$
4. $(0, \hat{m}_H)$ where $\hat{m}_H$ solves

$$m_H = \frac{D(q_b(\text{👍}, 0, m_H; T)) - D(q_b(\text{👎}, 0, m_H; T))}{4bc_H} \tag{A.86}$$

5. $(\hat{m}_L, \hat{m}_H)$ where $\hat{m}_H$ and $\hat{m}_L$ solve the system

$$\begin{cases} m_H = \frac{D(q_b(\text{👍}, m_L, m_H; T)) - D(q_b(\text{👎}, m_L, m_H; T))}{4bc_H} \\ m_L = \frac{D(q_b(\text{👍}, m_L, m_H; T)) - D(q_b(\text{👎}, m_L, m_H; T))}{4bc_L} \end{cases} \tag{A.87}$$

6. $(\overline{m}_L, \hat{m}_H)$ where $\hat{m}_H$ solves

$$m_H = \frac{D(q_b(\text{👍}, \overline{m}_L, m_H; T)) - D(q_b(\text{👎}, \overline{m}_L, m_H; T)}{4bc_H} \tag{A.88}$$

7. $(0, \overline{m}_H)$

8. $(\hat{m}_L, \overline{m}_H)$ where $\hat{m}_L$ solves

$$m_L = \frac{D(q_b(👍, m_L, \overline{m}_H; T)) - D(q_b(👎, m_L, \overline{m}_H; T))}{4bc_L} \tag{A.89}$$

9. $(\overline{m}_L, \overline{m}_H)$

I now proceed to give the conditions under which each candidate is an equilibrium. The level of the threshold is determinant for the type of profiles that can be equilibrium.

Let $\hat{m}_x(q_b(s, m, T))$ be the solution to

$$\hat{m}_x(q_b(s, m; T)) = \frac{D(q_b(👍, m_L, m_H, T) - D(q_b(👎, m_L, m_H, T)}{4bc_x}, \tag{A.90}$$

then, the next table summarizes the zones for which each pair can be an equilibrium:

(m_L, m_H)	0	$\hat{m}_H(q_b(s, m, T))$	$\overline{m}_H(T)$
0	$\{\mathcal{T}_1, \mathcal{T}_5\}$	$\{\mathcal{T}_1, \mathcal{T}_2\}$	$\{\mathcal{T}_1, \mathcal{T}_2\}$
$\hat{m}_L(q_b(s, m, T))$	$\mathcal{T}_4$	$\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$	$\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$
$\overline{m}_L(T)$	$\mathcal{T}_4$	$\emptyset$	$\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$

Note that the solution $\hat{m}_x(q_b(s, m; T))$ in (A.90) is admissible only if it belongs to the set $(\underline{m}_x(T), \overline{m}_x(T))$. Table A.2 and Table A.3 give the conditions for this to be satisfied.

A.5.2 Conditions under which each candidate is an equilibrium

A.5.2.1 $(0, 0)$

Lemma A.4 together with the condition for the admissibility of $\hat{m}_x$ imply that the high type cannot choose $m_H = 0$ in $\mathcal{T}_2$ and $\mathcal{T}_3$ and the low type cannot choose $m_L = 0$ in $\mathcal{T}_3$ and $\mathcal{T}_4$. Thus $(0, 0)$ can possibly occur only in $\mathcal{T}_1$ and $\mathcal{T}_5$.

- In $\mathcal{T}_1$, when $(m_H, m_L) = (0, 0)$ the consumers beliefs are $q_b(👎; T) = q$ and $q_b(👍; T) = z \in [0, 1]$ (out of equilibrium specification). The assumption that $z \leq q$ is sufficient to insure that $(0, 0)$ is an equilibrium. For $z > q$, then

  i. if the deviation to $\hat{m}_x = \frac{D(q_b(👍; T) - D(q_b(👎; T)}{4bc_x}$ is admissible, then it is not profitable when

$$\frac{D(z) - D(q)}{8b(T - \theta_x - 2b)} \leq c_x \tag{A.91}$$

ii. otherwise, the only potentially profitable deviation is $\overline{m}_x$ and it is not profitable when

$$\frac{D(z) - D(q)}{(T - \theta_x)^2} \leq c_x \tag{A.92}$$

- In $\mathcal{T}_5$, no type has an incentive to manipulate the reviews since the signal 👍 is published with probability 1. The only sustainable pair is $(0, 0)$.

A.5.2.2 $(\hat{m}_L, 0)$

In this case, there is no out-of-equilibrium to specify since the choice of the low type insures that both 👍 and 👎 can be observed.

- In $\mathcal{T}_1$, the entire support of the high type is below T when $m_H = 0$ and the support of the low type is partially below T only. Hence by Lemma 2.4, the low type can profitable deviate by decreasing its manipulation effort.

- In $\mathcal{T}_2$ and $\mathcal{T}_3$, Lemma A.4 implies that it cannot be optimal for the high type to make no manipulation effort.

- In $\mathcal{T}_4$, the entire support of the high type is above T when $m_H = 0$, there is no profitable deviation for him. It is implicit that $\hat{m}_L$ is admissible because, the low type is choosing it. Then, by Lemma A.4, there is no profitable deviation for the low type.

A.5.2.3 $(\overline{m}_L, 0)$

- In $\mathcal{T}_1$, the entire support of the high type is below T when $m_H = 0$ and the support of the low type is entirely above T. Hence by Lemma 2.4, the low type can profitably deviate by decreasing its manipulation effort.

- In $\mathcal{T}_2$ and $\mathcal{T}_3$,the support of the high type is partially below T when $m_H = 0$ and the support of the low type is entirely above T. Hence by Lemma 2.4, the low type can profitable deviate by decreasing its manipulation effort.

- In $\mathcal{T}_4$, the entire support of both the high and the low types are below T, such that only 👍 is observed. Assume that out-of-equilibrium beliefs when seeing 👎 are that quality is high with probability z. There is no potentially profitable deviation for the high type. For the low type, it is necessary that $z < q$ for him not to deviate to $m_L < \overline{m}_L$. Moreover, it is required that

i. $\hat{m}_L$ is not admissible

ii. $\overline{m}_L \gtrsim 0$

$$\frac{D(q) - D(z)}{2b(T - \theta_L)} \geq c_L$$

A.5.2.4 $(0, \hat{m}_H)$

By Lemma A.4, the low type cannot choose $m_L = 0$ in $\mathcal{T}_3$ and $\mathcal{T}_4$. Moreover, it is impossible that the high type manipulates in $\mathcal{T}_5$. Thus $(0, \hat{m}_H)$ can occur only in $\mathcal{T}_1$ and $\mathcal{T}_2$.

When $(m_L, m_H) = (0, \hat{m}_H)$, the consumers beliefs are $q_b(\text{👍}; T) = 1$ and

$$q_b(\text{👎}, \hat{m}_H) = \frac{1}{1 + \left(\frac{1-q}{q}\right)\left(\frac{2b}{T - \theta_H - \hat{m}_H}\right)}. \tag{A.93}$$

The high type manipulation $\hat{m}_H$ is the solution to

$$\hat{m}_H = \frac{D(1) - D(q_b(\text{👎}, 0, \hat{m}_H; T))}{4bc_H}, \tag{A.94}$$

which has the following closed form solution

$$\begin{aligned} \hat{m}^*_H(T) &= \frac{c_H\theta_H \cdot R_1 \pm \sqrt{c_H\theta_H^2 \left(c_H \cdot R_1^2 - (1-2b)(\theta_H - \theta_L)(1-q)q\right)}}{2c_H\theta_H^2 q} \\ & \quad \text{with } R_1 = 2b\theta_L(1-q) + \theta_H q \left(T - \theta_H\right). \end{aligned} \tag{A.95}$$

Deviations

The low type could potentially deviate to $\overline{m}_L$ or to an interior level of manipulation $\hat{m}_L$, where

$$\hat{m}_L = \frac{D(1) - D(q_b(\text{👎}, 0, \hat{m}_H; T))}{4bc_L}. \tag{A.96}$$

If the deviation to $\hat{m}_L$ is not profitable (provided that $\hat{m}_L$ is admissible), then it will be the case that the deviation $\overline{m}_L$ is also not profitable by Lemma A.2.

The high type can deviate to 0 or to $\overline{m}_H$. The assumption that $\hat{m}_H$ is admissible together with Lemma A.2 imply that $\overline{m}_H$ is not a profitable deviation.

- In $\mathcal{T}_1$, $(0, \hat{m}_H)$ is an equilibrium if
  - $\hat{m}_H$ is admissible:
$$\frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - \theta_H)} \leq c_H \leq \frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - 2b - \theta_H)}. \tag{A.97}$$
  - $\hat{m}_H \gtrsim 0$:
$$c_H \leq \frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{8b(T - 2b - \theta_H)}. \tag{A.98}$$

  and either one of the following hold:
  - $\hat{m}_L$ is admissible and $0 \gtrsim \hat{m}_L$:
$$\frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - \theta_L)} \leq c_L \leq \frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - 2b - \theta_L)}, \tag{A.99}$$
$$c_L > \frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{8b(T - 2b - \theta_L)}, \tag{A.100}$$
  - $\hat{m}_L$ is not admissible and $0 \gtrsim \overline{m}_L$:
$$\frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - \theta_L)} > c_L \quad \text{or} \quad c_L > \frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - 2b - \theta_L)}, \tag{A.101}$$
$$c_L > \frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{(T - \theta_L)^2}. \tag{A.102}$$
- In $\mathcal{T}_2$, the conditions that guarantees that the low type has no profitable deviation from 0 are exactly the same as in $\mathcal{T}_1$. For the high type to choose $\hat{m}_H$, it is required that
  - $\hat{m}_H \gtrsim 0$: which is true for all $c_H > 0$.
  - $\hat{m}_H$ be admissible:
$$\frac{D(1) - D(q_b(\text{☟}, 0, \hat{m}_H; T))}{4b(T - \theta_H)} \leq c_H \tag{A.103}$$

A.5.2.5 $(\hat{m}_L, \hat{m}_H)$

It is impossible that the high type manipulates in $\mathcal{T}_4$ and $\mathcal{T}_5$. Thus the pair $(\hat{m}_H, \hat{m}_L)$ can occur only in $\{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3\}$.

