THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 1 Capitole

Présentée et soutenue par

Willy LEFEZ

Le 22 juin 2021

Essays in Industrial Organization

Ecole doctorale : **TSE - Toulouse Sciences Economiques**

Spécialité : **Sciences Economiques - Toulouse**

Unité de recherche :
UMR 5314 -TSE-R

Thèse dirigée par
Bruno JULLIEN et Yassine LEFOUILI

Jury

M. Rolland STRAUSZ, Rapporteur
M. In-Uck PARK, Rapporteur
Mme Özlem BEDRE-DEFOLIE, Examinatrice
M. Bruno JULLIEN, Directeur de thèse
M. Yassine LEFOUILI, Co-directeur de thèse
M. Daniel GARRETT, Président

Essays in Industrial Organization

WILLY LEFEZ

PhD thesis in ECONOMICS

June 2021

IN-UCK PARK	Professor, University of Bristol	Rapporteur
ROLAND STRAUSS	Professor, Humboldt Universität	Rapporteur
ÖZLEM BEDRE-DEFOLIE	Associate Professor, ESMT Berlin	Examinatrice
DANIEL F. GARRETT	Professor, TSE	Examinateur
BRUNO JULLIEN	Research Director, CNRS, TSE	Directeur de thèse
YASSINE LEFOUILI	Associate Professor, TSE	Co-directeur de thèse

Acknowledgments

I am very grateful to Bruno Jullien and Yassine Lefouili who supervised my thesis. Their numerous advice guided me through my PhD. Their unconditional support and trust allowed me to overcome many of the difficulties a PhD student has. Their commitment and hard work are out of the norm.

I would like to express my sincere gratitude to In-Uck Park, Roland Strausz, Özlem Bedre-Defolie and Daniel F. Garrett who agreed to be part of the jury of my thesis.

I owe a lot to Esther who stood by my side during this journey. She shared much of her thoughtful advice, strength and love. I am deeply grateful for all her help in all aspects of my thesis.

I warmly thank my family for their constant support, my parents Patricia and Christophe, my brother Romain, his wife Alice and my two sisters Marie and Zoé. I also want to thank Esther's family who are always very welcoming and especially the two wonderful recent additions Marcelle and Athénaïs.

I give a special mention to my friends Lucas, Quentin and Yacine with whom Beamish tastes good and games of coinche have no end, which is great.

I also thank all my friends from Toulouse for their continuous support and interest. Sadly, I can't mention all of them, except Pite who insisted to be mentioned in the acknowledgment and Alex and Léa who helped me with my spiel.

Finally, I would like to thank Jacint, Jérôme and Miguel with whom I shared living space during these six years, movie nights and the wonderful conversations four PhD students in economics could have.

Contents

Introduction

1 Coordination in Collusion

The theory of repeated games captures the crucial trade-off that enables cooperation among privately motivated agents, namely whether long-run cooperative gains offset short-run gains from deviations. However, it fails to capture how agents coordinate to start cooperation. Agents may not a priori agree on cooperation schemes to play, especially since repeated games feature many equilibria; and agents may not know what cooperation schemes are the best. Understanding the coordination problem is crucial for collusion, since how firms coordinated is more important from a legal perspective than whether firms cooperated.

Two assumptions inhibit the coordination problem among firms: (i) the Nash equilibrium assumption, i.e. players anticipate each others' strategies, and (ii) common knowledge of the game which implies that players have common knowledge of the best collusive equilibria. Relaxing the first assumption and the problem of equilibrium selection has been studied in Harrington (2019). Athey and Bagwell (2001) and Athey, Bagwell, and Sanchirico (2004) relax the second assumption. In Athey and Bagwell (2001) for instance, firms have private information with respect to their marginal costs. To achieve the highest profits, colluding firms must allocate production to the lowest-cost firms of that period. However, all firms are willing to pretend being a low-cost firm and claim profits during this period. Under some conditions, firms are not be able to coordinate production efficiently, and all firms produce each period regardless of costs. In the same vein, Chapter 1 of this dissertation relaxes the second assumption and studies collusion when firms have private information on their discount factors. Firms with low discount factors value less long run collusive payoffs relative to undercutting profits and thus sustain lower prices in equilibrium. As a result, this assumption creates a coordination problem for firms in that they do not know a priori (and potentially disagree on) what is the best collusive scheme.

Relaxing either of the two assumptions creates a coordination problem that firms must solve before cooperating. However, in practice, many unmodelled factors provide a basis for firms to coordinate and agree on which collusive schemes to use. The history of detected cartels has shown that social norms, common business practices, personal contacts and connections and "focal points" are used by firms to start a collusive scheme.[1] Famously, in the 1950s, General Electric and Westinghouse even used a phases-of-the-moon system to assign low bids on electrical equipment. In some sense collusion does not take place in a vacuum, and taking all these factors in consideration diminishes the coordination problem that stems from relaxing the Nash equilibrium assumption. In contrast, these factors do not help firms to disclose truthfully private information. Communicating truthfully private information is costly, creates rents and distorts outcomes. Consequently, the coordination problem that stems from relaxing common knowledge seems more severe and robust. Chapter 1 shows that private information on the discount factor can explain a recurring pattern, observed in many discovered cartels, that prices increase gradually at the start of collusion.

2 Regulation of Data and the Allocation of Information

The regulation of data is vividly debated in the digital economy. Despite the rapidly growing literature on the topic, there is no consensus on how to model data nor on how to capture the extent of its effects on market outcomes and market structures. In Chapters 2 and 3,

[1] As defined in Schelling (1980).

I capture data as information structures, that is, as distributions of signals that reduce the uncertainty of the environment for the decision maker. Chapters 2 and 3 treat the question of the socially desirable use and collection of data as an issue of efficient allocation of information.

Many papers in the digital economy literature capture data as information. Yet, two different definitions of information are used. Information is often considered as a non-rival good that improves or expands agents' choices or production possibilities. In this framework, the efficient allocation of information is straightforward: "The cost of transporting a given body of information is frequently very low. If it were zero, then optimal allocation would obviously call for unlimited distribution of the information without cost."[2] In this view, data sharing is welfare improving and prevents, for instance, data bottlenecks or dominant market positions claimed by incumbents that have a larger access to data.

In contrast, Chapters 2 and 3 capture information as the decision maker's information structure or type in a Bayesian game. In this framework, information also expands the strategy set of the decision maker,[3] yet there are two crucial differences with the first definition. First, information structures cannot be seen as non-rival commodities in a strategic setting. Consider, for instance, the model of chapter 2 in which a platform, that has information about buyers, interacts with sellers. Suppose that a social planner shares the platform's buyer information with sellers. Formally the social planner discloses the realizations of the platform's signal (which affect the platform's beliefs about buyers) to sellers. Sellers observe these signal realizations, update their beliefs about buyers, and can adapt their prices accordingly. However, the platform's information structure has also changed in the process. The signal realizations that affect the platform's beliefs about buyers now also affect the platform's beliefs about sellers, and, in fact, these realizations also affect the sellers' beliefs about the platform. In a Bayesian game, changing the information structure (or the type) of one player changes the information structure of all players. Consequently, information

[2] See Arrow (2015)

[3] This is a consequence of Blackwell theorem.

cannot be transmitted in the same way non-rival commodities are traded. Second, more information for players does not always lead to better outcomes. While it is the case for decision problems under uncertainty, this is not true in a strategic setting. In Chapters 2 and 3 for instance, disclosing all buyer information to sellers is not welfare optimal. To achieve the efficient allocation of information, the social planner partially informs sellers about buyers, in a way that minimizes transaction inefficiencies.

In an environment where strategic interactions are prevalent, as it is the case between a platform and its users, defining data as information structures better captures the potential distortions regarding its use. This definition prompts two questions: (i) How to model information transmission (as it cannot be traded simply as a non-rival good), and (ii) How to determine the efficient allocation of information. In chapter 2 and 3, I use the framework of mechanism/information design to capture how a platform can transmit its information to sellers. Further, to determine whether the platform's use of data is efficient, I compare the platform mechanism with the welfare maximizing mechanism.

To capture distortions in data collection, I compare the platform's demand for data with the social planner's demand for data. In this environment, strategic interactions are less prevalent, and, therefore, data is captured as a commodity. To compute the platform's demand for data, i.e the marginal value of information structures, I use duality analysis. Duality analysis has often been used to study the allocation of goods in markets. In Chapters 2 and 3, I compare the platform's willingness to pay for data with to the social planner's one to determine distortions.

All in all, Chapters 2 and 3 use two frameworks to study the problem of allocation of information. Regarding the use of data, which entails strategic interactions, I use a mechanism design approach. Regarding the collection of data, in the context of a data market for instance, I use a duality analysis to compute its marginal value.

3 Summary of the Chapters

3.1 Summary of Chapter 1

Chapter 1 presents a new mechanism that explains the gradual increase in prices at the start of collusion, based on firms having private information about their discount factors. To elicit firms' true discount factors and allow patient firms to collude at the highest sustainable prices, optimal collusive strategies take advantage of the differences in time preferences across types. Patient firms delay the period during which they set the highest collusive prices to induce impatient firms to undercut early and reveal their types.

I consider an infinitely repeated Bertrand duopoly. In each period, firms set prices, and the lowest price firm serves all of the demand, which is composed of a unit mass of consumers with value v. In case of a tie, firms can choose to split the market nonstrategically. The competitive equilibrium yields 0 profits and is used in the dynamic game to punish deviations. Firm A's discount factor is high, δ_H, and is known by both firms. In contrast, firm B's discount factor is private information and is either high, δ_H, or low, $\delta_L < \delta_H$. Think of A as an incumbent firm with a known low cost of capital and B as an entrant for which access to capital is not common knowledge. A is unsure about the interest rate that B is facing and therefore does not know the highest collusive price that B can sustain in equilibrium.

I first characterize the Pareto frontier under complete information.[4] I assume that $\delta_L + \delta_H < 1$ so that no collusion is possible between B_L and A; whereas, I assume that $\delta_H \geq \frac{1}{2}$ so that, all prices are sustainable in equilibrium for B_H and A. As a result, if B is of type δ_H, all Pareto optimal equilibria involve the monopoly price v for every period.

Then, I assume that B's discount factor is private information, and A believes that B is of type δ_H with probability ρ_0. First, I study pooling equilibria, in which B_L and B_H set the same prices in each information set reached on the path, thus prices set on path must be sustainable for all types in particular for the least patient firm. I show pooling equilibria

[4] I use results of Harrington (1989) and Obara and Zincenko (2017), who analyzed this game under complete information.

yield 0 profits for all firms. Then, I study separating equilibria in pure strategies, that is, equilibria in which B_H's and B_L's do not set the same price on some information sets on path. The first time that this scenario occurs, A learns B's type, which is the separation period. After separation, A and B_H play the collusive path, whereas A and B_L cannot sustain collusive prices and play the competitive price. I show that, from the separation period, Pareto optimal profits for patient firms are constructed from collusive schemes that feature a transition phase, where for a finite number of periods, prices played by A and B_H increase gradually before reaching the monopoly price v. I compute the Pareto frontier for patient firms.

To do so, I first characterize all continuation profits achievable in a separating equilibrium using promise keeping, sustainability and incentive compatibility conditions, and I introduce a state variable to capture incentive compatibility in a recursive form. This step simplifies the optimization problem since it reduces the search space from the space of strategy profiles to the space of continuation profit sequences. Then, I characterize the Pareto frontier after and from the separation period. I show that B_L's profit is positively related to B_H's and A's profit; therefore, I only consider patient firms for the Pareto frontier. The main issue for patient firms regarding the implementation of collusive schemes is to prevent B_L from mimicking B_H. In the separation period, B_L undercuts the market price to obtain a small rent. During this period, the market price cannot be too high since A expects future collusive profit only with probability ρ_0. Therefore, prices cannot increase too rapidly on the patient firms collusive path; otherwise, B_L would mimic B_H to undercut later at a higher price. I show that prices cannot grow at a rate larger than $\frac{1}{\delta_L}$.

This model captures the transition phase as a way for firms to elicit the true discount factors of their rivals and to collude at the highest sustainable prices. It applies to both tacit and explicit collusion in the sense that cheap talk or direct forms of communication do not restore complete information since one type strictly benefits from mimicking the other.

3.2 Summary of Chapter 2

Chapter 2 studies how a platform can affect its sellers' pricing decisions by using price recommendations. I investigate whether platforms design price recommendations and collect buyer information used for this purpose in a socially desirable way. Price recommendations allow platforms to communicate strategically their buyer information to sellers so as to influence the outcome of the market. As the best transaction price for platforms may not be the same as the preferred transaction price for sellers,[5] the platform benefits from price recommendations. I identify potential distortions in the collection and use of personal data and argue that the extent to which a platform's incentives are (mis)aligned with social incentives crucially depends on its business model.

I consider two business models: "paid" and "free" platforms. Paid platforms charge participation fees on both the buyer and the seller side. Free platforms provide free access to buyers and charge a participation fee on the seller side. Under both business models, the platform draws informative signals about buyers' valuations and correlates these signals with price recommendations to influence sellers' pricing decisions. Joining sellers receive a price recommendation and set a price for their good while joining buyers observe their value of the good and their matching seller's price and then purchase the good or not.

I determine the most profitable way for a platform to influence seller prices through price recommendations, and compare it to the socially optimal way of influencing seller prices. I then use the results to discuss the value of data for the platform and resulting incentives to collect data.

First, I show that paid platforms use data efficiently, whereas free platforms use data inefficiently. As often in platform models, paid platforms extract the entire surplus of the interaction via entry fees.[6] This implies that paid platforms use data to maximize surplus per trade. On the contrary, free platforms only make profits on the seller side. They use

[5] Sellers are typically too small in these markets to take into account network externalities.

[6] See e.g. Rochet and Tirole (2006).

data to help sellers to extract buyer surplus, which may destroy surplus relative to a no-data benchmark. These two different uses of information by paid platforms and free platforms imply different incentives to collect information.

To capture the platform's incentives to collect data, I compute the platform's marginal value of information; in other words, by how much the platform's profit changes when their information structures marginally change. In my model, information is an input of price recommendations. Using sensitivity analysis, I characterize paid and free platforms' willingness to pay for information by computing the shadow price of this input. The marginal value of information provides rich comparative statics on information structures which copes with the high dimensionality of this input.

Then, I compare the platform's willingness to pay to collect information with the social planner's and identify several distortions. I show that paid platforms, despite using information efficiently, under-value any additional information. Paid platforms set inefficiently high entry fees, which implies that learning impacts less trades than under the social planner's trade mechanism. Therefore, the paid platforms' willingness to collect information is proportionally lower than the social planner's by a factor that only depends on the elasticity of demand, which can be estimated empirically. By contrast, free platforms have a biased demand for information. Free platforms value more learning about mark up opportunities than learning about trade opportunities, although the latter is the efficient way to learn. The bias holds regardless of the source of information and, therefore, free platforms' incentives to collect data are inefficiently oriented.

3.3 Summary of Chapter 2

The third chapter is an extension of Chapter 2 to the case of competition between platforms. It shows that distortions in platforms' incentives to collect data are mitigated as the degree of competition between platforms increases.

In this chapter, I consider two competing platforms that intermediate trade between buyers and sellers. Platforms charge entry fees to users on both sides of the market (compared to Chapter 2, I only study the "paid" platform business model). Each platform draws informative signals about buyers' valuations and correlates these signals with price recommendations to influence sellers' pricing decisions. Sellers, who can multi-home, receive a price recommendation when joining a platform and set a price for their good. On the buyer side, platforms are located at both ends of a Hotelling segment. Buyers, that are uniformly distributed over the segment, choose which platform to join, if any, and incur a linear transportation cost to join a platform. As in Bénabou and Tirole (2016), I assume that buyers have two outside options located at both ends of the Hotelling segment. This assumption allows the transportation cost to only determine the degree of competition between firms (i.e. how many market shares a platform gains by reducing the buyer entry fee) and not market participation (the trade-off between joining a platform or collecting the outside option payoff). As a result, the transportation cost is identified to the inverse of the degree of competition between platforms as it impacts the demand elasticites within the market but not outside the market. Furthermore, the "co-located" outside options version of the Hotelling model is better suited for welfare analysis than the standard version.[7]

First, I study the competitive equilibrium in which platforms set user entry fees and a price recommendation rule. In equilibrium, platforms design the price recommendation rule that maximizes the surplus per transaction to attract as many users on both sides, and set entry fees to generate profit and compete for market shares. Compared to the efficient outcome, however, the equilibrium user fees are too high and, as a result, the mass of transactions in equilibrium is inefficiently low. Increasing the degree of competition (i.e. reducing the transportation cost) induces platforms to charge lower entry fees in equilibrium, which increases market participation and welfare.

Second, I capture platforms' incentives to collect data by computing their marginal value

[7]Assuming that the market is covered, the welfare analysis in a standard Hotelling model is limited to minimizing the total transportation cost of the economy. See Bénabou and Tirole (2016) that discusses this assumption.

for buyer information. Platforms' signals about buyers' valuations are the inputs of the price recommendations. I compute the change in platforms' equilibrium profits when changing marginally their distributions of signals. Platforms value additional information as it improves their price recommendations, the surplus per transaction and therefore the mass of users platforms' attract. However, since less users join platforms under a competitive equilibrium than under the efficient outcome, additional information for platforms benefits less transactions. As a result, platforms have a lower marginal value for information than what is socially desirable. Furthermore, I consider a benevolent information provider that maximizes welfare by choosing the platforms' information structures but not their trade mechanisms. I show that conditional on the mass of users joining platforms the benevolent information provider's marginal value for information is larger than the platforms' one. The benevolent information provider values the increase in welfare coming from the readjustment of user entry fees by platforms when changing their information structures. Consequently, platforms undervalue additional information compared to what is socially optimal which suggests that platforms undercollect data. However, increasing the degree of competition increases the marginal value of information for platforms. In a more competitive market, each improvement of the price recommendations allows a platform to increase the surplus per transaction from which the platform gains market shares at a higher rate. Therefore, increasing the degree of competition reduces the distortions in platforms' incentives to collect data.

Bibliography

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990). “Toward a theory of discounted repeated games with imperfect monitoring”. In: *Econometrica: Journal of the Econometric Society*, pp. 1041–1063.

Acemoglu, Daron et al. (2019). *Too much data: Prices and inefficiencies in data markets*. Tech. rep. National Bureau of Economic Research.

Aghadadashli, Hamid and Patrick Legros (2020). “Let’s Collude”. In:

Andreoni, James, Michael A Kuhn, and Larry Samuelson (2019). “Building rational cooperation on their own: Learning to start small”. In: *Journal of Public Economic Theory* 21(5), pp. 812–825.

Armstrong, Mark (2006). “Competition in two-sided markets”. In: *The RAND Journal of Economics* 37(3), pp. 668–691.

Arrow, Kenneth J (2015). *Economic welfare and the allocation of resources for invention*. Princeton University Press.

Athey, Susan and Kyle Bagwell (2001). “Optimal collusion with private information”. In: *RAND Journal of Economics*, pp. 428–465.

Athey, Susan, Kyle Bagwell, and Chris Sanchirico (2004). “Collusion and price rigidity”. In: *The Review of Economic Studies* 71(2), pp. 317–349.

Aumann, Robert J (1987). “Correlated equilibrium as an expression of Bayesian rationality”. In: *Econometrica: Journal of the Econometric Society*, pp. 1–18.

Barvinok, Alexander (2002). *A course in convexity*. Vol. 54. American Mathematical Soc.

Bénabou, Roland and Jean Tirole (2016). "Bonus culture: Competitive pay, screening, and multitasking". In: *Journal of Political Economy* 124(2), pp. 305–370.

Bergemann, Dirk and Alessandro Bonatti (2015). "Selling cookies". In: *American Economic Journal: Microeconomics* 7(3), pp. 259–94.

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2020). "The economics of social data". In:

Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin (2018). "The design and price of information". In: *American Economic Review* 108(1), pp. 1–48.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris (2015). "The limits of price discrimination". In: *American Economic Review* 105(3), pp. 921–57.

Bergemann, Dirk, Tibor Heumann, and Stephen Morris (2020). "Information, market power and price volatility". In:

Bergemann, Dirk and Stephen Morris (2016). "Bayes correlated equilibrium and the comparison of information structures in games". In: *Theoretical Economics* 11(2), pp. 487–522.

Blackwell, David (1953). "Equivalent comparisons of experiments". In: *The annals of mathematical statistics*, pp. 265–272.

Blonski, Matthias and Daniel André Probst (2001). *The emergence of trust*. University of Mannheim.

Bonatti, Alessandro and Gonzalo Cisternas (2020). "Consumer scores and price discrimination". In: *The Review of Economic Studies* 87(2), pp. 750–791.

Bos, Iwan, Wilko Letterie, and Dries Vermeulen (2015). "Antitrust as facilitating factor for collusion". In: *The BE Journal of Economic Analysis & Policy* 15(2), pp. 797–814.

Bourreau, Marc, Bernard Caillaud, and Romain De Nijs (2018). "Taxation of a digital monopoly platform". In: *Journal of Public Economic Theory* 20(1), pp. 40–51.

Bourreau, Marc and Germain Gaudin (2018). "Streaming platform and strategic recommendation bias". In:

Caffarra, Cristina et al. (2020). "Designing regulation for digital platforms: Why economists need to work on business models". In: *VoxEU. Available from URL: https://voxeu.org/article/designing-regulation-digital-platforms*.

Caillaud, Bernard and Bruno Jullien (2003). "Chicken & egg: Competition among intermediation service providers". In: *RAND journal of Economics*, pp. 309–328.

Calzolari, Giacomo and Alessandro Pavan (2006). "Monopoly with resale". In: *The RAND Journal of Economics* 37(2), pp. 362–375.

Carroni, Elias, Guiseppe Pignataro, and Alessandro Tampieri (2020). "Recommendation Service as a Coordination Device among Peers". In:

Colombo, Luca, Gianluca Femminis, and Alessandro Pavan (2014). "Information acquisition and welfare". In: *The Review of Economic Studies* 81(4), pp. 1438–1483.

Crémer, Jacques and Richard P McLean (1988). "Full extraction of the surplus in Bayesian and dominant strategy auctions". In: *Econometrica: Journal of the Econometric Society*, pp. 1247–1257.

Crémer, Jacques, Yves-Alexandre de Montjoye, and Heike Schweitzer (2019). "Competition policy for the digital era". In: *Report for the European Commission*.

De Corniere, Alexandre and Greg Taylor (2020). "Data and Competition: a General Framework with Applications to Mergers, Market Structure, and Privacy Policy". In:

Dimakopoulos, Philipp D and Slobodan Sudaric (2018). "Privacy and platform competition". In: *International Journal of Industrial Organization* 61, pp. 686–713.

Dizdar, Deniz and Eugen Kováč (2020). "A simple proof of strong duality in the linear persuasion problem". In: *Games and Economic Behavior*.

Dworczak, Piotr (2016). "Mechanism Design with Aftermarkets: On the Impossibility of Pure Information Intermediation". In:

Dworczak, Piotr (2020). "Mechanism design with aftermarkets: Cutoff mechanisms". In: *Available at SSRN 2859206*.

Dworczak, Piotr and Anton Kolotilin (2019). “The persuasion duality”. In: *Available at SSRN 3474376.*

Dworczak, Piotr and Giorgio Martini (2019). “The simple economics of optimal persuasion”. In: *Journal of Political Economy* 127(5), pp. 1993–2048.

Fleder, Daniel M and Kartik Hosanagar (2007). “Recommender systems and their impact on sales diversity”. In: *Proceedings of the 8th ACM conference on Electronic commerce*, pp. 192–199.

Furman, Jason et al. (2019). “Unlocking digital competition: Report of the digital competition expert panel”. In: *UK government publication, HM Treasury.*

Galperti, Simone and Jacopo Perego (2018). “A dual perspective on information design”. In: *Available at SSRN 3297406.*

Gomes, Renato and Alessandro Pavan (2019). “Price customization and targeting in matching markets”. In:

Hagiu, Andrei and Hanna Hałaburda (2014). “Information and two-sided platform profits”. In: *International Journal of Industrial Organization* 34, pp. 25–35.

Hagiu, Andrei and Bruno Jullien (2014). “Search diversion and platform competition”. In: *International Journal of Industrial Organization* 33, pp. 48–60.

Hagiu, Andrei and Julian Wright (2020). *Data-enabled learning, network effects and competitive advantage.* Tech. rep. Working Paper.

Harrington, Joseph Emmett (2006). *How do cartels operate?* Vol. 5. Now Publishers Inc.

Harrington Jr, Joseph E (1989). “Collusion among asymmetric firms: The case of different discount factors”. In: *International Journal of Industrial Organization* 7(2), pp. 289–307.

Harrington Jr, Joseph E (2017). “A theory of collusion with partial mutual understanding”. In: *Research in Economics* 71(1), pp. 140–158.

Harrington Jr, Joseph E and Joe Chen (2006). “Cartel pricing dynamics with cost variability and endogenous buyer detection”. In: *International Journal of Industrial Organization* 24(6), pp. 1185–1212.

Harrington Jr, Joseph E, Roberto Hernan Gonzalez, and Praveen Kujal (2016). "The relative efficacy of price announcements and express communication for collusion: Experimental findings". In: *Journal of Economic Behavior & Organization* 128, pp. 251–264.

Harrington Jr, Joseph E and Wei Zhao (2012). "Signaling and tacit collusion in an infinitely repeated Prisoners' Dilemma". In: *Mathematical Social Sciences* 64(3), pp. 277–289.

Hua, Xiameng and Joel Watson (2021). "Starting Small and Renegotiation in Discrete-Time Relationships with a Continuum of Types". In:

Johnson, Justin, Andrew Rhodes, and R. Matthijs Wildenbeest (2020). "Platform Design When Sellers Use Pricing Algorithms". In:

Jullien, Bruno and Alessandro Pavan (2019). "Information management and pricing in platform markets". In: *The Review of Economic Studies* 86(4), pp. 1666–1703.

Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian persuasion". In: *American Economic Review* 101(6), pp. 2590–2615.

Kartal, Melis (2018). "Honest equilibria in reputation games: The role of time preferences". In: *American Economic Journal: Microeconomics* 10(1), pp. 278–314.

Kolotilin, Anton (2014). "Optimal information disclosure: Quantity vs. quality". In: *UNSW Australian School of Business Research Paper*(2013-19).

Kolotilin, Anton (2018). "Optimal information disclosure: A linear programming approach". In: *Theoretical Economics* 13(2), pp. 607–635.

Kolotilin, Anton et al. (2017). "Persuasion of a privately informed receiver". In: *Econometrica* 85(6), pp. 1949–1964.

Le Cam, Lucien (1996). "Comparison of experiments: A short review". In: *Lecture Notes-Monograph Series*, pp. 127–138.

Milgrom, Paul and Ilya Segal (2002). "Envelope theorems for arbitrary choice sets". In: *Econometrica* 70(2), pp. 583–601.

Montes, Rodrigo, Wilfried Sand-Zantman, and Tommaso Valletti (2019). “The value of personal information in online markets with endogenous privacy”. In: *Management Science* 65(3), pp. 1342–1362.

Mouraviev, Igor (2013). “Explicit collusion under antitrust enforcement”. In: *Institute of Mathematical Economics Working Paper*(494).

Myerson, Roger B (1982). “Optimal coordination mechanisms in generalized principal–agent problems”. In: *Journal of mathematical economics* 10(1), pp. 67–81.

Obara, Ichiro and Federico Zincenko (2017). “Collusion and heterogeneity of firms”. In: *The RAND Journal of Economics* 48(1), pp. 230–249.

Oliveira, Henrique de (2018). “Blackwell’s informativeness theorem using diagrams”. In: *Games and Economic Behavior* 109, pp. 126–131.

Pathak, Bhavik et al. (2010). “Empirical analysis of the impact of recommender systems on sales”. In: *Journal of Management Information Systems* 27(2), pp. 159–188.

Pavan, Alessandro, Xavier Vives, et al. (2015). “Information, coordination, and market frictions: An introduction”. In: *Journal of Economic Theory* 158(Part B), pp. 407–426.

Pavlov, Vladimir and Ron Berman (2019). *Price Manipulation in Peer-to-Peer Markets*. Tech. rep. Working paper.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz (2019). “The cost of information”. In:

Rauch, James E and Joel Watson (2003). “Starting small in an unfamiliar environment”. In: *International Journal of industrial organization* 21(7), pp. 1021–1042.

Roberts, Gilbert and James S Renwick (2003). “The development of cooperative relationships: an experiment”. In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(1530), pp. 2279–2283.

Rochet, Jean-Charles and Jean Tirole (2003). “Platform competition in two-sided markets”. In: *Journal of the european economic association* 1(4), pp. 990–1029.

Rochet, Jean-Charles and Jean Tirole (2006). “Two-sided markets: a progress report”. In: *The RAND journal of economics* 37(3), pp. 645–667.

Rothschild, Michael and Joseph E Stiglitz (1970). "Increasing risk: I. A definition". In: *Journal of Economic theory* 2(3), pp. 225–243.

Schelling, Thomas C (1980). *The Strategy of Conflict: With a New Preface by The Author.* Harvard university press.

Scott Morton, Fiona et al. (2019). "Committee for the study of digital platforms: Market structure and antitrust subcommittee report". In: *Chicago: Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business.*

Stahl, Dale O (1989). "Oligopolistic pricing with sequential consumer search". In: *The American Economic Review*, pp. 700–712.

Taneva, Ina (2019). "Information design". In: *American Economic Journal: Microeconomics* 11(4), pp. 151–85.

Teh, Tat-How and Julian Wright (2020). "Intermediation and steering: Competition in prices and commissions". In: *American Economic Journal: Microeconomics.*

Watson, Joel (1999). "Starting small and renegotiation". In: *Journal of economic Theory* 85(1), pp. 52–90.

Watson, Joel (2002). "Starting small and commitment". In: *Games and Economic Behavior* 38(1), pp. 176–199.

Weyl, E Glen (2010). "A price theory of multi-sided platforms". In: *American Economic Review* 100(4), pp. 1642–72.

Yang, Kai Hao (2020). *Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences.* Tech. rep. Working Paper.

Ye, Maoliang et al. (2020). "One Step at a Time: Does Gradualism Build Coordination?" In: *Management Science* 66(1), pp. 113–129.

Chapter 1

Collusion Under Incomplete Information on the Discount Factor

Abstract. The gradual increase in prices at the start of collusion is a recurrent pattern that has been observed in many discovered cartels. When firms have private information regarding their respective discount factors, I show that optimal collusive schemes feature a transition phase during which prices increase gradually. Impatient firms, for which sustainable collusive prices are low, are willing to mimic patient firms to undercut them at a high market price. To elicit firms' true discount factors, optimal collusive strategies take advantage of the differences in time preferences across types. Patient firms delay the period during which the highest collusive prices are set to force impatient firms to undercut early and reveal their type. In addition, patient firms find it optimal to reach the highest collusive prices by employing a gradual price path. I characterize the Pareto payoff frontier and compute the optimal speed of price increases during the transition phase.

Keywords : Collusion, incomplete information, transition phase.

JEL classification : C73, D43, D82.

1 Introduction

Identifying patterns in collusive agreements allows competition authorities to detect and prosecute illegal practices. A well-recognized pattern is the gradual increase in prices during a cartel's formation.

Many discovered cartels have been explicit about gradually raising prices at the start of their collusive agreements. At a meeting of the choline chloride cartel on November 16, 1992, members agreed about sequences of gradual price increases for several of their products.[1] In the case of the carbonless paper cartel, at a general cartel meeting, "it was agreed that the price would be increased in two stages on 1 July and on 1 September 1994, both times by 5%."[2] Similarly, the vitamin cartel raised the price of vitamins A and E in increments of 5%, and a gradual price increase also happened for vitamin B1: " From 1991 until about 1993, the price of vitamin B1 was gradually increased by the cartel. In 1991, the producers raised the market price from below DEM 65 to DEM 68/kg."[3] This "transition phase", during which prices increase gradually, has also been observed in laboratory experiments on collusion. In Kujal, Harrington and Hernan-Gonzalez (2013), for instance, in symmetric duopoly groups in which participants set prices each period without communication, prices increase gradually before reaching a high supra-competitive level.[4] This widespread pattern of collusive agreements echoes findings in the experimental literature on cooperation that suggest that "start small" or "raising-the-stakes" strategies promote cooperation.[5] In fact, it is common wisdom that "if a number of preparatory bargains can be struck on a small scale, each may be willing to risk small investment to create the tradition of trust."[6] However, this pattern lacks game theoretic foundation, and in standard collusion models, the highest collusive price is typically set from the start of the agreement.

[1] See table 1 in the appendix and Harrington (2006)

[2] EC report Carbonless paper cartel 211.

[3] EC report vitamins cartel, 182 and 255.

[4] See table 1 in the appendix.

[5] In these strategies, participants gradually increase their levels of investment in the partnership. See, e.g., Roberts and Renwick (2003), Andreoni, Kuhn, and Samuelson (2019) or Ye et al. (2020).

[6] See Schelling (1980)

My paper presents a new mechanism that explains the gradual increase in prices at the start of collusion, based on firms having private information about their discount factors. To elicit firms' true discount factors and allow patient firms to collude at the highest sustainable prices, optimal collusive strategies take advantage of the differences in time preferences across types. Patient firms delay the period during which they set the highest collusive prices to induce impatient firms to undercut early and reveal their types. Delay is necessary to enable separation. Furthermore, it is optimal for patient firms to reach the highest collusive prices by employing a gradually increasing price path. In other words, the "transition phase" is a distortion of first-best collusive schemes that enables firms to screen each others' discount factors to collude at the highest sustainable prices.

To be sure, prices might be raised gradually for several reasons. Small price increases can accommodate buyers and prevent buyer resistance or simply might be a response to an increase in costs. In Chen and Harrington (2006), the transition phase allows firms to reduce the probability of buyer detection and to avoid suspicion. This paper studies an alternative and complementary rationale based on the discount factor – the parameter that critically determines the prices that are sustainable in a collusive agreements. The discount factor is a composite parameter that compounds many interpretations, and there are many reasons why it can vary across firms and might not be common knowledge. First, firms' access to and cost of capital can differ. For instance, a well-established firm could have low financing costs, whereas an entrant firm might face higher interest rates, and such interest rates are private information to the firm and the bank/shareholders. Second, the incentives of the managers, which are in practice the instigators of collusive agreements, can vary drastically across firms.[7] The private contract binding shareholders to the manager is affected by many factors: the level of bonuses, whether shareholders are short-sighted investors, information asymmetries, etc. Third, the discount factor also reflects the general belief in the sustainability of the market. Consider a market in which new innovations often replace

[7] See Aghadadashli and Legros (2020).

older generations of products (smartphones, computers, video game consoles, etc.). In these markets, some firms might invest more in R&D or be more optimistic about the launch of a new innovation in the next period, resulting in firms having private and asymmetric beliefs about whether current products will continue to sell.

Direct forms of communication cannot always restore complete information. Indeed, the collusive profits at the highest sustainable price for impatient firms is lower than the deviation profit at the highest sustainable price for patient firms.[8] Therefore, impatient firms are typically willing to mimic patient firms to undercut at high prices. Furthermore, in this model, I show that low type firms are better off under incomplete information. As a result, firms are unable to truthfully communicate their discount factors at cartel meetings or disclose them publicly via other means.[9] All in all, at the start of collusive schemes, firms are unsure about the highest sustainable price of the cartel. Facing this problem, this paper shows that it is Pareto optimal for firms to start collusion with a gradual increase in prices to elicit this information.

I consider an infinitely repeated Bertrand duopoly. In each period, firms set prices, and the lowest price firm serves all of the demand, which is composed of a unit mass of consumers with value v. In case of a tie, firms can choose to split the market nonstrategically.[10] The competitive equilibrium yields 0 profits and is used in the dynamic game to punish deviations. Firm A's discount factor is high, δ_H, and is known by both firms. In contrast, firm B's discount factor is private information and is either high, δ_H, or low, $\delta_L < \delta_H$. Think of A as an incumbent firm with a known low cost of capital and B as an entrant for which access to capital is not common knowledge. A is unsure about the interest rates that B is facing and therefore does not know the highest collusive price that B can sustain in equilibrium.

[8] By construction, the highest price that impatient firms can sustain is such that the deviation profit equals the continuation collusive profit. Hence, for any price greater than that level, the deviation profit outweighs the continuation collusive profit.

[9] Firms cannot reliably cheap talk their private discount factors to others. However, if cartel meetings are costly (for instance, if meetings increase the probability of detection), then they can be used as signaling devices for firms to communicate private information; see Mouraviev (2013); Bos, Letterie, and Vermeulen (2015).

[10] See Harrington Jr (1989) and Obara and Zincenko (2017).

Throughout the paper, I focus on Bayesian perfect equilibria in pure strategies.

In the first part of the paper, I characterize the Pareto frontier under complete information. I build on the results of Harrington (1989) and Obara and Zincenko (2017), who analyzed this game under complete information. I assume that $\delta_L + \delta_H < 1$ so that, if B is of type δ_L, no collusion is possible with A. In contrast, I assume that $\delta_H \geq \frac{1}{2}$ so that, if B is of type δ_H, all prices up to the monopoly price v are sustainable in equilibrium. If B is of type δ_H, all Pareto optimal equilibria involve the monopoly price v for every period.

In the second part of the paper, I assume that B's discount factor is private information, and A believes that B is of type δ_H with probability ρ_0. First, I show that pooling equilibria, in which B_L and B_H set the same prices in each information set reached on the path, are sub-optimal. Prices set on a pooling equilibrium path must be sustainable for all types and so are restricted by the least patient firm. In this case, pooling equilibria yield 0 profits for all firms since collusion is unsustainable for A and B_L. Then, I study separating equilibria in pure strategies, that is, equilibria in which B_H's and B_L's do not set the same price on some information sets on path. The first time that this scenario occurs, A learns B's type, which is the separation period. After separation, A and B_H play the collusive path, whereas A and B_L cannot sustain collusive prices and play the competitive price. I show that, from the separation period, Pareto optimal profits for patient firms are constructed from collusive schemes that feature a transition phase, where for a finite number of periods, prices played by A and B_H increase gradually before reaching the monopoly price v. I characterize the speed of price increases during the transition phase, and I compute the Pareto frontier for patient firms.

To do so, I first characterize all continuation profits achievable in a separating equilibrium using promise keeping, sustainability and incentive compatibility conditions, and I introduce a state variable to capture incentive compatibility in a recursive form. This step simplifies the optimization problem since it reduces the search space from the space of strategy profiles to the space of continuation profit sequences. Then, I characterize the Pareto frontier after and from the separation period. I show that B_L's profit is positively related to B_H's and A's profit; therefore, I only consider patient firms for the Pareto frontier. The main issue for patient firms regarding the implementation of collusive schemes is to prevent B_L from mimicking B_H. In the separation period, B_L undercuts the market price to obtain a small rent. During this period, the market price cannot be too high since A expects future collusive profit only with probability ρ_0. As a result, prices cannot increase too rapidly on the patient firms collusive path; otherwise, B_L would mimic B_H to undercut later at a higher price. The speed of the price increase is therefore limited by the impatient firms' discount factors. Specifically, proposition 4.2 bounds the geometric speed of the price increase to less than $\frac{1}{\delta_L}$.

My model captures the transition phase as a way for firms to elicit the true discount factors of their rivals and to collude at the highest sustainable prices. This mechanism applies to both tacit and explicit collusion in the sense that cheap talk does not restore complete information since one type strictly benefits from mimicking the other.

Related Literature

This paper relates to the literature on collusion. The complete information game has been studied for stationary strategies in Harrington (1989) and for general strategies in Obara and Zincenko (2017). Collusion under private information on the discount factor has been studied in Harrington and Zhao (2012); Bos, Letterie, and Vermeulen (2015); and Aghadadashli and Legros (2020). Harrington and Zhao (2012) and Bos, Letterie, and Vermeulen (2015) both studied a repeated prisoners dilemma game. In Bos, Letterie, and Vermeulen (2015), firms exploit the cost of explicit collusion, related to the existence of antitrust laws, as a means to signal their discount factors. Harrington Jr and Zhao (2012) studied equilibria in which firms gradually learn each other's discount factors. In contrast, I only restrict the strategies of firms by assuming pure strategies, and my paper focuses on the gradual increase in prices and not on beliefs. Aghadadashli and Legros (2020) studied collusion among managers with private information about their discount factors. The authors analyzed how antitrust fines affect the ability of managers to truthfully communicate their types.

In a different framework, Harrington and Chen (2006) and Harrington (2017) presented mechanisms that capture the gradual increase in prices at the start of collusion. In Harrington and Chen (2006), small price increases reduce buyers' suspicion and the probability of the cartel being detected. Harrington (2017) relaxed the Nash equilibrium assumption that firms anticipate their rivals' strategies. In each period, firms observe each other prices and gradually learn each other strategies, generating a transition phase in the class of strategies studied. In contrast, my paper uses standard Nash-related equilibrium concepts but relaxes the common knowledge assumption about the discount factor.

My paper also relates to the broader literature on cooperation, particularly papers that study gradualism as a means to promote cooperation; see, e.g., Watson (1999); Blonski and Probst (2001); Watson (2002); Rauch and Watson (2003); and Hua and Watson (2021). The closest paper to mine is Kartal (2018). The authors studied a repeated principal agent relationship under limited commitments, in which the principal's discount factor is private information. The paper shows that a patient principal's optimal contract features a gradual increase in the agent's effort levels and that this transition phase is finite. Similarly, my paper captures the gradual increase in prices as a feature of optimal collusive schemes for patient firms, and I also characterize the optimal speed of price increases.

The paper is organized as follows. Section 2 presents the model. Section 3 characterizes the Pareto frontier of the game under complete information. Section 4 analyzes the game under incomplete information, characterizes the Pareto frontier for patient firms from the separation period and provides the properties of the transition phase. Section 5 concludes.

2 The Model

Firms A and B meet in periods $t = 0, 1, ..., \infty$. Firms A and B engage every period in Bertrand competition in a homogeneous-good market with no costs. Demand is inelastic, and there is a unit mass of identical consumers with value v, where $v > 0$. In a period t, firms A and B simultaneously select prices $p_{A,t}$ and $p_{B,t}$ in $[0, v]$. The firm with the lower price serves the market. In period t, if $p_{A,t} = p_{B,t}$, then firms can split the market nonstrategically, where $\alpha_t \in [0, 1]$ denotes A's market share, and $(1 - \alpha_t)$ denotes B's market share.[11] I assume that, for all price $p_k \in (0, v]$ $k \in A, B$, there exists an optimal undercutting price p_k^u that yields profit p_k for the undercutting firm.[12]

Information Structure. Firm B's private discount factor is either low, δ_L, with probability $1 - \rho_0$, or high, δ_H, with probability ρ_0, where $\delta_H > \delta_L$ are both in $(0, 1)$. Firm A's discount factor is $\delta_H \in (0, 1)$ and is publicly known. Firm A has a prior belief $\rho_0 \in (0, 1)$ on firm B's being of type δ_H. I assume that $\delta_L + \delta_H < 1$ and that $\delta_H \geq \frac{1}{2}$. Section 3 demonstrates that these two assumptions imply that no collusion is possible if B is impatient[13], and if B is patient, collusion is possible at the monopoly price v.

Notations. Given a path of play $\{p_{A,t}, p_{B,t}, \alpha_t\}_{t \geq 0}$, $\pi_{A,t}$ denotes the current profit of firm A in period t, which is equal to $p_{A,t}\mathbf{1}_{\{p_{A,t} < p_{B,t}\}} + \mathbf{1}_{\{p_{A,t} = p_{B,t}\}}\alpha_t p_{A,t}$, and $\Pi_{A,t}$ denotes the continuation profit of firm A from period t onward, which is equal to $(1 - \delta_H)\sum_{s \geq t} \delta_H^s \pi_{A,s}$. Similarly, $\pi_{B,t}$ is the current profit of firm B in period t and is equal to $\pi_{B,t}\mathbf{1}_{\{p_{B,t} < p_{A,t}\}} + (1 - \alpha_t)p_{B,t}\mathbf{1}_{\{p_{B,t} = p_{A,t}\}}$, and $\Pi_{B,t}$ is the continuation profit of firm B from period t onward and is equal to $(1 - \delta_B)\sum_{s \geq t} \delta_B^s \pi_{B,s}$. The subscript L (resp. H) refers to firm B of type δ_L (resp. δ_H).

[11] Harrington Jr (1989) and Obara and Zincenko (2017) used the same assumption. In an asymmetric context, allowing firms to share the market unequally improves the tractability of the model.

[12] This assumption facilitates the analysis of separating equilibria of the incomplete information game and has no impact on the complete information game analysis.

[13] There is a unique SPE, in which A and B_L play the NE in all histories.

Strategies and Equilibrium Concepts. In the complete information repeated game, a firm's strategy maps the set of histories to the set of prices. Payoffs are the discounted by the sum of each period's profit. In section 3, I focus on subgame perfect equilibria (SPE).

In the incomplete information dynamic game, a strategy for firm B maps the set of histories and types to the set of prices. Firm B's payoff is the discounted sum of profits given B's type. Firm A's strategy maps the set of histories to the set of prices. Firm A's payoff is the expected sum of discounted profits given its prior belief ρ_0 about firm B's type.

I focus on perfect Bayesian equilibria (PBE) in pure strategies, consisting of a sequentially rational profile of strategies for each firm, together with consistent beliefs for firm A. Consistent beliefs about firm B's type are updated according to Bayes' rule in all information sets reached with positive probability. A strategy is sequentially rational if it maximizes the players payoff at each information set.

The stage game has a unique Nash equilibrium (NE) that yields $(0, 0)$ for both firms. The NE has two crucial properties: (i) it is an equilibrium regardless of firms' types or beliefs; and (ii) it yields the minmax payoff for each firm regardless of its type or belief. Therefore, in the dynamic game, punishing any detectable deviation by playing the NE forever provides the best incentives to both firms and types and is a continuation equilibrium itself, regardless of firms' beliefs and types.[14] Consequently, there is no need to refine beliefs off path (for information sets reached with 0 probability).

3 Collusion under complete information

This section characterizes the Pareto optimal equilibrium payoffs for the complete information game. A version of this game was analyzed in Harrington (1989) for stationary strategies and in Obara and Zincenko (2017) for general strategies. First, I characterize the sequences

[14]Not all deviations are observable by all firms; in particular, when one firm mimics another type's strategy, this deviation is not observed by rival firms.

of continuation profits achievable in equilibrium. Then, I determine the Pareto optimal equilibrium profits for firms A and B. Importantly, in all Pareto optimal equilibria, firms set the monopoly price v for every period. That is, the gradual increase inf prices is specific to the incomplete information game. For the next subsection, firm A's (resp. B's) discount factor is $\delta_A \in (0, 1)$ (resp. δ_B).

