DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Essays on Consumer Search“

verfasst von / submitted by

Anton Sobolev

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 370 140
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Doctor of Philosophy-Doktoratsstudium Wirtschaftswissenschaften: Volkswirtschaftslehre (Economics)
Betreut von / Supervisor:	Univ.-Prof. Dr. Maarten Janssen
Mitbetreut von / Co-Supervisor:	Univ.-Prof. Dipl.-Math. Karl Schlag, PhD

Contents

Acknowledgements

First of all, I would like to express sincere gratitude to my main advisor Maarten Janssen for his exceptional advice and support. I was very lucky to meet Maarten during my master studies in Moscow where he introduced me to modern economic research. I am deeply indebted to Maarten for his encouragement, patience and tremendous effort that helped to develop my economic thinking.

I am extremely grateful to my secondary advisor Karl Schlag who continues teaching me how to think out of the box. His insightful and innovative look at various economic problems significantly contributed to my own view on economic research.

A special thank you to my master advisor, my co-author and my good friend Alexei Parakhonyak. It was a pleasure to work together on a project that resulted in the first published paper in my life. I am very happy that we have many new ideas for future research and this number is only growing.

I would like to acknowledge several other people who contributed to my work by providing valuable comments and suggestions. I am very grateful to Yongmin Chen, Daniel Garcia, Eeva Mauring, Philipp Schmidt-Dengler, Yossi Spiegel, Matan Tsur, Chris Wilson and my colleagues at the Vienna Graduate School of Economics.

Financial support from the University of Vienna and the Austrian Academy of Sciences (project number: FA 374005) is gratefully acknowledged.

1 Introduction

Traditionally, economics has studied the functioning of frictionless markets, where consumers can obtain prices and all relevant product information at no cost. While this approach was satisfactory initially, many markets clearly did not fit these strong assumptions. As a result, consumer search literature emerged to provide the rationale of information frictions on consumers' side of the market. It models consumers who do not know prices and product characteristics and have to incur a search cost to find them out. Consumer search theory was greatly appreciated by economists for its explanation of consumer search behavior as well as the extent to which the firms' pricing is affected by imperfect information of consumer. This led to an increasingly sophisticated understanding of information frictions, with high relevance for firm strategy and public policy. In the seminal paper on consumer search, George Stigler (1961) was the first who showed the importance of information that consumers have when they make their purchases. After almost 40 years of continuous research in this area, in 2010 Peter Diamond, Dale Mortensen, and Christopher A. Pissarides were awarded the Nobel Prize for their analysis of markets with search frictions.

The theory has now passed the initial stage of development and it is time to reflect on common assumptions and inquire about new applications. Central to the consumer search literature is the idea that consumers do not know the products that are sold in the market and the prices charged and engage in costly search. However, other aspects of the market are considered to be known: the number of firms, their costs, etc. Similarly, firms are assumed to know each others' costs and consumers' search strategies. This dissertation includes three related projects that extend consumer search theory by relaxing some of the informational assumptions made and by inquiring into the incentives of firms to share information.

In the first chapter of my dissertation, I explore the incentives of firms to share their private information in search markets. The second and the third chapters considerably relax the informational assumptions in consumer search theory by exploring the robust search strategies that perform well for any model specifications. Precisely, the second chapter explores an oligopolistic model of competition in which consumers search without priors. The third chapter inquires into the question of how to search and learn without relying on the initial information.

1.1 Information Exchange and Consumer Search

In the first chapter, I explore the impact of information sharing on social welfare in search markets and inquire into the incentives of firms to share their private information with each other. In many financial and nonfinancial markets, third parties and trade associations collect private information of firms and calculate the aggregated statistics that is disseminated across the industry. The revealed information often takes a form of index, rate or industry report and is often referred to as a benchmark. There are many markets, in which benchmarks are extensively used. Benchmarks help consumers to navigate their search and firms use them as a basis for the price setting. Benchmarks, such as LIBOR and EURIBOR, underlie a significant fraction of business loans, commercial and retail mortgages, syndicated loans and approximately 100 trillions of OTC derivatives. Moreover, almost every industry has at least one trade association that produces summary industry reports that are extensively used by firms and buyers (e.g The Aluminium Association). Motivated by LIBOR and EURIBOR, I build a model, in which firms have to decide whether or not to reveal their private costs and contribute to the benchmark formation.

In their paper, Duffie *et al.* (2017) show that benchmarks can improve social welfare because of the two following effects. First, they encourage consumers to enter the market. Second, if consumers know the benchmark, it becomes more difficult for firms to set higher prices. In contrast, I show that benchmarks that are formed based on the exchange of information can reduce total welfare. Homogenization of information on the firms' side leads to prices that become less dispersed. Therefore, consumers do not find it beneficial to compare products and often buy the very first product they find. This allows firms to set higher prices and earn higher profits. Moreover, less active search unambiguously reduces total welfare as consumers often buy less preferred products from less efficient firms.

We show that the firms' incentives to contribute their private information about costs are highly dependent on the size of the search friction. Precisely, firms find it optimal to contribute to the benchmark when search friction is sufficiently high, whereas they prefer to conceal their private information when search friction is low. If search costs are low and many consumers decide to compare prices before buying, then information sharing will lead to intensified competition, which puts a downward pressure on prices. However, when we consider consumers with relatively high search costs, the incentives to share information change dramatically. In this case, consumers will search less, allowing firms to increase prices. We show that if the fraction of consumers with high search costs is large enough, then the gains from search deterrence outweigh the losses from intensified competition for shoppers.

1.2 Non-Reservation Price Equilibrium and Search Without Priors

In this joint work with Alexei Parakhonyak, we explore robust search by enriching the standard model in consumer search theory to account for model uncertainty. In the classical models of consumer search, uncertainty is reduced to risk: consumers have priors over parameters of the model and expected utility theory is well employed. But there is no recipe of how to choose a prior over the set of model specifications and therefore, it becomes important to characterize a robust search rule that performs well for any given prior. So the main objective of this chapter is to consider consumers who care about the model uncertainty and investigate how this non-strategic behavior affects the outcomes in search markets.

We derive two types of results: *i)* the optimal stopping rule, which can be easily used in the analysis of various economic models with search, and *ii)* the application of this rule to the model of oligopolistic competition in a homogeneous goods market. This optimal stopping rule can be easily used in the analysis of various economic models of search markets. Under uncertainty, the optimal consumer search rule is no longer characterized by a reservation price: there is a range of prices for which consumers stop with a probability between zero and one.

Regarding the results related to oligopolistic market outcomes, we get two types of results. First, in traditional search models, even a small positive search cost leads to monopoly pricing (Diamond (1971)). As all firms charge the monopoly price consumers have no incentives to search, and as consumers do not search there is no reason to undercut. Stahl (1996) showed that in the equilibrium with consumers following a reservation price strategy a necessary condition for price dispersion is to have a strictly positive mass of consumers with zero search costs. Due to the model uncertainty in our setup, this is, however, not necessary. As consumers follow a randomized search rule, instead of a reservation price rule, price dispersion may result in the market even if all consumers have a positive search cost. Second, we expect that market prices tend to be lower under model uncertainty as active search leads to more consumers making price comparisons thus putting downward pressure on prices. This would imply that less information (about the workings of the model) may improve consumer welfare.

1.3 Search and Learning Without Priors

In the joint work with Karl Schlag, we explore the problem of sequential search under minimal assumptions on the environment. More precisely, we are interested in the question of how to search and make inferences without relying on a specific initial information called priors. Consider a consumer or an unemployed worker who is engaged in costly search based on her initial information. In practice, this searcher does not know the principles of how the market functions and has vague information about the market fundamentals. Most of the search literature assumes that the searcher forms her leaning and searching decisions based on priors. However, these strategies are complex and are very sensitive to the misspecified initial information. Therefore, the searcher might want to have a search rule that would perform well even if her initial information is imprecise. A major obstacle is how to process new information and make inferences without priors. We avoid a formalization of how to learn by benchmarking to a set of rational agents of which we know something about their prior. The main idea is to search for a rule that is almost optimal for each of these agents who search and learn from the newly acquired information.

In this paper, we focus on a simple setting with binary-valued payoffs. In this setting, we characterize a dynamically robust search rule, that performs well for all priors. We identify two important properties of the robust rule. First, it is stochastic and prescribes to randomize between searching and buying. Second, this rule is adaptive to the newly acquired information and incorporates the pessimism of finding good alternatives over time. In addition, we characterize a myopic and a linear rule that can approximate the robust one very well. Finally, we show that the initial information that lies at the heart of the dynamically robust search rule restricts the set of possible environments that are consistent with the observed history. This unambiguously improves the performance in comparison to any stationary search rule.

2 Information Exchange and Consumer Search

2.1 Introduction

In many industries, third parties actively collect market-sensitive information and disseminate the aggregate statistics across market participants. This information is usually collected and shared on behalf of firms themselves and takes the form of industry reports and/or benchmarks.[1] Consumers rely on benchmarks to improve their search behavior and firms take benchmarks into account when they set prices. In this paper, we explore the implications of this form of information sharing in consumer search markets. Consumer search is an important aspect of our study, as information sharing affects firms' behavior not only directly, but also indirectly, through its impact on consumer search: consumers will search differently under the knowledge that the benchmark exists. We explain how information sharing affects search, competition and total welfare depending on the size of the search friction. We also inquire into the incentives of firms to share private information with each other.

Benchmarks are used extensively in Over-the-Counter (OTC) markets. According to a report of the Market Participant Group (MPG), trillions of dollars in loans are tied to the LIBOR and EURIBOR benchmarks. Among them are business loans (1,2$ TNs), commercial and retail mortgages (1,4$ TNs each)[2], syndicated loans (3,4$ TNs) and variable-rate bonds (1,2$ TNs). The volume of the

[1] The European Commission defines a *benchmark* as any commercial index or published figure calculated by the application of a formula to the value of underlying assets or prices, including estimated prices, interest rates or other values.

[2] In 2012, the Federal Reserve Bank of Cleveland found that 45% of prime adjustable-rate

OTC derivatives linked to LIBOR is approximately 100 trillion dollars. LIBOR and EURIBOR are computed as the average of the interbank borrowing rates (the costs of raising funds) of the member banks.[3] In addition to these financial markets, there are more than 7000 trade associations in various industries across the US and approximately 2000 in the UK that publish benchmarks and industry reports based on the private information of firms.[4] Due to recent legal cases concerning the manipulation of a variety of benchmarks, many authorities have begun to thoroughly consider the problem of benchmark formation based on firms' private information.[5]

As in the case of LIBOR and EURIBOR, we consider a market in which firms have private information about their respective costs. Information revelation improves market transparency both for the competing firms and for consumers, and has an impact on market outcomes in two important ways. First, better informed consumers search more efficiently, which leads to greater consumer surplus. Second, benchmarks may also deter search as increased transparency makes firms use benchmarks as the basis for their pricing. As a result, prices become more aligned and the expected gains yielded by searching diminish, allowing firms to set higher prices. We show the conditions on the market structure under which information revelation transfers welfare from consumers to firms and reduces matching efficiency, as well as total welfare.

We build a model in which in the first-stage firms decide whether or not to truthfully submit their private costs to the trade association (TA), which calcu-

mortgages (ARMs, 22% of outstanding volume) and 80% of subprime ARMs used LIBOR as a benchmark.

[3] There are 17 banks that currently contribute to the fixing of the US Dollar LIBOR, such as Bank of America, Barclays, JP Morgan etc. LIBOR is administrated by the Intercontinental Exchange and EURIBOR is published by the European Money Markets Institute.

[4] Average Wholesale Price (AWP) is a benchmark that refers to the average price at which drugs are purchased at the wholesale level. LBMA administrates benchmarks for gold (25$ BNs per month) and silver (3$ BNs per month) based on the daily auctions in which large traders determine the price. The Aluminium Association collects data on the costs incurred by the main producers and publishes monthly reports: www.aluminum.org

[5] See the reports on *Principles for Financial Benchmarks* by IOSCO, *Regulation of Indices* by EC and *Principles for Benchmark-Setting Processes* by ESMA and EBA.

lates the average costs of all firms that have submitted information.[6] The TA then disseminates this benchmark across all firms and consumers. In the second stage of the game, consumers search and firms, anticipating the search decision of consumers, simultaneously set their prices based on private costs and the benchmark. Consumers are heterogeneous in their valuation of acquiring the products and have different search costs. We assume that there is a positive fraction of consumers with zero search costs which we refer to as *shoppers*. Shoppers always compare products and prices before buying. On the other hand, *non-shoppers* engage in costly search and decide how many firms to visit based on the size of the search costs and the revealed benchmark. It is also important to note that consumers take into account not only the benchmark itself, but the fact that firms have exchanged information on their private costs in order to create the benchmark.

We show that the size of the search friction plays a crucial role in a firm's decision whether or not to share private information. Given that benchmarks affect firms' pricing and consumers' search behavior, the reactions of both sides of the market play important roles. More accurate information on rivals' costs allows firms to choose the prices that best reflect the underlying market conditions and this unambiguously increases profits. On the other hand, the pooling of information leads firms' pricing strategies to be increasingly correlated. Indeed, if the benchmark reveals that competitors' costs are higher than the expected level, then a firm has an incentive to increase its price. Firms with relatively high costs will decrease their prices.[7] Therefore, due to the strategic complementarity of prices, information exchange decreases variance in pricing, which, has two important effects. First, less dispersed prices intensify competition for shoppers. Second, it

[6] We assume that the trade association provides the exogenous mechanism of truthful reporting. The LIBOR scandal, in which banks manipulated reporting rates, induced governments and trade associations to redesign the benchmarks to discourage untruthful reports by banks. Benchmarks design has been explored in recent papers by Duffie and Dworczak (2014) and Coulter and Shapiro (2015).

[7] It can also be the case that all firms have relatively low (high) costs in comparison to the expected level and they will all decrease (increase) their prices. What is important in this case is that firms with higher costs will decrease (increase) prices by more (less) than firms with low cost. Therefore, prices become less dispersed. Technically, this is a case of the natural assumption of increasing differences in prices, see Vives (2000).

decreases non-shoppers' incentives to search. We will refer to this second effects as the *variance effect*[8]. In addition to these two main effects, the benchmark provides better information and shrinks the range of prices that consumers might have expected, thereby reducing the incentives to search even further.

Given these different effects, our first result shows that firms find it optimal to reveal their private costs when the search friction is sufficiently high, whereas they prefer to conceal their private information when the search friction is sufficiently low. If the fraction of shoppers is high and many consumers decide to compare prices before buying, then information sharing will lead to intensified competition (on average), putting downward pressure on prices. Thus, firms prefer to conceal their private costs and the benchmark is not formed. This result is in line with the standard literature on information exchange in oligopoly where all consumers are informed about prices.[9] However, when we consider consumers with relatively high search costs, the incentives to share dramatically change. In this case, consumers will search less, allowing firms to increase prices. We show that if the fraction of consumers with high search costs is large enough, then the gains from search deterrence outweigh the losses from intensified competition for shoppers and firms decide to reveal.

Our second result is a welfare result. We show that if the search friction is sufficiently high, information revelation reduces consumer surplus and total welfare. The increase in correlation of prices, as well as the improved information on the average cost, allows consumers to avoid wasteful search, which is good for total welfare. On the other hand, less-active search increases matching inefficiency. Since more consumers decide not to compare prices before buying, high-cost firms enjoy a greater market share and more consumers end up buying from inefficient firms. This generates lower gains from trade that cannot be recovered by the savings accrued in terms of search costs. Moreover, because of the variance effect, firms increase their prices thereby reducing consumer surplus. As a result, the total welfare loss from information revelation in markets with sufficiently high search costs is paid by the consumers.

[8] Our terminology is in line with Benabou and Gertner (1993).

[9] See Gal-Or (1986), Raith (1996), Vives (2000).

Our final result is that, qualitatively, the above results continue to hold for the case of *private* information exchange, in which consumers only know that information exchange took place, but do not know the content of the information that was exchanged. We show that the variance effect by itself can severely deter search such that the losses from matching inefficiency outweigh the benefits from search costs savings.

A recent paper by Duffie *et al.* (2017) also explores how benchmarks affect social welfare. The authors build a model in which firms have identical cost and explore the question of whether or not welfare is greater if consumers know this cost. If there is no benchmark, consumers learn about the underlying costs from prices. Based on the mechanism of learning described by Janssen *et al.* (2011), they show that benchmarks improve social welfare. However, they do not study how the benchmarks come about or whether or not firms have an incentive to share information that would result in the creation of a benchmark. In contrast, the focus of our paper is benchmark formation. As in the case of LIBOR, benchmarks are often created on the basis of information provided by firms, where each of the firms may face a different (possibly correlated) cost. To draw a comparison with Duffie *et al.* (2017), two additional points become important: (i) whether or not consumers know that information has been exchanged and (ii) whether or not firms have an incentive to share this information. Thus, consumers have an additional source of learning: if information is exchanged, prices become less dispersed and consumers search less. In order to focus on these two points, we eliminate the learning-from-prices channel and consider non-sequential search.[10]

More generally, this paper is related to two strands of literature. First, it relates to the literature on *learning in search markets*. Most of this literature considers firms' information structures as exogenously given. In some studies, all firms have the same information on the parameters of the model (Varian (1980), Stahl (1989),

[10] In their paper, De los Santos *et al.* (2012) observe that people may act as if they search non-sequentially. Morgan and Manning (1985) show that depending on the search problem, both sequential and non-sequential protocols can be optimal. The non-sequential search protocol better fits markets with economies of scale in the search trials (searching online, using price aggregators). In some OTC markets, and especially in their interdealer market segments, there are brokers that provide summarized information about prices, resulting in economies of scale.

Wolinsky (1986), Dana Jr (1994) and Janssen *et al.* (2011), Janssen *et al.* (2017)). In this case price dispersion arises as a result of firms employing mixed strategies. In other approaches, price dispersion can arise as a result of asymmetry of information between competitors (Reinganum (1979), Benabou and Gertner (1993)). In these models, firms have private information about their own costs. The static modelling in these papers sheds some light on what strategies are employed in a particular asymmetric information game. The question as to why a specific type of asymmetry arises has not yet received any attention in the search literature. In order to fill this gap, we allow for information exchange between firms before the pricing game takes place.

Second, there is a large literature on *information sharing* in markets. In much of this literature, firms reveal their private costs to a trade association that discloses the average cost that we refer to as a benchmark. The literature on information sharing in oligopoly (Gal-Or (1985), Gal-Or (1986), Vives (1990), Raith (1996) and Vives (2010)) underlines that the most important determinant of the incentives to reveal information is the reaction of competitors. Key to this literature is the assumption that consumers are able to perfectly observe prices. Under this assumption, consumers' purchasing strategy does not depend on (i) whether or not firms exchange information and (ii) whether or not this information is also revealed to consumers. Therefore, the problem of information exchange is reduced to the exploration of the reaction of competitors. However, in search markets such as OTC markets, information exchange also affects the behavior of consumers. We show that results from this standard literature do *not* hold when the search friction is sufficiently large.

Information exchange between firms is of primary importance to antitrust authorities and is regulated by the Guidelines on Horizontal Agreements. Information exchange has never been analysed in search markets. Therefore, we also contribute to the growing body of work that analyses the impact of consumer search on traditional competition policy. Mergers in search markets have been analysed by Moraga-González and Petrikaitė (2013), collusion by Petrikaitė (2016) and vertical restraints by Janssen and Shelegia (2015).

The rest of the paper is organized as follows. Section 2 describes the model.

In section 3, we analyse firms' pricing and the search behavior of consumers in each possible information revelation scenario. In section 4, we characterize the equilibrium revelation strategy and provide conditions on the size of the search friction under which benchmarks are formed. In section 5, we explore how information revelation affects social welfare. Section 6 concludes. Technical proofs are presented in the Appendix.

2.2 Model

Consider an oligopolistic market with $N \geq 2$ firms that produce a differentiated good and compete in prices.[11] Importantly, firms have asymmetric production costs c_i, $i = 1, \ldots, N$ which is their private information. We assume that c_i is independently and identically distributed on the compact support $[\underline{c}, \overline{c}]$ according to the distribution $G(c)$ with a variance $\sigma_c^2 > 0$.[12] Although we do not explicitly model the wholesale market, one may view the asymmetry of costs coming from the fact that firms might work either with different wholesalers or with the same wholesaler offering different contracts. In banking, cost asymmetry is a result of different sources from where banks raise their funds.

Due to this asymmetry, firms can decide whether or not to make an agreement to share their costs with other market participants. The information exchange is performed by the trade association (TA), which collects all submissions and calculates the aggregate statistics that we refer to as a *benchmark*. If $K \subseteq \{1, \ldots, N\}$ is the non-empty set of firms which decided to join the trade association, then the benchmark b is modelled as the average of submitted costs, so $b = 1/|K| \sum_{j \in K} c_j$.

We will assume that the trade association *publicly* announces the benchmark and that all market participants can observe it. Later on, we will distinguish

[11] Many financial markets in which benchmarks are extensively used are characterized by substantial product differentiation as it has been documented by various studies on credit markets (Matutes and Vives (1996), Kim *et al.* (2005)) and mortgage markets (Allen *et al.* (2012), Del Prete *et al.* (2017)). The differentiation comes from the variety of complex contracts that might be suitable for some borrowers but not for others.

[12] I separately analyse the model with correlated costs and show that the results qualitatively remain the same. The preliminary version of this extension is available upon request.

between *public* and *private* information exchange. In contrast, if information is shared privately then consumers cannot see the benchmark but might observe or believe that information is exchanged between firms. We postpone the discussion of how the observability of benchmarks by consumers affects our results.

There are two important assumptions we make on how firms exchange information. First, as in the literature on information exchange mediated by third parties, we assume that firms have to decide whether or not to reveal costs *before* the actual arrival of private information. Therefore, we consider the incentives to reveal costs unconditionally on the information itself. This assumption naturally describes the formation of many benchmarks such as LIBOR or EURIBOR, in which banks commit themselves to submit their costs on a regular basis. Second, we assume that the costs of manipulating the reports are high enough so that the information is revealed *truthfully.* Otherwise, firms would always try to exaggerate their costs to induce their competitors to set higher prices, and the formation of benchmarks would be impossible.

The demand side of the market is represented by an infinite number of consumers, each of whom wishes to buy one unit of at most one of N products. The heterogeneity of consumers' preferences is described by the Chen and Riordan (2007) spokes model. This model captures the nonlocalized oligopolistic competition in which any firm directly competes with each of its $N-1$ rivals based on the Hotelling framework. For concreteness, we assume that each consumer cares only about two products, and values all the $N-2$ remaining products at zero. All consumers assign valuation v to these two products and we assume that v is large enough so the market is fully covered. The preference relation between two products is described by the Hotelling line of unit length, with two firms located at the endpoints. We say that consumer $\ell \in L_{ij}$ (L stands for the Hotelling line) if she cares about products i and j. Consumers are uniformly distributed on the corresponding lines and have linear transportation costs measured by t. The location of consumer ℓ relative to the firm i is denoted by $x_{\ell i}$ ($x_{\ell i} + x_{\ell j} = 1$). There is a mass of $2/(N-1)$ consumers located between two randomly selected firms i and j, so the total demand of each firm cannot exceed 2.[13]

[13]The total demand in the market is $C_N^2 2/(N-1) = N$, so 1 per firm.

Although a random consumer ℓ knows which two products she is interested in, we assume that her location and all prices are ex ante unknown. Consumers have to engage in costly search and are heterogeneous in search costs. We assume that in each Hotelling line there is a fraction of *shoppers* $\lambda \in [0, 1]$ that have zero search cost and visit both firms before buying. Notice, that if $\lambda = 1$ then the model coincides with a generalized version of Gal-Or (1986).[14] The remaining $1 - \lambda$ are *non-shoppers* and have positive search costs which are distributed according to the uniform distribution $s_\ell \sim U[0, \overline{s}]$. It is assumed that consumers search *non-sequentially* and become informed about prices and their locations by paying search costs. Otherwise, consumers choose in which firm to buy and randomize with probability 1/2 if they are indifferent.

In order to avoid the boundary problem in which some firms cannot attract informed consumers because of high prices, we impose the following constraint on the relation between t and the support of the cost distribution:

Assumption 2.1. $\overline{c} - \underline{c} < 2t$.

This assumption implies that the variance of costs is not too high $\sigma_c^2 \leq (\overline{c} - \underline{c})^2 < 4t^2$. As long as t is high enough, meaning that the product differentiation is sufficiently strong, the informed consumers will not react too drastically to the difference in prices. This will ensure that the location of the indifferent consumer lies in the interior of the respective Hotelling line. Therefore, firms with the highest possible cost $\overline{c}$ will sell to some informed consumers even if they compete with the most efficient firms with costs $\underline{c}$.

The timing goes as follows. In the first stage, firms simultaneously decide whether to join the trade association and reveal their costs. Then the private costs are generated and the TA conducts the transmission of the private information according to the firms' commitments. Afterwards, the TA calculates the average cost b and publicly announces $\{b, k\}$, where k is the number of contributing firms.[15]

[14] In Gal-Or(1986), the author considers the information exchange problem in a duopoly, assuming that costs are normally distributed. In contrast, we analyse an oligopolistic market and work with a general distribution of production costs.

[15] We assume that consumers observe k but cannot see the identities of firms that made the

In the second stage, firms simultaneously set prices p_i, $i = \{1, \ldots, N\}$ and consumers search. Consumer ℓ cannot observe prices and makes the search decision $a_\ell \in A = \{0, 1\} = \{\textit{do not search}, \textit{search}\}$ simultaneously with other consumers. The resulting fraction of informed consumers interested in product i and j is $\mu_{ij} = \int_{L_{ij}} a_\ell d\ell > \lambda$.

Thus, the profit of firm i that charges p_i and has cost c_i is given by

$$\pi_i(p_i, p_{-i}, a) = (p_i - c_i) \sum_{j \neq i} \frac{2}{N-1} \int_{L_{ij}} \phi_\ell(p_i, p_j, a_\ell) d\ell,$$

where the summation is taken over the competing products j and ϕ_ℓ equals to 1 if consumer ℓ buys the product i and 0 otherwise. The net surplus of consumer ℓ who decided to buy from firm i is given by:

$$u_\ell(p_i, x_{\ell i}) = v\mathbb{I}_{\ell \in \cup_{j \neq i} L_{ij}} - (p_i + t x_{\ell i} + a_\ell s_\ell)$$

The strategy profile of firm i includes any mapping p from the production cost c_i to the actual price: $p = p_i(c_i, b)$.[16] Consumers search strategy is described by $q_\ell(\cdot; b, s_\ell) \in \Delta(A)$, so consumers can randomize their search decisions.

We look for a profile of strategies in which: *1)* firms optimally decide whether or not to contribute to the benchmark formation $e_i \in \Delta E = \Delta\{0, 1\} = \Delta\{\textit{reveal}, \textit{conceal}\}$ and *2)* observing this benchmark, consumers search optimally given the pricing strategies of the firms and firms maximize their profits given the number of consumers that become informed. We formally define the equilibrium notion below:

Definition 2.1. *A Perfect Bayesian Equilibrium (PBE) in the model of benchmark formation in search markets is the profile $(e_i, p_i(c_i, b))$ for all i and $q_\ell(\cdot)$ for all ℓ such that:*

1. *Each firm i chooses e_i to maximize its expected profits given the revelation strategies of other firms e_{-i} and anticipating that consumers will search according to $q_\ell(\cdot; b, s_\ell)$ and firms will set prices according to $p_j(\cdot, b)$;*

submission. In the Appendix we show that this assumption is not restrictive and was made for the model simplicity. In reality, trade associations might reveal the list of contributing firms and consumers might use it.

[16] We do not exclude price distributions from the consideration but, since the pricing game is based on the Hotelling framework, it is easy to see that firms will never use mixed strategies. We prove this fact in the next session.

2. *Using the Bayes' Rule, consumers update their beliefs about the underlying price dispersion based on the realized benchmark b. Then consumers optimally choose their search strategies q_ℓ given their beliefs about prices and the search strategies of the other consumers. The resulting total fraction of informed consumers in each market L_{ij} is $\mu_{ij}(b)$;*

3. *Based on the benchmark b, each firm i optimally sets its price according to the function $p_i(c_i, b)$, given the price strategies of other firms $p_j(\cdot, b)$ and the belief about the total fraction of informed consumers $\mu^e_{ij}(b)$ which is correct in equilibrium.*

2.3 Analysis: The Second Stage

We analyse the model by solving it backwards and start with the second stage of the game in which firms set prices and consumers search. Suppose that $K \subseteq \{1, \ldots, N\}$ is a non-empty set of $k = |K|$ firms which decided to reveal their costs and form a benchmark $b = 1/k \sum_{j \in K} c_j$. If all firms decide to conceal their cost we set $K = \emptyset$ and $k = 0$, so the benchmark is not formed. In this section, we characterize the equilibria in every possible subgame described by the outcome of the preceding benchmark formation stage $\{b, k\}$.

2.3.1 Firms' Pricing

We start from the investigation of the optimal pricing of firms which revealed or concealed their costs for a given search behavior of consumers. Since consumers follow the non-sequential search and decided whether to be informed or not, the outcome of their decision is a fraction of informed consumers. Since consumers do not know the identities of the firms which shared information, we have that the resulting fraction of informed consumers must be the same for all direct markets L_{ij}, so $\mu_{ij}(b) = \mu(b)$. Firms do not know the resulting $\mu(b)$ but have the correct expectation in equilibrium.

In the degenerate case when $\mu(b) = 0$, firms are local monopolists which set the same monopoly price v. Despite this, there is a positive mass of consumers with very small search costs who will still search in order to learn their locations. Therefore, there is an $\epsilon > 0$, such that all $\mu(b) < \epsilon$ cannot constitute an equilibrium. We focus on the positive $\mu(b)$ and assume that v is high enough that all consumers buy.

Consider how the information of consumers affects firms' demand. Suppose an informed consumer who has to choose between firm i and j and is located at the distance x from product i. This consumer will buy product i if p_i is low or she is located close enough to firm i: $p_i + tx \leq p_j + t(1-x)$ or, equivalently:

$$x \leq \overline{x}_{ij} = \frac{1}{2} + \frac{p_j - p_i}{2t}$$

,where $\overline{x}_{ij}$ is the indifferent consumer in the market L_{ij}. Thereafter, we will assume that the indifferent consumer is located at the interior of the corresponding Hotelling line L_{ij}, so $\overline{x}_{ij} \in (0,1)$, and we will show that this is the case under Assumption 2.1. Uninformed consumers $1-\mu(b)$ split evenly among the firms and buy there.

The expected demand of firm i setting price p_i is given by:

$$D(p_i) = \sum_{j \neq i} \frac{2}{N-1} \left\{ \mu(b) \left(\frac{1}{2} + \frac{\mathbb{E}[p_j|I_i] - p_i}{2t} \right) + \frac{1-\mu(b)}{2} \right\} \tag{2.1}$$

The expression in the curly brackets represents the demand of consumers choosing between products i and j. The first term is the expected fraction of informed consumers who are located closer to firm i than the indifferent consumer $\overline{x}_{ij}$. The expectation are taken conditionally on the information I_i if firm i that includes the benchmark $\{b,k\}$ and whether firm i contributed to it. The summation is taken over all $N-1$ Hotelling lines, each of which contains the mass of $2/(N-1)$ consumers.

By simplifying the demand expression, the profit function can be written in the following way:

$$\pi_i(p_i) = \frac{\mu(b)}{t}(p_i - c_i) \left(\frac{t}{\mu(b)} + \frac{1}{N-1} \sum_{j \neq i} \mathbb{E}[p_j|I_i] - p_i \right) \tag{2.2}$$

There are two important points about the functional form of the profit given in (2.2). First, a change in the fraction of informed consumers $\mu(b)$ acts as a change in transportation costs t but with a different magnitude. Precisely, a decrease in $\mu(b)$ by $x\%$ is equivalent to an increase in t by $x/(1-x/100)\% > x\%$. So the market power of firms is higher for lower $\mu(b)$. Second, since the profit function is quadratic we are able to solve for the equilibrium pricing analytically using the similar technique as in Li (1985) and Shapiro (1986) who analyse the information sharing in Cournout oligopoly and abandon the assumption of normality of costs distribution.

The first order condition of firm i's profit maximization problem is:

$$2p_i = t_1 + c_i + \frac{1}{N-1}\sum_{j\neq i}\mathbb{E}[p_j|I_i] \tag{2.3}$$

The linearity of (2.3) with respect to c_i hints that there is an equilibrium in linear strategies. Assume that firms follow the symmetric linear strategies with respect to their revelation decisions. Define the candidate strategies as:

$$\begin{cases} p_R(c_i) = \alpha_R + \beta_R b + \gamma_R c_i, \text{ if } i \in K \\ p_C(c_i) = \alpha_C + \beta_C b + \gamma_C c_i, \text{ if } i \in N \setminus K \end{cases} \tag{2.4}$$

,where indexes R and C stand for "reveal" and "conceal" respectively. The next proposition shows that there is a unique profile of pricing strategies satisfying (2.3) and it has the linear form given by (2.4).