Let

$$q_b(\text{👍}, m_L, m_H, T) = \frac{1}{1 + \frac{(1-q)}{q}\left(\frac{2b+\theta_L+m_L-T}{2b+\theta_H+m_H-T}\right)} \tag{A.104}$$

$$q_b(\text{👎}, m_L, m_H, T) = \frac{1}{1 + \frac{(1-q)}{q}\left(\frac{T-\theta_L-m_L}{T-\theta_H-m_H}\right)}, \tag{A.105}$$

then the interior levels $\hat{m}_H$ and $\hat{m}_L$ are found by solving the system

$$\begin{cases} m_L = \frac{D(q_b(\text{👍}, m_L, m_H, T) - D(q_b(\text{👎}, m_L, m_H, T)}{4bc_L} \\ m_H = \frac{D(q_b(\text{👍}, m_L, m_H, T) - D(q_b(\text{👎}, m_L, m_H, T)}{4bc_L} \end{cases} \tag{A.106}$$

which is equivalent to solving

$$\begin{cases} m_L = \left(\frac{1-2b}{8bc_L}\right)\left(\frac{1}{\theta_L + (\theta_H - \theta_L) q_b(\text{👎}, m_H, m_L, T)} - \frac{1}{\theta_L + (\theta_H - \theta_L) q_b(\text{👍}, m_H, m_L, T)}\right) \\ m_H = \left(\frac{c_L}{c_H}\right) m_L. \end{cases} \tag{A.107}$$

For $(\hat{m}_L, \hat{m}_H)$ to be an equilibrium, it is required that $\hat{m}_x$ be admissible and that $\hat{m}_x \gtrsim 0$ for $x \in \{H, L\}$.

- In $\mathcal{T}_1$, the level $\hat{m}_x$ is admissible if

$$\frac{D(q_b(\text{👍}, \hat{m}_L, \hat{m}_H; T)) - D(q_b(\text{👎}, \hat{m}_L, \hat{m}_H; T))}{4b(T - \theta_x)} \le c_x \le \frac{D(q_b(\text{👍}, \hat{m}_L, \hat{m}_H; T)) - D(q_b(\text{👎}, \hat{m}_L, \hat{m}_H; T))}{4b(T - 2b - \theta_x)}. \tag{A.108}$$

  A seller with quality θ_x prefers $\hat{m}_x$ over 0 if

$$c_x \le \frac{D(q_b(\text{👍}, \hat{m}_L, \hat{m}_H; T)) - D(q_b(\text{👎}, \hat{m}_L, \hat{m}_H; T))}{8b(T - 2b - \theta_x)} \tag{A.109}$$

- In $\mathcal{T}_2$, the conditions that guarantees that the low type has no profitable deviation from $\hat{m}_L$ are exactly the same as in $\mathcal{T}_1$.

  For the high type, $\hat{m}_H$ is preferred to 0 for any $c_H > 0$ and is now admissible if

$$c_H \ge \frac{D(q_b(\text{👍}, \hat{m}_L, \hat{m}_H; T)) - D(q_b(\text{👎}, \hat{m}_L, \hat{m}_H; T))}{4b(T - \theta_H)}. \tag{A.110}$$

- In $\mathcal{T}_3$, the conditions that guarantees that the high type has no profitable deviation from $\hat{m}_H$ are exactly the same as in $\mathcal{T}_2$.

  For the low type, $\hat{m}_L$ is admissible if

  $$c_L \geq \frac{D(q_b(\ThumbsUp, \hat{m}_L, \hat{m}_H; T)) - D(q_b(\ThumbsDown, \hat{m}_L, \hat{m}_H; T))}{4b(T - \theta_L)}. \tag{A.111}$$

  and is preferred to 0 for any $c_L > 0$.

A.5.2.6 $(\overline{m}_L(T), \hat{m}_H)$

This case is not an equilibrium for any set of parameters. To see this, suppose that the seller is choosing $(m_L, m_H) = (\overline{m}_L, \hat{m}_H)$. This implies that the support of the low type is completely above the threshold and the the high type's support is above only partially. Then, we will have $q_b(\ThumbsDown, \overline{m}_L, \hat{m}_H, T) > q_b(\ThumbsUp, \overline{m}_L, \hat{m}_H, T)$. Therefore there is a profitable deviation for the low type which is to decrease m_L.

A.5.2.7 $(0, \overline{m}_H(T))$

By Lemma A.4, the low type cannot choose $m_L = 0$ in $\mathcal{T}_3$ and $\mathcal{T}_4$. Moreover, it is impossible that the high type manipulates in $\mathcal{T}_5$. Thus $(0, \overline{m}_H(T))$ can occur only in $\mathcal{T}_1$ and $\mathcal{T}_2$.

When $(m_L, m_H) = (0, \overline{m}_H(T))$, the consumers beliefs are

$$q_b(\ThumbsUp, 0, \overline{m}_H; T) = 1 \quad \text{and} \quad q_b(\ThumbsDown, 0, \overline{m}_H; T) = 0.$$

In other words, there is perfect revelation of quality.

<u>*Deviations*</u>

A type can potentially deviate to an interior level of manipulation $\hat{m}_x$, where

$$\hat{m}_x = \frac{D(1) - D(0)}{4bc_x}. \tag{A.112}$$

For the low type, there will be no profitable deviation if $0 \gtrsim \hat{m}_L$ and $\hat{m}_L$ is admissible or if $0 \gtrsim \overline{m}_L$ and $\hat{m}_L$ is not admissible.

For the high type, Lemma A.2 implies that $\hat{m}_H$ cannot be admissible, otherwise it is preferred to $\overline{m}_L$. There will be no profitable deviation if $\overline{m}_H \gtrsim 0$ and $\hat{m}_H$ is not admissible.

- In $\mathcal{T}_1$, $(0, \overline{m}_H)$ is an equilibrium if
  - $\hat{m}_H$ is not admissible:
$$\frac{D(1)-D(0)}{4b(T-\theta_H)} > c_H \quad \text{or} \quad c_H > \frac{D(1)-D(0)}{4b(T-2b-\theta_H)}. \tag{A.113}$$
  - $\overline{m}_H \gtrsim 0$:
$$c_H \leq \frac{D(1)-D(0)}{(T-\theta_H)^2}. \tag{A.114}$$

  and either one of the following hold:

  - $\hat{m}_L$ is admissible and $0 \gtrsim \hat{m}_L$:
$$\frac{D(1)-D(0)}{4b(T-\theta_L)} \leq c_L \leq \frac{D(1)-D(0)}{4b(T-2b-\theta_L)}. \tag{A.115}$$
$$c_L > \frac{D(1)-D(0)}{8b(T-2b-\theta_L)}, \tag{A.116}$$
  - $\hat{m}_L$ is not admissible and $0 \gtrsim \overline{m}_L$:
$$\frac{D(1)-D(0)}{4b(T-\theta_L)} > c_L \quad \text{or} \quad c_L > \frac{D(1)-D(0)}{4b(T-2b-\theta_L)}. \tag{A.117}$$
$$c_L > \frac{D(1)-D(0)}{(T-\theta_L)^2}. \tag{A.118}$$

- In $\mathcal{T}_2$, the conditions that guarantees that the low type has no profitable deviation from 0 are exactly the same as in $\mathcal{T}_1$. For the high type to choose $\overline{m}_H(T)$, it is required that :
  - $\hat{m}_H$ is not admissible:
$$\frac{D(1)-D(0)}{4b(T-\theta_H)} > c_H \tag{A.119}$$
  - $\overline{m}_H \gtrsim 0$:
$$c_H \leq \frac{D(1)-D(0)}{2b(T-\theta_H)}. \tag{A.120}$$

A.5.2.8 $(\hat{m}_L, \overline{m}_H(T))$

The pair $(\hat{m}_L, \overline{m}_H(T))$ can occur in $\mathcal{T}_1, \mathcal{T}_2$, and $\mathcal{T}_3$ with $\overline{m}_H(T) > 0$. In $\mathcal{T}_4$, $\overline{m}_H(T) = 0$ and the pair becomes $(\hat{m}_L, 0)$ a case already covered in A.5.2.2.