Equilibrium Profits

This subsection characterizes all sequences of continuation profits $\{\Pi_{A,t}, \Pi_{B,t}\}_{t\geq 0}$ achievable in a subgame perfect equilibrium in pure strategies. This step simplifies the construction of the Pareto frontier since it reduces the search space from the space of strategies to the space of payoffs. Following Abreu, Pearce and Stacchetti (1990), equilibrium payoffs are characterized by promise keeping, sustainability and transversality conditions. Promise keeping and transversality conditions ensure that the sequence of continuation profits can be constructed recursively from stage game profits. Sustainability conditions ensure that one-shot deviations are not profitable.

Lemma 3.1. *A sequence of continuation profits $\{\Pi_{A,t}, \Pi_{B,t}\}_{t\geq 0}$ is achievable in a subgame perfect equilibrium if and only if, for all $t \geq 0$, there are $\pi_{A,t}, \pi_{B,t} \geq 0$ with $\pi_{A,t} + \pi_{B,t} \leq v$ such that:*

$$\Pi_{A,t} = (1 - \delta_A)\pi_{A,t} + \delta_A \Pi_{A,t+1} \qquad PK(A, t)$$

$$\Pi_{B,t} = (1 - \delta_B)\pi_{B,t} + \delta_B \Pi_{B,t+1} \qquad PK(B, t)$$

$$\Pi_{A,t} \geq (1 - \delta_A)(\pi_{A,t} + \pi_{B,t}) \qquad S(A, t)$$

$$\Pi_{B,t} \geq (1 - \delta_B)(\pi_{A,t} + \pi_{B,t}) \qquad S(B, t)$$

$$\{\Pi_{A,t}, \Pi_{B,t}\}_{t\geq 0} \textit{ is bounded.} \qquad (Transversality)$$

Proof. See appendices. □

The transversality condition excludes exponentially growing sequences of continuation profits.[15] The best one-shot deviation for a firm is to undercut slightly its rival to attract all consumers and capture the entire industry profits. The best way to disincentivize a firm from undercutting its rival is to punish it with the play of the worst SPE (in which firms play the NE in all histories), yielding a continuation profit of 0. In other words, collusion is sustainable if, during any period and for all firms, the collusive continuation profit is larger than the current industry profits. Rewriting sustainability conditions in terms of continuation profits yields:

$$\begin{aligned} \Pi_{B,t} &\leq \delta_B \Pi_{B,t+1} + \frac{1-\delta_A}{1-\delta_B}\delta_A \Pi_{A,t+1} && S(A,t) \\ \Pi_{A,t} &\leq \delta_A \Pi_{A,t+1} + \frac{1-\delta_B}{1-\delta_A}\delta_B \Pi_{B,t+1} && S(B,t). \end{aligned}$$

For A's sustainability condition to be satisfied, B's continuation profit must not be too high compared to a weighted sum of both firms' future continuation profits. Equivalently, in matrix form:[16]

$$\begin{pmatrix} \Pi_{A,t} \\ \Pi_{B,t} \end{pmatrix} \leq \underbrace{\begin{pmatrix} \delta_A & \frac{1-\delta_B}{1-\delta_A}\delta_B \\ \frac{1-\delta_A}{1-\delta_B}\delta_A & \delta_B \end{pmatrix}}_{S} \begin{pmatrix} \Pi_{A,t+1} \\ \Pi_{B,t+1} \end{pmatrix}$$

Sustainability conditions hold for all periods so the period 0 vector of equilibrium profit is necessarily less than S^t multiplied by the period t vector of equilibrium profit. The matrix S has two eigenvalues of 0 and $\delta_A + \delta_B$. If $\delta_A + \delta_B < 1$ only a continuation payoff of 0 can be sustained in equilibrium.

Stated differently, a sequence of continuation profits $\{\Pi_{A,t}, \Pi_{B,t}\}_{t \geq 0}$ is sustainable only if it grows at a geometric rate of at least $\frac{1}{\delta_L + \delta_H}$. Thus, if $\delta_L + \delta_H < 1$ only exponentially

[15] In this game, boundedness is equivalent to the more standard transversality condition $\lim_{t\to\infty} \delta_K^t \Pi_{K,t} = 0$ for $K \in \{A, B\}$.

[16] Where $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ iff $x_1 \leq y_1$ and $x_2 \leq y_2$.

increasing sequences are sustainable. However, these sequences are unfeasible since demand is bounded by v.

Proposition 3.2. *If $\delta_A + \delta_B < 1$, positive profits cannot be sustained in equilibrium. The set of equilibrium payoffs is $\{(0,0)\}$.*

Proof. See Harrington (1989) or Obara and Zincenko (2017) (theorem 3.1) for proofs of this result. Appendix A.2 presents an alternative proof. □

Positive payoffs cannot be sustained in equilibrium if the mean of both firms' discount factors $\frac{\delta_A+\delta_B}{2}$ is less than $\frac{1}{2}$, echoing the standard condition under homogeneous discounting of $\delta < \frac{1}{2}$. In the remainder of this paper, I assume that firm A's discount factor is δ_H and that $\delta_H > \frac{1}{2}$, whereas if firm B is of type δ_L, I assume that $\delta_L + \delta_H < \frac{1}{2}$ so that collusion is not sustainable.

The next subsection characterizes the Pareto frontier of equilibrium profits for δ_H firms.

Pareto Frontier

To construct the Pareto Frontier, I optimize the weighted sum of A's and B's profits achievable in equilibrium with parameter $\gamma \in (0,1)$. Using lemma 3.1, this problem is:

$$
\begin{aligned}
\mathcal{P}_1(\gamma) \quad : \quad & \max_{\{\Pi_{A,t},\Pi_{B,t}\}_{t\geq 0}} \gamma\Pi_{A,0} + (1-\gamma)\Pi_{B,0} \\
& \text{subject to: } \Pi_{A,t} \geq \delta_H \Pi_{A,t+1} \quad \Pi_{B,t} \geq \delta_H \Pi_{B,t+1} && C_1(t), \quad C_2(t) \\
& \Pi_{A,t} + \Pi_{B,t} \leq (1-\delta_H)v + \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] && C_3(t) \\
& \Pi_{B,t} \leq \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] && S(A,t) \\
& \Pi_{A,t} \leq \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] && S(B,t)
\end{aligned}
$$

The Pareto frontier F_1 is the set of solutions to $\mathcal{P}_1(\gamma)$ for all $\gamma \in (0,1)$. These solutions are the Pareto optimal collusive profits achievable in a repeated Bertrand duopoly game

$\delta_H \geq \frac{1}{2}$. In this game, it is well known that firms can sustain the monopoly price v on the collusive path. This subsection presents an alternative method for solving this problem that extends to the analysis of an incomplete information game. It reestablishes that, to achieve the greatest collusive profits, firms set the monopoly price v every period, and it presents the divisions of the resulting industry profits that are part of the Pareto frontier. These results and those in appendix $A.2$ are also used in section 4.

To construct the Pareto frontier, I first discuss the prices that are sustainable on the collusive path for firms with discount factor $\delta_H \geq \frac{1}{2}$. The market price that correspond to $\Pi_{A,t} + \Pi_{B,t} - \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}]$ is either restricted by sustainability conditions, i.e., the current price cannot be too high compared to future collusive profits, or by the demand, i.e., when more expensive than v, no consumers buy the good. Consider a sequence of continuation industry profits $\{\Pi_t\}_{t\geq 0}$. Industry profits are limited either by sustainability concerns or v:

$$\begin{aligned} \Pi_t &\leq 2\delta_H \Pi_{t+1} && S(A,t) + S(B,t) \\ \Pi_t - \delta_H \Pi_{t+1} &\leq (1-\delta_H)v && C_3(t) \end{aligned}$$

Since $\delta_H \geq \frac{1}{2}$, all nondecreasing sequences of continuation industry profits $\{\Pi_t\}_{t\geq 0}$ satisfy sustainability conditions. What restricts industry profits is the consumers' willingness to pay v captured by $C_3(t)$. Therefore, a candidate solution is to choose a strategy profile that pays the monopoly price v in all periods on the collusive path and so yields a constant industry profit of v during each period. The industry profit must be shared fairly to be sustainable, that is, with $\alpha v \leq \delta_H v$ (from $S(A,t)$) and with $(1-\alpha)v \leq \delta_H v$ (from $S(B,t)$). This outcome constructs a segment on the plane of firms profits: $\{(v - \Pi_{B,0}, \Pi_{B,0}) \ : \ (1-\delta_H)v \leq \Pi_{B,0} \leq \delta_H v\}$, which is part of the Pareto frontier and maximizes the industry profit.

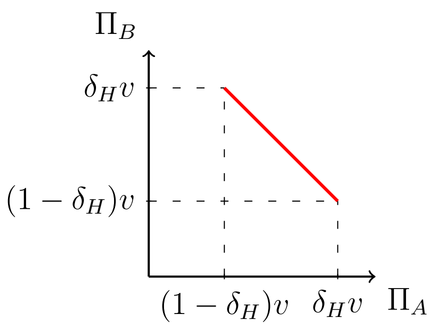
In fact, the Pareto frontier coincides with this segment. Indeed, firms' current profits are constrained by the future industry profits. Therefore, to maximize one firm's profit, the industry profits must also be maximized.

Proposition 3.3. *The Pareto frontier of equilibrium profits is:*

$$F_1 = \{(v - \Pi_{B,0}, \Pi_{B,0}) \ : \ (1-\delta_H)v \leq \Pi_{B,0} \leq \delta_H v\}$$

To achieve a Pareto optimal payoff, firms must set the monopoly price v in all periods.

Proof. A general version of this Pareto frontier (more players and discount factors) is characterized in Obara and Zinceko (2017). Appendix A.2 characterizes the frontier for this case. □



If firms are more patient, then less equal divisions of the industry profits are sustainable. The appendix presents various ways of dynamically sharing the market that construct a single point of the Pareto frontier. However, all profits on the Pareto frontier are achieved by setting the monopoly price v in each period. The next section shows that, under incomplete information about B's discount factor, Pareto optimal collusive paths start with a gradual increase in prices.

4 Collusion under incomplete information

For this section, $\Pi_{L,t}$ denotes B_L's continuation profit, and $\Pi_{H,t}$ denotes B_H's continuation profit. I assume that $\delta_L < \frac{1}{2}$ and that $\delta_H \geq \frac{1}{2}$. That is, if B is of type δ_L, and firm A knows B's type, then no price greater than 0 can be sustained in equilibrium. In contrast, if B is of type δ_H, and A knows B's type, then all prices in $[0, v]$ can be sustained in equilibrium.

This section shows that, in a pure strategy separating BPE, profits on the Pareto frontier from the separation period feature a gradual increase in prices on the patient firms' path. It proceeds in four steps. First, I show that pooling equilibria, in which B_L and B_H set the same prices on the path, are suboptimal and yield 0 profits for all firms. Second, I study an example of a separating equilibrium that yields positive profits to all firms and types. This example shows that, to separate B_L from B_H in an incentive compatible manner, patient firms must delay the time at which they pay the highest sustainable price v . Third, this section characterizes the payoffs achievable in a separating equilibrium using promise keeping, sustainability conditions and incentive compatibility conditions. Incentive compatibility conditions ensure that types are in best response by separating and not mimicking each other. Finally, this section characterizes the Pareto frontier from the separation period. I do not address the question of the optimal separation period; rather, I show that, from the separation period, any Pareto optimal profits are constructed using a transition phase during which prices increase gradually.

Pooling Equilibria

This subsection shows that pooling equilibria yield 0 profits for all firms and types. In a pooling equilibrium, B_L and B_H pay the same price on path, and so A's belief remains at ρ_0 on the path. A pooling equilibrium path must be sustainable for all firms and types; that is, the highest sustainable price in a pooling equilibrium is pinned down by the least patient firms. In this case, A and B_L cannot sustain supra-competitive prices. Therefore, a pooling equilibrium yields 0 profits for all firms.

Proposition 4.1. *Any pooling BPE yields* 0 *profits for all firms and types.*

Proof. See appendices. □

This result extends to mixed strategies. For profiles of mixed strategies in which, during every period, there is a history reached with positive probability (even vanishingly small) during which B_L and B_H have set the same prices for all periods. For this particular path, the proposition applies, and so it must yield 0 profits to all firms and types. Consequently, since firms are in best response employing mixed strategies, any other path yields 0 profits.

To reach higher prices, patient firms must screen out impatient firms from collusive schemes. In fact, as the next subsection shows, all types of firms can be made better by separating.

Separating Equilibria: An Example

This subsection presents a separating equilibrium that yields positive payoffs for all firms and types and so dominates any pooling equilibria. Importantly, it shows that, to separate impatient firms from patient firms, A and B_H must delay the period during which they set the highest prices. This delay is necessary to prevent B_L from mimicking B_H and undercutting at a high collusive price. Introducing this delay costs more to a mimicking B_L than to a patient firm since it discounts profits at rate δ_L rather than δ_H. That is, it increases the difference in profits between types and enables separation.

Consider the following strategy profile with separation in period 0. During period 0, firms A and B_H set the same positive price $p_0 > 0$, while firm B_L undercuts and obtains a profit of p_0. At the end of period 0, firm A knows B's type. If B is of type δ_L, then it plays the NE for all future histories, which is the only equilibrium continuation. If B is of type δ_H, then after a delay of τ periods, firms A and B_H set v for all future periods. During this delay, they set a price of 0. Any detected deviations are punished by playing the NE forever.[17] Firms A and B_H split the market according to a fixed $\alpha \in [1-\delta_H, \delta_H]$. This strategy profile yields positive profits for all firms; moreover, there is a $p_0 > 0$ and a $\tau \in \mathbb{N}^*$ for which this strategy profile constitutes a BPE.

The strategy profile delays the best collusive price by τ to prevent B_L from mimicking firm B_H in period 0 and undercutting at price v instead of p_0. This first equilibrium condition

[17] An undetected deviation is, for instance, if B_L mimics B_H at $t = 0$.

is:

$$p_0 \geq (1-\alpha)p_0 + \delta_L^\tau v. \tag{1.1}$$

B_L's undercutting profit p_0 must be greater than the deviation profit from mimicking B_H in 0 to undercut after a delay of τ periods at the monopoly price v. In this deviation, B_L claims B_H's share of the industry profit $(1-\alpha)p_0$ in period 0 and obtains the entire industry profit at the monopoly price in period τ. A higher p_0 relaxes this condition, particularly if $p_0 = v$ then (1) holds. However, since in period 0, firm A expects collusion only with probability ρ_0, then p_0 cannot be too high; otherwise, firm A might find it profitable to undercut B_L in period 0. This second equilibrium condition is:

$$\rho_0\left(\alpha p_0 + \frac{\delta_H^\tau}{1-\delta_H}\alpha v\right) \geq p_0. \tag{1.2}$$

The expected continuation collusive profit for A in period 0 must be greater than the undercutting profit of p_0.[18] Since collusion is uncertain for A, the price before separation cannot be too high. For instance, if $\alpha = \delta_H$ and $\rho_0 < \frac{1-\delta_H}{\delta_H}$, then $p_0 = v$ violates (2). In that case, p_0 is limited by $A's$ sustainability condition (2). In turn, the price in period 1 on the collusive path for patient firms cannot be too high compared to p_0; otherwise, B_L might mimic B_H in period 0 to undercut at a higher price in period 1. Consequently, if $\rho_0 < \frac{\delta_L}{\delta_H^2+\delta_L}$,[19] a delay of at least two periods is necessary to separate B_L from B_H. Conditions (1) and (2) imply that:

$$\delta_L^\tau \frac{v}{\alpha} \leq p_0 \leq \delta_H^\tau \frac{\alpha v \rho_0}{(1-\delta_H)(1-\alpha\rho_0)}$$

For any parameter values, since $\lim_{\tau\to\infty}\left(\frac{\delta_L}{\delta_H}\right)^\tau = 0$, there are sufficiently large values of τ and a p_0 such that the strategy profile satisfies (1) and (2) and therefore constitutes a

[18] As mentioned in section 2, I assume that, for all prices p, there exists an undercutting price p^u that captures the entire market and yields a profit of p.

[19] This condition is obtained in appendix C and is discussed later in this section.

BPE of the dynamic game. Using the mechanism design terminology, firms' profits have the "strict increasing difference" property in the delay's length τ. Increasing τ is more costly to B_L, which discounts future undercutting profits at rate δ_L, than to a firm of type δ_H, which enables separation.

Although delay is necessary for separation, setting prices of 0 during the delay is suboptimal. The next subsections show that the Pareto optimal way to implement this delay is with a gradually increasing price path.

Separating Equilibrium Profits

This subsection characterizes all profits achievable in a pure strategy separating BPE. This step simplifies the construction of the Pareto frontier since it reduces the search space from the space of strategies to the space of continuation profits. A separating BPE is a strategy profile such that the strategies of B_L and B_H differ on at least one information set reached with positive probability. In pure strategies, separation occurs in only one period: on the first information set, reached on a path, for which σ_L and σ_H differ. Let T denote the separation period, in which A learns B's type. Specifically, before period T, firm A believes B is B_H with probability ρ_0. Between T and $T+1$, A's belief moves from ρ_0 to 0 (if B is of type δ_L) or to 1 (if B is of type δ_H) and remains at these levels afterward. Consequently, A's consistent beliefs are entirely pinned down by the separation period T and matter for the following analysis only in computing A's expected continuation profits. In the remainder of this section, I discuss the conditions for which continuation profits are achievable in a BPE. In contrast to the complete information case, sustainability and promise keeping conditions are necessary but not sufficient; therefore, I introduce incentive compatibility conditions.

Consider a strategy profile that separates B_H from B_L in period $T \in \mathbb{N}$. After separation, there are two equilibrium paths: one designed for patient firms A and B_H; and the other designed for A and B_L. In A and B_H's path, sustainability and promise keeping conditions characterize the sequences of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}\}_{t>T}$ achievable in equilibrium, as in lemma 3.1. In A and B_L's path, since $\delta_L + \delta_H < 1$, the only continuation equilibrium is the repetition of the NE for all future histories. A and B_L's equilibrium profits on that path are 0. As a result, B_L must be in static best response in period T, which must undercut A's price. During the pooling period, sustainability and promise keeping conditions characterize the sequence of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}, \Pi_{L,t}\}_{0 \leq t < T}$ achievable in equilibrium, as in lemma 3.1.

The patient firms' sequence of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}\}_{t \geq 0}$ corresponds to A's and B_H's profits during the pooling period and on the collusive path after separation, whereas the sequence $\{\Pi_{L,t}\}_{0 \leq t \leq T}$ corresponds to B_L's continuation profit during the pooling period and at the separation periods (B_L's profit must be 0 after separation). Using these notations, promise keeping and sustainability conditions outside the separation period entail that, for all of period $t \neq T$, there are nonnegative current profits $\pi_{A,t}, \pi_{B,t}$ with $\pi_{A,t} + \pi_{B,t} \leq v$ such that:

$$
\begin{aligned}
&\Pi_{A,t} = (1-\delta_H)\pi_{A,t} + \delta_H \Pi_{A,t+1} && PK(A,t) \\
&\Pi_{H,t} = (1-\delta_H)\pi_{B,t} + \delta_H \Pi_{H,t+1} && PK(H,t) \\
&\Pi_{L,t} = (1-\delta_L)\pi_{B,t} + \delta_L \Pi_{L,t+1} && PK(L,t) \quad \text{for } t < T \\
&\Pi_{A,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) && S(A,t) \\
&\Pi_{H,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) && S(H,t) \\
&\Pi_{L,t} \geq (1-\delta_L)(\pi_{A,t} + \pi_{B,t}) && S(L,t) \quad \text{for } t < T.
\end{aligned}
$$

During the separation period, B_H's promise keeping condition is unchanged, whereas A's promise keeping condition considers that future collusive profits occur with probability ρ_0. In the separation period, B_L's only action supported in equilibrium is to undercut the market price. Therefore, in this period, A and B_L's promise keeping conditions are:

$$\Pi_{A,T} = \rho_0 \left[(1-\delta_H)\pi_{A,T} + \delta_H \Pi_{A,T+1}\right] \qquad PK(A,T)$$

$$\Pi_{L,T} = (1-\delta_L)(\pi_{A,T} + \pi_{B,T}) \qquad PK(L,T).$$

Promise keeping conditions ensure that continuation profits can be constructed from stage game profits. Sustainability conditions ensure that *detectable* one shot deviations are not profitable. Any *detected* deviations are punished by playing the NE repeatedly, which yields 0 payoffs.[20] Undercutting the current market price is the best one shot deviation. Therefore, sustainability conditions hold if, for all firms and types and during each period, the expected continuation profit is weakly larger than the current industry profit. At T, there are no sustainability conditions for B_L since its equilibrium strategy is to undercut the current price.

These conditions are, however, insufficient to characterize equilibrium profits. In the incomplete information game, not all deviations are *observable*. Indeed, when B_H mimics B_L or vice-versa, A does not detect the deviation and believes the play continues on the equilibrium path, yet A is wrong about B's true type. To capture these deviations into equilibrium conditions, I introduce incentive compatibility conditions (IC) and a law of motion (LM) that ensure B is in best reply separating in period T. The incentive compatibility conditions ensure that B_L's deviation profit when mimicking B_H is not greater than its equilibrium profit from separating at T. The law of motion allows for a recursive formulation of the incentive compatibility conditions that will simplify the optimization process and the construction of the Pareto frontier.

First, consider B_H's undetected deviation to mimic B_L at T. That means B_H undercut A's price at T. In this case, A believes that B is of type δ_L and plays according to the equilibrium strategy: play NE for all future periods. Therefore, if B_H mimics B_L, it captures

[20]This continuation is a BPE regardless of A's belief or B's type and yields the minmax payoff to all firms and types.

industry profit at T but has no profits in the future. This incentive compatibility condition is:

$$\Pi_{H,T} \geq (1-\delta_H)(\pi_{A,T}+\pi_{B,T}).$$

By abuse of notation and because it is functionally equivalent to a sustainability condition, I label this condition $S(H,T)$ in the analysis.

Second, consider B_L's undetected deviation to mimic B_H from T to, say, $t-1$ and undercut the market price at t. In this case, B_L plays according to B_H's strategy; the play continues on the patient firms' path and B_L claims B_H's market shares from T to $t-1$. Formally, mimicking B_H from T to $t-1$ and undercutting at t do not result in a profitable deviation for B_L if:

$$\pi_{A,T}+\pi_{B,T} \geq \sum_{s=T}^{t} \delta_L^{s-T}\pi_{B,s} + \delta_L^{t-T}(\pi_{A,t}+\pi_{B,t}) \qquad IC(t).$$

The separating profit for B_L must be greater than the profit generated by B_H's market shares every for period plus the undercutting profit at t. B_L's best response is separating at T if and only if, for all $t>T$, the incentive compatibility conditions $IC(t)$, hold. Proposition 4.2 shows that sustainability and promise keeping conditions, together with the incentive compatibility conditions, are necessary and sufficient for sequences of continuation profits to be achievable in a BPE.

Before stating the result of proposition 4.2, I rewrite the incentive compatibility constraint in a recursive form, which simplifies the construction of the Pareto frontier. I consider the sequence of state variable $\{\Pi_{L,t}\}_{t>T}$, where $\Pi_{L,t}$ is the B_L's promised profit on the patient collusive path at t, rendering B_L indifferent between separating at T or mimicking B_H in periods $T, T+1, .., t-1$.

Definition 1. *Consider sequences of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}\}_{t\geq 0}$ and $\{\Pi_{L,t}\}_{0\leq t\leq T}$. The sequence of state variables $\{\Pi_{L,t}\}_{t\geq T+1}$ is the sequence of B_L's promised continuation* *profits on the patient firms path, making B_L's indifferent between separating at T or mimicking B_H. Formally, for $t \geq T+1$:*

$$\underbrace{(1-\delta_L)\sum_{s=T}^{t-1}\delta_L^{s-T}\pi_{B,s}}_{\text{mimicking profit until } t-1} + \delta_L^t \Pi_{L,t} = \underbrace{(1-\delta_L)(\pi_{A,T}+\pi_{B,T})}_{\text{separating profit at } T}$$

The right hand side of the equation captures B_L's separation profit at T, B_L undercuts at T and claims the entire industry profit. The left hand side represents B_L's mimicking profit in which it claims B_H's market shares from T to $t-1$ plus the state variable $\Pi_{L,t}$ that ensures both sides are equal. To be sure $\Pi_{L,t}$ does not correspond to a continuation profit that is obtainable in the game. Rather, the state variable is the upper bound on deviation profits above which B_L is not in best response separation at T. For instance, consider a deviation in which B_L mimics B_H from T to $t-1$ and undercut the market price at t. This deviation is not strictly profitable if $IC(t)$ holds. Using the defining equation of $\Pi_{L,t}$, $IC(t)$ becomes:

$$\Pi_{L,t} \geq (1-\delta_L)(\pi_{A,t}+\pi_{B,t}) \qquad IC(t).$$

In other words, if the promised profit at t, which renders B_L indifferent from mimicking B_H until $t-1$ and separating at T, is larger than the undercutting profit in period t, then mimicking B_H until $t-1$ and undercutting at t is not a profitable deviation for B_L.

In addition, the sequence of state variables $\{\Pi_{L,t}\}_{t>T}$ can be characterized recursively. From the defining equation of $\Pi_{L,t}$ and $\Pi_{L,t+1}$ one has:

$$(1-\delta_L)\sum_{s=T}^{t-1}\delta_L^{s-T}\pi_{B,s} + \delta_L^t \Pi_{L,t} = (1-\delta_L)\sum_{s=T}^{t}\delta_L^{s-T}\pi_{B,s} + \delta_L^{t+1}\Pi_{L,t+1}$$

$$\iff$$

$$\Pi_{L,t+1} = \frac{\Pi_{L,t}-(1-\delta_L)\pi_{B,t}}{\delta_L} \qquad LM(t).$$

If $\Pi_{L,t}$ is what must be promised for B_L to be indifferent from mimicking B_H up to t or separating, then what must be promised in $t+1$ is less by $(1-\delta_L)\pi_{B,t}$ (the benefit of mimicking in t) or is greater by a factor of $\frac{1}{\delta_L}$ to compensate discounting. Remark that, in period T:

$$\Pi_{L,T+1} = \frac{(1-\delta_L)\pi_{A,t}}{\delta_L} \overset{PK}{=} \frac{\Pi_{L,T} - (1-\delta_L)\pi_{B,T}}{\delta_L} \qquad LM(T)$$

The law of motion $LM(t)$ constructs $\{\Pi_{L,t}\}_{t \geq T+1}$ inductively from the starting value $\Pi_{L,T} = (1-\delta_L)(\pi_{A,T} + \pi_{B,T})$ defined by B_L's separation profit.

All in all, (IC) and (LM) ensure that undetected deviations, in which B_L mimics B_H, are not profitable. The next proposition shows that adding these two conditions to the standard sustainability and promise keeping conditions characterizes all of the continuation profits achievable in equilibrium.

Proposition 4.2. *Sequences of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}\}_{t\geq 0}$,$\{\Pi_{L,t}\}_{0 \leq t \leq T}$ and the sequence of state variables $\{\Pi_{L,t}\}_{t \geq T+1}$ are achievable in a BPE with separation at $T \geq 0$ if and only if, for all $t \geq 0$, there exists $\pi_{A,t}, \pi_{B,t} \geq 0$ with $\pi_{A,t} + \pi_{B,t} \leq v$ such that:*

$$\Pi_{H,t} = (1-\delta_H)\pi_{B,t} + \delta_H \Pi_{H,t+1} \qquad PK(H,t)$$

$$\Pi_{A,t} = (1-\delta_H)\pi_{A,t} + \delta_H \Pi_{A,t+1} \qquad PK(A,t) \text{ for } t \neq T$$

$$\Pi_{A,T} = \rho_0 \left[(1-\delta_H)\pi_{A,T} + \delta_H \Pi_{A,T+1}\right] \qquad PK(A,T)$$

$$\Pi_{L,t} = (1-\delta_L)\pi_{B,t} + \delta_L \Pi_{L,t+1} \qquad PK(L,t) \ / \ LM(t)$$

$$\Pi_{L,T} = (1-\delta_L)(\pi_{A,T} + \pi_{B,T}) \qquad PK(L,T)$$

$$\Pi_{H,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) \qquad S(H,t)$$

$$\Pi_{A,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) \qquad S(A,t)$$

$$\Pi_{L,t} \geq (1-\delta_L)(\pi_{A,t} + \pi_{B,t}) \qquad S(L,t) \ / \ IC(t) \text{ for } t \neq T.$$

$$\{\Pi_{A,t}, \Pi_{H,t}\}_{t \geq 0} \text{ is bounded.} \qquad (Transversality)$$

Proof. See appendices. □

There are, after all, few differences between conditions that characterize continuation profits achievable in a pooling equilibrium and conditions that characterize continuation profits achievable in a separating equilibrium. For B_H, nothing changes; for A the promise keeping condition at T changes since A expects collusion with probability ρ_0. For B_L, in all periods, $\{\Pi_{L,t}\}_{t\geq T+1}$ satisfies conditions functionally equivalent to sustainability and promise keeping conditions. However, the interpretation of these conditions is different. From proposition 3.2, there are no bounded $\{\Pi_{L,t}\}_{t\geq T+1}$ that can satisfy these conditions. That is, $\{\Pi_{L,t}\}_{t\geq T+1}$ is typically unbounded, which is not an issue since $\{\Pi_{L,t}\}_{t\geq T+1}$ value is anchored in period T: $\Pi_{L,T} = (1-\delta_L)(\pi_{A,T} + \pi_{B,T})$, and after T, this value does not capture profits achievable in the game. Consequently, after a finite number of periods, the value of the state variable increases to greater than the monopoly price level. Since $\Pi_{L,t}$ restricts the current price (from (IC)), then subsequently, stage patient firms can collude at the monopoly price level, whereas before that stage, prices are restricted to less than v, and these periods correspond to the transition phase.

The next subsection computes and reports the main property of the Pareto frontier on the patient firms collusive path (after T) for a given starting state variable $\Pi_{L,T}$. It characterizes the speed of the price increase during the transition phase.

Pareto Frontier After Separation

This subsection characterizes the Pareto frontier of profits achievable in a separating BPE after a given separation period T for a given starting state variable. To simplify the notations, I assume the separation period is 0 so that I optimize the weighted sum of firm A and B_H's profits from period 1 onward with a parameter $\gamma \in (0,1)$.[21] Using proposition 4.2 this

[21] This assumption is inconsequential since profits from t onward are independent of past profits.

problem is:

$$
\begin{aligned}
&\mathcal{P}_2(\gamma, \Pi_{L,1}): \\
&\max_{\{\Pi_{A,t},\Pi_{H,t}\}_{t\geq 1},\{\Pi_{L,t}\}_{t\geq 1}} \gamma\Pi_{A,1} + (1-\gamma)\Pi_{H,1} \\
&\text{subject to: } \Pi_{A,t} \geq \delta_H \Pi_{A,t+1} \quad \Pi_{H,t} \geq \delta_H \Pi_{H,t+1} && C_1(t), \quad C_2(t) \\
&\Pi_{A,t} + \Pi_{H,t} \leq (1-\delta_H)v + \delta_H[\Pi_{H,t+1} + \Pi_{A,t+1}] && C_3(t) \\
&\Pi_{H,t} \leq \delta_H[\Pi_{H,t+1} + \Pi_{A,t+1}] && S(A,t) \\
&\Pi_{A,t} \leq \delta_H[\Pi_{H,t+1} + \Pi_{A,t+1}] && S(B,t) \\
&\delta_L \Pi_{L,t+1} = \Pi_{L,t} - \frac{1-\delta_L}{1-\delta_H}[\Pi_{H,t} - \delta_H \Pi_{H,t+1}] && LM(t) \\
&\Pi_{L,t} \geq \frac{1-\delta_L}{1-\delta_H}[\Pi_{A,t} + \Pi_{H,t} - \delta_H(\Pi_{A,t+1} + \Pi_{H,t+1})] && IC(t)
\end{aligned}
$$

Solutions to $\mathcal{P}_2(\gamma, \Pi_{L,1})$ maximize firm A and B_H's equilibrium profits after separation while restricting B_L's mimicking profit to $\Pi_{L,1}$. This problem is similar to the complete information problem with the additional conditions of $LM(t)$ and $IC(t)$. As in section 3, I neglect the transversality condition and remove ex post unbounded solutions to the problem. To solve $\mathcal{P}_2$, I proceed in two steps. First, I show that solutions to $\mathcal{P}_2$ must, for feasibility reasons, feature a transition phase during which, for a finite number of periods, prices are less than the monopoly level. Second, I show that, to obtain Pareto optimal payoffs, prices gradually increase during the transition phase.

Definition 2. *The transition phase of a solution to $\mathcal{P}_2$ corresponds to the periods in which the prices played on the patient firm collusive path are less than the monopoly price v. Let τ denote the length of the transition phase.*

In each period, the market price is restricted by sustainability issues ($S(A)$ and $S(B)$), by the demand (C_3) or by incentive compatibility concerns (IC). In section 3, I showed that C_3 is more stringent than sustainability issues for non decreasing paths of continuation payoffs. Whether IC is more stringent than C_3 depends on the value of the state variable

$\Pi_{L,t}$:

$$\underbrace{\Pi_{A,t} + \Pi_{H,t} - \delta_H(\Pi_{A,t+1} + \Pi_{H,t+1})}_{\text{period } t \text{ market price}} \leq \begin{cases} (1-\delta_H)v & C_3(t) \\ \frac{1-\delta_L}{1-\delta_H}\Pi_{L,t} & IC(t) \end{cases}$$

If $\Pi_{L,t} > (1-\delta_L)v$, then the demand constrains the market price and not incentive compatibility. Indeed, in the ideal scenario, B_L mimics B_H to undercut at the monopoly price if this ideal undercutting profit is less than the promised profit that renders B_L indifferent between separating or mimicking B_H; then, incentive compatibility is no longer an issue. In this case, $\mathcal{P}_2$ can be shown to be equivalent to the complete information problem $\mathcal{P}_1$. In contrast, if $\Pi_{L,t} < (1-\delta_L)v$, the price is restricted by the incentive compatibility condition. If patient firms set a price higher than $\frac{1-\delta_L}{1-\delta_H}\Pi_{L,t}$, then B_L has a strictly profitable deviation to mimic B_H and undercut at t when this price is played. For these periods, the price played by patient firms must be less than the monopoly price.

From proposition 4.3, no bounded sequences $\{\Pi_{L,t}\}_{t>0}$ satisfy $LM(t)$ and $IC(t)$.[22] Therefore, $\{\Pi_{L,t}\}_{t>0}$ must be unbounded and so $(1-\delta_L)v$ after a finite number of periods. In other words, after a finite number of periods, patient firms set the monopoly price v in all periods. The transition phase, during which the price is restricted to less than the monopoly price, lasts a finite number of periods. Next, I discuss why during the transition phase the price gradually increases.

The law of motion $LM(t)$ captures how B_H's market shares determines the evolution of the state variable $\Pi_{L,t}$ over time. Reducing B_H's current market share makes it less profitable for B_L to mimic B_H, increasing the future state variable by a factor of $\frac{1}{\delta_L}$. In turn, the market price in the next period can be increased by a factor of $\frac{1}{\delta_L}$. Since patient firms' profits are discounted at a higher rate δ_H, the industry profits in the next period increase by a factor of $\frac{\delta_H}{\delta_L} > 1$. Therefore, reducing B_H's market shares to 0 during the transition

[22] The law of motion and incentive compatibility conditions are functionally equivalent to the sustainability and promise keeping conditions.

phase increases the industry profits and causes the state variable, and therefore the market price, to grow at a rate of $\frac{1}{\delta_L}$ during every period. However, this trade-off does not always benefit B_H. B_H bears the entire costs of reducing B_L's mimicking profit but obtains only a fraction of the gains since industry profits must be shared among patient firms to ensure sustainability. For solutions to $\mathcal{P}_2$ that assign a high weight to B_H's profit, B_H has positive market shares during the transition phase. As a result, prices grow at a slower rate of $\frac{1}{\delta_L+\delta_H}$. The next proposition summarizes these results:

Proposition 4.3. *If $0 < \Pi_{L,1} < (1-\delta_L)v$, then all solutions to $\mathcal{P}_2$ feature a transition phase of finite length $\tau \geq 1$ such that:*

1. *During the transition phase, for $t \in \{1, ..., \tau\}$, the market price is restricted to less than the monopoly price and equals $\frac{\Pi_{L,t}}{1-\delta_L} < v$;*

2. *During the transition phase, the market price strictly increases at a geometric rate between $\frac{1}{\delta_L+\delta_H}$ and $\frac{1}{\delta_L}$. For all $t \in \{1, ..., \tau-1\}$:*

$$\frac{1}{\delta_L + \delta_H} \leq \frac{\Pi_{L,t+1}}{\Pi_{L,t}} \leq \frac{1}{\delta_L}$$

   ; and

3. *After the transition phase, from period $\tau + 1$ onward, patient firms set the monopoly price in all periods.*

Proof. See the appendix. □

If the starting value of the state variable is 0, then B_L's profit from mimicking B_H cannot be higher than 0; therefore, the only feasible price is 0. In other words, B_L must obtain a positive profit during the separation period, a rent, to separate types and enable patient firms to play the highest sustainable prices. If the starting value of the state variable is greater than $(1-\delta_L)v$, then incentive compatibility is not an issue from the first period onward, and patient firms play the monopoly price in all periods. Next, I characterize Pareto optimal payoffs after the separation period for a given starting state variable value $\Pi_{L,1}$.

Pareto Frontier. A transition phase yielding Pareto optimal payoff obeys the following trade-off. Reducing current B_H's profit accelerates the price increase (with the highest geometric speed of $\frac{1}{\delta_L}$) and increases all future prices during the transition phase. However, B_H bears entirely the cost of accelerating the price increase but captures only a share of the gains due to sustainability issues. This trade-off is always profitable for firm A and profitable for firm B_H only if the remaining length of the transition phase is sufficiently long, that is, only if the acceleration of the price increase affects sufficiently many prices. As a result, a transition phase that yields Pareto optimal profits takes different forms: (i) if the weight on A's profit is large, then the price increase at the highest rate of $\frac{1}{\delta_L}$ and B_H's market shares are 0 during the transition phase; and (ii) if the weight on B_H's is large, however, then a transition phase of length τ has two parts. During the first part of length τ_f, in which the remaining length of the transition phase is sufficiently great, the price increases at the highest rate of $\frac{1}{\delta_L}$. During the second part of length $\tau_s = \tau - \tau_f$, in which there are a few periods of the transition phase remaining, B_H has positive market shares and the price increase at the smallest rate of $\frac{1}{\delta_L+\delta_H}$.

The parameter γ determines the length τ_s of the "slow" part of the transition phase. The higher that $1-\gamma$ is, the more weight that is assigned on B_H's profit in the objective, and the longer that the slow part of the transition phase is during which B_H has positive market shares. The value of the starting state variable $\Pi_{L,1}$ determines the total length of the transition phase. The smaller that $\Pi_{L,1} is$, that is, the smaller that the upper bound is on B_L's mimicking profit on the patient firm collusive path, the longer that the transition phase is; i.e., patient firms further delay playing the monopoly price.

Proposition 4.4. *All solutions to $\mathcal{P}_2$ feature a transition phase that starts with a fast part of length τ_f and ends with a slow part of length τ_s, with $\tau_f + \tau_s = \tau$.*

1. *During the fast part, B_H's market shares are 0, and prices increase at a rate of $\frac{1}{\delta_L}$.*

2. *During the slow part, B_H's market shares equal $\frac{\delta_H}{\delta_L+\delta_H}$, and prices increase at a rate of $\frac{1}{\delta_H+\delta_L}$.*

The length of the fast part τ_f and of the slow part τ_s are characterized by[23]*:*

$$(\delta_L+\delta_H)^{\tau_s-1}\delta_L^{\tau_f} > \frac{\Pi_{L,0}}{(1-\delta_L)v} \geq (\delta_L+\delta_H)^{\tau_s}\delta_L^{\tau_f}$$
$$1-\sum_{t=1}^{\tau_s-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t \geq \frac{\gamma}{1-\gamma} \geq 1-\sum_{t=1}^{\tau_s}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t$$

Proof. See the appendix. □

The transition phase can have only a slow part $\tau_s = \tau$ or only a fast part $\tau_f = \tau$. For instance, the latter case happens if $\frac{\gamma}{1-\gamma} \geq 1 \iff \gamma \geq \frac{1}{2}$. In other words, if the objective weakly favors firm A, the optimal transition phase only has a fast part, which also maximizes the industry profits. The next proposition presents the Pareto optimal payoffs for patient firms for a transition phase of length $\tau_f + \tau_s$ and for a starting value of the state variable of $\Pi_{L,1}$. The formulas differ for the industry efficient solution, for which $\tau_s = 0$.

Proposition 4.5. *Pareto optimal profits constructed from a transition phase of length $\tau = \tau_f + \tau_s$ with $\tau_s > 0$ equal:*

$$\Pi_{A,1} = \frac{1-\delta_H}{1-\delta_L}\Pi_{L,1}\sum_{s=0}^{\tau_f}\left(\frac{\delta_H}{\delta_L}\right)^s$$
$$\Pi_{H,1} = \frac{1-\delta_H}{1-\delta_L}\Pi_{L,1}\left(\frac{\delta_H}{\delta_L}\right)^{\tau_f}\sum_{s=1}^{\tau_s-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^s + \delta_H^{\tau+1}v$$

Pareto optimal profits constructed from the industry efficient transition phase of length $\tau_f = \tau$

[23] With the assumption that $\sum_{t=1}^{\tau_s-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t = 0$ if $\tau_s = 0$.

equal:

$$\Pi_{A,1} = \frac{1-\delta_H}{1-\delta_L} \sum_{s=0}^{\tau-1} \left(\frac{\delta_H}{\delta_L}\right)^s \Pi_{L,1} + \delta_H^{\tau+1} v$$

$$\Pi_{H,1} = \delta_H^{\tau}(1-\delta_H) v$$

Proof. See appendices. □

In the solution that maximizes the industry profit, B_H's has 0 market shares during the transition phase and therefore only makes profit from period $\tau + 1$ onward, when the monopoly price is played.

In a separating equilibrium, the starting value of the state variable is determined endogenously. It depends on A's belief ρ_0 about the probability of colluding with B_H. Reducing ρ_0 reduces the highest price sustainable by A during the separation period.

Pareto Frontier from the Separation Period

This subsection presents the Pareto frontier from the separation period given A's belief ρ_0. I assume that separation occurs in period 0; although the period during which firms separate has no impact on the Pareto frontier, I do not discuss the question of the optimal separation period.

During the separation period, B_L undercuts, claims the entire industry profit and obtains 0 profit afterward. Increasing the market price during the separation period benefits B_L, and also benefits patient firms by allowing them to set higher prices on the patient firm path. Therefore, to construct the Pareto frontier, it is sufficient to maximize the weighted sum of A and B_H only since B_L's profit is positively related to the patient firms' profits.

To construct the frontier from the separation period, future profits must be part of the frontier after the separation period. Patient firms' continuation profits from period 1 onward must be a solution to $\mathcal{P}_2$ for a starting value of the state variable endogenously determined in period 0. From $LM(0)$ and $PK(0)$ the value of $\Pi_{L,1}$ is:

$$\begin{aligned}\Pi_{L,1} &= \frac{\Pi_{L,0} - (1-\delta_L)\pi_{B,0}}{\delta_L} \\ &= \frac{(1-\delta_L)\pi_{A,0}}{\delta_L}\end{aligned}$$

Therefore, to construct the Pareto frontier from the separation period, I optimize the weighted sum of A's and B_H's profits from period 0 with a parameter $\gamma \in (0,1)$. Using proposition 4.2, this problem is:

$$\mathcal{P}_3(\gamma):$$

$$\max_{\pi_{H,0},\pi_{A,0},\Pi_{A,1},\Pi_{H,1}} \gamma\left((1-\delta_H)\pi_{A,0} + \delta_H \Pi_{A,1}\right) + (1-\gamma)\left((1-\delta_H)\pi_{H,0} + \delta_H \Pi_{H,1}\right)$$

$$\text{subject to}: \ \pi_{H,0}, \pi_{A,0} \geq 0\ , \quad (\Pi_{A,1}, \Pi_{H,1}) \in F_2\left(\frac{1-\delta_L}{\delta_L}\pi_{A,0}\right)$$

$$(1-\delta_H)(\pi_{H,0} + \pi_{A,0}) \leq (1-\delta_H)v \qquad C_3(0)$$

$$\rho_0\left((1-\delta_H)\pi_{A,0} + \delta_H \Pi_{A,1}\right) \geq (1-\delta_H)(\pi_{H,0} + \pi_{A,0}) \qquad S(A,0)$$

$$\delta_H \Pi_{H,1} \geq (1-\delta_H)\pi_{A,0} \qquad S(H,0)$$

Increasing the price during period 0 and A's market shares increases the value of the state variable and thus both firms' continuation profits from period 1 onward. However, A's sustainability condition restricts the period 0 market price. Since A expects collusion to happen only with probability ρ_0, if the period 0 price is too high, then A might be better-off undercutting this price. In fact if ρ_0 is sufficiently high, i.e., greater than $\frac{\delta_L}{\delta_H^2+\delta_L}$, then the highest price that A can sustain in period 0 allows patient firms to play the monopoly price v from period 1 onward.[24] In this subsection, I focus on the case $\rho_0 < \frac{\delta_L}{\delta_H^2+\delta_L}$, for which, from

[24] See appendices for these conditions and the computations.

period 1, there is a transition phase.

Increasing A's profit in period 0 increases the value of the state variable from period 1, which in turn benefits all players. Therefore, at the solution, B_H's market shares are 0 in period 0, and the period 0 market price is set to the highest level that A can sustain. Profits from period 1 on are taken from the Pareto frontier of $\mathcal{P}_2$; thus, solutions to $\mathcal{P}_3$ feature similar transition phases of length $\tau_f + \tau_f$. However, increasing A's future profits after period 1 relaxes A's sustainability condition in period 0 and therefore allows patient firms to set a higher price in period 0. That is, increasing A's profits after separation increases the value of the state variable from period 1, in turn increasing patient firms' profits.

Proposition 4.6. *Assume $\rho_0 < \frac{\delta_L}{\delta_H^2 + \delta_L}$. All solutions to $\mathcal{P}_3$ feature, from period 1 onward, a transition phase of length $\tau_f + \tau_s$. Where τ_s and τ_f are determined by A's belief:*

$$C(\tau_f, \tau_s) > \rho_0 \geq \frac{1}{1 + \sum_{s=1}^{\tau_f+1} \left(\frac{\delta_H}{\delta_L}\right)^s}$$

Where:

$$C(\tau_f, \tau_s) = \frac{\sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s \left(\frac{\delta_H - \delta_L}{\delta_H \delta_L} - \left(\frac{\delta_H}{\delta_L + \delta_H}\right)^{\tau_s - 1}\right)}{\left(\frac{\delta_H}{\delta_L}\right)^{\tau_f} + \sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s \left(\frac{\delta_H - \delta_L}{\delta_H \delta_L} - \left(\frac{\delta_H}{\delta_L + \delta_H}\right)^{\tau_s - 1}\right)}$$

Proof. See appendices. □

Reducing ρ_0 reduces the period 0 price and the undercutting profit of B_L and, in turn, increases the length of the transition phase.