Proposition 2.1. *Given any outcome of the benchmark formation in the first stage $\{b, k\}$, there is a unique Bayesian Nash Best Response of firms to any search behavior described by $\mu(b) > 0$. Depending on whether firm i revealed or concealed its costs, it sets the price according to the linear pricing strategies given by:*

$$\begin{aligned} p_R(c_i) =& \frac{t}{\mu(b)} + \frac{N-k}{2N-1}c^e + \frac{k}{2N-1}b + \frac{N-1}{2N-1}c_i \\ p_C(c_i) =& \frac{t}{\mu(b)} + \left(\frac{1}{2} - \frac{k}{2N-1}\right)c^e + \frac{k}{2N-1}b + \frac{1}{2}c_i \end{aligned} \tag{2.5}$$

The proof of proposition 2.1 consists of two steps. First, when we substitute the linear candidate to the first order condition described by (2.3), we get a linear system of six equations and six unknown coefficients from (2.4). We show that this system has a unique solution. Second, we show that the FOC (2.3) connects the pricing strategies in a way that precludes the existence of more than one equilibrium (lemma 2.5 in the appendix).

In (2.1) it was assumed that all indifferent consumers are located in the interior of the corresponding Hotelling lines. It is easy to verify that the difference in any two equilibrium prices given by (2.5) cannot be higher than $(\bar{c} - \underline{c})/2$. Thus, according to Assumption 2.1, the indifferent consumer is in the interior: $|p_i - p_j| < t$ and $x_{ij} \in (0, 1)$ for all L_{ij}.

There are several important points concerning the equilibrium price strategies given in (2.5). First, all firms set higher prices for a lower fraction of informed consumers $\mu(b)$ and pass through the benchmark b at the same rate which is more profound if more firms have contributed to the benchmark. Second, the expected prices are the same: $p^e_R = p^e_C$. So, the revelation by itself does not change the relative market power of firms from an ex ante perspective. Third, the revelation changes the price dispersion measured by $\mathrm{Var}(p_1 - p_2)$, that captures both the variances of p_1 and p_2 and the correlation between them.[17] Notice that the fraction of informed consumers $\mu(b)$ and the benchmark b do not affect the price dispersion and determine only the expected price level relative to which prices vary (terms dependent on $\mu(b)$ and b cancel out in $p_1 - p_2$). Importantly, the price dispersion is determined by the revelation decisions of firms and the following corollary shows that the revelation makes prices less dispersed:

Corollary 2.1. *Price dispersion of any two given firms i and j is lower if i and/or j reveal their costs.*

Proof. The direct comparison of the variances of price differences establishes the result:

$$\mathbb{E}(p_C(c_i) - p_C(c_j))^2 = \frac{1}{2}\sigma_c^2 > \left(\frac{1}{4} + \left(\frac{N-1}{2N-1}\right)^2\right)\sigma_c^2 = \mathbb{E}(p_R(c_i) - p_C(c_j))^2 >$$
$$> 2\left(\frac{N-1}{2N-1}\right)^2 \sigma_c^2 = \mathbb{E}(p_R(c_i) - p_R(c_j))^2$$

[17] In the standard search literature (Varian(1980) or Stahl(1989)), the dispersion is usually associated with the variance of equilibrium price distribution, since prices are independently and identically distributed. In the general case, price correlation should be taken into account. As an alternative to our approach, one could measure price dispersion in the expected positive price differences $\mathbb{E}\max\{p_1 - p_2, 0\} = \mathbb{E}\,|p_1 - p_2|$.

□

We will refer to the decrease of price dispersion due to information sharing as the *variance effect*. The intuition behind the variance effect comes from the fact that holding benchmark b fixed, a revealing firm passes through its cost at a lower rate than a concealing firm. In other words, revealing firms base their pricing on the public information to a greater extent than firms that conceal information which results in more similar prices. For $N = 2$ revealing firms pass through 33% of their private information (c_i) and 66% of the public information (c^e and b). In contrast, concealing firms pass through 50% of their cost. Notice, that the variance effect becomes less severe with an increase in N as the pass through rate of revealing firms approaches the pass through rate of concealing firms: $\lim_{N\to\infty}(N-1)/(2N-1) = 1/2$.

2.3.2 Consumer Search

On the demand side, consumers cannot observe prices and their locations, so they have to decide to either pay search cost s_ℓ and inspect two products or save it and visit only one firm and buy there. They base their decisions on the benchmark $\{b, k\}$ that is publicly announced by TA and on their beliefs about firms' pricing. The expected gain of being informed is represented by the difference between the expected effective price and the expected minimum of two effective prices that include the transportation costs of consumers. Consumer ℓ will find it optimal to search if the gain of search compensates her for the search costs incurred:

$$s_\ell \leq \tilde{s}_k(b) = \mathbb{E}[p_i + tx_i|b] - \mathbb{E}[\min\{p_i + tx_i, p_j + t(1-x_i)\}|b]$$

, where expectation is taken over two prices and consumers' distance x_i to firm i and $\tilde{s}_k(b)$ is the consumer who is indifferent between searching and staying uninformed. Using the identity $-\min\{a, b\} = \max\{-a, -b\}$ we rearrange the expected gain of search in the following form:

$$\tilde{s}_k(b) = \mathbb{E}[\max\{p_i - p_j + t(2x_i - 1), 0\}|b] = 2t\mathbb{E}\left[x_i - \overline{x}_{ij}, 0\}|b\right]$$

, where $\overline{x}_{ij}$ is the location of the indifferent consumer in L_{ij}. The gain of search is positive if the consumer is located closer to firm j. Next, we take the expectation with respect to x_i:

$$\mathbb{E}_{x_i}[x_i - \overline{x}_{ij}, 0] = \int_{\overline{x}_{ji}}^{1} (x_i - \overline{x}_{ij}) dx_i = \frac{1}{2}(1 - \overline{x}_{ij})^2 = \frac{1}{2}\overline{x}_{ji}^2.$$

and get that the location of the indifferent consumer is given by:

$$\tilde{s}(b) = t\mathbb{E}_p[\overline{x}_{ji}^2 | b] = t\mathbb{E}_p\left[\left(\frac{1}{2} + \frac{p_i - p_j}{2t}\right)^2 \middle| b\right] \tag{2.6}$$

From (2.6) we see that search becomes more valuable if the variance of the indifferent consumer location $\overline{x}_{ji}$ is high, since in this case there is a high chance that a more desired product is not inspected. The variance of the indifferent consumer location is driven by the price dispersion. Therefore, the price dispersion is the key driver of consumer search.

It is important to notice that the price difference $p_i - p_j$ given by (2.5) does not depend on $\mu(b)$ in equilibrium, thus it does not affect the price dispersion. So, the right side of (2.6) does not depend on $\mu(b)$ and the equilibrium $\mu(b)$ exists and is uniquely determined. Search costs are distributed according to $s_\ell \sim U[0, \overline{s}]$, so the probability that search costs are lower than $\tilde{s}(b)$ is given by $\tilde{s}(b)/\overline{s}$. Since there is a $1-\lambda$ fraction of uninformed consumers, the total fraction of informed consumers is given by:

$$\mu_k(b) = \lambda + (1-\lambda)\tilde{s}(b)/\overline{s}$$

Next proposition shows that if firms reveal their costs and form the benchmark, then less consumers are expected to search: the ex ante fraction of informed consumers $\mu_k = \mathbb{E}\mu_k(b)$ (at the stage of benchmark formation) decreases in k.

Proposition 2.2. *The ex ante equilibrium fraction of informed consumers is lower for a higher number of revealing firms k: $\tilde{s}_k > \tilde{s}_{k+1}$ and $\mu_k > \mu_{k+1}$.*

The intuition of this claim goes as follows. It was shown in corollary 2.1 that if more firms decide to contribute to the benchmark, then their prices become less dispersed due to the variance effect of information exchange. As a result, the expected gain of search given by 2.6 becomes lower and consumers are less likely to find a better deal. Diminished necessity of search decreases the fraction of informed consumers μ and pushes firms' prices upwards.

2.4 Information Sharing

We explored the equilibrium pricing and search behavior of consumers for any outcome of the benchmark formation stage and now we are prepared to inquire into the firms' incentives to reveal their private costs. We consider the benchmark formation stage from the ex ante perspective, so all firms simultaneously decide whether to join the trade association before knowing their costs. In this section, we show that firms find it optimal to reveal information when search friction is sufficiently strong: i) the fraction of informed consumers λ is low and/or ii) the search costs characterized by the upper bound $\overline{s}$ are sufficiently high.

2.4.1 No search friction

We start the analysis of the first stage by considering the incentives to exchange information in the market without search friction when $\lambda = 1$. Let $\Pi^R(k)$ and $\Pi^C(k)$ be the ex ante profits of revealing and concealing firms respectively, when k firms contribute to the benchmark. These profits are given by the following expressions:

$$\Pi^R(k) = \mathbb{E}(p_R(c_i) - c_i)^2/t \qquad\qquad \Pi^C(k) = \mathbb{E}(p_C(c_i) - c_i)^2/t$$

, where p_R and p_C are the equilibrium pricing strategies given by (2.5). Information sharing allows firms to choose prices that better correspond to the real market conditions and this unambiguously improves their profits. At the same time, information sharing makes pricing more correlated and prices become less dispersed by corollary 2.1. To describe the main tradeoff we consider the profit function in the following general form:

$$\Pi(k) = \mathbb{E}(p_i - c_i)(1 - p_i/t) + \frac{1}{t(N-1)} \sum_{j \neq i} [\underbrace{\mathbb{E}[p_j p_i]}_{+} - \underbrace{\mathbb{E}[p_j c_i]}_{-}] \tag{2.7}$$

The increased price correlation improves profits since the second term becomes higher due to the strategic complementarity of pricing. On the contrary, higher costs make rivals set higher prices leading to a higher demand that raises firms' total costs. The following proposition shows that the negative effect always dominates and benchmarks formation is impossible in markets without search friction.

Proposition 2.3. *It is optimal to conceal private costs in the market without search friction* ($\lambda = 1$)*:* $\Pi^C(k) > \Pi^R(k+1)$ *for all* $k = 0, 1, \dots N-1$*.*

This result is a generalization of Gal-Or (1986) to oligopolistic markets and bounded distributions of production costs. The "no-revelation" equilibrium is in dominant strategies because firms find it optimal to deviate and conceal their costs for any given number of revealing firms: $\Pi^C(k) > \Pi^R(k+1)$. So, in markets without search friction, price dispersion improves firms' profits.

2.4.2 Search friction

Consider the problem of benchmark formation in search markets with positive fraction of non-shoppers $\lambda < 1$.

Denote $\tilde{t}(b) = t/\mu(b)$ as the effective transportation cost that results from the formation of benchmark b. The equilibrium ex ante firms' profit can be written similar to expression (2.7):

$$\Pi(k) = \mathbb{E}(p_i - c_i)\left(1 + \frac{1}{N-1}\sum_{j\neq i}(p_j - p_i)/\hat{t}(b)\right) =$$
$$= \underbrace{\mathbb{E}(p_i - c_i)(1 - p_i/\tilde{t}(b))}_{+} + \frac{1}{(N-1)}\sum_{j\neq i}\underbrace{(\mathbb{E}[p_j p_i] - \mathbb{E}[p_j c_i])}_{-}\Big/\tilde{t}(b) \quad (2.8)$$

First, the tradeoff for the markets with $\lambda = 1$ described in (2.7) is present in the search markets as well(the second term in (2.8) includes the second term in (2.7)). According to proposition 2.3, firms find it optimal to conceal information when they compete only for informed consumers. Second, the information exchange decreases price dispersion and consumers search less, which leads to a higher effective transportation cost $\tilde{t}(b)$. From equation (2.8), we can see that a higher $\tilde{t}(b)$ increases the first term, since the own-price elasticity of demand becomes lower due to a higher fraction of non-shoppers $1 - \mu(b)$. Moreover, a higher $\tilde{t}(b)$ decreases the negative effect of less dispersed prices described by the second term.

Next, we explore the interplay of these two effects and determine the conditions, under which firms find it optimal to contribute to the benchmark. The analysis is heavily based on the comparison between *public* and *private* information sharing. Under *private* information sharing, consumers *can* see the number of firms that revealed their costs k but *cannot* observe the realized benchmark b. In this case, consumers do not learn anything from the benchmark, but rather update their beliefs about prices based on the fact that firms exchange information. We will denote Π_1 and Π_0 as the equilibrium profits in the case of public and private information sharing respectively. If information is shared publicly, then the equilibrium profit of firm i can be written as the expected profit in the second stage given by (2.2) using the first order condition (2.3):

$$\Pi_1(k) = \mathbb{E}\left[\frac{\mu(b)}{t}\left(\frac{t}{\mu_k(b)} + \frac{1}{N-1}\sum_{j\neq i}(p_j - p_i)\right)^2\right] = t\mathbb{E}_b\left[\frac{1}{\mu_k(b)} + \mu_k(b)T_k(b)\right] \tag{2.9}$$

, where $T_k(b)$ describes the term that increases when prices become more dispersed:

$$T_k(b) = \mathbb{E}\left[\left(\frac{1}{N-1}\sum_{j\neq i}(p_j - p_i)/t\right)^2 \middle| b\right] \tag{2.10}$$

and $p_i = p_C$ or $p_i = p_C$ depending on whether firm i revealed or concealed its cost. By assumption (2.1) we have that each intermediate consumer x_{ij} is in the interior of the corresponding Hotelling line L_{ij} and therefore we can evaluate $T_k(b)$ from above. By proposition 2.1 the difference between any two prices cannot be larger than $(\bar{c} - \underline{c})/2)$, thus

$$T_k(b) < \mathbb{E}\left[\left(\frac{1}{N-1}\sum_{j\neq i}(\bar{c} - \underline{c})/2t\right)^2 \middle| b\right] < 1.$$

If information is exchanged privately, then consumers cannot condition the expected gains of search on the benchmark b and decide to search if $s < \tilde{s}_k = \mathbb{E}\tilde{s}_k(b)$. The equilibrium profit of firm i in this case is given by:

$$\Pi_0(k) = t\left[\frac{1}{\mu_k} + \mu_k \mathbb{E}T_k(b)\right] \tag{2.11}$$

, where $\mu_k = \mathbb{E}\mu_k(b)$ is the fraction of searching consumers who cannot observe the benchmark but can see that k firms exchanged information.

To solve for the equilibria in the benchmark formation game we introduce some useful lemmas that significantly help to describe the nature of revelation decisions. First, we determine how consumers' learning from the benchmark affects the equilibrium profit and compare (2.9) and (2.11). The first term in (2.9) is a convex function of $\mu_k(b)$, and by the Jenssen inequality, any uncertainty about consumers' searching behavior improves the profit in comparison to the expected μ_k. Second, since a higher price dispersion described by higher $T_k(b)$ increases the fraction of searching consumers $\mu_k(b)$, we have that $\mu_k(b)$ and the profit from price variability $T_k(b)$ are positively correlated. Therefore, the second term in (2.9) is higher than the corresponding second term in (2.11). The next lemma summarizes this discussion and shows that the equilibrium profits are higher if information is exchanged publicly rather than privately.

Lemma 2.1. *Irrespectively of whether firm $i \in K$ or $i \in N \setminus K$, it is better off if information sharing is public rather than private: $\Pi_1^R(k) > \Pi_0^R(k)$ and $\Pi_1^C(k) > \Pi_0^C(k)$ for all $k > 0$ (equal if $k = 0$).*

Second, we determine how the number of revealing firms in K influences the incentives of concealing firms to deviate and reveal their cost in the case of private information exchange. The next lemma shows that if a firm finds it optimal to reveal its costs when there are k revealing firms in the market, then it will do so when there are more than k revealing firms. In other words, the situation in which many firms join the TA and reveal their costs makes it more attractive for other firms to join the TA too.

Lemma 2.2. *If information sharing is performed privately, then the revelation is more profitable for higher number of revealing firms k: $\Pi_0^R(k+1) - \Pi_0^C(k) > \Pi_0^R(k) - \Pi_0^C(k-1)$ for all $k = 0, 1, \ldots N-1$.*

The intuition goes as follows. According to proposition 2.2, if more firms decide to reveal information, then prices become less dispersed, so less μ_k consumers decide to search. We show that a change in the total fraction of informed consumers $\mu_k - \mu_{k+1}$ does not depend on k, when one additional firm reveals information. Importantly, the effective transportation cost t/μ increases more rapidly when μ is low. Therefore, the marginal gain from revelation that results in a decrease in market competitiveness is higher when more firms reveal information and search is already low.

The next proposition extends this lemma to the situation of public information revelation, in which consumers can observe the benchmark.

Lemma 2.3. *If information sharing is performed publicly, then the revelation is more profitable for higher number of revealing firms k:* $\Pi_1^R(k+1) - \Pi_1^C(k) > \Pi_1^R(k) - \Pi_1^C(k-1)$ *for all* $k = 0, 1, \ldots N-1$.

Proof. By lemma 2.1, we have that firms are better off if information is revealed publicly: $\Pi_1^R(k) > \Pi_0^R(k)$ and $\Pi_1^C(k) > \Pi_0^C(k)$ for all k. Using these inequalities, we have that

$$\Pi_1^R(k+1) + \Pi_1^C(k-1) > \Pi_0^R(k+1) + \Pi_0^C(k-1)$$

and

$$\Pi_1^C(k) + \Pi_1^R(k) > \Pi_0^C(k) + \Pi_0^R(k)$$

We subtract the second inequality from the first one and apply lemma 2.2:

$$\Pi_1^R(k+1) - \Pi_1^C(k) - (\Pi_1^R(k) - \Pi_1^C(k-1)) >$$
$$> \Pi_0^R(k+1) - \Pi_0^C(k) - (\Pi_0^R(k) - \Pi_0^C(k-1)) > 0.$$

□

This lemma establishes the fact that a higher number of revealing firms makes the revelation strategy more attractive for others. Similarly, it is important to explore how the fraction of informed consumers λ changes the profitability of the revelation strategy. The next lemma shows that if the fraction of shoppers is low, then firms have more incentives to reveal their cost.

Lemma 2.4. $\Pi_1^R(k+1) - \Pi_1^C(k)$ *is decreasing in* λ *for all* $k = 0, 1, \ldots N-1$.

In fact, the reasoning behind this lemma is quite similar to the intuition of lemma 2.3. If the fraction of informed consumers is low, then there are less consumers who decide to search. Since the marginal reduction of searching consumers softens competition more significantly when $\mu_k(b)$ is low, then the revelation becomes more attractive in markets with a lower fraction of λ.

So far we showed that,everything else being equal, firms have more incentives to contribute to the benchmark if a higher number of firms reveal their costs and the fraction of informed consumers is low. Denote λ_k as the fraction of informed consumers at which firms in $N \setminus K$ are indifferent between revealing and concealing if k firms decide to contribute to the benchmark. We define λ_k as the solution of the following equation:

$$\Pi_1^R(k+1,\lambda) - \Pi_1^C(k,\lambda) = 0.$$

We do not know yet for which parameters of the model λ_k exists. We will come to this question at the end of this section. Now we continue by assuming that there exists λ_k for some k. We define $\tilde{k}$ as the maximum number of revealing firms which find it optimal to deviate to concealment for all λ. Therefore, a group consisting of $\tilde{k}+1 < N$ firms will keep revealing their costs if λ is sufficiently low. We describe $\tilde{k}$ more formally in the next definition.

Definition 2.2. *Suppose that $\tilde{k} \in K$ is a number of revealing firms such that i) $\tilde{k}+1$th firm strictly prefers to reveal for some λ, so $\exists \tilde{\lambda}$ $\Pi_1^R(\tilde{k}+1,\tilde{\lambda}) > \Pi_1^C(\tilde{k},\tilde{\lambda})$ and ii) if $\tilde{k} \geq 1$ each firm in K prefers to conceal for all λ: $\Pi_1^R(\tilde{k}) < \Pi_1^C(\tilde{k}-1)$.*

By proposition 2.3 firms prefer to conceal their costs if $\lambda = 1$: $\Pi_1^R(\tilde{k}+1,1) < \Pi_1^C(\tilde{k},1)$. If $k = \tilde{k}+1$ and λ is sufficiently low, then firms in K will not deviate to concealment $\Pi_1^R(\tilde{k}+1,\lambda) > \Pi_1^C(\tilde{k},\lambda)$. Since the gain from revelation is smaller for higher λ (lemma 2.4), then by the fixed point argument there exists the unique $\lambda_{\tilde{k}}$ such that firms will reveal if $\lambda \leq \lambda_{\tilde{k}}$ and conceal otherwise. Recall, that firms find it more profitable to reveal their costs for higher k. Therefore, if $\lambda < \lambda_{\tilde{k}}$ and it is optimal to join the TA consisting of $\tilde{k}$ firms, then it is optimal to do so if the TA consists of larger number of firms $k > \tilde{k}$.

Proposition 2.4. *For all $k \geq \tilde{k}$, $\lambda_k \in (0,1)$ are uniquely determined and*
i) λ_k increases in k: $\lambda_{\tilde{k}} < \lambda_{\tilde{k}+1} < \ldots < \lambda_N$
ii) λ_k increases in $\overline{s}$: $\partial\lambda_k/\partial\overline{s} > 0$.

Since both larger k and lower λ make the revelation decision more attractive, then firms are indifferent between revelation and concealment if there is a low fraction of informed consumers and few revealing firms or there is large k and high λ. Therefore λ_k must increase in k. Second, when search costs represented by the upper bound $\overline{s}$ are high and less consumer decide to search, the market becomes more competitive. Thus, to make firms indifferent in their revelation decisions the

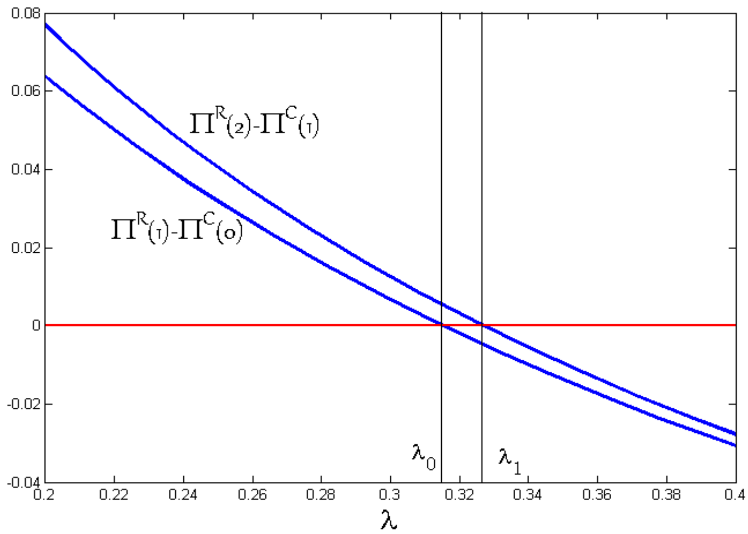


Figure 2.1: The difference in profits from revelation and concealment: $N = 2$, $\sigma_c^2 = 3$, $t = 2$, $s \sim U[0, 4]$.

competition must be intensified, so the fraction of informed consumers has to be higher.

Figure 1 illustrates the proposition for $N = 2$. There exist λ_0 and λ_1 such that firms are indifferent between revealing and concealing for $k = 1$ and $k = 2$ consequently. The upper blue line represent the change in profits when $k = 1$ and the second firm reveals it costs: $\Pi_1^R(2) - \Pi_1^R(1)$. The lower blue curve represents the change in profits of a firm that join the TA if the other firm conceals. We can see from the picture that it becomes more profitable to reveal information if there is a low fraction of informed consumers and more firms decide to join the trade association.

Based on proposition 2.4 we are ready to characterize the equilibrium in the benchmark formation game. Consider the case in which $\tilde{k} = 0$ and λ_0 exists as in Figure 1. If $\lambda < \lambda_0$, firms profits are highly dependent on the search strategy of non-shoppers. Therefore, firms will find it optimal to contribute to the benchmark and decrease the search activity of non-shoppers. Less active search allows firms to set higher prices because more consumers buy without comparing the products. Since the majority of consumers are non-shoppers, the losses from shoppers that we considered in the section about no search friction, are compensated by higher profits from non-shoppers. In this case, firms will follow the dominant strategy and reveal their costs irrespectively of k.

On the contrary, if the fraction of informed consumers is high and $\lambda > \lambda_N$, then firms will value shoppers more when deciding about revelation strategies. In this case, revelation would increase prices not enough to compensate the shrink of price dispersion. Since prices become less dispersed, firms earn less from shoppers which is the main source of their profits in this case. Therefore, firms will conceal information if the fraction of informed consumers is significantly high.

In the intermediate case when $\lambda_k < \lambda \leq \lambda_{k+1}$ firms play a coordination game in which there are two pure strategy equilibria and one equilibrium is in mixed strategies. Importantly, we show that in this region firms find it more profitable to coordinate on revelation rather than concealment. If $\tilde{k} > 0$ then the intuition remains the same, but in this case, there is no region where revelation is the only equilibrium. The following proposition summarizes the above discussion:

Proposition 2.5. *Consider the benchmark formation game in which firms decide whether to reveal their costs or not.*

1. If $\tilde{k} = 0$ (high $\overline{s}$), then i) if $\lambda \leq \lambda_0$ then all firms reveal their costs, $k^* = N$. ii) If $\lambda_0 < \lambda \leq \lambda_N$ then there are two pure strategy equilibria, $k_1^* = 0$ and $k_2^* = N$. In this case, firms are better off by coordinating on revelation: $\Pi_1^R(N) > \Pi_1^C(0)$. iii) If $\lambda > \lambda_N$ then all firms conceal their costs, $k^* = 0$.

2. If $\tilde{k} > 0$ (intermediate $\overline{s}$), then j) if $\lambda \leq \lambda_N$ then there are two pure strategy equilibria, $k_1^* = 0$ and $k_2^* = N$. In this case, firms are better off by coordinating on revelation: $\Pi_1^R(N) > \Pi_1^C(0)$. jj) If $\lambda > \lambda_N$ then all firms conceal their costs, $k^* = 0$.

3. If $\nexists \tilde{k}$ (low $\overline{s}$) described by definition 2.2, then firms will conceal their costs, $k^* = 0$.

It remains to be explored how other parameters of the model affect the decision to reveal information. The following proposition establishes the fact that if search costs are sufficiently high, then there exists an equilibrium with information revelation for sufficiently low λ.

Proposition 2.6. $\lambda_0 > 0$ *exists and firms reveal their costs for* $\lambda < \lambda_0$ *if search*

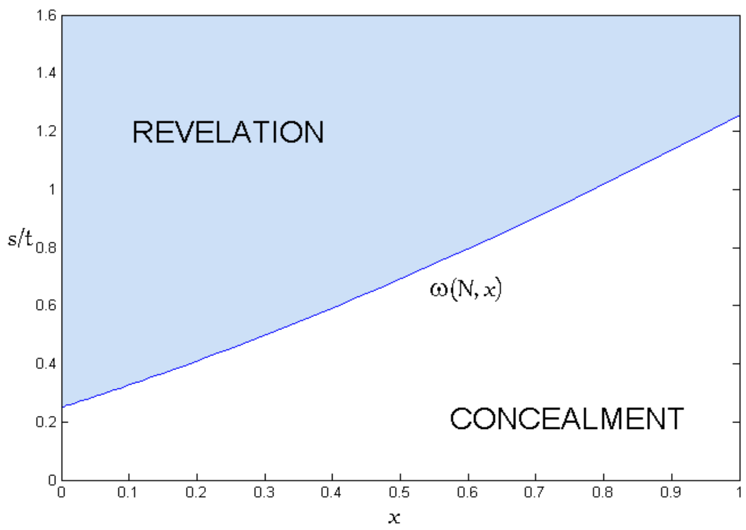


Figure 2.2: Incentives to deviate from no sharing for small λ: $N = 2$, $s \sim U[0, \overline{s}]$.

costs are high enough: $\overline{s}/t > \omega(N, \sigma_c^2/4t^2)$*, where*

$$\omega^2(N, x) = \left(\frac{x}{2} + N\left(\frac{1}{4} + z(N)x\right)\right)\left(\frac{1}{2} + x\right)\left(\frac{1}{4} + z(N)x\right)$$

and

$$z(N) = \frac{1}{2}\left(1 - \left(1 - \left(\frac{2N-2}{2N-1}\right)^2 / N\right)\right).$$

Figure 2 illustrate proposition 2.6. Firms find it optimal to deviate from the no benchmark case and reveal their costs if the relative search costs $\overline{s}/t$ are high and/or the relative variance of production costs $x = \sigma_c^2/4t^2$ is low. Intuitively, if the variance of costs is high and consumers search more, then the revelation can significantly deter search if and only if search costs are high. On the contrary, if the variance of production cost is low and consumer search less, firms find it optimal to reveal even if search is cheap. In the shaded region of parameters firms prefer to reveal information for sufficiently low λ. We can see that everything else equal, a higher search friction is more attractive for benchmark creation.

2.5 Welfare Analysis

In this section we explore the impact of information exchange on the total welfare. We show that if search costs are high, then the information sharing about private costs leads to a reduction of the total welfare. Since the revelation decision are made before the realization of costs, it is more natural to consider the welfare from an ex ante perspective.

Consider a consumer ℓ who decided to buy product i and has to pay p_i. Clearly, the price does not affect the total surplus and simply divides the gains from trade between firm i and consumer ℓ. The gains from trade can be written in the following form:

$$W_{\ell i} = v - [a_\ell s_\ell + tx_{\ell i} + c_i] \tag{2.12}$$

, where a_ℓ equals 1 if the consumer decided to search and 0 otherwise. From 2.12, there are three sources of inefficiency. First, the search friction does not allow consumers to buy the most preferred product and they have to incur search costs. Second, inefficiency comes from the preference mismatch. Consumers with higher search costs find it optimal not to search and end up buying a less preferred product incurring higher transportation costs. Third, consumers who do not search might buy products that are produced by less efficient firms. We will refer to it as a cost inefficiency.

Notice, that the market without search friction has the maximal possible total welfare for a given production costs distribution. In this case, consumers will always compare all products and minimize the effective price $tx_{\ell i} + p_i$. Since less efficient firms set higher prices, then consumers' purchasing strategies minimize the matching and cost inefficiencies $tx_{\ell i} + c_i$ and therefore total welfare is maximized.[18] Consequently, there is no reduction of the gains from trade on the side of informed consumers. Therefore, we have to characterize the search strategy of *non-shoppers* that maximizes the social welfare. As it was pointed out earlier, the social maximizer cares about firms' costs rather than prices. The ex ante gain of search of the social maximizer that incurs search cost s_ℓ is given by:

$$s_\ell \leq \tilde{s}^* = \mathbb{E}[c_i + tx_i] - \mathbb{E}[\min\{c_i + tx_i, c_j + t(1 - x_i)\}]$$

[18] In Duffie et al(2017), benchmarks encourage entry by consumers and increase realized gains from trade. In our welfare analysis we do not consider entry decision, rather we focus on how benchmarks change the welfare through the impact on the search behaviour of consumers.

which can be rearranged as in the case of consumers' optimal search strategy (2.6)

$$\tilde{s}_k^* = t\mathbb{E}_c[\overline{x}_{ji}^2] = t\mathbb{E}_c\left(\frac{1}{2} + \frac{c_i - c_j}{2t}\right)^2 = \frac{t}{4} + \frac{1}{4t}\mathbb{E}(c_i - c_j)^2 \tag{2.13}$$

The total welfare attains its maximum when the incurred search costs are counterbalanced with the improved preference match as well as the increased cost efficiency. If consumers are pushed to search too much, then the generated gains from better information will not recover the search costs. Otherwise, if consumer are not active enough, their search costs savings do not compensate for the foregone reduction in mismatch and cost inefficiencies. According to proposition 2.1, firms pass through their costs at a rate lower than 1. Suppose k firms reveal their costs then:

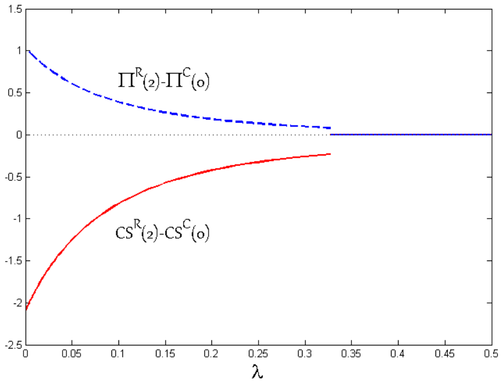
$$\tilde{s}^* > \frac{t}{4} + \frac{1}{4t}\mathbb{E}_k(p_i - p_j)^2 = \tilde{s}_k$$

This inequality establishes the fact that in equilibrium consumers search less than needed to maximize the total welfare. The variance effect of information exchange diminishes the search activity even further away from the optimal level $\tilde{s}^*$, reducing the total welfare: $\tilde{s}^* > \tilde{s}_0 > \tilde{s}_N$. The previous section established the fact that information is exchanged if the fraction of shoppers is sufficiently low: $\lambda < \lambda_N$. Therefore, the welfare is the lowest in our model, when the search friction is high because i) a higher fraction of consumers incurs search costs and ii) firms exchange information making consumers search too little.