At $(\hat{m}_L, \overline{m}_H(T))$, the consumers posterior belief are

$$q_b(\text{👎}, \hat{m}_L, \overline{m}_H, T) = 0 \quad \text{and} \quad q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T) = \frac{1}{1 + \frac{(1-q)}{q}\left(\frac{2b+\theta_L+\hat{m}_L-T}{2b}\right)}$$

The low type level $\hat{m}_L$ is the solution to

$$\hat{m}_L = \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4bc_L}, \tag{A.121}$$

which has the following closed form solution

$$\hat{m}_L^* = \frac{-c_L\theta_L \cdot R_0 \pm \sqrt{c_L\theta_L^2\left(c_L \cdot R_0^2 + (1-2b)(\theta_H - \theta_L)(1-q)q\right)}}{2c_L\theta_L^2(1-q)}, \tag{A.122}$$

$$\text{with } R_0 = 2b\left(\theta_L(1-q) + \theta_H q\right) - \theta_L(1-q)\left(T - \theta_L\right), \tag{A.123}$$

<u>*Deviations*</u>

The high type can potentially deviate to an interior level of manipulation $\hat{m}_H$, where

$$\hat{m}_H = \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4bc_H}. \tag{A.124}$$

By assumption that $\hat{m}_L$ is the chosen level by type L, $\hat{m}_L$ must be admissible and thus by Lemma A.2, only the deviation to 0 needs to be checked.

For the high type, for him to choose $\overline{m}_H$, Lemma A.2 also implies that $\hat{m}_H$ must not be admissible. Then, the level $\overline{m}_H$ must be preferred to 0.

- In $\mathcal{T}_1$, $(\hat{m}_L, \overline{m}_H(T))$ is an equilibrium if
  - $\hat{m}_H$ is not admissible:

$$c_H < \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4b(T - \theta_H)} \quad \text{or} \quad c_H > \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4b(T - 2b - \theta_H)}. \tag{A.125}$$

– $\overline{m}_H \gtrsim 0$:

$$c_H \leq \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{(T - \theta_H)^2}. \tag{A.126}$$

– $\hat{m}_L$ is admissible:

$$\frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4b(T - \theta_L)} \leq c_L \leq \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4b(T - 2b - \theta_L)}. \tag{A.127}$$

– $\hat{m}_L \gtrsim 0$:

$$c_L \leq \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{8b(T - 2b - \theta_L)}. \tag{A.128}$$

- In $\mathcal{T}_2$, the conditions that guarantees that the low type has no profitable deviation from $\hat{m}_L$ are exactly the same as in $\mathcal{T}_1$. For the high type to choose $\overline{m}_H(T)$, it is required that:

– $\hat{m}_H$ is not admissible:

$$c_H < \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4b(T - \theta_H)} \tag{A.129}$$

– $\overline{m}_H(T) \gtrsim 0$:

$$c_H \leq \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{2b(T - \theta_H)}. \tag{A.130}$$

- In $\mathcal{T}_3$, the conditions that guarantees that the high type has no profitable deviation from $\overline{m}_H(T)$ are exactly the same as in $\mathcal{T}_2$. For the low type, as long as $\hat{m}_L$ is admissible than it is preferred to 0 for any $c_L > 0$,

– $\hat{m}_L$ is admissible:

$$c_L \geq \frac{D(q_b(\text{👍}, \hat{m}_L, \overline{m}_H, T)) - D(0)}{4b(T - \theta_L)}. \tag{A.131}$$

A.5.2.9 $(\overline{m}_L, \overline{m}_H)$

The pair $(\overline{m}_L, \overline{m}_H)$ can only occur in $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$. For $\mathcal{T}_4, \mathcal{T}_5$, we have $\overline{m}_x = 0$ for at least one $x \in \{H, L\}$ and the case is already covered in A.5.2.3 and A.5.2.1.

At $(\overline{m}_L, \overline{m}_H)$, the consumers' posterior beliefs are $q_b(\text{👎}, \overline{m}_L, \overline{m}_H, T) = z$ and $q_b(\text{👍}, \overline{m}_L, \overline{m}_H, T) = q$, where $z \in [0, q)$ is an out-of-equilibrium specification. Notice that the profile $(\overline{m}_L, \overline{m}_H)$ cannot be an equilibrium for $z \geq q$, since there will be a profitable deviation for both types to decrease the level of manipulation.

<u>*Deviations*</u>

A type can possible deviate to an interior level $\hat{m}_x$ with $\hat{m}_x = \frac{D(q)-D(z)}{4bc_x}$. Lemma A.2 implies that $\hat{m}_x$ must not be admissible for $\overline{m}_x$ to be chosen in equilibrium. Then, for both types, the deviation to $m_x = 0$ must not be profitable.

- In $\mathcal{T}_1$, $(\overline{m}_L, \overline{m}_H)$ is an equilibrium if for $x \in \{H, L\}$
  - $\hat{m}_x$ is not admissible:
  $$\frac{D(q)-D(z)}{4b(T-\theta_x)} > c_x \quad \text{or } c_x > \frac{D(q)-D(z)}{4b(T-\theta_x-2b)} \tag{A.132}$$
  - $\overline{m}_x \gtrsim 0$:
  $$\frac{D(q)-D(z)}{(T-\theta_x)^2} \geq c_x \tag{A.133}$$
- In $\mathcal{T}_2$, the conditions for the low type are the same as in $\mathcal{T}_1$. For the high type, no profitable deviation exist if
  - $\hat{m}_H$ is not admissible:
  $$\frac{D(q)-D(z)}{4b(T-\theta_H)} > c_H \tag{A.134}$$
  - $\overline{m}_H \gtrsim 0$:
  $$\frac{D(q)-D(z)}{2b(T-\theta_H)} \geq c_H \tag{A.135}$$
- In $\mathcal{T}_3$, the conditions for the high type are the same as in $\mathcal{T}_2$. For the low type, no profitable deviation exist if
  - $\hat{m}_L$ is not admissible:
  $$\frac{D(q)-D(z)}{4b(T-\theta_L)} > c_L \tag{A.136}$$
  - $\overline{m}_L \gtrsim 0$:
  $$\frac{D(q)-D(z)}{2b(T-\theta_L)} \geq c_L. \tag{A.137}$$

Appendix B

Appendix to Chapter 3

B.1 Proofs

Let $\mathbb{E}$ and $\mathbb{V}$ be the *expectation* and *variance* operators, respectively.

Proof of Proposition 3.1. Using Definition 3.1, we proceed in three steps. First, given the uninformed buyers' updating rule, we solve for the firm's optimal price strategies. Second, we derive the distribution of the posterior beliefs that follows from the firm's price strategies and the prior beliefs. Finally, we check that the uninformed buyers's updating rule and the distribution of the price-signals are mutually consistent.

1. Given (3.9), $\mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B] = \delta^*_0 + \delta^*_1 P_A + \delta^*_2 P_B$.

   Plugging (3.1), (3.2), and $\mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B]$ into (3.3) yields

$$\max_{P_A,P_B} \{P_A \cdot (\mu - P_A + \eta_A) + P_B \cdot (\lambda(\gamma\mu - P_B) + (1-\lambda)(\gamma(\delta^*_0 + \delta^*_1 P_A + \delta^*_2 P_B) - P_B) + \eta_B)\}. \tag{B.1}$$

   Taking the first-order conditions with respect to prices yields

$$P_A : \mu - 2P_A + \eta_A + (1-\lambda)\gamma\delta^*_1 P_B = 0, \tag{B.2}$$

$$P_B : \lambda(\gamma\mu - 2P_B) + (1-\lambda)(\gamma(\delta^*_0 + \delta^*_1 P_A + 2\delta^*_2 P_B) - 2P_B) + \eta_B = 0. \tag{B.3}$$

   Given the expressions for δ^*_1 and δ^*_2 given in (3.11) and (3.12), the Hessian matrix is negative definite. Solving (B.2) and (B.3) for the price strategies yields

$$P^*_{\mathcal{D},A}(\mu, \eta_A, \eta_B) = \frac{\delta^*_0\delta^*_1\gamma^2(1-\lambda)^2 + (2 - 2\delta^*_2\gamma(1-\lambda) + \delta^*_1\gamma^2\lambda(1-\lambda))\mu}{4 - \delta^{*2}_1\gamma^2(1-\lambda)^2 - 4\delta^*_2\gamma(1-\lambda)} + \frac{(2 - 2\delta^*_2\gamma(1-\lambda))\eta_A + \delta^*_1\gamma(1-\lambda)\eta_B}{4 - \delta^{*2}_1\gamma^2(1-\lambda)^2 - 4\delta^*_2\gamma(1-\lambda)} \tag{B.4}$$

   and

$$P^*_{\mathcal{D},B}(\mu, \eta_B, \eta_A) = \frac{2\delta^*_0\gamma(1-\lambda) + (\delta^*_1\gamma(1-\lambda) + 2\gamma\lambda)\mu + \delta^*_1\gamma(1-\lambda)\eta_A + 2\eta_B}{4 - \delta^{*2}_1\gamma^2(1-\lambda)^2 - 4\delta^*_2\gamma(1-\lambda)}. \tag{B.5}$$