All in all, there are two mechanisms that produce gradualism in this model. First, patient firms employ a gradually increasing path of prices to separate B_L from B_H. Delaying the period during which the monopoly price is played on the patient firm collusive path reduces the incentives more for B_L to mimic B_H than the profits earned by patient firms.

5 Conclusion

This paper studies a mechanism that explains the gradual increase in prices at the start of collusion. I analyze a repeated Bertrand pricing game, in which one firm is privately informed about its own discount factor. I show that, in pure strategies, Pareto optimal equilibria feature a transition phase during which prices increase gradually after the separation period. It is necessary for patient firms to delay the period during which they play the highest sustainable prices to separate themselves from impatient firms. In addition, I identify two channels that affect the length of the transition phase. First, if A's belief is that B is patient is lower, that is, A expects collusion with a lower probability, then the separation profit of the impatient firm is lower; thus, the transition phase's length increases. Second, the patient firm B (B_H) could reduce its market shares during the transition phase to deter the impatient firm (B_L) from mimicking and accelerating the price increase. However, B_H cannot appropriate all the gains since they are shared with A to ensure sustainability. As a result, B_H might prefer to claim positive market shares during the transition phase, increasing its length.

If B can signal its type, then the first channel that generates the transition phase can be shut down. For instance, we assume that negative (or below cost) prices are feasible. Then, B_H can set a price that is sufficiently low during the first period such that B_L cannot profitably mimic; therefore, patient firms can set the monopoly price from the second period onward. Alternatively, if we assume that B and A can meet and bargain at the beginning of the game, then in this case, B_H can burn an amount of money large enough to signal its type. The resulting equilibrium does not depend on A's beliefs, and there could be an equilibrium path without a transition phase. However, it is still suboptimal for firm B_H to reduce the length of the transition phase beyond the length of the short part τ_s. For B_H, the same trade-off applies: it can either reduce its current market share to accelerate the price increase or set a lower below-cost price to shorten the transition phase.

Date planned of price increase	50% dry	60 % dry	75 % liquid
January 1993	1000	1200	1000
July 1993	1100	1320	1100
January 1994	1100	1320	1200

Table 1.1: Cartel Prices – Choline Chloride.

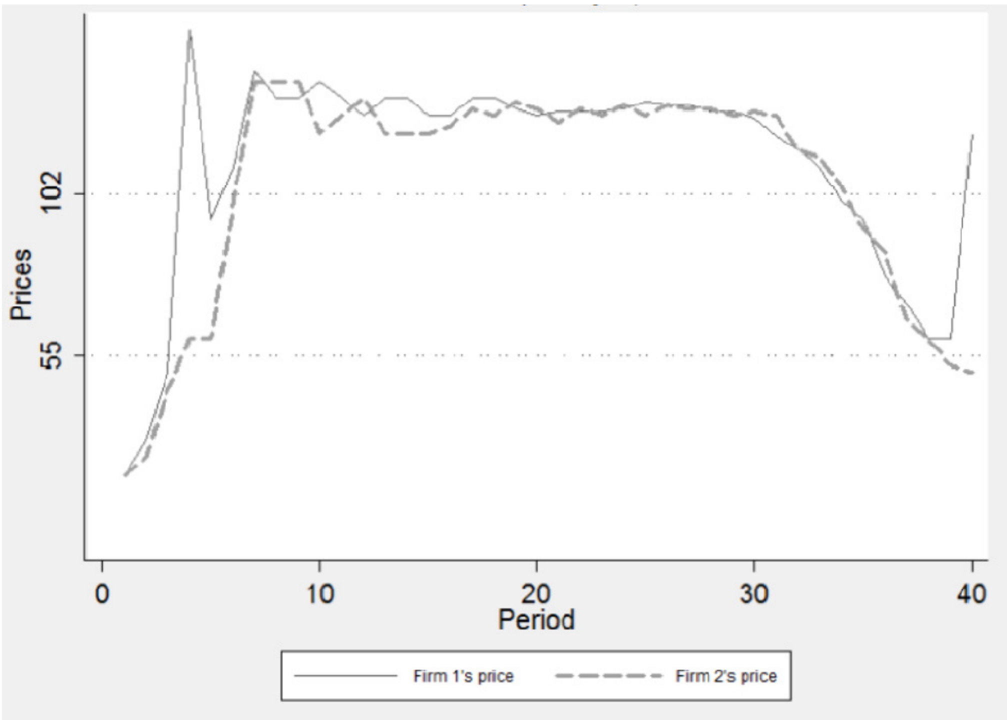


Figure 1.1: Price serie for a symetric duopoly, from Harrington, Gonzales and Kujal (2017)

6 Appendices

6.1 Collusion under Complete Information

Proof of Proposition 3.1

Sufficiency ($\Leftarrow$). Define the following Grim-Trigger strategy for $k = A, B$:

$$\sigma_{GT}^{k} \; : \; h_t \mapsto \begin{cases} 0 & if \; h_t \neq \{p_{A,t}, p_{B,t}\}_{t\geq 0} \\ p_{k,t} & if \; h_t = \{p_{A,t}, p_{B,t}\}_{t\geq 0} \end{cases}$$

Where for all $t \geq 0$:

$$p_{A,t} = p_{B,t} = \pi_{A,t} + \pi_{B,t}$$
$$\alpha_t = \frac{\pi_{A,t}}{\pi_{A,t} + \pi_{B,t}}$$

Since $0 \leq p_{k,t} \leq v$, by following the Grim-Trigger strategy profile each firms obtains in each period a flow profit of $\alpha_t p_{A,t} = \pi_{A,t}$ and $(1-\alpha_t)p_{B,t} = \pi_{B,t}$.

Using the promise keeping conditions $PK(k,t)$ and the fact that $\Pi_{A,t}, \Pi_{B,t}$ are bounded one has that for all t:

$$\Pi_{A,t} = (1-\delta_A)\sum_{s\geq t} \pi_{A,s}$$
$$\Pi_{B,t} = (1-\delta_B)\sum_{s\geq t} \pi_{B,s}$$

That is, the strategy profile defined above achieves the sequence of continuations profits for each firm. It remains to be shown that the strategy profile constitute a SPE of the complete information game. From the one-shot deviation principle, $(\sigma^A_{GT}, \sigma^B_{GT})$ is an SPE if and only if for all t

$$\Pi_{A,t} \geq (1-\delta_A) \sup_{p' \neq p_{A,t}} p' \mathbf{1}_{\{p' < p_{k,t}\}}$$
$$\Pi_{B,t} \geq (1-\delta_B) \sup_{p' \neq p_{B,t}} p' \mathbf{1}_{\{p' < p_{k,t}\}}$$
$$\iff S(A,t), S(B,t).$$

Therefore, the conditions of proposition 3.1 are sufficient for $\{\Pi_{A,t}, \Pi_{B,t}\}_{t\geq 0}$ to be achievable in equilibrium.

Necessity ($\Rightarrow$) Let the sequence $\{\Pi_{A,t}, \Pi_{B,t}\}_{t\geq 0}$ be achievable in equilibrium. That is, there is a strategy profile:

$$k = A, B \qquad \sigma^k \ : \ h_t \mapsto p_k(h_t) \in [0, v]$$

and a map $h_t \mapsto \alpha(h_t)$ such that for all h_t on path (noted h_t^p):

$$\Pi_{A,t} = (1-\delta_A) \sum_{s\geq t} \left(p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) < p_B(h_s^p)\}} + \alpha(h_s^p) p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) = p_B(h_s^p)\}} \right) \qquad (i)$$

$$\Pi_{B,t} = (1-\delta_B) \sum_{s\geq t} \left(p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) < p_A(h_s^p)\}} + (1-\alpha(h_s^p)) p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) = p_A(h_s^p)\}} \right) \qquad (ii)$$

and for all h_t the one shot deviation principle holds. Define for all t:

$$\pi_{A,t} = p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) < p_B(h_s^p)\}} + \alpha(h_s^p) p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) = p_B(h_s^p)\}}$$

$$\pi_{B,t} = p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) < p_A(h_s^p)\}} + (1-\alpha(h_s^p)) p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) = p_A(h_s^p)\}}$$

Notice, $\pi_{k,t} \in [0, v]$ and $\pi_{A,t} + \pi_{B,t} \leq v$.

From (i) and (ii), $\{\Pi_{A,t}, \Pi_{B,t}\}_{t\geq 0}$ is bounded. Further, with the previous definitions of $\pi_{A,t}, \pi_{B,t}$, $\Pi_{A,t}$ and $\Pi_{B,t}$ satisfy $PK(A,t)$ and $PK(B,t)$ for all t.

Last, the one shot deviation principle holds for all h_t, in particular, for all h_t^p:

$$\begin{aligned}
&\Pi_{A,t} \geq (1-\delta_A) \sup_{p' \neq p_A(h_t^p)} p' \mathbf{1}_{\{p' < p_B(h_t^p)\}} \\
&\Pi_{B,t} \geq (1-\delta_B) \sup_{p' \neq p_B(h_t^p)} p' \mathbf{1}_{\{p' < p_A(h_t^p)\}} \\
\Longrightarrow& \\
&\Pi_{A,t} \geq (1-\delta_A)(\pi_{A,t} + \pi_{B,t}) \\
&\Pi_{B,t} \geq (1-\delta_B)(\pi_{A,t} + \pi_{B,t}) \\
\Longleftrightarrow\ & S(A,t), S(B,t).
\end{aligned}$$

$\square$

Proof of Proposition 3.2

Because $S(t)$ is true for all t this implies that for any t:

$$\begin{aligned}
&\begin{pmatrix} \Pi_{A,0} \\ \Pi_{B,0} \end{pmatrix} \leq \begin{pmatrix} \delta_A & \frac{1-\delta_B}{1-\delta_A}\delta_B \\ \frac{1-\delta_A}{1-\delta_B}\delta_A & \delta_B \end{pmatrix}^t \begin{pmatrix} \Pi_{A,t} \\ \Pi_{B,t} \end{pmatrix} \\
\Longleftrightarrow &\begin{pmatrix} \Pi_{A,0} \\ \Pi_{B,0} \end{pmatrix} \leq P \begin{pmatrix} 0 & 0 \\ 0 & (\delta_A + \delta_B)^t \end{pmatrix} P^{-1} \begin{pmatrix} \Pi_{A,t} \\ \Pi_{B,t} \end{pmatrix}
\end{aligned}$$

Where :

$$P = \begin{pmatrix} \delta_A(1-\delta_B) & 1-\delta_B \\ -\delta_B(1-\delta_A) & 1-\delta_A \end{pmatrix}$$

If $\delta_A + \delta_B < 1$, and because continuation profits are bounded, then the only payoff satisfying sustainability conditions is the Nash equilibrium payoff

Pareto Optimal Profits

$$\begin{aligned}
\mathcal{P}_1(\gamma) \ : \ & \max_{\{\Pi_{A,t},\Pi_{B,t}\}_{t\geq 0}} \gamma\Pi_{A,0} + (1-\gamma)\Pi_{B,0} \\
& \text{subject to: } \Pi_{A,t} \geq \delta_H \Pi_{A,t+1} \quad \Pi_{B,t} \geq \delta_H \Pi_{B,t+1} && C_1(t), \quad C_2(t) \\
& \Pi_{A,t} + \Pi_{B,t} \leq (1-\delta_H)v + \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] && C_3(t) \\
& \Pi_{B,t} \leq \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] && S(A,t) \\
& \Pi_{A,t} \leq \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] && S(B,t)
\end{aligned}$$

I neglect the transversality condition, I rule out ex-post unbounded solutions to the problem. The problem $\mathcal{P}_1(\gamma)$ is linear and can be solved using a dual approach. The dual problem is:

$$\begin{aligned}
\mathcal{D}_1(\gamma) \ : \ & \min_{\{\lambda(t)\}_{t\geq 0}} (1-\delta_H)v \sum_{t\geq 0} \delta_H^t \lambda_3(t) \\
& \text{subject to: } \gamma + \lambda_1(0) = \lambda_3(0) + \lambda_B(0) && C_4(0) \\
& 1-\gamma+\lambda_2(0) = \lambda_3(0) + \lambda_A(0) && C_5(0) \\
& \forall t \geq 1 : \\
& \lambda_1(t) + \lambda_3(t-1) + \lambda_A(t-1) + \lambda_B(t-1) = \lambda_1(t-1) + \lambda_3(t) + \lambda_B(t) && C_4(t) \\
& \lambda_2(t) + \lambda_3(t-1) + \lambda_A(t-1) + \lambda_B(t-1) = \lambda_2(t-1) + \lambda_3(t) + \lambda_A(t) && C_5(t)
\end{aligned}$$

The next proposition gives the optimality conditions:

Proposition 6.1. *If feasible sequences of primal and dual variables*

$\{\Pi_{A,t}, \Pi_{B,t}, \lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_A(t), \lambda_B(t)\}_{t\geq 0}$ *are such that:*

$$\lambda_1(t)(\Pi_{A,t} - \delta_H \Pi_{A,t+1}) = 0 \qquad CS_1(t)$$

$$\lambda_2(t)(\Pi_{B,t} - \delta_H \Pi_{B,t+1}) = 0 \qquad CS_2(t)$$

$$\lambda_3(t)((1-\delta_H)v + \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] - \Pi_{A,t} - \Pi_{B,t}) = 0 \qquad CS_3(t)$$

$$\lambda_A(t)(\delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] - \Pi_{B,t}) = 0 \qquad CS_A(t)$$

$$\lambda_B(t)(\delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] - \Pi_{A,t}) = 0 \qquad CS_B(t)$$

Then $\{\Pi_{A,t}, \Pi_{B,t}, \lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_A(t), \lambda_B(t)\}_{t\geq 0}$ *is a solution to the primal and dual problem.*

First, the following lemma shows how constraints interact.

Lemma 6.2. *1. $C_1(t)$ and $S(B,t)$ can't be both binding.*

2. $C_2(t)$ and $S(A,t)$ can't be both binding.

3. If $\Pi_{A,t+1} + \Pi_{B,t+1} < \frac{1-\delta_H}{\delta_H}v$ then $S(A,t)$ and $S(B,t)$ implies $C_3(t)$.

4. If $\Pi_{A,t+1} + \Pi_{B,t+1} \geq \frac{1-\delta_H}{\delta_H}v$ then $S(A,t)$ and $S(B,t)$ can't be both binding.

5. If $\Pi_{A,t+1} \geq \frac{1-\delta_H}{\delta_H}v$ then $C_1(t)$ and $C_3(t)$ implies $S(A,t)$. If $\Pi_{A,t+1} < \frac{1-\delta_H}{\delta_H}v$ then $C_1(t)$ and $C_3(t)$ can't be both binding.

6. If $\Pi_{B,t+1} \geq \frac{1-\delta_H}{\delta_H}v$ $C_2(t)$ and $C_3(t)$ implies $S(B,t)$. If $\Pi_{B,t+1} < \frac{1-\delta_H}{\delta_H}v$ $C_2(t)$ and $C_3(t)$ can't be both binding.

Proof. 1. If it is the case, then $\Pi_{B,t+1} = 0$ which implies that $\Pi_{A,t+1} = 0$.

2. Same as 1.

3. Summing $S(A,t)$ and $S(B,t)$ yields:

$$\Pi_{A,t} + \Pi_{B,t} \leq 2\delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] \qquad (S)$$

(S) implies $C_3(t)$ if $\Pi_{A,t+1} + \Pi_{B,t+1} < \frac{1-\delta_H}{\delta_H}v$. 4. Consider (S) binding, $C_3(t)$ is violated if $\Pi_{A,t+1} + \Pi_{B,t+1} \geq \frac{1-\delta_H}{\delta_H}v$.
5. Summing $C_1(t)$ and $C_3(t)$ yields:

$$\Pi_{B,t} \leq (1-\delta_H)v + \delta_H \Pi_{B,t+1} \qquad C_{12}$$

If $\Pi_{A,t+1} \geq \frac{1-\delta_H}{\delta_H}v$ then C_{12} implies $S(A,t)$. If $\Pi_{A,t+1} < \frac{1-\delta_H}{\delta_H}v$ then C_{12} binding violates $S(A,t)$. □

Consider the case where $\Pi_{A,t} + \Pi_{B,t} \geq \frac{1-\delta_H}{\delta_H^2}v$, for all t, that is point 4. of lemma $A.1$ applies.

Assume $\gamma > \frac{1}{2}$.

For $C_4(0)$ and $C_5(0)$ to be satisfied $\lambda_3(0)$ must be positive, since from lemma $A.1$ 4., $\lambda_A(0)$ and $\lambda_B(0)$ can't be both positive. From 6., either $\lambda_B(0) > 0$ if $\Pi_{B,t+1} < \frac{1-\delta_H}{\delta_H}v$ (case 1(0)) or $\lambda_2(0) >$ if $\Pi_{B,t+1} \geq \frac{1-\delta_H}{\delta_H}v$ (case 2(0)).
In case 1(0) one has: $\lambda_3(0) = \gamma$ and $\lambda_2(0) = 2\gamma - 1$, and $\Pi_{A,0} = \delta_H(\Pi_{A,1} + \Pi_{B,1})$ and $\Pi_{B,0} = (1-\delta_H)v$.
In case 2(0) one has: $\lambda_3(0) = 1-\gamma$ and $\lambda_B(0) = 2\gamma - 1$, and $\Pi_{A,0} = (1-\delta_H)v + \delta_H \Pi_{A,1}$ and $\Pi_{B,0} = \delta_H \Pi_{B,1}$ $(\geq (1-\delta_H)v)$ by assumption.

In case 1(0), $C_4(1)$ and $C_5(1)$ becomes:

$$\lambda_1(1) + \gamma = \lambda_3(1) + \lambda_A(1)$$

$$\lambda_2(t) + \gamma = \lambda_3(1) + \lambda_B(1)$$

From the same arguments of lemma $A.1$ 1., 2. and 4. $\lambda_3(1)$ must be positive and equal to γ. Iterating for all t $C_4(t)$ and $C_5(t)$ are the same as $C_4(1)$ and $C_5(1)$, therefore:

$$\lambda_B(0) = 2\gamma - 1$$

$$\lambda_3(0) = 1 - \gamma$$

$$\lambda_3(t) = \gamma \;\; \forall\, t \geq 0$$

All other dual variables equal 0

From $CS_3(t)$, $C_3(t)$ binds for all t:

$$\forall t \geq 0 \qquad \Pi_{A,t} + \Pi_{B,t} = (1 - \delta_H)v + \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}]$$

Because the sequence is bounded, $\Pi_{A,t} + \Pi_{B,t} = v$ for all $t \geq 0$, and so not smaller than $\frac{1-\delta_H}{\delta_H}v$ as assumed previously.

From $CS_B(0)$, $S(B,0)$ binds:

$$\begin{aligned}\Pi_{A,0} &= \delta_H(\Pi_{A,1} + \Pi_{B,1}) \\ &= \delta_H v\end{aligned}$$

Hence, from $C_3(0)$: $\Pi_{B,0} = (1 - \delta_H)v$.

In case 2(0), $C_4(1)$ and $C_5(1)$ are thus the same as $C_4(0)$ and $C_5(0)$:

$$\lambda_1(1) + \gamma = \lambda_3(1) + \lambda_B(1)$$

$$\lambda_2(1) + 1 - \gamma = \lambda_3(1) + \lambda_A(1).$$

Hence the same reasoning applies. Suppose the solution stays in case 1 for all t, then:

$$\Pi_{A,t} = (1-\delta_H)v + \Pi_{A,t+1}$$
$$\Pi_{B,t} = \delta_H \Pi_{B,t+1}$$

Since sequences are bounded that means:

$$\Pi_{A,t} = v$$
$$\Pi_{B,t} = 0$$

Which is unfeasible.

Instead, consider a solution of type in case 2 for τ periods then of type case 1.[25] One has:

$$\text{For } 0 \leq t \leq \tau :$$
$$\Pi_{A,t} = \delta_H[\Pi_{A,t+1} + \Pi_{B,t+1}]$$
$$\Pi_{B,t} = \delta_H \Pi_{B,t+1} \geq (1-\delta_H)v$$
$$\text{For } t \geq \tau+2 :$$
$$\Pi_{A,t} + \Pi_{B,t} = v$$
$$\Pi_{A,\tau+1} = \delta_H v$$
$$\Pi_{B,\tau+1} = (1-\delta_H)v$$

To be in case 2 for τ implies $\Pi_{B,\tau+1} \geq \frac{1-\delta_H}{\delta_H^{\tau+1}}v$. Feasible if $1 \geq \frac{1}{\delta_H^{\tau+1}}$, that is only if $\tau = 0$.

[25]Once the solution reaches case 1 it stays in case 1 for subsequent periods.

Therefore, if $\gamma > \frac{1}{2}$ all feasible profit sequences such that:

$$\Pi_{A,0} = \delta_H v$$
$$\Pi_{B,0} = (1-\delta_H)v$$
$$\Pi_{A,t} + \Pi_{B,t} = v \qquad \forall t \geq 1$$

are solutions to $\mathcal{P}(\gamma)$.

For $\gamma < \frac{1}{2}$ symmetric arguments holds. All feasible sequences such that:

$$\Pi_{A,0} = (1-\delta_H)v$$
$$\Pi_{B,0} = \delta_H v$$
$$\Pi_{A,t} + \Pi_{B,t} = v \qquad \forall t \geq 1$$

are solutions to $\mathcal{P}(\gamma)$.

If $\gamma = \frac{1}{2}$, then only $\lambda_3(t)$ is positive for all $t \geq 0$. Therefore $\Pi_{A,t} + \Pi_{B,t} = v$ for all t. A feasible solution such that for all $t \geq 0$: $\{(v - \Pi_{B,t}, \Pi_{B,t}) \mid \Pi_{B,t} \in [(1-\delta_H)v, \delta_H v]\}$ satisfy the complementary slackness condition and thus is solution to $\mathcal{P}(\gamma)$.

6.2 Collusion under Incomplete Information

Proof of proposition 4.1

In a pooling equilibrium, for all h_t^p $\sigma_L(h_t^p) = \sigma_H(h_t^p)$. Firm's A profit on path does not depend on B's type, nor on its beliefs about B's type.

Therefore, lemma 3.1 applies to pooling equilibria. A bounded sequence of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}, \Pi_{L,t}\}_{t\geq 0}$ is a achievable in a pooling BPE if and only if for all t there are

$\pi_{A,t}, \pi_{B,t} \geq 0$ with $\pi_{A,t} + \pi_{B,t} \geq 0$ such that:

$$\Pi_{A,t} = (1-\delta_A)\pi_{A,t} + \delta_A \Pi_{A,t+1} \qquad PK(A,t)$$

$$\Pi_{H,t} = (1-\delta_H)\pi_{B,t} + \delta_H \Pi_{H,t+1} \qquad PK(H,t)$$

$$\Pi_{L,t} = (1-\delta_L)\pi_{B,t} + \delta_L \Pi_{L,t+1} \qquad PK(L,t)$$

$$\Pi_{A,t} \geq (1-\delta_A)(\pi_{A,t} + \pi_{B,t}) \qquad S(A,t)$$

$$\Pi_{H,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) \qquad S(H,t)$$

$$\Pi_{L,t} \geq (1-\delta_L)(\pi_{A,t} + \pi_{B,t}) \qquad S(L,t)$$

Because $\delta_L + \delta_A < \frac{1}{2}$, the only sequence of continuation profits that satisfies all conditions is 0 for all t.

Therefore, equilibria that are pooling on path yield 0 profits for all firms.

Proof of proposition 4.2

Necessity $(\Rightarrow)$ Let the sequences of continuation profits $\{\Pi_{A,t}, \Pi_{H,t}\}_{t\geq 0}$, $\{\Pi_{L,t}\}_{0\leq t\leq T}$ be achievable in equilibrium. That is, there is a strategy profile:

$$k = A, H, L \qquad \sigma_k \ : \ h_t \mapsto p_k(h_t) \in [0, v]$$

and a map $h_t \mapsto \alpha(h_t)$ such that for all h_t on the collusive path (noted h_t^p):

$$\Pi_{A,t} = \begin{cases} (1-\delta_H) \sum_{s \geq t} \delta_H^{s-t} (\mathbf{1}_{\{s<T\}} + \rho_0 \mathbf{1}_{\{s \geq T\}}) \left(p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) < p_B(h_s^p)\}} + \alpha(h_s^p) p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) = p_B(h_s^p)\}} \right) & \text{for} \\ (1-\delta_H) \sum_{s \geq t} \delta_H^{s-t} \left(p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) < p_B(h_s^p)\}} + \alpha(h_s^p) p_A(h_s^p) \mathbf{1}_{\{p_A(h_s^p) = p_B(h_s^p)\}} \right) & \text{for} \end{cases}$$

$$\Pi_{H,t} = (1-\delta_H) \sum_{s \geq t} \delta_H^{s-t} \left(p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) < p_A(h_s^p)\}} + (1-\alpha(h_s^p)) p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) = p_A(h_s^p)\}} \right)$$

$$\Pi_{L,t} = (1-\delta_L) \sum_{t \leq s \leq T-1} \delta_L^{s-t} \left(p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) < p_A(h_s^p)\}} + (1-\alpha(h_s^p)) p_B(h_s^p) \mathbf{1}_{\{p_B(h_s^p) = p_A(h_s^p)\}} \right)$$
$$+ (1-\delta_L) \delta_L^T \min\{p_A(h_T^p), p_B(h_T^p)\}$$

For each t define:

$$\pi_{A,t} + \pi_{B,t} = \Pi_{A,t} + \Pi_{B,t} - \delta_H [\Pi_{A,t+1} + \Pi_{B,t+1}]$$
$$\pi_{A,t} = \frac{\Pi_{A,t} - \delta_H \Pi_{A,t+1}}{\Pi_{A,t} + \Pi_{B,t} - \delta_H [\Pi_{A,t+1} + \Pi_{B,t+1}]}$$

Thus $\pi_{A,t}, \pi_{B,t} \geq 0$ and $\pi_{A,t} + \pi_{B,t} \leq v$. Moreover one has for all t:

$$\begin{aligned}
\Pi_{H,t} &= (1-\delta_H)\sum_{s\geq t}\delta_H^{s-t}\pi_{B,s} \\
&= (1-\delta_H)\pi_{B,t} + \delta_H \Pi_{H,t+1}
\end{aligned}$$

$$\begin{aligned}
\Pi_{A,t} &= (1-\delta_H)\sum_{s\geq t}\delta_H^{s-t}\pi_{A,s} \\
&= (1-\delta_H)\sum_{s\geq t}\delta_H^{s-t}\pi_{A,s} \\
&= \begin{cases} (1-\delta_H)\pi_{A,t} + \delta_H \Pi_{A,t+1} & t \neq T \\ \rho_0((1-\delta_H)\pi_{A,T} + \delta_H \Pi_{A,T+1}) & \end{cases} \\
\Pi_{L,t} &= (1-\delta_L)\sum_{0\leq s\leq T-1}\delta_L^{s-t}\pi_{B,s} + \delta_L^T(\pi_{A,t} + \pi_{B,t}) \\
&= \begin{cases} (1-\delta_L)\pi_{B,t} + \delta_L \Pi_{L,t+1} & \text{for } t \leq T-1 \\ (1-\delta_L)(\pi_{A,T} + \pi_{B,T}) & \end{cases}
\end{aligned}$$

Hence, promise keeping conditions hold.

For these sequences to be achievable in equilibrium, the one shot deviation principle must hold, at best assuming min-max punishment payoff of 0. On the collusive path this implies:

$$\begin{aligned}
\Pi_{B,t} &\geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) && \forall t \\
\Pi_{A,t} &\geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t}) && \forall t \\
\Pi_{L,t} &\geq (1-\delta_L)(\pi_{A,t} + \pi_{B,t}) && \text{for } 0 \leq t < T
\end{aligned}$$

Consider the best-response condition at h_T^p for B_L. B_L must find it profitable to undercut at T rather than wait more periods and undercut later, that is for all $t \geq 0$:

$$(1-\delta_L)(\pi_{A,T}+\pi_{B,T}) \geq (1-\delta_L)\sum_{0\leq s\leq t}\delta_L^s\pi_{B,T+s} + (1-\delta_L)\delta_L^{T+t}(\pi_{A,T+t}+\pi_{B,T+t})$$

$$\Longleftrightarrow$$

$$\Pi_{L,t} \geq (1-\delta_L)\sum_{0\leq s\leq t}\delta_L^s\pi_{B,T+s} + (1-\delta_L)\delta_L^{T+t}(\pi_{A,T+t}+\pi_{B,T+t})$$

Define $\{\Pi_{L,t}\}_{t\geq T+1}$:

$$\Pi_{L,t} \overset{def}{=} \frac{1}{\delta_L^t}\left(\Pi_{L,T} + (1-\delta_L)\sum_{T\leq s\leq t}\delta_L^{s-T}\pi_{B,T+s}\right)$$

$$\Longleftrightarrow$$

$$\Pi_{L,t} = (1-\delta_L)\pi_{B,t} + \delta_L\Pi_{L,t+1}$$

Rewritting best response conditions at h_T^p, for all $t \geq T+1$:

$$\Pi_{L,t} \leq (1-\delta_L)(\pi_{A,t}+\pi_{B,t})$$

Sufficiency ($\Leftarrow$) Consider the following strategy profile:

$$(i)\ \sigma_A \ :\ h_t \mapsto \begin{cases} 0 & \text{if } h_t \neq \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \\ p_{A,t} & \text{if } h_t = \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \end{cases}$$

$$(ii)\ \sigma_H \ :\ h_t \mapsto \begin{cases} 0 & \text{if } h_t \neq \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \\ p_{B,t} & \text{if } h_t = \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \end{cases}$$

$$(iii)\ \sigma_L \ :\ h_t \mapsto \begin{cases} 0 & \text{if } h_t \neq \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \\ p_{B,t} & \text{if } h_t = \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \quad t < T \\ p^u_{A,t} \text{ or } p_{B,t} & \text{if } h_t = \{p_{A,s}, p_{B,s}\}_{0\leq s\leq t} \quad t \geq T \end{cases}$$

Such that for all $t \geq 0$:

$$p_{A,t} = p_{B,t} = \pi_{A,t} + \pi_{B,t}$$
$$\alpha_t = \frac{\pi_{A,t}}{\pi_{A,t} + \pi_{B,t}}$$

The strategy profile plays the NE repeatedly off-path which is BPE continuation. On-path, following the strategy is better than one shot deviation if:

$$\Pi_{A,t} \geq (1-\delta_H) \sup_{p' \neq p_A(h_t^p)} p' \mathbf{1}_{\{p' < p_B(h_t^p)\}}$$
$$\Pi_{H,t} \geq (1-\delta_H) \sup_{p' \neq p_B(h_t^p)} p' \mathbf{1}_{\{p' < p_A(h_t^p)\}}$$
$$\Longrightarrow$$
$$\Pi_{A,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t})$$
$$\Pi_{H,t} \geq (1-\delta_H)(\pi_{A,t} + \pi_{B,t})$$
$$\Longleftrightarrow S(A,t), S(H,t).$$

For B_L for $t < T$:

$$\begin{aligned}
&\Pi_{L,t} \geq (1-\delta_L) \sup_{p' \neq p_B(h_t^p)} p' \mathbf{1}_{\{p' < p_A(h_t^p)\}} \\
\implies& \\
&\Pi_{L,t} \geq (1-\delta_L)(\pi_{A,t} + \pi_{B,t}) \\
\iff& S(A,t), S(B,t).
\end{aligned}$$

On path after separation equilibrium, B_L's strategy isn't specified. At h_T^p, undercutting A's price must yields higher profits than playing as B_H and then playing the best continuation. That is, undercutting at h_T^p must yield more profit than undercutting in h_t^p for all $t > T$:

$$(1-\delta_L)(\pi_{A,t} + \pi_{B,t}) \geq (1-\delta_L) \sum_{T \leq s < t} \delta_l^{s-T} \pi_{B,s} + \delta_L^{t-T}(1-\delta_L)(\pi_{A,t} + \pi_{B,t})$$

Defining state variables as:

$$\begin{aligned}
\Pi_{L,T+1} &= \frac{\Pi_{L,T} - (1-\delta_L)\pi_{B,T}}{\delta_L} \\
\Pi_{L,t+1} &= \frac{\Pi_{L,t} - (1-\delta_L)\pi_{B,t}}{\delta_L} \qquad \text{for } t > T
\end{aligned}$$

Previous conditions boils down to:

$$\Pi_{L,t} \geq (1-\delta_L)(\pi_{A,t} + \pi_{B,t}) \qquad \forall t > T$$

□

Pareto Frontier

$$\begin{aligned}
\mathcal{P}_2(\gamma)\ :\ \max_{\{\Pi_{A,t},\Pi_{B,t},\Pi_{L,t}\}_{t\geq 0}} & \gamma\Pi_{A,0} + (1-\gamma)\Pi_{B,0} & \\
\text{subject to: } & \Pi_{A,t} \geq \delta_H \Pi_{A,t+1} \quad \Pi_{B,t} \geq \delta_H \Pi_{B,t+1} & C_1(t), \quad C_2(t) \\
& \Pi_{A,t} + \Pi_{B,t} \leq (1-\delta_H)v + \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] & C_3(t) \\
& \Pi_{B,t} \leq \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] & S(A,t) \\
& \Pi_{A,t} \leq \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] & S(B,t) \\
& \delta_L \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t+1} = \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t} - [\Pi_{B,t} - \delta_H \Pi_{B,t+1}] & LM(t) \\
& \Pi_{A,t} + \Pi_{B,t} \leq \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t} + \delta_H(\Pi_{A,t+1} + \Pi_{B,t+1}) & IC(t)
\end{aligned}$$

Lemma 6.3. *1. If $\Pi_{L,t} \geq (1-\delta_L)v$ then $C_3(t)$ implies $IC(t)$. If $\Pi_{L,t} < (1-\delta_L)v$ then $IC(t)$ implies $C_3(t)$.*

2. If $\frac{1-\delta_H}{1-\delta_L}\Pi_{L,t} \geq \delta_H(\Pi_{A,t+1} + \Pi_{B,t+1})$ then $S(A,t) + S(B,t)$ implies $IC(t)$. If $\frac{1-\delta_H}{1-\delta_L}\Pi_{L,t} < \delta_H(\Pi_{A,t+1} + \Pi_{B,t+1})$ then IC(t) implies $S(A,t) + S(B,t)$.

3. If $\delta_H \Pi_{A,t+1} \geq \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t}$ then $C_1(t)$ and $IC(t)$ implies $S(A,t)$. If $\delta_H \Pi_{A,t+1} < \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t}$ then $C_1(t)$ and $IC(t)$ can't be both binding.

4. If $\delta_H \Pi_{B,t+1} \geq \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t}$ $C_2(t)$ and $IC(t)$ implies $S(B,t)$. If $\delta_H \Pi_{B,t+1} < \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t}$ $C_2(t)$ and $IC(t)$ can't be both binding.

Proof. 1. $IC(t)$ is more stringent than $C_3(t)$ iff $\frac{1-\delta_H}{\delta_L}\Pi_{L,t} < (1-\delta_L)v$.

2., 3., 4. and 5. are the same proof as lemma A.1 replacing $(1-\delta_H)v$ by $\frac{1-\delta_H}{\delta_L}\Pi_{L,t}$.

□

Proposition 6.4. *1. If $\Pi_{L,0} \geq (1-\delta_L)v$ then the solution to $\mathcal{P}_2(\gamma)$ is the same as $\mathcal{P}_1(\gamma)$.*

2. If $\Pi_{L,0} = 0$ then the only solution is $\Pi_{A,t} = \Pi_{B,t} = \hat{\Pi}_{L,t} = 0$ for all $t \geq 0$.

Proof. 1. From lemma B.1 6. and 1., $LM(t)$ and $IC(t)$ are redundant and thus $\mathcal{P}_2(\gamma)$ coincides with $\mathcal{P}_1(\gamma)$.

2. If $\Pi_{L,0} = 0$, by $IC(0)$, $C_1(0)$ and $C_2(0)$ then $\Pi_{L,1} = 0$, $\Pi_{A,0} = \delta_H \Pi_{A,1}$ and $\Pi_{B,0} = \delta_H \Pi_{B,1}$. Iterating for future periods and since profits are bounded implies that $\Pi_{A,t} = \Pi_{B,t} = \hat{\Pi}_{L,t} = 0$ for all $t \geq 0$. □

$\mathcal{P}_2(\gamma)$ is a linear problem solved by duality:

$$\mathcal{D}_2(\gamma) \ : \ \min_{\{\lambda(t)\}_{t\geq 0}} \frac{1-\delta_H}{1-\delta_L} \Pi_{L,0}(\Lambda_{IC}(0) + \lambda_{LM}(0)) + (1-\delta_H) v \sum_{t \geq 0} \delta_H^t \lambda_3(t)$$

$$\begin{aligned}
\text{subject to: } & \gamma + \lambda_1(0) = \lambda_{IC}(0) + \lambda_3(0) + \lambda_B(0) && C_4(0) \\
& 1 - \gamma + \lambda_2(0) - \lambda_{LM}(0) = \lambda_{IC}(0) + \lambda_3(0) + \lambda_A(0) && C_5(0) \\
& \forall t \geq 1 : \\
& \lambda_1(t) + \lambda_{IC}(t-1) + \lambda_3(t-1) + \lambda_A(t-1) + \lambda_B(t-1) \\
= & \lambda_1(t-1) + \lambda_{IC}(t) + \lambda_3(t) + \lambda_B(t) && C_4(t) \\
& \lambda_2(t) - \lambda_{LM}(t) + \lambda_{IC}(t-1) + \lambda_3(t-1) + \lambda_A(t-1) + \lambda_B(t-1) \\
= & \lambda_2(t-1) - \lambda_{LM}(t-1) + \lambda_{IC}(t) + \lambda_3(t) + \lambda_A(t) && C_5(t) \\
& \lambda_{IC}(t) + \lambda_{LM}(t) = \frac{\delta_L}{\delta_H} \lambda_{LM}(t-1) && C_6(t)
\end{aligned}$$

The next proposition gives the optimality conditions:

Proposition 6.5. *If feasible sequences of primal and dual variables*

$\{\Pi_{A,t}, \Pi_{B,t}, \lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_A(t), \lambda_B(t)\}_{t\geq 0}$ *are such that:*

$$\lambda_1(t)(\Pi_{A,t} - \delta_H \Pi_{A,t+1}) = 0 \qquad CS_1(t)$$

$$\lambda_2(t)(\Pi_{B,t} - \delta_H \Pi_{B,t+1}) = 0 \qquad CS_2(t)$$

$$\lambda_3(t)((1-\delta_H)v + \delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] - \Pi_{A,t} - \Pi_{B,t}) = 0 \qquad CS_3(t)$$

$$\lambda_A(t)(\delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] - \Pi_{B,t}) = 0 \qquad CS_A(t)$$

$$\lambda_B(t)(\delta_H[\Pi_{B,t+1} + \Pi_{A,t+1}] - \Pi_{A,t}) = 0 \qquad CS_B(t)$$

$$\lambda_{IC}(t)\left(\frac{1-\delta_H}{1-\delta_L}\Pi_{L,t} + \delta_H(\Pi_{A,t+1} + \Pi_{B,t+1}) - \Pi_{A,t} + \Pi_{B,t}\right) = 0 \qquad CS_{IC}(t)$$

Then $\{\Pi_{A,t}, \Pi_{B,t}, \lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_A(t), \lambda_B(t)\}_{t\geq 0}$ *is a solution to the primal and dual problem.*

First, I express Λ_{LM} as a function of λ_{IC} using C_6. Define for all $t \geq 0$ $C_6'(t)$:

$$C_6'(t) \iff \sum_{s\geq t+1} \left(\frac{\delta_H}{\delta_L}\right)^{s-t} C_6(s)$$
$$\iff \sum_{s\geq t+1} \left(\frac{\delta_H}{\delta_L}\right)^{s-t} \lambda_{IC}(s) = \lambda_{LM}(t)$$

$C_6'(t)$ has the following interpretation: Increasing $\Pi_{L,t}$ relaxes all future incentive compatibility conditions at rate $\frac{1}{\delta_L}$ with benefit discounted by δ_H.

Assume $\Pi_{L,0} < (1-\delta_L)v$ and $\frac{1-\delta_H}{1-\delta_L}\Pi_{L,0} < \delta_H(\Pi_{A,1} + \Pi_{B,1})$. Consider $C_4(0)$ and $C_5(0)$. By lemma $B.1$ $\lambda_{IC}(0)$ must be positive. From $C_6'(t)$ $\lambda_{LM}(0) \geq 0$. If $1-\gamma-\lambda_{LM}(0) < \gamma$ there are two cases:

1. $\lambda_B(0)$ is positive and $\delta_H \Pi_{B,1} < \frac{1-\delta_H}{1-\delta_L}\Pi_{L,0}$

2. $\lambda_2(0)$ is positive and $\delta_H \Pi_{B,1} \geq \frac{1-\delta_H}{1-\delta_L}\Pi_{L,0}$

If $1-\gamma-\lambda_{LM}(0)>\gamma$ the only possibility is:

3. $\lambda_A(0)>0$

Assuming case 1., one has:

$$\lambda_{IC}(0)+\lambda_B(0)=\gamma$$
$$\lambda_{IC}(0)+\lambda_{LM}(0)=1-\gamma$$

For primal variables $IC(0)$ and $S(B,0)$ binds:

$$\Pi_{A,0}=\delta_H[\Pi_{A,1}+\Pi_{B,1}]$$
$$\Pi_{B,0}=\frac{1-\delta_H}{1-\delta_L}\Pi_{L,0}$$
$$\Pi_{L,1}=\frac{\delta_H(1-\delta_L)}{\delta_L(1-\delta_H)}\Pi_{B,1}$$

$C_4(1)$ and $C_5(1)$ thus become:

$$\lambda_1(1)+\gamma=\lambda_{IC}(1)+\lambda_B(1)+\lambda_3(1)$$
$$\lambda_2(1)+\frac{\delta_H-\delta_L}{\delta_H}\lambda_{LM}(0)+\gamma=\lambda_A(1)+\lambda_3(1)$$

Thus, $\lambda_{IC}(1)$ must be 0, and so is $\lambda_{LM}(0)$. From period 1 on the solution is as in complete information:

$$\Pi_{A,1}+\Pi_{B,1}=v$$
$$(1-\delta_H)v\leq\Pi_{B,1}\leq\min\left\{\delta_H v,\frac{1-\delta_H}{\delta_H(1-\delta_L)}\Pi_{L,0}\right\}$$
$$\lambda_3(1)=\gamma$$

From the assumption of case 1:

$$\gamma \geq \frac{1}{2}$$
$$\delta_H(1-\delta_L)v \leq \Pi_{L,0} < (1-\delta_L)v$$

Assuming case 2.

$$\begin{aligned}
\lambda_{IC}(0) &= \gamma \\
\lambda_2(0) - \lambda_{LM}(0) &= 2\gamma - 1 \\
\Pi_{A,0} &= \frac{1-\delta_H}{1-\delta_L}\Pi_{L,0} + \delta_H \Pi_{A,1} \\
\Pi_{H,0} &= \delta_H \Pi_{B,1} \\
\Pi_{L,1} &= \frac{\Pi_{L,0}}{\delta_L}
\end{aligned}$$

$C_4(1)$ and $C_5(1)$ becomes:

$$\begin{aligned}
\lambda_1(1) + \gamma &= \lambda_{IC}(1) + \lambda_B(1) + \lambda_3(1) \\
\lambda_2(1) - \lambda_{LM}(1) + 1 - \gamma &= \lambda_{IC}(1) + \lambda_3(1) + \lambda_A(1)
\end{aligned}$$

Thus are the same as in $t = 0$. If one stay in this case for τ one has for $0 \le t \le \tau - 1$

$$\begin{aligned}
\Lambda_{IC}(t) &= \gamma \\
\lambda_{LM}(t) &= \gamma \sum_{s=t+1}^{\tau-2} \left(\frac{\delta_H}{\delta_L}\right)^{s-t} + \left(\frac{\delta_H}{\delta_L}\right)^{\tau-t-1} \lambda_{LM}(\tau-1) \\
\lambda_2(t) &= -1 + 2\gamma + \gamma \sum_{s=t+1}^{\tau-2} \left(\frac{\delta_H}{\delta_L}\right)^{s-t} + \left(\frac{\delta_H}{\delta_L}\right)^{\tau-t-1} \lambda_{LM}(\tau-1) \\
\Pi_{A,0} &= \frac{1-\delta_H}{1-\delta_L} \sum_{s=0}^{\tau-1} \left(\frac{\delta_H}{\delta_L}\right)^{s} \Pi_{L,0} + \delta_H^{\tau+1} v \\
\Pi_{H,0} &= \delta_H^{\tau}(1-\delta_H) v \\
\Pi_{L,\tau} &= \frac{\Pi_{L,0}}{\delta_L^{\tau}} \\
1 - \gamma - \lambda_{LM}(t) &< \gamma
\end{aligned}$$

From the assumptions of case 2 that means:

$$\begin{aligned}
\lambda_2(t) = -1 + 2\gamma + \gamma \sum_{s=t+1}^{\tau-1} \left(\frac{\delta_H}{\delta_L}\right)^{s-t} + \left(\frac{\delta_H}{\delta_L}\right)^{\tau-t} \lambda_{LM}(\tau) &\ge 0 \\
\delta_H \Pi_{H,t+1} &\ge \frac{1-\delta_H}{1-\delta_L} \Pi_{L,t} \\
\delta_H [\Pi_{A,t+1} + \Pi_{H,t+1}] &> \frac{1-\delta_H}{1-\delta_L} \Pi_{L,t} \\
1 - \gamma - \lambda_{LM}(t) &< \gamma
\end{aligned}$$

Suppose at τ $\frac{\Pi_{L,0}}{\delta_L^\tau} = \Pi_{L,\tau} \geq (1-\delta_L)v$ that means:

$$\begin{aligned}
\lambda_2(0) &= -1 + 2\gamma + \gamma \sum_{s=0}^{\tau-1} \left(\frac{\delta_H}{\delta_L}\right)^s \\
\Pi_{A,0} &= \frac{1-\delta_H}{1-\delta_L} \sum_{s=0}^{\tau-1} \left(\frac{\delta_H}{\delta_L}\right)^s \Pi_{L,0} + \delta_H^{\tau+1} v \\
\Pi_{H,0} &= \delta_H^\tau (1-\delta_H) v \\
\gamma &\geq \frac{1}{2} \\
\delta_L^{\tau-1}\delta_H(1-\delta_L)v \geq \hat{\Pi_{L,0}} &\geq \delta_L^\tau (1-\delta_L) v
\end{aligned}$$

It could be that $\delta_L^{\tau-1}\delta_H(1-\delta_L)v \leq \hat{\Pi_{L,0}} \leq \delta_L^{\tau-1}(1-\delta_L)v$ instead in which case period $\tau - 1$ is in case 1, one has:

$$\begin{aligned}
\Pi_{A,0} &= \frac{1-\delta_H}{1-\delta_L} \sum_{s=0}^{\tau-2} \left(\frac{\delta_H}{\delta_L}\right)^s \Pi_{L,0} + \delta_H^\tau v \\
\Pi_{H,0} &= \left(\frac{\delta_H}{\delta_L}\right)^{\tau-1} \frac{1-\delta_H}{1-\delta_L} \Pi_{L,0} \\
\gamma &\geq \frac{1}{2}
\end{aligned}$$

Where $\Pi_{A,\tau} + \Pi_{H,\tau} = v$ and $\Pi_{A,\tau} = \delta_H v$ if $\gamma > \frac{1}{2}$ and $\Pi_{A,\tau} = (1-\delta_H)v$ if $\gamma < \frac{1}{2}$. This case holds if:

$$\begin{aligned}
(1-\delta_L)\delta_L^\tau v \leq \Pi_{L,0} &< (1-\delta_L)\delta_L^{\tau-1} v \\
\frac{1-\gamma}{\gamma} &\leq \sum_{s=0}^{\tau-1} \left(\frac{\delta_H}{\delta_L}\right)^s
\end{aligned}$$

Alternatively, at τ the solution might go in case 3. Assuming a case 3 that lasts τ' periods.