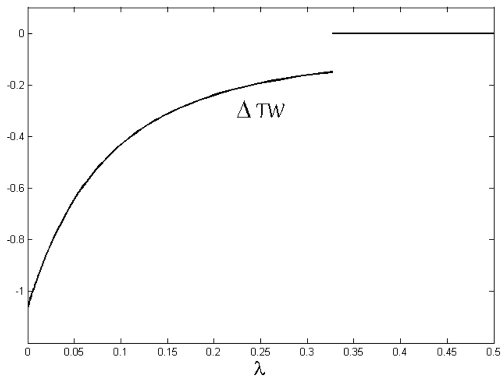
It is important to notice that the reduction of welfare is paid by consumers. If k firms reveal their costs, then the ex ante consumer surplus is given by:

$$CS_k = v - \mathbb{E}[p + tx] + \left\{\lambda\tilde{s}_k + (1-\lambda)\int_0^{\tilde{s}_k}(\tilde{s}_k - s)ds\right\} \tag{2.14}$$

The informed consumers λ buy at the lowest price which is equivalent to saying that they buy at the expected price corrected for the expected gain of search: $\mathbb{E}[p + tx] - \tilde{s}_k$. Uninformed consumers with sufficiently low search costs $s_\ell < \tilde{s}_k$ decide to search and save less than shoppers: $\tilde{s}_k - s_\ell$. Those who decided not to search pay the expected price. It is easy to show that (2.14) is an increasing function in $\tilde{s}_k$, since the savings of search costs are higher. Firms set higher prices and consumers search less when the benchmark is formed, so we can deduce that consumer surplus is higher if firms conceal information: $CS_0 > CS_N$. We showed



(a) The impact of cost sharing on profits and CS



(b) The impact of cost sharing on welfare

Figure 2.3: Welfare loss from Information Sharing: $N = 2$, $\sigma_c^2 = 3$, $t = 2$, $s \sim U[0, 4]$

in the previous section, that prices and profits are higher when firms exchange information, $\Pi_1(N) > \Pi_1(0)$ for $\lambda < \lambda_N$, therefore the reduction of total welfare is paid by consumers. When information is exchanged, the welfare is transferred from consumers to firms with subsequent dead weight losses. The following proposition summarizes the discussion:

Proposition 2.7. *The equilibrium in which firms form the benchmark yields higher profits of firms but reduces consumer surplus as well as total welfare.*

This propositions is illustrated in figures 3(a) and 3(b). We compare the equilibrium profits and consumer surplus to the case where the formation of benchmarks is not possible and firms always conceal information. According to proposition 2.5, there is $\tilde{\lambda}$ such that for all $\lambda < \tilde{\lambda}$ firms will reveal information. Therefore, the difference in welfare can be observed only for a small fraction of informed consumers. From figure 3(a), we can see that revelation increases profits of firms and decreases consumer surplus. Importantly, the consumers surplus decrease is not compensated by higher profits of firms and this is depicted in figure 3(b): $\Delta TW < 0$.

2.6 Conclusions

Many benchmarks are used as the basis for price setting in various financial and nonfinancial markets, such as OTC markets, in which consumer search is of great importance. Benchmarks and/or industry reports are often formed based on the private information of firms that are collected and disseminated by third parties. This form of information sharing affects the level of uncertainty in markets, which, in turn, has an impact on firms' pricing and consumers' search behavior. In this paper, we explore the impact of benchmarks on competition and total welfare and inquire into the incentives to contribute to benchmark formation. Information sharing improves firms' information regarding competitors' costs and prices become more dependent on benchmarks. We show that this leads to price alignment and consumers have less incentive to search because they believe that a significantly lower price is less likely to be found. Firms find it optimal to contribute to a benchmark when search friction is significantly strong, so that information sharing maximally deters the search of consumers. Although firms' profits increase, we find that benchmarks reduce total welfare. As a result, consumers search less and buy less desired products from inefficient firms.

In a recent paper, Duffie *et al.* (2017) show that benchmarks improve welfare as they encourage entry by consumers and intensify competition. They do not consider information sharing of firms and their result is based on consumers learning about costs. If consumers know firms' costs, firms cannot manipulate consumers' beliefs and cannot set a high price because consumers will attribute it to a higher margin and continue to search. In our paper, we uncover another important impact of benchmarks. Precisely, benchmarks facilitate price alignment due to improved information on the firms' side. This unambiguously deters search and reduces total welfare. It is important to investigate the interplay of these two impacts of benchmarks and consider a general model that would incorporate consumers' learning from prices as well as information sharing between firms.

From a more general perspective, in this paper we show that the presence of players who are not involved in communication plays an important role. In their paper, Okuno-Fujiwara *et al.* (1990) explore strategic information revelation, and provide sufficient conditions for information revelation. They consider a Bayesian game where players can publicly reveal their type and then take actions. However, there are often some players who are not involved in communication, but their actions influence the payoffs of everybody else. Therefore, when a player reveals her information, she must take into account both the reactions of those players who receive it and of those who do not, but believe or observe that information transmission takes place. Thus, it would be interesting to examine a general model of strategic information revelation in the presence of players who do not communicate.

2.7 Appendix

Lemma 2.5. *Consider the vectors of random variables X_i and functions h_i which satisfy the following equation for all $i = 1, \ldots, N$:*

$$2h_i(X_i) = \frac{1}{N-1} \sum_{j \neq i} \mathbb{E}[h_j(X_j)|X_i] \tag{2.15}$$

Then $h_i(X_i) = 0$ almost surely for all i.

Proof. By taking the expectation of (2.15) conditional on $h_i(X_i)$ and using the tower property of conditional expectation, we get:

$$2h_i(X_i) = \frac{1}{N-1} \sum_{j \neq i} \mathbb{E}[\mathbb{E}[h_j(X_j)|X_i]|h_i(X_i)] = \frac{1}{N-1} \sum_{j \neq i} \mathbb{E}[h_j(X_j)|h_i(X_i)]$$

Notice that now the equation (2.15) depends only on the functions $h_i(X_i)$, so we can denote $h_i(X_i)$ as Z_i and rewrite:

$$2Z_i = \frac{1}{N-1} \sum_{j \neq i} \mathbb{E}[Z_j|Z_i] \tag{2.16}$$

Next, we multiply both sides of (2.16) by Z_i, take the expectation using the law of iterated expectations and add $\mathbb{E}[Z_i^2]/(N-1)$ to both sides:

$$\frac{2N-1}{N-1}\mathbb{E}[Z_i^2] = \frac{1}{N-1} \sum_{j \neq i} \Big(\mathbb{E}[Z_i \mathbb{E}[Z_j|Z_i]] + \mathbb{E}[Z_i^2]\Big) = \frac{1}{N-1} \sum_{j=1}^{N} \mathbb{E}[Z_i Z_j]$$

Applying the Cauchy-Schwarz inequality to the right-hand side of the above equation and subtracting the left hand side from it, we get that the following inequality holds for all i:

$$\sqrt{\mathbb{E}[Z_i^2]} \left(\frac{1}{N-1} \sum_{j=1}^{N} \sqrt{\mathbb{E}[Z_j^2]} - \frac{2N-1}{N-1} \sqrt{\mathbb{E}[Z_i^2]} \right) \geq 0 \tag{2.17}$$

Without loss of generality, let's assume that $\mathbb{E}[Z_i^2] > 0$, so the second multiplier of (2.17) is non-negative for all i. Summing up for all i we get that

$$\frac{N}{N-1} \sum_{j=1}^{N} \sqrt{\mathbb{E}[Z_j^2]} - \frac{2N-1}{N-1} \sum_{i=1}^{N} \sqrt{\mathbb{E}[Z_i^2]} = -\sum_{i=1}^{N} \sqrt{\mathbb{E}[Z_i^2]} \geq 0$$

which holds if and only if $\mathbb{E}[(h_i(X_i))^2] = 0$. Therefore, $h_i(X_i) = 0$ almost surely for all $i = 1, \ldots, N$. □

Proposition 2.1. *Given any outcome of the benchmark formation in the first stage $\{b, k\}$, there is a unique Bayesian Nash Best Response of firms to any search behavior described by $\mu(b)$. Depending on whether firm i revealed or concealed its costs, it sets the price according to the linear pricing strategies given by:*

$$p_R(c_i) = \frac{t}{\mu(b)} + \frac{N-k}{2N-1}c^e + \frac{k}{2N-1}b + \frac{N-1}{2N-1}c_i$$
$$p_C(c_i) = \frac{t}{\mu(b)} + \left(\frac{1}{2} - \frac{k}{2N-1}\right)c^e + \frac{k}{2N-1}b + \frac{1}{2}c_i$$

Proof. Denote the resulting transportation costs as $t_1 = t/\mu(b)$. The profit function of firm i, conditional on the information it has $I_i = \{b, k\} \cup c_i$, is

$$\mathbb{E}_{c_j}[\pi_i(p_i)|I_i] = \frac{1}{t_1}(p_i - c_i)\left(t_1 + \frac{1}{N-1}\sum_{j\neq i}\mathbb{E}[p_j|I_i] - p_i\right) \tag{2.18}$$

By taking the FOC of (2.18) we get that optimal p_i satisfies:

$$2p_i = t_1 + c_i + \frac{1}{N-1}\sum_{j\neq i}\mathbb{E}[p_j|I_i]$$

Suppose for now that firms in K follow the symmetric pricing strategy $p_R(c)$ and firms $N \setminus K$ follow symmetric pricing strategy $p_C(c)$. It is clear, that if a firm j didn't reveal its cost, $j \in N \setminus K$, then the cost of firm i, does not contain anything about j's price, so $\mathbb{E}[p_j|I_i] = \mathbb{E}[p_C|b]$. Moreover, since I_i does not allow to distinguish between any two firms in K, we must have that any two expected prices of firms $j_1, j_2 \in K$ are the same: $\mathbb{E}[p_{j_1}|I_i] = \mathbb{E}[p_{j_2}|I_i] = \mathbb{E}[p_R|I_i]$. So, for the symmetric equilibrium we can rewrite FOC in the following way:

$$2p_i = t_1 + c_i + \frac{1}{N-1}\sum_{j\in N\setminus K, j\neq i}\mathbb{E}[p_C|b] + \frac{1}{N-1}\sum_{j\in K, j\neq i}\mathbb{E}[p_R|I_i] \tag{2.19}$$

The equilibrium candidate is in linear strategies and given by (2.4):

$$\begin{cases} p_R(c_i) = \alpha_R + \beta_R b + \gamma_R c_i, \text{ if } i \in K \\ p_C(c_i) = \alpha_C + \beta_C b + \gamma_C c_i, \text{ if } i \in N \setminus K \end{cases}$$

We will verify that there are uniquely determined coefficients of p_R and p_C constituting the symmetric equilibrium in linear strategies. Later on, we show that this problem cannot have more than one equilibrium.

So, to compute FOC, we substitute the suggested linear strategies given by (2.4) into (2.19). First, suppose firm $i \in K$. The expected price of concealing and revealing firms are given by $\mathbb{E}[p_C|b] = \alpha_C + \beta_C b + \gamma_C c^e$ and $\mathbb{E}[p_R|I_i] = \alpha_R + \beta_R b + \gamma_R \mathbb{E}[c_j|I_i]$ respectively. Let we compute the expected costs of revealing firm j given I_i. Take the sum of $\mathbb{E}[c_j|I_i]$ for all $j \in K$:

$$\sum_{j \in K} \mathbb{E}[c_j|I_i] = \mathbb{E}[kb|I_i] = kb$$

On the other hand, since the strategies are symmetric for any $j_1, j_2 \in K$ we have that $\mathbb{E}[c_{j_1}|I_i] = \mathbb{E}[c_{j_2}|I_i]$ and $\sum_{j \in K} \mathbb{E}[c_j|I_i] = c_i + (k-1)\mathbb{E}[c_j|I_i]$. Therefore, it is natural that the expected costs of competitors is simply the average of $k-1$ costs without taken into account the cost of firm i:

$$\mathbb{E}[c_j|I_i] = \frac{kb - c_i}{k-1}$$

The FOC takes the following form:

$$2p_R(c_i) = t_1 + c_i + \frac{N-k}{N-1}\left(\alpha_C + \beta_C + \gamma_C c^e\right) + \frac{k-1}{N-1}\left(\alpha_R + \beta b + \gamma_R \frac{kb - c_i}{k-1}\right)$$

Rearranging to the linear form given by (2.4), we get first three equation for the coefficients:

$$\begin{aligned} p_R(c_i) = {} & \frac{1}{2}\left(t_1 + \frac{N-k}{N-1}(\alpha_C + \gamma_C c^e) + \frac{k-1}{N-1}\alpha_R\right) + \\ & + \frac{1}{2}\left(\frac{N-k}{N-1}\beta_C + \frac{k-1}{N-1}\beta_R + \frac{k}{N-1}\gamma_R\right) b + \frac{1}{2}\left(1 - \frac{\gamma_R}{N-1}\right) c_i = \\ & = \alpha_R + \beta_R b + \gamma_R c_i \end{aligned}$$

Second, consider a firm $j \in N \setminus K$. This firms has the same expectation about the price of concealing firms $\mathbb{E}[p_C|b] = \alpha_C + \beta_C b + \gamma_C c^e$. Since j didn't contribute to b, its estimation of p_R will be less precise: $\mathbb{E}[c_{i \in K}|I_j] = b$ and so $\mathbb{E}[p_R|I_i] = \alpha_R + (\beta_R + \gamma_R)b$. The FOC of firm j is:

$$2p_C(c_j) = t_1 + c_j + \frac{N-k-1}{N-1}(\alpha_C + \beta_C b + \gamma_C c^e) + \frac{k}{N-1}(\alpha_R + (\beta_R + \gamma_R)b)$$

Rearranging to the linear form given by (2.4), we obtain

$$\begin{aligned} p_C(c_i) = {} & \frac{1}{2}\left(t_1 + \frac{N-k-1}{N-1}(\alpha_C + \gamma_C c^e) + \frac{k}{N-1}\alpha_R\right) + \\ & + \frac{1}{2}\left(\frac{N-k-1}{N-1}\beta_C + \frac{k}{N-1}(\beta_R + \gamma_R)\right) b + \frac{1}{2}c_j = \alpha_C + \beta_C b + \gamma_C c_j \end{aligned}$$

The two FOCs for firms $i \in K$ and $j \in N \setminus K$ result in the system of linear equation with 6 unknowns. The solution of this system is:

$$\alpha_R = t_1 + \frac{N-k}{2N-1}c^e \qquad \beta_R = \frac{k}{2N-1} \qquad \gamma_R = \frac{N-1}{2N-1}$$
$$\alpha_C = t_1 + \left(\frac{1}{2} - \frac{k}{2N-1}\right)c^e \qquad \beta_C = \frac{k}{2N-1} \qquad \gamma_C = \frac{1}{2}$$

Uniqueness. Denote $p^* = (p_R, p_C)$ as an equilibrium pricing strategy. We want to show that there does not exist any other Bayesian equilibrium for any given I_i and μ. Suppose that there is another one equilibrium profile of pricing strategies $\tilde{p}$, possible asymmetric. Then according to the FOC it must hold that for any firm i:

$$2\tilde{p}_i = t_1 + c_i + \frac{1}{N-1}\sum_{j \neq i} \mathbb{E}[\tilde{p}_j | I_i]$$

Subtracting the FOC of p^* from this equation we get:

$$2(\tilde{p}_i(I_i) - p_i^*(I_i)) = \frac{1}{N-1}\sum_{j \neq i} \mathbb{E}[\tilde{p}_j(I_j) - p_j^*(I_j) | I_i]$$

We apply *lemma 2.5* to this equation by setting $h_i(X_i) = \tilde{p}_i(X_i) - p_i^*(X_i)$, where $X_i = I_i$ and get that $h_i(X_i) = 0$ for all $i = 1, \ldots, N$:

$$\tilde{p}_i = p_i^*$$

Therefore, this problem has the unique equilibrium for every μ that is characterized by the linear pricing strategies p_R and p_C. □

Proposition 2.2. *The ex ante equilibrium fraction of informed consumers is lower for a higher number of revealing consumers k: $\tilde{s}_k > \tilde{s}_{k+1}$ and $\mu_k > \mu_{k+1}$.*

Proof. The expected fraction of informed consumers μ_k depends on the expected search cost of the indifferent consumer $\tilde{s}_k = \mathbb{E}\tilde{s}_k(b)$:

$$\tilde{s}_k = t\mathbb{E}_p\left(\frac{1}{2} + \frac{p_i - p_j}{2t}\right)^2 = \frac{t}{4} + \frac{1}{4t}\mathbb{E}_p(p_i - p_j)^2 \tag{2.20}$$

Since consumers do not know which firms revealed information they consider possible three situation. Consider a consumer who is interested in products i and j. Denote the probability that two (both i and j), one (either i or j) or none firms revealed their cost as $q_2(k)$, $q_1(k)$ and $q_0(k)$ correspondingly. Denote the variance of the price difference as $z_m = E_m(p_i - p_j)^2$ if $m = \{0, 1, 2\}$ firms revealed costs. According to the corollary 2.1, z_m decreases in m since prices become less dispersed. Consider the last term in (2.20):

$$\mathbb{E}_p(p_i - p_j)^2 = \sum_{m=0}^{2} q_m(k) z_m$$

such that $\sum_{m=0}^{2} q_m(k) = 1$ for any k. Comparing the difference we get:

$$\tilde{s}_{k+1} - \tilde{s}_k = \sum_{m=0}^{2} (q_m(k+1) - q_m(k)) z_m < (q_0(k+1) - q_0(k)) z_0 +$$
$$+ z_1 \sum_{m=1}^{2} (q_m(k+1) - q_m(k)) = \underbrace{(z_0 - z_1)}_{>0} \underbrace{(q_0(k+1) - q_0(k))}_{<0} < 0$$

The first inequality comes from the fact that the probability of dealing with two revealing firms $q_2(k) = C_k^2 / C_N^2$ is higher for higher k and $z_1 > z_2$. Then we use the fact the sum of probabilities $q_0 + q_1 + q_2 = 1$ for all k and since the chance to deal with two concealing firms $q_0(k) = C_{N-k}^2 / C_N^2$ gets lower for higher k, we get the last inequality.

Since $\mu_{k+1} - \mu_k = (1 - \lambda)(\tilde{s}_{k+1} - \tilde{s}_k)/\overline{s}$ we have that μ_k decreases in k as well. □

Proposition 2.3. *It is optimal to conceal private costs in the market without search friction* ($\lambda = 1$)*:* $\Pi_1^C(k) > \Pi_1^R(k+1)$ *for all* $k = 0, 1, \ldots N - 1$.

Proof. Since $\lambda = 1$, all consumers have zero search costs and buy after inspecting two firms, $\mu = 1$. Importantly, the learning of consumers does not take place, so private and public information exchange regimes coincide and will proceed without indexing the equilibrium profit functions: $\Pi_1^R(k) = \Pi_0^R(k) = \Pi^R(k)$. We start from the situation in which k firms decided to reveal their costs and form the revealing group K.

The equilibrium profit (multiplied by t) of firm $i \in N \setminus K$ that reveals its costs, is given by:

$$t\Pi^R(k+1) =$$
$$= \mathbb{E}(p_R(c_i) - c_i)^2 = t^{-1}\mathbb{E}\left(t + \frac{N-k-1}{2N-1}(c^e - c_i) + \frac{k+1}{2N-1}(b_{k+1} - c_i)\right)^2 =$$

By taking into account that $\mathbb{E}[c^e - c_i] = \mathbb{E}[b_{k+1} - c_i] = 0$, we can rearrange the last expression in the following way:

$$= t^2 + \left(\frac{N-k-1}{2N-1}\right)^2 \sigma_c^2 + 2\frac{(N-k-1)(k+1)}{(2N-1)^2}\mathbb{E}(c^e - c_i)(b_{k+1} - c_i) + \\ + \left(\frac{k+1}{2N-1}\right)^2 \mathbb{E}(b_{k+1} - c_i)^2 \tag{2.21}$$

To compute (2.21) we have to separately consider the second and the third terms. In the derivation of the second term we use the fact that the costs are distributed independently:

$$\mathbb{E}(c^e - c_i)(b_{k+1} - c_i) = \frac{1}{k+1}\sum_{j\neq i}\mathbb{E}(c^e - c_i)(c_j - c^e) + \frac{1}{k+1}\sum_{j\neq i}\mathbb{E}(c^e - c_i)^2 = \frac{k}{k+1}\sigma_c^2$$

Next we compute the third term of (2.21) by substituting the expression for b_{k+1} and simplifying of the square expression:

$$\mathbb{E}(b_{k+1} - c_i)^2 =$$

$$= \frac{1}{(k+1)^2}\mathbb{E}\left(\sum_{j\in K}(c_j - c_i)\right)^2 = \frac{1}{(k+1)^2}\left(\sum_{j\in K}\mathbb{E}(c_j - c_i)^2 + 2\sum_{\substack{j,z\in K\\ j\neq z}}\mathbb{E}(c_j - c_i)(c_z - c_i)\right)$$

It is easy to check that $\mathbb{E}(c_j - c_i)^2 = 2\sigma_c^2$ and $\mathbb{E}(c_j - c_i)(c_z - c_i) = \sigma_c^2$. Since there are k firms in K and C_k^2 combination of two randomly taken firms from K, we get the expression for the third term:

$$\mathbb{E}(b_{k+1} - c_i)^2 = \frac{1}{(k+1)^2}(2\sigma_c^2 k + 2C_k^2\sigma_c^2) = \frac{k}{k+1}\sigma_c^2$$

We notice that the second and the third terms coincide and substitute them back to (2.21) and get a simple expression for the equilibrium profit:

$$t\Pi^R(k+1) =$$

$$= t^2 + \frac{(N-k-1)^2 + 2(N-k-1)k + (k+1)k}{(2N-1)^2}\sigma_c^2 = t^2 + \left(\left(\frac{N-1}{2N-1}\right)^2 + \frac{k}{(2n-1)^2}\right)\sigma_c^2$$

Now consider the equilibrium profit (multiplied by t) of concealing firm $j \in N\backslash K$

$$t\Pi^C(k) = \mathbb{E}(p_C(c_j) - c_j)^2 = \mathbb{E}\left(t + \frac{k}{2N-1}(b_k - c_j) + \frac{1}{2}(c^e - c_j)\right)^2$$

Notice that firm j does not reveal its cost, therefore $\mathbb{E}(b_k - c_j)^2 = \sigma_c^2/k$. By simplifying we get:

$$t\Pi^C(k) = t^2 + \left(\frac{1}{4} + \frac{k}{(2N-1)^2}\right)\sigma_c^2 \tag{2.22}$$

We prove the proposition by comparing the equilibrium profits:

$$\Pi^C(k) - \Pi^R(k+1) = \left(1 - \left(\frac{2N-2}{2N-1}\right)^2\right)\frac{\sigma_c^2}{4t} > 0$$

□

Lemma 2.6. *For all $k < N$, the difference $\Delta\mu = \mu_k - \mu_{k+1}$ does not depend on k.*

Proof. To show that $\Delta\mu$ does not depend on k, it suffice to show that the difference in the search costs of indifferent consumers does not depend on k: $\Delta\mu = (1-\lambda)(\tilde{s}_k - \tilde{s}_{k+1})/\overline{s}$. Denote q_m as the probability that $m \in \{0,1,2\}$ firms revealed their costs consequently. Expected gain of search is the expected effective price difference given this weights:

$$\tilde{s}_k/t = q_2\mathbb{E}\left(\frac{1}{2} + \frac{N-1}{2N-1}(c_i - c_j)/2t\right)^2 +$$
$$+ q_1\mathbb{E}\left(\frac{1}{2} + \frac{1}{2}(c_i - c^e)/2t + + \frac{N-1}{2N-1}(c^e - c_j)/2t\right)^2 + q_0\mathbb{E}\left(\frac{1}{2} + \frac{1}{2}(c_i - c_j)/2t\right)^2$$

Denote $\psi = \left(\frac{2N-2}{2N-1}\right)^2 < 1$. By computing the probabilities of these three events and taking the expectations, we get

$$\tilde{s}_k/t = \frac{1}{4} + \frac{1}{2N(N-1)}[k(k-1)\psi + k(N-k)(1+\psi) + (N-k)(N-k-1)]\frac{\sigma^2}{4t^2}$$

which can be simplified further and the final expression is as follows:

$$\tilde{s}_k = \frac{t}{4} + \frac{1}{2N}\left(N - k(1-\psi)\right)\frac{\sigma_c^2}{4t}$$

Next, we compute the difference between the search costs of indifferent consumers in case of k and $k+1$ numbers of revealing firms and see that it does not depend on k

$$\tilde{s}_k - \tilde{s}_{k+1} = \frac{1}{2N}(1-\psi)\frac{\sigma_c^2}{4t} > 0.$$

Therefore, $\Delta\mu$ does not depend on k either. □

Lemma 2.2. *If information sharing is performed privately, then the revelation is more profitable (less unprofitable) for higher number of revealing firms* k*:* $\Pi_0^R(k+1) - \Pi_0^C(k) > \Pi_0^R(k) - \Pi_0^C(k-1)$ *for all* $k = 0, 1, \dots N-1$.

Proof. The effect of information revelation on the profits can be split in two parts. First, information revelation increases the price level (square brackets) at which firms base their prices. Second, it changes the price variability which negatively affects profits (curly brackets). The profit difference can be written in the following form:

$$\Delta\Pi_0(k) = \Pi_0^R(k+1) - \Pi_0^C(k) = t\left[\frac{1}{\mu_{k+1}} - \frac{1}{\mu_k}\right] + \{\mu_{k+1}(\psi + \phi(k)) - \mu_k(1 + \phi(k))\}\frac{\sigma_c^2}{4t} \tag{2.23}$$

, where $\psi = \left(\frac{2N-2}{2N-1}\right)^2 < 1$ and $\phi(k) = \frac{4k}{(2N-1)^2}$. We want to show that the increase (decrease) in profits is larger (smaller) for higher k: $\Delta\Pi_0(k) > \Delta\Pi_0(k-1)$. We will prove separately that i) the positive effect of higher price basis described by the curly brackets is larger for higher k and ii) the negative effect of price variability is lower for higher k.

i)Using lemma (2.6) and the fact that $\mu_{k+1} < \mu_k < \mu_{k-1}$ we have that:

$$\left[\frac{1}{\mu_{k+1}} - \frac{1}{\mu_k}\right] - \left[\frac{1}{\mu_k} - \frac{1}{\mu_{k-1}}\right] = \frac{\Delta\mu}{\mu_{k+1}\mu_k\mu_{k-1}}(\mu_{k-1} - \mu_{k+1}) > 0$$

ii) Denote $V(k)$ as the size of the negative effect of price variability described by the curly brackets in (2.23). We rearrange the terms in $V(k)$ and use lemma (2.6):

$$V(k) = \mu_k(1 + \phi(k)) - \mu_{k+1}(\psi + \phi(k)) = \Delta\mu(\psi + \phi(k)) + \mu_k(1 - \psi)$$

To show that the negative effect is stronger for smaller k we have to show that $V(k)$ is decreasing in k. By comparing the difference we prove the proposition:

$$V(k-1) - V(k) = \Delta\mu(\phi(k-1) - \phi(k) + (1 - \psi)) > 0$$

since for any $N \geq 2$ we have that two multipliers are positive:

$$\phi(k-1) - \phi(k) + (1 - \psi) = -\frac{4}{(2N-1)^2} + 1 - \left(\frac{2N-2}{2N-1}\right)^2 = \frac{4N-7}{(2N-1)^2} > 0$$

□

Lemma 2.4. $\Pi_1^R(k+1) - \Pi_1^C(k)$ *is decreasing in* λ *for all* $k = 0, 1, \ldots N-1$.

Proof. To prove this lemma let's first show that the expected total fraction of consumers μ_k is increasing in λ faster for higher number of revealing firms k. Notice, that we can represent the derivative of μ_k' as the function of μ_k:

$$\mu_k' = 1 - \tilde{s}_k/\bar{s} = \frac{1-\mu_k}{1-\lambda}$$

Therefore, the sign of the difference between the derivatives is the opposite sign of the difference between the functions: $sign(\mu_{k+1}' - \mu_k') = sign(\mu_k - \mu_{k+1}) > 0$, since μ_k is decreasing in k. The equilibrium profit of firm i can be written in the following general form:

$$\Pi_1(k) = t\mathbb{E}_b\left[\frac{1}{\mu_k(b)} + \mu_k(b)T_k(b)\right]$$

, where $T_k(b) < 1$. Next, we show that the derivative of the profits difference is decreasing. Denote $z_1 = \min_b\{\mu_k^2(b), \mu_{k+1}^2(b)\}$ and $z_2 = \max\{T_{k+1}^R(b), T_k^C(b)\}$. It is clear that $z_1 < 1$ as long as $\lambda < 1$ and $z_2 < 1$ given that $T_k(b) < 1$ for all b and k.

$$\begin{aligned}\frac{\partial}{\partial\lambda}\left(\Pi_1^R(k+1) - \Pi_1^C(k)\right) = t\mathbb{E}\left[\frac{\mu_k'(b)}{\mu_k^2(b)} - \frac{\mu_{k+1}'(b)}{\mu_{k+1}^2(b)} + \mu_{k+1}'T_{k+1}^R(b) - \mu_k'T_k^C(b)\right] \leq \\ \leq t\mathbb{E}\left[\frac{\mu_k'(b) - \mu_{k+1}'(b)}{z_1} + (\mu_k'(b) - \mu_{k+1}'(b))z_2\right] < \\ < t\mathbb{E}(\mu_k'(b) - \mu_{k+1}'(b))(1 - 1/z_1) = t\,\underbrace{(\mu_k' - \mu_{k+1}')}_{>0}\,\underbrace{(1 - 1/z_1)}_{<0} < 0\end{aligned}$$

□

Lemma 2.1. *Irrespectively of whether firm* $i \in K$ *or* $i \in N \setminus K$*, it is better off if information sharing is public rather than private:* $\Pi_1^R(k) > \Pi_0^R(k)$ *and* $\Pi_1^C(k) > \Pi_0^C(k)$ *for all* $k > 0$ *(equal if* $k = 0$*).*

Proof. The equilibrium profit function in case of public information revelation can be represented as

$$\Pi_1(k) = t\mathbb{E}_b\left[\frac{1}{\mu_k(b)} + \mu_k(b)T_k(b)\right] \tag{2.24}$$

,where $T_k(b) < 1$ for all b. The equilibrium profit in the case of private revelation is given by:

$$\Pi_0(k) = t\left[\frac{1}{\mu_k} + \mu_k T_k\right] \tag{2.25}$$

,where $\mu_k = \mathbb{E}_b \mu_k(b)$ and $T_k = \mathbb{E}_b T_k(b)$ are the expectation taken before the benchmark is formed. By Jenssen inequality we have that: $\mathbb{E}\mu_k^{-1}(b) > (\mathbb{E}\mu_k(b))^{-1} = \mu_k^{-1}$. Second, since $\text{Cov}(\mu(b), T_k(b)) > 0$, we have that $\mathbb{E}_b[\mu_k(b)T_k(b)] > \mu_k T_k$. Since each of the term in (2.24) is higher than in (2.25) we get that $\Pi_1(k) > \Pi_0(k)$ for all $k > 0$.

□

Proposition 2.4. *For all* $k \geq \tilde{k}$, $\lambda_k \in (0, 1)$ *are uniquely determined and increase in* k*:*

$$\lambda_{\tilde{k}} < \lambda_{\tilde{k}+1} < \ldots < \lambda_N$$

Proof. Consider the situation under which $k > \tilde{k}$ firms decide to reveal their costs and the benchmark b_k is formed.