2. Next, given the firm's price strategies, we solve for the buyers' posterior beliefs. Specifically, using the expressions for $P^*_{\mathcal{D},A}$ and $P^*_{\mathcal{D},B}$, let

$$\tilde{z}_A \equiv 2\left(P^*_{\mathcal{D},A}(\mu,\tilde{\eta}_A,\tilde{\eta}_B) - \frac{\delta_0^*\delta_1^*\gamma^2(1-\lambda)^2}{D_0}\right)\cdot\left(\frac{D_0}{D_0+\delta_1^{*2}\gamma^2(1-\lambda)^2+2\delta_1^*\gamma^2\lambda(1-\lambda)}\right) \tag{B.6}$$

$$= \mu + \left(\frac{2-2\delta_2^*\gamma(1-\lambda)}{D_1}\right)\tilde{\eta}_A + \left(\frac{\delta_1^*\gamma(1-\lambda)}{D_1}\right)\tilde{\eta}_B, \tag{B.7}$$

and

$$\tilde{z}_B \equiv \left(P^*_{\mathcal{D},B}(\mu,\tilde{\eta}_B,\tilde{\eta}_A) - \frac{2\delta_0^*\gamma(1-\lambda)}{D_0}\right)\left(\frac{D_0}{\delta_1^*\gamma(1-\lambda)+2\gamma\lambda}\right) \tag{B.8}$$

$$= \mu + \left(\frac{\delta_1^*\gamma(1-\lambda)}{\delta_1^*\gamma(1-\lambda)+2\gamma\lambda}\right)\tilde{\eta}_A + \left(\frac{2}{\delta_1^*\gamma(1-\lambda)+2\gamma\lambda}\right)\tilde{\eta}_B, \tag{B.9}$$

where

$$D_0 \equiv 4 - 4\delta_2^*\gamma(1-\lambda) - \delta_1^{*2}\gamma^2(1-\lambda)^2, \tag{B.10}$$

$$D_1 \equiv 2 - 2\delta_2^*\gamma(1-\lambda) + \delta_1^*\gamma^2\lambda(1-\lambda). \tag{B.11}$$

From (B.7) and (B.9), $\tilde{\mathbf{z}}|\mu \equiv [\tilde{z}_A, \tilde{z}_B]'|\mu$ is jointly normally distributed. Hence, given the prior distribution $\tilde{\mu} \sim N(\rho, \sigma_\mu^2)$, the posterior distribution of quality μ upon observing $\mathbf{z}$ (i.e., upon observing $\{P_A, P_B\}$) is

$$\tilde{\mu}^*_{\mathcal{D}}|\mathbf{z} \sim N(\rho + \sigma_\mu^2 \mathbb{1}\mathbf{\Sigma}^{-1}(\mathbf{z}-\rho\mathbb{1}'), \sigma_\mu^2 - \sigma_\mu^4\mathbb{1}\mathbf{\Sigma}^{-1}\mathbb{1}') \tag{B.12}$$

where $\mathbb{1}$ is a 1×2 vector of ones and

$$\mathbf{\Sigma} \equiv \begin{bmatrix} \sigma_\mu^2 + \sigma_\eta^2\frac{4(1+\delta_2\gamma(\lambda-1))^2+\delta_1^2\gamma^2(1-\lambda)^2}{D_1^2} & \sigma_\mu^2 + \sigma_\eta^2\frac{2\delta_1\gamma(1-\lambda)(2+\delta_2\gamma(\lambda-1))}{D_1(\delta_1\gamma(1-\lambda)+2\gamma\lambda)} \\ \sigma_\mu^2 + \sigma_\eta^2\frac{2\delta_1\gamma(1-\lambda)(2+\delta_2\gamma(\lambda-1))}{D_1(\delta_1\gamma(1-\lambda)+2\gamma\lambda)} & \sigma_\mu^2 + \sigma_\eta^2\left(\frac{\delta_1^2\gamma^2(1-\lambda)^2+4}{(\delta_1\gamma(1-\lambda)+2\gamma\lambda)^2}\right) \end{bmatrix}. \tag{B.13}$$

Simplifying (B.12) yields

$$\mathbb{E}[\tilde{\mu}^*_{\mathcal{D}}|P_A,P_B] = \frac{\rho\sigma_\eta^2 - \delta_0^*\gamma^2\lambda(1-\lambda)\sigma_\mu^2 + (2-\delta_1^*\gamma^2\lambda(1-\lambda))\sigma_\mu^2 P_A}{\sigma_\eta^2+(1+\gamma^2\lambda^2)\sigma_\mu^2} + \frac{(2\gamma\lambda(1-\delta_2^*\gamma(1-\lambda)) - \delta_1^*\gamma(1-\lambda))\sigma_\mu^2 P_B}{\sigma_\eta^2+(1+\gamma^2\lambda^2)\sigma_\mu^2}, \tag{B.14}$$

and

$$\mathbb{V}[\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B] = \frac{\sigma_\eta^2 \sigma_\mu^2}{\sigma_\eta^2 + (1+\gamma^2\lambda^2)\sigma_\mu^2}. \tag{B.15}$$

3. Setting (B.14) equal to $\delta_0^* + \delta_1^* P_A + \delta_2^* P_B$ and solving for δ_0^*, δ_1^* and δ_2^* yields (3.10), (3.11), and (3.12) Since δ_0^*, δ_1^* and δ_2^* uniquely exist, the posterior beliefs are normally distributed as defined by (3.9) and are consistent with (B.14) and (B.15). Moreover, from (3.7) and (3.8), the price-signals are jointly normally distributed.

Proof of Proposition 3.2. The proof of Proposition 3.2 follows the same steps of the proof of Proposition 3.1. Using Definition 3.2, we proceed as follows.

1. Given (3.16), $\mathbb{E}[\tilde{\mu}^*_{\mathcal{U}}|P] = \beta_0^* + \beta_1^* P$. Plugging (3.1), (3.2), and $\mathbb{E}[\tilde{\mu}^*_{\mathcal{U}}|P]$ into (3.4) yields

$$\max_P \ \{P \cdot ((\mu - P + \eta_A) + \lambda(\gamma\mu - P) + (1-\lambda)(\gamma(\beta_0^* + \beta_1^* P) - P) + \eta_B)\}. \tag{B.16}$$

Taking the first-order condition with respect to P yields

$$(1+\gamma\lambda)\mu + \eta_A + \eta_B + \beta_0^*\gamma(1-\lambda) - 2P(2 - \beta_1^*\gamma(1-\lambda)) = 0. \tag{B.17}$$

Given the expression for β_1^* given in (3.18), the second-order condition holds, i.e., $-2\left(2 - \beta_1^*\gamma(1-\lambda)\right) < 0$. Then, solving (B.17) for the price strategy yields

$$P^*_{\mathcal{U}}(\mu, \eta_A, \eta_B) = \frac{\beta_0^*\gamma(1-\lambda) + (1+\gamma\lambda)\mu + \eta_A + \eta_B}{4 - 2\beta_1^*\gamma(1-\lambda)}. \tag{B.18}$$

2. Next, given the firm's price strategy, we solve for the buyers' posterior beliefs. Specifically, using P^*, let

$$\tilde{z} \equiv \frac{2(2 - \beta_1^*\gamma(1-\lambda))P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B) - \beta_0^*\gamma(1-\lambda)}{1+\gamma\lambda} \tag{B.19}$$

$$= \mu + \frac{\tilde{\eta}_A + \tilde{\eta}_B}{1+\gamma\lambda} \tag{B.20}$$

such that $\tilde{z}|\mu$ is normally distributed with mean μ and variance $\sigma_z^2 \equiv 2\sigma_\eta^2/(1+\gamma\lambda)^2$. Given the prior distribution, $\tilde{\mu} \sim N(\rho, \sigma_\mu^2)$, the posterior belief upon observing z (i.e., upon observing P) is

$$\tilde{\mu}^*_{\mathcal{U}}|z \sim N\left(\frac{\rho\sigma_z^2 + z\sigma_\mu^2}{\sigma_z^2 + \sigma_\mu^2}, \frac{1}{1/\sigma_z^2 + 1/\sigma_\mu^2}\right). \tag{B.21}$$

Hence, simplifying (B.21), the posterior mean and variance are

$$\mathbb{E}[\tilde{\mu}^*_{\mathcal{U}}|P] = \frac{2\rho\sigma^2_\eta + ((4 - 2\beta^*_1\gamma(1-\lambda))P - \beta^*_0\gamma(1-\lambda))(1+\gamma\lambda)\sigma^2_\mu}{2\sigma^2_\eta + (1+\lambda\gamma)^2\sigma^2_\mu} \tag{B.22}$$

and

$$\mathbb{V}[\tilde{\mu}^*_{\mathcal{U}}|P] = \frac{2\sigma^2_\eta\sigma^2_\mu}{2\sigma^2_\eta + (1+\lambda\gamma)^2\sigma^2_\mu}. \tag{B.23}$$

3. Setting (B.22) equal to $\beta^*_0 + \beta^*_1 P$ and solving for β^*_0 and β^*_1 yields (3.17) and (3.18). Since β^*_0 and β^*_1 uniquely exist, the posterior beliefs are normally distributed as defined by (3.16) and are consistent with (B.22) and (B.23). Finally, from (3.15), the price-signal is normally distributed.

B.2 Equilibrium Definition

Definition B.1 states the Perfect Bayesian Equilibrium. The equilibrium consists of the firm's strategy (a segmentation decision at stage 1 and prices at stage 2), the distribution of the price-signals conditional on any quality x, and the uninformed buyers' posterior beliefs about the quality upon observing any prices.[1] In equilibrium, the posterior beliefs are consistent with Bayes' rule and the equilibrium distribution of prices.