One has:

$$\Pi_{A,0} = \frac{1-\delta_H}{1-\delta_L}\Pi_{L,0}\sum_{s=0}^{\tau}\left(\frac{\delta_H}{\delta_L}\right)^s$$
$$\Pi_{H,0} = \frac{1-\delta_H}{1-\delta_L}\Pi_{L,0}\left(\frac{\delta_H}{\delta_L}\right)^{\tau}\sum_{s=1}^{\tau'-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^s + \delta_H^{\tau+\tau'}v$$

Such that:

$$(1-\delta_L)v(\delta_L+\delta_H)^{\tau'-1}\delta_L^{\tau} > \Pi_{L,0} \geq (1-\delta_L)v(\delta_L+\delta_H)^{\tau'-1}\delta_L^{\tau}$$
$$\frac{1-2\gamma}{1-\gamma} > \sum_{t=1}^{\tau'-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t \iff 1-\sum_{t=1}^{\tau'-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t > \frac{\gamma}{1-\gamma}$$
$$\frac{\delta_L}{\delta_L+\delta_H} - \sum_{t=2}^{\tau'}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t < \frac{\gamma}{1-\gamma} \iff 1-\sum_{t=1}^{\tau'}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t < \frac{\gamma}{1-\gamma}$$

Assuming case 3.

$$\lambda_{IC}(0) = \gamma$$
$$\lambda_A(0) + \lambda_{LM}(0) = 1-2\gamma$$
$$\Pi_{B,0} = \delta_H[\Pi_{A,1} + \Pi_{B,1}]$$
$$\Pi_{A,0} = \frac{1-\delta_H}{1-\delta_L}\Pi_{L,0}$$
$$\Pi_{L,1} = \frac{\Pi_{L,0} - \delta_H\frac{1-\delta_L}{1-\delta_H}\Pi_{A,1}}{\delta_L}$$

$C_4(1)$ and $C_5(1)$ becomes:

$$\lambda_1(1) + 1 - \gamma - \lambda_{LM}(0) = \lambda_{IC}(1) + \lambda_B(1) + \lambda_3(1)$$
$$\lambda_2(1) - \lambda_{LM}(1) + 1 - \gamma = \lambda_{IC}(1) + \lambda_3(1) + \lambda_A(1)$$

From C_6': $1-\gamma-\lambda_{LM}(0) \leq 1-\gamma-\lambda_{LM}(1)$ thus we are again in case 3 with:

$$\begin{aligned}
\lambda_{IC}(1) &= 1-\gamma-\lambda_{LM}(0) \\
\lambda_A(1)+\Lambda_{LM}(1) &= \lambda_{LM}(0) \\
\Pi_{H,1} &= \delta_H[\Pi_{A,2}+\Pi_{H,2}] \\
\Pi_{A,1} &= \frac{1-\delta_H}{1-\delta_L}\Pi_{L,1} \\
\Pi_{L,2} &= \frac{\Pi_{L,1}-\delta_H\frac{1-\delta_L}{1-\delta_H}\Pi_{A,2}}{\delta_L}
\end{aligned}$$

$C_4(2)$ and $C_5(2)$ becomes:

$$\lambda_1(2)+1-\gamma-\lambda_{LM}(1) = \lambda_{IC}(2)+\lambda_B(2)+\lambda_3(2)$$

$$\lambda_2(2)-\lambda_{LM}(2)+1-\gamma = \lambda_{IC}(2)+\lambda_3(2)+\lambda_A(2)$$

Which are the same as in $t=1$. Staying in case 3 for τ periods implies that:

$$\begin{aligned}
\Pi_{A,\tau-1} &= \frac{(1-\delta_H)\Pi_{L,0}}{(1-\delta_L)(\delta_L+\delta_H)^{\tau-1}} \\
\Pi_{H,\tau-1} &= \frac{(1-\delta_H)\Pi_{L,0}}{1-\delta_L}\sum_{t=1}^{\tau-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t + \delta_H^\tau(\Pi_{A,\tau}+\Pi_{B,\tau}) \\
\Pi_{L,\tau-1} &= \frac{\Pi_{L,0}}{(\delta_L+\delta_H)^{\tau-2}} \\
\lambda_{LM}(0) &= (1-\gamma)\sum_{t=1}^{\tau-1}\left(\frac{\delta_H}{\delta_H+\delta_L}\right)^t + \left(\frac{\delta_H}{\delta_L+\delta_H}\right)^{\tau-1}\lambda_{LM}(\tau-1)
\end{aligned}$$

From the assumption of case 3:

$$(1-\gamma)\sum_{t=1}^{\tau-1}\left(\frac{\delta_H}{\delta_H+\delta_L}\right)^t + \left(\frac{\delta_H}{\delta_L+\delta_H}\right)^{\tau-1}\lambda_{LM}(\tau-1) < 1-\gamma$$

$$\delta_H[\Pi_{A,t+1}+\Pi_{H,t+1}] > \frac{1-\delta_H}{1-\delta_L}\Pi_{L,t}$$

$$\Pi_{L,t} > (1-\delta_L)v$$

Case 3 ends when $\Pi_{L,\tau} \geq (1-\delta_L)v$, in which case one ends up in a $\mathcal{P}_1(\gamma)$ solution. In this case:

$$\begin{aligned}
\Pi_{A,0} &= \frac{(1-\delta_H)\Pi_{L,0}}{1-\delta_L} \\
\Pi_{H,0} &= \frac{(1-\delta_H)\Pi_{L,0}}{1-\delta_L}\sum_{t=1}^{\tau-1}\left(\frac{\delta_H}{\delta_L+\delta_H}\right)^t + \delta_H^\tau v \\
\Pi_{L,\tau} &= \frac{\Pi_{L,0}}{\delta_L(\delta_L+\delta_H)^{\tau-1}} - \frac{\delta_H(1-\delta_L)}{\delta_L}v \\
\lambda_{LM}(0) &= (1-\gamma)\sum_{t=1}^{\tau-1}\left(\frac{\delta_H}{\delta_H+\delta_L}\right)^t
\end{aligned}$$

Holds if:

$$\begin{aligned}
&(1-\delta_L)(\delta_L+\delta_H)^{\tau-1}v > \Pi_{L,0} \geq (1-\delta_L)(\delta_L+\delta_H)^\tau v \\
&\sum_{t=1}^{\tau-1}\left(\frac{\delta_H}{\delta_H+\delta_L}\right)^t < \frac{1-2\gamma}{1-\gamma}
\end{aligned}$$

Pareto frontier from separation

$$\begin{aligned}
\mathcal{P}_3(\gamma): \quad & \max_{\pi_B,\Pi_L,\Pi_A,\Pi_B} \gamma\left((1-\delta_H)\pi_A + \delta_H\Pi_A\right) + (1-\gamma)\left((1-\delta_H)\pi_B + \delta_H\Pi_B\right) \\
& \text{subject to}: \ \pi_B, \pi_A \geq 0\ , \quad (\Pi_A, \Pi_B) \in F\left(\frac{1-\delta_L}{\delta_L}\pi_A\right) \\
& \qquad (1-\delta_H)(\pi_B+\pi_A) \leq (1-\delta_H)v \qquad\qquad C_3 \\
& \qquad \rho_0\left((1-\delta_H)\pi_A + \delta_H\Pi_A\right) \geq (1-\delta_H)(\pi_B+\pi_A) \qquad S(A) \\
& \qquad \delta_H\Pi_B \geq (1-\delta_H)\pi_A \qquad\qquad S(B)
\end{aligned}$$

First if $\pi_A = 0$, by proposition 4.2 $\Pi_A = \Pi_B = 0$ and thus the objective equals 0. At the solution $\pi_A > 0$. The objective is increasing in π_B, hence at the solution either C_3 or $S(A)$ binds. If C_3 binds, then the first price is equal to v and there is no transition phase. This solution is feasible for ρ_0 sufficiently large:

$$\rho_0 \Pi_A \geq (1-\delta_H)v$$
$$v - \Pi_A \geq (1-\delta_H)v$$
$$\Longrightarrow$$
$$\rho_0 \geq \frac{1-\delta_H}{\delta_H}$$

Assume $\rho_0 < \frac{1-\delta_H}{\delta_H}$ so the first price is below v but further assume it is v after that period. If both binds: $(S(A) + \rho_0 S(B)$:

$$\pi_A + \pi_B = \rho_0 \frac{\delta_H}{1-\delta_H} v$$
$$\delta_H v \leq \pi_A \leq \min\left\{\rho_0 \frac{\delta_H}{1-\delta_H} v, \frac{\delta_H^2}{1-\delta_H} v\right\}$$

Thus this is unfeasible S_B can't bind with S_A. Thus it must be that S_B is slack and S_A binds:

$$\rho_0 \delta_H \Pi_A = (1-\delta_H)\pi_B + (1-\rho_0)(1-\delta_H)\pi_A$$
$$\frac{\rho_0}{1-\rho_0} \delta_H^2 v \geq \frac{\rho_0}{1-\rho_0} \delta_H \Pi_A \geq \pi_A \geq \delta_L v$$

It is feasible to have a price of v in the next period iff $\rho_0 \geq \frac{\delta_L}{\delta_H^2 + \delta_L}$

If $\rho_0 < \frac{\delta_L}{\delta_H^2 + \delta_L}$, (Π_A, Π_B) are picked in the frontier of $\mathcal{P}_2$. Moreover, this implies that $S(B)$ and C_3 are slack. Since the objective is increasing in π_B, $S(A)$ must bind, so $\mathcal{P}_3(\gamma)$ becomes:

$$\mathcal{P}_3(\gamma): \max_{\pi_A,\Pi_A,\Pi_B} \gamma\left((1-\delta_H)\pi_A + \delta_H\Pi_A\right) + (1-\gamma)\left(\rho_0\delta_H\Pi_A - (1-\rho_0)(1-\delta_H)\pi_A + \delta_H\Pi_B\right)$$
$$\text{subject to} :(\Pi_A,\Pi_B) \in F\left(\frac{1-\delta_L}{\delta_L}\pi_A\right)$$
$$\rho_0\delta_H\Pi_A - (1-\rho_0)(1-\delta_H)\pi_A \geq 0$$

From proposition 4.4, $(\Pi_A,\Pi_B) \in F\left(\frac{1-\delta_L}{\delta_L}\pi_A\right)$ are functions of π_A, τ_f, τ_s. For a given τ_f, τ_s the objective is linear in π_A. The derivative of the objective is:

$$(1-\delta_H)(\gamma-(1-\gamma)(1-\rho_0)) + \delta_H\frac{\partial\Pi_A}{\partial\pi_A}(\gamma+(1-\gamma)\rho_0) + \delta_H\frac{\partial\Pi_B}{\partial\pi_A}(1-\gamma)$$

If the objective is increasing in π_A then $(1-\delta_H)\pi_A = \frac{\rho_0\delta_H\Pi_A}{(1-\rho_0)}$. In this case one needs to find the feasible τ_s, τ_f that maximizes the objective. In the case of $\tau_s = 0$ and staying in case 2 one has:

$$(1-\delta_H)\pi_A = \frac{\rho_0\delta_H\Pi_A}{(1-\rho_0)}$$
$$\implies \pi_A = \frac{\rho_0\delta_H^{\tau_f+1}v}{1-\rho_0\sum_{s=0}^{\tau_f}\left(\frac{\delta_H}{\delta_L}\right)^s}$$
$$\delta_L^{\tau_f+1}v < \pi_A < \delta_L^{\tau_f}\delta_H v$$
$$\implies \delta_L^{\tau_f+1} < \frac{\rho_0\delta_H^{\tau_f+1}}{1-\rho_0\sum_{s=0}^{\tau_f}\left(\frac{\delta_H}{\delta_L}\right)^s} < \delta_L^{\tau_f}\delta_H$$
$$\iff \frac{1}{\sum_{s=0}^{\tau_f+1}\left(\frac{\delta_H}{\delta_L}\right)^s} < \rho_0 < \frac{1}{\sum_{s=0}^{\tau_f}\left(\frac{\delta_H}{\delta_L}\right)^s + \left(\frac{\delta_H}{\delta_L}\right)^{\tau_f}}$$

For a (τ_f, τ_s) transition phase to be feasible, ρ_0 must be such that:

$$\pi_A \leq \frac{\rho_0 \delta_H \pi_A}{(1-\rho_0)\delta_L} \sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s$$
$$\implies \rho_0 \geq \frac{1}{1 + \sum_{s=1}^{\tau_f+1} \left(\frac{\delta_H}{\delta_L}\right)^s}$$

A point on the frontier in between (τ_f, τ_s) and $(\tau_f + 1, \tau_s - 1)$. payoffs are:

$$\Pi_A = \frac{1-\delta_H}{\delta_L} \pi_A \sum_{s=0}^{\tau_f - 1} \left(\frac{\delta_H}{\delta_L}\right)^s + (1-\delta_H)\delta_H^{\tau_f} \pi_{A,\tau_f} \left(\frac{\delta_L + \delta_H}{\delta_L}\right)$$
$$\Pi_B = (1-\delta_H)\delta_H^{\tau_f} \left(\frac{\pi_A}{\delta_L^{\tau_f}} - \pi_{A,\tau_f}\right) + \frac{\delta_H^{\tau_f+1}}{\delta_L} \pi_{A,\tau_f} \sum_{s=1}^{\tau_s-2} \left(\frac{\delta_H}{\delta_L + \delta_H}\right)^s + \delta_H^{\tau_f+\tau_s} v$$

Using assumptions of that case:

$$(\delta_L + \delta_H)^{\tau_s - 2} \delta_L^{\tau_f} v > \pi_A - \delta_L^{\tau_f} \pi_{A,\tau_f} \geq (\delta_L + \delta_H)^{\tau_s - 1} \delta_L^{\tau_f} v$$
$$\pi_A \left[1 - \rho_0 \sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s\right] = \rho_0 \delta_H^{\tau_f+1} \frac{\delta_H + \delta_L}{\delta_L} \pi_{A,\tau_f}$$

Replacing π_{A,τ_f} one has:

$$\delta_H \Pi_A = \frac{(1-\delta_H)(1-\rho_0)}{\rho_0} \pi_A$$
$$\Pi_B = (1-\delta_H)\pi_A \left[\left(\frac{\delta_H}{\delta_L}\right)^{\tau_f} - \frac{\delta_L}{\delta_H} \frac{1 - \rho_0 \sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s}{\rho_0(\delta_H + \delta_L)}\right]$$
$$+ \frac{\pi_A}{\rho_0(\delta_L + \delta_H)} \left[1 - \rho_0 \sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s\right] \sum_{s=1}^{\tau_s - 2} \left(\frac{\delta_H}{\delta_L + \delta_H}\right)^s + \delta_H^{\tau_f + \tau_s} v$$
$$= (1-\delta_H)\pi_A \left[\left(\frac{\delta_H}{\delta_L}\right)^{\tau_f} - \frac{1 - \rho_0 \sum_{s=0}^{\tau_f} \left(\frac{\delta_H}{\delta_L}\right)^s}{\rho_0(\delta_H + \delta_L)} \left(\sum_{s=1}^{\tau_s - 2} \left(\frac{\delta_H}{\delta_L + \delta_H}\right)^s - \frac{\delta_L}{\delta_H}\right)\right] + \delta_H^{\tau_f + \tau_s} v$$

Chapter 2

Price Recommendation and the Value of Data: A Mechanism Design Approach.

Abstract. I study how a platform can affect its sellers' pricing decisions by using price recommendations. Sellers are too small to take account of network effects in their pricing decisions. To alleviate this problem, the platform influences seller prices by strategically disclosing demand information. I then study how the platform values additional information to improve price recommendations. I identify distortions in the way the platform uses and collects information and relate the nature and extent of these distortions with its business model. A platform that makes profits on both sides of the market uses information efficiently, but values information less to what is socially desirable. A platform that makes profits only on the seller side uses information to help sellers extract buyers surplus, which is inefficient. Further, such platforms value different types of information than the social planner.

Keywords: price recommendations, information design, two-sided markets.

JEL Codes: D82, D83, L21, L81.

1 Introduction

Data is a key input of the digital economy, often referred to as the "new oil."[1] The rise of large platforms and their rapidly expanding user bases has lead to an unprecedented accumulation and exploitation of personal data. This paper identifies potential distortions in the collection and use of personal data and argues that the extent to which a platform's incentives are (mis)aligned with social incentives crucially depends on its business model.

The competitive strength of platforms is increasingly determined by the amount of data available to them. E-commerce platforms assist sellers with descriptive statistics about their demand and help them set the best price or discount. Consumer relationship management (CRM) providers or online platforms such as Google or Facebook display personalized ads based on consumer characteristics. For e-commerce platforms in particular, empirical evidence shows that recommender systems have a significant impact on sellers prices and sales (Fleder and Hosanagar (2007), Pathak et al. (2010)). Whether this new form of influence is used to improve welfare is widely debated by policy makers and scholars,[2] and is a critical issue in a growing market for data worth 216 billion dollars in the US and 72.3 billion euros in the EU.[3] For instance, Executive Vice President of the European Commission Margrethe Vestager has stressed the need to "regulate the way that companies collect and use and share data – so that it benefits all of society."[4]

Against this background, my paper investigates whether platforms design price recommendations and collect the information used for this purpose in a socially desirable way. Although price recommendations do not constrain sellers in any way, they allow platforms

[1]See, for example, the speech of former Consumer Commissioner Kuneva at the Roundtable on Online Data Collection, Targeting and Profiling, 31 March 2009, SPEECH/09/156.

[2]See Crémer et al. (2019).

[3]See the European Data Market Monitoring Tool report page 9.

[4]Quoted from the commitments made at the hearing of Margrethe Vestager before the European parliament, in October 2019.

to strategically disclose their information about buyers thereby affecting sellers' beliefs and prices.[5] As the best transaction price for platforms may not be the same as the preferred transaction price for sellers, platforms benefit from price recommendations, which is consistent with the significant investments they make to launch a recommender system.[6] I study the link between the influence of the platform's recommendations on seller prices and its information about buyers to capture the platform's incentives to collect data about demand. In line with the recent calls for platform regulation,[7] I investigate how a platform's business model determines the nature of the distortions in the platform's use and collection of data.

I consider two business models: "paid" and "free" platforms. Paid platforms charge participation fees on both the buyer and the seller side. Free platforms provide free access to buyers and charge a participation fee on the seller side. Under both business models, the platform draws informative signals about buyers' valuations and correlates these signals with price recommendations to influence sellers' pricing decisions. Joining sellers receive a price recommendation and set a price for their good while joining buyers observe their value of the good and their matching seller's price and then purchase the good or not.

The two business models are commonly encountered among e-commerce platforms.[8] Moreover, they are motivated by a mechanism design approach. I show that they achieve the same set of outcomes as optimal mechanisms in a bilateral trade framework where neither the sellers' prices nor the buyers' decisions to purchase the good are contractible.[9] A paid

[5]Price recommendations are the direct representation of any committed communication strategy, see Myerson (1982). Other forms of information disclosure policy, such as displaying the predicted probability of selling, descriptive statistics about the market, and color schemes, can be viewed as indirect implementations of price recommendations.

[6]For instance, Ebay acquired Terapeak a data analytics firm in 2017 with estimated market value of $5 million. Mercari launched on kaggle in 2017 the "Mercari price suggestion challenge", rewarding the best price suggestions algorithm's programmers with $100,000 prize money.

[7]See Caffarra et al. (2020).

[8]Many e-commerce platforms, like Amazon Marketplace or eBay, provide free access to buyers. By doing so buyers can browse all sellers' products without providing payment-enabling information or creating accounts. However, platforms like Veepee and Showroomprive charge participation fees for premium members (V'pass and Infinity). Paid platforms may also charge a negative participation fee to buyers. For instance, some platforms offer discounts to new members or propose deals like "sponsor a friend" that can be interpreted as a negative participation fee.

[9]For instance, I show that the platform does not gain from screening the buyers' valuations to improve its recommendations.

platform implements transfers for buyers and sellers, whereas free platforms cannot implement transfers for buyers. This is the case for matching platforms that only interfere in the transactions by disclosing information.[10]

I determine the most profitable way for a platform to influence seller prices through price recommendations, and compare it to the socially optimal way of influencing seller prices. I then use the results to discuss the value of data for the platform and resulting incentives to collect data.

In the first part of the paper, I show that paid platforms use data efficiently, whereas free platforms use data inefficiently. As often in platform models, paid platforms extract the entire surplus of the interaction via entry fees.[11] This implies that paid platforms use data to maximize surplus per trade. On the contrary, free platforms only make profits on the seller side. They use data to help sellers to extract buyer surplus, which may destroy surplus relative to a no-data benchmark. These two different uses of information by paid platforms and free platforms imply different incentives to collect information.

As an illustration, consider an e-commerce platform that offers sellers the possibility to use a smart discount device. This device automatically shows a discounted price to some specific buyers browsing the seller's product, and shows the full price to other buyers.[12] Consider the following two data collection strategies. Strategy (a) collects the type of data that suggests that buyers have a high valuation for the good. Think of a browsing history that shows the buyer has purchased this good at a high price on a regular basis, or social media data that shows the buyer liked this good. Strategy (b) collects the type of data that suggests buyers have a low valuation for the good. Think of the geographic location, whether the buyer is a student (which would indicate that the buyer is financially constrained), or simply a browsing history that shows the buyer only purchases this good when discounted. Neither strategy is perfectly accurate: many buyers remain undetected and none of these

[10] This abstracts from the possibility to charge fees in the transaction.

[11] See e.g. Rochet and Tirole (2006).

[12] Cdiscount marketplace proposes Smart Discount Voucher to its sellers which is precisely what is presented in the example. Many other e-commerce platforms offer similar devices to their sellers.

types of data reveals exactly buyers' valuations. Strategy (a) is best at generating surplus from trade: the platform can show the full price to detected buyers and the discounted price to all others. Doing so minimizes the probability that buyers with low valuations face no discounted price, which would lead to no trade. By contrast, strategy (b) is best at extracting buyer surplus: the platform shows the discount to detected buyers and the full price to all others. Doing so maximizes the probability that buyers with high valuations see the full price and, therefore, increases the probability of correctly marking-up buyers. Paid platforms prefer investing in strategy (a), which is efficient. However, the level of that investment is inefficiently low. By contrast, free platforms' demand for data is biased towards strategy (b). Therefore, they collect/purchase inefficient types of data.

To capture the platform's incentives to collect data, I compute the platform's marginal value of information; in other words, by how much the platform's profit changes when their information structures marginally change. In my model, information is an input of price recommendations. Using sensitivity analysis, I characterize paid and free platforms' willingness to pay for information by computing the shadow price of this input. The marginal value of information provides rich comparative statics on information structures which copes with the high dimensionality of this input.

Then, I compare the platform's willingness to pay to collect information with the social planner's and identify several distortions. I show that paid platforms, despite using information efficiently, under-value any additional information. Paid platforms set inefficiently high entry fees, which implies that learning impacts less trades than under the social planner's trade mechanism. Therefore, the paid platforms' willingness to collect information is proportionally lower than the social planner's by a factor that only depends on the elasticity of demand, which can be estimated empirically. By contrast, free platforms have a biased demand for information. Free platforms value more learning about mark up opportunities than learning about trade opportunities, although the latter is the efficient way to learn. The bias holds regardless of the source of information and, therefore, free platforms' incentives to collect data are inefficiently oriented.

Platforms either collect data through their websites (e.g. by requiring users to identify with personal accounts or by tracking purchase histories or geographical locations) or purchase data from data brokers. My model exhibits the distorted incentives of platforms to collect information. The different nature of the distortions between paid and free platforms argues in favor of a regulation based on the business model of platforms. Moreover, I develop a rigorous framework, which may prove useful beyond the issues studied in this paper. This framework is based on: (i) a mechanism design approach to capture how information affects market outcomes, and (ii) a duality approach that treats information as an input and computes its marginal value.

Related Literature

This paper relates to the recent strand of the literature on platforms that analyzes how intermediaries use their information to increase profits (see e.g. Hagiu and Hałaburda (2014), Gomes and Pavan (2019), Bourreau and Gaudin (2018), Jullien and Pavan (2019), and Carroni, Pignataro, and Tampieri (2020)).[13] The closest papers to mine are Pavlov and Berman (2019) and Bonatti and Cisternas (2020). In both papers the platform's information is used to affect seller prices. Pavlov and Berman (2019) compare multiple pricing regimes by an e-commerce platform, including a price recommendation regime, in a cheap-talk environment. Instead, I assume that the platform commits to the trade mechanism. Bonatti and Cisternas (2020) study how platforms can affect seller prices and sales via consumer scores. These authors show that a platform can smooth interactions in the market by strategically designing scores transparency and updating rate. In contrast, I explore the issue of data collection and study the platform's incentives to collect data for general information structures.

From a methodological perspective, my paper relates to the Bayesian persuasion, infor-

[13] For the literature on platforms see Caillaud and Jullien (2003), Armstrong (2006), Rochet and Tirole (2003) and Rochet and Tirole (2006))

mation design literature (see Kamenica and Gentzkow (2011), Bergemann and Morris (2016) and Taneva (2019)). In my paper the platform is a mechanism designer that commits to an information disclosure policy to sellers. Several other papers use the Bayesian persuasion framework to study data brokers, e.g. Calzolari and Pavan (2006), Bergemann and Bonatti (2015), Bergemann, Bonatti, and Smolin (2018) and Yang (2020). Contrary to a data broker, in my model, the platform contracts with both sides, and uses information to balance the allocation of transaction surplus across the two sides.

My paper also addresses the broader question of the efficient acquisition and allocation of information in markets. Many recent papers have studied this issue, including Colombo, Femminis, and Pavan (2014), Pavan, Vives, et al. (2015), Montes, Sand-Zantman, and Valletti (2019), Bergemann, Bonatti, and Gan (2020), Acemoglu et al. (2019), Hagiu and Wright (2020), De Corniere and Taylor (2020), and Bergemann, Heumann, and Morris (2020). Most of them warn against potential distortions and compare welfare under different informational regimes. My paper contributes to our understanding of this issue by providing a comparative static analysis based on the marginal value of information, which not only identifies the socially desirable information structures, but also the incentives to collect information.

The duality analysis I provide relates to a recent strand of the information design literature, see e.g. Kolotilin (2018), Dworczak and Kolotilin (2019). The dual problem is set up in an alternative way to perform a sensitivity analysis on the information structure. I also provide a different proof of strong duality, than the one provided in Dworczak and Martini (2019), Galperti and Perego (2018), and Dizdar and Kováč (2020). In addition, I show that, in the context of my model, the dual variable associated to the informational constraint can be interpreted as the marginal value of information for the platform.

The paper is organized as follows. Section 2 presents the model. Section 3 characterizes the most profitable price recommendations for the platform and discusses the static welfare distortions under both business models. Section 4 characterizes the platform's marginal value of information. Section 5 discusses the resulting distortions in the platforms' incentives to collect data. Section 6 concludes.

2 Model

Environment. Consider a platform that intermediates trade between buyers and sellers. The platform sets participation fees on both sides and recommends a price to each seller. Sellers produce a good at marginal cost c which is valued at either v_l or v_h by their matching buyer, with $v_l < v_h$. I normalize outside option to 0 and I assume that $c < v_l$ so that trade is always efficient. For the buyers and sellers that decide to join the platform, transactions unfold as follows: buyers inspect sellers' goods to find their matching seller[14] as well as their valuation for its good (v_l or v_h). Following a successful match, sellers receive a price recommendation and then set a price. Buyers observe their matching seller's price and then decide to purchase the good or not.

Information Structure. The buyers' types (v, b) have two independent dimensions. The value of the matching seller's good is either high (v_h), with probability ρ_0, or low (v_l), with probability $1 - \rho_0$. The stand-alone valuation b is derived by the buyer if he visits the platform. It compounds many interpretations: b includes benefits derived from additional services provided on the platform but also costs due to search or lack of privacy and advertising nuisances. The stand-alone valuation b can be positive or negative and is distributed according to a continuously differentiable and log-concave distribution Q supported on an interval in $\mathbb{R}$. For instance, a buyer that joins the platform and purchases his matching seller's good gets a payoff of $b + v - p - t_b$, where p is the seller's price and t_b is the buyer entry fee. Buyers learn v only after joining the platform and inspecting their matching seller's good, whereas their stand-alone valuation b is known before.

[14] Matching is one to many: For each buyer there is only one valuable seller.

For each joining buyer, the platform receives informative signals about his valuation, from which it forms posterior beliefs about whether the buyer is of type v_h or v_l. Think of the platform with access to consumer-level data such as histories of past transactions, location or cookies and the related means of recording browsing data. This data may be collected by the platform or obtained from a data broker (Nielsen, Acxiom or Epsilon for instance). Once a buyer joins the platform, the platform may observe some of its characteristics and updates its beliefs about whether the buyer is a low or high type.

Without loss of generality, the platform's information structure is represented as the distribution of its posterior beliefs. For each buyer, the platform independently draws signals $\rho_s = P(v = v_h|\rho_s)$ that are normalized to the platform's posterior belief that the buyer is a high type. These posterior beliefs ρ_s are drawn according to a distribution F with mean ρ_0.[15]

Timing and Decisions. First, the platform sets entry fees t_s and t_b and commits to a price recommendation rule, which I describe later. Second, buyers observe their stand-alone valuation b. Then, buyers and sellers decide whether to join the platform or not. Third, the platform observes each match, each signal ρ_s and recommends a price to each matched seller. Fourth, buyers observe v and their seller's price and then decide whether or not to buy.

Solution Concept. The analysis focuses on perfect Bayesian equilibria, where players hold rational expectations, are risk neutral and expected-payoff maximizers.

Price Recommendation Rules and Entry Fees. I analyze two platform business models: paid platforms and free platforms. Paid platforms set entry fees t_b on the buyer side and t_s on the seller side, $t_b, t_s \in \mathbb{R}$. Free platforms set entry fees only on the seller side and provide free access to buyers ($t_b = 0$).[16] The seller entry fee is paid per buyer joining,

[15] F has mean ρ_0 to be consistent with the prior distribution of buyers' types, see for instance Kamenica and Gentzkow (2011) proposition 1.

[16] Free platforms capture the case of CRM provides and many e-commerce platforms that do not charge entry fees on the buyer side.

in other words, sellers pay t_s times the mass of buyers joining.[17] Both types of platforms commit to a price recommendation rule that maps signals ρ_s with a private price recommendation to sellers. Instead of capturing the price recommendation rule as the probability of recommending a price conditional on signals, it is more convenient to define it with $\mu(S)$ the joint probability of recommending a low price v_l and receiving signals in S:[18]

$$\begin{aligned} \mu \ : \ & \mathcal{B}[0,1] \to [0,1] \\ & S \mapsto \mu(S) = \int_S P(\text{recom } v_l \mid \rho_s) dF(\rho_s). \end{aligned}$$

With complementary probability the platform recommends v_h to sellers. Since buyers are either of type v_l or of type v_h, there are only two potentially optimal pricing strategies. Sellers either sell to both buyer types at price v_l, leaving the low type without any surplus, or sell only to the high types at price v_h leaving them with no surplus. Therefore, considering price recommendation rules that only recommend to sellers these two prices v_l or v_h is without loss of generality.

A price recommendation rule is feasible for an information structure F if it satisfies the informational constraint (In):

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) = F(S). \tag{In}$$

(In) captures the ability of the platform to match price recommendations to the true buyers' valuations. If $\mu(S) = F(S)$ the platform always recommends a low price when the signal falls in the set S. However, if $F(S) = 0$ for some S the platform cannot make recommendations as it does not receive signals in this range. Therefore, the informational constraint captures the relationship between the quality of the platform information structure F and the precision of the price recommendations.

[17] This entry fee corresponds to the *per-interaction price* introduced in Rochet and Tirole (2006) which links the modelling approach of Armstrong (2006) and Rochet and Tirole (2003).

[18] S is a measurable set in $\mathcal{B}[0,1]$, the Borel $\sigma-$algebra on $[0,1]$.

This completes the description of the model. The next two paragraphs discuss two properties that motivate the modelling choice of the trade mechanism and the price recommendation rule.

Alternative Trade Mechanisms. Appendix D shows that by setting entry fees and designing a price recommendation rule, a platform can achieve all possible outcomes among all mechanisms in the context of bilateral trade under private transaction terms. I assume that the platform can be viewed as a mechanism designer organizing multiple independent bilateral trade interactions between a seller and a buyer except that the transaction price and the allocation of the good are not contractible decisions.[19] This assumption captures the idea that platforms facilitate transactions without taking part in the transactions. The key result of this section is to show that the platform cannot improve its recommendations to sellers by screening buyers' types. The information gain on the buyer side cannot be passed on the seller side, and, therefore the quality of the platform's recommendations solely relies on its ex-ante private information. As the platform does not gain from screening buyers, posting entry fees on both sides can achieve the largest profit. This result also holds when considering the class of mechanisms that are free for buyers. Therefore, both mechanisms studied are without loss of generality as they achieve the same set of outcomes as the direct mechanism in the context of bilateral trade with private transaction terms.

Interpretation of the Price Recommendation Rule. Many e-commerce platforms such as Amazon Marketplace, E-bay or Mercari use data about demand to suggest prices or discounts to their sellers, akin to the price recommendation rule. Additionally, a free platform in my model can be interpreted as a CRM provider offering devices to sellers that personalize discounts, or target specific consumers. For instance, users of Facebook Ads, Google Ads, or Voucherify can launch personalized coupon campaigns associated to specific queries, search

[19] If agents actions are non-contractible or private, meaning that these actions cannot be an output of the mechanism, the principal may still send messages to agents to influence their actions. See Myerson (1982).

histories, cookies related to buyer characteristics. In the context of this model, think of a cookie as providing some information, say ρ_s, about the user of the e-commerce platform. These companies have access (through their user base directly, or via third parties) to partial information about consumers whereby ρ_s is distributed according to some cdf F. Given their information, these platforms offer sellers personalized pricing/discount strategies tying prices (here v_l or v_h) to queries ρ_s. Sellers may or may not follow the offered pricing or discounting strategies, in which case absent of any information a uniform price is set to all consumers. In the context of the price recommendation rule, if a seller disobey, say price at v_l when recommended v_h, then she is effectively pricing uniformly at v_h for all consumers.[20] In this framework, free platforms' price recommendation rule can be interpreted as the discounting strategies offered by online platforms or CRM providers.

3 Optimal price recommendation rules

This section first analyzes the optimal price recommendation rule for paid platforms and then presents the case of free platforms.

The paid platform case

Buyers value trade only if the platform recommends a low price and if they are a high type. Facing an entry fee of t_b there is a marginal buyer $\tilde{b}$ such that all buyers with stand-alone valuation higher than $\tilde{b}$ join the platform and all buyers with stand-alone valuation lower than $\tilde{b}$ do not. The mass of buyers joining the platform is given by $1 - Q(\tilde{b})$ and $\tilde{b}$ is characterized by:

$$\tilde{b} = t_b - \int_0^1 \rho_s \mu(d\rho_s)(v_h - v_l). \tag{2.1}$$

[20] See the description of the sellers' incentive compatibility for more details about this claim.

The price recommendation rule μ is the joint probability of recommending a low price v_l and receiving a signal in a measurable set S. As the signal $\rho_s = P(v = v_h \mid \rho_s)$ is normalized to the posterior probability of a high type buyer, the probability $\int_0^1 \rho_s d\mu(\rho_s) = P(\text{recom } v_l,\ v = v_h)$ is the joint probability of recommending a low price to a high type buyer.

If a seller joins the platform, she trades at price v_l when she receives a low price recommendation, and at price v_h when she receives a high price recommendation and the buyer has a high valuation. Sellers want to match their price with the buyer's type. If she is recommended a low price but the buyer is type v_h she loses the mark-up $v_h - v_l$. If she is recommended a high price but the buyer is type v_l, then she misses a trade opportunity of value $v_l - c$. Upon entering, a seller pays a total entry fee of $t_s(1 - Q(\tilde{b}))$. To maximize the platform's profit, t_s is set to extract the entire seller's profit:

$$t_s = \int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c).$$

In each trade, sellers trade at a low price, which yields profit of $(v_l - c)$, when a low price is recommended, with probability $\int_0^1 \mu(d\rho_s)$. Sellers trade at a high price, which yields $(v_h - c)$, when a high price is recommended and the buyer is a high type, that is with probability:

$$\rho_0 - \int_0^1 \rho_s \mu(d\rho_s) = P(v = v_h) - P(\text{recom } v_l,\ v = v_h) = P(\text{recom } v_h,\ v = v_h).$$

The platform benefits from the transaction only through the entry fees:

$$\Pi = (1 - Q(\tilde{b}))(t_s + t_b).$$

Using equation (1) and replacing the seller entry fee by their profit, the platform's objective becomes:

$$\Pi = (1 - Q(\tilde{b}))\left[\int_0^1 (1 - \rho_s)\mu(d\rho_s)(v_l - c) + \rho_0(v_h - c) + \tilde{b}\right].$$

Via the entry fees, the platform captures the surplus from trade. With probability ρ_0 the buyer is of type v_h, which yields a surplus of $\rho_0(v_h - c)$ regardless of the price. If the buyer is type v_l, transaction happens only if the platform correctly recommends a low price, generating an expected surplus of $\int_0^1 (1-\rho_s)\mu(d\rho_s)(v_l - c)$. When the platform designs the price recommendation rule, the platforms avoiding recommending a high price to sellers matched with low type buyers as much as possible.

Formulation of the problem. The platform chooses the profit maximizing price recommendation rule and the associated entry fees, subject to incentive compatibility and feasibility conditions.

Incentive compatibility entails that sellers must find it optimal to follow the platform's price recommendations:

$$\int_0^1 \mu(d\rho_s)(v_l - c) \geq \int_0^1 \rho_s \mu(d\rho_s)(v_h - c). \tag{IC_l}$$

$$\left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c) \geq \left[1 - \int_0^1 \mu(d\rho_s)\right](v_l - c). \tag{IC_h}$$

Sellers have four options: follow both price recommendations, disobey both price recommendations, or disobey one recommendation and follow the other. If sellers disobey one recommendation and follow the other, they effectively follow a uniform pricing strategy. If disobeying both recommendations is optimal, then one of the uniform pricing strategies is better than following both recommendations.[21] Ensuring that obeying both recommendations yields more profit than both uniform pricing strategies is equivalent to incentive compatibility:

$$\int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c) \geq \max\{v_l - c, \rho_0(v_h - c)\}. \tag{IC}$$

[21] See Kolotilin et al. (2017) for this result.

Sellers either follow the platform recommended strategy, or reject it and set a uniform price. This formalism of the incentive compatibility constraints corresponds to many CRM providers and platforms practices. Sellers, via CRM providers or online platforms, can launch ads which display discounts that are personalized to consumer characteristics. On Cdiscount, for instance, sellers can use the "smart discount voucher" device that automatically displays a discounted price to a selected group of buyers (based on the platform's available data) and display the full price to others.

Given the prior distribution of high type buyers in the population, one of the two uniform pricing strategies is better than the other. If $\rho_0 > \frac{v_l - c}{v_h - c}$ then setting a uniform price v_h is strictly more profitable than v_l for sellers. In this case sellers are said to be *optimistic*. If $\rho_0 \leq \frac{v_l - c}{v_h - c}$, then pricing uniformly at v_l is more profitable, and sellers are said to be *pessimistic*.

The platform pins down the marginal buyer joining $\tilde{b}$ via the choice of the buyer entry fee. Together with the choice of the price recommendation rule μ, the platform maximizes its profit, under the (IC) and (In) conditions:

$$\max_{\tilde{b},\ \mu}(1 - Q(\tilde{b}))\left[\int_0^1 (1-z)\mu(dz)(v_l - c) + \rho_0(v_h - c) + \tilde{b}\right]$$

subject to:

$$\text{Incentive Compatibility} \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

The optimal price recommendation rule can be deduced from interpreting the recommendations of the platform as the result of a hypothesis testing. Let "recommending v_l" correspond to accepting H_0 and let "recommending v_h" correspond to rejecting H_0. The platform then chooses as a function of its test statistic ρ_s whether to reject H_0. In this formulation the incentives are captured by type I and II errors. Sellers want to minimize both types of errors: they want to match their price with their buyer's type. The platform wants to avoid type I errors, that is, avoid recommending a high price if the buyer's valuation is low as it reduces trade efficiency.

The platform and sellers agree that the type I error must be as small as possible. Therefore, the optimal price recommendation rule must have the following property: for any level of the type II error chosen, the level of the type I error must be minimized.

This property is common in statistics and econometrics and shared by all standard hypothesis test. In this case, the tests that satisfy this property follows a cutoff rule. The platform chooses a cutoff $\rho_t \in [0, 1]$ such that for all signals ρ_s above the cutoff the platform recommends a high price and for all the signals below the cutoff the platform recommends a low price.

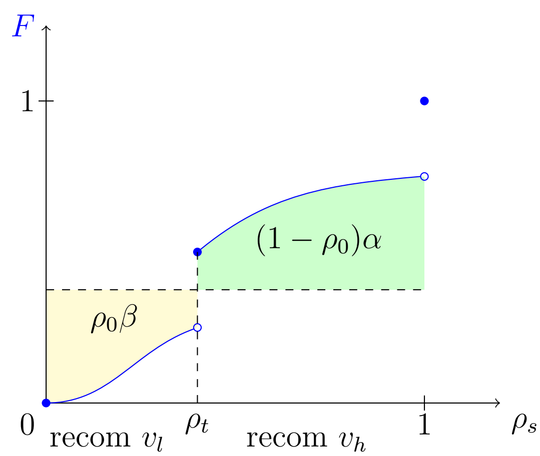


Figure 2.1: Cutoff rule.

Lemma 3.1. *The optimal price recommendation rule takes the form of a* ***cutoff rule****. The* *platform chooses a cutoff* $\rho_t \in [0,1]$ *and then recommends prices as follows:*

(i) *For all* $\rho_s < \rho_t$, *the platform recommends a low price.*

The platform sets for all measurable $S \subset [0, \rho_t)$, $\quad \mu(S) = F(S)$.

(ii) *For all* $\rho_s > \rho_t$, *the platform recommends a high price.*

The platform sets for all measurable $S \subset (\rho_t, 1]$, $\quad \mu(S) = 0$.

(iii) *If* F *has a mass point at* ρ_t, *the platform randomizes its recommendations at the cutoff.*

The platform picks $\mu(\{\rho_t\}) \in [0, dF(\rho_t)]$, *with* $dF(\rho_t) = F(\rho_t) - \lim_{x \uparrow \rho_t} F(x)$.

Proof. This is a standard result in the Bayesian Persuasion literature see e.g. Kamenica and Gentzkow (2011) and Dworczak (2020). A new proof (using duality techniques) is given in the appendix A.2 and B.2. □

Figure 1 pictures a cutoff rule by the platform in the case where F has a mass point at the cutoff. On the figure the platform mixes equally the recommendation at the cutoff ρ_t. The shaded areas represent the associated type I and II errors. In the text, equations are presented assuming no mass point at the cutoff, while the appendix presents the general case.

Optimal cutoffs. The optimal cutoff is set to maximize surplus from trade. The ideal cutoff is $\rho_t = 1$ as in this case the platform always recommends a low price hence trade is efficient. However, this price recommendation rule provides no information to sellers about their matched buyer's type.

When sellers are *pessimistic* the ideal cutoff $\rho_t = 1$ is incentive compatible: With their prior belief, pessimistic sellers' optimal price is v_l. In this case, the platform always recommends a low price which provides no additional information to sellers, and trade is efficient.

However, when sellers are *optimistic* the ideal cutoff $\rho_t = 1$ is not incentive compatible. The price recommendation rule must feature a cutoff low enough so that low price recommendations are sufficiently informative about the buyer's type and followed by sellers. Coincidentally, to limit trade inefficiencies the platform aims to increase the cutoff. The optimal price recommendation rule is therefore the highest cutoff up to which the sellers' incentive compatibility constraint is binding.

Formally, the optimal cutoff ρ_t^* is set to generate a posterior belief of $\frac{v_l-c}{v_h-c}$ for sellers such that at the low price recommendation they are indifferent between setting a low price or a high price:

$$\frac{\int_0^{\rho_t^*} \rho_s dF(\rho_s)}{F(\rho_t^*)} = \frac{v_l - c}{v_h - c}.$$

Given the optimal cutoff ρ_t^* the platform optimally sets t_b to balance out the revenue extracted on infra-marginal buyers versus the revenue made on both sides by attracting more buyers. Formally, for pessimistic buyers, the optimal entry fee t_b^* induces an optimal $\tilde{b}^*$ that satisfies:

$$1 - Q(\tilde{b}^*) = Q'(\tilde{b}^*)\left[(1-\rho_0)(v_l - c) + \rho_0(v_h - c) + \tilde{b}^*\right]. \tag{2.2}$$

For optimistic buyers, the optimal entry fee t_b^* induces an optimal $\tilde{b}^*$ that solves:[22]

$$1 - Q(\tilde{b}^*) = Q'(\tilde{b}^*)\left[\int_0^{\rho_t^*}(1-\rho_s)dF(\rho_s)(v_l - c) + \rho_0(v_h - c) + \tilde{b}^*\right]. \tag{2.3}$$

As in many platform models, t_b^* may be negative, since the platform considers the revenue an extra buyer generates on the seller side. Platforms may have the possibility to price negatively via vouchers for new accounts or sponsor a friend deals.[23]

The next proposition summarizes the results on the optimal paid platforms entry fees

[22] In case of a mass point at the threshold see appendices.