First, we prove that λ_k exists and unique. From the definition of $\tilde{k}$ there exists $\tilde{\lambda}$ such that $\tilde{k}+1$ firm find it optimal to contribute to the benchmark formation: $\Pi_1^R(\tilde{k}+1, \tilde{\lambda}) - \Pi_1^C(\tilde{k}, \tilde{\lambda}) > 0$. Since firms gain form revealing costs under the coalition of $\tilde{k}$ firms, we have that by proposition 2.3, it will be optimal to reveal costs for all coalitions including more than $\tilde{k}$ firms at $\lambda = \tilde{\lambda}$:

$$\Pi_1^R(k+1, \tilde{\lambda}) - \Pi_1^C(k, \tilde{\lambda}) > \Pi_1^R(\tilde{k}+1, \tilde{\lambda}) - \Pi_1^C(\tilde{k}, \tilde{\lambda}) > 0$$

Since $\Pi_1^R(k+1) - \Pi_1^C(k)$ is decreasing in λ by lemma 2.4, and firms conceal their costs when $\lambda = 1$ (by proposition 2.3) and decide to reveal when $\lambda = \tilde{\lambda} < 1$, then due to the fixed point argument we have that there exists unique $\lambda_k \in (\tilde{\lambda}, 1)$ such that firms are indifferent between revelation and concealment of their costs: $\Pi_1^R(k+1, \lambda_k) - \Pi_1^C(k, \lambda_k) = 0$.

Second, we show that λ_k is increasing in k. From the definition of λ_k and proposition 2.3 we have that it is optimal to reveal at $\lambda = \lambda_k$ if the coalition consists of $k+1$ firms:

$$\Pi_1^R(k+2, \lambda_k) - \Pi_1^C(k+1, \lambda_k) > \Pi_1^R(k+1, \lambda_k) - \Pi_1^C(k, \lambda_k) = 0$$

Applying the fixed point argument, we find that $\lambda_{k+1} \in (\lambda_k, 1)$ and therefore λ_k is an increasing function in k.

Second, we show that λ_k is an increasing function in $\overline{s}$. From the implicit function theorem we have that:

$$\frac{\partial \lambda_k}{\partial \overline{s}} = -\frac{\partial}{\partial \overline{s}}\left(\Pi_1^R(k+1,\lambda_k) - \Pi_1^C(k,\lambda_k)\right) \Big/ \frac{\partial}{\partial \lambda}\left(\Pi_1^R(k+1,\lambda_k) - \Pi_1^C(k,\lambda_k)\right)$$

By lemma 2.4, we have that the denominator is positive, so it remains to explore the sign of the numerator. The derivative of $\mu_k(b)$ with respect to $\overline{s}$ is given by:

$$\mu_k'(b) = \frac{\partial \lambda_k}{\partial \overline{s}}\mu_k = -\frac{(1-\lambda)\tilde{s}_k(b)}{\overline{s}^2}$$

Denote $\mu_k' = \mathbb{E}_b \mu_k'(b)$, then $\mu_k' > \mu_{k+1}'$ since $\tilde{s}_k < \tilde{s}_{k+1}$. Denote $z_1 = \max_b\{\mu_k^2(b)\} < 1$, $z_2 = \max\{T_{k+1}^R(b), T_k^C(b)\} < 1$.

$$\frac{\partial}{\partial \overline{s}}\left(\Pi_1^R(k+1,\lambda_k) - \Pi_1^C(k,\lambda_k)\right) = t\mathbb{E}\left[\frac{\mu_k'(b)}{\mu_k^2(b)} - \frac{\mu_{k+1}'(b)}{\mu_{k+1}^2(b)} + \mu_{k+1}'T_{k+1}^R(b) - \mu_k' T_k^C(b)\right] >$$
$$> t\mathbb{E}\left[\frac{\mu_k'(b) - \mu_{k+1}'(b)}{z_1} - (\mu_k'(b) - \mu_{k+1}'(b))z_2\right] = t(\mu_k' - \mu_{k+1}')(1/z_1 - z_2) > 0$$

Numerator is positive, therefore $\partial \lambda_k / \partial \overline{s} > 0$.

□

Proposition 2.5. *Consider the benchmark formation game in which firms decide whether to reveal their costs or not.*

1. If $\tilde{k} = 0$ (high $\overline{s}$), then i) if $\lambda \leq \lambda_0$ then all firms reveal their costs, $k^* = N$. ii) If $\lambda_0 < \lambda \leq \lambda_N$ then there are two pure strategy equilibria, $k_1^* = 0$ and $k_2^* = N$. In this case, firms are better off by coordinating on revelation: $\Pi_1^R(N) > \Pi_1^C(0)$. iii) If $\lambda > \lambda_N$ then all firms conceal their costs, $k^* = 0$.

2. If $\tilde{k} > 0$ (intermediate $\overline{s}$), then j) if $\lambda \leq \lambda_N$ then there are two pure strategy equilibria, $k_1^* = 0$ and $k_2^* = N$. In this case, firms are better off by coordinating on revelation: $\Pi_1^R(N) > \Pi_1^C(0)$. jj) If $\lambda > \lambda_N$ then all firms conceal their costs, $k^* = 0$.

3. If $\nexists \tilde{k}$ (low $\overline{s}$) described by definition 2.2, then firms will conceal their costs, $k^* = 0$.

Proof. 1. First, consider the case in which $\tilde{k} = 0$. i) By proposition 2.4, there is a positive fraction of informed consumers λ_0, such that it would be optimal to reveal if all firms conceal information. By proposition 2.3 we find that for all $\lambda \leq \lambda_0$ there is no stable coalition of $m < N$ revealing firms, because each time a concealing firm would deviate to revelation:

$$\Pi_1^R(N,\lambda)-\Pi_1^C(N-1,\lambda) > \ldots > \Pi_1^R(m+1,\lambda)-\Pi_1^C(m,\lambda) > \ldots > \Pi_1^R(1,\lambda)-\Pi_1^C(0,\lambda) > 0$$

ii) Consider an intermediate $\lambda_0 < \lambda \leq \lambda_N$. By proposition 2.4 there exists λ_k such that $\lambda_{k-1} < \lambda \leq \lambda_k$. Since $\lambda_{k-1} < \lambda$ we have that each coalition consisting of k or lower number of revealing firms $1 \leq m \leq k$ is not an equilibrium, because each firm will deviate to concealment:

$$0 = \Pi_1^R(k,\lambda_{k-1})-\Pi_1^C(k-1,\lambda_{k-1}) \geq \Pi_1^R(k,\lambda)-\Pi_1^C(k-1,\lambda) > \ldots \Pi_1^R(m,\lambda)-\Pi_1^C(m-1,\lambda)$$

Moreover, since $\lambda \leq \lambda_k$ we have that all revealing coalitions that consists of $k < m < N$ firms cannot constitute an equilibrium because firms from $N \setminus K$ will prefer to reveal:

$$0 = \Pi_1^R(k+1,\lambda_k)-\Pi_1^C(k,\lambda_k) \leq \Pi_1^R(k+1,\lambda)-\Pi_1^C(k,\lambda) < \ldots < \Pi_1^R(m,\lambda)-\Pi_1^C(m-1,\lambda)$$

Therefore, in the intermediate range of $\lambda_0 < \lambda \leq \lambda_N$, there are two equilibria in pure strategies, in which firms either conceal or reveal their costs, $k^* = 0$ or $k^* = N$.

There are also mixed strategy equilibria in each interval $\lambda_{k-1} < \lambda < \lambda_k$ which we do not characterize. Instead, we show that if firms reveal their costs, the industry profit attains its maximum over all other equilibria including the ones in mixed strategies.

Let's check that the profit of concealing firm $\Pi_1^C(k)$ is higher for higher k. The equilibrium profit can be represented in the following form (see the proof of lemma 2.4):

$$\Pi_1(k) = t\mathbb{E}_b\left[\frac{1}{\mu_k(b)} + \mu_k(b)T_k(b)\right]$$

,where $T_k(b) < 1$ for all b.

Denote $z_1 = \min_b \mu_{k-1}(b) < 1$ and $z_2 = \max_b\{T_k(b), T_{k-1}(b)\} < 1$ and compare the difference in profits:

$$\Pi_1^C(k) - \Pi_1^C(k-1) = t\mathbb{E}_b\left[\frac{1}{\mu_k(b)} - \frac{1}{\mu_{k-1}(b)} + \mu_k(b)T_k(b) - \mu_{k-1}(b)T_{k-1}(b)\right] >$$

$$> t\mathbb{E}_b\left[\frac{\mu_{k-1}(b)-\mu_k(b)}{z_1} - (\mu_{k-1}(b)-\mu_k(b))z_2\right] = \underbrace{(\mu_{k-1}(b)-\mu_k(b))}_{>0}\underbrace{(1/z_1 - z_2)}_{>0} > 0$$

Therefore, since $\Pi_1^C(k)$ is increasing in k we have that $\Pi_1^C(N-1) > \Pi_1^C(0)$. We showed above that for all $\lambda_0 < \lambda < \lambda_N$ we have that it is optimal to reveal costs of other $N-1$ firms reveal. Combining these two facts we get that the equilibrium profit under revelation is higher than under concealment:

$$\Pi_1^R(N) - \Pi_1^C(0) > \Pi_1^R(N) - \Pi_1^C(N-1) > 0$$

iii) Consider the case where there is a big fraction of informed consumers $\lambda > \lambda_N$. In this case firms will always deviate to concealing their private costs:

$$0 = \Pi_1^R(N, \lambda_N) - \Pi_1^C(N-1, \lambda_N) > \Pi_1^R(N, \lambda) - \Pi_1^C(N-1, \lambda) > \ldots >$$
$$> \Pi_1^R(m, \lambda) - \Pi_1^C(m, \lambda)$$

2. If $\tilde{k} > 0$ then $\Pi_1^C(m-1) > \Pi_1^R(m)$ for all λ if $m \leq \tilde{k}$. Therefore, similar to the proof of case 2, one can show that there are two equilibria if $\lambda \leq \lambda_N$ and unique equilibrium $k^* = 0$ if $\lambda > \lambda_N$.

3. Third, suppose that there is no $\tilde{k}$ defined in 2.2, then all revelation coalitions including $k = 1, \ldots, N$ firms are not stable because firms will deviate to concealment: for all λ we have that $\Pi_1^C(k, \lambda) > \Pi_1^R(k, \lambda)$ (from definition). Therefore, firms conceal costs in equilibrium and $k^* = 0$.

□

Proposition 2.6. $\lambda_0 > 0$ *exists and firms reveal their costs for* $\lambda < \lambda_0$ *if search costs are high enough:* $\overline{s}/t > \omega(N, \sigma_c^2/4t^2)$*, where*

$$\omega^2(N, x) = \left(\frac{x}{2} + N\left(\frac{1}{4} + z(N)x\right)\right)\left(\frac{1}{2} + x\right)\left(\frac{1}{4} + z(N)x\right)$$

and

$$z(N) = \frac{1}{2}\left(1 - \left(1 - \left(\frac{2N-2}{2N-1}\right)^2 / N\right)\right).$$

Proof. Firms find it optimal to deviate from the no sharing regime and reveal their private costs if $\Pi_1^R(1, \lambda) > \Pi_1^C(0, \lambda)$. By lemma 2.4, λ_0 exists if it is optimal to reveal in the market without shoppers: $\Pi_1^R(1, \lambda = 0) > \Pi_1^C(0, \lambda = 0)$. Applying lemma 2.1 we can show that if firms want to reveal under private information sharing, they will do so if information is hared publicly:

$$\Pi_1^R(1, \lambda = 0) - \Pi_1^C(0, \lambda = 0) > \Pi_0^R(1, \lambda = 0) - \Pi_0^C(0, \lambda = 0)$$

Explore the sign of the difference in profits $\Delta\Pi_0(1)$ in the private case.

$$t^{-1}\Delta\Pi_0(1) = \left(\frac{1}{\mu_1} - \frac{1}{\mu_0}\right) - (\mu_0 - \psi\mu_1)x = (\mu_0 - \mu_1)\left(\frac{1}{\mu_1\mu_0} - x\right) - (1-\psi)\mu_1 x \tag{2.26}$$

, where $\psi = (2N-2)^2/(2N-1)^2 < 1$ and $x = \sigma_c^2/4t^2$. By lemma 2.6, we can find the analytical expression for equilibrium fraction of informed consumers when $\lambda = 0$:

$$\mu_k = \frac{t}{\overline{s}}\left(\frac{1}{4} + \frac{1}{2N}(N - k + \psi k)x\right) \tag{2.27}$$

To explore the sign of profit difference, we denote $z(N) = \frac{1}{2}\left(1 - \frac{1-\psi}{N}\right)$ and substitute the expressions for μ_0 and μ_1 given by (2.27) to (2.26)

$$t^{-1}\Delta\Pi_0(1) = \left(\frac{1}{2} - z(N)\right)\left[\frac{(\overline{s}/t)^2}{\left(\frac{1}{4}+\frac{1}{2}x\right)\left(\frac{1}{4}+zx\right)} - x\right] - x(1-\psi)\left(\frac{1}{4}+zx\right)$$

By simplifying, we get that it is optimal to reveal and $\Delta\Pi_0(1) > 0$ if and only if the search costs are high enough and the following condition holds:

$$\left(\frac{\overline{s}}{t}\right)^2 > \left(\frac{x}{2} + N\left(\frac{1}{4} + z(N)x\right)\right)\left(\frac{1}{2} + x\right)\left(\frac{1}{4} + z(N)x\right) = \omega^2(N, x).$$

□

3 Non-Reservation Price Equilibrium and Search Without Priors

3.1 Introduction

Uncertainty lies at the heart of any optimal search problem. Most of the search literature reduces uncertainty to risk, by assuming that options are sampled from a known distribution. In the consumer search literature this risk comes from mixed pricing strategies employed by firms (e.g. Stahl (1989), Janssen *et al.* (2005)), or from a random component in utility (e.g. Wolinsky (1986), Anderson and Renault (1999), Armstrong and Zhou (2010), Zhou (2014)). In the recent literature a more sophisticated approach is often employed: consumers are unaware of some parameters of the model and 'estimate' them in a Bayesian way using the observed prices (e.g. Dana Jr (1994), Janssen *et al.* (2011)). All these models require consumers to have certain 'priors' – distributional assumptions over prices, utilities or other parameters of the model.

As the consumer search literature deals with situations where consumers do not know prices, it is natural to assume that in such situations they also do not know how many firms operate in the market, and what those firms' costs and market shares might be. Therefore in this paper we propose a robust search procedure, where the solution of the optimal stopping problem does not rely either on priors about parameters of the model, or on the market model structure itself. As consumers are agnostic about the distribution of options from which they sample, they cannot base their decisions on traditional expected utility. In our paper consumers use the optimal stopping rule which minimises the loss in comparison with an (imaginary) informed searcher (who knows the distribution of the prices) in the *worst case scenario*. By doing so consumers are guaranteed that the expected difference between their utility and the utility obtained under ex ante optimal behaviour is below a certain bound.

In our paper we look at a market with a homogenous good in which consumers search for the best price. However our results can be applied to a wide variety of settings. Following Salop and Stiglitz (1977) and Dana Jr (1994) we use the so-called ‘newspaper search protocol’: after observing the first price for free, consumers can learn at some cost all the prices in the market by visiting a price aggregator website (or by buying a newspaper with advertisements). A traditional sequential search approach is very complicated in the setting of search without priors and is prone to a dynamic inconsistency problem, as pointed out by Hayashi (2009): consumers might want to change their search strategy as they gather information about prices. As the first search is costless in our paper, there is effectively one point in time where consumers make a search decision, and, hence the dynamic inconsistency issue vanishes. This approach also allows us to consider equilibrium settings with more than two firms. Newspaper search protocol has its limitations: in markets with physical transportation costs sequential search is a more realistic model. Both protocols coincide in duopoly markets.

It turns out that the optimal stopping rule is stochastic: for sufficiently high prices consumers randomise between stopping and continuing to search. Once a consumer has observed a price, she suffers from both *greed* (a concern that by not searching, she may miss out on a lower price) and *fear* (a concern that searching may not yield a lower price, and thus be wasteful). In principle the consumer has to think about all possible distributions, then figure out what an informed searcher would do for each distribution and estimate the maximum loss. The analysis is simplified by the result that there are only two worst case distributions: one associated with greed and another associated with fear. As the minimax regret approach is equivalent to playing a fictitious strictly competitive game against malicious nature (which maximises consumer’s regret by its choice of price distribution), it results in a mixed strategy search behaviour. An important feature of our approach is that the search behaviour is robust to model specification and distributional assumptions over parameters. As such it can be applied to various market models without any change.

We then look at the consequences of the minimax optimal stopping rule for market outcomes. In a market model, firms are strategic players who can correctly anticipate the behaviour of consumers and maximise their profit, while consumers are agnostic about the market primitives and minimise regret in the worst case scenario. We allow for a general distribution of search costs and solve for the market equilibrium. We find that for all parameter values and search costs distributions there is a unique pricing equilibrium, which is characterised by price dispersion. An intriguing feature of this result is that price dispersion arises in equilibrium even if all consumers have strictly positive costs, i.e. our model avoids the Diamond (1971) paradox. We find that the average *listed* price increases with the number of firms and approaches the monopoly price as the number of firms grows without bound. However the average price *paid* by consumers stays constant. As the number of firms grows, more consumers decide to search and, conditional on deciding to search, they observe more price quotations. As under the newspaper search protocol the price paid is the order statistic, consumers end up paying the same prices although the listed prices go up.

Finally we compare our results with Stahl (1989) model, in which consumers hold a correct belief about firms' pricing strategies. It turns out that average price is lower when searchers do not have a prior about the price distribution. The reason is that in our model consumers actively search, whereas in Stahl's model they do not.

Our paper applies the minimax regret approach[1] to the problem of search under uncertainty.[2] The closest article to ours is Bergemann and Schlag (2011b). Although we use a similar minimax regret approach to the search problem, our paper differs from theirs in several ways. We follow the consumer search literature by studying an environment with both search costs and perfect recall. Bergemann

[1] See Hayashi (2008) and Stoye (2011) for axiomatisation of minimax regret, Bergemann and Schlag (2008), Bergemann and Schlag (2011a)) for an application of this approach to the monopoly pricing problem and Manski (2007) and Stoye (2012) for an application to treatment with missing data.

[2] Alternative approaches to this problem can be found in Chou and Talmain (1993), Nashimura and Ozaki (2004), Riedel (2009) and Parakhonyak (2014).

and Schlag (2011b) on the other hand consider a setup more common in the labour search literature, without search costs and recall, and only briefly discuss costly search. We use the so-called 'newspaper' search protocol, when consumers can obtain information about all other options via a price aggregator. This allows us to look at environments with more than two alternatives. Another important difference is that we do not restrict attention to just the optimal stopping problem, but we also consider its consequences for oligopolistic competition.

Our paper makes several contributions to the literature on consumer search. First, we provide a tractable framework that allows for an arbitrary search cost distribution and that can be applied to various market structure specifications. Second, our paper does not differentiate between incomplete information (e.g., about production costs or some other parameters of the model) and strategic uncertainty which arises from equilibrium mixed strategies in complete information games. This significantly simplifies the analysis of models with asymmetric information about production costs (see e.g., Dana Jr (1994), Tappata (2009), Janssen *et al.* (2011)). Models with unknown production costs are characterised by reservation price search strategies, but when the heterogeneity in production costs is sufficiently large a pricing equilibrium does not exist. Janssen *et al.* (2017) propose a non-reservation price equilibrium in order to avoid this problem. Our paper also uses non-reservation price search strategies, though derived from different considerations. The departure from the rationality paradigm brings with it the advantage that we can study a more general model than Janssen *et al.* (2017) we allow for multiple firms, while they consider only duopoly; we allow for general search cost distributions while they look only at binary distribution of search costs (zero and positive). Moreover, the equilibrium in our model is unique, while they have multiple equilibria.

Our model has interesting comparative statics. Textbook oligopoly models suggest that an increase in the number of firms should lead to lower prices. In the search literature this result is often reversed. Stahl (1989) showed that as the number of firms increases, listed prices approach the monopoly price, and the effective price paid by all consumers first increases (when the search cost binds) and then stays stable (when the valuation binds). In our model listed prices also increase and approach the monopoly price as the number of firms approaches infinity. However, the average price is constant regardless of the level of search costs. The reason is that although the listed prices are higher, consumers search more and each search pays off better because of our choice of search protocol.

Arguably our model has more appealing empirical properties than models of sequential consumer search which have the reservation price property. In those models, consumers stop searching after the first search, which is not true in reality (see, De los Santos *et al.* (2012)). Our model explains why consumers actively search in markets for homogeneous goods. Moreover, the empirical prediction of our model is that for the same search history consumers can make different decisions, depending on the realization of their mixed strategy. This result comes from the absence of priors and loss aversion of consumers, rather than from some fundamental irrationality.

The rest of the paper is organised as follows. Section 2 presents the model setup, including the formal statement of the objective function. In Section 3 we derive the optimal stopping rule. Section 4 considers the pricing equilibrium in a general model and studies its main properties. Section 5 provides a detailed comparison with the Stahl (1989) model for the case of duopoly. Section 6 provides a discussion of the robustness of our results and briefly concludes. Technical proofs are presented in the Appendix.

3.2 Model

There are $N \geq 2$ firms in the industry. Firms produce a homogeneous good at marginal cost $m > 0$ and compete in prices. There is a unit mass of consumers. Each consumer has a unit demand for the good, and values it at v which we normalise to 1. These consumers are split evenly among the firms and know the price at their firm but not at the others.[3] After observing the price at their firm, consumers can pay a cost c and get access to all the prices in the market via a price aggregator. Search costs are distributed on $[\underline{c}, \overline{c}]$ according to distribution function $\mu(c)$. We assume that $\underline{c} < 1$. Consumers can return to the previously sampled

[3] Janssen *et al.* (2005) relax the assumption that the first observation is free.

option without incurring any additional costs.[4] A consumer is agnostic about the underlying mechanisms of the market: she is unaware both of the market structure and possible parameter distribution. She holds a belief about the support of the price distribution $[\underline{q}, \overline{q}]$, but does not know the actual shape of the distribution function for any of the firms. The usual way to model beliefs is to assume that the consumer knows the true boundaries of the support of the theoretical distribution (see, e.g. Bergemann and Schlag (2011b)). This assumption is not plausible in a consumer search context, since the derivation of these boundaries requires knowledge of all the parameters of the model. In this paper we derive the optimal stopping rule for an arbitrary belief $[\underline{q}, \overline{q}]$, assuming that it is not possible to observe 'surprising' prices, i.e. $p < \underline{q}$ or $p > \overline{q}$. Then, for the equilibrium analysis we make the assumption $\underline{q} = 0$ and $\overline{q} = 1$, i.e. consumers only rule out the prices at which a firm cannot make any profit. Although we look at a symmetric equilibrium of the model, from an uninformed searcher perspective there is no reason to assume that the distributions are the same, i.e. that the firms are ex ante symmetric or play a symmetric equilibrium. It is also possible that an uniformed searcher does not know the number of firms in the market. We assume that an uninformed searcher holds the belief that there are up to M firms operating in the market, and that this belief is correct, i.e. $N \leq M$.

Note that as consumers do not know the shape of the distribution, the objective function of consumers cannot take the form of an expected utility function, simply because the expectation can only be defined when the distribution function is known. We assume that consumers are regret minimisers, in the sense that they want to minimise their maximum loss compared to what they would get if they knew the true distribution. More formally, let mapping $s(p) : [0, \infty] \to [0, 1]$ be a stopping rule which indicates the probability of accepting a price p.[5] Let $\mathbb{F} = \{F_i\}_{i=1}^{M}$ be a collection of probability distributions over the prices of all M firms which can be potentially active in the market. We require that the price distributions of truly existing firms $F_i(\cdot)$, $i = \overline{1, N}$ have supports consistent with

[4] Janssen and Parakhonyak (2014) relax this assumption and explicitly model return costs.

[5] Standard reservation price optimal stopping rules take the form $s(p) = 0$ for all prices above the reservation price, and $s(p) = 1$ for all prices less or equal to the reservation price.

the belief $[\underline{q}, \overline{q}]$, i.e. with $F_i(\underline{q} - \varepsilon) = 0$ and $F_i(\overline{q} + \varepsilon) = 1$ for any $\varepsilon > 0$. This implies that there is no 'surprise' price discovery. We denote this collection of probability distributions by $\mathbb{F}_N = \{F_i\}_{i=1}^N$. For the purpose of characterising the optimal stopping rule, we model the remaining $M-N$ that are not actually present in the market as artificial active firms which charge prices above the valuation: for all $M \geq j > N$ $F_j(p) = \mathbb{I}_{p \geq p_j}$, $p_j > 1$. These artificial 'firms' do not affect the behaviour of informed searchers, which stays the same both under collection $\mathbb{F}$ and $\mathbb{F}_N$. However when uninformed searchers make their decision, they take into account the possibility that these $M-N$ firms can play non-degenerate strategies. Suppose an uninformed searcher observes the price in store j and has to decide whether to continue searching. Denote $\mathbb{F}_{-j} = \{F_i\}_{i \neq j}$. The loss function is defined by

$$
\begin{aligned}
L(s, F_j, \mathbb{F}_{-j}) = \quad & -\int_{\underline{q}}^{\rho_{\mathbb{F}_{-j}}} p_j dF_j(p_j) - \int_{\rho_{\mathbb{F}_{-j}}}^{\overline{q}} \int_{\underline{q}}^{\overline{q}} \min(p_j + c, y + c) dF_j(p_j) dH_{\mathbb{F}_{-j}}(y) + \\
+ \quad & \int_{\underline{q}}^{\overline{q}} \int_{\underline{q}}^{\overline{q}} \{s(p_j)p_j + [1 - s(p_j)] \min(p_j + c, y + c)\} \, dF_j(p_j) dH_{\mathbb{F}_{-j}}(y)
\end{aligned} \tag{3.1}
$$

where $y = \min(p_1, \ldots, p_{j-1}, p_{j+1}, \ldots, p_M)$ and $H_{\mathbb{F}_{-j}}(y) = 1 - \prod_{i \neq j, i \leq M}[1 - F_i(y)]$, and $\rho_{\mathbb{F}_{-j}}$ is a reservation price if consumers knew the distribution function which is defined by:

$$
\rho_{\mathbb{F}_{-j}} = c + H_{\mathbb{F}_{-j}}(\rho_{\mathbb{F}_{-j}}) \mathbb{E}_{H_{\mathbb{F}_{-j}}}(y | y \leq \rho_{\mathbb{F}_{-j}}) + [1 - H_{\mathbb{F}_{-j}}(\rho_{\mathbb{F}_{-j}})] \rho_{\mathbb{F}_{-j}}
$$

i.e. $\rho_{\mathbb{F}_{-j}}$ is the price at which consumers are indifferent between buying now and searching.

Note, that the first two integrals in (3.1) define the expected price paid in the informed consumer case: what a consumer pays if the first price is less than $\rho_{\mathbb{F}_{-j}}$, and what she pays if it is greater and one more search is conducted. The third integral defines the expected price paid by an uninformed consumer who uses stopping rule $s(p)$. By rearranging terms in (3.1) we get

$$
\begin{aligned}
L(s, F_j, \mathbb{F}_{-j}) = -\int_{\underline{q}}^{\overline{q}} \int_{\underline{q}}^{\overline{q}} \Big\{ & \mathbb{I}_{p \leq \rho_{\mathbb{F}_{-j}}} [1 - s(p)][p - \min(p, y) - c] + \\
& + \mathbb{I}_{p > \rho_{\mathbb{F}_{-j}}} s(p)[\min(p, y) + c - p] \Big\} \, dH_{\mathbb{F}_{-j}}(y) dF_j(p)
\end{aligned}
$$

Denote $M_{\mathbb{F}_{-j}}(p) = \mathbb{E}_{H_{\mathbb{F}_{-j}}} \min(p, y) + c$. Then,

$$L(s, F_j, \mathbb{F}_{-j}) = \int_{\underline{q}}^{\overline{q}} \left\{ \mathbb{I}_{p \leq \rho_{\mathbb{F}_{-j}}} [1 - s(p)][M_{\mathbb{F}_{-j}}(p) - p] + \mathbb{I}_{p > \rho_{\mathbb{F}_{-j}}} s(p)[p - M_{\mathbb{F}_{-j}}(p)] \right\} dF_j(p) \tag{3.2}$$

The first summand represents the loss from too much search – if the distributions were known it would be better to stop. The second term represents the loss from stopping too early when it is in fact optimal to continue searching. In the full information case, we have $\rho_{\mathbb{F}_{-j}} = M(\rho_{\mathbb{F}_{-j}})$. It is straightforward to verify that this solution always exists and is unique.

Let $\mathcal{S}$ be the set of all stopping rules $s(p)$ and let $\mathcal{F}$ be a set of all possible collections of probability distributions $\mathbb{F} = \{F_j, \mathbb{F}_{-j}\}$. The objective of a consumer is to find the optimal stopping rule which minimises the loss function in the worst case scenario (over all possible collections of probability distributions):

$$s \in \operatorname*{Arg\,min}_{s \in \mathcal{S}} \max_{\mathbb{F} = \{F_j, \mathbb{F}_{-j}\} \in \mathcal{F}} L(s, F_j, \mathbb{F}_{-j})$$

We look at equilibrium pricing in our model under the assumption that firms are strategic. We do this even though consumers are not strategic, as they minimise their regret in a situation of ambiguity, and therefore they do not necessarily play best-responses to firms' strategies in a normal game-theoretical sense. More formally, we define equilibrium in the following way. Let $F(p), G(p) : [0, \infty) \to [0, 1]$ be pricing strategies (possibly degenerate distributions over prices) and s be a search strategy (defined for each consumer with search costs $c \in [\underline{c}, \overline{c}]$).

Definition 3.1. *An equilibrium of the model consists of pricing strategies $F(p)$ and search strategies s such that:*

1. *A firm's profit is maximised by $F(p)$ given the strategies of consumers and other firms: $\pi_i[F_i(p), F_{-i}(p), s] \geq \pi_i[G_i(p), F_{-i}(p), s]$.*

2. *A consumer starts at a random firm j and follows the strategy which minimises regret: $s \in \operatorname{Arg\,min}_{s \in \mathcal{S}} \max_{\mathbb{F} \in \mathcal{F}} L(s, F_j, \mathbb{F}_{-j})$.*

3.3 Optimal Stopping Rule

We start our analysis by deriving the optimal stopping rule. Later we use this optimal stopping rule for the equilibrium analysis. We define the maximum loss as

$$\overline{L} = \min_{s \in \mathcal{S}} \max_{\mathbb{F} \in \mathcal{F}} L(s, F_j, \mathbb{F}_{-j})$$

The following Theorem defines the optimal stopping rule.

Theorem 3.1. *The stopping rule*

$$s(p; c, \underline{q}) = \begin{cases} 1 & \text{if} \;\; p \leq \underline{q} + c \\ \frac{c}{p - \underline{q}} & \text{if} \;\; \underline{q} + c < p \leq \min(\overline{q} + c, 1) \\ 0 & \text{if} \;\; p > \min(\overline{q} + c, 1) \end{cases} \tag{3.3}$$

minimises the maximum loss and guarantees that $\overline{L} = \Big[1 - c/(\overline{q} - \underline{q})\Big]\, c$.

Note, that the optimal stopping rule does not possess the standard reservation price property: for intermediate prices, consumers stop with certain probability. This means that unlike in many search models with homogeneous goods[6] (based on Stahl (1989)) there is active search in equilibrium. To understand the reasons behind the stochastic optimal stopping rule, suppose that the consumer faces price p. In making her decision, she considers two worst case scenarios, which for expositional purposes we associate with two basic emotions:

- *Greed*: if the consumer decides to stop, she is concerned that the supports of (some of) other price distributions include only the the lower bound $\underline{q}$, and it was optimal to continue;

- *Fear*: if the consumer decides to search, she is concerned that the support of other price distributions contain only the current price p (and possibly prices above p), so the search is wasteful.

[6] See, for example, Dana Jr (1994), Janssen *et al.* (2011), Janssen and Parakhonyak (2013), Parakhonyak (2014)

The higher the probability of stopping, the larger is the regret in the first scenario. Likewise the lower the probability of stopping, the higher is the regret in the second scenario. The maximum of these two regrets is minimised by some intermediate probability of stopping, for which both regrets are equal.

Another notable feature of the optimal stoping rule is that it depends (when it is between 0 and 1) only on the lower bound of the support of the distribution, while the upper bound plays a passive role. This is also the case in Parakhonyak (2014). In both cases the reason is the same: given the assumption of perfect recall, when determining whether to stop or not, only the distribution of prices below the current price matters. This implies that if the consumer does not have any prior, only the lower bound of the distribution (together with search costs) determines the optimal stopping rule. Note that the maximum loss is defined by $[1 - s(\overline{q}; c, \underline{q})]c$, which is strictly better than the naive strategy 'always to search', for which the maximum loss equals c. Note, that whether consumers know the number of firms ($N = M$) or not ($N < M$) does not affect the optimal stopping rule.