Definition B.1. *The tuple* $\{\{M^*, \{\{P^*_{\mathcal{U}}(\mu,\eta_A,\eta_B), \{P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B), P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A)\}\}\}\}$, $\{\hat{\xi}^*_{\mathcal{U}}(\cdot|P_A,P_B), \hat{\xi}^*_{\mathcal{D}}(\cdot|P)\}\}$ *is a perfect Bayesian equilibrium if, for all* $\mu > 0$,

1. *At stage 2,*

   (a) *For* $M^* = \mathcal{U}$,

      i. *Given* $\hat{\xi}^*_{\mathcal{U}}(\cdot|P)$, *and for any* η_A *and* η_B, *the firm's price strategy is*

$$P^*_{\mathcal{U}}(\mu,\eta_A,\eta_B) = \arg\max_P \left\{ P \cdot \Big(Q_A(P,\mu,\eta_A) + Q_B(P,\mu,\hat{\xi}^*_{\mathcal{U}}(\cdot|P),\eta_B)\Big) \right\}. \tag{B.24}$$

      ii. *Given the distribution of* $\{\tilde{\eta}_A,\tilde{\eta}_B\}$, $\phi^*_{\mathcal{U}}(P|x)$ *is the p.d.f. of the random price-signal* $P^*_{\mathcal{U}}(x,\tilde{\eta}_A,\tilde{\eta}_B)$ *conditional on any quality* x.

[1] The variable μ refers to the true quality whereas x is used as a dummy variable for quality.

iii. *Given* $\phi^*_{\mathcal{U}}(P|\cdot)$ *and prior beliefs* $\xi(\cdot)$*, the uninformed buyers' posterior beliefs upon observing any* P *is* $\tilde{\mu}^*|P$ *with p.d.f.*

$$\hat{\xi}^*_{\mathcal{U}}(x|P) = \frac{\xi(x)\phi^*_{\mathcal{U}}(P|x)}{\int_{x'\in\mathbb{R}} \xi(x')\phi^*_{\mathcal{U}}(P|x')\mathrm{d}x'}, \quad \forall x \in \mathbb{R}. \tag{B.25}$$

(b) *For* $M^* = \mathcal{D}$*,*

i. *Given* $\hat{\xi}^*_{\mathcal{D}}(\cdot|P_A, P_B)$*, and for any* η_A *and* η_B*, the firm's price strategies are*

$$\Big\{P^*_{\mathcal{D},A}(\mu,\eta_A,\eta_B), P^*_{\mathcal{D},B}(\mu,\eta_B,\eta_A)\Big\} = \arg\max_{P_A,P_B} \Big\{P_A \cdot Q_A(P_A,\mu,\eta_A) + P_B \cdot Q_B(P_B,\mu,\hat{\xi}^*_{\mathcal{D}}(\cdot|P_A,P_B),\eta_B)\Big\}. \tag{B.26}$$

ii. *Given the distribution of* $\{\tilde{\eta}_A, \tilde{\eta}_B\}$*,* $\phi^*_{\mathcal{D}}(P_A, P_B|x)$ *is the p.d.f. of the random price-signals* $\Big\{P^*_{\mathcal{D},A}(x,\tilde{\eta}_A,\tilde{\eta}_B), P^*_{\mathcal{D},B}(x,\tilde{\eta}_B,\tilde{\eta}_A)\Big\}$ *conditional on any quality* x*.*

iii. *Given* $\phi^*_{\mathcal{D}}(P_A, P_B|\cdot)$ *and prior beliefs* $\xi(\cdot)$*, the uninformed buyers' posterior beliefs about quality upon observing* P_A *and* P_B *is* $\tilde{\mu}^*_{\mathcal{D}}|P_A, P_B$ *with the p.d.f.*

$$\hat{\xi}^*_{\mathcal{D}}(x|P_A,P_B) = \frac{\xi(x)\phi^*_{\mathcal{D}}(P_A,P_B|x)}{\int_{x'\in\mathbb{R}} \xi(x')\phi^*_{\mathcal{D}}(P_A,P_B|x')\mathrm{d}x'}, \quad \forall x \in \mathbb{R}. \tag{B.27}$$

2. *At stage 1,*

$$M^* = \arg\max_{M\in\{\mathcal{U},\mathcal{D}\}} \mathbb{1}_{[M=\mathcal{U}]} \cdot \mathbb{E}[\Pi^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)] + \mathbb{1}_{[M=\mathcal{D}]} \cdot \mathbb{E}[\Pi^*_{\mathcal{D}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)] \tag{B.28}$$

where

$$\mathbb{E}[\Pi^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)] = \mathbb{E}\Big[P^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B) \cdot \Big(Q_A(P^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B),\tilde{\mu},\tilde{\eta}_A) + Q_B(P^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B),\tilde{\mu},\hat{\xi}^*_{\mathcal{U}}(\cdot|P^*_{\mathcal{U}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)),\tilde{\eta}_B)\Big)\Big] \tag{B.29}$$

and

$$\mathbb{E}[\Pi^*_{\mathcal{D}}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B)] = \mathbb{E}\Big[P^*_{\mathcal{D},A}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B) \cdot Q_A(P^*_{\mathcal{D},A}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B),\tilde{\mu},\tilde{\eta}_A)\Big] + \mathbb{E}\Big[P^*_{\mathcal{D},B}(\tilde{\mu},\tilde{\eta}_B,\tilde{\eta}_A) \cdot Q_B(P^*_{\mathcal{D},B}(\tilde{\mu},\tilde{\eta}_B,\tilde{\eta}_A),\tilde{\mu},\hat{\xi}^*_{\mathcal{D}}(\cdot|P^*_{\mathcal{D},A}(\tilde{\mu},\tilde{\eta}_A,\tilde{\eta}_B),P^*_{\mathcal{D},B}(\tilde{\mu},\tilde{\eta}_B,\tilde{\eta}_A)),\tilde{\eta}_B)\Big]. \tag{B.30}$$

B.3 Probability of Exclusion

In this appendix, we study whether the presence of uninformed buyers (or the informational externality) decreases or increases the probability of excluding market B under uniform pricing. In market B, the informed buyers and the uninformed buyers with unbiased prior beliefs do not buy the good if the price is above the reservation price, i.e., $P > \gamma\mu$. Using (3.15), we compare the probability of such an event under the two scenarios of complete and incomplete information, i.e., $P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|_{\lambda=1}$ and $P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|_{\lambda\in(0,1)}$. In Figure B.1, the shaded area encompasses the points $\{\gamma, \lambda\}$ for which the presence of uninformed buyers increases the probability of exclusion, i.e., $\mathbb{P}[\gamma\mu < P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|_{\lambda=1}] < \mathbb{P}[\gamma\mu < P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|_{\lambda\in(0,1)}]$.[2] An increase in the variance of the demand shock increases the size of the shaded area.

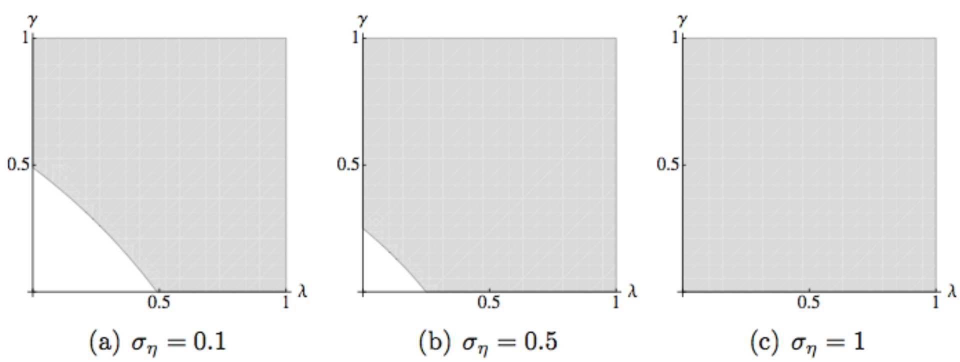


Figure B.1: $\mathbb{P}[\gamma\mu < P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|_{\lambda=1}] < \mathbb{P}[\gamma\mu < P^*_{\mathcal{U}}(\mu, \tilde{\eta}_A, \tilde{\eta}_B)|_{\lambda\in(0,1)}]$ in the shaded area.

[2]To generate Figure B.1, we set $\mu = 1$ and $\sigma^2_\mu = 1$.

Appendix C

Appendix to Chapter 4

C.1 Multivariate First-Order Stochastic Dominance and Dependence Orderings

In this section, we provide equivalent definitions of our dependence ordering, based on multivariate first order stochastic dominance, which we define next. Note that the following definition specializes to the usual first-order stochastic dominance in the univariate case.

Definition C.1 (Multivariate first-order stochastic dominance)**.**

*i. Let f and g be two multivariate probability distribution functions (pdf) on the support $\mathcal{X}^k$. We say that g **first-order stochastically dominates (FOSD)** f if or all increasing L, we have*

$$\sum_{x\in L} f(x) \leq \sum_{x\in L} g(x).$$

*Moreover, we say that g **strictly FOSD** f if g FOSD f, but f does not FOSD g.*

*ii. Let $X = (X_1, ..., X_k)$ and $X' = (X_1', ..., X_k')$ be two random vectors on the support $\mathcal{X}^k$. We say that X **FOSD** X' if the pdf of X FOSD the pdf of X'. Moreover we say that X **strictly FOSD** X' if X FOSDX', but X' does not FOSD X.*

The following result due to Lehman (1955), Levhari, Paroush and Peleg (1975) and Østerdal (2010) provides four alternative and equivalent definitions of multivariate stochastic dominance.