[23] Deliveroo, Ubereats, Cdiscount and many other platforms offer these types of deals.

and price recommendation rules:

Proposition 3.2. *Optimal paid platform mechanism.*

1. *Consider the case of pessimistic sellers ($\rho_0 \leq \frac{v_l - c}{v_h - c}$). The platform recommends a low price for all signals, and sets entry fees $t_s = v_l - c$ and t_b according to the first order condition (2).*

2. *Consider the case of optimistic sellers ($\rho_0 > \frac{v_l - c}{v_h - c}$). The platform recommends prices according to a cutoff rule, where ρ_t^* binds (IC). It sets the entry fee for sellers t_s to capture their entire profit and the entry fee for buyers t_b according to the first order condition (3).*

Proof. See appendices. □

The paid platform price recommendation rule is socially optimal: A social planner would also pick the same cutoff rule as it minimizes trade inefficiencies.

In this paper, the social planner is a decision maker motivated by total surplus that can use the same instruments as paid platforms (respectively as free platforms in the next subsection). The social planner designs the trade mechanism to maximize total surplus:

$$\max_{\tilde{b},\ \mu} \int_{\tilde{b}} \left(\int_0^1 (1-z)\mu(dz)(v_l - c) + \rho_0(v_h - c) + b \right) dQ(b)$$

subject to:

$$\text{Incentive Compatibility} \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

Both paid platforms and the social planner value the total surplus from trade. Therefore, the planner uses the same price recommendation rule as paid platforms.

However, the social planner also values the total buyer surplus that comes from the stand-alone valuation, i.e. $\int_{\tilde{b}} b dQ(b)$, while the platform captures this surplus only at the marginal stand-alone valuation, i.e. $(1 - Q(\tilde{b}))\tilde{b}$. Consequently, the social planner sets a lower buyer entry fee than the platform. The social planner reduces the buyer entry fee to the point where the cost of attracting an additional buyer compensates the surplus it generates on both sides:

$$-\tilde{b}^{sp} = \int_0^{\rho_t^*} (1 - \rho_s) dF(\rho_s)(v_l - c) + \rho_0(v_h - c). \tag{3}'$$

Therefore, under the social planner's trade mechanism, a larger mass of buyers joins generating more trades and so the total surplus is larger.

Proposition 3.3. *Paid platforms' price recommendation rule is socially optimal as it maximizes trade efficiency. However, it sets a buyer entry fee higher than what is socially optimal.*

Proof. See appendices. □

Although paid platforms disclose information efficiently, it does not imply that paid platforms collect information efficiently. Section 4 demonstrates that paid platforms' marginal value of information is lower than that of the social planner, leading to sub-efficient incentives to collect data.

Distortions in the incentives to collect data are worse if in addition the platform does not disclose information efficiently. The case of free platforms introduces a new friction which distorts the way free platforms use information.

The free platform case

The analysis follows the same steps as the case of paid platforms. As buyers face no entry fees, the marginal buyer is:

$$\tilde{b}_f = -\int_0^1 \rho_s \mu(d\rho_s)(v_h - v_l). \tag{2.4}$$

To increase the mass of buyers joining $Q(\tilde{b}_f)$, free platforms can only increase the probability of recommending a low price when buyers have a high valuation.

As in the previous case, free platforms set the seller entry fee to capture their profits:

$$t_{s,f} = \int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c).$$

The incentive compatibility condition and the informational constraint remain unchanged. Free platforms choose the price recommendation rule μ to maximize profits:

$$\max_{\mu}(1 - Q(\tilde{b}_f))\left[\int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c)\right]$$

subject to:

Incentive Compatibility (IC)

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

When setting the price recommendation rule free platforms trade off sellers' profits, which directly increases their revenue through the entry fee, with buyer surplus, which affects the mass of buyers joining and the total mass of trades.

Similar to the case of paid platforms, it is useful to interpret the design of the price recommendation rule as the design of a hypothesis test. The type I error corresponds to recommending a high price v_h if the buyer's valuation is low v_l, and the type II error corresponds to recommending a low price v_l if the buyer's valuation is high v_h. Sellers want to avoid both types of errors and match their price with the buyer's type. Free platforms want to avoid type I errors as it leads to no trade. However, free platforms face a tradeoff regarding the optimal level of type II errors. More type II errors reduces the sellers' profit and the platform's revenue, but it also increases buyer surplus from trade, and therefore attracts more buyers to the platform.

Overall, free platforms and sellers agree that the level of the type I error must be as low as possible. The optimal price recommendation rule follows the property shared by many hypothesis tests: for any level of the type II error chosen, the level of the type I error must be minimized.

The free platforms optimal price recommendation rule follows a cutoff rule. Free platforms choose a cutoff $\rho_{t,f} \in [0, 1]$ such that for all signals ρ_s below the cutoff, free platforms recommend a low price and for all signals above the cutoff recommend a high price.[24]

Optimal cutoff. For now I ignore the sellers' incentive compatibility condition. The ideal, unconstrained, cutoff for the platform trades off the revenue made on the seller side by attracting $(v_h - v_l)\rho_t Q'(\tilde{b}_f)$ more consumers (increasing the cutoff by 1 unit), and the loss of revenue on the infra marginal trades from pricing at v_l instead of v_h at the cutoff. Formally, the ideal cutoff solves:

$$\begin{aligned} &\rho_{t,f}(v_h - v_l)Q'(\tilde{b}_f)\left(F(\rho_{t,f})(v_l - c) + \int_{\rho_{t,f}}^{1} \rho_s dF(\rho_s)(v_h - c)\right) \\ =&(1 - Q(\tilde{b}_f))\left(\rho_{t,f}(v_h - c) - (v_l - c)\right). \end{aligned} \tag{5}$$

The incentive compatibility condition of pessimistic sellers holds at the ideal cutoff. To be convinced to set a high price, the pessimistic seller's posterior belief must be higher than $\frac{v_l - c}{v_h - c}$. However, the ideal cutoff cannot be lower than $\frac{v_l - c}{v_h - c}$, even in the case of inelastic Q. For pessimistic sellers, the ideal cutoff pinned down by (5) is optimal.

For optimistic sellers there are two cases. Either the ideal cutoff is low enough so the sellers' (IC) holds, in which case the ideal cutoff is optimal. Or the ideal cutoff violates the sellers' (IC), in which case the platform chooses the highest cutoff until the sellers' incentive compatibility condition is binding.

The next proposition summarizes the results on the optimal free platform entry fee and price recommendation rule:

[24]This result is standard in the Bayesian Persuasion literature. For completeness, a formal new proof (using duality) of this claim is presented in appendix B.2.

Proposition 3.4. *Optimal free platform mechanism.*

Free platforms set the seller entry fee to capture their entire profits. Additionally, the optimal price recommendation rule is as follows:

1. *Consider the case of pessimistic sellers ($\rho_0 \leq \frac{v_l - c}{v_h - c}$). The platform recommends prices according to a cutoff rule. The optimal cutoff is the solution to equation (5).*

2. *Consider the case of optimistic sellers ($\rho_0 \leq \frac{v_l - c}{v_h - c}$). The platform recommends prices according to a cutoff rule. There are two cases: either (a) ρ_t^* binds the sellers' (IC), or (IC) is slack and the optimal cutoff is the solution to equation (5).*

Proof. Appendix B.2 demonstrates the optimality of cutoff rules. Appendix B.3 completes the characterization of the optimal price recommendation rule. □

A social planner, constrained to set $t_b = 0$, would set the same cutoff as paid platforms. That is, set the highest cutoff feasible, and permitted by the sellers' (IC). For the social planner there is no trade-off: increasing the cutoff increases both trade efficiency and the total mass of trades.

Free platforms' price recommendation rule is efficient only if buyers are optimistic and the optimal cutoff binds the sellers' (IC), in which case it coincides with the socially optimal cutoff. However, if the optimal cutoff is interior, free platforms recommend high prices too often leading to a lower surplus generated in each trade and a lower total mass of buyers joining. In the case of pessimistic sellers, the social planner always recommends a low price, that is, it uses *no data* and achieves trade efficiency. However, in this case, the free platforms use data to design a price recommendation rule with an interior cutoff. When sellers are pessimistic, free platforms destroy surplus relative to no-data.

Proposition 3.5. 1. *Consider the case of pessimistic sellers ($\rho_0 \leq \frac{c_l - c}{v_h - c}$). The free platform price recommendation rule is inefficient, and even reduces total surplus relative to a no data benchmark.*

2. *Consider the case of optimistic sellers ($\rho_0 > \frac{c_l - c}{v_h - c}$). The free platform price recommendation rule is inefficient if the cutoff is interior and is efficient if the cutoff binds the sellers' (IC).*

Proof. See appendix B.4 □

Paid platforms, free platforms and social planners all use cutoff rules. From Blackwell (1953), these cutoff rules could not be replicated with less informative signal structures. Despite the distorted free platforms' price recommendation rule that led to inefficient outcomes, it nonetheless wastes no information in the sense of Blackwell. Section 4 however shows that the platform does not value information in the same way the social planner does. The next section presents the marginal value of information for the platform, and then compares it with the social planner's marginal value of information.

4 The Marginal Value of Information

The value of information for a decision maker is the best payoff achievable given his information structure.

Definition 3. *The platform's value of information maps each information structure into the optimal profit that can be achieved under it. $V_p(F)$ (resp. $V_f(F)$) denotes a paid platform's (resp. a free platform's) value of information.*

This section focuses on the platform's marginal value of information: how V changes when the information structure marginally changes. Formally, it is defined as the gradient of V at F.

Definition 4. *The platform's marginal value of information is defined as the gradient of the platform's value of information: $\nabla_F V$.*

Studying the platform's marginal value of information provides rich comparative statics on the platform's willingnesses to pay to acquire additional information. Changing the distribution of signals F for the platform has many economic interpretations. It can correspond to modifying the platform's contracts with data brokers. It can also correspond to a change in how much data is collected on the platform's website. For instance, many platforms require users to be logged in with personal accounts to interact with sellers and choose how much information is necessary to create such personal accounts. Identifying users with personal account makes it easier for platforms to track their purchase histories or to merge this data with external information sources. Many platforms' practices result in the collection of user information. For instance, addresses are necessary for deliveries, gift cards provide information on the nature of the purchase.

This section computes the platform's willingness to pay to acquire more information which partly captures the platform choice of data acquisition strategy. The next section compares these demands for information with the social planner's to study distortions.

Sensitivity Analysis

The platform's marginal value is constructed as the shadow price associated to the informational constraint (In): The impact on the platform's profit when relaxing this constraint. Formally the shadow price of (In) is the Lagrangian multiplier, i.e. the dual variable, associated to the constraint (In). Consider the platform's optimal price recommendation rule problem:

$$\mathcal{P}: \qquad \max_{\mu \in V_+} (1 - Q(\tilde{b})) \left(\int_0^1 (1-\rho_s) d\mu(\rho_s)(v_l - c) + \rho_0 (v_h - c) + \tilde{b} \right)$$

subject to:

$$\text{Incentive Compatibility} \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq F(S) \qquad (In)$$

Paid platforms also choose $\tilde{b}$ via the buyer entry fee and free platforms $\tilde{b}_f$ is pinned down by the price recommendation rule (see (4)). In both business models, (In) restricts the choice of price recommendation rule. Both are solved using duality[25] which computes the dual variable associated to the constrained (In) noted Λ and Λ_f respectively. Since the dual space of the space of regular measure is the space of continuous functions, the dual variables Λ and Λ_f are continuous functions on $[0, 1]$. The sensitivity analysis proves that these dual variables can be interpreted as the change in the optimal platform's profit obtained by relaxing (In), that is by changing F (see proposition 4.1 below). From this, the platform's marginal value of information is defined as the dual variable associated to (In), formally $\nabla_F V = \Lambda$ computed in appendix A.4 for paid platforms and appendix B.4 for free platforms.

In the reminder of this paper, I assume that F has full support. The full support assumption is crucial to define derivatives of the value function at F and to guarantee that dual solutions are unique.[26] From a technical standpoint, the main difference compared to standard sensitivity analysis is that V is defined on an infinitely dimensional space, the space of regular measures. This paper shows that V is directionally differentiable at F.

In this space, directional derivatives can be interpreted as follows. Consider an alternative information structure H with mean ρ_0. Acquiring information in the direction of H, starting from F, means that the new information structure of the platform is the mixture $(1-\epsilon)F+\epsilon H$ with a small probability ϵ.[27] Using arbitrary H and F captures all possible changes in information structures, since any information structure can be represented by a distribution of posterior beliefs.[28]

For instance, consider an e-commerce platform that decides to purchase social media data. This type of data produces many useful signals for platforms (current occupation,

[25]see appendicesA and B respectively.

[26]Under this assumption Dworczak and Kolotilin (2019) theorem 3 together with proposition 4 applies. See technical appendix for another proof.

[27]Note that because both F and H have mean ρ_0, the mixture also has mean ρ_0 and thus is a Bayes plausible information structure.

[28]However notice that the notion of marginal variation may not be equivalent in the space of posterior belief distributions and in the space of experiments. The first is endowed with the weak-* topology which need not be equivalent to the topology chosen on the space of experiments.

geographical location, liked pages, etc). Social media platforms like Facebook also define categories reflecting user preferences based on profiles. Typically, this data is correlated with what the platform already knows. Assume that for each buyer identified by a signal ρ_s, observing the his social media data of that buyer generates new posteriors distributed according to $G(.|\rho_s)$ with mean ρ_s. By purchasing social media data for all its buyers, the platform has access to a new information structure H equals to:[29]

$$H(S) = \int_0^1 G(S|\rho_s) dF(\rho_s).$$

In practice, it may be difficult or very costly for the platform to transform social media data into predictions on buyers' valuations. However, G need not correspond to the true distribution of predictions, rather $G(.|\rho_s)$ can be also interpreted as the platform's imperfect predictions (although belief consistent) of buyers' valuations after observing social media data.

The marginal change of F towards H can be interpreted as follows: The platform purchases a social media data set from a broker, that contains a mass ϵ of the platform's buyers. Although, the platform cannot identify in advance which buyers are in the data set. In that case, the new platform's information structure is:

$$(1-\epsilon)F + \epsilon H = F + \epsilon \int_0^1 (G(.|\rho_s) - F(.)) dF(\rho_s).$$

For marginal variations only F is required to have full support, G may be diffuse or discrete or both. In general, see proposition below, any marginal changes in the direction of H from F can be computed from the platform's marginal value of information as follows:

$$\int_0^1 \nabla_F V(\rho_s) d(H - F)(\rho_s).$$

[29]In this example, G is a dilation which implies that H is an MPS of F. The converse is also true, if H is a MPS of F then such dilation exists. See Le Cam (1996) theorem 1. Therefore, this construction captures without loss of generality any gain in informativeness.

Using the example, the marginal gain in profit from purchasing social media data to one extra buyer equals:

$$\int_0^1 \left(\int_0^1 \nabla_F V(z) G(dz|\rho_s) - \nabla V(\rho_s) \right) dF(\rho_s).$$

Since all directional derivatives are computed from the marginal value of information (the dual variable associated to (In)), the marginal value of information is the gradient of V at F. All these results are sumarized in the proposition below:

Proposition 4.1. *Assume that F has full support.*

1. *V is directionally differentiable at F in any directions. For a change towards H starting from F, the derivative equals:*

$$D_{H-F}V(F) = \int_0^1 \nabla_F V(\rho_s) d(H-F)(\rho_s).$$

2. *If H has full support, then:*

$$V(H) - V(F) = \int_0^1 \int_0^1 \nabla V_{\{(1-\epsilon)F+\epsilon H\}}(\rho_s) d(H-F)(\rho_s) d\epsilon.$$

Proof. See technical appendix. □

The last point of the proposition completes the construction: The difference in the platform's value of information between H and F can be obtained by integrating the platform's willingness to pay to learn in the direction $H - F$ along the path $(1-\epsilon)F + \epsilon H$, $\epsilon \in [0, 1]$.

The marginal value of information is $\nabla_F V_p$ for paid platform at F and is $\nabla_F V_f$ for free platforms at F. Based on the platform's marginal value of information in each model, the next subsections study the platform's incentives to collect additional information and compare these incentives with the social planner's.

The platform's marginal value of information

Appendix A (resp. appendix B) presents the dual analysis of the paid platform problem (resp. the free platform problem). Each business model's marginal value of information is computed with the dual variable associated to (In). The next proposition presents these results:

Proposition 4.2. *1. The paid platform marginal value of information is:*

$$\nabla_F V_p(\rho_s) = (1 - Q(\tilde{b})) \left(\rho_s(v_h - c) + \frac{(v_h - v_l)(v_l - c)}{(v_h - c)\rho_t - (v_l - c)} \max\{\rho_t - \rho_s, 0\} \right).$$

2. The free platform marginal value of information is:

$$\nabla_F V_f(\rho_s) = (1 - Q(\tilde{b})) \left((1 + \lambda_f) \frac{(v_l - c)}{\rho_t} \max\{\rho_t - \rho_s, 0\} + \rho_s(v_h - c) \right).$$

Where λ_f is the gain in free platform profits when relaxing the sellers' (IC) by 1.

Proof. See appendix A.4 for the paid platform case, and appendix B.4 for the free platform case. □

First, if the optimal cutoff is 1 then trade is efficient. In this case, additional information has no value. The linearity of the platform's marginal value of information captures this fact: Consider marginally changing in the platform's information structure from F in the direction of H. The formula of proposition 4.1 point 2. computes the corresponding change in profit and with the fact that H and F have the same mean ρ_0 and total mass 1, this change under both business models yields 0 profit.[30]

The construction of the marginal value of information also allows for variations of F outside of the space probability measures.[31] That is why the marginal value of information does not equal the constant map 0 in this case. Variations outside of the probability

[30] Remark that $\lambda_f = 0$ in this case.

[31] This is useful for F to be considered as an interior point and thus define derivatives. Otherwise, no point in the space of probability measures is interior.

spaces may be interpreted economically in other models as follows: Consider launching an ad campaign which changes the valuations for the good of some buyers, thus reaching a new distribution H with mean $\rho_h > \rho_0$ or which induces some buyers to purchase more than one good, thus reaching a new distribution with a mass higher than 1. These variations are also supported by the analysis, in the case where the proportion of high valuations buyers in the population increases to ρ_h, the corresponding change to paid platforms' profit equals $(1 - Q(\tilde{b}))(v_h - v_l)(\rho_h - \rho_0)$ (still assuming the cutoff is at 1), as they capture the surplus from trade. For free platforms' the change in profit is the same. Locally, free platforms' readjustment of the price recommendation rule is negligible. The extra mass on high posteriors is used to set a high price and the lost mass on low posteriors reduces the probability of recommending a low price so the change in profit is the same.

In contrast, if the cutoff is interior, the platform's marginal value of information, under both business models, has two relevant properties. First, $\nabla_F V$ is continuous and convex which is a consequence of the Blackwell theorem: The platforms' profit increases under more informative signal structures and decreases under less informative signal structures. A distribution of posterior beliefs H is generated by a more informative signal structure than another distribution F if and only if H is a mean preserving spread of F.[32] Then, from the definition mean preserving spreads, integrating any continuous and convex functions under H yields larger value than under F. Therefore, by the formula of proposition 4.1 point 2. the platform's profit increases if the information structure F marginally changes in the direction of H.

The converse is also true: If $\nabla_F V$ is convex for any F, then for all MPS H_1 of H_2, by the third point of proposition 4.1 $V(H_1) \geq V(H_2)$. Therefore, the convexity of the gradient at all F is not only a consequence of the Blackwell theorem, it is also a first order characterization of the Blackwell order.

Second, $\nabla_F V$ is piece-wise linear around the cutoff value ρ_t. Any variations on the dis-

[32] See Kolotilin (2014) proposition 1 and Blackwell (1953).

tribution of posterior beliefs F which affects only one side of the cutoff has no effect on the platform's profit. Learning buyers' types is valuable only if it affects paid platforms' price recommendations.

The next paragraphs use this concept to capture a paid platform's incentives to collect data. The next results assume that the optimal cutoff is interior, which implies that sellers are optimistic. Otherwise, as discussed above, the platform has no incentives to collect additional information.

The platform's incentives to collect data

By incentives to collect data, this paper refers to the platform's willingness to pay to change its information structure. The formula of proposition 4.1 point 2. computes a paid platform's willingness to pay in the direction of H by applying the marginal value of information $\nabla_F V$ to the direction $(H - F)$.

The platform's marginal value of information captures incentives to collect data in any direction starting from any information structure. It copes with the high dimensionality of information, yet its structure solely relies on the platform's profit. As shown in the next propositions, it reflects the way the platform uses its information.

Paid Platforms. Paid platforms use information to maximize surplus from trade while meeting the sellers' obedience condition. The paid platforms' willingness to pay to learn in the direction of H reflects these two needs: (i) the gain in trade efficiency from learning towards H, and (ii) the gain in relaxing the sellers' obedience condition.

The next proposition formally computes this gain in profit using the formula of proposition 4.1 point 2.:

Proposition 4.3. *Paid platforms' willingness to pay to learn in direction of H is characterized by:*

(i) how much H increases trade efficiency compared to F,

(ii) how much H relaxes the sellers' incentive compatibility constraint compared to F.

This willingness to pay equals:

$$(1-Q(\tilde{b}))\underbrace{(v_l-c)\left(\int_{\rho_t}^{1}(1-\rho_s)dF(\rho_s)-\int_{\rho_t}^{1}(1-\rho_s)dH(\rho_s)\right)}_{\text{Change in trade efficiency}}$$
$$+(1-Q(\tilde{b}))\underbrace{\lambda\left((v_l-c)H(\rho_t)-(v_h-c)\int_{0}^{\rho_t}\rho_s dH(\rho_s)\right)}_{\text{Change in relaxing incentive compatibility}}.$$

Where λ is the dual variable associated to the sellers' (IC):

$$\lambda=\frac{(1-\rho_t)(v_l-c)}{\rho_t(v_h-c)-(v_l-c)}.$$

Proof. See appendix A.5. □

The change in trade efficiency corresponds to the change in the probability of recommending a high price when the buyer has a low value. That is, the reduction in the mass of low type buyers above the cutoff when learning towards H: $\int_{\rho_t}^{1}\rho_s d(H-F)(\rho_s)$. Weighted by $(1-Q(\tilde{b}))(v_l-c)$ this variation gives the change in the total surplus from trade. In this model, this change in surplus from trade is first captured by sellers, as they now correctly price down these low type buyers. Then, captured by the platform as it increases the seller entry fee by the same amount their profits increase. Since the surplus from trade has changed, the platform adjusts the buyer entry fee as well, but this has no effect on profit due to standard envelop arguments.

Learning towards an information structure H also changes the platform's profit via relaxing the sellers' incentive compatibility constraint. The difference in the sellers' profit if they follow low price recommendations $(v_l-c)H(\rho_t)$ compared to if they do not $(v_h-c)\int_{0}^{\rho_t}\rho_s dH(\rho_s)$ corresponds to the play in the sellers' (IC) under H the platform can exploit to increase the cutoff. By increasing the cutoff the platform now recommends a low price around ρ_t which increases the surplus by $(1-\rho_t)(v_h - c)$ but decreases the sellers' profit (hence tightens the sellers' (IC)) by $\rho_t(v_h - c) - (v_l - c)$. The ratio of the two equals the dual variable associated to the sellers' (IC) that gives the increase in the paid platforms' profit when the sellers' (IC) the relaxed by one unit.

A paid platform's profit changes from learning toward H via only two channels (i) the change in the surplus from trade and (ii) the change in the sellers' (IC). Next, I describe a free platform's incentives to collect data.

Free Platforms. Let us focus on the case when sellers' (IC) is slack. In this case, free platforms' use information to (i) assist sellers with capturing buyer surplus and (ii) provide surplus to buyers to attract more of them. These two channels shape their willingness to pay to learn in the direction of H.

Proposition 4.4. *Assume the sellers' (IC) is slack.*
A free platform's willingness to pay to learn in the direction of H is characterized by:
(i) how much H increases the sellers' profit compared to F.
(ii) how much H increases the mass of buyers joining compared to F.

This willingness to pay equals:

$$(1-Q(\tilde{b}))\underbrace{\left((v_l - c)(H(\rho_t) - F(\rho_t)) + (v_h - c)\int_{\rho_t}^{1} \rho_s d(H-F)(\rho_s)\right)}_{\textit{Profit gained on infra-marginal trades}}$$
$$+\underbrace{(v_h - v_l)Q'(\tilde{b})\int_0^{\rho_t} \rho_s d(H-F)(\rho_s)}_{\textit{Gain in the mass of buyers joining}}\underbrace{\left((v_l - c)\int_0^{\rho_t} dF(z) + (v_h - c)\int_{\rho_t}^{1} z dF(z)\right)}_{\textit{expected profit per trade}}.$$

Proof. See appendix B.5 □

A free platform captures the change in the sellers' profit from the entry fee. The sellers' probability to trade at a low price changes by $(H(\rho_t) - F(\rho_t))$. The change in the mass of buyers below the cutoff and to trade at a high price changes by $\int_{\rho_t}^{1} \rho_s d(H - F)(\rho_s)$. Combining both gives the total change in the sellers' profit per buyer.

Learning in the direction of H also affects the mass of joining buyers. The probability for high type buyers to face a low price and obtain a surplus of $(v_h - v_l)$, changes by $\int_0^{\rho_t} \rho_s d(H - F)(\rho_s)$. Each additional unit of buyer surplus lowers the marginal buyer type joining $\tilde{b}$. This generates $Q'(\tilde{b})$ more trades, of which free platforms appropriate the sellers' profit.

The next section studies the distortions in the platform's incentives to collect data under each business model. It exhibits the biases in the way platforms collect data compared to what is socially optimal.

5 Distortions in Data Collection

This section compares the platform's incentives to collect data with the social planner's. When collecting or purchasing data, the platform compares its incentives to collect data with the marginal cost of collecting data. Assuming the social planner and the platform face the same cost of collecting information, the differences in their incentives to collect data captures the differences in their information acquisition decisions.

How the platform evaluates information structures locally determines their decision to acquire information. A first approach to study distortions is to consider whether the platform locally ranks information structures efficiently. This concept is formalized in the following definition:

Definition 5. *The platform's ranking of information structures is locally efficient at the distribution F if these two statements are equivalent:*

1. *Total surplus increases more in the direction of H_1 rather than of H_2.*

2. *The platform's profit increases more in the direction of H_1 rather than of H_2.*

Equivalently, it means that for all information structures H_1 and H_2, there is a $\epsilon > 0$ such that:

$$V((1-\epsilon)F + \epsilon H_1) \geq V((1-\epsilon)F + \epsilon H_2) \iff V^{sp}((1-\epsilon)F + \epsilon H_1) \geq V^{sp}((1-\epsilon)F + \epsilon H_2).$$

If the platform locally ranks information structure efficiently, then the only remaining distortions are a matter of intensity. The platform and social planner agree on the direction of the investment in information but may disagree on the size of that investment. Formally, there is a positive scalar γ_F such that for all information structures H:[33]

$$\int_0^1 \nabla_F V^{sp}(\rho_s) d(H-F)(\rho_s) = \gamma_F \int_0^1 \nabla_F V(\rho_s) d(H-F)(\rho_s).$$

The platform's incentives to collect data only depends on the two average posterior beliefs conditional on recommending a low price or a high price. Since all distributions have the same mean, one of the average posterior determines the other. Therefore, comparing learning directions is equivalent to compare the allocation of the mass on one side of the cutoff. Consequently, if the platform uses the same cutoff as the social planner, then the platform's ranking of information structure is locally efficient.[34] This leads to the next proposition:

Proposition 5.1. *1. Paid platforms' ranking of information structures is locally efficient at any distribution F.*

2. Free platforms' ranking of information structures is locally efficient for the distribution F where (IC) binds and locally inefficient for the distribution F where (IC) is slack.

Proof. See appendix A.6 □

[33] Considering variations outside the space of probability distributions local efficiency implies that the gradient of the platform is an affine transformation of the gradient of the social planner. The constant terms is irrelevant when considering Bayes plausible probability distributions.

[34] This argument holds because the platform recommends only two prices. If more actions are recommended it may not be true that a platform that uses information efficiently also locally ranks information structures efficiently.

The remainder of this section is split into two parts. First, the next subsection analyzes the case when the platform's ranking of information structures is locally efficient. In the case when the sellers' (IC) binds, the platform's incentives to collect data is proportionally lower than what is socially optimal. Second, the last subsection analyzes distortions in the case of free platforms if (IC) is slack. Here, the distortions are also a matter of orientation and the analysis shows that for any additional source of information free platforms are biased to learn at the bottom: in the region where posteriors are low.

Distortions in investment intensity

This subsection considers the case of paid platforms and free platforms when the seller's (IC) binds. In this case, the platform's ranking of information structure is locally efficient, regardless of its business model. Yet, this section shows that in both business models the platform under-values additional information.

Paid platforms set a higher buyer's entry fee compared to what is efficient. Accordingly, fewer buyers join which implies that any additional information is used on a lower mass of trades as compared to under a social planner's trade mechanism. Thus, Paid platforms value additional information proportionally less than what is efficient by a factor of $\frac{1-Q(\tilde{b})}{1-Q(\tilde{b}^{sp})}$.

Free platforms use the same price recommendation rule as the social planner. However, free platforms do not take into account the buyer surplus when valuing additional information. If learning relaxes the sellers' (IC), then free platforms can increase the cutoff to increase their profit. Free platforms under value this change as they dismiss the corresponding gain in the buyer surplus. Increasing the buyer surplus also attracts more buyers and generates more trades. Although, free platforms and the social planner value this gain in the same way. The additional buyer joining does not increases buyer surplus since his stand-alone valuation offsets his expected surplus from trade. Overall, free platforms undervalue additional information by a factor of $\frac{1+\lambda_f}{1+\lambda_f^{sp}}$, where λ_f gives the gain in the platform's profit from relaxing the sellers' (IC) by one unit. Similarly, λ_f^{sp} gives the gain in total surplus from relaxing the sellers' (IC) by one unit. These results are summarized in the next proposition:

Proposition 5.2. *1. A paid platform's willingness to pay is proportionally lower than the socially optimal one by a factor of* $\frac{1-Q(\tilde{b})}{1-Q(\tilde{b}^{sp})}$.

2. Consider the case where the sellers' (IC) binds. A free platform's willingness to pay to is proportionally lower than the social planner's by a factor of $\frac{Q'(\tilde{b}_f)\rho_0(v_h-c)}{(1-Q(\tilde{b}_f))+Q'(\tilde{b}_f)\rho_0(v_h-c)}$.

Proof. See appendix A.5. □

For a paid platform, the distortion between social versus private incentives to collect information is only a matter of the total mass of trades generated on the platform. However, it does not imply that, conditional on the mass of buyers joining, a paid platform's willingness to pay to learn is efficient.

To demonstrate, consider a benevolent information provider who chooses the information structure of the platform but not the trade mechanism. In this case, the benevolent information provider values the gain in surplus coming from the platform's readjustment of the buyer's entry fee when learning in the direction of H.[35] If learning increases the surplus from trade per buyer then a paid platform responds to it by lowering the buyers' entry fee, as each additional buyer is more valuable. Yet, a paid platform only captures this increase in the buyers' surplus at the marginal buyer, since it posts a fixed entry fee for all stand-alone valuations. By contrast, the benevolent information provider values the total change in the buyers' surplus.

Example. Interpret the buyers' stand-alone valuations as search costs and let b be uniformly distributed on $(\underline{b}, 0]$ where $\underline{b}$ is assumed to be low enough for each optimal $\tilde{b}$ to be interior.

Paid platforms set the buyer entry fee to pin down the marginal buyer joining according

[35] These gains were previously negligible as paid platforms were choosing t_b optimally.

to (3):

$$-\tilde{b}^* = \frac{1}{2}\left[\int_0^{\rho_t^*}(1-z)dF(z)(v_l - c) + \rho_0(v_h - c)\right].$$

But, from (3)′ the total surplus maximizing marginal buyers' type joining is:

$$-\tilde{b}^{sp} = \int_0^{\rho_t^*}(1-z)dF(z)(v_l - c) + \rho_0(v_h - c).$$

In this example, half the buyers join the platform as compared to the planner's case. In turn, the paid platforms' willingness to pay to collect data is half the social planner's.

Free platforms set the same cutoff rule that binds the sellers' (IC). The marginal buyer is:

$$-b_f = (v_h - v_l)\int_0^{\rho_t^*}\rho_s dF(\rho_s).$$

Accordingly, the free platforms' willingness to pay is lower than the social planner's by a factor of $\frac{\rho_0(v_h - c)}{(1-Q(\tilde{b}^{sp}))}$.

Since distortions are just a matter of intensity in that case, they can be reduced by a per unit subsidy. For instance, if $C(F)$ is the total cost (or market price) to collect the information structure F and if the social planner takes a share $1-\gamma_F$ of the total cost with γ_F being the factor by which platform's incentives differ from social incentives, then the platform's decision to collect or purchase information is locally efficient.[36] This share depends on the information structure F but it is observable. Indeed, for paid platforms' it only depends on the ratio between the mass of buyers joining the platform under the social planner or under paid platforms. Thus, computing this share consists in estimating the elasticity of the mass of buyers joining. Knowing this elasticity would be enough also to compute the share

[36] Remark that in this model, V is concave in F, thus under suitable conditions for the cost function C, local optimality characterizes the optimum.

for free platforms.

By contrast, the next subsection demonstrates that, for free platforms, these distortions are also a matter of direction.

Distortions in investment orientation

Consider the case of free platforms if the sellers' (IC) is slack. In this case, the free platforms' ranking of information structure is locally inefficient: distortions are both a matter of intensity and direction. I study a specific class of directions to highlight distortions.

Consider again the example where the platform purchases its buyers' social media data. Assume that for a buyer ρ_s, this data purchase generates new posterior beliefs distributed according to $G(.|\rho_s)$ with mean ρ_s. The main question is: compared to the social planner, does the platform prefer buying the social media data for low ρ_s or for high ρ_s?

First, notice that any additional source of information can be represented by a conditional distribution $G(.|\rho_s)$ that has mean ρ_s for all ρ_s.[37] Additional information simply allows the platform to spreads again its beliefs. The following analysis holds for any source of additional information.

Intuitively, free platforms value learning relatively more for low ρ_s as compared to a social planner. For ρ_s lower than their respective cutoff, trade is efficient since a low price is recommended. Thus, for the social planner, the only gain in surplus comes from relaxing the sellers' (IC). However, by learning below the cutoff, free platforms can now correctly mark-up high type buyers, which increases the sellers' profit and their profit. By contrast, learning for high ρ_s is more valuable for the social planner. Learning in the region where a high price is recommended improves trade efficiency and relaxes the sellers' (IC). Yet, in the case the sellers' (IC) is slack at optimal cutoff, free platforms do not value this last channel.

[37] distribution H is a mean preserving spread of F if and only if there is such a conditional distribution G such that: $H = \int_0^1 G(.|\rho_s)dF(\rho_s)$. See Le Cam (1996) theorem 1.

Consider the following change in the information structure: After receiving a signal in $(0, s)$, there is a probability ϵ that the platform observes buyers' social media data. Let $T_b(s)$ denote the platform willingness to pay to learn the social media data for $\rho_s \in (0, s)$, that is at the bottom. Alternatively, consider another variation: After receiving a signal in $(s, 1)$, there is a probability ϵ that the platform observes its buyers' social media data. Let $T_t(s)$ denote the free platforms' willingness to pay to learn the social media data for $\rho_s \in (s, 1)$, that is at the top.

If $s = 1$ (resp $s = 0$) then $T_b(1)$ (resp $T_t(0)$) corresponds to willingness to pay to observe buyers' social media data for all $\rho_s \in (0, 1)$. Conversely, $T_b(0) = T_t(1) = 0$. The main question is to determine if free platforms prefer learning at the bottom or at the top relative to the social planner. To remove the distortions in intensity, I normalize with the price of learning the social media data over the whole interval $T_b(1) - T_t(1)$ and $T_b^{sp}(1) - T_t^{sp}(1)$ respectively.

Proposition 5.3. *Assume F has a density. For any additional source of information G one has:*

$$\forall \ s \in (0,1) \ : \frac{T_b(s) - T_t(s)}{T_b(1) - T_t(1)} \geq \frac{T_b^{sp}(s) - T_t^{sp}(s)}{T_b^{sp}(1) - T_t^{sp}(1)}.$$

Proof. See appendix B.6. □

A free platform values marking up buyers to increase the sellers' profit. However, the social planner is only motivated by marking down low type buyers to avoid inefficiencies. Consequently, a free platform collects data or purchases data to refine its predictions when recommending a low price, whereas the social planner collects or purchases data to refine its predictions when recommending a high price.

This bias in the incentives to collect data is robust to the source of the additional information which leads to different information acquisition strategies by free platforms and the social planner. In turn, free platforms construct data bases which enriches their predictions over buyers in the wrong region.

6 Conclusion

This paper studies how e-commerce platforms design price recommendations to influence the market price and their incentives to collect information. It studies two business models: paid platforms and free platforms, and establishes a connection between a platform's business model and the way it values and uses information. To make this connection, this paper builds the platform's marginal value of information, i.e. how a platform's profit varies when its information structure varies. Using this concept, I show that a paid platform uses information efficiently and locally ranks the different directions to learn efficiently. However, a paid platform under-values additional information as compared to what is socially optimal. By contrast, a free platform may rank inefficiently learning directions. A free platform prefers learning when recommending a low price to avoid missing potential mark-ups. What is efficient, however, is to learn when recommending a high price to avoid missing trade opportunities on low type buyers.

This paper warns against potentially distorted incentives of platforms to collect information. The platform's marginal value of information exhibits these distortions and identifies these with the platforms' business model. Since the marginal value of information relates how the platform uses information with how the platform collects information, it provides two entry points to implement and study regulation. Consider implementing a regulation on the downstream market (e.g. by promoting competition between platforms). The consequences of this regulation on the platform's demand for data and on upstream market (the data market) can be analyzed through the platform's marginal value of information. Alternatively, consider a regulation on the data market (e.g. making some data public, changing the privacy regulation, taxing data, etc.). The consequence of this regulation on the downstream market can be studied through the marginal value of information.

Finally, I derive the marginal value of information using duality and sensitivity analysis. This approach may prove useful to shed light on related issues regarding the efficient allocation of information.

7 Appendices

7.1 Paid Platforms Problem

Primal and Dual program

The paid platforms' problem is:

$$\max_{\tilde{b},\ \mu}(1-Q(\tilde{b}))\left[\int_0^1 (1-z)\mu(dz)(v_l-c)+\rho_0(v_h-c)+\tilde{b}\right]$$

subject to:

$$\text{Incentive Compatibility} \qquad (IC)$$
$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

The optimal price recommendation rule can be solved independently from the optimal buyer entry fee.

Solving only for the price recommendation rule with optimistic sellers is a linear program:

$$\mathcal{P}: \qquad \max_{\mu\in V_+}\int_0^1 (1-\rho_s)d\mu(\rho_s)(v_l-c)$$

subject to:

$$\int_0^1 \mu(d\rho_s)(v_l-c)+\left[\rho_0-\int_0^1 \rho_s\mu(d\rho_s)\right](v_h-c) \geq \rho_0(v_h-c) \qquad (IC)$$
$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

Let C_+ refers to the set of non-negative continuous functions defined on $[0,1]$, and V_+ the set of non-negative measures defined on $[0,1]$. The dual writes:

$$\mathcal{D}: \qquad \min_{\Lambda \in C_+,\ \lambda \in \mathbb{R}_+} \int_0^1 \Lambda(\rho_s) dF(\rho_s)$$

subject to:

$$\forall\ \rho_s \in [0,1] \quad (v_l - c)(1+\lambda) - \rho_s[(v_l - c) + \lambda(v_h - c)] \leq \Lambda(\rho_s) \qquad (\star)$$

Technical appendix 4.1 describes the construction of the dual as well as proof that strong duality holds:

Proposition 7.1. *Problems $\mathcal{P}$ and $\mathcal{D}$ are strong duals:*

1. *Both problems have a solution and $val\{\mathcal{P}\} = val\{\mathcal{D}\}$.*

2. *Let μ and Λ, λ be feasible then:*

μ is and optimal solution of $\mathcal{P}$ and (Λ, λ) is an optimal solution of $\mathcal{D}$ if and only if

$$\begin{cases} \int \Lambda d(F - \mu) = 0 & (C1) \\ \lambda \left(\int_0^1 d\mu(\rho_s)(v_l - c) - \int_0^1 \rho_s d\mu(\rho_s)(v_h - c) \right) = 0 & (C2) \\ \int_0^1 \Lambda(\rho_s) - (v_l - c)(1+\lambda) + \rho_s[(v_l - c) + \lambda(v_h - c)] d\mu(\rho_s) = 0 & (C3) \end{cases}$$

Proof. See technical appendix 4.1. $\square$

Additionally, the planner's program is:

$$\max_{\tilde{b},\ \mu} \int_{\tilde{b}} \left(\int_0^1 (1-z)\mu(dz)(v_l - c) + \rho_0(v_h - c) + b \right) dQ(b)$$

subject to:

$$\text{Incentive Compatibility} \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1] : \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

The linear part of this problem boils down the problem $\mathcal{P}$ stated above. Therefore, the analysis of the optimal paid platforms' price recommendation rule μ applies also to the planner.

Optimality of the cutoff rule

This subsection proves lemma 3.1 for the case of paid platforms.

Pessimistic Sellers

Pessimistic sellers prefer setting a low price at their prior belief. Thus, recommending a low price for all values of the signal is (IC) is incentive compatible. This also coincide with the platform's objective. Therefore, in this case the platform uses a cutoff with cutoff $\rho_t = 1$. That is $\mu = F$ for all measurable sets included in $[0, 1]$

Optimistic Sellers

On top of the complementary slackness conditions, the solution must satisfy primal feasibility:

$$\begin{cases} \mu \in V_+ \\ \mu \leq F & (In) \\ \int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c) \geq \rho_0(v_h - c) & (IC) \end{cases}$$

As well as dual feasibility:

$$\begin{cases} \Lambda_0 \in C_+ \\ \lambda \in \mathbb{R}_+ \\ \forall\ \rho_s \in [0,1] \ \ (v_l - c)(1+\lambda) - \rho_s[(v_l - c) + \lambda(v_h - c)] \leq \Lambda(\rho_s) & (\star) \end{cases}$$

The LHS of equation $(\star)$ is an affine function of z. Since $\lambda \in \mathbb{R}_+$ this affine function starts positive at 0 and ends up non-positive at 1. It crosses the x axis at:

$$\rho_t \stackrel{def}{=} \frac{(v_l - c)(1+\lambda)}{v_l - c + (v_h - c)\lambda} \in \left(\frac{v_l - c}{v_h - c}, 1\right]$$

Because Λ is a non negative map, for all $z \in (\rho_t, 1]$ $(\star)$ is slack:

$$(1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z) < 0$$

Thus, using (C3):

$$\text{For all measurable } B \subset (\rho_t, 1] \ \ \mu(B) = 0$$

Hence, from (C1):

$$\forall\ z \ \in supp(F) \cap (\rho_t, 1], \quad \Lambda(z) = 0$$

Because Λ is continuous:

$$\forall\ z\ \in \overline{supp(F) \cap (\rho_t, 1]}, \quad \Lambda(z) = 0$$

Second, the LHS of $(\star)$ is strictly positive for $z \in [0, \rho_t)$. Thus $(\star)$ implies that for all $z \in [0, \rho_t), \quad \Lambda(z) > 0$.

So, using this in (C1):

$$\forall \text{ measurable } \ B \subset supp(F) \cap [0, \rho_t), \quad \mu(B) = F(B)$$

But on $B \subset supp(F)^c \cap [0, \rho_t)$, primal feasibility implies $\mu(B) = 0 = F(B)$. Therefore:

$$\forall \text{ measurable } \ B \subset [0, \rho_t), \quad \mu(B) = F(B)$$

By the third complementary slackness condition one has again:

$$\forall \text{ measurable } \ B \subset \overline{supp(F) \cap [0, \rho_t)}, \quad \Lambda = (1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)]$$

To sum up we have so far:

$$\begin{aligned} &\forall \text{ measurable } B : \\ &\qquad B \subset [0, \rho_t) \ \ \mu(B) = F(B) \\ &\qquad B \subset (\rho_t, 1] \ \ \mu(B) = 0 \end{aligned}$$

Additionally, from the $\sigma-$additivity property of measures, μ is pinned down up to the choice of mass at $\{\rho_t\}$.

That is, receiving a signal below ρ_t always riggers a low price recommendation, and above ρ_t

always leads to a high price recommendation. The platform can also mix recommendation at the cutoff. Which concludes the proof of for paid platforms of lemma 3.1. □

The analysis has also established the relationship between the threshold and the value of dual variables:

$$\rho_t \stackrel{def}{=} \frac{(v_l - c)(1+\lambda)}{v_l - c + (v_h - c)\lambda} \in \left(\frac{v_l - c}{v_h - c}, 1\right]$$

Equivalently:

$$\lambda = \frac{(1-\rho_t)(v_l - c)}{\rho_t(v_h - c) - (v_l - c)} \in \mathbb{R}_+$$

In addition, it has determined the value of the dual variable Λ on the closure of the support of F. To perform the sensitivity analysis, Λ is chosen outside the support to be continuous to small perturbations of F, if a perturbed F had vanishingly small mass on the entire interval then:

$$\Lambda(z) = \begin{cases} (v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)] & \text{if } \rho_s \leq \rho_t \\ 0 & \text{if } \rho_s \geq \rho_t \end{cases}$$

Optimal price recommendation rule and Fees

The case of Paid Platforms.

Proof of proposition 3.2.

The complementary slackness condition (C2) is associated to (IC):

$$\lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right) = 0$$

Case 1: Assume (IC) is slack at the solution, from $(C2)$: $\lambda = 0$.

Using the formula for ρ_t that implies $\rho_t = 1$. (IC) is indeed slack with $\rho_t = 1$ if there is a $\mu(\{1\}) \in [0, dF(1)]$ such that:

$$\begin{aligned}
&\int_0^1 dF(z)(v_l - c) - \int_0^1 z dF(z)(v_h - c) - (dF(1) - \mu(\{1\}))(v_l - v_h) > 0 \\
\iff &\frac{v_l - c}{v_h - c} + \frac{dF(1) - \mu(\{1\})}{v_h - c}(v_h - v_l) > \rho_0
\end{aligned}$$

As$\frac{v_l - c}{v_h - c} < \rho_0$, if F doesn't have a mass point at 1 the previous inequality cannot hold.

However, F may have a mass point at 1, for instance a fully informed platform has a distribution of posterior F with a mass point at 1 of size ρ_0.