The following proposition summarises the properties of the optimal stopping rule:

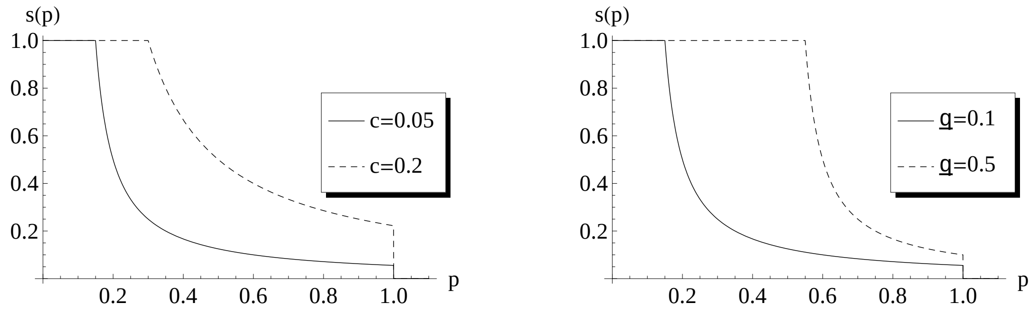
Proposition 3.1. *For all $p \in [\underline{q} + c, \overline{q}]$ the optimal stopping probability $s(p; c, \underline{q})$ (i) decreases in price p, (ii) increases in search costs c and (iii) increases in the belief about the lower bound of the support $\underline{q}$.*

Thus, consumers are more likely to stop if the search cost is high, or if price is low in comparison with the lower bound. For the rest of the paper we will omit the list of parameters wherever it does not affect understanding and just use notation $s(p)$. Figure 3.1 shows what the optimal stopping rule looks like for different sets of parameters.

3.4 Equilibrium Analysis

In this section we use the optimal stopping rule derived in Theorem 3.1 to solve for market equilibrium. In order to put more structure on the model, in this

Figure 3.1: $s(p)$ for different parameter values



(a) $\underline{q} = 0.1$, $c \in \{0.05, 0.2\}$ (b) $c = 0.05$, $\underline{q} \in \{0.1, 0.5\}$

section we impose that $\underline{q} = 0$ and $\overline{q} = 1$, i.e. consumers rule out only those prices that are strictly dominated from the firms' perspective. Recall, that consumers do not know the marginal cost of production m and therefore can assume that there might be prices lower than the actual value of the marginal cost. As long as $\underline{q}$ is an exogenous parameter, this assumption is not critical for our results, but allows us to reduce the dimensionality of the parameter space and make our results more comprehensive. In our analysis we concentrate on symmetric market equilibria.

Define the probability that a randomly chosen consumer decides to stop after she observes a price p by

$$\sigma(p) = \int_{\underline{c}}^{\overline{c}} s(p; c) d\mu(c) \tag{3.4}$$

where $s(p, c)$ is the optimal stopping rule in Theorem 3.1 for the case $\underline{q} = 0$ and $\overline{q} = 1$. Note, that for all $p \leq \underline{c}$ we have $\sigma(p) = 1$ and for all prices $p \in (\underline{c}, 1]$ we have $\sigma(p) \in (0, 1)$. Before we are ready to write down the profit function, we need to establish a few preliminary results: (i) there is no pure strategy pricing equilibria, (ii) the equilibrium distribution of prices is continuous, and (iii) the support of the equilibrium distribution is a convex set. The following Lemma shows that any symmetric pricing equilibrium is characterised by price dispersion.

Lemma 3.1. *There is no symmetric pure strategy pricing equilibria in the game. The support of the equilibrium price distribution $F(p)$ is a convex set and $F(p)$ continuous on its support.*

The result of Lemma 3.1 is different from the classical results in the consumer search literature. Diamond (1971) showed that if all consumers have the same positive search cost, and if the first search is for free, there is a unique equilibrium in which firms charge the monopoly price. Stahl (1996) showed that this result holds for an arbitrary distribution of search costs, unless there is a strictly positive mass of consumers with zero search costs. Note that Lemma 3.1 requires neither $\underline{c}$ to be equal to zero, nor $\mu(c)$ to have an atom at $\underline{c}$. Therefore, search without priors is a new way to avoid Diamond's paradox. Stiglitz (1987) avoided it by assuming that consumers know the empirical distribution of prices in the market, but do not know which firm charges which price. This approach implies that consumers possess much more information than in the original Diamond (1971) model, while our paper assumes the opposite: consumers are less informed about the primitives of the model. Rhodes (2014) avoided the paradox by considering multi-product firms. Now we are ready to write down the profit function of a firm.

$$\pi(p) = \frac{p-m}{N}\Big\{\sigma(p) + [1-\sigma(p)][1-F(p)]^{N-1} \\ + (N-1)[1-F(p)]^{N-2}\int_p^{\overline{p}}[1-\sigma(x)]dF(x)\Big\} \quad (3.5)$$

There is a fraction $1/N$ of consumers who start searching at our firm which charges a price p. $\sigma(p)$ of these consumers decide to stop and buy the good. $[1-\sigma(p)]$ decide to search and they buy from our firm only if all other prices are higher. Finally, there are consumers who start their search from each of the remaining $N-1$ firms. After observing price x at some other firm, they decide to search with probability $[1-\sigma(x)]$, and buy from our firm if both x and all other prices are higher than our price. Now we can proceed with the characterisation of the support of the equilibrium price distribution.

Lemma 3.2. *The upper bound of the equilibrium price distribution equals* 1.

This Lemma illustrates a very interesting choice faced by the firms. Note that at the upper bound consumers who leave the firm after their first search never come back. Thus, if a firm faces only fresh consumers, the threat of losing them is not strong enough to charge a price which is lower than the monopoly price. The motivation to charge lower prices is sufficient only when a firm takes into account those consumers who first visited other stores. Lemma 3.2 allows us to pin down the equilibrium level of profits:

$$\overline{\pi} = \frac{\sigma(1)}{N}(1-m)$$

Now we are ready to establish the existence and uniqueness of equilibrium.

Theorem 3.2. *For any number of firms N and any search cost distribution $\mu(c)$ there is a unique symmetric equilibrium.*

Although search models with homogeneous goods under complete information usually have a unique symmetric equilibrium (see, e.g. Stahl (1989), Janssen and Parakhonyak (2014)) this is not typical for models with incomplete information. For example, models with common production costs which are unknown to consumers (e.g., Dana Jr (1994), Tappata (2009), Janssen *et al.* (2011)) suffer from non-existence of equilibria whenever there is large dispersion in production costs. An attempt to solve this issue by considering non-reservation price equilibria by Janssen *et al.* (2017) encounters the opposite problem – multiplicity of equilibria. Our paper gives a framework which can be directly used for studying this type of problem: as the demand side of the model is not affected by production cost uncertainty, the fact that the marginal cost of production is not known to consumers does not change (3.5). Theorem 3.2 is still applicable, with the interpretation that m is a realisation from some distribution of production costs and the model still has a unique equilibrium which always exists.

For the rest of this Section we concentrate on the properties of the equilibrium. We start with comparative statics with respect to N. Note that as we have a unit mass of consumers and unit demand, the expected price paid by consumers equals the total profit collected by the firms plus total production cost: $\mathbb{E}p = N\overline{\pi} + m = (1-m)\sigma(1) + m$. As $\sigma(1) < 1$ the expected price is increasing in the marginal cost of production. As $\sigma(1)$ does not depend on N the average price paid by consumers is constant in the number of firms *for any distribution of search costs.* This result is different from comparative statics of the Stahl (1989) model, where average price paid by all consumers first increases with the number of firms, and then remains constant once the reservation price hits the valuation. In our model the upper bound of the support of the equilibrium price distribution equals 1 for any distribution of search costs and thus total industry profit (and the average price) is constant in N. The distribution in Stahl's model approaches $\mathbb{I}_{p\geq 1}$ as the number of firms approaches infinity, thus consumers with positive search costs tend to pay higher prices as the number of firms goes up. As shown in the following Proposition, our price distribution has a similar limit property.

Proposition 3.2. *As the number of firms becomes larger, the probability mass concentrates in a neighbourhood around 1:*

$$\lim_{N\to\infty} F(p) = \mathbb{I}_{p\geq 1}$$

Even though the listed prices are higher, the average price paid by consumers is constant. The reason is that consumers who decided not to search pay higher prices, but consumers who decided to search pay lower prices as they compare more prices. Somewhat similar result is obtained in Varian (1980), although search decisions in his model are not endogenised, as search costs in his model are either zero or infinite.

Now we consider comparative statics with respect to search costs. As we work with general cost distributions, rather than common search costs, we formulate our result in terms of first order stochastic dominance.

Proposition 3.3. *Consider two possible distributions of search costs: $\mu_L(c)$ and $\mu_H(c)$. Suppose $\mu_H(c) \leq \mu_L(c)$ for all c. If $F_L(p)$ and $F_H(p)$ are equilibrium price distributions for $\mu_L(c)$ and $\mu_H(c)$ correspondingly, then $F_H(p)$ stochastically dominates $F_L(p)$.*

Following an increase in search costs consumers search less, there is less competition and firms charge higher prices. This result is common to most search models. Unlike higher N, higher search costs reduce search activity and do not affect the benefits of search, except via changes in the price distribution. Consequently higher search costs lead to higher expected prices in the sense of first order stochastic dominance. Finally, we define $\mu_\alpha(c)$ to be a distribution of search costs on $[\alpha\underline{c}, \alpha\bar{c}]$ such that $\mu_\alpha(\alpha c) = \mu(c)$. As $\alpha \to 0$ the aggregate stopping probability $\sigma(1) \to 0$ and $F(p) \to \mathbb{I}_{p \geq m}$, i.e. as search costs approach zero the market approaches perfect competition.

3.5 Comparison with Stahl (1989)

In this section we compare our results with a seminal paper by Stahl (1989). Stahl considers a model with two types of consumers: a fraction λ are shoppers (with zero search cost) whilst the remaining fraction are searchers (with positive search cost c_0). Thus,

$$\mu_{\text{Stahl}}(c) = \begin{cases} \lambda & \text{if} \quad c \in [0, c_0) \\ 1 & \text{if} \quad c \geq c_0 \end{cases} \tag{3.6}$$

Stopping probabilities are then defined in the following way:

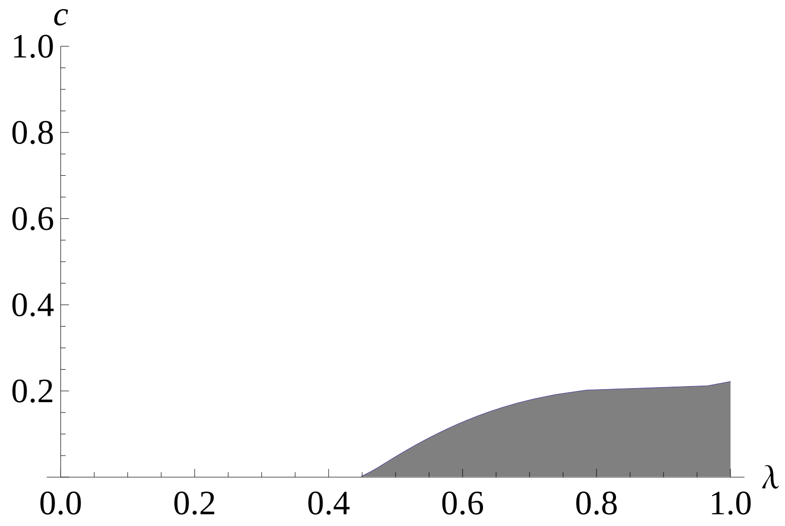
$$\sigma_{\text{Stahl}}(p) = \begin{cases} 1 - \lambda & \text{if} \quad p \in [0, c_0) \\ (1 - \lambda)s(p; c_0) & \text{if} \quad p \in [c_0, 1] \end{cases} \tag{3.7}$$

For this section we concentrate on the case of $N = 2$, as in this case the equilibrium price distribution can be explicitly characterised both in our model and in Stahl's. Note that for $N = 2$ the price-aggregator search protocol coincides with standard sequential search, which makes a direct comparison with Stahl (1989) possible. We have discussed the comparative statics of the two models with respect to N in the previous section: price distributions in both models tend to degenerate at 1 as the number of firms approaches infinity, while the total industry profit (and expected price paid) in Stahl's model first increases for sufficiently small search costs, while in our model it does not depend on N.

Notice that the results of Lemma 3.2 and Theorem 3.2 are directly applicable here. Therefore there is a unique symmetric mixed strategy equilibrium. Unlike in Stahl (1989), in our model the upper bound of the support always equals the valuation and is not affected by search costs. This pushes the expected price upwards in comparison with the reference models. However active search in our model makes it less attractive to charge higher prices, therefore they are charged with lower probability, which pushes the expected price down. The following proposition establishes the interrelation between these two forces.

Proposition 3.4. *For any $(c_0, \lambda) \in (0,1)^2$ the expected price in the model of search without priors is lower that in the model with informed searchers.*

Figure 3.2: Total Cost of Buying in Comparison with Stahl (1989), $m = 0.2$



The average price is always lower in our model than in the reference model by Stahl (1989). However in his model consumers do not search, while in ours they do and therefore they incur extra search costs. Figure 3.2 illustrates the total expected cost of acquiring the good, which includes price and search cost. In the gray area non-shoppers end up paying more for acquiring the good in our model than in Stahl's. Thus if the number of shoppers is sufficiently small, excessive search puts downward pressure on prices and this is enough to compensate consumers for their search cost (regardless how large that search cost might be). If the number of shoppers is sufficiently large the upper bound of the support of the price distribution in Stahl (1989) is far from the valuation and consumers end up paying less in his model.

Note that Proposition 3.3 formulated for a general model is applicable in this particular case. It implies that higher search costs c_0 or lower λ lead to higher prices in the model. This result is in line with the model with informed searchers.

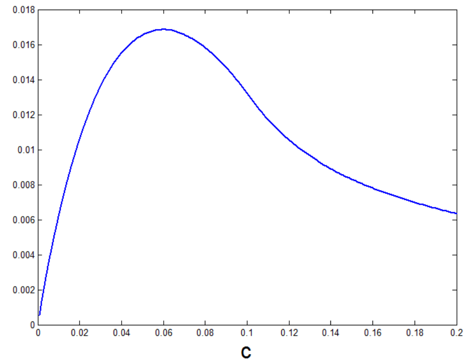
Now we look at the cost of being agnostic about the pricing distribution. We make two types of comparisons. First, we look at how much an (atomistic) informed searcher would be better off if she were put in our equilibrium environment. Second, we look how much an (atomistic) uninformed searcher would be worse off if she is put in an environment with informed searchers, i.e. the Stahl (1989) model. The results are represented in Figure 3.3: the left panel is for an informed searcher in our model, the right one is for an uninformed searcher is Stahl's model. It turns out that the loss is not dramatic: in all scenarios it does not exceed a few percentage points. As the fraction of shoppers grows, prices tend to concentrate near the lower bound of the support, uninformed searchers search very rear (while informed do not search at all) and the loss decreases. As the marginal cost of production increases, the price distribution shifts upwards and uninformed searchers search suboptimally often. In Stahl (1989) it is optimal not to search, so it is natural that the loss is increasing in search cost. In our model although consumers search a lot for small search cost, this cost is small and the loss is small. As c grows the loss increases. For large c both informed and uninformed searchers reduce their search activity, thus the loss decreases.

3.6 Discussion and Conclusions

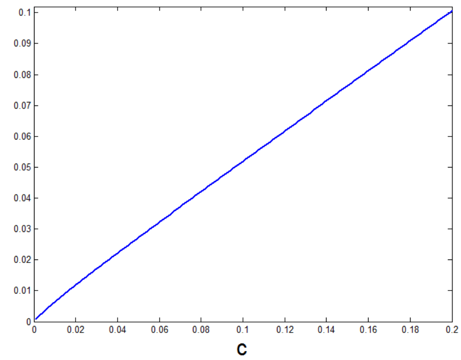
In this paper we proposed a framework for search without priors: consumers do not have information about the distribution from which they sample and try to minimise the maximum regret. The optimal stopping rule does not possess the reservation price property, i.e. consumers stop searching with certain probability. This implies that unlike in standard search models with homogeneous goods there is active search in equilibrium. This has two important implications for empirical research: consumers search beyond the first option (which coincides with findings of De los Santos *et al.* (2012)) and might have different choices for the same price histories.

As the optimal stopping behaviour does not require any information about the underlying market model and its parameters, our optimal search strategy is applicable to a wide variety of settings. Moreover, the optimal search strategy is

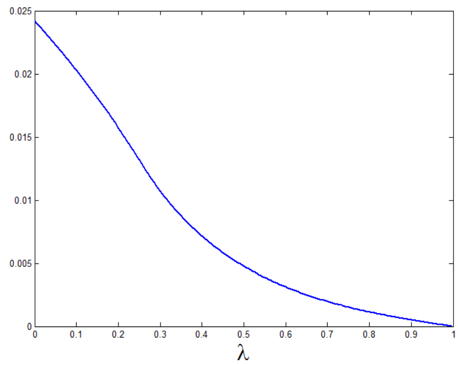
Figure 3.3: Equilibrium Loss



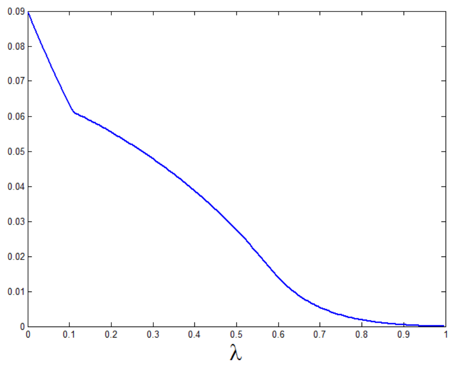
(a) Uninformed Model, $\lambda = 0.25$, $m = 0.05$



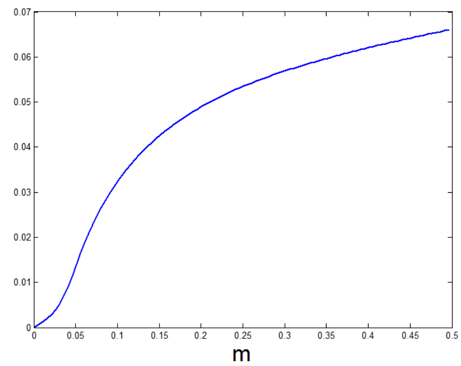
(b) Informed Model, $\lambda = 0.25$, $m = 0.05$



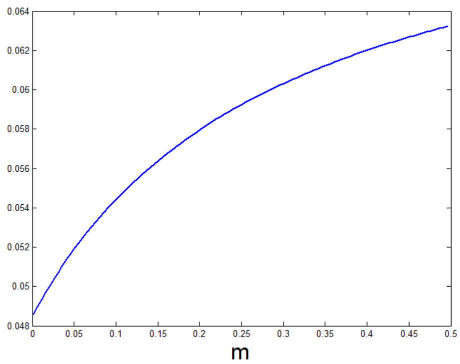
(c) Uninformed Model, $c = 0.1$, $m = 0.05$



(d) Informed Model, $c = 0.1$, $m = 0.05$



(e) Uninformed Model, $c = 0.1$, $\lambda = 0.25$



(f) Informed Model, $c = 0.1$, $\lambda = 0.25$

invariant to the choice of a market model. This approach allows us to solve the non-existence problem in a consumer search model with production cost information asymmetry.

In this paper we considered a model with heterogeneous searchers, which is quite often a difficult task in models with homogeneous goods (see, e.g. Stahl (1989)). We showed that for all parameter values there is a unique equilibrium in the model. Average prices paid by consumers in this equilibrium are lower than in Stahl (1989) (for the same parameter values). This equilibrium exhibits interesting comparative statics in the number of firms: listed prices increase and effective prices remain constant as the number of firms grows.

Throughout the paper we have made several assumptions, which we discuss briefly here. Firstly, we assumed that consumers do not know the number of firms in the market. This assumption can be relaxed by setting $M = N$ and none of our results change. Secondly, we assumed so-called 'newspaper' search protocol. Sequential search is more common choice for consumer search problems. However, under search without priors this protocol is prone to a dynamic inconsistency problem and in order to make the analysis tractable one has to assume that consumers commit ex ante to their search strategy (see Hayashi (2009)). Our modelling choice allows us to restrict consumers' decisions to one point in time and thus avoid this problem. Thirdly, one might ask the question: what happens if consumers know the boundaries of the support of the distribution? In this case $\underline{q}$ would depend on firms' equilibrium strategies. It turns out that the results from Section 5 do not change considerably: now it is the case that the upper bound of the support can be lower than the valuation, but the equilibrium still exists and is unique. Finally, we assumed that uninformed searchers do not impose any symmetry assumptions on equilibrium pricing distributions. We find this assumption reasonable in most situations. In order to relax this assumption the family of the least favourable distributions has to be altered: the distribution associated with fear remains the same, while the support of the distribution associated with greed now must contain the observed price. Although it is clear that the optimal stopping rule remains stochastic (as the searcher plays a strictly competitive game against malicious nature) we were not able to obtain a closed form solution for the stopping probability.

We think that our approach to search without priors opens up a broad research agenda. It creates new possibilities for empirical research: as the optimal stopping rule is invariant to the choice of the market model, it is possible to directly compare search costs in different markets. From a theoretical perspective it would be very interesting to have a fully robust model of a search market, in which consumers do not know pricing strategies of the firms, and firms do not know the distribution of search costs (or, like in Bergemann and Schlag (2008) valuations) of consumers.

3.7 Appendix

Theorem 3.1. *The stopping rule*

$$s(p;c,\underline{q}) = \begin{cases} 1 & \text{if } \; p \leq \underline{q}+c \\ \frac{c}{p-\underline{q}} & \text{if } \; \underline{q}+c < p \leq \min(\overline{q}+c,1) \\ 0 & \text{if } \; p > \min(\overline{q}+c,1) \end{cases} \tag{3.8}$$

minimises the maximum loss and guarantees that $\overline{L} = \left[1 - c/(\overline{q}-\underline{q})\right] c$.

Proof. We consider the non-trivial case where search costs are sufficiently small such that $\underline{q}+c<1$.

If $p<\underline{q}+c$ the consumer obviously wants to stop: the best price she can get is $\underline{q}$, but she also has to pay search costs of c.

If $p>\min(\overline{q}+c,1)$ the consumer is either guaranteed to get a better offer, or rejects it, because the price is higher than the valuation of the good. In either case no purchase is made.

Now, we consider the case of $p \in [\underline{q}+c, \min(\overline{q}+c,1)]$. Note, that if $p > (\leq) \rho_{\mathbb{F}_{-j}}$ then $M_{\mathbb{F}_{-j}}(p) < (\geq) p$.
Let's consider the degenerate distribution functions $\underline{F} = \{\underline{q}\}$ and $\overline{F}_p = \{p\}$ that can help us to compute the loss function. Let's denote $\underline{\mathbb{F}}_{-j} = \{F_i : i \neq j, \exists k : F_k = \underline{F}\}$ and $\overline{\mathbb{F}}_{-j,p} = \{F_i : \forall i \neq j, F_i = \overline{F}_p\}$. Clearly $M_{\underline{\mathbb{F}}_{-j}}(q) \leq M_{\mathbb{F}_{-j}}(q) \leq M_{\overline{\mathbb{F}}_{j,q}}(q)$ for any $q \in [\underline{q}, \overline{q}]$. It is easy to verify that $\rho_{\underline{\mathbb{F}}_{-j}} \leq \rho_{\mathbb{F}_{-j}}$ and $\rho_{\overline{\mathbb{F}}_{-j,p}} = p+c$.

Since $M_{\mathbb{F}_{-j}}(p) \leq M_{\overline{\mathbb{F}}_{-j,p}}(p)$ we have that $[1-s(p)][M_{\mathbb{F}_{-j}}(p)-p] \leq [1-s(p)][M_{\overline{\mathbb{F}}_{-j,p}}(p)-p]$ for any p. Thus, the loss from excessive search is higher for the worst possible distribution for searching $\overline{\mathbb{F}}_{-j,p}$ rather than for $\mathbb{F}_{-j}$. Similarly the loss from stopping too early is maximized by $\underline{\mathbb{F}}_{-j}$. Therefore we get that

$$L(s,F_j,\mathbb{F}_{-j}) \leq \int_{\underline{q}}^{\overline{q}} \left\{ \mathbb{I}_{p \leq \rho_{\mathbb{F}_{-j}}} [1-s(p)][M_{\overline{\mathbb{F}}_{-j,p}}(p)-p] + \mathbb{I}_{p > \rho_{\mathbb{F}_{-j}}} s(p)[p - M_{\underline{\mathbb{F}}_{-j}}(p)] \right\} dF_j(p)$$

Notice that from the definition of loss function for any $p > \rho_{\mathbb{F}_{-j}}$ (and therefore for $p > \rho_{\underline{\mathbb{F}}_{-j}}$) we have that $L(s, \mathbb{I}_{q \geq p}, \underline{\mathbb{F}}_{-j}) = s(p)[p - M_{\underline{\mathbb{F}}_{-j}}(p)]$. For any $p \leq \rho_{\overline{\mathbb{F}}_{-j,p}}$, including those $p \leq \rho_{\mathbb{F}_{-j}}$, we obtain $L(s, \mathbb{I}_{q \geq p}, \overline{\mathbb{F}}_{-j,p}) = [1-s(p)][M_{\overline{\mathbb{F}}_{-j,p}}(p)-p]$.

Finally we derive the upper bound for the loss function:

$$L(s,F_j,\mathbb{F}_{-j}) \leq \int_{\underline{q}}^{\overline{q}} \Big\{\mathbb{I}_{p\leq\rho_{\mathbb{F}_{-j}}}[1-s(p)][M_{\overline{\mathbb{F}}_{-j,p}}(p)-p]+\mathbb{I}_{p>\rho_{\mathbb{F}_{-j}}}s(p)[p-M_{\underline{\mathbb{F}}_{-j}}(p)]\Big\}\,dF_j(p) =$$
$$= \int_{\underline{q}}^{\overline{q}} \Big[\mathbb{I}_{p\leq\rho_{\mathbb{F}_{-j}}}L(s,\mathbb{I}_{q\geq p},\overline{\mathbb{F}}_{-j,p})+\mathbb{I}_{p>\rho_{\mathbb{F}_{-j}}}L(s,\mathbb{I}_{q\geq p},\underline{\mathbb{F}}_{-j})\Big]\,dF_j(p) \leq$$
$$\leq \int_{\underline{q}}^{\overline{q}} \max_p \Big[\mathbb{I}_{p\leq\rho_{\mathbb{F}_{-j}}}L(s,\mathbb{I}_{q\geq p},\overline{\mathbb{F}}_{-j,p})+\mathbb{I}_{p>\rho_{\mathbb{F}_{-j}}}L(s,\mathbb{I}_{q\geq p},\underline{\mathbb{F}}_{-j})\Big]\,dF_1(p) =$$
$$= \max_p\{\max[L(s,\mathbb{I}_{q\geq p},\overline{\mathbb{F}}_{-j,p}),L(s,\mathbb{I}_{q\geq p},\underline{\mathbb{F}}_{-j})]\}$$

This holds for any F_j and $\mathbb{F}_{-j}$ therefore

$$\max_{F_j,\mathbb{F}_{-j}} L(s,F_j,\mathbb{F}_{-j}) \leq \max_p \max_{\mathbb{G}\in\{\underline{\mathbb{F}}_{-j},\overline{\mathbb{F}}_{-j,p}\}} L(s,\mathbb{I}_{q\geq p},\mathbb{G}).$$

Denote $\mathcal{F}$ as the set of all possible combination of any M distribution functions on $[\underline{q},\overline{q}]$. As $\{\{\mathbb{I}_{q\geq p},\underline{\mathbb{F}}_{-j}\},\{\mathbb{I}_{q\geq p},\overline{\mathbb{F}}_{-j,p}\}\}\subset\mathcal{F}$ for any p we have that

$$\max_{\mathbb{F}\in\mathcal{F}} L(s,F_j,\mathbb{F}_{-j}) \geq \max_p \max_{\mathbb{G}\in\{\underline{\mathbb{F}}_{-j},\overline{\mathbb{F}}_{-j,p}\}} L(s,\mathbb{I}_{q\geq p},\mathbb{G})$$

So we can simplify our problem by narrowing the set of the candidates to be the worst distributions:

$$\max_{\mathbb{F}\in\mathcal{F}} L(s,F_j,\mathbb{F}_{-j}) = \max_p \max_{\mathbb{G}\in\{\underline{\mathbb{F}}_{-j},\overline{\mathbb{F}}_{-j,p}\}} L(s,\mathbb{I}_{q\geq p},\mathbb{G})$$

Now we solve the problem $\min_s \max_{\mathbb{G}\in\{\underline{\mathbb{F}}_{-j},\overline{\mathbb{F}}_{-j,p}\}} L(s,\mathbb{I}_{q\geq p},\mathbb{G})$ for a given p. Note, that our problem is equivalent to a zero-sum game between a consumer and malicious nature: a searcher tries to minimize the loss function, while the nature maximises it. The consumer chooses whether to stop or not, while nature chooses distribution $\underline{\mathbb{F}}_{-j}$ and $\overline{\mathbb{F}}_{-j,p}$ and possibly mixture of them. Then,

$$L(s,\mathbb{I}_{q\geq p},\underline{\mathbb{F}}_{-j}) = s(p-\underline{q}-c) \tag{3.9}$$

as the consumer gets a price p with probability 1, then stops with probability s and pays p instead of the expected minimal price $\underline{q}$ plus the search cost c. If the distribution is $\overline{\mathbb{F}}_{-j,p}$, then

$$L(s,\mathbb{I}_{q\geq p},\overline{\mathbb{F}}_{-j,p}) = (1-s)c \tag{3.10}$$

as the consumer incurs losses only if she faces price p and decides to continue searching, when it is optimal to stop. Note that the zero-sum game we consider admits only a mixed-strategy equilibrium. Therefore, the consumer must choose the optimal stopping probability in such a way, that the nature is indifferent between playing $\underline{\mathbb{F}}_{-j}$ and $\overline{\mathbb{F}}_{-j,p}$:

$$L(s, \mathbb{I}_{q \geq p}, \underline{\mathbb{F}}_{-j}) = L(s, \mathbb{I}_{q \geq p}, \overline{\mathbb{F}}_{-j,p})$$

which gives

$$s(p - \underline{q} - c) = (1 - s)c$$

Solving for s gives

$$s = \frac{c}{p - \underline{q}}$$

The saddle point probability μ of playing $\{\mathbb{I}_{q \geq p}, \underline{\mathbb{F}}_{-j}\}$ can be obtained from the fact that consumer is indifferent between stopping and continuing to search:

$$p = c + (1 - \mu)p + \mu\underline{q}$$

which gives

$$\mu = s = \frac{c}{p - \underline{q}}.$$

Clearly the maximum loss is achieved by choosing $F_j = \mathbb{I}_{q \leq \overline{q}}$ and equals to $\overline{L} = (1 - s(\overline{q}))c = \left[1 - c/(\overline{q} - \underline{q})\right] c$.

□

Proposition 3.1. *For all $p \in [\underline{q} + c, \overline{q}]$ the optimal stopping probability $s(p; c, \underline{q})$ (i) decreases in price p, (ii) increases in search costs c and (iii) increases in the belief about the lower bound of the support $\underline{q}$.*

Proof. By taking partial derivatives we get:

$$\frac{\partial s}{\partial p} = -\frac{c}{(p - \underline{q})^2} < 0$$
$$\frac{\partial s}{\partial \underline{q}} = \frac{c}{(p - \underline{q})^2} > 0$$
$$\frac{\partial s}{\partial c} = \frac{1}{p - \underline{q}} > 0$$

□

Lemma 3.1. *There is no symmetric pure strategy pricing equilibria in the game. The support of the equilibrium price distribution $F(p)$ is a convex set and $F(p)$ continuous on its support.*

Proof. Suppose, that the equilibrium price is $p^* > \underline{c}$. Each firm earns $\pi_0 = (p^* - m)/N$ in equilibrium. Note, that $1-\sigma(p^*)$ consumers search in equilibrium. Thus, a firm deviating from p^* to $p = p^* - \varepsilon$ earns $\pi_1 = \{1/N + [1-\sigma(p^*)](N-1)/N\}(p^* - \varepsilon - m)$. Then $\pi_1 - \pi_0 = [1-\sigma(p^*)](N-1)/N(p^* - \varepsilon - m) - \varepsilon/N$, which is clearly positive for ε small enough. Thus, there is a profitable deviation and $p^* > \underline{c}$ cannot be an equilibrium.

Now, suppose $p^* < \underline{c}$. In this case consumers do not search, and a firm can gain by deviating to $p = \underline{c}$.

Finally, suppose that $p^* = \underline{c}$. Consider a deviation to $p = 1$. Then a firm earns

$$\pi(1) = \frac{1-m}{N}\sigma(1) = \frac{1-m}{N}\int_{\underline{c}}^{\overline{c}} s(1,c)d\mu(c) \geq \frac{1-m}{N}s(1,\underline{c}) = \frac{1-m}{N}\underline{c} \geq \frac{\underline{c}-m}{N} = \pi(\underline{c})$$

The first inequality holds due to Proposition 3.1. Thus, there is no pure strategy equilibrium.