Theorem C.1. *Let X and Y be random vectors with respective pdfs f and g on the support $\mathcal{X}^k$. The following conditions are equivalent:*

i. Y FOSD X.

ii. For all decreasing L, we have

$$\sum_{x\in L} f(x) \geq \sum_{x\in L} g(x).$$

iii. For all nondecreasing mapping $W : \mathcal{X}^k \to \mathbb{R}$, $\mathbb{E}(W(Y)) \geq \mathbb{E}(W(X))$.

iv. There exist two random vectors X' and Y' with respective pdfs f and g such that Y' FOSD X'.

v. There exist a finite list of vector pairs $(x_t, y_t)_{t=1,...,T}$ *with* $x_t \leq y_t$ *and a list of reals* $(\Delta_t)_{t=1,...,m}$ *, with* $\Delta_t \in [0,1]$ *such that*

$$g(x) - f(x) = \sum_t \Delta_t \left(1_{\{y_t\}}(x) - 1_{\{x_t\}}(x)\right).$$

This enables us to provide the following definition equivalent to Definition (4.1) which is in turn used in Definition (4.2), using the notion of multivariate first-order stochastic dominance.

Definition C.2 (weakly greater conditional dependence)**.** *Let* $i \in I$ *and let* X_{-i} *be a profile of signals for all players different from i. For all* X_i *and* X_i' *in* $\mathbb{X}_i$*, we say that* X_i' ***depends at least as much as*** X_i ***on*** X_{-i} ***conditionally on*** Θ*, if for all* (θ, x) *the conditional pdf* $\mathbb{P}(X_{-i} \mid X_i' \geq x,\ \Theta = \theta)$ *FOSD the conditional pdf* $\mathbb{P}(X_{-i} \mid X_i \geq x,\ \Theta = \theta)$*.*

C.2 Most Public Signal, Most Private Signal and Most d-dependent Signal

In this section, we show how to construct two examples of signal structures $(\mathbb{X}_1, ..., \mathbb{X}_N)$ that admit a most public signal, a most private signal and most d-dependent signal.

Example C.1. *Let* $I = \{1, 2\}$ *. Let* $\left(\Theta, X_1^*, X_2^*, X_P^*, Y_1^I, Y_1^{II}, Y_2^I, Y_2^{II}\right)$ *be a random vector distributed on* $\{-1,1\}^4 \times \{0,1\}^4$ *so that, the three vectors* (Θ, X_1^*) *,* (Θ, X_2^*) *and* (Θ, X_P^*) *are distributed as* (Θ, X_I) *in the binary information structure presented in Section 4.3. Moreover, let the random vector* $\left(X_1^*, X_2^*, X_P^*, Y_1^I, Y_1^{II}, Y_2^I, Y_2^{II}\right)$ *be independent conditionally on* $\Theta = \theta$*, for all* $\theta \in \{-1, 1\}$ *and let the vector* $\left(\Theta, Y_1^I, Y_1^{II}, Y_2^I, Y_2^{II}\right)$ *be independent. For each* $i \in \{1, 2\}$ *, we assume that* $\mathbb{P}\left(Y_i^I = 1\right) < \mathbb{P}\left(Y_i^{II} = 1\right)$ *holds. Last. for each* i*, let the set* $\mathbb{X}_i$ *consist of two signals*

$$\begin{cases} X_i^I &= X_P^* Y_i^I + X_i^* \left(1 - Y_i^I\right) \\ X_i^{II} &= X_P^* Y_i^{II} + X_i^* \left(1 - Y_i^{II}\right). \end{cases}$$

In the signal structure constructed in Example C.1, it is easily verified that for each i, the signals $\left(X_i^I, X_i^{II}\right)$ are such that $X_i^I \prec X_i^{II}$, regardless of what signals the other players choose. The signal X_i^I is player i's most private signal and the signal X_i^{II} is player i's most public signal. It is also clear that one can generalize this construction to more than two players, where each player i has a set of signals $\mathbb{X}_i = \{X_i^1, ..., X_i^{m_i}\}$, so that $X_i^1 \prec_i ... \prec_i X_i^{m_i}$. For each i, the signal is player i's most private signal and the signal $X_i^{m_i}$ is player i's most public signal. When generalizing this construction to more than two players and more than two signals per player, each player i still has a most private signal, namely X_i^1, and a most public signal, namely $X_i^{m_i}$, but for an arbitrary dependence vector d^i, it is not the case that player i has a most d^i-dependent signal.

We now provide an example of a signal structure $(\mathbb{X}_1, \mathbb{X}_2, \mathbb{X}_3)$ for three players, such that each of the three players i has a most d^i-dependent signal, for each dependence vector d^i.

Example C.2. *Let* $I = \{1, 2, 3\}$. *Let*

$$(\Theta, X_{12}^*, X_{23}^*, X_{13}^*, X_{11}^*, X_{22}^*, X_{33}^*, Y_1, Y_2, Y_3)$$

be random vector such that Θ *and the* X_{ij}^* *have support* $\{-1, 1\}$ *and the random variable* Y_i *has full support* $\{\{i\}, \{i, j\}, \{i, k\}, I\}$. *For all* i, j, *let the vector* $\left(\Theta, X_{ij}^*\right)$ *be distributed as* (Θ, X_I) *in the binary signal structure presented in Section 4.3. Moreover, let the random vector*

$$(X_{12}^*, X_{23}^*, X_{13}^*, X_{11}^*, X_{22}^*, X_{33}^*, Y_1, , Y_2, Y_3)$$

be independent conditionally on $\Theta = \theta$, *for all* $\theta \in \{-1, 1\}$ *and let the vector* (Θ, Y_1, Y_2, Y_3) *be independent. Last. for each* i, *let the set* $\mathbb{X}_i$ *consist of four signals defined as follows:*

$$\begin{cases} X_i^{\{i,j,k\}} & = & \mathbf{1}_{\{Y_i=ij\}} X_{ij}^* + \mathbf{1}_{\{Y_i=ik\}} X_{ik}^* + \mathbf{1}_{\{Y_i=i\}} X_{ii}^* \\ X_i^{ij} & = & \mathbf{1}_{\{Y_i=ij\}} X_{ij}^* + \mathbf{1}_{\{Y_i=ik\}} X_{ii}^* + \mathbf{1}_{\{Y_i=i\}} X_{ii}^*. \\ X_i^{\{i,k\}} & = & \mathbf{1}_{\{Y_i=ik\}} X_{ik}^* + \mathbf{1}_{\{Y_i=ij\}} X_{ii}^* + \mathbf{1}_{\{Y_i=i\}} X_{ii}^* \\ X_i^{\{i\}} & = & X_{ii}^* \end{cases}$$

In the signal structure constructed in Example C.2, for each player i and each dependence vector i, the signal X_i^S, with $S = \left\{j : d_j^i = 1\right\}$ is player i's most d^i-dependent signal. One can generalize this construction to more than two players and to more signals.

C.3 Proofs

Abusing the terminology defined in section (4.4.1), for all $k \geq 0$, and all $d \in \{-1,1\}^k$, we say that a mapping Φ from $\mathcal{X}^k$ to $\mathbb{R}$ is (strictly) d-**monotonic** if it is (strictly) increasing in each x_j such that $d_j = 1$ and (strictly) decreasing in each x_j such that $d_j = -1$.

The following Lemma is useful in the Proof of Theorem 4.4.

Lemma C.1. *Let X_{-i} be a profile of signals, and let X_i and X_i' be signals. Let $d \in \{-1,1\}^I$ be a dependence vector, with $d_i = 1$ and let Φ be a mapping from $\mathcal{X}^{I\setminus\{i\}}$ to $\mathbb{R}$.*

i. Suppose that X_i' conditionally d-depends as least as much on X_{-i} as X_i does and that Φ is d_{-i}-monotonic. Then for all $z \in \mathcal{X}$,

$$\mathbb{E}_{X_i,X_{-i}}\left[\Phi\left(X_{-i}\right) \mid X_i > z\right] \leq \mathbb{E}_{X_i',X_{-i}}\left[\Phi\left(X_{-i}\right) \mid X_i' > z\right]. \tag{C.1}$$

ii. Suppose that X_i' conditionally d-depends strictly more on X_{-i} than X_i does and that Φ is strictly d_{-i}-monotonic. Then for all $z \in \mathcal{X}$ such that $z \neq \max \mathcal{X}$,

$$\mathbb{E}_{X_i,X_{-i}}\left[\Phi\left(X_{-i}\right) \mid X_i > z\right] < \mathbb{E}_{X_i',X_{-i}}\left[\Phi\left(X_{-i}\right) \mid X_i' > z\right] \tag{C.2}$$

Proof of Lemma C.1: (i) By definition, since X_i' d-depends as least as much on X_{-i} as X_i does, the pdf of $(d_jX_j)_{j\neq i}$, conditional on $X_i' > z$, first-order stochastically dominates the pdf of $(d_jX_j)_{j\neq i}$, conditional on $X_i > z$. Then consider the function $\Gamma\left(x_{-i}\right) = \Phi\left(\left(d_jX_j\right)_{j\neq i}\right)$. Because Φ is d_{-i}-monotonic, the function Γ is increasing. Together, and by the equivalence between (i) and (iii) in Theorem C.1, these last two claims imply that

$$\mathbb{E}_{X_i,X_{-i}}\left[\Gamma\left(\left(d_jX_j\right)_{j\neq i}\right) \mid X_i > z\right] \leq \mathbb{E}_{X_i',X_{-i}}\left[\Gamma\left(\left(d_jX_j\right)_{j\neq i}\right) \mid X_i' > z\right],$$

which is equivalent to inequality (C.1).