If the platform has full information the inequality boils down to:

$$\begin{aligned}
&(1 - \rho_0)(v_l - c) - \mu(\{1\})(v_h - v_l) > 0 \\
\iff &\mu(\{1\}) < (1 - \rho_0)\frac{v_l - c}{v_h - v_l}
\end{aligned}$$

Together with $\mu(\{1\}) \geq 0$ that corresponds to an interval of solutions.

In a case of an arbitrary mass point such solutions are feasible if the mass point is large enough formally:

$$dF(1) > \frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}$$

$\mu(\{1\})$ can be optimally picked in the interval $\left[0, dF(1) - \frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}\right)$.

All these solutions are optimal because they are all efficient: when the platform recommends a high price the buyer has a high valuation with probability one. But we can also consider only the one that binds (IC) by choosing:

$$\rho_t = 1$$
$$\mu(\{1\}) = dF(1) - \frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}$$

So that it also corresponds to the description of 2. of proposition 3.2.

Case 2: Assume (IC) binds at the solution.

In this case we can compute precisely λ from (IC) and the formula on ρ_t:

$$\int_{[0,\rho_t)} v_l - c - z(v_h - c)dF(z) + \mu(\{\rho_t\})(v_l - c - \rho_t(v_h - c)) = 0$$

Because $\rho_t \in (\frac{v_l-c}{v_h-c}, 1]$, $\int_{[0,\rho_t)} v_l - c - z(v_h - c)dF(z)$ is strictly positive for ρ_t close to $\frac{v_l-c}{v_h-c}$ and strictly decreasing in ρ_t. In addition, because $\frac{v_l-c}{v_h-c} < \rho_0$ it is strictly negative at $\rho_t = 1$.

Therefore it changes sign only once, but it need not to be continuous as F may have mass points.

However, by (In): $\mu(\{\rho_t\}) \in [0, dF(\rho_t)]$. If F has a mass point at ρ_t, there is a unique $\mu(\{\rho_t\})$ that binds (IC), and if F doesn't have mass point at ρ_t then $\mu(\{\rho_t\}) = 0$, and there is a unique ρ_t which binds (IC).

In both scenarios there exist a unique pair $(\rho_t,\mu(\{\rho_t\}))$ that satisfies IC with equality. In turn, λ is determined by ρ_t.

Once μ^* is determined, the platform chooses tariffs optimally. $t_s = \int_0^1 d\mu^*(z)(v_l - c) + \int_0^1 zd(F - \mu^*)(z)(v_h - c)$ to capture all the seller's profit.

t_b is set to maximize the platform's profit with respect to the consumer marginal type $\tilde{b}$. Formally:

$$\max_{\tilde{b}}(1 - Q(\tilde{b}))\left[\int_0^1 (1-z)dF(z)(v_l - c) - (dF(\rho_t) - \mu(\{\rho_t\})(1-\rho_t)(v_l - c) + \rho_0(v_h - c) + \tilde{b}\right]$$

Because Q is continuously differentiable and log concave the maximum is characterized by the first order condition:

$$Q'(\tilde{b})\left[\int_0^1 (1-z)dF(z)(v_l - c) - (dF(\rho_t) - \mu(\{\rho_t\})(1-\rho_t)(v_l - c) + \rho_0(v_h - c) + \tilde{b}\right] = 1 - Q(\tilde{b})$$

Which completes the proof of proposition 3.2 □

The Planner's problem.

This subsection proves proposition.

The linear part of the planner's problem is identical to the linear part of the paid platform problem. Hence, the social planner uses the same price recommendation rule as presented in proposition 3.2.

However, the socially optimal buyer entry fee differs from the profit maximizing one. Indeed, given the optimal price recommendation rule $(\rho_t, \mu(\{\rho_t\})$:

$$-\tilde{b}^{sp} = \int_0^{\rho_t} (1-\rho_s)dF(\rho_s)(v_l - c) + \mu(\{\rho_t\})(v_l - c) + \rho_0(v_h - c)$$

Thus, compared to the profit maximizing t_b^* inducing $\tilde{b}^*$:

$$-\tilde{b}^{sp} = \frac{1 - Q(\tilde{b}^*)}{Q'(\tilde{b}^*)} - \tilde{b}^*$$

Thus:

$$-\tilde{b}^{sp} \geq -\tilde{b}^* \iff t_b^{sp} \leq t_b^*$$

That is, the planner set a lower buyer entry fee and attract more buyers joining.

Which completes the proof of proposition 3.3. □

Paid Platforms' value for Data

The dual variable from the dual problem is:

$$\Lambda(z) = \begin{cases} (v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)] & \text{if } \rho_s \leq \rho_t \\ 0 & \text{if } \rho_s \geq \rho_t \end{cases}$$

This problem was ignoring constant terms, in particular using strong duality the value of the problem (with the constant terms) is:

$$Val(\mathcal{P}) = (1 - Q(\tilde{b}))\left(\int_0^1 \max\{(v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)], 0\} + \int_0^1 \rho_s dF(\rho_s) + \tilde{b}\right)$$

Thus the gradient is:

$$\nabla_F V(\rho_s) = (1 - Q(\tilde{b}))\left(\rho_s(v_h - c) + \max\{(v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)], 0\}\right)$$

In the previous subsections, solving the paid platforms' problem yields the following dual variables:

$$(1-Q(\tilde{b}))\Lambda(\rho_s) = \nabla_F V_p(\rho_s) = (1-Q(\tilde{b}))\begin{cases}(v_l-c)(1+\lambda)-\rho_s[(v_l-c)+\lambda(v_h-c)] & \text{if } \rho_s \le \rho_t \\ 0 & \text{if } \rho_s \ge \rho_t\end{cases}$$

And:

$$\rho_t \stackrel{def}{=} \frac{(v_l-c)(1+\lambda)}{v_l-c+(v_h-c)\lambda} \in \left(\frac{v_l-c}{v_h-c}, 1\right]$$

Equivalently:

$$\lambda = \frac{(1-\rho_t)(v_l-c)}{\rho_t(v_h-c)-(v_l-c)} \in \mathbb{R}_+$$

Under pessimistic sellers, or when the optimal cutoff is at 1 (in both cases $\lambda = 0$ the platform's marginal value of information is:

$$\nabla_F V(\rho_s) = (1-Q(\tilde{b}))\left(\rho_s(v_h-c)+(1-\rho_s)(v_l-c)\right)$$

When ρ_t is interior ($\lambda > 0$) and using the formula relating λ with ρ_t one has:

$$\nabla_F V_p(\rho_s) = (1-Q(\tilde{b}))\left(\rho_s(v_h-c)+\frac{(v_h-v_l)(v_l-c)}{(v_h-c)\rho_t-(v_l-c)}\max\{\rho_t-\rho_s,0\}\right)$$

Which completes the proof of proposition 4.2 □

Directional derivatives

This subsection proves .

The change in the paid platform's profit in a direction $H - F$ is given by:

$$(1 - Q(\tilde{b}^*)) \int_0^1 \nabla_F V(\rho_s) d(H - F)(\rho_s)$$

For clarity lets compute it per buyer (dividing by $1 - Q(\tilde{b}^*)$):

$$\int_0^{\rho_t} (v_l - c)(1 - \rho_s) d(H - F)(\rho_s) + \lambda \int_0^{\rho_t} ((v_l - c) - \rho_s(v_h - c)) d(H - F)(\rho_s)$$

Because H and F have a total mass of 1 and a mean of ρ_0:

$$\int_0^{\rho_t} (1 - \rho_s) d(H - F)(\rho_s) = - \int_{\rho_t}^1 (1 - \rho_s) d(H - F)(\rho_s)$$

Using this, and $(C3)$:

$$\int_{\rho_t}^1 ((v_l - c)(1 - \rho_s) d(F - H)(\rho_s) + \lambda \int_0^{\rho_t} ((v_l - c) - \rho_s(v_h - c)) dH(\rho_s) + (1 - \rho_t)(v_l - c)[dF(\rho_t) - \mu(\{\rho_t$$

$$=(v_l - c) \left(\left(\int_{[\rho_t, 1]} (1 - \rho_s) dF(\rho_s) - \mu(\{\rho_t\}) \right) - \int_{\rho_t}^1 (1 - \rho_s) dH(\rho_s) \right) + \lambda \int_0^{\rho_t} ((v_l - c) - \rho_s(v_h - c)) dH(\rho_s)$$

Which is the formula of proposition 4.1 allowing for mass points at the threshold. Thus concluding the proof of proposition 4.1 □

Moreover, since the formula was computed per buyer (dividing by the mass of buyer joining). The same formula can be used to obtain proposition 4.2.

Ranking Directions

This section proves proposition 4.3

First, I prove the following lemma:

Lemma 7.2. *Consider a planner or a platform under any business model, let ρ_t^* be their respective optimal cutoff. Then, learning in direction H_1 is more valuable than learning in direction H_2 if and only if:*

$$\int_0^{\rho_t^*} H_1(\rho_s)d\rho_s \geq \int_0^{\rho_t^*} H_2(\rho_s)d\rho_s$$

Proof. Under each business model the platform's marginal value of information is continuous and piece-wise linear around the cutoff, let $L_1(\rho_s) = a_1\rho_s + b_1$ and $L_2(\rho_s) = a_2\rho_s + b_2$. Continuity implies $a_1\rho_t + b_1 = a_2\rho_t + b_2$ Now compare two directions H_1 and H_2:

$$\begin{aligned}&\int_0^1 \nabla_F V(\rho_s)d(H_1 - F)(\rho_s) - \int_0^1 \nabla_F V(\rho_s)d(H_2 - F)(\rho_s)\\ =&\int_0^{\rho_t} (a_1\rho_s + b_1)d(H_1 - H_2)(\rho_s) + \int_{\rho_t}^1 (a_2\rho_s + b_2)d(H_1 - H_2)(\rho_s)\end{aligned}$$

Using IBP together with the fact that H_1 and H_2 have the same mean and mass yields:

$$=(b_1 - b_2 - (a_1 - a_2)\rho_t)\int_0^{\rho_t} d(H_1 - H_2)(\rho_s) - (a_1 - a_2)\int_0^{\rho_t} (H_1(\rho_s) - H_2(\rho_s))d\rho_s$$

The first term is 0 because the marginal value of information is continuous at ρ_t. Further, if $a_1 < a_2$ (the marginal value is strictly convex) then:

$$\int_0^1 \nabla V(\rho_s)d(H_1 - F)(\rho_s) \geq \int_0^1 \nabla_F V d(H_2 - F)(\rho_s) \iff \int_0^{\rho_t} H_1(\rho_s)d\rho_s \geq \int_0^{\rho_t} H_2(\rho_s)d\rho_s$$

Which completes the proof of the lemma. □

The lemma implies that, the platform ranks learning direction in the same way as the social planner if and only if it uses the same cutoff rule as the social planner.

In particular, it shows that the marginal value of information for the social planner and the platform are proportional at F by a factor of $\frac{a_2^{sp}-a_1^{sp}}{a_2^{*}-a_1^{*}}$. Indeed, using the formula obtained in each case these coefficients are:

$$\begin{aligned} \text{For paid platforms} &: \frac{1-Q(\tilde{b}^{sp})}{1-Q(\tilde{b}^{*})} \\ \text{For free platforms} &: \frac{1+\lambda_f^{sp}}{1+\lambda_f} \end{aligned}$$

It remains to show that this is equivalent to locally rank information structures the same way a social planner would.

7.2 The Free Platforms' Problem

Platforms' and Planner's programs and optimality conditions

The free platforms' problem is:

$$\max_{\mu}(1-Q(\tilde{b}_f))\left[\int_0^1 \mu(dz)(v_l-c)+\left[\rho_0-\int_0^1 z\mu(dz)\right](v_h-c)\right]$$

subject to:

$$\text{Incentive Compatibility} \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

where:

$$\tilde{b}_f = -(v_h - v_l)\int_0^1 \rho_s dF(\rho_s)$$

The Lagrangian is formally:

$$\mathcal{L} = (1 - Q(\tilde{b}))(1+\lambda)\left[(v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right]$$
$$- (1 - Q(\tilde{b}))\lambda \max\{v_l - c, \rho_0(v_h - c)\} + \int_0^1 \Lambda(z) d(F - \mu)(z) + \int_0^1 \Lambda_0(z) d\mu(z)$$

The KKT conditions are:

$$\begin{cases} 0 \leq \mu \leq F & (K1) \\ \Lambda_f, \Lambda_0 \in C_+, \;\; \lambda_l \geq 0 & (K2) \\ (v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right) \geq \max\{\rho_0(v_h - c), v_l - c\} & (K3) \\ \int_0^1 \Lambda_f(z) d(F - \mu)(z) = 0, \;\; \int_0^1 \Lambda_0(z) d\mu(z) = 0 & (K4) \\ \lambda_f\left[(v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right) - \max\{\rho_0(v_h - c), v_l - c\}\right] = 0 & (K5) \\ \forall \rho_s \in [0,1], \quad \rho_s(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ +(1 - Q(\tilde{b}))(1 + \lambda_f)\left((v_l - c) - \rho_s(v_h - c)\right) = \Lambda_f(\rho_s) - \Lambda_0(\rho_s) & (K6) \end{cases}$$

These conditions are direct extensions of standard KKT conditions to Hilbert spaces.[38]

The fact that $(K6)$ indeed corresponds to the gradient of Lagrangian by μ is proved in the technical appendix.

The planner's problem.

[38] See e.g. proposition 2.13. in Lecture Notes, 285J Martin Burger UCLA.

The planner's problem, when constrained to set $t_b = 0$, is:

$$\max_{\mu} \int_{\tilde{b}} \left(\int_0^1 (1-z)\mu(dz)(v_l - c) + \rho_0(v_h - c) + b \right) dQ(b)$$

subject to:

Incentive Compatibility (IC)

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq \int_S dF(\rho_s) \qquad (In)$$

The Lagrangian is formally:

$$\mathcal{L} = \int_{\tilde{b}} \left(\left((v_l - c) \int_0^1 (1-z)d\mu(z) + \rho_0(v_h - c) + b \right) \right) dQ(b) + \int_0^1 \Lambda(z)d(F-\mu)(z) + \int_0^1 \Lambda_0(z)d\mu(z)$$
$$-(1-Q(\tilde{b}))\lambda \left((v_l - c) \int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z) \right) - \max\{\rho_0(v_h - c), v_l - c\} \right)$$

Only the stationarity condition (K6) changes:

$$\forall \rho_s \in [0,1], \quad \rho_s(v_h - v_l)Q'(\tilde{b}) \left((v_l - c) \int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z) \right) \right)$$
$$+(1-Q(\tilde{b}))((1-\rho_s)(v_l - c) + \lambda_f((v_l - c) - \rho_s(v_h - c)) = \Lambda_f(\rho_s) - \Lambda_0(\rho_s) \quad (K6)^{sp}$$

Optimality of the cutoff rule The proof procedes as in appendix A.2.

The LHS of (K6) is decreasing and affine in ρ_s and positive at $\rho_s = 0$. It may be negative at $\rho_s = 1$, in which case it crosses the x-axis at:

$$\rho_{t,f} \stackrel{def}{=} \frac{(1-Q(\tilde{b})(1+\lambda_f)(v_l - c)}{(1-Q(\tilde{b})(1+\lambda_f)(v_h - c) - (v_h - v_l)Q'(\tilde{b}) \left((v_l - c) \int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z) \right) \right)}$$

Below $\rho_{t,f}$ the LHS is strictly negative and above strictly positive thus (K2) implies:

$$\Lambda_f(\rho_s) > 0 \;\; \forall \rho_s < \rho_{t,f}$$

$$\Lambda_0(\rho_s) > 0 \;\; \forall \rho_s > \rho_{t,f}$$

Hence from (K4):

$$\text{for all measurable } S \subset (\rho_{t,f}, 1], \quad \mu(S) = 0$$

$$\text{for all measurable } S \subset [0, \rho_{t,f}), \quad \mu(S) = F(S)$$

Therefore, by the $\sigma-$additivity of measures, only the mass at $\{\rho_t\}$ remains to be determined. Hence, the optimal price recommendation rule is a cutoff rule: above the cutoff the platform recommends v_h and below it recommends v_l.

If the LHS never cross the x-axis, then from (K2), for all ρ_s :

$$\Lambda_f > 0$$

$$\Lambda_0 = 0$$

Thus, by (K4) for all measurable subsets: $\mu(S) = F(S)$. That is the optimal rule is an extreme price recommendation rule with cutoff at 1.

The Social Planner

Compared to the Free platforms' problem, the social planner's problem has a different $(K6)^{sp}$.

The LHS is a decreasing affine function that crosses 0 at a point:

$$\rho_{t,f}^{sp} \overset{def}{=} \frac{(1-Q(\tilde{b})(1+\lambda_f)(v_l-c)}{(1-Q(\tilde{b})((v_l-c)+\lambda_f(v_h-c))-(v_h-v_l)Q'(\tilde{b})\left((v_l-c)\int_0^1 d\mu(z)+(v_h-c)\left(\rho_0-\int_0^1 zd\mu(z)\right)\right)}$$

The same arguments follow, the planner uses a cutoff rule around $\rho_{t,f}^{sp}$.

Which concludes the proof of lemma 3.1. □

Value of the Dual Variables

The value of the cutoff pins down the value of the dual variable associated to (IC). For free platforms that value is:

$$\lambda_f = \frac{\rho_{t,f}(v_h-v_l)Q'(\tilde{b})\left((v_l-c)\int_0^1 d\mu(z)+(v_h-c)\left(\rho_0-\int_0^1 zd\mu(z)\right)\right)-(1-Q(\tilde{b}))(\rho_t(v_h-c)-(v_l-c))}{(1-Q(\tilde{b}))(\rho_t(v_h-c)-(v_l-c))}$$

Additionally, choosing the maps Λ_f and Λ_0 continuous to small perturbation of the support of F (which also correspond to the smallest norm dual variables), yields:

$$\Lambda_f(\rho_s) = \begin{cases} \rho_s(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ +(1 - Q(\tilde{b}))(1 + \lambda_f)\left((v_l - c) - \rho_s(v_h - c)\right) & \text{for } \rho_s \leq \rho_t \\ 0 & \text{otherwise} \end{cases}$$

$$\Lambda_0(\rho_s) = \begin{cases} 0 & \text{for } \rho_s \leq \rho_t \\ -\rho_s(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ -(1 - Q(\tilde{b}))(1 + \lambda_f)\left((v_l - c) - \rho_s(v_h - c)\right) & \text{otherwise} \end{cases}$$

For the **social planner** this values are respectively:

$$\lambda_f = \frac{\rho_{t,f}(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) + (1 - Q(\tilde{b}))(1 - \rho_t)(v_l - c)}{(1 - Q(\tilde{b}))(\rho_t(v_h - c) - (v_l - c))}$$

And:

$$\Lambda_f(\rho_s) = \begin{cases} \rho_s(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ +(1 - Q(\tilde{b}))((1 - \rho_s)(v_l - c) + \lambda_f\left((v_l - c) - \rho_s(v_h - c)\right)) & \text{for } \rho_s \leq \rho_t \\ 0 & \text{otherwise} \end{cases}$$

$$\Lambda_0(\rho_s) = \begin{cases} 0 & \text{for } \rho_s \leq \rho_t \\ -\rho_s(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ -(1 - Q(\tilde{b}))((1 - \rho_s)(v_l - c) + \lambda_f\left((v_l - c) - \rho_s(v_h - c)\right)) & \text{otherwise} \end{cases}$$

Optimal price recommendation rule and Sellers' Fee.

The case of Free Platforms.

This subsection proves proposition 3.4. It makes use of the Complementary slackness condition associated to the sellers' (IC):

$$\lambda_f \left[(v_l - c) \int_0^1 d\mu(z) + (v_h - c) \left(\rho_0 - \int_0^1 z d\mu(z) \right) - \max\{\rho_0(v_h - c), v_l - c\} \right] = 0 \quad (K5)$$

Case 1: Assume (IC_l) is slack at solution.

(K5) implies $\lambda_f = 0$. Therefore, $(K6)^{sp}$ implies $\rho^{sp}_{t,f} > 1$.

Therefore, $\mu = F$ so the planner, if (IC) is slack, always recommend a low price.

This is consistent with (IC) being slack only if sellers are pessimistic that is if $\rho_0 \leq \frac{v_l - c}{v_h - c}$. But if sellers are optimistic then a planner's solution in which (IC) is slack is not feasible.

$$\tilde{b} = -(v_h - v_l) \left(\int_{[0,\rho_t)} z dF(z) + \rho_t \mu(\{\rho_t\}) \right)$$

Case 2: Assume (IC) binds at the solution.

Because $\rho^{sp}_{t,f} > \frac{v_l - c}{v_h - c}$ then:

$$(v_h - c) \int_{\rho_{t,f}} \rho_s dF(\rho_s) > (v_l - c) \int_{\rho_{t,f}} dF(\rho_s)$$

Thus, (IC) is never binding for pessimistic sellers. So when sellers are pessimistic the only solution is the one described in case 1.

Case 2 is feasible only for optimistic sellers, that is if $\rho_0 \geq \frac{v_l - c}{v_h - c}$.

The analysis is now the same as appendix A.3. There is a unique pair $(\rho_t, \mu(\{\rho_t\})$ that binds (IC):

$$\int_{[0,\rho_t)} v_l - c - z(v_h - c)dF(z) + \mu(\{\rho_t\})(v_l - c - \rho_t(v_h - c)) = 0$$

If F has a mass point of at least $\frac{\rho_0(v_h-c)-(v_l-c)}{v_h-v_l}$ at 1, then $\rho_{t,f} = 1$.

The pair $(\rho_t, \mu(\{\rho_t\})$ that binds (IC) in turn characterize λ_f:

$$\lambda_f = \frac{\rho_{t,f}(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_{[0,\rho_{t,f})} dF(z) + (v_h - c)\int_{[\rho_{t,f},1]} zdF(z) + \mu(\{\rho_{t,f}\})((v_l - c) - \rho_{t,f}(v_h - c))\right)}{(1 - Q(\tilde{b}))(\rho_{t,f}(v_h - c) - (v_l - c))}$$

λ_f is the ration of the platform's profit variation over the sellers' profit variation when increasing the threshold (or marginally increasing the mass of low price recommendations at the threshold if it is a mass point).

Whether the solution is slack or binding, one can consider the value that binds (IC) on $(5)'$:

$$\rho_t{}^*(v_h - v_l)Q'(\tilde{b})\rho_0(v_h - c) - (1 - Q(\tilde{b})(\rho_t(v_h - c) - (v_l - c))$$

Rearranging, (IC) slack if and only if:

$$\frac{1 - Q}{Q'} > 1 + \lambda$$

Which concludes the proof of proposition 3.4. □

The planner's problem.

This subsection presents the planner's solution when constrained to set $t_b = 0$, which is used in proposition 3.5. It makes use of the Complementary slackness condition associated to the sellers' (IC):

$$\lambda_f \left[(v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right) - \max\{\rho_0(v_h - c), v_l - c\}\right] = 0 \quad (K5)$$

Case 1: Assume (IC_l) is slack at solution.

From (K5): $\lambda_f = 0$, one must find a pair $(\rho_t, \mu(\{\rho_t\})$ with $\mu(\{\rho_t\}) \in [0, dF(\rho_t)]$ that satisfies:

$$\begin{aligned}&\rho_{t,f}(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_{[0,\rho_t)} dF(\rho_s) + \int_{[\rho_t,1]} \rho_s dF(\rho_s)(v_h - c) + \mu(\{\rho_t\})((v_l - c) - \rho_t(v_h - c))\right)\\ =&(1 - Q(\tilde{b}))(\rho_t(v_h - c) - (v_l - c)) \quad (K6\end{aligned}$$

Where:

$$\tilde{b} = -(v_h - v_l)\left(\int_{[0,\rho_t)} z dF(z) + \rho_t \mu(\{\rho_t\})\right)$$

Considering $(K6)$ at $\rho_{t,f} = 0$, the LHS is larger than the RHS, and at $\rho_{t,f} = 1$, the LHS is smaller than the RHS if:

$$1 - Q(-\rho_0(v_h - v_l)) > Q'(-\rho_0(v_h - v_l))(v_l - c)$$

If this is the case, then from the log-concavity of Q ($\frac{Q'}{1-Q}$ is decreasing in $\rho_{t,f}$) there is a unique pair $(\rho_{t,f}, \mu(\{\rho_{t,f}\}))$ solving the $(K6)$, which characterize the maximum in this case. In addition if there is no mass point at the cutoff, $(K6)$ becomes then (5) as displayed in proposition 3.4.

If otherwise $1 - Q(-\rho_0(v_h - v_l)) < Q'(-\rho_0(v_h - v_l))(v_l - c)$, then the solution is to always recommend a low price. This case only happen if sellers are pessimistic $\rho_0 \leq \frac{v_l - c}{v_h - c}$ otherwise (IC) would be violated.

Case 2: Assume (IC) binds at the solution.

Because $\rho_{t,f} > \frac{v_l - c}{v_h - c}$ then:

$$(v_h - c)\int_{\rho_{t,f}} \rho_s dF(\rho_s) > (v_l - c)\int_{\rho_{t,f}} dF(\rho_s)$$

Thus, (IC) is never binding for pessimistic sellers.

This case is feasible only for optimistic sellers, that is $\rho_0 \geq \frac{v_l - c}{v_h - c}$.

The analysis is now the same as appendix A.3. There is a unique pair $(\rho_t, \mu(\{\rho_t\})$ that binds (IC):

$$\int_{[0,\rho_t)} v_l - c - z(v_h - c)dF(z) + \mu(\{\rho_t\})(v_l - c - \rho_t(v_h - c)) = 0$$

If F has a mass point of at least $\frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}$ at 1, then $\rho_{t,f} = 1$.

The pair $(\rho_t, \mu(\{\rho_t\})$ that binds (IC) in turn characterize λ_f:

$$\lambda_f = \frac{\rho_{t,f}(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_{[0,\rho_{t,f})} dF(z) + (v_h - c)\int_{[\rho_{t,f},1]} z dF(z) + \mu(\{\rho_{t,f}\})((v_l - c) - \rho_{t,f}(v_h - c))\right)}{(1 - Q(\tilde{b}))(\rho_{t,f}(v_h - c) - (v_l - c))}$$
$$+ \frac{(1 - \rho_{t,f})(v_l - c)}{(\rho_{t,f}(v_h - c) - (v_l - c))}$$

λ_f is the ration of the total surplus variation over the sellers' profit variation when increasing the threshold (or marginally increasing the mass of low price recommendations at the threshold if it is a mass point).

This solution will be used for the proof of proposition 3.4. □

Value of Information

Free platforms' value of Information.

The value of the dual variable is:

$$\Lambda_f(\rho_s) = \begin{cases} \rho_s(v_h - v_l)Q'(\tilde{b}) \left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ +(1 - Q(\tilde{b}))(1+\lambda_f)\left((v_l - c) - \rho_s(v_h - c)\right) & \text{for } \rho_s \leq \rho_t \\ 0 & \text{otherwise} \end{cases}$$

Using strong duality, the value of the dual is the value of the problem:

$$V_f(F) = \int_0^1 \Lambda_f(\rho_s) + (1 - Q(\tilde{b}))(v_h - c)\int_0^1 \rho_s dF(\rho_s)$$

Using $(K6)$ the gradient is thus:

$$\nabla_F V_f(\rho_s) = (1 - Q(\tilde{b}))\left((1+\lambda_f)\frac{(v_l - c)}{\rho_t}\max\{\rho_t - \rho_s, 0\} + \rho_s(v_h - c)\right)$$

where:

$$\lambda_f = \frac{\rho_{t,f}(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) - (1 - Q(\tilde{b}))(\rho_t(v_h - c) - (v_l - c))}{(1 - Q(\tilde{b}))(\rho_t(v_h - c) - (v_l - c))}$$

Planners' value of Information.

The value of the dual variable is:

$$\Lambda_f(\rho_s) = \begin{cases} \rho_s(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) & \\ +(1 - Q(\tilde{b}))((1 - \rho_s)(v_l - c) + \lambda_f\left((v_l - c) - \rho_s(v_h - c)\right)) & \text{for } \rho_s \leq \rho_t \\ 0 & \text{otherwise} \end{cases}$$

Using $(K6^{sp})$:

$$\nabla_F V_f^{sp}(\rho_s) = (1 - Q(\tilde{b}))(1 + \lambda_f^{sp})\frac{(v_l - c)}{\rho_t}\max\{\rho_t - \rho_s, 0\}$$

where:

$$\lambda_f^{sp} = \frac{\rho_{t,f}(v_h - v_l)Q'(\tilde{b})\left((v_l - c)\int_0^1 d\mu(z) + (v_h - c)\left(\rho_0 - \int_0^1 z d\mu(z)\right)\right) + (1 - Q(\tilde{b}))(1 - \rho_t)(v_l - c)}{(1 - Q(\tilde{b}))(\rho_t(v_h - c) - (v_l - c))}$$

Directional derivatives

For free platforms.

This subsection proves .

From the gradient of V, the change in the free platform's profit in a direction $H - F$ is given by:

$$\int_0^1 \nabla_F V_f(\rho_s) d(H - F)(\rho_s)$$

In the case where the optimal cutoff is interior:

$$\int_0^{\rho_t} \rho_s(v_h - v_l)Q'(\tilde{b}) \left((v_l - c) \int_{[0,\rho_t)} dF(\rho_s) + \int_{[\rho_t,1]} \rho_s dF(\rho_s)(v_h - c) + \mu(\{\rho_t\})((v_l - c) - \rho_t(v_h - c)) \right)$$
$$(1 - Q(\tilde{b}))((v_l - c) - \rho_s(v_h - c))d(H - F)(\rho_s)$$

$$= Q'(\tilde{b})(v_h - v_l) \int_0^{\rho_t} \rho_s d(H - F)(\rho_s) \left((v_l - c) \int_{[0,\rho_t)} dF(\rho_s) + \int_{[\rho_t,1]} \rho_s dF(\rho_s)(v_h - c) + \mu(\{\rho_t\})((v_l - c) - \right.$$
$$\left. (1 - Q(\tilde{b})) \left((v_l - c)(H(\rho_t) - F(\rho_t)) + (v_h - c) \int_{\rho_t}^1 d(H - F)(\rho_s) \right) \right.$$

Which is the equation presented in proposition 4. including mass points.

For the case where (IC) is binding:

$$\begin{aligned}
&\int_0^{\rho_t} \rho_s(v_h - v_l)Q'(\tilde{b})\rho_0(v_h - c) + (1 - Q(\tilde{b}))(1 + \lambda_f)((v_l - c) - \rho_s(v_h - c))d(H - F)(\rho_s) \\
= &\int_0^{\rho_t} \rho_s d(F - H)(\rho_s)(v_h - v_l)Q'(\tilde{b})\rho_0(v_h - c) \\
&+ (1 - Q(\tilde{b})) \left((v_l - c)(H(\rho_t) - F(\rho_t)) + (v_h - c) \int_{\rho_t}^1 \rho_s d(H - F)(\rho_s) \right) \\
&+ (1 - Q(\tilde{b}))\lambda_f \left((v_l - c)H(\rho_t) - (v_h - c) \int_0^{\rho_s} \rho_s dH(\rho_s) \right)
\end{aligned}$$

Using the fact that H and F have a total mass of 1 and a mean of ρ_0:

$$\int_0^{\rho_t} (1 - \rho_s)d(H - F)(\rho_s) = -\int_{\rho_t}^1 (1 - \rho_s)d(H - F)(\rho_s)$$

Thus concluding the proof of proposition 4.. $\square$

For the social planner.

When (IC) doesn't bind the gradient is 0 and so are directional derivatives.

When (IC) binds at the planner's solution:

$$\int_0^{\rho_t} (\rho_s(v_h - v_l)Q'(\tilde{b})\rho_0(v_h - c) + (1 - Q(\tilde{b}))((1-\rho_s)(v_l - c) + \lambda_f^{sp}((v_l - c) - \rho_s(v_h - c))))d(H-F)(\rho_s)$$
$$= \int_0^{\rho_t} \rho_s d(H-F)(\rho_s)(v_h - v_l)Q'(\tilde{b})\rho_0(v_h - c)$$
$$+ (1 - Q(\tilde{b}))(v_l - c)\left(\int_{\rho_t}^1 (1-\rho_s)dF(\rho_s) - \int_{\rho_t}^1 (1-\rho_s)dH(\rho_s)\right)$$
$$(1 - Q(\tilde{b}))\lambda_f^{sp}\left((v_l - c)H(\rho_t) - (v_h - c)\int_0^{\rho_t} \rho_s dH(\rho_s)\right)$$

Which is the equation presented in proposition 4.. □

The price of reducing uncertainty at the top and at the bottom

Consider the dilation G, that is G is a Markov kernel such that $\int_0^1 zG(dz|\rho_s) = \rho_s)$ for all ρ_s.

From Le Cam (1996) theorem 1, H is a MPS of F if and only if there is a dilation G such that: $\int_0^1 G(S|\rho_s)F(d\rho_s) = H(S)$ for all measurable set S. Now, lets define the price of learning at the bottom and at the top:

Definition 6. *Let $T_b(s)$ (resp. $T_b^{sp}(s)$) be the platform's (resp planner's) incentives to learn G at the bottom, i.e conditional on $\rho_s \in (0, s]$.*

Similarly, Let $T_t(s)$ (resp. $T_t^{sp}(s)$)be the platform's (resp planner's) incentives to learn G

at the top, i.e conditional on $\rho_s \in (s, 1)$. Formally:

$$T_b(s) = \int_0^s \left(\int_0^1 \nabla_F V(z) G(dz|\rho_s) - \int_0^1 \nabla_F V(z) \right) F(d\rho_s)$$
$$T_t(s) = \int_s^1 \left(\int_0^1 \nabla_F V(z) G(dz|\rho_s) - \int_0^1 \nabla_F V(z) \right) F(d\rho_s)$$

That is, the platform learns G only for $\rho_s \leq s$. Because F has full support, these variations for all $s \in (0, 1)$ are feasible.

In each case, the marginal value is proportional to $\max\{\rho_t - \rho_s, 0\}$. In the ratio, these factors vanishes. Thus one can focus one:

$$T_b(s) \propto \int_0^s \int_0^1 \mathbf{1}_{[0,\rho_t]}(z)(\rho_t - z) G(dz|\rho_s) F(d\rho_s) - \int_0^s \mathbf{1}_{[0,\rho_t]}(z)(\rho_t - z) F(dz)$$
$$= \int_0^s \int_0^1 \mathbf{1}_{(\rho_t,1]}(z)(z - \rho_t) G(dz|\rho_s) F(d\rho_s) - \int_0^s \mathbf{1}_{(\rho_t,1]}(z)(z - \rho_t) F(dz)$$

Because G is a dilation, similarly:

$$T_t(s) \propto \int_s^1 \int_0^1 \mathbf{1}_{(\rho_t,1]}(z)(z - \rho_t) G(dz|\rho_s) F(d\rho_s) - \int_s^1 \mathbf{1}_{(\rho_t,1]}(z)(z - \rho_t) F(dz)$$

Since F is absolutely continuous, the slope of $T_b - T_t$ is s is:

$$2f(s) \left[\int_0^1 \mathbf{1}_{(\rho_t,1]}(z)(z - \rho_t) G(dz|s) - \mathbf{1}_{(\rho_t,1]}(s)(s - \rho_t) \right]$$

Which is positive by Jensen's inequality since $z \mapsto \mathbf{1}_{(\rho_t,1]}(z)(z - \rho_t)$ is convex.
The cross derivative is (in ρ_t and ρ_s):

$$2f(s) \left(G(\rho_t|s) - \mathbf{1}_{(0,\rho_t]}(z) \right)$$

That is the slope is decreasing before the cutoff ρ_t in ρ_t and increasing after the cutoff in

ρ_t. Further, the cross derivatives increases in s. Together with the facts that $\rho_{t,f} < \rho_{t,f}^{sp}$ and that at $s = 1$ and $s = 0$ both ratios are equal. Thus:

1. For $s \in (0, \rho_t^{sp})$ the platform's ratio increases more than the planner's one, starting at the same point.

2. For $s \in (\rho_t^{sp}, 1)$ the platform's ratio increases less than the planner's one, ending at the same point.

That is the ratios never cross and the platform's one is always higher than the planner's one. $\square$

7.3 Comparison of Business Models

Comparison of price recommendation rule

The optimal t_b^* for paid platforms solves:

$$\frac{1 - Q(\tilde{b})}{Q'(\tilde{b})} - t_b - \rho_0(v_h - c) = 0$$

Since Q is log concave, hence the LHS is decreasing. Further, fixing the entry fee to 0 thus having $\tilde{b}_f$:

$$sign\left\{\frac{1 - Q(\tilde{b}_f)}{Q'(\tilde{b}_f)} - \rho_0(v_h - c)\right\} = sign\{t_b^*\}$$

For free platforms, consider the LHS of $(K6)$ at the threshold value and for ρ_t binding (IC) and $\lambda_f = 0$ one has:

$$\rho_t(v_h - v_l)Q'(\tilde{b}_f)\rho_0(v_h - c) - (1 - Q(\tilde{b}_f))(\rho_t(v_h - c) - (v_l - c))$$

Rearranging:

$$\frac{\rho_t(v_h - v_l)}{\rho_t(v_h - c) - (v_l - c)}\rho_0(v_h - c) - \frac{1 - Q(\tilde{b}_f)}{Q'(\tilde{b}_f)}$$

Which is decreasing in ρ_t as Q is log concave. Thus, if this quantity is positive, then to satisfy $(K6)$ ρ_t must be reduced, thus (IC) slack at the solution.

If it is negative, then ρ_t should be increased, thus (IC) binds at the solution.

In addition $\frac{\rho_t(v_h - v_l)}{\rho_t(v_h - c) - (v_l - c)} \geq 1$ and equal to 1 if and only if $\rho_t = 1$. Thus, there are ρ_t such that:

$$\frac{\rho_t(v_h - v_l)}{\rho_t(v_h - c) - (v_l - c)}\rho_0(v_h - c) - \frac{1 - Q(\tilde{b}_f)}{Q'(\tilde{b}_f)} > 0$$

$$\frac{1 - Q(\tilde{b}_f)}{Q'(\tilde{b}_f)} - \rho_0(v_h - c) > 0$$

Thus there exists $\overline{t_b} > 0$ such that if $t_b^* > \overline{t_b}$ then (IC) slack at the free platforms' solution. If $t_b^* \leq \overline{t_b}$ then (IC) binds at the free platforms' solution.

7.4 Bilateral Trade under Private Terms of Transaction

This section gives a rational for the choice of the trade mechanism. While the platform only sets entry fees, and does not screen buyers' types, this section shows that, in the context of private transaction terms, these restrictions are without loss. I focus on paid platform in this section, but the proof of proposition D.1 also holds (in a more direct way) in the case where the platform cannot set transfers on buyers.

In this model, since buyers' types are drawn independently, and the match is one to many with each buyer having a unique matching seller for which trade is valuable, then the platform can be seen as a designer running many independent bilateral trade interactions. Therefore, this subsection focuses on one arbitrary bilateral trade interaction with one buyer and one seller, to demonstrate the main result.

This section analyses a bilateral trade set up where nor the seller's price and nor the buyer's decision to purchase are not contractible. That is, these decisions cannot be an output of the trade mechanism. Following, Myerson (1982), in a direct mechanism, the designer makes recommendations to players with non-contractible actions. These recommendations are the direct representation of any possible communication strategies by the designer. Therefore, platforms that display descriptive statistics about demand, display predicted probability of trading, or use color schemes to influence sellers' pricing decision, can be viewed as using indirect implementations of price recommendations.

Additionally, this assumption captures the idea that platforms facilitates transaction but do not directly choose the terms of trade. As described in the European law, Online intermediation services "allow business users to offer goods or services to consumers, with a view to facilitating the initiating of direct transactions between those business users and consumers, irrespective of where those transactions are ultimately concluded."[39]

This is arguably an extreme view: Most platform retain some control over the transaction terms. However, in the standard bilateral trade set up the designer has full control over the allocation and the price of the good, which is also excessive. This assumption undershoots the influence of platforms on transactions; but more importantly it makes this influence on the transaction purely a matter of information, which is the focus of this paper.

The analysis of this section also relies on a critical assumption that prevents the platform from using the transfer schemes studied in Crémer and McLean (1988). In this model, the buyer's type and the platform's type are correlated. Hence, the platform can design lotteries of transfers which would induce the buyer to report its type without paying any information rents.

Platforms do not use this type of transfers schemes in practice; however, this paper

[39]Article 2(b), Directive (EU) No 2015/1535

doesn't capture why this is the case. This fact is taken for granted in this analysis.[40]

In this context, the next paragraphs present formally the environment, describe direct mechanisms, and study the properties of the trade mechanisms analyzed in this paper.

Underlying Interaction.

Information structure: As described previously the buyer has a type (v, b) and the platform a type ρ_s.

Action sets: The seller posts a price $\mathbf{A}_s = \{p \in \mathbb{R}\}$.

The buyer chooses whether to buy or not after observing the seller's price $\mathbf{A}_b = \{\sigma_b : \mathbb{R} \times \{v_l, v_h\} \to \{\text{buy, not buy}\}\}$.

The platform is able to shutdown trade or not, and sets transfers to the buyer and seller $\mathbf{A}_p = \{q \in \{0, 1\}, t_b, t_s \in \mathbb{R}\}$.

Mechanism: The platform is a designer in the sense that it controls the choice of a general coordination mechanism. Following Myerson (1982), it is without loss of generality to confine attention to incentive compatible direct mechanism in which the platform's asks the buyer to report his type truthfully and as a function of the reported type and ρ_s recommends (possibly randomly) an action in $\mathbf{A}_b$ to the buyer and in $\mathbf{A}_s$ to the seller and picks (possibly randomly) an action in $\mathbf{A}_p$.

[40] There are many factors that could explain why platforms don't rely on these schemes. First, as these schemes relies on large payments and rewards risk aversion, commitment or limited liabilities issues may diminish their impact and applicability. Second, the exact shape of the transfer schemes relies on the common knowledge assumption of second order beliefs between the platform and the buyer. That is, there must be an exact common understanding on how the correlation in types is for these transfers schemes to work since small changes in terms of the correlation between the types may imply large changes in terms of transfers. That is why, these types of transfers are implicitly ruled out when studying robust mechanism design.

Formally, a direct mechanism is a conditional distribution:

$$
\begin{aligned}
\nu \quad : \quad & \mathbf{\Theta}_b \times \mathbf{\Theta}_p \times \mathcal{A}_b \times \mathcal{A}_s \times \mathcal{A}_p \to [0,1] \\
& (v, b, \rho_s, A_b, A_s, A_p) \mapsto \nu(A_b, A_s, A_p | v, b, \rho_s)
\end{aligned}
$$

Timing:

1. The platform commits to a mechanism ν.
2. Buyer observes b, and both the buyer and seller choose to participate to the mechanism (i.e join the platform).
3. Buyer observes v and reports (v, b).
4. According to ν, actions are recommended to the buyer and the seller, transfers are set, and trade is allowed or shutdown.
5. The seller observes her recommended action and picks a price.
6. The buyer observes his recommended action and the seller price, then decides whether to purchase the good or not.
7. Payoffs are realized.

The main results relies on a critical assumption:

Assumption 1: The transfers described in the mechanism can't depend on the platform type. Formally, for all $\rho_s, \rho'_s, t_b, t_s, v, b$: $\nu(dt_b, dt_s | v, b, \rho_s) = \nu(dt_b, dt_s | v, b, \rho'_s)$.

The role of this assumption is to prevent the platform from using type of transfer schemes presented in Crémer and McLean (1988).Their result applies in this set up as the platform's type and the buyer's type v are correlated. As long as the platform's type distribution F is not δ_{ρ_0}, then the platform can make the expected transfers of a buyer that have reported v_l conditional on v_h arbitrarily high while keeping constant the expected transfers of a buyer having reported v_l conditional on v_l. Hence, by deterring any buyer's types to misreport, the platform extracts the entire surplus at no cost in terms of surplus formation.

Under this set up and the assumption 1 discussed above I can state the main result of this section:

Proposition 7.3. *Consider the underlying interaction of bilateral trade with private terms of transaction. Under assumption 1 price recommendation mechanisms achieve the same set of outcomes as direct mechanisms.*

Proof. The underlying game is played by two player a buyer and a seller, and a designer: the platform.

The buyer has a two dimensional type $v \in \{v_l, v_h\}$ where $v = v_h$ with probability ρ_0, and b independently distributed according to Q in $\mathbb{R}$. The platform has a type corresponding to its belief about the buyer's type $\rho_s = P(v = v_h|\rho_s) \in [0,1]$ distributed according to a cdf F with mean ρ_0.
The buyer's type corresponds both to his valuation for the good and to his belief about the platform's type. Indeed, observing his type of say v_h, the buyer updates his belief about the platform's type as follow:

$$\begin{aligned} P(\rho_s \in B \mid v_h) &= \frac{1}{\rho_0} P(\rho_s \in B,\ v = v_h) \\ &= \frac{1}{\rho_0} \int_B \rho_s dF(\rho_s) \end{aligned}$$

The seller posts a transaction price $\mathbf{A}_s = \{p \in \mathbb{R}\}$, and the buyer chooses whether to buy or not given the transaction price $\mathbf{A}_b = \{\sigma_b \ : \ \mathbb{R} \mapsto \{\text{buy, not buy}\}\}$. The platform's is able to match or not the buyer and seller, and set transfers to both $\mathbf{A}_p = \{q \in \{0,1\}, t_s, t_b \in \mathbb{R}\}$. As a designer, the platform also picks a general coordination mechanism. From Myerson 1982, we focus our attention to direct mechanism in which: (a) the buyer reports his type to the platform (privately). (b) based on the report the platform recommends an action (privately) to the seller and the buyer, as well as defining transfers. (c) both players take an action.

Prior to the game, both the buyer and the seller can choose to participate in the mechanism. If they don't, they collect their outside option payoff of 0.

The buyer's utility is quasilinear in money (for p and t_b) and he obtains $v \in \{v_l, v_h\}$ if he purchases the good.
The seller's profit (gross of transfers) when she sells the good is $p - c$.
The platform's objective corresponds to the money it collects from the transfers.

A mechanism noted ν is therefore a map:

$$\nu \ : \ \{v_l, v_h\} \times \mathbb{R} \times [0,1] \rightarrow \Delta\{\mathbf{A}_s \times \mathbf{A}_b \times \mathbf{A}_p\}$$
$$(v, b, \rho_s) \mapsto \nu(., ., .|v, \rho_s)$$

Where $\nu(dq, dp, d\sigma_b, dt_s, dt_b|v, b, \rho_s)$ is the probability of recommending action p to the seller, action σ_b to the buyer and setting transfers t_s, t_b and the probiability of trade to q.