Now we prove that $F(p)$ is continuous on its support. It makes sense to consider the case where $F(\underline{c}) < 1$. Otherwise, if $F(\underline{c}) = 1$ all consumers buy immediately and there is no reason to play mixed strategy. Such a candidate would be a pure strategy and according to Lemma 3.1 there is no pure strategy equilibrium.

Suppose that there is a point mass at price p. We show that there is an optimal deviation to a slightly lower price $p - \epsilon$ which allows the firm to avoid a tie at p. Consider the difference in profits before and after the deviation:

$$\Delta\Pi = Pr(p_i > p-\epsilon, p_i \neq p, \forall i)(p-\epsilon-m)\left\{1/N + (N-1)/N\int_{p-\epsilon}^{\overline{p}}[1-\sigma(q)]dF(q)\right\} -$$
$$- Pr(p_i > p)(p-m)\left\{1/N + (N-1)/N\int_{p}^{\overline{p}}[1-\sigma(q)]dF(q)\right\} +$$
$$+Pr(p_j < p-\epsilon \text{ for some } j)(p-\epsilon-m)\sigma(p-\epsilon)/N - Pr(p_j < p \text{ for some } j)p\sigma(p)/N +$$
$$+\sum_{k=2}^{N} Pr(p_i \geq p-\epsilon, p_i = p \text{ for } k \text{ stores})(p-\epsilon-m)\left\{1/N + (N-1)/N\textstyle\int_{p-\epsilon}^{\overline{p}}[1-\sigma(q)]dF(q)\right\} -$$
$$-\sum_{k=2}^{N} Pr(p_i \geq p, p_i = p \text{ for } k \text{ stores})(p-m)\left\{1/N + (N-k)/N\textstyle\int_{p-\epsilon}^{\overline{p}}[1-\sigma(q)]dF(q)\right\}$$

Clearly, when ϵ approaches 0 the sum of the four first terms approaches 0, but the sum of the two last terms remains positive.

Finally, we show that the support of $F(p)$ is a convex set. Suppose there is a gap, i.e. there is an interval of prices $[p_1, p_2]$, such that $F(p_1) = F(p_2)$. Consider a price $p' \in (p_1, p_2)$. Then

$$N[\pi(p') - \pi(p_2)] = [\sigma(p') - \sigma(p_2)] \left\{1 - [1 - F(p_2)]^{N-1}\right\} (p' - m) > 0$$

as $\sigma(p)$ is a decreasing function. Thus the deviation to p' is profitable. □

Lemma 3.2. *The upper bound of the equilibrium price distribution equals* 1.

Proof. Note that

$$\pi(\overline{p}) = \frac{\overline{p} - m}{N}\sigma(\overline{p}) = \frac{1}{N}\int_{\underline{c}}^{\overline{c}} (\overline{p} - m)s(\overline{p}; c)d\mu(c)$$

As for any c and $\overline{p} < 1$ we have

$$\frac{\partial[(\overline{p} - m)s(\overline{p}; c)]}{\partial\overline{p}} > 0$$

we obtain $\partial\pi/\partial\overline{p} > 0$. Thus, it is optimal to set the upper bound to be equal to 1. □

Theorem 3.2. *For any number of firms N and any search cost distribution $\mu(c)$ there is a unique symmetric equilibrium.*

Proof. We establish existence of equilibrium by proving that equation (3.5) together with the boundary condition $\overline{\pi} \equiv \pi(1) = (1 - m)\sigma(1)/N$ has a unique solution, which gives a well-defined distribution function.

Denote $\Phi(p) = [1 - F(p)]^{N-2}$. Then (3.5) together with the boundary condition can be rewritten as

$$\frac{\overline{\pi}N}{(p - m)\Phi(p)} = \frac{\sigma(p)}{\Phi(p)} + [1 - \sigma(p)]\Phi(p)^{\frac{1}{N-2}} - (N - 1)\int_p^1 [1 - \sigma(x)]d\Phi(x)^{\frac{1}{N-2}} \quad (3.11)$$

Denote

$$\phi(p) \equiv \frac{d\Phi(p)}{dp}, \qquad \xi(p) \equiv \frac{d\Phi(p)^{\frac{1}{N-2}}}{dp} = \frac{\Phi(p)^{\frac{1}{N-2}}\phi(p)}{(N - 2)\Phi(p)}.$$

Then by differentiating (3.11) we obtain:

$$-\frac{\overline{\pi}N[\Phi(p)+(p-m)\phi(p)]}{(p-m)^2\Phi(p)^2}=\left[\frac{\sigma'(p)}{\Phi(p)}-\frac{\sigma(p)\phi(p)}{\Phi(p)^2}\right]+N[1-\sigma(p)]\xi(p)-\sigma'(p)\Phi^{\frac{1}{N-2}} \tag{3.12}$$

By plugging in the expression for $\xi(p)$ and rearranging terms we obtain:

$$\phi(p)=\Phi(p)\frac{(p-m)^2\sigma'(p)+N\overline{\pi}-\sigma'(p)\Phi(p)^{\frac{N-1}{N-2}}}{(p-m)^2\sigma(p)-N(p-m)\overline{\pi}-\frac{N}{N-2}[1-\sigma(p)]\Phi(p)^{\frac{N-1}{N-2}}} \tag{3.13}$$

Now we have a boundary value problem (3.13) with the initial condition $\Phi(1)=0$. Both the numerator and denominator are differentiable almost everywhere on $p\in(0,1]$. As $\sigma'(p)<0$ the numerator of (3.13) is positive. Note, that

$$(p-m)^2\sigma(p)-N(p-m)\overline{\pi}=N(p-m)\left[\frac{(p-m)\sigma(p)}{N}-\overline{\pi}\right]$$

The term in brackets is negative (see the proof of Lemma 3.2), thus the numerator of (3.13) is negative. As the numerator of (3.13) is positive and the denominator is negative for all $p<1$ if a solution to (3.13) exists, then $H(p)$ is decreasing, so $F(p)$ is a well-defined distribution function.

Now we proceed with the existence result. Denote the denominator of 3.13 as $w(p)$, so

$$w(p)=(p-m)^2\sigma(p)-N(p-m)\overline{\pi}-\frac{N}{N-2}[1-\sigma(p)]\Phi(p)^{\frac{N-1}{N-2}}$$

Then,

$$\begin{aligned}w'(p)&=2(p-m)\sigma(p)+(p-m)^2\sigma'(p)-N\overline{\pi}-\\&\qquad\frac{N}{N-2}\left\{[1-\sigma(p)]\frac{N-1}{N-2}\Phi(p)^{\frac{1}{N-2}}\phi(p)-\sigma'(p)\Phi(p)^{\frac{N-1}{N-2}}\right\}=\\&=2(p-m)\sigma(p)+(p-m)^2\sigma'(p)-N\overline{\pi}+\sigma'(p)\frac{w(p)-(p-m)^2\sigma(p)+N(p-m)\overline{\pi}}{1-\sigma(p)}+\\&\frac{N-1}{N-2}\left[w(p)-(p-m)^2\sigma(p)+N(p-m)\overline{\pi}\right]\frac{(p-m)^2\sigma'(p)+N\overline{\pi}-\sigma'(p)\frac{w(p)-(p-m)^2\sigma(p)+N(p-m)\overline{\pi}}{\frac{N}{N-2}[1-\sigma(p)]}}{w(p)}\equiv\\&\equiv\alpha(p)+\beta(p)w(p)+\gamma(p)\frac{1}{w(p)}\end{aligned} \tag{3.14}$$

where $\alpha(p), \beta(p), \gamma(p)$ are corresponding coefficients, which are all are bounded continuously differentiable almost everywhere on $[0,1]$. Now let $\omega(p) = w^2(p)$, then using $\omega' = 2ww'$ we can rewrite our equation as:

$$\omega' = 2\alpha(p)\sqrt{\omega} + 2\beta(p)\omega + 2\gamma(p) \tag{3.15}$$

The boundary condition is $\omega_1 = 0$. The right hand side of (3.15) is continuous both in p and ω, thus, by Peano's theorem there exists a solution to differential equation (3.15), and, therefore, for (3.13).

Now we proceed with establishing uniqueness of the solution. Let's prove the result by contradiction. We know that there is at least one solution to this equation. Assume that there is more than one solution to this problem and consider two such functions $F_1(p), F_2(p)$ satisfying (3.5) and terminal condition $F_1(1) = F_2(1) = 1$. Without loss of generality, let price $\tilde{p}$ be such that $F_2(\tilde{p}) > F_1(\tilde{p})$ and for all $p > \tilde{p}$ we have $F_2(p) \geq F_1(p)$. Obviously such a $\tilde{p}$ exists.

Consider the case where $\tilde{p} \geq \underline{c}$, so $\sigma(\tilde{p}) < 1$. Define $\Delta(\tilde{p}) = [\pi(\tilde{p}, F_1) - \pi(\tilde{p}, F_2)]/(\tilde{p} - m)$. As $\pi(1, F_1) = \pi(1, F_2)$ it must be the case that $\Delta(\tilde{p}) = 0$. Now we write down the expression for $\Delta(\tilde{p}) = 0$:

$$\Delta(\tilde{p}) = \frac{1}{N}[1 - \sigma(\tilde{p})]\left\{[1 - F_1(\tilde{p})]^{N-1} - [1 - F_2(\tilde{p})]^{N-1}\right\} +$$
$$\frac{N-1}{N}\left\{[1 - F_1(\tilde{p})]^{N-2}\int_{\tilde{p}}^{1}[1 - \sigma(x)]dF_1(x) - [1 - F_2(\tilde{p})]^{N-2}\int_{\tilde{p}}^{1}[1 - \sigma(x)]dF_2(x)\right\}$$

Note, that the first summand of $\Delta(\tilde{p})$ is clearly positive. Now, consider the second one:

$$[1 - F_1(\tilde{p})]^{N-2}\int_{\tilde{p}}^{\overline{p}}[1 - \sigma(x)]dF_1(x) - [1 - F_2(\tilde{p}))]N - 2\int_{\tilde{p}}^{\overline{p}}[1 - \sigma(x)]dF_2(x) =$$
$$\left\{[1 - F_1(\tilde{p})]^{N-2} - [1 - F_2(\tilde{p})]^{N-2}\right\}\int_{\tilde{p}}^{1}[1 - \sigma(x)]dF_1(x) +$$
$$[1 - F_2(\tilde{p})]^{N-2}\int_{\tilde{p}}^{1}[1 - \sigma(x)]d[F_1(x) - F_2(x)] =$$
$$\left\{[1 - F_1(\tilde{p})]^{N-2} - [1 - F_2(\tilde{p})]^{N-2}\right\}\int_{\tilde{p}}^{1}[1 - \sigma(x)]dF_1(x) +$$
$$[1 - F_2(\tilde{p})]^{N-2}\left\{[F_2(\tilde{p}) - F_1(\tilde{p})][1 - \sigma(\tilde{p})] + \int_{\tilde{p}}^{1}\sigma'(x)[F_1(x) - F_2(x)]dx\right\} > 0$$

as $F_2(\tilde{p}) > F_1(\tilde{p})$, for all $p > \tilde{p}$ $F_2(p) > F_1(p)$ and $\sigma'(p) < 0$. Thus, $\Delta(\tilde{p}) > 0$ and we arrived at contradiction.

Now consider the case of $\tilde{p} < \underline{c}$. In this case we have

$$\Delta(\tilde{p}) = \frac{N-1}{N}\left\{[1-F_1(\tilde{p})]^{N-2}\int_{\tilde{p}}^{1}[1-\sigma(x)]dF_1(x) - [1-F_2(\tilde{p})]^{N-2}\int_{\tilde{p}}^{1}[1-\sigma(x)]dF_2(x)\right\}$$

which is positive, as we have shown before. Therefore solution to (3.5) is unique.

□

Proposition 3.2. *As the number of firms becomes larger, the probability mass concentrates in a neighbourhood around 1:*

$$\lim_{N\to\infty} F(p) = \mathbb{I}_{p\geq 1}$$

Proof. Define

$$G(p, F(\cdot), N) = \sigma(p)+[1-\sigma(p)][1-F(p)]^{N-1}+(N-1)[1-F(p)]^{N-2}\int_{p}^{1}[1-\sigma(x)]dF(x)-\frac{N\overline{\pi}}{p-m}$$

Let $F_N(p)$ be the equilibrium price distribution, i.e. $G(p, F_N(\cdot), N) = 0$. Now define

$$\tilde{G}(p, F(\cdot), N) = \sigma(p) + N[1-\sigma(1)][1-F(p)]^{N-1} - \frac{N\overline{\pi}}{p-m}$$

It is straightforward to verify that $\tilde{G}(p, F(\cdot), N) \geq G(p, F(\cdot), N)$. Let $\tilde{F}_N(p)$ be a price distribution, which solves $\tilde{G}(p, \tilde{F}_N(\cdot), N) = 0$. Note that if $F_1(p) \geq F_2(p)$ for all p, then $\tilde{G}(p, F_1(\cdot), N) \leq \tilde{G}(p, F_2(\cdot), N)$.

Now we show that for all p $F_N(p) \leq \tilde{F}_N(p)$. Suppose this is not true and there is $\hat{p}$ such that $F_N(\hat{p}) > \tilde{F}_N(\hat{p})$. Then,

$$\tilde{G}(\hat{p}, F_N(\hat{p}), N) < \tilde{G}(\hat{p}, \tilde{F}_N(\hat{p}), N) = 0 = G(\hat{p}, F_N(.), N)$$

which contradicts the fact that $\tilde{G}(p, F(\cdot), N) \geq G(p, F(\cdot), N)$. Thus for all p $F_N(p) \leq \tilde{F}_N(p)$. Note that

$$\tilde{F}_N(p) = 1 - \left\{\frac{(1-m)\sigma(1) - p\sigma(p)}{N[1-\sigma(1)]}\right\}^{\frac{1}{N-1}}$$

As $\sigma(1) < 1$ and $(1-m)\sigma(1) - p\sigma(p) > 0$ (see the proof of Lemma 3.2 for any $p < 1$ $\lim_{N\to\infty} \tilde{F}_N(p) = 0$, thus the same holds for $F_N(p)$. □

Proposition 3.3. *Consider two possible distributions of search costs: $\mu_L(c)$ and $\mu_H(c)$. Suppose $\mu_H(c) \leq \mu_L(c)$ for all c. If $F_L(p)$ and $F_H(p)$ are equilibrium price distributions for $\mu_L(c)$ and $\mu_H(c)$ correspondingly, then $F_H(p)$ stochastically dominates $F_L(p)$.*

Before we prove the proposition we need the following technical Lemma.

Lemma A.1. *Suppose that the random variables L and H are respectively distributed according to $\mu_L(c)$ and $\mu_H(c)$ with supports $[\underline{c}_L, \overline{c}_L]$ and $[\underline{c}_H, \overline{c}_H]$. Suppose μ_H first order stochastically dominates μ_L. Let $\gamma : [\underline{c}_L, \overline{c}_H] \to R$ is an increasing and differentiable function. Then $\mathbb{E}[\gamma(H)] \geq \mathbb{E}[\gamma(L)]$.*

Proof. Assume that the difference of distribution functions $\theta(c) = \mu_H(c) - \mu_L(c)$ has M points of discontinuity c_i, $i = 1, \dots, M$ on $(\underline{c}_L, \overline{c}_H]$. We will consider $\theta(c)$ on each of the interval $[c_i, c_{i+1}]$, where there is only one point of discontinuity at c_{i+1} in each of them. Then the part of the difference in expected values taken for i-th interval is give by:

$$\xi([c_i, c_{i+1}]) = \int_{c_i}^{c_{i+1}} \gamma(c) d\theta(c) + \Big[\gamma(c)\theta(c)\Big]_{c_{i+1-0}}^{c_{i+1}}$$

Using integration by parts for piecewise smooth function we get:

$$\begin{aligned}\xi([c_i, c_{i+1}]) = \Big[\gamma(c)\theta(c)\Big]_{c_{i+1-0}}^{c_{i+1}} + \Big[\gamma(c)\theta(c)\Big]_{c_i}^{c_{i+1-0}} - \int_{c_i}^{c_{i+1}} \frac{d\gamma(c)}{dc}\theta(c)dc = \\ = \Big[\gamma(c)\theta(c)\Big]_{c_i}^{c_{i+1}} - \int_{c_i}^{c_{i+1}} \frac{d\gamma(c)}{dc}\theta(c)dc\end{aligned}$$

Taking the sum over all intervals, where $c_0 = \underline{c}_L$ and $c_{M+1} = \overline{c}_H$ we get:

$$\xi([\underline{c}_L, \overline{c}_H]) = \sum_{i=0}^{M} \Big[\gamma(c)\theta(c)\Big]_{c_i}^{c_{i+1}} - \int_{\underline{c}}^{\overline{c}} \frac{d\gamma(c)}{dc}\theta(c)dc = -\gamma(\underline{c})\theta(\underline{c}) - \int_{\underline{c}}^{\overline{c}} \frac{d\gamma(c)}{dc}\theta(c)dc$$

Since $\theta(c) \leq 0$ and γ is increasing, we get that $\xi([\underline{c}_L, \overline{c}_H]) \geq 0$ or equivalently $\mathbb{E}[\gamma(H)] \geq \mathbb{E}[\gamma(L)]$. □

Now we proceed with the proof of Proposition 3.3.

Proof. All prices in the equilibrium support must be higher then m. Consider the case $p > \underline{c}$. Notice that $s(p, c)$ increases in c for $p > \underline{c}$. Therefore using Lemma

A.1 for $\gamma(c) = s(p, c)$ it is easy to see that for stochastically dominant distribution of costs the aggregated probability to stop is not smaller:

$$\Delta\sigma(p) = \sigma_H(p) - \sigma_L(p) \geq 0$$

Let's denote rearranging profit function as G_i for $i = L, H$:

$$G_i(p, F) = \frac{1}{N}(p - m)[1 - \sigma_i(p)][1 - F(p)]^{N-1} + \\ \frac{N-1}{N}(p-m)[1-F(p)]^{N-2}\int_p^{\overline{p}}[1-\sigma_i(q)]dF(q) - \frac{1}{N}\left[\sigma_i(\overline{p})(\overline{p} - m) - \sigma_i(p)(p - m)\right]$$

Consider the following difference for arbitrary distribution function F:

$$\Delta G(p, F) = G_L(p, F) - G_H(p, F) = \frac{1}{N}\Delta\sigma(p)(p - m)[1 - F(p)]^{N-1} + \\ + \frac{N-1}{N}(p - m)\int_p^{\overline{p}} \Delta\sigma(p)dF(q) + \psi(p)$$

where $\psi(p) = [\sigma_H(\overline{p})(\overline{p} - m) - \sigma_H(p)(p - m)] - [\sigma_L(\overline{p})(\overline{p} - m) - \sigma_L(p)(p - m)]$. Let's show that the sign of $\psi(p)$ is not negative:

$$\psi(p) = [\sigma_H(\overline{p})(\overline{p} - m) - \sigma_L(\overline{p})(\overline{p} - m)] - [\sigma_H(p)(p - m) - \sigma_L(p)(p - m)] = \\ = \{[\sigma_H(x)(x - m)]' - [\sigma_L(x)(x - m)]'\}(\overline{p} - p) = \{E[\tilde{\gamma}(H)] - E[\tilde{\gamma}(L)]\}(\overline{p} - p)$$

where $\tilde{\gamma}(c) = [s(x, c)(x - m)]'_x$ and $p \leq x \leq \overline{p}$. Since for $x > \underline{c}$ we have that

$$\frac{\partial}{\partial c}\tilde{\gamma}(c) = \frac{m}{x^2} > 0$$

and therefore $\tilde{\gamma}(c)$ is increasing function. Using Lemma A.1 we get that $E[\tilde{\gamma}(H)] \geq E[\tilde{\gamma}(L)]$ and therefore $\psi(p) \geq 0$ for any $p > \underline{c}$. Consequently we can deduce that $G_L(p, F) \geq G_H(p, F)$ for any $p > \underline{c}$. Now assume that there is $\hat{p} > \max(\underline{c}, m)$ such that $F_H(\hat{p}) > F_L(\hat{p})$. Then

$$G_L(\hat{p}, F_L) - G_L(\hat{p}, F_H) = \frac{1}{N}[1-\sigma_L(\hat{p})](\hat{p}-m)\left\{[1 - F_L(\hat{p})]^{N-1} - [1 - F_H(\hat{p})]^{N-1}\right\} + \\ + \frac{N-1}{N}(\hat{p}-m)\left\{ [1 - F_L(\hat{p})]^{N-2}\int_{\hat{p}}^{\overline{p}}[1 - \sigma_L(q)]dF_L(q) - [1 - F_H(\hat{p})]^{N-2}\int_{\hat{p}}^{\overline{p}}[1 - \sigma_L(q)]dF_H(q)\right\} > \\ > \frac{N-1}{N}(\hat{p} - m)[1 - F_H(\hat{p})]^{N-2}\int_{\hat{p}}^{\overline{p}}[1 - \sigma_L(q)][f_L(q) - f_H(q)]dq > \\ > \frac{N-1}{N}(\hat{p} - m)[1 - F_H(\hat{p})]^{N-2}[1 - \sigma_L(\hat{p})][F_H(\hat{p}) - F_L(\hat{p})] > 0$$

Since F_L is equilibrium distribution for $\mu_L(c)$ then $G_L(\hat{p}, F_L) = 0$. Therefore

$$G_L(\hat{p}, F_H) < G_L(\hat{p}, F_L) = 0 = G_H(\hat{p}, F_H)$$

That contradicts the fact that $\Delta G(\hat{p}, F_H) \geq 0$. Thus $F_H(p) \leq F_L(p)$ for any $p > \underline{c}$.

Now consider the case where $m < p \leq \underline{c}$ or $\sigma(p) = 1$. In this case the function $G_i(p, F)$ will take the form:

$$G_i(p, F) = \frac{N-1}{N}(p-m)[1-F(p)]^{N-2} \int_{\underline{c}}^{\overline{p}} [1-\sigma_i(q)]dF(q) - \frac{1}{N}\left[\sigma_i(\overline{p})(\overline{p}-m) - (p-m)\right]$$

Consider the difference:

$$\Delta G(p, F) = G_L(p, F) - G_H(p, F) = \frac{N-1}{N}(p-m) \int_{c}^{\overline{p}} \Delta\sigma(p) dF(q) + \psi_1,$$

where $\psi_1 = [\sigma_H(\overline{p})(\overline{p}-m) - \sigma_L(\overline{p})(\overline{p}-m)] \geq 0$. Thus we get that $\Delta G(p, F) \geq 0$. Now assume that there is $\hat{p} \leq \underline{c}$ such that $F_H(\hat{p}) > F_L(\hat{p})$.Using the previous inequalities we can simply obtain:

$$G_L(\hat{p}, F_L) - G_L(\hat{p}, F_H) > \frac{N-1}{N}(p-m)[1-F_H(p)]^{N-2}[1-\sigma_L(c)][F_H(\hat{p}) - F_L(\hat{p})] = 0$$

Therefore we get the same contradiction $G_L(\hat{p}, F_H) < G_H(\hat{p}, F_H)$. □

Proposition 3.4. *For any $(c_0, \lambda) \in (0,1)^2$ the expected price in the model of search without priors is lower that in the model with informed searchers.*

Proof. Note that the expected price in Stahl's model is defined by $2\pi^S + m = (1-\lambda)(\rho - m) + m$, where ρ is the reservation price. In our model expected price is given by $2\pi + m = (1-\lambda)c_0(1-m) + m$. Thus, the expected price (and profit) is our model is lower if and only if $\rho > c_0(1-m) + m$. Note, that $\rho = \min[1, c_0/M(\lambda) + m]$ where

$$M(\lambda) = 1 - \frac{(1-\lambda)\ln\frac{1+\lambda}{1-\lambda}}{2\lambda}$$

As $M(0) = 0, M(1) = 1$ and $M'(\lambda) > 0$ we obtain $\rho > c_0 + m > c_0(1-m) + m$ which completes the proof. □

4 Search and Learning Without Priors

4.1 Introduction

Nowadays, search models lie at the heart of explaining the behaviour of agents in imperfectly competitive markets. Learning is an inherent part of any search problem. It was studied theoretically (Rothschild (1978), Benabou and Gertner (1993), Dana Jr (1994), Janssen *et al.* (2011) and tested empirically (Koulayev (2013), De los Santos *et al.* (2017)). Most of the literature on search and learning considers searchers endowed with some initial beliefs or priors and are updated according to the Bayes rule. However, this approach has two drawbacks. First, this is a complex problem and the optimal search and learning behaviour for a given prior are often hard to characterize. Second, the resulting search strategies are very sensitive to the initial information that might be imprecise practically (Gastwirth (1976)). In this paper, we explore the problem of search and learning when the searcher is worried about the correctness of her initial beliefs. We propose a *robust search rule* that performs well for a large set of priors and adapts to the newly acquired observations.

To be more explicit, consider two examples. For instance, suppose that an unemployed worker searches for a job in a specific industry. She might have some imprecise idea of how her education, work experience as well as labour supply and market conditions affect the chances of getting a job in this industry. Every job search outcome reveals some information that must be taken into account by the job seeker. Alternatively, suppose that a consumer is interested to buy a product with specific characteristics, but he does not know which firms offer this product.

He might have some imprecise idea of whether the demand for this product is high, so that the probability of finding this product is high as well. Every search trial will convey some new information about the chances to find the product. In both of these examples, the searchers do not know the underlying market mechanisms and therefore it is an extremely difficult task to specify reasonably "good" beliefs about the unknown distributions. Instead, searchers might want to design a strategy that would use the initial information and perform reasonably well even if their initial information is misspecified.[1]

We consider a setting with binary-valued alternatives that can take either a low or a high value. The searcher discovers alternatives sequentially and decides whether or not to stop. If the searcher decides to stop she can accept the outside option that is better than the low value. Initially, a searcher has some information about the probability of finding the high value. We introduce model uncertainty by *ε-contamination*, under which the searcher believes that her initial information is correct with probability ε. We are primarily interested in intermediate or small values of ε that describe sufficient degree of uncertainty. With the remaining probability of $1-\varepsilon$, the searcher does not know the true distribution. We assume all values are discounted over time and therefore search is costly.

In order to compare search rules it is important to determine how to measure their performance. Our approach is to compare the performance of a search rule after any history to the performance of the optimal rule of a Bayesian who has a prior that is contained in the set of possible priors of the searcher. This Bayesian searcher learns about the actual distribution by updating her beliefs according to the Bayes' rule. We define the *loss function* of a search rule for a given prior and a given history as the difference between the payoff of the optimal rule of the Bayesian who has this prior and the payoff that is obtained by the searcher who uses the specified search rule. The performance of a search rule of the searcher is evaluated by the maximal loss across all priors and all histories. We are dealing

[1] In his seminal paper on search and learning Rothschild (1978), emphasized that "...the problem of what the searcher should do if he does not know the price distribution must be attacked. This can be done in two ways: by exploring the properties of reasonable rules of thumb, or by characterizing optimal rules." In this paper we take the first approach and explore the rules of thumb that perform well for any prior.

with a dynamic problem and do not wish to impose any commitment possibilities. So the searcher has to anticipate her search strategy in the future when deciding about today's actions. This criterion that is based on comparing performance after any history was introduced by Schlag and Zapechelnyuk (2018). A rule that obtains a small maximial loss after any history can be interpreted as a rule that is good for any Bayesian. Related studies on robust search by Bergemann and Schlag (2011b) and Bergemann and Schlag (2011a) only compared performance from an ex ante perspective.[2]

We fully characterize a robust search rule in the case of binary-valued payoffs. It has two important properties. First, we find that the searcher has to *randomize* between searching and stopping up to a certain number of search periods. The intuition is as follows. It is easy to show that any search problem with binary valued payoffs is terminated after a finite time. Therefore, there exist a number of search periods, after which it is optimal to stop irrespectively of the priors. Next, suppose that the searcher decides to search or stop with certainty. Then the losses are maximal and come from either *excessive search* or *too early stopping*. The robust strategy must resolve the tradeoff between these two sources of losses.

Another key property of our robust search rule is that it is *adaptive* to the new information and it incorporates the pessimism of finding the high value over time. Intuitively, all Bayesian searchers who decided to search and observed the low value update their beliefs in favour of their initial knowledge. This will make them more pessimistic about finding the high value and more reluctant to search. We show that the robust search rule must mimic this optimal behaviour. It is interesting that when ε goes to zero, the robust search rule converges to the one that is stationary. The case without initial information, so where $\varepsilon = 0$ was analysed by Schlag and Zapechelnyuk (2018), who found that the robust rule must be stationary.

In addition, we propose a simple myopic search rule which is based on the idea that the today's searching strategy influences only current payoff and has no effect

[2] This idea was mentioned in the earlier papers on search as well. In a seminal paper on search and learning Rothschild (1978) states that "it should be possible to compute the loss from following ad hoc rules..." (page 700)

on the losses in the future. We show that the myopic rule prescribes to search more than the robust one. Moreover, we find that the losses from the proposed myopic rule represent a very good approximation of the original minimax losses. For instance, if $\varepsilon > 0.4$, then the losses of the robust search rule cannot be higher than 7,5% of the difference between the high and the low value. In addition, we show that the maximal losses can be also approximated by a rule which requires commitment and has a very simple linear form.

It is important to emphasize that the initial information reduces the losses in the worst case scenario for two reasons. First, it directly reduces the degree of the model uncertainty and makes the robust search strategy closer to the optimal one. Second, the initial information indirectly restricts the possible worst case environments that are not consistent with the history of observation. Therefore, the losses in our model are lower than in Schlag and Zapechelnyuk (2018) and decrease with respect to ε.

In the extensions we also discuss a more general case, in which the values of alternatives are continuously distributed on some interval. It turns out that the worst case scenario is characterized by a situation, in which the searcher becomes less confident in her initial information. So, there exist histories of observations, such that the searcher will become completely agnostic about the true environment. In this case the problem becomes identical to the one studied by Schlag and Zapechelnyuk (2018) and the losses reach their maximal levels. The intuition behind this is based on the possibility of strong positive learning[3] under which the searcher might stop relying on her initial information if some unexpected values were observed. In addition, even if strong positive learning is excluded, then in the worst case her initial information remains almost unchanged and learning might be very slow. As a result, the searcher remains active for more search periods and losses can be very high.

We highlight similarities and differences to Schlag and Zapechelnyuk (2018),

[3] Strong positive learning does not guarantee that the Bayesian search strategies are characterized by the reservation price property. Therefore, it becomes difficult to characterize the optimal strategies. Most of the standard search literature imposes an assumption on the initial prior of the searcher to eliminate strong positive learning (Rothschild (1978), Benabou and Gertner (1993), Dana Jr (1994), Janssen *et al.* (2011))

who also explore the search rules that perform well irrespectively of the priors. Our paper is different from theirs in several important aspects. First, Schlag and Zapechelnyuk (2018) focus on the extreme version of model uncertainty and consider a searcher who is completely agnostic about the true environment. They characterize a robust search procedure that is stationary and does not depend on the history of observations. In contrast, we consider a searcher who does not know the actual distribution, but believes that it is not too different from her initial information. Then, two points become important: *i)* how to make inferences without relying on specific priors and *ii)* whether processing of information is important for a good search strategy. We show that a robust search rule is no longer stationary and adapts to the acquired information. Second, Schlag and Zapechelnyuk (2018) use a payoff ratio as a measure of the search rule performance, whereas we focus on the losses approach, which has a better economic intuition.[4]

More generally, this paper is related to two strands of literature. First, there is a literature on *search and learning* about the unknown value distribution. Most of the literature is focused on the Bayesian learning that characterizes the optimal search strategies for the specified priors (Rothschild (1978), Rosenfield and Shapiro (1981), Dana Jr (1994), Janssen *et al.* (2011)). The choice of prior is a very difficult problem, so our aim is to explore the question of how to search without relying on specific priors. Chou and Talmain (1993) consider a consumer who uses a nonparametric procedure to estimate the probability distribution. Although this procedure does not rely on priors, they do not explore the question of designing an adaptive search rule that performs well for all priors.

This paper relates to the literature on *robustness* where performance of a heuristic strategy is measured with respect to the best strategy for the actual theoretical model. This approach was employed by *i)* Bergemann and Schlag (2011a) in robust monopoly pricing, by *ii)* Manski (2004), Stoye (2011) in statistical treatment choice and by *iii)* Bergemann and Schlag (2011b), Parakhonyak and Sobolev (2015), Schlag and Zapechelnyuk (2018) in robust search. We contribute to this literature by focusing on the question of how to learn without relying on specific

[4] See the discussion about measuring the performance using the payoff differences in the section 2.

priors. There is also a strand of literature that refers to robustness when aiming to maximize the minimal payoffs (Gilboa and Schmeidler (1989)). In this literature there is a debate on how to correctly learn from the past. Parallel to Schlag and Zapechelnyuk (2018) we avoid any debate by letting the Bayesians learn and then looking for a rule that they would accept as being close to optimal under any circumstances.