(ii) Under assumptions of point (ii), the inequalities hold strictly. ■

Proof of Theorem 4.1:

Let $A_{i,X} : \alpha_{-i} \longrightarrow \alpha_i$ be the action best-response of player i under information structure X. Then, the following lemma is useful for the proof of Theorem 4.1.

Lemma C.2. *Let* (α_1, α_2) *be a Nash Bayesian equilibrium of the game* Γ_X *with an exogenous information structure* X*, such that for all* $i = 1, 2$ *the following holds*

1) $X_1 = X_2$ *if and only if* $\alpha_1(1)\alpha_2(1)b_{ia} > 0$
$X_1 \neq X_2$ *if and only if* $\alpha_1(1)\alpha_2(1)b_{ia} < 0$

2) $A_{i,(X_i,X_{-i})}(\alpha_{-i})$ *has a sign that does not depend on* X_i.

Then, $(X, (\alpha_1, \alpha_2))$ *is a Nash Bayesian equilibrium of the game* Γ *with an endogenous information structure.*

Proof of Lemma C.2: First, for each $i = 1, 2$, because (α_1, α_2) is a Nash Bayesian equilibrium of Γ_X, player i does not have a profitable deviation of the form (X_i, α_i') with $\alpha_i' \neq \alpha_i$.

Suppose by contradiction that $(X, (\alpha_1, \alpha_2))$ is not a Nash Bayesian equilibrium of the game Γ with an endogenous information structure. Then this means that a player i has a profitable deviation (X_i', α_i') with $X_i' \neq X_i$. Then, $\left(X_i', A_{i,(X_i',X_{-i})}(\alpha_{-i})\right)$ is an even better deviation, thus it is also profitable. But because $A_{i,(X_i',X_{-i})}(\alpha_{-i})$ is of the same sign as α_i, and by 1), the deviation $\left(X_i, A_{i,(X_i',X_{-i})}(\alpha_{-i})\right)$ is even better and thus must be profitable. But then, this contradicts the assumption that (α_1, α_2) is a Nash Bayesian equilibrium of the game Γ_X. ■

We now proceed to prove Theorem 4.1. Consider the game given in Section 4.3 and assume that $b_{1\theta} = 4.8, b_{2\theta} = 5$, $b_{1a} = -0.8$, $b_{2a} = -1.2$ and $\varepsilon = 0.26$.

Fix the information structure X, then by taking the first-order condition to (4.2) with respect to $\alpha_i(1)$ and $\alpha_i(-1)$ for $i = 1, 2$, we can observe that $-\alpha_i(1) = \alpha_i(-1)$ and that the best response functions are

$$\begin{cases} \alpha_1(1) = \frac{288}{125} - \frac{4}{5}\alpha_2(1) \text{ and } \alpha_2(1) = \frac{12}{5} - \frac{6}{5}\alpha_1(1) & \text{if } X_1 = X_2 \\ \alpha_1(1) = \frac{288}{125} - \frac{576}{3125}\alpha_2(1) \text{ and } \alpha_2(1) = \frac{12}{5} - \frac{864}{3125}\alpha_1(1) & \text{if } X_1 \neq X_2 \end{cases}. \tag{C.3}$$

Given the information structure is $X_1 \neq X_2$, the unique equilibrium in action strategies is $((1.9616, -1.9616), (1.85766, -1.85766))$. Whereas, when the information structure is $X_1 = X_2$, the unique equilibrium in action strategies is $\left(\left(\frac{48}{5}, -\frac{48}{5}\right), \left(-\frac{228}{25}, \frac{228}{25}\right)\right)$.

Then, by checking that conditions 1) and 2) of Lemma C.2 hold using (C.3), we can show that the profile $(X, (1.9616, -1.9616), (1.85766, -1.85766))$ with $X_1 \neq X_2$ and the profile $\left(X, \left(\frac{48}{5}, -\frac{48}{5}\right), \left(-\frac{228}{25}, \frac{228}{25}\right)\right)$ with $X_1 = X_2$ are both Nash Bayesian equilibria of the game Γ.

Proof of Theorem 4.2: It suffices to give an example where the planner chooses an information structure that differs from what the players choose in the decentralized game. Suppose that player i's payoff is given by Equation (4.6) and that $b_a = -3, b_\theta = 1, b_{\theta a} = -2, b_{aa} = 0.75$ and $\varepsilon = 0.25$. In this case, an equilibrium exists and according to Proposition 4.2, the players will choose to obtain private information. The social planner, however, will prefer to impose public information. ■

Proof of Theorem 4.4: *Proof of* (i). Let $i \in I$ and let c^i be a complementarity vector for i. Suppose that u_i has c^i-increasing differences in own and others' actions. Let $m \in \{-1, 1\}^I$ be a monotonicity profile and let X_{-i} be a profile of signals such that $X_j \in \mathbb{X}_j$ for all j. Suppose that X_i and X_i' are two signals in $\mathbb{X}_i$ such that X_i' d^i-depends as least as much on X_{-i} as X_i does, where d^i is the dependence vector such that the relation (4.7) holds for all j. Let α be profile of pure m-monotonic action strategies. We will show that

$$\mathbb{E}_{\Theta, X_i, X_{-i}} \left(u_i \left(\alpha \left(X_i, X_{-i} \right), \Theta \right) \right) \leq \mathbb{E}_{\Theta, X_i', X_{-i}} \left(u_i \left(\alpha \left(X_i', X_{-i} \right), \Theta \right) \right) \tag{C.4}$$

holds.

Let $z^1 < ... < z^m$ be the elements of $\mathcal{X}$. Also, for each $k \in \{1, .., m-1\}$, and each $\theta \in T$, let $\Phi_{k,\theta}$ be the function from $\mathcal{X}^{I \setminus \{i\}}$ to $\mathbb{R}$ such that

$$\Phi_{k,\theta} \left(x_{-i} \right) = u_i \left(\alpha \left(z^{k+1}, x_{-i} \right), \theta \right) - u_i \left(\alpha \left(z^k, x_{-i} \right), \theta \right).$$

The proof is in four steps.

Step 1: The function $\Phi_{k,\theta}$ is d^i_{-i}-monotonic.

Proof of Step 1: Since for each θ, the function u_i has c^i-monotonic differences in x, and the function α is m-monotonic, it follows that $\Phi_{k,\theta}$ is d^i_{-i}-monotonic. □

Step 2: For all $k \in \{1, .., m-1\}$, and all $\theta \in T$, we have

$$\mathbb{E}_{X_i, X_{-i}} \left[\Phi_{k,\theta} \left(X_{-i} \right) \mid X_i > z^k, \ \Theta = \theta \right] \leq \mathbb{E}_{X_i', X_{-i}} \left[\Phi_{k,\theta} \left(X_{-i} \right) \mid X_i' > z^k, \ \Theta = \theta \right]. \tag{C.5}$$

Proof of Step 2: By Step 1, the function $\Phi_{k,\theta}$ is d^i_{-i}-monotonic. By assumption, X_i' d-depends more on X_{-i} than X_i does, thus by Lemma C.1, inequality (C.5) holds. □

Step 3: For all $k \in \{1, .., m-1\}$, and all $\theta \in T$, we have

$$u_i \left(\alpha \left(X_i, X_{-i} \right), \theta \right) = \sum_{k=0}^{m-1} \mathbb{1}_{\{X_i > k\}} \Phi_{k,\theta} \left(X_{-i} \right) \tag{C.6}$$

and

$$u_i \left(\alpha \left(X_i', X_{-i} \right), \theta \right) = \sum_{k=0}^{m-1} \mathbb{1}_{\{X_i' > k\}} \Phi_{k,\theta} \left(X_{-i} \right). \tag{C.7}$$

Proof of Step 3: These identities are easily verified. We leave them to the reader. □

Step 4:

$$\mathbb{E}_{\Theta,X_i,X_{-i}}\left[u_i\left(\alpha\left(X_i,X_{-i}\right),\Theta\right)\mid\Theta=\theta\right]\leq\mathbb{E}_{\Theta,X_i',X_{-i}}\left[u_i\left(\alpha\left(X_i',X_{-i}\right),\Theta\right)\mid\Theta=\theta\right]. \tag{C.8}$$