A mechanism is said to be incentive compatible (IC) if (a) the buyer finds it optimal to report his true type, (b) the seller and the buyer find it optimal to follow the recommended action by the platform. Before proving the proposition, I simplify the analysis with the following observations:

The buyer's decision is made after observing his type and the seller's price; thus the platform have to recommend to the buyer to buy when $v > p$ and not buy when $v < p$. In the threshold case: $p = v$ all recommendations are incentive compatible since the buyer is indifferent, but only the buy recommendation will maximize the platform's profit. I now restrict attention (wlog) to both the platform's recommendation and buyer's actual strategy to be buy when $v \geq p$ and don't buy when $v < p$.

The seller's thus has only two candidates best response as a function of his posterior belief: $p \in \{v_l, v_h\}$. By the same token, for a price recommendation to be incentive compatible it must be that the recommendation lies in $\{v_l, v_h\}$. I use the notation ν_l when the mechanism recommends v_l and ν_h when it recommends v_h.

The platform cannot screen for the buyer's stand-alone valuation since b appears additively Single crossing does not hold (the cross derivative is 0). Thus b is dropped in the mechanism thereafter. Buyer's truthtelling condition are simply on v.

The truthtelling condition of the buyer are:

$$\int_0^1 \int -t_b d\nu(q, \hat{p}, t_b | \rho_s, \hat{v_l}) dF(\rho_s | v_l) \geq \int_0^1 \int -t_b d\nu(q, \hat{p}, t_b | \rho_s, \hat{v_h}) dF(\rho_s | v_l)$$

$$\int_0^1 \int [q(v_h - p)\mathbf{1}_{\{v \geq p\}} - t_b] d\nu(q, \hat{p}, t_b | \rho_s, \hat{v_h}) dF(\rho_s | v_h) \geq \int_0^1 \int [q(v_h - p)\mathbf{1}_{\{v \geq p\}} - t_b] d\nu(q, \hat{p}, t_b | \rho_s, \hat{v_l}) dF(\rho_s | v$$

Recall $dF(\rho_s | v_h) = \frac{\rho_s}{\rho_0} dF(\rho_s)$ and that $dF(\rho_s | v_l) = \frac{1-\rho_s}{1-\rho_0} dF(\rho_s)$. Using assumption 2 (IC_l) becomes:

$$\int -t_b d\nu(q, \hat{p}, t_b | \hat{v_l}) \geq \int -t_b d\nu(q, \hat{p}, t_b | \hat{v_h})$$

Since the buyer's payoff is quasilinear randomizing transfers has no impact. The same is true for the platform's profit. Thus, there is no loss to have transfers only function on the reported type.

Summing both constraints:

$$\int_0^1 \frac{\rho_s}{\rho_0}(v_h - v_l)\nu_l(\{q=1\}|\rho_s, \hat{v_h})dF(\rho_s) \geq \int_0^1 \frac{\rho_s}{\rho_0}(v_h - v_l)\nu_l(\{q=1\}|\rho_s, \hat{v_l})dF(\rho_s)$$
$$\int_0^1 \frac{\rho_s}{\rho_0}\nu_l(\{q=1\}|\rho_s, \hat{v_h})dF(\rho_s) \geq \int_0^1 \frac{\rho_s}{\rho_0}\nu_l(\{q=1\}|\rho_s, \hat{v_l})dF(\rho_s)$$

Randomizing the probability of trade may not be effective, however linking it to the buyer's report is useful for incentive compatibility. To further simplify, I assume (wlog) that when the platform block trade it always recommends a high price to the seller. That is:

$$\nu_l(\{q=1\}|\rho_s, \hat{v}) = \nu_l(\rho_s, \hat{v})$$

The condition becomes:

$$\int_0^1 \frac{\rho_s}{\rho_0}\nu_l(\rho_s, \hat{v_h})dF(\rho_s) \geq \int_0^1 \frac{\rho_s}{\rho_0}\nu_l(\rho_s, \hat{v_l})dF(\rho_s) \qquad (nc)$$

Condition (nc) puts a restriction on the informativeness of the recommendations. In order to show it, view the platform's recommendation as a statistical test with null "the buyer's type is low". The type one error is the probability to recommend a high price when the buyer's type is low, and the type two error the probability of recommending a low price but the buyer's type is high. Errors can be expressed using our notations:

$$\alpha = P(\hat{p} = v_h|v_l) = \frac{1}{1-\rho_0}\int \nu_h(\rho_s, \hat{v_l})(1-\rho_s)dF(\rho_s)$$
$$\beta = P(\hat{p} = v_l|v_h) = \frac{1}{\rho_0}\int_0^1 \nu_l(\rho_s, \hat{v_h})\rho_s dF(\rho_s)$$

Thus:

$$\begin{aligned}\alpha + \beta &= \int_0^1 \left[\frac{1-\rho_s}{1-\rho_0} + \frac{\rho_s}{\rho_0}\nu_l(\rho_s, \hat{v}_h) - \frac{1-\rho_s}{1-\rho_0}\nu_l(\rho_s, \hat{v}_l) \right] dF(\rho_s) \\ &\overset{(nc)}{\geq} \int_0^1 \left[\frac{1-\rho_s}{1-\rho_0} + \left(\frac{\rho_s}{\rho_0} - \frac{1-\rho_s}{1-\rho_0} \right) \nu_l(\rho_s, \hat{v}_l) \right] dF(\rho_s)\end{aligned}$$

Remark that the quantity in the RHS of the inequality is the sum of the type one and type two errors achieved by a mechanism that do not screen the buyer's type and set $\nu_l'(\rho_s) = \nu_l(\rho_s, \hat{v}_l)$. More formally a non screening mechanism defined by: $\nu'(|\rho_s) = \nu(|\rho_s, \hat{v}_l)$ achieves lower type 1 and 2 errors when it recommends prices to the seller. Hence, for any given level of one type of error it can achieve a lower level of error of the other type compared to the screening mechanism. Thus, by Blackwell (1953) the non screening mechanism is more informative for the seller. Therefore by Blackwell (1953) the non screening mechanism can achieve more outcomes than the screening mechanism. $\square$

If the platform does not screen the buyer's type, then transfers don't depend on reported type and thus can be front loaded as entry fees. However, it is not trivial that the platform should not screen the buyer's type, as it would allow the platform to design more precise price recommendation. The key insight of the previous proposition is that the buyer's truthtelling condition restricts the informativeness of the price recommendations so the platform cannot pass information gained on the buyer's side to the seller's side. Therefore, without ex ante private information the platform cannot affect the terms of transaction. Further, screening could allow the platform to increase its profit by price discriminating buyers of different types. However, this is unnecessary since the buyer's participation is ex-ante.

Technical Appendix

Strong Duality

Although, proposition 4.1 is an application of linear programming results, for completeness, I provide a proof in three steps. In the first step, I construct the dual $\mathcal{D}$. In the second step, I recall the weak duality principle, that the complementary slackness conditions are sufficient to characterize the optima. The third step shows that strong duality holds, that is both problems have solutions and have the same values (1.) and that the slackness conditions are necessary conditions for optimality.

Preliminary Facts.

1. The integral is a duality between C and V: $(f,\nu) \mapsto \int f d\nu$ is bilinear and non degenerate on $(C \times V)$.

2. The spaces of $\sigma-$additive measures V and continuous functions C on $[0,1]$ are dual spaces.[41]

3. If B is a cone in C, $B^* = \{\mu \in V \ ; \ \int f d\mu \geq 0 \ \ \forall f \in B\}$ is the dual cone of B. For instance, V_+ is the dual cone of C_+.

4. The uniform norm in C, noted $\|.\|_\infty$ can be used to define a norm in the dual space V:[42]

$$\|\nu\| := \sup_{f \in C} \left\{ \left| \int f d\nu \right| \ ; \ \|f\|_\infty \leq 1 \right\}$$

Step 1: Construction of the dual.

[41] Riezs-Markov-Kakutani representation Theorem.
[42] Which correspond to the weak* topology on V.

To construct the dual problem, I use the Lagrangian a function from $V \times C_+ \times \mathbb{R}_+$ to $\mathbb{R}$ defined by:

$$\mathcal{L} = \int_0^1 (1-z)\mu(dz)(v_l - c) + \int_0^1 \Lambda(z)d(F-\mu)(z) \\ + \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right)$$

Lemma 7.4.

$$(i)\ \mu \leq F \iff \min_{\Lambda \in C_+} \int_0^1 \Lambda(z)d(F-\mu)(z) = 0$$
$$(ii)\ (IC_l)\quad holds \iff \min_{\lambda \in \mathbb{R}_+} \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right) = 0$$

Proof. (i) Necessity.

$$\mu \leq F \iff F - \mu \in V_+ \iff \forall f \in C_+, \int f d(F-\mu) \geq 0$$

Since C_+ is a cone $0 \in C_+$ for which $\int 0 d(F-\mu) = 0$. Therefore:

$$\min_{\Lambda \in C_+} \int_0^1 \Lambda(z)d(F-\mu)(z) = 0$$

(i) Sufficiency.

$$\min_{\Lambda \in C_+} \int_0^1 \Lambda(z)d(F-\mu(z)) = 0 \implies \forall \Lambda \in C_+, \int_0^1 \Lambda(z)d(F-\mu)(z) \geq 0 \iff F-\mu \in V_+$$

(iii) Necessity.

If (IC_l) holds, then the map $\lambda \mapsto \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right)$ is linear and has a weakly positive slope. Thus, the minimum on $\mathbb{R}_+$ is 0 at 0.

(iii) Sufficiency.

$$
\begin{aligned}
&\min_{\lambda \in \mathbb{R}_+} \lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c) \right) = 0 \\
\Longrightarrow\ &\forall \lambda \in \mathbb{R}_+,\ \lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c) \right) \geq 0 \\
\Longleftrightarrow\ &\int_0^1 \mu(dz)(v_l - c) \geq \int_0^1 z\mu(dz)(v_h - c)
\end{aligned}
$$

Thus (IC_l) holds, which ends the proof of lemma .4. □

Using lemma .4:

$$val\{\mathcal{P}\} = \max_{\mu \in V} \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \mathcal{L}$$

Let the dual problem $\mathcal{D}$ be defined as:

$$\mathcal{D}: \qquad \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \max_{\mu \in V} \mathcal{L}$$

The following lemma completes the construction of the dual problem.

Lemma 7.5. *The dual problem $\mathcal{D}$ boils down to:*

$$\min_{\Lambda, \lambda} \int_0^1 \Lambda(z) dF(z)$$

subject to:

$$\forall\ z \in [0,1] \quad (v_l - c)(1 + \lambda) - z[(v_l - c) + \lambda(v_h - c)] \leq \Lambda(z) \qquad (\star)$$

$$\Lambda \in C_+, \quad \lambda \in \mathbb{R}_+$$

Proof.

$$\begin{aligned}\mathcal{L} &= \int_0^1 (1-z)\mu(dz)(v_l - c) + \int_0^1 \Lambda(z) d(F-\mu)(z) \\ &\quad + \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right) \\ &= \int_0^1 \Lambda(z) dF(z) \\ &\quad + \int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz)\end{aligned}$$

The Lagrangian can be written as a Linear function of the control μ. In the dual problem, the Lagrangian is first unconstrained maximized by $\mu \in V_+$, hence if the slope is positive the objective can be increased arbitrarily high. Since then the objective is minimized by the dual variables, it is optimal to set λ, Λ_0 and Λ such that the slope is non positive.

Indeed, assume there is a $z \in [0,1]$ where $(1-z)(v_l - c) - \Lambda(z) + \lambda[v_l - c - z(v_h - c)] > 0$. By choosing:

$$\mu_n(B) = n\delta_z(B) = \begin{cases} 0 & \text{if } z \notin B \\ n & \text{if } z \in B \end{cases}$$

the objective can be made arbitrarily large: $\lim_{n\to\infty} \mathcal{L}(\mu_n) = \infty$.

However, if dual variables λ and Λ are set such that the slope is non positive for all $z \in [0,1]$, then the objective is finite. Therefore, the solution of the dual problem must feature dual variables that makes the slope non positive for all z.

Since the slope is non positive, for all $\mu \in V_+$:

$$\int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz) \leq 0$$

Hence when maximizing by μ, one should pick a μ such that:

$$\int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz) = 0$$

Therefore, the dual problem becomes:

$$\min_{\Lambda,\lambda} \int_0^1 \Lambda(z)dF(z)$$

subject to:

$$\forall\ z \in [0,1] \quad (v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)] \leq \Lambda \qquad (\star)$$

$$\Lambda \in C_+, \quad \lambda \in \mathbb{R}_+$$

Which completes the proof of lemma .5. □

Step 2: Weak Duality.

To make the proof self-contained I provide a proof the weak duality principle in this appendix adapted to this problem.

Lemma 7.6. *Weak Duality*

1. *$val\{\mathcal{D}\} \geq val\{\mathcal{P}\}$.*

2. *Take feasible primal and dual variable. If $\int_0^1 \Lambda(z)dF(z) = \int_0^1 (1-z)\mu(dz)(v_l - c)$, then*

the primal and dual variables are optimal in both problems.

3. *Take feasible primal and dual variable such that:*

$$\begin{cases} \int \Lambda d(F-\mu) = 0 \\ \lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c) \right) = 0 \\ \int_0^1 ((v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)] - \Lambda(z))\mu(dz) = 0 \end{cases}$$

Then the primal and dual variable are optimal in both problems.

Proof. Point 1.

$$\begin{aligned} & \forall \mu, \Lambda \in C_+, \lambda \in \mathbb{R}_+ \ : \ \mathcal{L} \geq \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \mathcal{L} \\ \implies \quad & \forall \Lambda \in C_+, \lambda \in \mathbb{R}_+ \ : \ \max_{\mu \in V} \mathcal{L} \geq \max_{\mu \in V} \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \mathcal{L} \\ \implies \quad & val\{\mathcal{D}\} \geq val\{\mathcal{P}\} \end{aligned}$$

Point 2.

If $\int_0^1 \Lambda(z) dF(z) = \int_0^1 (1-z)\mu(dz)(v_l - c)$, the dual variables Λ, λ are such that (using inequality 2 from point 1.):

$$\forall \Lambda \in C_+, \lambda \in \mathbb{R}_+ \ : \ \max_{\mu \in V} \mathcal{L} \geq \max_{\mu \in V} \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \mathcal{L} = \int_0^1 \Lambda(z) dF(z)$$

therefore, the dual variables are optimal.

Similarly, one can show that (same argument as point 1.):

$$\forall \ \mu \in V \qquad \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \mathcal{L} \leq \min_{\Lambda \in C_+, \lambda \in \mathbb{R}_+} \max_{\mu \in V} \mathcal{L} = \int_0^1 (1-z)\mu(dz)(v_l - c)$$

Thus μ is optimal.

Point 3.

Note that there are two ways of writing the Lagrangian, and using the assumptions one has:

$$\mathcal{L} = \int_0^1 (1-z)\mu(dz)(v_l - c) + \int_0^1 \Lambda(z) d(F-\mu)(z) \\ + \lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c) \right) = \int_0^1 (1-z)\mu(dz)(v_l - c)$$

$$\mathcal{L} = \int_0^1 \Lambda(z) dF(z) \\ + \int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz) = \int_0^1 \Lambda(z) dF(z)$$

That is in this case:$\int_0^1 \Lambda(z)dF(z) = \int_0^1 (1-z)\mu(dz)(v_l - c)$.

Therefore from point 2. both the primal and dual variables are optimal. □

Step 3: Strong Duality.

Lemma 7.7. *1. Both problems have a solution.*

2. $val\{\mathcal{D}\} = val\{\mathcal{P}\}$

Proof. Dworczak and Martini (2019) discuss the possibility of adapting optimal transport results on strong duality to the persuasion literature, which is done in Dizdar and Kováč (2020).

I provide a proof based on the proof sketched p.202 Barvinok (2002) used in optimal transport problems. The main benefit of this type of proof is that it does not rely on a generalize slater condition, and thus it holds for any F. On the contrary, generalized slater condition would imply that F must have full support on $[0,1]$, which is unrealistic in this model.

Strong duality ensures not only that both problems have the same values, but also that optimal solutions are necessarily found using the complementary slackness conditions.

The first step is to write the linear program in its canonical form.

For this proof only lets denote by μ_l, μ_h the joint measure of recommending a low (resp high) prices over signals $[0,1]$. The following linear program:

$$\max_{\mu_l,\mu_h,\beta} \int_0^1 (1-z)d\mu_l(z)(v_l-c)$$

subject to

$$A(\mu_l,\mu_h) = b$$

$$(\mu_l,\mu_h,\beta) \in V_+^2 \times \mathbb{R}_+$$

with , $b = (F,0) \in V \times \mathbb{R}$ and

$$A \ : \ \left((\mu_l,\mu_h) \mapsto \mu_l + \mu_h, \quad (\mu_l,\mu_h,\beta) \mapsto \int_0^1 d\mu_l(z)(v_l-c) - \int_0^1 z d\mu_l(z)(v_h-c) - \beta \right)$$

is the problem $\mathcal{P}$ in canonical form.

The second step consists in applying lemma 7.3 p.171 in Barvinok (2002), in order to show that the conditions to apply theorem 7.2 p.168 in Barvinok (2002) "strong duality theorem" are met.

Define the linear map $\hat{A}(\mu_l, \mu_h, \beta) = (A(\mu_l, \mu_h, \beta),\ \int_0^1 (1-z) d\mu_l(z)(v_l - c))$.

$\hat{A}$ is a continuous linear function because $i)$ the sum is continuous in V and $ii)$ the integral (as a bilinear form) is continuous in its second argument.

The cone V_+^2 as a compact base that is formally:

$$B = \left\{ \nu \in V^2 \ ;\ \nu \geq 0, \quad \int_{[0,1]^2} d\nu = 1 \right\}$$

Thus the cone $V_+^2 \times \mathbb{R}_+$ has a compact base.

Moreover:

$$ker(\hat{A}) = \left\{ (\mu_l, \mu_h, \beta);\ \mu_l = -\mu_h\ ,\ \int_0^1 d\mu_l(z)(v_l - c) - \int_0^1 z d\mu_l(z)(v_h - c) - \beta = 0,\ \int_0^1 (1-z) d\mu_l(z)(v_l -$$

Therefore:

$$ker(\hat{A}) \cap V_+^2 \times \mathbb{R}_+ = \{0\}$$

Because $\mu_l = -\mu_h$ and $\mu_l, \mu_h \geq 0$ implies $\mu_l = \mu_h = 0$, and in turn $\beta = 0$.

Thus by lemma 7.3 p.171 in Barvinok (2002), $\hat{A}(V_+^2 \times \mathbb{R}_+)$ is a closed convex cone.

Therefore, the conditions for theorem 7.2 p.168 in Barvinok (2002) are met and so there is no duality gap.

Which concludes the proof of $\square$

It remains to show that all solutions of the primal and the dual problems must satisfy the complementary slackness conditions.

Assume μ and Λ, λ are solutions (thus are feasible):

$$\begin{aligned}\int_0^1 \Lambda(z)dF(z) &\geq \int_0^1 \Lambda(z)dF(z) + \int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz) \\ &= \int_0^1 (1-z)\mu(dz)(v_l - c) + \int_0^1 \Lambda(z)d(F-\mu)(z) \\ &\quad + \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right)\end{aligned}$$

Because both plans are feasible one has: $\int \Lambda_0 d\mu \geq 0$, $\int \Lambda d\mu \geq 0$, and $\lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c\right.$
Therefore the last equation:

$$\begin{aligned}&\int_0^1 (1-z)\mu(dz)(v_l - c) + \int_0^1 \Lambda(z)d(F-\mu)(z) \\ &+ \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right) \\ &\geq \int_0^1 (1-z)\mu(dz)(v_l - c)\end{aligned}$$

But because both problems have the same value, inequalities above where equalities and in particular:

$$\int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz) = 0$$

$$\int_0^1 \Lambda_0(z)\mu(dz) + \int_0^1 \Lambda(z)d(F-\mu)(z) + \lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right) = 0$$

Therefore, as all quantities are non-negative:

$$\begin{cases} \int \Lambda d(F-\mu) = 0 \\ \lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c) \right) = 0 \\ \int_0^1 ((1-z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z))\mu(dz) = 0 \end{cases}$$

Which completes the proof of step 3 and of proposition 4.1. $\square$

Sensitivity Analysis

Consider the perturbed problem, for a $t \in \mathbb{R}$ and $h \in V$:

$$\mathcal{P}(t): \qquad \max_{\mu \in V_+} \int_0^1 (1-\rho_s) d\mu(\rho_s)(v_l - c)$$

subject to:

$$\int_0^1 \mu(d\rho_s)(v_l - c) \geq \int_0^1 \rho_s \mu(d\rho_s)(v_h - c) \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu(S) \leq (F + th)(S) \qquad (In)$$

Because F has full support, for there is an open interval u of t around 0 such that: $\{\mu : 0 \leq \mu \leq F + th\}$ is non empty. Notice that F and $F + th$ need not to be probability measures for strong duality to holds.

Because for all F $\rho_t > \frac{v_l - c}{v_h - c}$ λ is finite and thus there is an upper bound on λ noted $\overline{\lambda}$. So $0 \leq \Lambda \leq (v_l - c)(1 + \overline{\lambda})$ the dual variables are restricted to a compact set.

Moreover, μ is restricted to a compact set (the space of probability measures), and Λ^* is unique when F has full support.

Additionally:

$$\mathcal{L}(t) = \int_0^1 (1-z)\mu(dz)(v_l - c) + \int_0^1 \Lambda(z) d(F + th - \mu)(z)$$
$$+ \lambda \left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c) \right)$$

And

$$\mathcal{L}_t(t) = \int_0^1 \Lambda(z) dh(z)$$

Are continuous, as linear functions with bounded slopes.

Theorem 5 of Milgrom and Segal (2002) applies and V(t) is differentiable for $t \in u$:

$$V'(t) = \int_0^1 \Lambda(z) dh(z)$$

Therefore, for any $h = H - F$, $V(F)$ is Gateaux differentiable at $t = 0$ and:

$$D_h V(F) = \int_0^1 \Lambda_F(z) dh(z)$$

As Λ_F uniquely characterize all directional derivative for an interior F then Λ_F is the gradient of $V(F)$ and noted $\nabla_F V$.

For any $h = H - F$, the directional derivative of V is:

$$D_{H-F} V(F) = \int_0^1 \nabla_F V(\rho_s) d(H - F)(\rho_s)$$

If H also has full support, then Theorem 5 of Milgrom and Segal (2002) applies and one has:

$$V(H) - V(F) = \int_0^1 \int_0^1 \nabla((1-t)F + tH)(\rho_s)d(H-F)(\rho_s)dt$$

Which completes the proof of proposition 4.1. $\square$

Chapter 3

Price Recommendation and the Value of Data: Competition.

Abstract. I study whether competition between e-commerce platforms reduces the distortion in their incentives to collect data. In a Hotelling model with co-located outside option, I study competitive equilibria between two platforms charging participation fees and using price recommendations. Both platforms disclose strategically their information about buyers' valuations to sellers thereby influencing their sellers' pricing decisions. Platforms gain market shares by lowering the average seller prices, hence increasing buyer expected surplus from trade. Results show that increasing the degree of competition decreases the distortions in the platform's entry fees and incentives to collect data.

Keywords: price recommendations, information design, two-sided markets.

JEL Codes: D82, D83, L21, L81.

1 Introduction

The unprecedented collection of data by large platforms has led to vivid debates about competition and regulation in the digital economy.[1] The debate focuses on two related platforms' practices: the usage of data that enables, for instance, price discrimination, ad-targeting or product personalization; and the collection of data, a critical issue in a growing market for data worth 216 billion dollars in the US and 72.3 billion euros in the EU.[2] This paper examines how competition between e-commerce platforms reduces the distortion in platforms' incentives to collect data, where platforms use data to help sellers price discriminate buyers.

Many e-commerce platforms such as Amazon Marketplace, E-bay or Mercari use data about demand to suggest prices or discounts to their sellers. Consumer Relationship Management firms (CRM) offer devices to sellers that personalize discounts, or target specific consumers. For instance, users of Facebook Ads, Google Ads, or Voucherify can launch personalized coupon campaigns associated to specific queries, search histories, cookies related to buyer characteristics. Empirical evidences shows that recommender systems have a significant impact on sellers prices and sales.[3] The influence of price recommendations on market outcomes generates a demand for data for platforms and there is a question of which market structures foster efficient data collection by platforms.

This paper shows that distortions in platforms' incentives to collect data are mitigated as the degree of competition between platforms increases. Price recommendations are a communication device and do not constrain sellers. They allow platforms to strategically disclose their user data to sellers and influence sellers' pricing decisions.[4] Since the preferred sellers' price may not be the preferred platforms' price[5] platforms benefit from price

[1]See, for instance, recent policy reports by Crémer, Montjoye, and Schweitzer (2019), Scott Morton et al. (2019), Furman et al. (2019).

[2]See the European Data Market Monitoring Tool report page 9.

[3]See e.g. Fleder and Hosanagar (2007) and Pathak et al. (2010).

[4]Price recommendations are the direct representation of any committed communication strategy, see Myerson (1982).

[5]Platforms and sellers benefit differently from each transactions. Furthermore, cross network externalities are not taken into account by small sellers.

recommendations. Consequently, platforms value additional buyer data to improve their recommendations, enhance transactions and, in turn, attract more users. I study how competition between platforms foster a welfare enhancing platform demand for data.

I consider two competing platforms that intermediate trade between buyers and sellers. Platforms charge entry fees to users on both sides of the market. Each platform draws informative signals about buyers' valuations and correlates these signals with price recommendations to influence sellers' pricing decisions. Sellers, that can multi-home, receive a price recommendation when joining a platform and set a price for their good. On the buyer side, platforms are located at both ends of a Hotelling segment. Buyers, that are uniformly distributed over the segment, choose which platform to join, if any, and incur a linear transportation cost to join a platform. As in Bénabou and Tirole (2016), I assume that buyers have two outside options located at both ends of the Hotelling segment. This assumption allows the transportation cost to only determine the degree of competition between firms (i.e. how many market shares a platform gains by reducing the buyer entry fee) and not market participation (the trade off between joining a platform or collecting the outside option payoff). As a result, the transportation cost is identified to the inverse of the degree of competition between platforms as it impacts the demand elasticites within the market but not outside the market. Furthermore, the "co-located" outside options version of the Hotelling model is better suited for welfare analysis than the standard version.[6]

In the first part of the paper, I study the competitive equilibrium in which platforms set user entry fees and a price recommendation rule. In equilibrium, platforms design the price recommendation rule that maximizes the surplus per transaction to attract as many users on both sides, and set entry fees to generate profit and compete for market shares. Compared to the efficient outcome, however, the equilibrium user fees are too high and, as a result, the mass of transactions in equilibrium is inefficiently low. Increasing the degree

[6]Assuming that the market is covered, the welfare analysis in a standard Hotelling model is limited to minimizing the total transportation cost of the economy. See Bénabou and Tirole (2016) that discusses this assumption.

of competition (i.e. reducing the transportation cost) induces platforms charge lower entry fees in equilibrium which increases market participation and welfare.

In the second part of the paper, I capture platforms' incentives to collect data by computing their marginal value for buyer information. Platforms' signals about buyers valuation are the inputs of the price recommendations. I compute the change in platforms' equilibrium profits when changing marginally their distributions of signals. Platforms value additional information as it improves their price recommendations, the surplus per transaction and therefore the mass of users platforms' attract. However, since less users join platforms under a competitive equilibrium than under the efficient outcome, additional information for platforms benefits less transactions. As a result, platforms have a lower marginal value for information than what is socially desirable. Furthermore, I consider a benevolent information provider that maximizes welfare by choosing the platforms' information structures but not their trade mechanisms. I show that conditional on the mass of users joining platforms the benevolent information provider's marginal value for information is larger than the platforms one. The benevolent information provider values the increase in welfare coming from the readjustment of user entry fees by platforms when changing their information structures.[7] Consequently, platforms undervalue additional information compared to what is socially optimal which suggests that platforms under-collect data. However, increasing the degree of competition increases the marginal value of information for platforms. In a more competitive market, each improvement of the price recommendations allows a platform to increase the surplus per transaction from which the platform gains market shares at a higher rate. Therefore, increasing the degree of competition reduces the distortions in platforms' incentives to collect data.

My paper examines how the market structure affects platforms' incentives to collect data. I develop a framework based on: (i) information design which captures how information affect seller prices and market outcomes, and (ii) duality analysis that treats information as

[7] Readjusting entry fees is worth 0 by platforms due to envelop arguments.

an input and computes its marginal value.

Related Literature

This paper relates to the literature on digital economics that studies data and competition, see e.g. De Corniere and Taylor (2020) for a recent presentation of the literature on this topic. Many papers analyze how intermediaries use their information to increase profits (see e.g. Hagiu and Hałaburda (2014), Gomes and Pavan (2019), Bourreau and Gaudin (2018), Jullien and Pavan (2019), Bonatti and Cisternas (2020), and Carroni, Pignataro, and Tampieri (2020)).[8] Bourreau, Caillaud, and De Nijs (2018) and Dimakopoulos and Sudaric (2018) studies competition and data collection for platforms that use data to improve ad-targeting on their users. Pavlov and Berman (2019) compare multiple pricing regimes by an e-commerce platform, including a price recommendation regime, in a cheap-talk environment. In contrast, I use a Bayesian persuasion set up to capture the usage of data by platforms.

From a methodological perspective, my paper relates to the Bayesian persuasion, information design literature (see Kamenica and Gentzkow (2011), Bergemann and Morris (2016) and Taneva (2019)). In my paper price recommendations correspond to the direct implementation of any committed information disclosure policy to sellers. Several other papers use the Bayesian persuasion framework to study data brokers, e.g. Calzolari and Pavan (2006), Bergemann and Bonatti (2015), Bergemann, Bonatti, and Smolin (2018) and Yang (2020). Contrary to a data broker, in my model, platforms engage with both sides, and uses information to balance the allocation of transaction surplus across the two sides.

The duality analysis I provide relates to a recent strand of the information design literature, see e.g. Kolotilin (2018), Galperti and Perego (2018), Dworczak and Kolotilin (2019), Dworczak and Martini (2019), and Dizdar and Kováč (2020). The dual problem is set up in an alternative way to perform a sensitivity analysis on the information structure. I show

[8] For the literature on platforms see Caillaud and Jullien (2003), Armstrong (2006), Rochet and Tirole (2003) and Rochet and Tirole (2006))

that, in the context of my model, the dual variable associated to the informational constraint can be interpreted as the marginal value of information for platforms.

The paper is organized as follows. Section 2 presents the model. Section 3 characterizes the competitive equilibrium and discusses how the degree of competition impacts welfare. Section 4 characterizes platforms' marginal value of information and compares this value with the benevolent information provider's marginal value of information. Section 5 concludes.

2 Model

Environment. Consider two platforms A and B that intermediate trade between buyers and sellers. Platforms set participation fees on both sides of the market and recommend a price to each seller. On the buyer side, platforms are located at both ends of a Hotelling line $[0, 1]$ (A at 0 and B at 1). Buyers are uniformly distributed over this line and face linear transportation cost τ. Buyers choose to join one of the two platforms (single-homing), or collect one of the two outside option payoffs located at both ends of the line, at 0 and at 1.[9] Sellers view platforms A and B as identical ex-ante and can join either platform or both (multi-homing), or collect their outside option payoff normalized to 0. Sellers produce a good at marginal cost c which is valued at either v_l or v_h by their matching buyer, with $v_l < v_h$, and I assume that $c < v_l$ so that trade is always efficient. For the buyers and sellers that decide to join a platform, transactions unfold as follows: buyers inspect sellers' goods displayed on that platform to find their matching seller[10] as well as their valuation for its good (v_l or v_h). Following a successful match, sellers receive a price recommendation and then set a price. Buyers observe their matching seller's price and then decide to purchase the good or not.

[9]The buyer side is represented with a Hotelling model with co-located outside option, see e.g. Bénabou and Tirole (2016).

[10]Matching is one to many: For each buyer there is only one valuable seller.

Information Structure. The buyers' types (v, b, x) have three independent dimensions. The value of the matching seller's good is either high (v_h), with probability ρ_0, or low (v_l), with probability $1 - \rho_0$. The stand-alone valuation b and the buyer location x on the Hotelling segment together capture the utility the buyer derives from joining a platform. The stand-alone valuation b, common to both platforms, captures benefits of additional services provided on both platforms or costs (e.g. advertising nuisance, privacy costs or opportunity costs). Hence, b can be either positive or negative and is distributed according to a continuously differentiable and log-concave distribution Q_b supported on an interval in $\mathbb{R}$. The buyer's location $x \in [0, 1]$ captures his relative preference from joining one platform instead of the other. Assuming trade happens, the utility of a buyer joining platform A is $v - p + b - \tau x - t_{b,A}$ and of a buyer joining platform B is $v - p + b - \tau(1 - x) - t_{b,B}$, where p is the seller's price and $t_{b,A}$ (resp. $t_{b,B}$) is the buyer entry fee set by platform A (resp. B). Buyers learn v after joining the platform, but learn their stand-alone valuation b and location x before joining.

For each joining buyer, each platform receives informative signals about his valuation, from which it forms posterior beliefs about whether the buyer is of type v_h or v_l. Think for instance of both platforms having access to consumer-level data such as histories of past transactions, location or cookies and the related means of recording browsing data. Platforms may observe some of the buyers' characteristics and updates their beliefs about whether he is a low or high type.

Without loss of generality, each platform's information structure is represented as the distribution of its posterior beliefs. For each buyer joining, platforms draw signals $\rho_s = P(v = v_h | \rho_s)$ that are normalized to the posterior belief that the buyer is a high type. I assume that each platform draws posterior beliefs ρ_s according to the same distribution F with mean ρ_0.[11]

[11] Whether draws are correlated across platforms has no impact in the game. F has mean ρ_0 to be consistent with the prior distribution of buyers' types, see for instance Kamenica and Gentzkow (2011)

Sellers can multi-home and view A or B as ex-ante identical. All sellers produce a homogeneous good at a constant marginal cost c. To join a platform sellers incur a fixed cost κ (think of a listing cost for instance) distributed according to a continuously differentiable log-concave distribution Q_s supported in an interval of $\mathbb{R}$. A seller that joins platform $i \in \{A, B\}$ and trades one good at price p obtains profit: $p - c - \kappa - t_{s,i}$.

Timing and Decisions. First, each platform $i \in \{A, B\}$ sets entry fees $t_{s,i}$ and $t_{b,i}$ and commits to a price recommendation rule, which I describe later. Second, buyers observe their stand-alone valuation b and location x. Then, sellers decide whether to join or not each platform. Buyers decide which platform to join if any, and find their matching seller. Third, each platform draws a signal ρ_s for each buyer and recommends a price to each matched seller. Fourth, buyers observe v, their seller price and decide whether to buy.

Solution Concept. The analysis focuses on perfect Bayesian equilibria, where players hold rational expectations, are risk neutral and expected-payoff maximizers.

Price Recommendation Rules and Entry Fees. Platform A (resp. B) sets entry fees $t_{b,A}$ (resp. $t_{b,B}$) on the buyer side and $t_{s,A}$ (resp. $t_{s,B}$) on the seller side, $t_{b,A}, t_{s,A} \in \mathbb{R}$. Platform $i \in \{A, B\}$ commits to a price recommendation rule that maps signals ρ_s with a private price recommendation to sellers. Instead of capturing the price recommendation rule as the probability of recommending a price conditional on signals, it is more convenient to define it with $\mu_i(S)$ the joint probability of recommending a low price v_l and receiving

proposition 1.

signals in S:[12]

$$\mu_i \; : \; \mathcal{B}[0,1] \to [0,1]$$
$$S \mapsto \mu_i(S) = \int_S P(\text{recom } v_l \mid \rho_s) dF(\rho_s).$$

With complementary probability platform i recommends v_h to sellers. Since buyers are either of type v_l or of type v_h, there are only two potentially optimal pricing strategies. Sellers either sell to both buyer types at price v_l, leaving the low type without any surplus, or sell only to the high types at price v_h leaving them with no surplus. Therefore, considering price recommendation rules that only recommend to sellers these two prices v_l or v_h is without loss of generality.

A price recommendation rule μ_i for $k \in \{A, B\}$ is feasible for an information structure F if it satisfies the informational constraint (In):

$$\forall S \in \mathcal{B}[0,1] : \qquad \mu_i(S) \leq \int_S dF(\rho_s) = F(S). \qquad (In)$$

(In) captures the ability of the platform to match price recommendations to the true buyers' valuations. If $\mu_i(S) = F(S)$ platform i always recommends a low price when the signal falls in the set S. However, if $F(S) = 0$ for some S the platform cannot make recommendations as it does not receive signals in this range. Therefore, the informational constraint captures the relationship between the quality of the platform information structure F and the precision of the price recommendations.

3 Platform Competition

This section presents the demand of each platform on both side of the market and characterizes the competitive equilibrium. Then, this section discusses the interaction between the

[12]S is a measurable set in $\mathcal{B}[0,1]$, the Borel $\sigma-$algebra on $[0,1]$.

degree of competition $\frac{1}{\tau}$ and welfare.

Demand on the Seller Side

A seller can multi-home and her decision to join platform A can be treated independently from her decision to join platform B.

Consider a seller that has already joined platform $i \in \{A, B\}$. If a transaction occurs, she trades at a low price yielding $(v_l - c)$ profit, if a low price is recommended, that is, with probability $\int_0^1 \mu_i(d\rho_s)$; and she trades at a high price yielding $(v_h - c)$ profit, if a high price is recommended and the buyer is a high type, that is with probability:

$$P(\text{recom } v_h,\ v = v_h) = \rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s).$$

A seller's expected profit per transaction on platform i equals:

$$\int_0^1 \mu_i(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s)\right](v_h - c).$$

Let D_i be the mass of buyers joining platform i. Consequently, a seller trades on platform i with probability D_i. There is a marginal seller $\tilde{\kappa}_i$ that is indifferent between joining platform i or not, such that all sellers with listing cost κ lower than $\tilde{\kappa}_i$ join the platform and all sellers with listing cost higher than $\tilde{\kappa}_i$ do not. The mass of sellers joining platform i is given by $Q_s(\tilde{\kappa})$ and $\tilde{\kappa}_i$ is characterized by:

$$\tilde{\kappa}_i = D_i \left(\int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c) - c \right) - t_{s,i}. \tag{3.1}$$

To increase its demand on the seller side, platform i can use multiple instruments. It can reduce the seller entry fee, increase the mass of buyers joining, or recommending prices in a way that benefits sellers.

Demand on the Buyer Side

A buyer that joins a platform values trade only if the platform recommends a low price and he is of a high type. Let V_i denotes the buyer's expected value from trade on platform $k \in \{A, B\}$:

$$V_i = Q_s(\tilde{\kappa}_i)(v_h - v_l) \int_0^1 \rho_s \mu_i(d\rho_s).$$

The price recommendation rule μ is the joint probability of recommending a low price v_l and receiving a signal in a measurable set S. Integrating the signal ρ_s, the probability that the buyer is of type v_h, over its support $[0, 1]$ gives the joint probability of recommending a low price to a high type buyer: $\int_0^1 \rho_s d\mu(\rho_s)$. In this case, the buyer obtains a transaction surplus of $v_h - v_l$. Finally, a buyer finds its matching seller with probability $Q_s(\tilde{\kappa}_i)$.

There is a buyer located at $\tilde{x} \in [0, 1]$ that is indifferent between joining platform A or platform B:

$$\tilde{x} = \frac{1}{2} + \frac{V_A - t_{b,A} - (V_B - t_{b,B})}{2\tau}.$$

If the value of joining platform i, $V_i - t_{b,i}$ increases, then the indifferent consumer is located further away from i, that is, platform i gains market shares. The demand of firm i corresponds to buyers located closer to i than $\tilde{x}$ with a stand alone valuation b high enough such that joining i generates more value than both outside options.

Outside options are located at 0 and 1 on the Hotelling segment. A buyer located at $x < \tilde{x}$ joins firm A if neither of the outside options yields more value than A:

$$b \geq \max\{t_{b,A} - V_A, \tau(2x - 1) + t_{b,A} - V_A\}.$$

Symmetrically, a buyer located $x < \tilde{x}$ joins firm A if neither of the outside options yields more value than A:

$$b \geq \max\{t_{b,B} - V_B, \tau(1-2x) + t_{b,B} - V_B\}.$$

There is a consumer indifferent between joining i and collecting the outside option payoff at i's location denoted $\tilde{b}_i$ such that:

$$\tilde{b}_i = t_{b,i} - V_i. \tag{3.2}$$

Using this notation, the masses of buyers joining platforms A and B are respectively given by:

$$D_A(\tilde{b}_A, \tilde{b}_B) = \begin{cases} \left(1 - Q_b(\tilde{b}_A)\right)\left(\frac{1}{2} + \frac{\tilde{b}_B - \tilde{b}_A}{2\tau}\right) & \text{if } \tilde{x} \leq \frac{1}{2} \\ \frac{1}{2}\left(1 - Q_b(\tilde{b}_A)\right) + \int_{\frac{1}{2}}^{\tilde{x}} (1 - Q_b(\tau(2x-1) + \tilde{b}_A))dx & \text{if } \tilde{x} \geq \frac{1}{2}. \end{cases}$$

$$D_B(\tilde{b}_B, \tilde{b}_A) = \begin{cases} \left(1 - Q_b(\tilde{b}_B)\right)\left(\frac{1}{2} + \frac{\tilde{b}_A - \tilde{b}_B}{2\tau}\right) & \text{if } \tilde{x} \geq \frac{1}{2} \\ \frac{1}{2}\left(1 - Q_b(\tilde{b}_B)\right) + \int_{\tilde{x}}^{\frac{1}{2}} (1 - Q_b(\tau(1-2x) + \tilde{b}_B))dx & \text{if } \tilde{x} \leq \frac{1}{2}. \end{cases}$$

If the indifferent consumer $\tilde{x}$ is further away from platform A's location than from B's location, then consumers located between $\frac{1}{2}$ and $\tilde{x}$ arbitrage between joining A or collecting the outside option payoff at B's location. These consumers join platform A if $x \leq \tilde{x}$ and $b \geq \tau(2x-1) + \tilde{b}_A$. Consumers located at $x \in [0, \frac{1}{2}]$ arbitrage between joining A or collecting the outside option payoff at A's location, and, therefore join A if $b \geq \tilde{b}_A$.

Platform's Best Response Problem

Platform i generates profit from entry fees. The buyer entry fee affects directly the mass of buyers joining by changing their net utility from joining and affects indirectly the mass of sellers joining by changing the seller's probability of trading. In addition platform i's price recommendations affect the mass of users joining on each side. Platform i may design price recommendations to help sellers price discriminate buyers, increase their profit, and, therefore, increase the mass of sellers joining. Alternatively, platform i may design price recommendations to sway sellers to price lower on average, which increases the buyer expected surplus from transactions, and, therefore, increases the mass of buyer joining.

Platform i designs an incentive compatible price recommendation rule. That is, sellers must find it optimal to follow platform i's price recommendations:

$$\int_0^1 \mu_i(d\rho_s)(v_l - c) \geq \int_0^1 \rho_s \mu_i(d\rho_s)(v_h - c). \tag{IC_l}$$

$$\left[\rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s)\right](v_h - c) \geq \left[1 - \int_0^1 \mu_i(d\rho_s)\right](v_l - c). \tag{IC_h}$$

Sellers have four options: follow both price recommendations, disobey both price recommendations, or disobey one recommendation and follow the other. If sellers disobey one recommendation and follow the other, they effectively follow a uniform pricing strategy. If disobeying both recommendations is optimal, then one of the uniform pricing strategies is better than following both recommendations.[13] Ensuring that obeying both recommendations yields more profit than both uniform pricing strategies is equivalent to incentive compatibility:

$$\int_0^1 \mu_i(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s)\right](v_h - c) \geq \max\{v_l - c, \rho_0(v_h - c)\}.$$

Throughout the paper, I assume than $\rho_0 > \frac{v_l - c}{v_h - c}$, which implies that at their prior belief sellers prefer setting a high price.[14] Consequently incentive compatibility is reduced to a

[13]See Kolotilin et al. (2017).

[14]This assumption focuses on the case where platform's preferred price and the seller's preferred price differ. If instead $\rho_0 \leq \frac{v_l - c}{v_h - c}$ always recommending a low price (regardless of the signal) is incentive compatible and maximizes the transaction surplus. In this case there is no value for a platform to collect additional information.

single condition:

$$\int_0^1 \mu_i(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s)\right](v_h - c) \geq \rho_0(v_h - c). \tag{IC}$$

Given platform j's strategy, platform i chooses the buyer entry fee, seller entry fee and an incentive compatible price recommendation rule that maximizes its profit. Platform i's profit is generated entirely from the entry fees, $D_i t_{b,i}$ on the buyer side and $Q_s t_{s,i}$. I assume that, for each pair of entry fees $(t_{b,i}, t_{s,i})$ there is a unique pair of marginal users $(\tilde{b}_i, \tilde{\kappa}_i)$ that satisfy equations (1) and (2). As a result, I express the platform's best response problem in terms of the marginal buyer $\tilde{b}_i$ and and marginal seller $\tilde{\kappa}_i$ which are determined the entry fees.

$$\max_{\tilde{b}_i, \tilde{\kappa}_i, \mu_i} \Pi_i = D_i(\tilde{b}_i, \tilde{b}_j)\tilde{b}_i - Q_s(\tilde{\kappa}_i)\tilde{\kappa}_i + D_i(\tilde{b}_i, \tilde{b}_j)Q_s(\tilde{\kappa}_i) \underbrace{\left[\int_0^1 (1-\rho_s)\mu_i(d\rho_s)(v_l - c) + \rho_0(v_h - c)\right]}_{\text{Expected Transaction Surplus}}$$

subject to:

$$\int_0^1 \mu_i(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s)\right](v_h - c) \geq \rho_0(v_h - c) \qquad (IC)$$
$$\forall S \in \mathcal{B}[0,1]: \qquad \mu_i(S) \leq \int_S dF(\rho_s) \qquad (In)$$

From posted entry fees, platform i captures each side's non-trade related benefit from joining at its marginal users' levels and the entire surplus from trade. The platform is able to extract the buyer and seller surplus from each transaction because buyers and sellers decide to join the platform based on the ex-ante value of trade. Sellers do not know the recommended price before joining and buyers learn their valuation after joining. With probability ρ_0 the seller is of a high type and trade occurs regardless of the recommended price generating a surplus of $v_h - c$. Trade happens with a low type buyer, generating a surplus of $v_l - c$, only if a low price has been recommended, that is with joint probability $\int_0^1 (1-z)\mu_i(dz)$. The mass of transactions on platform i under one-to-many matching is $D_i(\tilde{b}_i, \tilde{b}_j)Q_s(\tilde{\kappa}_i)$.