The rest of the paper is organized as follows. Section 2 describes the model with binary-valued payoffs and the performance criterion. In section 3, we characterize a dynamically robust search rule for given initial information. In addition, we explore a myopic and linear rule and compare their relative performance. In section 4, we discuss an extension of the model with the general distribution of payoffs. Section 5 concludes. Technical proofs are presented in the Appendix.

4.2 Model

Preliminaries. Consider a decision maker who is engaged in costly sequential search. The searcher is interested in finding the highest value of alternatives $x_1, x_2, \ldots$, where each alternative x_i can take either a high or a low value that we normalize to 0 and 1 respectively, so $x_i \in \{0, 1\}$[5]. Prior to search the decision maker has an outside option with the positive value $a > 0$. We assume that alternatives that were rejected can be accepted later, so there is *perfect recall*.

The cost of search is modelled similar as in Schlag and Zapechelnyuk (2018). All found alternatives are discounted over time at a discount factor $\delta \in (0, 1)$. So the value of the outside option in the search period k is $\delta^k a$. Therefore, if a is not accepted in the period k and the decision maker continues to search, then she effectively pays $(1 - \delta)a$ which can be seen as the search cost.

Let $\mathbb{B}$ be the set of all probability distributions on the binary support $\{0, 1\}$, so $\mathbb{B} = \Delta(\{0, 1\})$. We assume that the alternatives x_i are i.i.d. distributed according to some Bernoulli distribution $B^\lambda \in \mathbb{B}$, such that x_i takes the value 1 with probability λ and the value 0 with probability $1 - \lambda$. We will refer to B^λ as the

[5] Without loss of generality, the model can be similarly adapted to any set of binary distributions on some predetermined support $\{\omega, z\}$, where $\omega < z$.

environment. Importantly, this distribution is unknown to the searcher.

Suppose that the searcher has some initial belief described by a distribution over elements in $\mathbb{B}$. Since a compound lotteries are equivalent to simple lotteries, the initial information can be equivalently represented by an element of $\mathbb{B}$. Let assume that the searcher's initial belief is described by $[B^{\lambda_0}]$, where $[\cdot]$ denotes a point belief. The searcher understands that her initial information might not be a precise description of the true environment. We introduce model uncertainty by considering ε-*contamination* of the searcher's initial knowledge B^{λ_0}. Precisely, the searcher believes that the values are distributed according to B^{λ_0} with probability that cannot be lower than $\varepsilon \in [0, 1]$. With the remaining probability the correct prior $\mu \in \Delta(\mathbb{B})$ is not known. Notice, that ε measures the degree of model uncertainty. If ε is high, then the degree of uncertainty if low. In this case, the searcher is significantly confident in his initial information and the model uncertainty has a moderate impact on her decisions. If $\varepsilon = 0$, then the searcher is fully agnostic about the true environment.

The set of all environments for a given initial belief B^{λ_0} and the degree of model uncertainty ε is:

$$\mathcal{F}_{\varepsilon}(B^{\lambda_0}) = \{F = \varepsilon[B^{\lambda_0}] + (1 - \varepsilon)\mu | \mu \in \Delta(B)\}$$

In other words, $\mathcal{F}_{\varepsilon}(B^{\lambda_0})$ is the $\varepsilon-$contamination neighbourhood of B^{λ_0}. Notice that the $\varepsilon-$contamination neighbourhood for a lower ε includes the $\varepsilon-$contamination neighbourhood of a higher ε.

Denote $\gamma = \frac{a(1-\delta)}{\delta(1-a)}$. We assume that absent the model uncertainty and based on the initial beliefs $[B^{\lambda_0}]$, the searcher would prefer to stop and accept the outside option a:

Assumption 4.1. *$\lambda_0 < \gamma$.*

Assumption 4.1 implies that the expected gain of search is higher than the outside option[6], when the true environment is represented by B^{λ_0}.

---

[6] The expected gain of search is equal to $\delta(\lambda_0 + (1 - \lambda_0)a)$ and higher then a if and only if $\lambda_0 < \gamma$.

Performance Criterion

The searcher takes into account that her initial information might be misspecified and she is interested in finding a search strategy that would perform approximately optimal for all priors in $\mathcal{F}_\varepsilon(B^{\lambda_0})$. In other words, the searcher tries to find a search rule that would approximate the performance of the optimal rule of an "informed" searcher who knows the true $\mu(\lambda)$ and learns over time.

In every search period, the searcher has to decide whether to stop or continue to search. Let $h_i = \{x_1, ...x_i\}$ be a history of observation prior to round i. Every search strategy s can be represented as the profile of stopping probabilities that depend on the number of search period and the history of observations. Let $s = (s(h_k))_{k=0}^{\infty}$ be a search strategy, in which $s(h_k)$ determines the probability of stopping and let $\mathcal{S}$ be a set of all possible stopping rules. Notice, that in the case of binary valued payoffs, the stopping probability s_k depends only on the number of the period k. Obviously, if the searcher observes 1, she will find it optimal to stop irrespectively of her priors. The only case in which the searcher considers to continue her search in the period k is when the history h_k consists of k zeros. Given $s \in \mathcal{S}$, denote $s^k = (s_i)_{i=k}^{\infty}$ as the profile of stopping probabilities starting from the search the period k. The objective of the searcher is to design a search strategy s that would approximate the maximal payoff obtained after any history by the bayesian searcher who knows the true underlying environment.

Consider some prior $F \in \mathcal{F}_\varepsilon(B^{\lambda_0})$. Denote $U_k^*(s^k, F)$ as the expected discounted payoff of the search rule s^k if the searcher's prior is described by F. Then, the optimal payoff for F is given by:

$$V_k^*(F) = \sup_{s^k \in \mathcal{S}} U_k^*(s^k, F).$$

We will measure the performance of any search rule s for a given prior by measuring the difference between its performance and the maximal performance that could have been obtained. We refer to this difference as the loss function:

$$L_k^*(s^k, F) = V_k^*(F) - U_k^*(s^k, F) \tag{4.1}$$

The aim of the searcher is to design a search rule that would minimize the losses for all possible priors in $\mathcal{F}_\varepsilon$ in each period of time after all possible history of observations. We assume that the searcher does not have commitment power and has to anticipate her search strategy in the future. Therefore, we determine a *dynamically robust* search strategy $\hat{s}$ recursively. In each round the searcher decides upon the stopping probability s_k, by taking the future robust strategy $\hat{s}^{k+1}$ as given. A dynamically robust search strategy must approximate the optimal one in terms of its performance for all possible priors $F \in \mathcal{F}_\varepsilon(B^{\lambda_0})$. next we formally define the dynamically robust search strategy:

Definition 4.1. *A dynamically robust search strategy minimizes the maximal losses given by (4.1) in all search periods by taking the future strategy as given:*

$$\hat{s}_k \in \operatorname*{argmin}_{s_k} \max_{F \in \mathcal{F}_\varepsilon(B^{\lambda_0})} L_k^*(s_k, \hat{s}^{k+1}, F) \qquad \forall k \tag{4.2}$$

We will refer to this problem as P^o.

Motivation. The the robust search strategy has to approximate the maximal possible performance irrespective of the true priors. Equilibrium concept involving the idea of being close to optimal was originally defined in economics as epsilon Nash Equilibitum (Radner (1982)) and subgame perfection (Mailath *et al.* (2005)). Both of these concepts use the approach of the payoff difference rather than the payoff ratio. It is noteworthy, that if the maximal loss from using the dynamically robust rule $\hat{s}$ given by (4.2) over all periods is small then $\hat{s}$ is an ε-optimal search strategy for any bayesian searcher with a prior in the set of feasible priors.

Since the losses are based on the difference, they can be interpreted as costs of having imprecise initial information. Moreover, the additivity allows to look at the expected losses when the searcher thinks about the future losses. The payoff ratios considered by Schlag and Zapechelnyuk (2018) are convenient when payoffs are unbounded or range is large. However, in our binary setting, the range is well defined and the difference approach can be well applicable.

Finally, the losses approach is related to the concept of regret (Milnor (1951)), except that we look at the ex ante problem and regret is modelled expost. These two approaches are identical if one considers priors to be states. Note, that regret as defined by Wald (1950) and axiomatized by Milnor (1951) is additive.

4.3 Analysis

We start our analysis by showing that when deriving a dynamically robust search rule $\hat{s}$ defined in (4.2), it is sufficient to consider only two-point priors.

Two point priors. To show that there is no loss of generality in restricting attention to two-point priors, we define a subset of $\mathcal{F}_\varepsilon(B^{\lambda_0})$, which includes the priors that take the form of B^{λ_0} with probability at least ε and some $B^\lambda \in \mathcal{B}$ with the remaining probability. So, let $\mathcal{F}^2_\varepsilon(B^{\lambda_0}) \subset \mathcal{F}_\varepsilon(B^{\lambda_0})$ be the $\varepsilon-$contamination neighbourhood of B^{λ_0}, which contains only two-point priors. Formally,

$$\mathcal{F}^2_\varepsilon(B^{\lambda_0}) = \{F = \varepsilon[B^{\lambda_0}] + (1-\varepsilon)[B^\lambda] | \mu \in B\}$$

Next proposition shows that a search strategy $s \in \mathcal{S}$ that guarantees a certain foregone payoff for *two-point* priors in the set $\mathcal{F}_\varepsilon$, can guarantee the same foregone payoff for the whole set $\mathcal{F}_\varepsilon(B^{\lambda_0})$.

Proposition 4.1. *For each search strategy $s \in \mathcal{S}$ and for all B^{λ_0}:*

$$\max_{F \in \mathcal{F}_\varepsilon} L^*_k(s^k, F) = \max_{F \in \mathcal{F}^2_\varepsilon} L^*_k(s^k, F) \qquad \forall k.$$

Intuitively, the maximal loss that is attained by the distribution over different environment, can be attained by some other environment. We will use this result to simplify the objective of the searcher determined by (4.2).

To simplify notations, let us denote $U_k(s^k, \lambda)$ and $V_k(\lambda)$ as the expected payoff of the search rule s^k and the maximal payoff respectively, if the searcher's prior is described by $F = \varepsilon[B^{\lambda_0}] + (1-\varepsilon)[B^\lambda]$. Consequently, define the losses as $L(s^k, \lambda) = L^*(s^k, F)$. Then the dynamically robust rule can be characterized as the solution of the following simplified problem:

$$\hat{s}_k \in \operatorname*{argmin}_{s_k} \max_{\lambda} L_k(s_k, \hat{s}^{k+1}, \lambda) \qquad \forall k \tag{4.3}$$

Next we show that search will terminate after a finite number of rounds when using a dynamically robust search strategy.

Finiteness of search. Consider a searcher who believes that the true environment is given by $F = \varepsilon[B^{\lambda_0}] + (1-\varepsilon)[B^\lambda]$. Notice, that if $\lambda \le \gamma$, then the expected payoff from searching is always smaller than the outside option a and it is optimal to stop [7]. If $\lambda > \gamma$, then the searcher might want to continue if ε is high enough. Denote $\overline{\varepsilon}(\lambda)$ as the threshold level of probability that the true distribution is B^{λ_0}, such that the searcher would continue to search for all $\varepsilon < \overline{\varepsilon}(\lambda)$ and stop if $\varepsilon \geq \overline{\varepsilon}(\lambda)$. The indifference condition of the searcher that equalizes the expected gain of search and the search costs is given by:

$$(\overline{\varepsilon}\lambda_0 + (1-\overline{\varepsilon})\lambda)\delta(1-a) = (1-\delta)a$$

The searcher will observe 1 in the next period with probability λ_0 if the true environment is B^{λ_0} and with probability λ if it is B^{λ}. Solving for $\overline{\varepsilon}$ we get that

$$\overline{\varepsilon}(\lambda) = 1 - \frac{\gamma - \lambda_0}{\lambda - \lambda_0}. \tag{4.4}$$

If $\lambda > \gamma$, then the cutoff level $\overline{\varepsilon}(\lambda) > 0$ and the searcher will continue to search for small enough ε. Consider the problem of the searcher who finds it optimal to search in round k after observing k zeros. Let $\varepsilon_k(\lambda)$ be the posterior probability of B^{λ_0} in the period k after observing k zeros:

$$\varepsilon_k(\lambda) = \frac{\varepsilon}{\varepsilon + (1-\varepsilon)\left(\frac{1-\lambda}{1-\lambda_0}\right)^k}$$

The following lemma shows that search cannot last for infinite time.

Lemma 4.1. *For every $F = \varepsilon[B^{\lambda_0}] + (1-\varepsilon)[B^{\lambda}] \in \mathcal{F}^2_{\varepsilon}(B^{\lambda_0})$, there exists $T(\lambda)$, such that it is never optimal to search more than $T(\lambda)+1$ times.*

As long as $\lambda > \gamma$, it is more likely to observe zero when the true environment is $[B^{\lambda_0}]$. Therefore, each additional zero makes the searcher more pessimistic about the finding 1 and she will terminated her search after a finite number of search periods. The class of environments with binary valued payoffs satisfies the assumption of decreasing reservation values. This assumption excludes strong positive learning and is very common in the literature on search with learning (Rothschild (1978), Benabou and Gertner (1993), Adam (2001)).

[7] To show this, consider a searcher with the a point prior $[B^{\lambda}]$. Let V be his value of searching, so $V = \max\{a, \delta(\lambda + (1-\lambda)V)\}$. By solving for V we get $V = \delta\lambda/(1-\delta(1-\lambda))$ which is lower or equal to the outside option a if and only if $\lambda \leq \gamma$. Since it is optimal to stop for priors $[B^{\lambda_0}]$ and $[B^{\lambda}]$, it is optimal to do so for any two-point prior that is a convex combination of these two.

Let T be the maximum number of zeros after which there still exist some bayesians characterized by λ, for whom it is optimal to search. In other words, $T+1$ is the smallest number of zeros, after which all searchers would prefer to stop searching and accept the outside option. T can be found by solving the following maximization problem

$$\max_{\lambda,T}\left\{\lfloor T\rfloor \text{ s.t. } \varepsilon_T(\lambda)=\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\left(\frac{1-\lambda}{1-\lambda_0}\right)^T}\leq\bar{\varepsilon}(\lambda)=1-\frac{\gamma-\lambda_0}{\lambda-\lambda_0}\right\}, \tag{4.5}$$

where $\lfloor x\rfloor$ is the floor of x. Consider some fixed λ. The searcher finds it optimal to continue if $\varepsilon_k(\lambda)<\bar{\varepsilon}(\lambda)$. We choose $T(\lambda)$ such that one additional 0 would induce the searcher to accept a, so $\varepsilon_T(\lambda)<\bar{\varepsilon}(\lambda)$ and $\varepsilon_{T+1}(\lambda)\geq\bar{\varepsilon}(\lambda)$. By equating $\varepsilon_T(\lambda)=\bar{\varepsilon}(\lambda)$, we find $T(\lambda)$ which is further maximized with respect to λ:

$$T=\max_{\lambda>\gamma}\left\lfloor \ln\left(\frac{\varepsilon}{1-\varepsilon}\frac{\gamma-\lambda_0}{\lambda-\gamma}\right)\Big/\ln\left(\frac{1-\lambda}{1-\lambda_0}\right)\right\rfloor. \tag{4.6}$$

Note, that if $\lambda<\gamma$, then the searcher would find it optimal to stop immediately and $T=0$. Consider the properties of T. If the searcher is dealing with the high degree of uncertainty (low ε), then she assigns a higher probability to the event of receiving 1 in the future and T becomes higher. Similarly, if the outside option a is low, then the search costs become lower and the expected gain of search is higher which leads to more search rounds. On the contrary, if δ is higher, then the searcher values future findings less and the maximal number of search rounds decreases. Similarly, if λ_0 increases, then it becomes less probable to receive 1 in the future which makes the searcher stop earlier.

The following lemma summarizes the properties of T with respect to all parameters of the model:

Lemma 4.2. *Let T be a solution of (4.6). Then T is weakly increasing in ε and a, and weakly increasing in δ and λ_0.*

4.3.1 Robust search rule

In this section we characterize the optimal rule defined in (4.3). For convenience, we denote the probability of receiving one after observing k zeros as:

$$\phi_k(\lambda)=(1-\varepsilon_k(\lambda))\lambda+\varepsilon_k(\lambda)\lambda_0. \tag{4.7}$$

Last period. Since searching cannot last for more than $T+1$ periods, it is optimal to stop in the period $T+1$ irrespectively of priors. The robust strategy must prescribe the searcher to stop, so $\hat{s}_{T+1} = 1$ and losses will be zero, so $L_{T+1}(s^{T+1}, \lambda) = 0$.

Search period k. To characterize the robust stopping probability $\hat{s}_k$ in the period k, we have to solve (4.3) by backwards induction. Suppose that we have characterized the robust strategy $\hat{s}^{k+1} = \hat{s}_{k+1}, \hat{s}_{k+2}, \ldots, \hat{s}_{T+1}$ and now we are interested in characterizing the robust probability of stopping $\hat{s}_k$ in round k.

We define the expected payoff of searching in the period k given that the searcher follows the optimal strategy from then on, as $X_k^V(\lambda)$. The upper index V indicates that the searcher plays optimally starting from the period $k+1$. In the period $k+1$, the searcher receives 1 with probability $\phi_k(\lambda)$ and continue to behave optimally with the remaining probability: $X_k^V(\lambda) = \delta(\phi_k(\lambda) + (1-\phi_k(\lambda))V_{k+1}(\lambda))$. The informed searcher compares the expected payoff from searching with the outside option so the maximal payoff in round k is $V_k(\lambda) = \max\{a, X_k^V(\lambda)\}$.

Denote $X_k^U(s^{k+1}, \lambda)$ as the expected payoff of searching in round k when the searcher follows s^{k+1} starting from the period $k+1$. The value of a searcher who plays the stopping rule s_k in round k is $U_k(s_k, \hat{s}^{k+1}, \lambda) = s_k a + (1-s_k)X_k^U(\hat{s}^{k+1}, \lambda)$. The loss function in the period k is the difference $L_k(s_k, \hat{s}^{k+1}, \lambda) = V_k(\lambda) - U_k(s_k, \hat{s}^k, \lambda)$ that can be rewritten in a very useful form:

$$L_k(s_k, \hat{s}^{k+1}, \lambda) = \underbrace{s_k \max\{X_k^V(\lambda) - a, 0\}}_{\text{early stopping}} + \underbrace{(1-s_k)\max\{a - X_k^V(\lambda), 0\}}_{\text{excessive searching}} + \\ + \underbrace{(1-s_k)\delta(1-\phi_k(\lambda))L_{k+1}(\hat{s}^{k+1}, \lambda)}_{\text{future losses}} \quad (4.8)$$

The first two terms describe the losses from not playing optimally in the period k only. The first term represents the losses from *too early stopping* and loosing the maximal value in the period k. The second term captures the losses from *excessive searching* in the period k when it is optimal to stop. The last term represents the *future expected losses* from playing strategy s^{k+1} instead of playing the optimal strategy. The future losses are positive if and only if 0 is observed in the period $k+1$, which happens with probability $1-\phi_k(\lambda)$.

Consider all Bayesian searchers characterized by different beliefs $[B^\lambda] \in \mathbb{B}$, for which it is optimal to search, so $V_k(\lambda) > a$. Among them, let us find the prior $[B^{\hat{\lambda}_k}]$, for which the losses are the maximal. In other words, $\hat{\lambda}_k$ maximizes (4.8) subject to $V_k(\lambda) > a$.

The loss function (4.8), in which all sources of losses are separated from each other, allows us to solve the searcher problem (4.3). In the next proposition we characterize the dynamically robust search rule and provide an upper bound for the maximal losses.

Proposition 4.2. *Let T be the last period in which it might be optimal to search (4.6). Then the dynamic robust search rule is given by:*

$$\begin{cases} \hat{s}_k = 1 - \frac{X_k^V(\hat{\lambda}_k)-a}{X_k^V(\hat{\lambda}_k)-X_k^V(0)+D} & \text{if } k \leq T, \\ \hat{s}_k = 1 & \text{if } k > T, \end{cases} \tag{4.9}$$

where D is the difference in the future losses:

$D = \delta[(1-\phi_k(0))L_{k+1}(\hat{s}_{k+1}, 0) - (1-\phi_k(\hat{\lambda}_k))L_{k+1}(\hat{s}_{k+1}, \hat{\lambda}_k)]$.

This rule minimizes the maximal losses that cannot be higher than

$$\hat{L}_k < (1-\delta)a \sum_{j=k}^{T} \left(\delta - \frac{\delta \sum_{i=k}^{j} \hat{s}_i}{j-k}\right)^{j-k}. \tag{4.10}$$

First, note that the dynamically robust search strategy (4.9) is not stationary and depends on k. This is different from the stationary robust search strategies described by Schlag and Zapechelnyuk (2018). The main difference is that in their paper, Schlag and Zapechelnyuk (2018) consider searchers who are fully agnostic about the environment. In our paper, we assume that searcher have some initial information. As a result, in each round, a searcher processes newly acquired information and makes inferences about the true environment. This, in turn, affects her search strategy.

Second, the robust search rule is stochastic. Precisely, after searching T times all informed searchers would stop and therefore the robust rule prescribes to stop as well, but in every period $k \leq T$, a searcher must randomize. This property was also observed in Bergemann and Schlag (2008), Parakhonyak and Sobolev (2015) and Schlag and Zapechelnyuk (2018). The intuition behind this property can be understood as follows. Suppose a searcher observed k zeros and has to decide whether to stop and accept a or continue in a hope of receiving 1 in the future. The optimal strategy of an informed searcher in the period k is deterministic. If a searcher decides to search and observes one more 0, then the maximal losses come from excessive searching in the period k and in the future. On the other hand, if the searcher decides to stop, then the maximal losses are based on the fact that she could have found 1 in the future. Clearly, any deterministic strategy that is used by the informed searcher is vulnerable to these two sources of losses and can be exploited against the searcher. Therefore, a dynamically robust strategy must resolve the trade-off between these two sources of losses.

The proof starts from the consideration of the last period T and goes backward to any $k < T$. If the searcher decided to continue then the losses are maximal when the chances of finding 1 are the lowest. This is the case when $\lambda = 0$. On the contrary, if the searcher decides to stop, then the worst case distribution maximizes the losses from too early stopping and the losses from the future.

Intuitively, the losses from stopping are not maximized when $\lambda = 1$. If $\lambda = 1$, then the searcher will believe with probability 1 that the true distribution is F_0 and it is optimal to stop, so $\varepsilon_k(\lambda = 1) = 1$ and $L_k = 0$. Instead, the nature should keep some masses on 0 to convince the searcher that the true distribution is F and at the same time, it should allocate the rest of the probability masses on 1 to make the expected payoff from searching as high as possible.

Interestingly, the searcher observing k zeros will never believe in finding 1 with very high probability and therefore the losses from stopping cannot be as big as in the case of $\varepsilon = 0$, considered in Schlag and Zapechelnyuk (2018). The initial information creates the tradeoff between convincing the searcher and making the environment attractive, which puts a restriction on the set of possible worst case environments. This unambiguously leads to the improved performance of the robust search rule. It is easy to see that the upper bound of the maximal losses given by (4.10), decreases in ε, since $\hat{s}_k$ are higher for higher ε.

Next, an important property of the search rule is that $\hat{s}_k$ increases in k, which means that the more zeros are observed by the searcher, the higher the losses from searching and therefore the searcher stops with a higher probability. Since ε_k increases in k with an increasing rate, the informed searcher's disappointment of finding 1 grows faster for every additional zero, because being unlucky and observing many zeros from a good environment becomes less and less likely. This intuition is carried over in the robust search rule that incorporates more pessimism of finding 1 over time. Figure 1 depicts the stopping probabilities of $\hat{s}_k$ given by (4.9). The probability of stopping increases over time and $\hat{s}_k$ is convex that implies that pessimism grows at increasing rate.

Let us consider the intuition behind the comparative statics of $\hat{s}_k$. First, if ε increases, then the searcher relies more on his initial information, which suggests that the true environment is F_0 and it is not that attractive for searching by Assumption 4.1. Therefore, in order to be protected against the increased losses from excessive searching, it is optimal to stop more often, so $\hat{s}_k$ becomes higher. Second, if the outside option a becomes more attractive, then the effective search costs become higher and searching becomes more expensive. The robust search strategy $\hat{s}_k$ takes this into account and prescribe to be less active in searching. Third, if the discount factor δ of the searcher increases and the future becomes more important, we show that the losses from stopping become more pronounced. In order to balance out the two sources of losses, the searcher has to search more frequently which implies that $\hat{s}_k$ decreases in δ. in addition, if λ_0 becomes higher, then searching becomes more attractive which lead to higher losses from stopping. To account for the increased searching incentives, $\hat{s}_k$ becomes higher.

Next, we explore the comparative statics of the minimax losses L_k with respect to all parameters of the model. The blue lines on the Figure 2 represent the minimax losses in dependence of all parameters of the model. Consider first how L_k depends on ε. If the degree of model uncertainty becomes lower (increased ε), then the size of the losses are lower because of the two reasons. First, a higher ε makes the search problems of the informed and uninformed searcher more similar, which directly shrinks the possible losses. Second, a higher ε restricts the possible worst cases even further, because the nature should keep even more probability masses on 0, to convince the searcher.

Figure 2(b) shows that mimimax losses increase in a for small a and then decrease in a for high a. If the outside option increases there are two effect on the maximin losses. First, the losses become higher due to higher search costs. Second, since it becomes more costly to search, the searcher stops with a higher probability, so the expected losses are lower. If a is small than the first effect dominates.

4.3.2 Myopic search rule

The robust search strategy (4.9) is given recursively and therefore the minimax losses L_k is hard to evaluate. In this subsection, we show that the problem, in which the searcher believes that the future losses are not affected by a current search strategy can provide a very good approximation of the minimax losses in the problem P^o.

Suppose that a searcher believes that tomorrow the nature will be able to change F, so the losses are higher. Let $\tilde{s}^{k+1}$ be the strategy that balances out the losses from stopping and searching in periods following period k. The loss function in this game is given by:

$$\begin{aligned}\tilde{L}_k(s_k, \tilde{s}^k, \lambda) = s_k \max\{X_k^V(\lambda) - a, 0\} + (1 - \tilde{s}_k) \max\{a - X_k^V(\lambda), 0\}+ \\ + (1 - s_k) \max_{\psi}\{\delta(1 - \phi_k(\psi))\tilde{L}_{k+1}(\tilde{s}^{k+1}, \psi)\}\end{aligned}$$

Since the probability of receiving zero is the highest when $\psi = 0$ and the losses in every period cannot be higher then the minimax losses from stopping, we have that:

$$\begin{aligned}\tilde{L}_k(s^k, \lambda) = s_k \max\{X_k^V(\lambda) - a, 0\} + (1 - s_k) \max\{a - X_k^V(\lambda), 0\}+ \\ + (1 - s_k)\delta(1 - \phi_k(0))\tilde{L}_{k+1}(\tilde{s}^{k+1}, 0) \quad (4.11)\end{aligned}$$

Notice, that the future losses in (4.11) do not depend on the true distribution characterized by λ.

We call $\tilde{s} = (\tilde{s}_k)_{k=1}^T$ as a "myopic" rule, since it does not take into account its influence on the future losses, by assuming that the true distribution can change in the future. Next we define it formally:

Definition 4.2. *A myopic search strategy minimizes the maximal losses given by (4.11) in all search periods by taking the future strategy as given:*

$$\tilde{s}_k = \operatorname*{argmin}_{s_k} \max_{\lambda} \tilde{L}(s_k, \tilde{s}^{k+1}, \lambda) \qquad \forall k \qquad (4.12)$$

We will refer to this problem as P^s. The myopic search rule $\tilde{s}_k$ should balance out the losses from searching and stopping without taking into account its influence on the future losses. Let $\tilde{\lambda}_k$ characterizes the distribution for which the value of informed searcher $V_k(\lambda)$ is maximal. Since the losses from searching is maximized when the probability of receiving 1 is minimal, we have that

$$\tilde{s}_k(V_k(\tilde{\lambda}_k) - a) = (1 - \tilde{s}_k)(a - X_k^V(0)) = (1 - \tilde{s}_k)(a - \delta(a + \phi_k(0)(1 - a)))$$

This equation determines the myopic search probability.

Similarly to the analysis of the robust search rule, denote $\tilde{\lambda}_k$ as the solution of the maximization problem of the losses given by (4.11) with respect to λ and subject to $V_k(\lambda) = a$. The next proposition fully characterizes the myopic search rule that solves (4.12).

Proposition 4.3. *The myopic rule that solves (4.12) is given by:*

$$\tilde{s}_k = \frac{a - \delta(a + \phi_k(0)(1 - a))}{V_k(\tilde{\lambda}_k) - \delta(a + \phi_k(0)(1 - a))}. \tag{4.13}$$

The myopic rule prescribes to search more than the dynamically robust search rule defined in (4.3): $\tilde{s}_k < \hat{s}_k$ for all k.

Moreover, the minmax losses (4.8) for the robust search rule (4.9) are bounded above by the minmax losses (4.11) for the myopic search rule (4.12).

There are several important properties of $\tilde{s}_k$. First, stopping probability $\tilde{s}_k$ in the period k does not depend on the stopping probabilities in the future $\tilde{s}^{k+1}$. Therefore, the myopic rule is much easier to calculate since the it does not depend on the future periods and there is no need to work backwards. Since $V_k(\lambda)$ is a decreasing function in k, the myopic rule $\tilde{s}_k$ incorporates more pessimism over time as well as $\hat{s}_k$.

Second, in the problem P^s the searcher has less concern about the future losses than in the problem P^o. Since the search equalize her losses from searching and stopping and the losses from searching in P^s remain the same as in the original problem P^o, we have that the searcher will search more, $\tilde{s}_k < \hat{s}_k$ for all k.

Figure 1 shows the robust search rule $\hat{s}_k$ in blue as well as the myopic one $\tilde{s}_k$ in red. It can be seen that the stopping probability of the myopic rule is lower than of

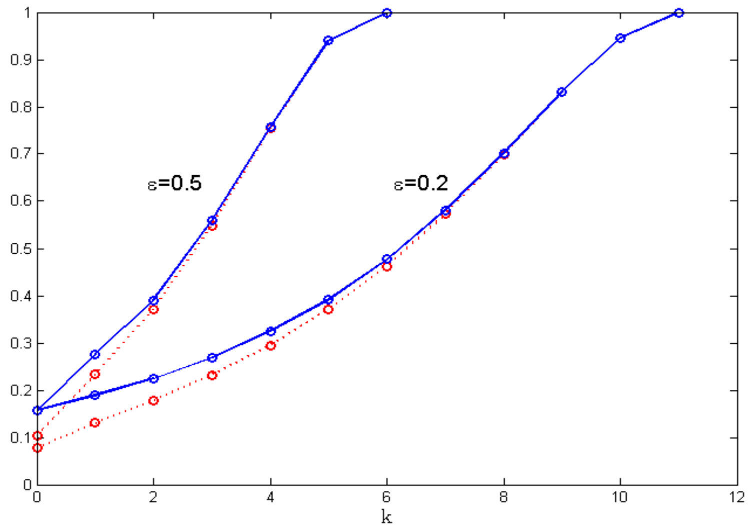


Figure 4.1: The optimal robust rule $\hat{s}_k$(blue) and the myopic rule $\tilde{s}_k$(red), $a = 0.2$, $\delta = 0.8$, $\lambda_0 = 0$.

the robust rule. The difference between the rules becomes more pronounced in the beginning of the search process because then the future losses are more important than in the last rounds.