Proof of Step 4: We know that

$$\begin{aligned}
& \mathbb{E}_{X_i,X_{-i}}\left[u_i\left(\alpha\left(X_i,X_{-i}\right),\Theta\right)\mid\Theta=\theta\right] \\
= \; & \mathbb{E}_{X_i,X_{-i}}\left[\sum_{k=0}^{m-1}\mathbb{1}_{\{X_i>k\}}\left(\Phi_{k,\theta}\left(X_{-i}\right)\right)\mid\Theta=\theta\right] \\
= \; & \sum_{k=0}^{m-1}\mathbb{E}_{X_i,X_{-i}}\left[\mathbb{1}_{\{X_i>k\}}\Phi_{k,\theta}\left(X_{-i}\right)\mid\Theta=\theta\right] \\
= \; & \sum_{k=0}^{m-1}\mathbb{E}_{X_{-i}}\left[\Phi_{k,\theta}\left(X_{-i}\right)\mid X_i>k\text{ and }\Theta=\theta\right]\mathbb{P}\left(X_i>k\mid\Theta=\theta\right) \\
\leq \; & \sum_{k=0}^{m-1}\mathbb{E}_{X_{-i}}\left[\Phi_{k,\theta}\left(X_{-i}\right)\mid X_i'>k\text{ and }\Theta=\theta\right]\mathbb{P}\left(X_i'>k\mid\Theta=\theta\right) \qquad \text{(C.9)} \\
= \; & \sum_{k=0}^{m-1}\mathbb{E}_{X_i',X_{-i}}\left[\mathbb{1}_{\{X_i'>k\}}\left(\Phi_{k,\theta}\left(X_{-i}\right)\right)\mid\Theta=\theta\right] \\
= \; & \mathbb{E}_{X_i',X_{-i}}\left[\sum_{k=0}^{m-1}\mathbb{1}_{\{X_i'>k\}}\left(\Phi_{k,\theta}\left(X_{-i}\right)\right)\mid\Theta=\theta\right] \\
= \; & \mathbb{E}_{X_i',X_{-i}}\left[u_i\left(\alpha\left(X_i',X_{-i}\right),\Theta\right)\mid\Theta=\theta\right]
\end{aligned}$$

where the first and sixth equalities follow from Step 3, and the inequality follows from Step 2 and from the assumption that (X_i,Θ) and (X_i',Θ) have the same joint marginal distribution. □

Final Step: Since for all θ,

$$\mathbb{E}_{X_i,X_{-i}}\left[u_i\left(\alpha\left(X_i,X_{-i}\right),\Theta\right)\mid\Theta=\theta\right]\leq\mathbb{E}_{X_i',X_{-i}}\left[u_i\left(\alpha\left(X_i',X_{-i}\right),\Theta\right)\mid\Theta=\theta\right], \tag{C.10}$$

taking expectations of both sides on Θ, we obtain

$$\mathbb{E}_{\Theta,X_i,X_{-i}}\left[u_i\left(\alpha\left(X_i,X_{-i}\right),\Theta\right)\right]\leq\mathbb{E}_{\Theta,X_i',X_{-i}}\left[u_i\left(\alpha\left(X_i',X_{-i}\right),\Theta\right)\right], \tag{C.11}$$

the desired conclusion. □■

Proof of (ii). Under the additional assumptions we will prove that the inequality (C.9) is strict at least for some realization θ. First, Step 1 can be modified as follows. since for each θ, the function u_i has strict c^i-complementarities in actions, and the function α is strictly m-monotonic, the function $\Phi_{k,\theta}$ is strictly d^i_{-i}-monotonic for all θ and k. Second, because X_i' d-depends more than X_i on X_{-i}, there exists a realization θ° and some integer k such that the pdf $\mathbb{P}\left(\left(d^i_j X_j\right)_{j\neq i} \mid d^i_i X_i' > k \text{ and } \Theta = \theta\right)$ strictly stochastically dominates the pdf $\mathbb{P}\left(\left(d^i_j X_j\right)_{j\neq i} \mid d^i_i X_i' > k \text{ and } \Theta = \theta\right)$. For this realization θ° and this integer k, the inequality (C.5) holds strictly. As a result, the inequality (C.10) holds strictly as well. Finally, since all realizations θ of Θ have positive probability, the inequality (C.11) holds strictly as well. ■

Proof of Corollary 4.4: Let $(X_1, X_2, \alpha_1, \alpha_2)$ be a pure Nash-Bayesian equilibrium of the game such that α is strictly isotonic (if the payoff complementarities in actions are strictly positive) or antitonic (if they are strictly negative) and suppose by contradiction that $X_1 \neq X_2$ with positive probability. Then by Theorem 4.4, the deviation (X_1', α_1) with $X_1' = X_2$ is strictly profitable for player 1, since X_1' depends more on X_2 than X_1. Therefore X must be public information. ■

Proof of Theorem 4.5. By changing variables $a_i' = m_i a_i$, the game is equivalent to one where $m = (1, ..., 1)$ and all u_i have increasing differences in actions and in θ. In the continuation we will thus restrict attention to the case where $m = (1, ..., 1)$ and $c^i = (1, ..., 1)$.

For any signal profile X, let Γ_X denote the game with exogenous information structure X, and Γ the game with endogenous information. The main result in Van Zandt and Vives (2007) implies that there exists an increasing action strategy profile α such that in the game Γ_X, the profile α is a Nash-Bayesian equilibrium of Γ_X. Let α be such a profile. We will now show that the profile (X, α) is a Nash-Bayesian equilibrium of the game with endogenous information Γ.

Suppose by contradiction that (X_i', α_i') is a profitable deviation for player i in this game. Let α_i'' be a player i's best response to α_{-i} in the game $\Gamma_{X_i', X_{-i}}$. By Proposition 11 in Van Zandt and Vives (2007), the action strategy α_i'' is increasing. Since (X_i', α_i') is a profitable deviation for player i from profile (X, α) in Γ, it follows that (X_i', α_i'') is also a profitable deviation for player i from profile (X, α) in Γ. But, because X_i depends more on X_{-i} than

X_i', the same argument used in Theorem 4.4 implies that (X_i, α_i'') is an at least as good profitable deviation for player i from profile (X, α) in Γ. But this implies that α_i'' is a profitable deviation for player i from profile α in Γ_X, which contradicts the statement that α is Nash-Bayesian equilibrium of Γ_X. Therefore no player has any profitable deviation from (X, α) in Γ, the desired conclusion. ■

Proof of Theorem 4.6. Suppose by contradiction that (X_i', α_i') is a profitable deviation from (X, α) for player i in this game Γ. Let α_i'' be a pure m_i-monotonic action strategy that is a best response for player i's to α_{-i} in the game $\Gamma_{X_i', X_{-i}}$. Such an action strategy exists by assumption (iv) and because α is m-monotonic by assumption (iii). Since (X_i', α_i') is a profitable deviation for player i from profile (X, α) in Γ, it follows that (X_i', α_i'') is an even better deviation in Γ, and is therefore also a profitable deviation for player i from (X, α) in Γ. But, because X_i d^i-depends more on X_i than X_i', and because $(\alpha_i'', \alpha_{-i})$ is m-monotonic, the same argument used in Theorem 4.4 implies that (X_i, α_i'') is at least as good as (X_i', α_i''), and therefore at least as good as (X_i', α_i'). Therefore (X_i, α_i'') is a profitable deviation for player i from profile (X, α) in Γ. But this implies that α_i'' is a profitable deviation for player i from profile α in Γ_X, which contradicts the statement that α is Nash-Bayesian equilibrium of Γ_X. Therefore no player has any profitable deviation from (X, α) in Γ, the desired conclusion. ■

C.4 PQD and SPM Dependence

For any two random vectors $X = (X_1, ..., X_N)$ and $Y = (Y_1, ..., Y_N)$ with identical marginals, and respective cdfs F and G, we define the following dependence orderings.

We say that X is **at least as Positive Quadrant Dependent (PQD) as** Y if for all $x \in \mathbb{R}^N$, we have

$$F(x) \leq G(x).$$

A function $u : \mathbb{R}^N \to \mathbb{R}$ is said to be supermodular if for any $x, y \in \mathbb{R}^N$ it satisfies

$$u(x) + u(y) \leq u(x \wedge y) + u(x \vee y),$$

where the operators $\wedge$ and $\vee$ denote coordinate-wise minimum and maximum respectively.

We say that X is **at least as Supermodular Dependent (SPM) as** Y if

$$\mathbb{E}_X\left(u\left(X\right)\right) \geq \mathbb{E}_Y\left(u\left(Y\right)\right)$$

for all supermodular functions $u : \mathbb{R}^N \to \mathbb{R}$.

C.5 Mixed Strategies

The results obtained in Theorem 4.4 generalize to mixed strategies, but they imply very few restrictions for Nash-Bayesian equilibria where players play non degenerate mixed strategies. For example, consider a game with two players and two signals, with a fixed information structure such that both players observe each of the two signals with equal probabilities (independent draws). Suppose that this game admits a pure Nash-Bayesian equilibrium in action strategies (they could be pure or not).

Then it is easy to see that the game with an endogenous information structure admits a Nash-Bayesian equilibrium, where both players randomize with equal probabilities between the two signals. To see this, suppose that player 2 uses this strategy. From the point of view of the player 1, the two signals are then equally informative in a Blackwell sense on the vector (θ, α_2), which is all he cares about. It is then a best response for him to play this half half mixed strategy and the same argument holds for player 2. This phenomenon is more general. A symmetric fully mixed equilibrium exists, for any number of players, if and only if the Bayesian game where this structure is fixed admits a Nash-Bayesian equilibrium. What is important for the result is that there are only two signals in $\mathbb{X}$. A more general result can be obtained for a larger number of signals in $\mathbb{X}$, provided that some symmetry condition, which automatically holds in the case of two signals, is imposed on the signal structure.

Theorem C.2. *Let $N \geq 2$ and $\mathbb{X}_i = \{X_I, X_{II}\}$ for all $i \in I$. Consider the game with an exogenous information structure, where each player observes X_I or X_{II} with probability $1/2$ (independent draws across players). Suppose that this game admits a pure Nash-Bayesian equilibrium in action strategies (pure or not). Then this action profile and this information structure form a Nash-Bayesian equilibrium of the game Γ where the information structure is endogenous.*