Platform i's choice of buyer and seller entry fee interacts with platform j's choice of buyer entry fee. However, platform i's optimal choice of the price recommendation rule is independent from j's decisions. Regardless of the mass of users joining platform i, it chooses the price recommendation rule that maximizes the surplus from each transaction. The next subsection characterize the dominant choice of price recommendation rule. The one following it characterizes the competitive equilibrium between platform A and B.

Optimal Price Recommendation Rules

The price recommendation rule that maximizes the surplus per transaction is a dominant strategy for both platforms. If the buyer is of a high type, transaction occurs regardless of the price recommendation, which generates a surplus of $v_h - c$. If the buyer is of a low type, trade occurs only if a low price is recommended, which generates a surplus of $v_l - c$. In other words, platform i maximizes the probability of correctly recommending a low price to low type buyers subject to the incentive compatibility and informational conditions:

$$\max_{\mu_i} \quad \int_0^1 (1-z)\mu_i(dz)(v_l - c) + \rho_0(v_h - c)$$

subject to:

$$\int_0^1 \mu_i(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu_i(d\rho_s)\right](v_h - c) \geq \rho_0(v_h - c). \qquad (IC)$$

$$\forall S \in \mathcal{B}[0,1]: \qquad \mu_i(S) \leq \int_S dF(\rho_s) \qquad (In)$$

The optimal price recommendation rule can be deduced from interpreting recommendations as the result of a hypothesis testing. Let "recommending v_l" correspond to accepting H_0 and let "recommending v_h" correspond to rejecting H_0. The platform then chooses as a function of its test statistic ρ_s whether to reject H_0. In this formulation the incentives are captured by type I and II errors. Sellers want to minimize both types of errors: they want to match their price with their buyer's type. Platform i wants to avoid type I errors, that is, avoid recommending a high price if the buyer's valuation is low because this reduces trade efficiency.

Platform i and sellers agree that the type I error must be as small as possible. Therefore, the optimal price recommendation rule must have the following property: for any level of the type II error chosen, the level of the type I error must be minimized.

This property is common in statistics and econometrics and shared by all standard hypothesis tests. In this case, tests that satisfy this property follow cutoff rules. Platform i chooses a cutoff $\rho_t \in [0, 1]$ such that for all signals ρ_s above the cutoff it recommends a high price and for all the signals below the cutoff it recommends a low price.

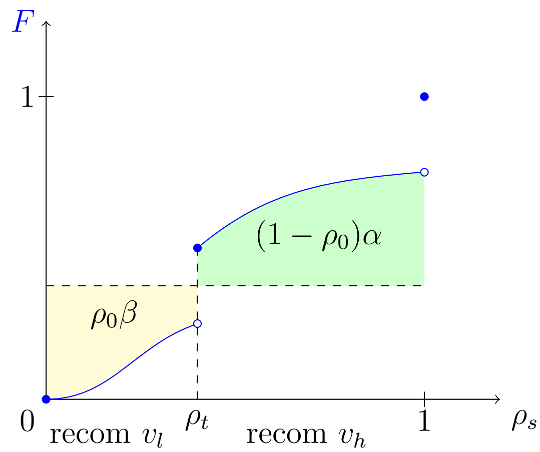


Figure 3.1: Cutoff rule.

The optimal cutoff maximizes surplus from trade. The ideal cutoff is $\rho_t = 1$ as in this case platform i always recommends a low price hence trade is efficient. However, this price recommendation rule provides no information to sellers about their matched buyer's type. As a result, the ideal cutoff $\rho_t = 1$ is not incentive compatible. The optimal cutoff must be low enough so that low price recommendations are sufficiently informative about the buyer's type and followed by sellers. However, to reduce trade inefficiencies platform i aims to increase the cutoff. Therefore, the optimal price recommendation rule uses the highest cutoff such that the incentive compatibility condition binds.

Proposition 3.1. *The dominant price recommendation rule takes the form of a* ***cutoff***

rule. *Each platform chooses the cutoff* $\rho_t^* \in [0,1]$ *and then recommends prices as follows:*

(i) *For all* $\rho_s < \rho_t,$ *the platform recommends a low price.*

The platform sets for all measurable $S \subset [0, \rho_t), \quad \mu(S) = F(S).$

(ii) *For all* $\rho_s > \rho_t,$ *The platform recommends a high price.*

The platform sets for all measurable $S \subset (\rho_t, 1], \quad \mu(S) = 0.$

(iii) *If* F *has a mass point at* $\rho_t,$ *the platform randomizes recommendations at the cutoff.*

The platform picks $\mu(\{\rho_t\}) \in [0, dF(\rho_t)],$ *with* $dF(\rho_t) = F(\rho_t) - \lim_{x \uparrow \rho_t} F(x).$

The optimal cutoff $\rho_t^* \in [0,1]$ *is the highest that binds the incentive compatibility condition:*

$$F(\rho_t^*)(v_l - c) = \int_0^{\rho_t^*} \rho_s dF(\rho_s)(v_h - c)$$

Proof. This is a standard result in the Bayesian Persuasion literature; see e.g. Kamenica and Gentzkow (2011) and Dworczak (2020). This specific lemma is proved in Lefez (2021). For completeness, a proof is given in appendix A.2. □

Figure 1 pictures a cutoff rule by a platform in the case where F has a mass point at the cutoff. On the figure it mixes equally the recommendation at the cutoff ρ_t. The shaded areas represent the associated type I and II errors. In the main text, equations are presented assuming no mass point at the cutoff, while the appendix presents the general case.

Competitive Equilibrium

Both platforms choose the dominant price recommendation rule with cutoff ρ_t^* and simultaneously choose their marginal sellers and marginal buyers joining via the entry fees. Given platform j's choice of entry fees, platform i 's optimal choice of entry fees trades off the extra profit generated on both sides by attracting one more users with the loss of profit on infra-marginal users. I impose a lower bound on τ for the existence of the symmetric equilibrium.[15] If the transportation cost is too small, only one platform serves the market in equilibrium.

Proposition 3.2. *The unique equilibrium* $(\tilde{b}^*, \tilde{\kappa}^*, \rho_t^*)$ *is symmetric. Each platform sets the dominant cutoff rule* ρ_t^* *and sets entry fees to pin down marginal users* $(\tilde{b}^*, \tilde{\kappa}^*)$ *such that:*

$$\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left[\frac{1}{2}(1-Q_b(\tilde{b}^*))\left(\int_0^{\rho_t^*}(1-\rho_s)dF(\rho_s)(v_l-c)+\rho_0(v_h-c)\right)-\tilde{\kappa}^*\right]=Q_s(\tilde{\kappa}^*) \tag{3.3}$$

$$\left(\frac{1-Q_b(\tilde{b}^*)}{\tau}+\frac{\partial Q_b}{\partial \tilde{b}_i}\right)\left[Q_s\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)+\tilde{b}^*\right]=1-Q_b(\tilde{b}^*) \tag{3.4}$$

If τ *increases the entry fees for both platform on both sides decreases. Formally* $\tilde{b}^*$ *decrease with* τ *and* $\tilde{\kappa}^*$ *increases with* τ*.*

Proof. See appendices. □

To set the buyer entry fee, platform i equates the gain of attracting additional buyers on both sides with the loss of profit on the infra-marginal buyers. Reducing the entry fee by 1 attracts $\frac{1}{2}\frac{\partial Q_b}{\partial \tilde{b}_i}$ additional buyers that were opting for the outside option, and $\frac{1-Q_b(\tilde{b}^*)}{2\tau}$ additional buyers that were choosing to join platform j. Reducing the transportation cost τ, increases the market shares gained from reducing the buyer entry fee and therefore reduces the buyer entry fee in equilibrium. This assertion is the focus of the next subsection.

To build some intuition, consider the two extreme cases, $\tau = +\infty$ and $\tau = 0$ (although in the last case the equilibrium does not exist). If $\tau = +\infty$, (3) and (4) characterizes the

[15] See appendices.

optimal entry fees for a monopoly platform. On the other hand, if $\tau = 0$, (4) boils down to:

$$Q_s \left(\int_0^{\rho_t^*} (1-\rho_s) dF\rho_s)(v_l - c) + \rho_0 (v_h - c) \right) + \tilde{b}^* = 0$$

That is, the platform reduces the buyer entry fee until the benefit generated on both sides from attracting an additional buyer goes to 0.

Competition and Welfare

In this paper, I restrict attention to symmetric outcomes. From a welfare perspective, there is a trade-off between serving the market with a monopoly platform which minimizes sellers' fixed cost κ, or serving the market with a duopoly which minimizes buyers' transportation costs. Given a symmetric allocation of the market, total welfare equals:

$$\begin{aligned} W(\tau) = & \int_{\tilde{b}} b dQ_b(b) - 2 \int^{\tilde{\kappa}} \kappa dQ_s(\kappa) - \frac{\tau}{4} \\ & + (1 - Q_b(\tilde{b})) Q_s(\tilde{\kappa}) \left[\int_0^1 (1-\rho_s) \mu(d\rho_s)(v_l - c) + \rho_0 (v_h - c) \right] \end{aligned}$$

The symmetric allocation minimizes the transportation costs in the economy that are equal to $\frac{\tau}{4}$ but doubles the listing costs for all sellers that are multi-homing, i.e. $2 \int^{\tilde{\kappa}} \kappa dQ_s(\kappa)$. In terms of welfare, users are valued at their average type level, and not at the marginal type level, which corresponds to the typical distortion of a posted price mechanism. Equilibrium entry fees are higher than the efficient entry fees. To maximize welfare entry fees are reduced until the benefit on both sides of attracting an additional user equals the cost of attracting this user.

As the degree of competition increases, the amount market shares gained by reducing the buyer entry fee increases, whereas the loss of profit on infra marginal consumers $(1 - Q(\tilde{b}))$ stays constant. As a result, the equilibrium buyer entry fee decreases. Additionally, the benefit of attracting an additional seller for platforms increases (as more buyers join), and therefore the seller entry fee decreases as well. All in all, as the degree of competition increases, the user entry fees gets closer to their efficient level and welfare increases.

Platforms value transaction the same way a social planner values transactions, as platforms extract the entire surplus of transactions. Consequently, platforms' optimal price recommendation rule is efficient.

Proposition 3.3.
1. *The platform price recommendation rule maximizes welfare for all degree of competition $\frac{1}{\tau}$.*
2. *Welfare increases in the degree of competition.*

Proof. See appendices. □

Although platforms' use of data (the price recommendation rule) is efficient, platforms' incentives to collect data is not. The next section shows that a platform's marginal value of information is lower than the social marginal value of information, and that increasing the degree of competition reduces this distortion.

4 Competition and the Value of Information

This section characterizes a platform's marginal value of information and the interaction between this value and the degree of competition τ. A platform's marginal value of information captures the change in the platform's equilibrium profit for a marginal change in its information structure.

Definition 7. *The platform's value of information $V(F)$ maps each information structure into the equilibrium profit.*

Definition 8. *The platform's marginal value of information is defined as the gradient of the platform's value of information: $\nabla_F V$.*

Studying the platform's marginal value of information provides rich comparative statics on the platform's willingnesses to pay to acquire additional information. This section shows the willingness to pay for information increases with the degree of competition. First, I provide an economic interpretation of a marginal change in F.

A Marginal Change in the Information Structure

Consider platform i purchasing a new data set G of "size" q. The size q captures the probability that a buyer is in the data set G. In other words, with probability $1-q$, the buyer is not in the data set G, and the platform draws a posterior belief according to F. With probability q the buyer is in the data set, and the platform draws a posterior belief ρ_s according to F and re-update this beliefs according to the data set $G(.|\rho_s)$ to ρ_s^u.[16]

$$\begin{tikzcd}
 & \rho_s \text{ dist. } F \text{ with mean } \rho_0. & \\
\rho_0 \arrow[ur, "1-q"] \arrow[dr, "q"'] & & \\
 & \rho_s \text{ dist. } F \arrow[r] & \rho_s^u \text{ dist. } G(.|\rho_s) \text{ with mean } \rho_s.
\end{tikzcd}$$

The re-updating process $G(.|\rho_s)$ depends on the first signal ρ_s which captures the fact that the new information contained in G may be correlated with the information the platform already has. The value of a data set G of size q is noted $\phi_G(q)$. Formally:

$$\phi_G(q) = V\left((1-q)F + q\int_0^1 G(.|\rho_s)dF(\rho_s)\right)$$

Consider the platform increasing the size q of the data set G by ϵ, that is the platform considers purchasing a slightly larger data set from the data broker. This corresponds to a marginal change in the platform's information structure of:

$$\epsilon\left(\int_0^1 G(.|\rho_s)dF(\rho_s) - F(.)\right)$$

[16]For all ρ_s $\int_0^1 G(dz|\rho_s) = \rho_s$ for the re-updating to be consistent with Bayes rule.

The platform observes one extra buyer in the data set, updates its belief about this buyer's valuation and potentially recommends different prices. The platform's value of information varies by:

$$\phi'_G(q) = \lim_{\epsilon \to 0} \frac{\phi_G(q+\epsilon)}{|\epsilon|}$$

Relation with Blackwell informativeness. This example of a platform purchasing a new data set G captures all variations that increase the platform's information in the sense of Blackwell. A platform information structure is more informative if F varies in directions of H such that H is a mean preserving spread of F. For marginal variation that means for a small $\epsilon > 0$ the platform new information structure is $F + \epsilon(H - F)$. In fact, any mean preserving spread H or F can be constructed from a re-updating process G:

Lemma 4.1. *H is a mean preserving spread of F if and only if there is a G with $\int_0^1 zG(dz|\rho_s) = \rho_s$ for all ρ_s such that $H(S) = \int_0^1 G(S|\rho_s)dF(\rho_s)$.*

Proof. See Le Cam (1996). □

In other words, any marginal variation that increases the platform's information is captured by ϕ'_G for some data set G. The remainder of this section characterizes ϕ'_G, shows it increases in the degree of competition but is lower than the efficient level.

Platform's Value for a Data Set

The platform's marginal value is constructed as the shadow price associated to the informational constraint (In): The impact on the platform's profit when relaxing this constraint. Formally the shadow price of (In) is the Lagrangian multiplier, and corresponds to the gradient of V the marginal value of information. I show that V is directionally differentiable at F, and that directional derivatives are computed from the gradient $\nabla_F V$.

I focus on marginal variations in directions in which a platform gains information in the sense of Blackwell. All these variations are interpreted as the platform marginally increasing the size a data set G.

Proposition 4.2. *Assume that F has full support.*

1. *V is directionally differentiable at F in any directions.*

2. *The marginal value of increase the size of a data set G equals:*

$$\phi'_G(q) = \int_0^1 \int_0^1 \nabla_{F+qG} V(\rho_s^u) dG(\rho_s^u|\rho_s) dF(\rho_s) - \int_0^1 \nabla_{F+qG} V(\rho_s) dF(\rho_s).$$

3. *The value of acquiring a data set G of size q equals:*

$$\phi_G(q) - \phi_G(0) = \int_0^q \left(\int_0^1 \int_0^1 \nabla_{F+\epsilon G} V(\rho_s^u) dG(\rho_s^u|\rho_s) dF(\rho_s) - \int_0^1 \nabla_{F+\epsilon G} V(\rho_s) dF(\rho_s) \right) d\epsilon.$$

Proof. See technical appendix. □

Proposition 4.2 shows that the increase in the platform's profit from increasing the size of the data set G is constructed from the gradient. In addition, the difference in profit between having a data set G of size q or not equals all the marginal gain in profit from increasing the size of G from 0 to q.

Appendix C computes the gradient of the platform's profit by computing the dual variable associated to the informational (In) condition. Using this value the platform's willingness to pay to increase the size of a data set G can be computed from the parameters of the model:

Proposition 4.3. *Platform i's willingness to pay to increase the size of the data set G, ϕ'_G*

equals:

$$\Bigg[\underbrace{(v_l - c)\left(\int_{\rho_t^*}^1 (1-\rho_s)dF(\rho_s) - \int_0^1 \int_{\rho_t^*}^1 (1-\rho_s^u)dG(\rho_s^u|\rho_s)dF(\rho_s)\right)}_{\text{Gain in trade efficiency}}$$
$$+ \underbrace{\lambda^*\left(\int_0^1 \left((v_l - c)G(\rho_t|\rho_s) - (v_h - c)\int_0^{\rho_t^*} \rho_s^u dG(\rho_s^u|\rho_s)\right) dF(\rho_s)\right)}_{\text{Gain in relaxing incentive compatibility}}\Bigg]$$
$$\times Q_s(\tilde{\kappa}^*)\frac{1 - Q_b(\tilde{b}^*)}{2}$$

Where λ is the dual variable associated to the sellers' (IC):

$$\lambda = \frac{(1-\rho_t^*)(v_l - c)}{\rho_t^*(v_h - c) - (v_l - c)}.$$

Proof. See appendices. □

Trade efficiency increases as the probability of recommending a high price when the buyer has a low value decreases, which is the probability that the low type buyers' signal falls above the cutoff after re-updating the platform's belief with G. The gain in transaction surplus is first captured by sellers, then captured by the platform via increasing the seller entry fee by the same amount their profits increase. Since the surplus from trade has changed, the platform adjusts the buyer entry fee as well, but this has no effect on profit due to standard envelop arguments.

Increasing the size of G also relaxes the sellers' incentive compatibility condition (IC), which allows the platform to increase the cutoff ρ_t. By increasing the cutoff the platform now recommends a low price around ρ_t which increases the surplus by $(1 - \rho_t)(v_h - c)$ but decreases the sellers' profit (hence tightens the sellers' (IC)) by $\rho_t(v_h - c) - (v_l - c)$. The ratio of the two is equal to the dual variable associated to the sellers' (IC) that gives the increase in the platforms' profit when the sellers' (IC) the relaxed by one unit.

All in all, a platform's profit increases from increasing the size of G via two channels: (i) the change in the surplus from trade and (ii) the change in the sellers' (IC). Next, I compare a platform's marginal value of information with the benevolent information provider's marginal value of information. Then, I show that increasing the degree of competition between platforms reduces the distortions in their marginal value of information.

Benevolent Information Provider

The difference between the value of increasing the size of the data set G for platforms or for a social planner is a matter of the total mass of trades on each platform. A social planner would attract more users, generate more trade, and, therefore, benefit more from additional information. However, this subsection shows that conditional on the mass of users joining, a platform's willingness to pay to increase the size of the data set is inefficient.

Consider a "benevolent information provider" that chooses the information structure for both platforms but does not choose their trade mechanisms. Compared to platforms, the benevolent information provider values increasing the size of the data set from a third extra channel: it values the gain in surplus coming from the platform's readjustment of the users' entry fee when increasing the size of G (these gains are negligible for platforms as entry fees are profit maximizing).

Proposition 4.4. *The benevolent information provider's willingness to pay to increase the size of the data set G is higher than a platform's willingness to pay by a multiplicative factor equal to:*

$$1 - \frac{\partial \tilde{b}^*}{\partial q}\left(\frac{1}{Q_s} - \frac{1}{\tau}\left(\int_0^{\rho_t^*}(1-\rho_s)dF(\rho_s)(v_l - c) + \rho_0(v_h - c)\right)\right) + 2\frac{\partial \tilde{\kappa}^*}{\partial q}\frac{1}{1-Q_b}$$

where λ is the dual variable associated to the sellers' (IC):

$$\lambda = \frac{(1-\rho_t^*)(v_l - c)}{\rho_t^*(v_h - c) - (v_l - c)}.$$

Proof. See appendices. □

Compared to platforms, the benevolent information provider values the readjustment of user entry fees. Each seller gains $\frac{1}{1-Q_b}$ (per transaction) from the reduction in both platforms' entry fee, and each buyer gains $\frac{1}{Q_s}$ (per transaction) from the reduction in the buyer entry fee, minus a term proportional to $\frac{1}{\tau}$. Indeed, platform's buyer entry fee is closer to the social optimum as the degree of competition increases. In turn, the difference between the benevolent information provider's and the platform's willingness pays decreases as the degree of competition increases.

The benevolent information provider's willingness to pay for information is proportional to the platform one. Consequently, they rank data sets in the same way. Precisely, consider two data sets G_1 and G_2. Then platforms value more increasing the size of G_1 compared to G_2 if and only if the benevolent information provider values more increasing the size of G_1 compared to G_2.

Proposition 4.5. *1. A platform's willingness to pay to increase the size of a data set is closer to the benevolent information provider's one as the degree of competition $\frac{1}{\tau}$ increases.*

2. The benevolent information provider and platforms' ranking of data sets coincide.

Proof. See appendices. □

Although platform data usage is efficient, platforms' incentives to collect data are not. However, increasing the degree of competition reduces the distortions in platform marginal value of information. Under a more competitive regime, platforms gain more market shares from reducing user entry fees. As a result, the equilibrium entry fees are lower and a larger mass of users join platforms. Therefore, additional information affects a larger mass of transactions which increases its value.

5 Conclusion

There is a vivid debate about the regulation of data the digital economy. There are two potential issues: whether data is misused by platforms or whether it is miss-collected. This paper examines the impact of the degree of competition on platforms' marginal value of information. Platforms use information to recommend prices to sellers and value additional information to improve their recommendations and attract more users. I use a Bayesian persuasion framework to capture how user information influence seller prices and a duality analysis to capture information as an input and compute its marginal value. In the context of my model, I show that platforms under value additional information compared to what is socially desirable. However, I show that increasing the degree of competition between platforms decreases this distortion.

This paper warns against potentially distorted platforms' incentives to collect data. Since my analysis relates how platforms use data with platforms' incentives to collect data, it provides multiple entry points for regulation. Regulation can be implemented on the market for data (facilitate access to data, making some data public, etc...) but it can also be implemented on the downstream market. For instance, promoting competition between platforms reduces the distortions in their incentives to collect data. Alternatively, in this model a social planner can subsidize user participation to increase the mass of users on platforms and, in turn, increase platforms marginal value for information.

6 Appendices

6.1 Platform Competition

Existence and Uniqueness of Competitive Equilibrium

Let $((\tilde{b}_A^*, \tilde{\kappa}_A^*), (\tilde{b}_B^*, \tilde{\kappa}_B^*))$ constitute a NE. Both firms are in best response and thus for A and B:

$$\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left[D_i\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)-\tilde{\kappa}_i\right]=Q_s(\tilde{\kappa}_i^*)$$
$$-\frac{\partial D_i}{\partial \tilde{b}_i}\left[Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)+\tilde{b}_i\right]=D_i(\tilde{b}_i^*, \tilde{b}_j^*)$$

Note that the profit function is sub-modular:

$$\frac{\partial^2 \Pi_i}{\partial \tilde{b}_i \partial \tilde{\kappa}_i}=\frac{\partial D_i}{\partial \tilde{b}_i}\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)<0$$

In other words, the larger the mass of sellers joining, the higher the gain by attracting new buyers. That is, users are complementary inputs for platform i. Therefore, from standard comparative static results, the optimal $\tilde{\kappa}_i$ decreases with respect to $\tilde{b}_i$.

I show that $(\tilde{b}_i^*, \tilde{\kappa}_i^*) = (\tilde{b}_j^*, \tilde{\kappa}_j^*)$. Fix j's strategy.

Assume that $\tilde{b}_i^* > \tilde{b}_j^*$. This implies that $D_i < D_j$. Consider i's profit's variation in the marginal seller:

$$\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left[D_i\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)-\tilde{\kappa}_i\right]-Q_s(\tilde{\kappa}_i^*)$$

From log concavity this expression is decreasing[17] in $\tilde{\kappa}_i$. Since, $D_i < D_j$ the derivative hits 0 before j's derivative, therefore $\tilde{\kappa}_i^* < \tilde{\kappa}_j^*$ and so $Q_s(\tilde{\kappa}_i^*) < Q_s(\tilde{\kappa}_j^*)$.

In this case, $|x_i - \tilde{x}| \leq \frac{1}{2}$, where x_i is the location of firm i, that is the indifferent consumer is closer to i than to j. The derivative on i's profit with respect to $\tilde{b}_i$ is:

$$\begin{aligned}
&\frac{\partial D_i}{\partial \tilde{b}_i}\left[Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i\right] + D_i(\tilde{b}_i, \tilde{b}_j) \\
&= -\left(\frac{1}{2\tau}(1 - Q_b(\tilde{b}_i)) + \frac{\partial Q_b}{\partial \tilde{b}_i}\left[\frac{1}{2} + \frac{\tilde{b}_j - \tilde{b}_i}{2\tau}\right]\right)\left[Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i\right] \\
&\quad + (1 - Q_b(\tilde{b}_i))\left(\frac{1}{2} + \frac{\tilde{b}_j - \tilde{b}_i}{2\tau}\right)
\end{aligned}$$

The derivative equals:

$$\begin{aligned}
&(1 - Q_b)\left[\frac{1}{2} + \frac{\tilde{b}_j - \tilde{b}_i}{2\tau}\right]\left[Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i\right] \\
&\left(-\frac{\frac{\partial Q_b}{\partial \tilde{b}_i}}{1 - Q_b} - \frac{1}{\tau + \tilde{b}_j - \tilde{b}_i} + \frac{1}{Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i}\right)
\end{aligned}$$

Assumption: Q_s, Q_b and τ are such that:

$$\begin{aligned}
&Q_s''\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s(v_l - c) + \rho_0(v_h - c)\right) \\
&< \left[2Q_s' - \frac{Q_b'}{1 - Q_b} - \frac{1}{\tau}\right]
\end{aligned}$$

This implies that $Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i$ is increasing in b_i along the path $\kappa^*(b_i)$ defined by (3). If τ is too small, only one platform serves both sides in equilibrium. Given this assumption, the RHS of the derivative is decreasing in $\tilde{b}_i$ (by log-concavity) along the path $\tilde{\kappa^*}(\tilde{b}_i)$. Consider plugging $\tilde{b}_i = \tilde{b}_j^*$, the derivative is 0, and negative for all $\tilde{b}_i > \tilde{b}_j^*$. Therefore, there cannot be an equilibrium with $\tilde{b}_i^* > \tilde{b}_j^*$.

[17] See **an1997log**

Consequently, the equilibrium must be symmetric. To show it exists, I check that for $\tilde{b}_i < \tilde{b}_j^*$ i's profit increases by increasing b_i.

Consider now candidates equilibrium where $\tilde{b}_A^* < \tilde{b}_B^*$. Using the same argument $\tilde{\kappa}_A^* > \tilde{\kappa}_B^*$ and so $Q_s(\tilde{\kappa}_A^*) > Q_s(\tilde{\kappa}_B^*)$. In that case, $\tilde{x} > \frac{1}{2}$. Recall in this case:

$$\begin{aligned}
D_A(\tilde{b}_A, \tilde{b}_B) &= \frac{1}{2}(1 - Q_b(\tilde{b}_A)) + \int_{\frac{1}{2}}^{\tilde{x}} (1 - Q_b(\tau(2x-1) + \tilde{b_A}))dx \\
\frac{\partial D_A}{\partial \tilde{b}_A} &= -\frac{1}{2}\frac{\partial Q_b}{\partial \tilde{b}_A}(\tilde{b}_A) - \frac{1}{2\tau}(1 - Q_b(\tau(2\tilde{x}-1) + \tilde{b}_A)) - \int_{\frac{1}{2}}^{\tilde{x}} \frac{\partial Q_b}{\partial \tilde{b}_A}(\tau(2x-1) + \tilde{b}_A)dx \\
&= -\frac{1}{2}\frac{\partial Q_b}{\partial \tilde{b}_A}(\tilde{b}_A) - \frac{1}{2\tau}(1 - Q_b(\tilde{b}_B)) - \frac{1}{2\tau}(Q_b(\tilde{b}_B) - Q_b(\tilde{b}_A)) \\
&= -\frac{1}{2}\frac{\partial Q_b}{\partial \tilde{b}_A}(\tilde{b}_A) - \frac{1}{2\tau}(1 - Q_b(\tilde{b}_A))
\end{aligned}$$

The derivative on A's profit with respect to $\tilde{b}_A$ is:

$$\begin{aligned}
&\left(-\frac{1}{2}\frac{\partial Q_b}{\partial \tilde{b}_A}(\tilde{b}_A) - \frac{1}{2\tau}(1 - Q_b(\tilde{b}_A))\right)\left[Q_s(\tilde{\kappa}_A)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i\right] + \frac{1}{2}(1 - Q_b(\tilde{b} \\
&+ \int_{\frac{1}{2}}^{\tilde{x}} (1 - Q_b(\tau(2x-1) + \tilde{b_A}))dx \\
=&\left(-\frac{1}{2}\frac{\partial Q_b}{\partial \tilde{b}_A}(\tilde{b}_A) - \frac{1}{2\tau}(1 - Q_b(\tilde{b}_A))\right)\left[Q_s(\tilde{\kappa}_A)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i\right] + D_i(\tilde{b}_i, \tilde{b}_j) \\
=&\frac{1}{2}(1 - Q_b(\tilde{b}_A)\left[Q_s(\tilde{\kappa}_A)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i\right] \\
&\times\left(-\frac{Q_b'}{1 - Q_b} - \frac{1}{\tau} + \frac{1 + 2\int_{\frac{1}{2}}^{\tilde{x}} \frac{(1 - Q_b(\tau(2x-1) + \tilde{b_A}))}{(1 - Q_b(\tilde{b}_A))}dx}{Q_s(\tilde{\kappa}_A)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}_i}\right)
\end{aligned}$$

Note that $\int_{\frac{1}{2}}^{\tilde{x}} \frac{(1 - Q_b(\tau(2x-1) + \tilde{b_A}))}{(1 - Q_b(\tilde{b}_A))}dx$ is decreasing in $\tilde{b}_A$ as log concavity of Q implies that the primitive of Q is log concave[18], together with the assumption the right part of the derivative is decreasing in $\tilde{b}_A$. Since the expression in 0 for $\tilde{b}_A = \tilde{b}_B^*$, this implies the FOC is positive

[18] See **an1997log**

for all $\tilde{b}_A < \tilde{b}_B^*$ and thus A must set $\tilde{b}_A = \tilde{b}_B^*$. □

Symmetric Equilibrium

Consider the system of equations that characterizes best responses:

$$\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left[D_i\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)-\tilde{\kappa}_i\right]=Q_s(\tilde{\kappa}_i^*)$$
$$-\frac{\partial D_i}{\partial \tilde{b}_i}\left[Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)+\tilde{b}_i\right]=D_i(\tilde{b}_i^*,\tilde{b}_j^*)$$

In a symmetric equilibrium one has:

$$\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left[\frac{1}{2}(1-Q_b(\tilde{b}_i))\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)-\tilde{\kappa}_i\right]=Q_s(\tilde{\kappa}_i^*)$$
$$\left(\frac{1}{\tau}(1-Q_b(\tilde{b}_i))+\frac{\partial Q_b}{\partial \tilde{b}_i}\right)\left[Q_s(\tilde{\kappa}_i)\left(\int_0^{\rho_t^*}(1-\rho_s)dF\rho_s)(v_l-c)+\rho_0(v_h-c)\right)+\tilde{b}_i\right]=(1-Q_b(\tilde{b}_i))$$

Price Recommendation Rule

Optimal Price Recommendation Rule

The optimal price recommendation rule problem is a linear program:

$$\mathcal{P}: \qquad \max_{\mu_i\in V_+}\int_0^1(1-\rho_s)d\mu_i(\rho_s)(v_l-c)$$

subject to:

$$\int_0^1 \mu_i(d\rho_s)(v_l-c)+\left[\rho_0-\int_0^1\rho_s\mu_i(d\rho_s)\right](v_h-c)\geq\rho_0(v_h-c) \qquad (IC)$$
$$\forall S\in\mathcal{B}[0,1]: \qquad \mu_i(S)\leq\int_S dF(\rho_s) \qquad (In)$$

Let C_+ refers to the set of non-negative continuous functions defined on $[0,1]$, and V_+ the set of non-negative measures defined on $[0,1]$. The dual writes:

$$\mathcal{D}: \qquad \min_{\Lambda \in C_+,\ \lambda \in \mathbb{R}_+} \int_0^1 \Lambda(\rho_s) dF(\rho_s)$$

subject to:

$$\forall\ \rho_s \in [0,1] \quad (v_l - c)(1+\lambda) - \rho_s[(v_l - c) + \lambda(v_h - c)] \leq \Lambda(\rho_s) \qquad (\star)$$

Technical appendix 4.1 describes the construction of the dual as well as proof that strong duality holds:

Proposition 6.1. *Problems $\mathcal{P}$ and $\mathcal{D}$ are strong duals:*

1. *Both problems have a solution and $val\{\mathcal{P}\} = val\{\mathcal{D}\}$.*

2. *Let μ and Λ, λ be feasible then:*

μ is and optimal solution of $\mathcal{P}$ and (Λ, λ) is an optimal solution of $\mathcal{D}$ if and only if

$$\begin{cases} \int \Lambda d(F - \mu) = 0 & (C1) \\ \lambda \left(\int_0^1 d\mu(\rho_s)(v_l - c) - \int_0^1 \rho_s d\mu(\rho_s)(v_h - c) \right) = 0 & (C2) \\ \int_0^1 \Lambda(\rho_s) - (v_l - c)(1+\lambda) + \rho_s[(v_l - c) + \lambda(v_h - c)] d\mu(\rho_s) = 0 & (C3) \end{cases}$$

Proof. See technical appendix 4.1. □

Optimality of the cutoff rule

This subsection proves proposition 3.1. On top of the complementary slackness conditions, the solution must satisfy primal feasibility:

$$\begin{cases} \mu \in V_+ & \\ \mu \leq F & (In) \\ \int_0^1 \mu(d\rho_s)(v_l - c) + \left[\rho_0 - \int_0^1 \rho_s \mu(d\rho_s)\right](v_h - c) \geq \rho_0(v_h - c) & (IC) \end{cases}$$

As well as dual feasibility:

$$\begin{cases} \Lambda_0 \in C_+ & \\ \lambda \in \mathbb{R}_+ & \\ \forall\ \rho_s \in [0,1]\ \ (v_l - c)(1+\lambda) - \rho_s[(v_l - c) + \lambda(v_h - c)] \leq \Lambda(\rho_s) & (\star) \end{cases}$$

The LHS of equation $(\star)$ is an affine function of z. Since $\lambda \in \mathbb{R}_+$ this affine function starts positive at 0 and ends up non-positive at 1. It crosses the x axis at:

$$\rho_t \stackrel{def}{=} \frac{(v_l - c)(1+\lambda)}{v_l - c + (v_h - c)\lambda} \in \left(\frac{v_l - c}{v_h - c}, 1\right]$$

Because Λ is a non negative map, for all $z \in (\rho_t, 1]$ $(\star)$ is slack:

$$(1 - z)(v_l - c) + \lambda[v_l - c - z(v_h - c)] - \Lambda(z) < 0$$

Thus, using (C3):

$$\text{For all measurable } B \subset (\rho_t, 1]\ \ \mu(B) = 0$$

Hence, from (C1):

$$\forall\ z\ \in supp(F) \cap (\rho_t, 1], \quad \Lambda(z) = 0$$

Because Λ is continuous:

$$\forall\ z\ \in \overline{supp(F) \cap (\rho_t, 1]}, \quad \Lambda(z) = 0$$

Second, the LHS of $(\star)$ is strictly positive for $z \in [0, \rho_t)$. Thus $(\star)$ implies that for all $z \in [0, \rho_t), \quad \Lambda(z) > 0$.

So, using this in (C1):

$$\forall \text{ measurable } \ B \subset supp(F) \cap [0, \rho_t), \quad \mu(B) = F(B)$$

But on $B \subset supp(F)^c \cap [0, \rho_t)$, primal feasibility implies $\mu(B) = 0 = F(B)$. Therefore:

$$\forall \text{ measurable } \ B \subset [0, \rho_t), \quad \mu(B) = F(B)$$

By the third complementary slackness condition one has again:

$$\forall \text{ measurable } \ B \subset \overline{supp(F) \cap [0, \rho_t)}, \quad \Lambda = (1 - z)(v_l - c) + \lambda[v_l - c - z(v_h - c)]$$

To sum up we have so far:

$$\forall \text{ measurable } B :$$

$$B \subset [0, \rho_t) \quad \mu(B) = F(B)$$

$$B \subset (\rho_t, 1] \quad \mu(B) = 0$$

Additionally, from the $\sigma-$additivity property of measures, μ is pinned down up to the choice of mass at $\{\rho_t\}$.

That is, receiving a signal below ρ_t always riggers a low price recommendation, and above ρ_t always leads to a high price recommendation. The platform can also mix recommendation at the cutoff.

The analysis has also established the relationship between the threshold and the value of dual variables:

$$\rho_t \stackrel{def}{=} \frac{(v_l - c)(1+\lambda)}{v_l - c + (v_h - c)\lambda} \in \left(\frac{v_l - c}{v_h - c}, 1\right]$$

Equivalently:

$$\lambda = \frac{(1-\rho_t)(v_l - c)}{\rho_t(v_h - c) - (v_l - c)} \in \mathbb{R}_+$$

In addition, it has determined the value of the dual variable Λ on the closure of the support of F. To perform the sensitivity analysis, Λ is chosen outside the support to be continuous to small perturbations of F, if a perturbed F had vanishingly small mass on the entire interval then:

$$\Lambda(z) = \begin{cases} (v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)] & \text{if } \rho_s \leq \rho_t \\ 0 & \text{if } \rho_s \geq \rho_t \end{cases}$$

The complementary slackness condition (C2) is associated to (IC):

$$\lambda\left(\int_0^1 \mu(dz)(v_l - c) - \int_0^1 z\mu(dz)(v_h - c)\right) = 0$$

Case 1: Assume (IC) is slack at the solution, from $(C2)$: $\lambda = 0$.

Using the formula for ρ_t that implies $\rho_t = 1$. (IC) is indeed slack with $\rho_t = 1$ if there is a $\mu(\{1\}) \in [0, dF(1)]$ such that:

$$\int_0^1 dF(z)(v_l - c) - \int_0^1 z dF(z)(v_h - c) - (dF(1) - \mu(\{1\}))(v_l - v_h) > 0$$
$$\iff \frac{v_l - c}{v_h - c} + \frac{dF(1) - \mu(\{1\})}{v_h - c}(v_h - v_l) > \rho_0$$

As $\frac{v_l - c}{v_h - c} < \rho_0$, if F doesn't have a mass point at 1 the previous inequality cannot hold.

However, F may have a mass point at 1, for instance a fully informed platform has a distribution of posterior F with a mass point at 1 of size ρ_0.

If the platform has full information the inequality boils down to:

$$(1 - \rho_0)(v_l - c) - \mu(\{1\})(v_h - v_l) > 0$$
$$\iff \mu(\{1\}) < (1 - \rho_0)\frac{v_l - c}{v_h - v_l}$$

Together with $\mu(\{1\}) \geq 0$ that corresponds to an interval of solutions.

In a case of an arbitrary mass point such solutions are feasible if the mass point is large enough formally:

$$dF(1) > \frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}$$

$\mu(\{1\})$ can be optimally picked in the interval $\left[0, dF(1) - \frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}\right)$.

All these solutions are optimal because they are all efficient: when the platform recommends a high price the buyer has a high valuation with probability one. But we can also consider only the one that binds (IC) by choosing:

$$\rho_t = 1$$
$$\mu(\{1\}) = dF(1) - \frac{\rho_0(v_h - c) - (v_l - c)}{v_h - v_l}$$

Case 2: Assume (IC) binds at the solution.

In this case we can compute precisely λ from (IC) and the formula on ρ_t:

$$\int_{[0,\rho_t)} v_l - c - z(v_h - c)dF(z) + \mu(\{\rho_t\})(v_l - c - \rho_t(v_h - c)) = 0$$

Because $\rho_t \in (\frac{v_l-c}{v_h-c}, 1]$, $\int_{[0,\rho_t)} v_l - c - z(v_h - c)dF(z)$ is strictly positive for ρ_t close to $\frac{v_l-c}{v_h-c}$ and strictly decreasing in ρ_t. In addition, because $\frac{v_l-c}{v_h-c} < \rho_0$ it is strictly negative at $\rho_t = 1$.

Therefore it changes sign only once, but it need not to be continuous as F may have mass points.

However, by (In): $\mu(\{\rho_t\}) \in [0, dF(\rho_t)]$. If F has a mass point at ρ_t, there is a unique $\mu(\{\rho_t\})$ that binds (IC), and if F doesn't have mass point at ρ_t then $\mu(\{\rho_t\}) = 0$, and there is a unique ρ_t which binds (IC).

In both scenarios there exist a unique pair $(\rho_t,\mu(\{\rho_t\}))$ that satisfies IC with equality. In turn, λ is determined by ρ_t.

Which completes the proof of proposition 3.1 □

Platform's value for Data

The dual variable from the dual problem is:

$$\Lambda(z) = \begin{cases} (v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)] & \text{if } \rho_s \leq \rho_t \\ 0 & \text{if } \rho_s \geq \rho_t \end{cases}$$

This problem was ignoring constant terms, in particular using strong duality the value of the problem (with the constant terms) is:

$$Val(\mathcal{P}) = (1 - Q(\tilde{b}))\left(\int_0^1 \max\{(v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)], 0\} + \int_0^1 \rho_s dF(\rho_s) + \tilde{b}\right)$$

Thus the gradient is:

$$\nabla_F V(\rho_s) = Q_s(\tilde{\kappa}^*)\frac{(1 - Q(\tilde{b}^*))}{2}\left(\rho_s(v_h - c) + \max\{(v_l - c)(1+\lambda) - z[(v_l - c) + \lambda(v_h - c)], 0\}\right)$$

In the previous subsections, solving the platforms' problem yields the following dual variables:

$$(1 - Q(\tilde{b}))\Lambda(\rho_s) = \nabla_F V_p(\rho_s) = Q_s(\tilde{\kappa}^*)\frac{(1 - Q(\tilde{b}^*))}{2}\begin{cases} (v_l - c)(1+\lambda) - \rho_s[(v_l - c) + \lambda(v_h - c)] & \text{if } \rho_s \leq \rho_t \\ 0 & \text{if } \rho_s \geq \rho_t \end{cases}$$

And:

$$\rho_t \overset{def}{=} \frac{(v_l - c)(1+\lambda)}{v_l - c + (v_h - c)\lambda} \in \left(\frac{v_l - c}{v_h - c}, 1\right]$$

Equivalently:

$$\lambda = \frac{(1 - \rho_t)(v_l - c)}{\rho_t(v_h - c) - (v_l - c)} \in \mathbb{R}_+$$

When ρ_t is interior ($\lambda > 0$) and using the formula relating λ with ρ_t one has:

$$\nabla_F V_p(\rho_s) = Q_s(\tilde{\kappa}^*)\frac{(1-Q(\tilde{b}^*))}{2}\left(\rho_s(v_h - c) + \frac{(v_h - v_l)(v_l - c)}{(v_h - c)\rho_t - (v_l - c)}\max\{\rho_t - \rho_s, 0\}\right)$$

□

6.2 Value of Data

Proof of Proposition 4.3

The change in the platform's profit in a direction $H - F$ is given by:

$$Q_s(\tilde{\kappa}^*)\frac{(1-Q(\tilde{b}^*))}{2}\int_0^1 \nabla_F V(\rho_s) d(H-F)(\rho_s)$$

For clarity lets compute it per buyer (dividing by $Q_s(\tilde{\kappa}^*)\frac{(1-Q(\tilde{b}^*))}{2}$:

$$\int_0^{\rho_t}(v_l - c)(1-\rho_s)d(H-F)(\rho_s) + \lambda\int_0^{\rho_t}((v_l - c) - \rho_s(v_h - c))d(H-F)(\rho_s)$$

Because H and F have a total mass of 1 and a mean of ρ_0:

$$\int_0^{\rho_t}(1-\rho_s)d(H-F)(\rho_s) = -\int_{\rho_t}^1(1-\rho_s)d(H-F)(\rho_s)$$

Using this, and (C3):

$$\int_{\rho_t}^1((v_l - c)(1-\rho_s)d(F-H)(\rho_s) + \lambda\int_0^{\rho_t}((v_l - c) - \rho_s(v_h - c))dH(\rho_s) + (1-\rho_t)(v_l - c)[dF(\rho_t) - \mu(\{\rho_t$$
$$=(v_l - c)\left(\left(\int_{[\rho_t,1]}(1-\rho_s)dF(\rho_s) - \mu(\{\rho_t\})\right) - \int_{\rho_t}^1(1-\rho_s)dH(\rho_s)\right) + \lambda\int_0^{\rho_t}((v_l - c) - \rho_s(v_h - c))dH(\rho_s)$$

Replacing H with $\int_0^1 G(.|\rho_s)dF(\rho_s)$ provides the formula of proposition 4.3 □

Benevolent Information provider value for data set G.

The benevolent information provider maximizes welfare given the competitive equilibrium:

$$\int_{\tilde{b}} b dQ_b(b) - 2\int^{\tilde{\kappa}} \kappa dQ_s(\kappa) - \frac{\tau}{4}$$
$$+ (1 - Q_b(\tilde{b}))Q_s(\tilde{\kappa}) \left[\int_0^1 (1-\rho_s)\mu(d\rho_s)(v_l - c) + \rho_0(v_h - c)\right]$$

The competitive equilibrium is given by

$$\frac{\partial Q_s}{\partial \tilde{\kappa}_i}\left[\frac{1}{2}(1 - Q_b(\tilde{b}^*))\left(\int_0^{\rho_t^*}(1-\rho_s)dF(\rho_s)(v_l - c) + \rho_0(v_h - c)\right) - \tilde{\kappa}^*\right] = Q_s(\tilde{\kappa}^*)$$
$$\left(\frac{1 - Q_b(\tilde{b}^*)}{\tau} + \frac{\partial Q_b}{\partial \tilde{b}_i}\right)\left[Q_s\left(\int_0^{\rho_t^*}(1-\rho_s)dF(\rho_s)(v_l - c) + \rho_0(v_h - c)\right) + \tilde{b}^*\right] = 1 - Q_b(\tilde{b}^*)$$

Denote by $h = \left(\int_0^{\rho_t^*}(1-\rho_s)dF(\rho_s)(v_l - c) + \rho_0(v_h - c)\right)$ the transaction surplus, and by dh the change in the transaction surplus that comes from a change in the information structure. The equilibrium marginal users vary by:

$$\frac{d\tilde{\kappa}^*}{dh}\left[Q_s''\left(\frac{1-Q_b}{2}h - \tilde{\kappa}^*\right) - 2Q_s'\right]\frac{2}{Q_s'(1-Q_b)} = \frac{d\tilde{b}^*}{dh}\frac{Q_b'}{1-Q_b}$$
$$\frac{d\tilde{b}^*}{dh}\frac{\left(-\frac{Q_b'}{\tau} + Q_b''\right)\left[Q_s h + \tilde{b}^*\right] + \frac{1-Q_b}{\tau} + 2Q_b'}{\left(\frac{1-Q_b}{\tau} + Q_b'\right)Q_s} = -\frac{d\tilde{\kappa}^*}{dh}\frac{Q_s'}{Q_s}$$