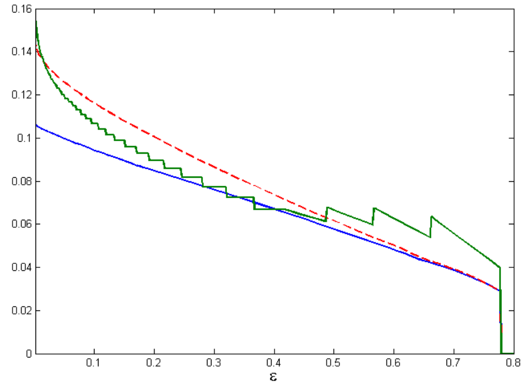
Third, the losses of the robust rule (4.8) are bounded above by the losses of the myopic rule (4.11). The minimax losses $\tilde{L}_k$ highest for $\lambda_0 = 0$ and are given by the recursive formula:

$$\tilde{L}_k = (1 - \tilde{s}_k)((1-\delta)a + \delta \tilde{L}_{k+1}) < (1-\delta)a \sum_{j=k}^{T} \left(\delta - \frac{\delta \sum_{i=k}^{j} \tilde{s}_i}{j-k} \right)^{j-k}.$$

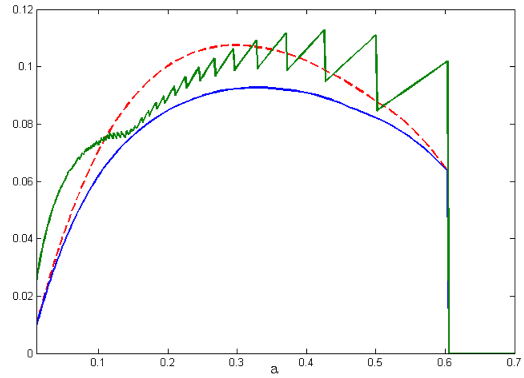
The minimax losses for the robust rule $\hat{s}_k$ and myopic rule $\tilde{s}_k$ are presented in figure 2 in blue and red colors respectively. It can be notices that the myopic rule provides a very good approximation of the robust search rule in terms of its performance for all parameters of the model. The difference between the red and the blue curves are the highest when the search costs $a(1-\delta)$ are the highest and when the number of the search rounds T is big (small ε, high δ).

4.3.3 Linear search rule

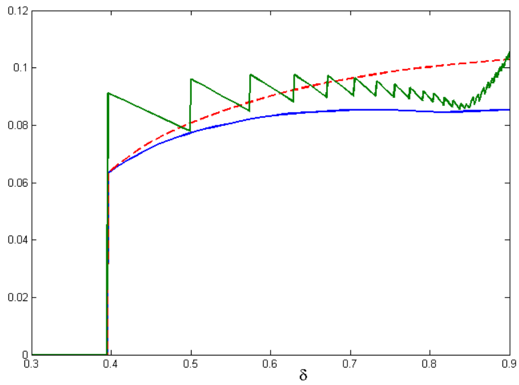
In the previous subsection we explored the myopic search rule which is simple and can well approximate the performance of the dynamically robust search rule. The



(a) Losses in ε, $a = 0.2$, $\delta = 0.8$, $\lambda_0 = 0$



(b) Losses in a, $\varepsilon = 0.2$, $\delta = 0.8$, $\lambda_0 = 0$



(c) Losses in δ, $\varepsilon = 0.2$, $a = 0.2$, $\lambda_0 = 0$

Figure 4.2: Minimax Losses (blue), Upper bound $\tilde{L}$ (red), Losses of the linear rule (green)

robust and myopic search rules do not require commitment of the searcher as in each round the future losses are computed based on the outcome of the minimax problems solved in the future.

In this section we assume that the search can commit to some search rule and follow it the course of search. We propose a very simple linear rule and we show numerically that this rule performs very well with respect to the robust search rule. Let s_k^ℓ be a linear stopping probability in the period k given by the following equation

$$s_k^\ell = \frac{k+1}{T+2} \tag{4.14}$$

We call this rule linear because the probability of stopping increases linearly with respect to the number of zeros in the history of observations. Second, the probability of stopping is inversely related to the number of maximal periods of search. Figure 2 depicts the performance of this rule with respect to the all parameters of our model. One can observe that s_ℓ performs very well with respect to the minimax losses of the robust search strategy $\hat{s}$. Its performance is worse for very small number of periods and when the number of search periods T is very large, so the optimal curve must be convex.

	Losses=V-U		
	$\varepsilon = 0.2$	$\varepsilon = 0.4$	$\varepsilon = 0.6$
	$T = 10$	$T = 5$	$T = 2$
Robust rule $\hat{s}$	0.0847	0.0669	0.0483
Myopic rule $\tilde{s}$	0.1003	0.0737	0.0501
Linear rule s^ℓ	0.0894	0.0667	0.0626

Table 4.1: Comparisons of the rules for $a = 0.2$, $\delta = 0.8$, $\lambda_0 = 0$.

The Table 1 shows how all considered rules perform for different degree of model uncertainty. Since the linear rule encompasses commitment, it performs well with respect to the robust and myopic rules.

4.4 Extension

In this extension we explore the case of the general payoffs distributed on the compact support $[0, 1]$. Let $\mathcal{F}$ is the set of all probability distribution on $[0, 1]$. As in the model with binary-valued payofs, assume that a searcher has some initial information described by $F_0 \in [0, 1]$ and believes that it is true with probability ε. We assume that if $\varepsilon = 1$, then F_0 is such that it is optimal to stop and accept the outside option a. With the remaining probability, she does not know the true distribution but believes that $F \in \mathcal{F}$. The two-point prior of a Bayesian searcher who knows F are given by:

$$F_\varepsilon(F) = \varepsilon[F_0] + (1 - \varepsilon)[F].$$

Using similar notations, we define $V_k(F_\varepsilon(F), h_k)$ as the maximal expected payoff that can be obtained in the period k when the priors are given by $F_\varepsilon(F)$ and history h_k was observed. If the searcher uses the search rule $s^k = (s_i)_{i=k}^{\infty}$, then the expected payoff is denoted as $U_k(s^k, F_\varepsilon(F), h_k)$. As in the previous section, the problem of the searcher is to maximally approximate the optimal performance or minimize the losses in the worst case scenario. The loss function represents the difference between these performances:

$$L_k(F_\varepsilon(F), h_k) = V_k(F_\varepsilon(F), h_k) - U_k(s^k, F_\varepsilon(F), h_k).$$

Denote $F_{(0,z,\sigma)}$ as a binary distribution on $[0, z]$, such that $\mathbb{P}(z) = \sigma$. It is known from Schlag and Zapechelnyuk (2018) that if $\varepsilon = 0$ and the search is completely agnostic about the actual environment, then for any history h the worst case distribution is represented by a binary distribution of type $F_{(0,z,\sigma)}$:[8]

$$\max_{F \in \mathcal{F}} L(F, h) = \max_{z,\sigma} L([F_{(0,z,\sigma)}], h). \tag{4.15}$$

The intuition of this result stems from the fact that the searcher eventually chooses between searching and stopping. Thus, the worst case distribution must consist only of two values. The lowest value 0 leads to the highest losses in case of searching as well as some high z is the worst case for the searcher who decided to stop.

The next proposition shows that there are histories of observations, for which the losses can reach their maximal possible value in case of no initial information.

Proposition 4.4. *For all $\eta > 0$ there exists some history h_η, after which:*

$$\max_{F} L(F_\varepsilon(F), h_\eta) > \max L([F_{(0,z,\sigma)}], h_\eta) - \eta.$$

Notice, that the minimax losses are always higher for a higher degree of uncertainty ε. This proposition establishes the fact that the losses in the general case for $\varepsilon > 0$ can reach the level of the maximal losses for the case of $\varepsilon = 0$, so eventually, the initial information becomes irrelevant.

The intuition behind this proposition is based on the possibility of strong positive learning and can be captured by a simple example that was illustrated by Rothschild (1978) in his seminal paper on search and learning. Suppose that a

[8] In their paper, Schlag and Zapechelnyuk (2018) work with payoff ratios, but their characterization of the worst case environments similarly applies to the case of payoff differences that is considered in this paper.

searcher observes some alternative with the value x_1. Then, it is always possible to concentrate enough mass on x_1, such that the searcher will believe almost with certainty that x_1 was not drawn from F_0. As a result, this observation creates the worst case scenario, in which the updated $\varepsilon_1 = 0$. Importantly, this can be achieved by a binary distribution. Therefore, the masses of the worst case distribution can be allocated in a way, that *i)* convinces the searcher in the inaccuracy of her initial information and *i)* maximizes the losses from stopping and searching when the searcher is completely agnostic, $\varepsilon = 0$.

Next, impose an assumption on the set of possible priors that excludes the positive learning patterns. The assumption of no strong positive learning was often used in the literature on search and learning (Rothschild (1978), Adam (2001), Dana Jr (1994),Janssen *et al.* (2011)). Under this assumption, the searcher will no longer excludes F_0 based on some history of observations and the maximal losses become dependent on F_0. It can be easily shown similarly to the proposition 4.4, that the losses are maximal if learning goes very slowly[9]. Intuitively, the losses are higher if the degree of model uncertainty remains high over time and foregone payoffs of searching and stopping are high. Depending on F_0 it might be not possible to achieve slow learning and high foregone payoffs at the same time, as it was demonstrated in the case of binary valued payoffs.

4.5 Conclusion

It is often the case that a consumer or a job seeker has to form her search strategy without a clear understanding of the market mechanism and based on little and vague information about the important market fundamentals. Most of the literature on search and learning reduces this uncertainty to risk and proposes the search strategies that are very sensitive to priors. As a result, these strategies perform very poorly if priors are misspecified. In this paper, we focus on the problem of a searcher who has doubts about the correctness of her initial information and therefore is interested in a search rule that would mitigate the losses from an imprecise understanding of the environment.

[9] By slow learning we mean that difference $\varepsilon_{k+1} - \varepsilon_k$ is positive but very close to 0.

In the setting with the binary-valued payoffs, we characterize a robust search rule that prescribes to randomize between searching and stopping up to a certain number of search periods. It is important that the search rule is adaptive and uses the initial information of the searcher. Precisely, it incorporates the pessimism of finding the desired alternative by prescribing to stop with an increasing probability over time. We conclude that the initial information reduces the losses by restricting the set of possible environments that are consistent with an observed history. In addition, we also show that the proposed robust rule can be very well approximated by a myopic and a linear search rules.

In the setting with general distributions of payoffs, we find that there are histories of observations, which make the searcher believe that her initial information is not correct. In this case, there is no advantage of having an initial information and the losses are maximal. This conclusion stems from the ex post analysis in which the losses are required to be minimal for every history. However, some histories might have low probability to occur. Therefore, it is important to explore the robust search rules that perform well for all histories, except maybe for the ones that happen with very low probability.

To sum up, this paper identifies new avenues for the future research. First, the general problem of optimizing that takes priors as an input is difficult and the optimal strategies can be very sensitive to priors. In many practical problems, agents do care about the correctness of their beliefs and might find it too risky to fully rely on them. Instead, they need to understand how specific about their information they have to be in order to get close to the optimum. This brings us to the need of understanding the ε-optimization under diffuse priors.

Second, learning is important in many economic applications. The optimal learning strategies that are based on the Bayes' rule crucially depend on priors that might be imprecise in practice. Therefore, it is very important to understand how to process information and how to make inferences without relying on specific priors.

The last but not least, there are many practical problems in economics that require simple strategies that do not rely on specific assumptions and can guarantee a good performance (e.g price-setting, search, auctions, investments, etc.). One can guess a good strategy and verify how it performs. Practically, it can be very costly and it is not clear how to proceed in case of failure. A more thoughtful approach is to consider a spectrum of heuristic strategies and introduce reasonable criteria to rank them. This approach would allow to search for good heuristic rules as well as to make the first steps towards the theory of rules of thumb.

4.6 Appendix

The appendix contains the proofs of all lemmas and propositions in the main text.

Proposition 4.1. *For each search strategy $s \in \mathcal{S}$ and for all B^{λ_0}:*

$$\max_{F \in \mathcal{F}_\varepsilon} L_k^*(s^k, F) = \max_{F \in \mathcal{F}_\varepsilon^2} L_k^*(s^k, F) \qquad \forall k.$$

Proof. Notice, that every element from $\mathcal{F}_\varepsilon$ can be described as a probability distribution over some two-point priors in $\mathcal{F}_\varepsilon^2$. Since $\mathcal{F}_\varepsilon^2 \subset \mathcal{F}_\varepsilon$, we have that $\mathcal{F}_\varepsilon = \Delta \mathcal{F}_\varepsilon^2$.

Consider the losses for priors $F \in \mathcal{F}_\varepsilon = \Delta \mathcal{F}_\varepsilon^2$:

$$L_k^*(s^k, F) = \int_{F_2 \in \mathcal{F}_\varepsilon^2} L_k^*(s^k, F_2) dF(F_2) \geq \max_{F_2 \in \mathcal{F}_\varepsilon^2} L_k^*(s^k, F_2)$$

The maximal losses $\max_{F \in \Delta \mathcal{F}_\varepsilon^2} L_k^*(s^k, F) \geq \max_{F_2 \in \mathcal{F}_\varepsilon^2} L_k^*(s^k, F_2)$. Since $\mathcal{F}_\varepsilon^2 \subset \mathcal{F}_\varepsilon$ the reverse inequality hold as well. So we established the fact that:

$$\max_{F \in \mathcal{F}_\varepsilon = \Delta \mathcal{F}_\varepsilon^2} L_k^*(s^k, F) = \max_{F \in \mathcal{F}_\varepsilon^2} L_k^*(s^k, F).$$

$\square$

Lemma 4.1. *For every $F = \varepsilon[B^{\lambda_0}] + (1-\varepsilon)[B^\lambda] \in \mathcal{F}_\varepsilon^2(B^{\lambda_0})$, there exists $T(\lambda)$, such that it is never optimal to search more than $T(\lambda) + 1$ times.*

Proof. If $\lambda < \gamma$, then the searcher will stop immediately. Otherwise, according to the assumption 4.1, we have that $\lambda > \lambda_0$. Therefore, $(1-\lambda)/(1-\lambda_0) < 1$ and $\varepsilon_k(\lambda)$ is an increasing function in k that goes to 1 when $k \to \infty$. So there exists a minimum number of zeros $T(\lambda)$ after which $\varepsilon_{T(\lambda)} > \overline{\varepsilon}(\lambda)$ and the searcher decides to stop. This $T(\lambda)$ exists for all λ. Therefore we showed that for all parameters of the model the searcher will stop after a finite number of periods. $\square$

Lemma 4.2. *Let T be a solution of (4.6). Then T is weakly increasing in ε and a, and weakly increasing in δ and λ_0.*

Proof. Consider a slightly different problem of $T(\lambda)$ maximazation, when $\lambda > \gamma$, in which $T(\lambda)$ can be a non-integer number.

$$T(\lambda) = \ln \left(\frac{\varepsilon}{1-\varepsilon} \frac{\gamma - \lambda_0}{\lambda - \gamma} \right) \Big/ \ln \left(\frac{1-\lambda}{1-\lambda_0} \right) \tag{4.16}$$

According the envelope theorem $dT(\lambda^*)/dp = \partial T(\lambda^*)/\partial p$, where p is some parameter. Next, by taking partial derivatives of $T(\lambda)$ with respect to all 4 parameters of the model ε, a, δ, λ_0, we get:

$$\begin{aligned}
&i) && \frac{\partial T(\lambda^*)}{\partial \varepsilon} = \frac{1}{(1-\varepsilon)^2}\left(\frac{\varepsilon}{1-\varepsilon}\ln\left(\frac{1-\lambda}{1-\lambda_0}\right)\right)^{-1} < 0 \\
&ii) && \frac{\partial T(\lambda^*)}{\partial a} = \frac{\gamma}{a(1-a)}\frac{\lambda-\lambda_0}{(\lambda-\gamma)^2}\left(\frac{\gamma-\lambda_0}{\lambda-\gamma}\ln\left(\frac{1-\lambda}{1-\lambda_0}\right)\right)^{-1} < 0 \\
&iii) && \frac{\partial T(\lambda^*)}{\partial \delta} = -\frac{\gamma}{\delta(1-\delta)}\frac{\lambda-\lambda_0}{(\lambda-\gamma)^2}\left(\frac{\gamma-\lambda_0}{\lambda-\gamma}\ln\left(\frac{1-\lambda}{1-\lambda_0}\right)\right)^{-1} > 0 \\
&iv) && \frac{\partial T(\lambda^*)}{\partial \lambda_0} = -\left(\frac{1}{\gamma-\lambda_0}+\frac{T(\lambda)}{1-\lambda_0}\right)\Big/\ln\left(\frac{1-\lambda}{1-\lambda_0}\right) > 0
\end{aligned}$$

T that solves (4.6) is obtained by downward rounding of $T(\lambda^*)$. Therefore, if a parameter p increases by some non-negligible amount, then T either remains the same or changes similarly as $T(\lambda^*)$.

□

Proposition 4.2. *Let T be the last period in which it might be optimal to search (4.6). Then the dynamic robust search rule is given by:*

$$\begin{cases} \hat{s}_k = 1 - \frac{X_k^V(\hat{\lambda}_k)-a}{X_k^V(\hat{\lambda}_k)-X_k^V(0)+D} & \text{if } k \leq T, \\ \hat{s}_k = 1 & \text{if } k > T, \end{cases}$$

where D is the difference in the future losses:

$D = \delta[(1-\phi_k(0))L_{k+1}(\hat{s}_{k+1}, 0) - (1-\phi_k(\hat{\lambda}_k))L_{k+1}(\hat{s}_{k+1}, \hat{\lambda}_k)]$.

This rule minimizes the maximal losses that cannot be higher than

$$\hat{L}_k < (1-\delta)a\sum_{j=k}^{T}\left(\delta - \frac{\delta\sum_{i=k}^{j}\hat{s}_i}{j-k}\right)^{j-k}.$$

Proof. We start the characterization of the robust rule by solving (4.3) from the last period T.

Last period. Suppose T is the maximum number of searches after which there still exist some Bayesians (λ) who are willing to search. Suppose that the searcher knows λ and can compare the expected payoff of searching $X_T^V(\lambda)$ and the outside option a. The expected payoff of searching in the period T is given by $X_T^V(\lambda) = \delta(\phi_T(\lambda) + (1-\phi_T(\lambda))a)$. Therefore, the maximal payoff in the period T is given by $V_T(\lambda) = \max\{a, X_T^V(\lambda)\}$. Suppose that an uninformed searcher who does not know the true λ stops with probability s_T. The performance of this strategy is given by $U_T(s,\lambda) = s_T a + (1-s_T)X_T^V(\lambda)$. Then the losses from playing s_T in round T:

$$L_T(s_T,\lambda) = V_T(\lambda) - U_T(s_T,\lambda) = s_T \max\{X_T^V(\lambda) - a, 0\} + (1-s_T)\max\{a - X_T^V(\lambda), 0\}$$

The first part of the losses is the complaint from *too early stopping*. The second part represents the losses from *excessive searching* when it was optimal to stop. The higher the probability of searching, the higher the losses from excessive search. Similarly, the higher the probability of stopping, the higher the losses from stopping.

Since the robust search strategy should perform well under any priors, it is important to understand for which λ the losses from searching and stopping are maximal. Let $\hat{\lambda}_T$ maximizes the losses from stopping, so $X_T^V(\hat{\lambda}_T) \geq X_T^V(\lambda)$ for all λ. it is easy to check that the losses from searching are maximal when the chances of finding 1 are the lowest, which is the case for $\lambda = 0$. A robust search strategy $\hat{s}_T$ must be chosen to balance out these two sources of losses $\hat{s}_T(X_T^V(\hat{\lambda}_T) - a) = (1-\hat{s}_T)(a - X_T^V(0))$, so

$$\hat{s}_T = \frac{a - X_T^V(0)}{X_T^V(\hat{\lambda}_T) - X_T^V(0)}$$

It is clear that the optimal strategy prescribes to randomize between searching and stopping[10], so $\hat{s}_T \in (0,1)$.

The period k. Suppose that we have characterized the robust stopping probabilities for all periods starting from $k+1$. Now we can solve for the optimal stopping probability in the period k by using the optimal future search rule $\hat{s}^{k+1}$ and the loss function given by (4.8).

Let us characterize the maximal losses for two cases: when the expected payoff of search is larger and smaller than the outside option. First, suppose that $X_k^V(\lambda) \leq a$ and the informed searcher finds it optimal to stop (subscript **S**). Then the losses

[10] T was defined as the round, in which some Bayesians (λ) want to search the last time. Therefore, there exists λ such that it is optimal to search and $X_T(\hat{\lambda}_T) > a$ which implies $0 < \hat{s}_T < 1$. One can notice that the zero-sum game between the malicious nature and the searcher admits only mixed strategy equilibria.

will be represented by the second and third terms of (4.8). Clearly, the losses when it is optimal to stop are maximized when $\lambda = 0$ and chances of finding 1 are minimal:

$$L_k^s(s_k, \hat{s}^{k+1}) = (1 - s_k)(a - X_k^U(\hat{s}^{k+1}, 0)).$$

Second, suppose that it is optimal to continue (subscript **C**), so $X_k^V(\lambda) > a$. Then the losses are represented by stopping too early and not searching optimally starting from the period $k+1$:

$$L_k^c(s_k, \hat{s}^{k+1}, \lambda) = s_k(V_k(\lambda) - a) + (1 - s_k)\delta(1 - \phi_k(\lambda))L_{k+1}(\hat{s}^{k+1}, \lambda)$$

Let $\hat{\lambda}_k$ maximizes the losses when it is optimal to search and the searcher stops with the optimal strategy $\hat{s}_k$, so $L_k^c(\hat{s}^k, \hat{\lambda}) \geq L_k^c(\hat{s}^k, \lambda)$ for all λ. The nature must be indifferent between these two sources of losses $L_k^c(\hat{s}^k, \hat{\lambda}) = L_k^s(\hat{s}^k)$.

The maximal losses. The maximal losses can be evaluated from above:

$$\hat{L}_k = (1 - \hat{s}_k)(a - X_k^U(\hat{s}^{k+1}, 0)) \leq (1 - \hat{s}_k)(a - \delta U_k(\hat{s}^{k+1}, 0)) =$$
$$= (1-\hat{s}_k)((1-\delta)a + \delta\hat{L}_k(\lambda_0 = 0)) = (1-\hat{s}_k)(1-\delta)a + (1-\hat{s}_k)(1-\hat{s}_{k+1}\delta(1-\delta)a + \cdots =$$
$$= (1 - \delta)a \sum_{j=k}^{T} \delta^{j-k} \Pi_{i=k}^{j}(1 - \hat{s}_i) \leq (1 - \delta)a \sum_{j=k}^{T} \left(\delta - \frac{\delta \sum_{i=k}^{j} \hat{s}_i}{j - k}\right)^{j-k}$$

The first inequality is the result of the fact that losses for a given strategy are when the probability of finding 1 is 0. $\square$

Proposition 4.3. *The myopic rule that solves (4.12) is given by:*

$$\tilde{s}_k = \frac{a - \delta(a + \phi_k(0)(1 - a))}{V_k(\tilde{\lambda}_k) - \delta(a + \phi_k(0)(1 - a))}.$$

The myopic rule prescribes to search more than the dynamically robust search rule defined in (4.3): $\tilde{s}_k < \hat{s}_k$ for all k.

Moreover, the minmax losses (4.8) for the robust search rule (4.9) are bounded above by the minmax losses (4.11) for the myopic search rule (4.12).

Proof. i) The myopic search rule can be found by equating the losses from searching and stopping. Let $\tilde{\lambda}_k$ maximizes the losses from searching: $V_k(\tilde{\lambda}_k) \geq V_k(\lambda)$ for all λ. Then $\tilde{s}_k$ is determined by $\tilde{L}_k(\tilde{s}, \tilde{\lambda}) =< \tilde{L}_k(\tilde{s}^k, 0)$. Since the future losses are identical for all λ, then they simply cancel out from this equation and we end up with the equality for today's losses:

$$\tilde{s}_k(V_k(\tilde{\lambda}_k)-a) = (1-\tilde{s}_k)(a-\delta(\phi_k(0)+(1-\phi_k(0))a)) = (1-s_k)a(1-\delta)(1-\phi_k(0)/\gamma)$$

By solving this equation with respect to $\tilde{s}_k$ we get (4.13).

ii) The losses from searching in the new game G^s are less affected by $\tilde{s}_k$ and the losses from stopping remain the same. THe future losses put no longer constraints on the searching strategy in the period k. It is formally shown below. Therefore, a searcher who has to balance out these two sources of losses, has more incentives to search than before, so $\tilde{s}_k < \hat{s}_k$.

iii) Let us prove that the minimax losses of the original game P^o are lower than the minimax losses of the problem P^s.

Let $L^c_k(\hat{s}^k, \lambda)$ and $L^s_k(\hat{s}^k)$ be the losses when it is optimal to continue and stop respectively. The losses from stopping are maximal when $\lambda = 0$, therefore it is omitted in L^s_k.

The proof goes by induction. Since P^o and P^s are equivalent in the period T ($\tilde{L}_{T+1} = L_{T+1} = 0$), then $\tilde{s}_T = \hat{s}_T$ and $\tilde{\lambda}_T = \hat{\lambda}_T$ and $\tilde{L}_T = L_T$.

Consider the period $T-1$. Since $L^c_k(\hat{s}^k, \lambda) < L^c_k(\hat{s}^k, 0)$ and $V_{T-1}(\tilde{\lambda}_{T-1}) \geq V_{T-1}(\hat{\lambda}_{T-1})$, then we can evaluate the losses in the original game from above:

$$\begin{aligned} L^s_{T-1}(\hat{s}^{T-1}) = L^c_{T-1}(\hat{s}_{T-1}, \hat{s}_T, \hat{\lambda}_{T-1}) < \hat{s}_{T-1}(V_{T-1}(\hat{\lambda}_{T-1}) - a) + \\ + (1-\hat{s}_{T-1})\delta(1-\phi_{T-1}(0))L_T(\hat{s}_T, 0) \leq \\ \leq \hat{s}_{T-1}(V_{T-1}(\tilde{\lambda}_{T-1}) - a) + (1-\hat{s}_{T-1})\delta(1-\phi_{T-1}(0))L_T(\tilde{s}_T, 0) = \tilde{L}^c_{T-1}(\hat{s}_{T-1}, \tilde{s}_T, \tilde{\lambda}_{T-1}) \end{aligned}$$

The indifference condition of the nature in the period $T-1$ of the game G^s is $\tilde{L}^c_{T-1}(\tilde{s}^{T-1}, \tilde{\lambda}_{T-1}) = \tilde{L}^s_{T-1}(\tilde{s}^{T-1})$ and it determines $\tilde{s}_{T-1}$.

Since $\tilde{L}^c_{T-1}(\hat{s}_{T-1}, \tilde{s}_T, \tilde{\lambda}_{T-1}) > L^s_{T-1}(\hat{s}^{T-1}) = L^s_{T-1}(\hat{s}_{T-1}, \tilde{s}_T)$ and $\tilde{L}^c_k - \tilde{L}^s_k$ is increasing in s_k, we get that $\tilde{s}_{T-1} < \hat{s}_{T-1}$. So the losses from searching are lower if the searcher stops with probability $\hat{s}_{T-1}$. In other words, in the period T the searcher who plays optimally in P^o, stops too often in the game G^s. Consequently, the losses in the period $T-1$ are higher in the supplementary game:

$$L^s_{T-1}(\hat{s}^{T-1}) < L^s_{T-1}(\tilde{s}_{T-1}, \hat{s}_T) = \tilde{L}^s_{T-1}(\tilde{s}_{T-1}, \tilde{s}_T)$$

Inductive Step. Next assume that $L^s_{k+1}(\hat{s}^{k+1}) < \tilde{L}^s_{k+1}(\tilde{s}^{k+1})$ and let us prove that the losses are higher in the game G^s in the period k as well: $L^s_k(\hat{s}^k) < \tilde{L}^s_k(\tilde{s}^k)$.

$$\begin{aligned} L^s_k(\hat{s}^k) &= \hat{s}_k(V_k(\hat{\lambda}_k) - a) + (1-\hat{s}_k)\delta(1-\phi_k(\hat{\lambda}_k))L_{k+1}(\hat{s}^{k+1}, \hat{\lambda}_k) < \\ &< \hat{s}_k(V_k(\hat{\lambda}_k) - a) + (1-\hat{s}_k)\delta(1-\phi_k(0))L^s_{k+1}(\hat{s}^{k+1}) < \\ &< \hat{s}_k(V_k(\tilde{\lambda}_k) - a) + (1-\hat{s}_k)\delta(1-\phi_k(0))\tilde{L}^s_{k+1}(\tilde{s}^{k+1}) = \tilde{L}^c_k(\hat{s}_k, \tilde{s}^k, \tilde{\lambda}_k) \end{aligned}$$

Suppose that $\Delta = \tilde{L}^c_k(\hat{s}_k, \tilde{s}^k, \tilde{\lambda}_k) - \tilde{L}^s_k(\hat{s}_k, \tilde{s}^k) > 0$. Since $\partial(\tilde{L}^c_k - \tilde{L}^s_k)/\partial s_k > 0$, then $\tilde{s}_k < \hat{s}_k$. Then, it is easy to show that the losses become higher in the period k:

$$\begin{aligned} L^s_k(\hat{s}^k) &= (1-\hat{s}_k)(a - X^V_k(0)) + (1-\hat{s}_k)\delta(1-\phi_k(0))L^s_{k+1}(\hat{s}_{k+1}) < \\ &< (1-\hat{s}_k)(a - X^V_k(0)) + (1-\hat{s}_k)\delta(1-\phi_k(0))\tilde{L}^s_{k+1}(\tilde{s}_{k+1}) < \\ &< (1-\tilde{s}_k)(a - X^V_k(0)) + (1-\tilde{s}_k)\delta(1-\phi_k(0))\tilde{L}^s_{k+1}(\tilde{s}_{k+1}) = \tilde{L}^s_k(\tilde{s}^k) \end{aligned}$$

It remains to show that $\Delta > 0$:

$$\Delta = \hat{s}_k(V_k(\tilde{\lambda}_k) - a) + (1-\hat{s}_k)(a - X^V_k(0)) > \hat{s}_k(V_k(\hat{\lambda}_k) - a) + (1-\hat{s}_k)(a - X^V_k(0)) > 0,$$

which follows from the definition of $\hat{s}_k$.

□

Proposition 4.4. *For all $\eta > 0$ there exists some history h_η, after which:*

$$\max L(F_\varepsilon(F_{(0,z,\sigma)}), h_\eta) > \max L([F_{(0,z,\sigma)}], h_\eta) - \eta.$$

Proof. Suppose that F_0 does not have a mass point on 0 and is continuous on some interval $[0, \kappa]$. Then if F has a mass point on 0, then after observing 0, the searcher will believe that the true distribution is F with probability 1. If we induce F that has two-mass points on 0 and z we can achieve the highest possible losses.

Let F_0 has a mass point on 0. Then we can choose any ω close enough to 0 such that $\mathbb{P}(X_{F_0} = \omega) = 0$ and consider the history $h = \omega$, for which $\varepsilon(F_{(\omega,z,\sigma)}, \omega) = 0$. By taking the limit of our loss function when ω goes to zero we establish the fact that the nature can achieve its best outcome with any precision:

$$\lim_{\omega \to 0} \max L(F_\varepsilon(F_{(\omega,z,\sigma)}), \omega) = \max L([F_{(0,z,\sigma)}], h),$$

since $F_\varepsilon(F_{(\omega,z,\sigma)}) \to F_{(0,z,\sigma)}$ when $\omega \to 0$. □

Bibliography

Adam, K. (2001). 'Learning while searching for the best alternative', *Journal of Economic Theory*, vol. 101(1), pp. 252–280.

Allen, J., Clark, R. and Houde, J.F. (2012). 'Price negotiation in differentiated products markets: Evidence from the canadian mortgage market', Bank of Canada Working Paper.

Anderson, S.P. and Renault, R. (1999). 'Pricing, product diversity and search costs: a Betrand-Chemberlin-Diamond model', *RAND Journal of Economics*, vol. 30(4), pp. 719–735.

Armstrong, M. and Zhou, J. (2010). 'Conditioning prices on search behaviour', ELSE Working Paper 351.

Benabou, R. and Gertner, R. (1993). 'Search with learning from prices: does increased inflationary uncertainty lead to higher markups?', *The Review of Economic Studies*, vol. 60(1), pp. 69–93.

Bergemann, D. and Schlag, K. (2008). 'Pricing without priors', *Journal of European Economic Association*, vol. 6, pp. 560–569.

Bergemann, D. and Schlag, K. (2011a). 'Robust monopoly pricing', *Journal of Economic Theory*, vol. 146, pp. 2527–2543.

Bergemann, D. and Schlag, K. (2011b). 'Should I stay or should I go? Search without priors', Mimeo.

Chen, Y. and Riordan, M.H. (2007). 'Price and variety in the spokes model', *The Economic Journal*, vol. 117(522), pp. 897–921.

Chou, C.F. and Talmain, G. (1993). ‘Nonparametric search’, *Journal of Economic Dynamics and Control*, vol. 17, pp. 771–784.

Coulter, B. and Shapiro, J.D. (2015). ‘A mechanism for libor’, .

Dana Jr, J.D. (1994). ‘Learning in an equilibrium search model’, *International Economic Review*, pp. 745–771.

De los Santos, B., Hortaçsu, A. and Wildenbeest, M.R. (2012). ‘Testing models of consumer search using data on web browsing and purchasing behavior’, *The American Economic Review*, vol. 102(6), pp. 2955–2980.

De los Santos, B., Hortaçsu, A. and Wildenbeest, M.R. (2017). ‘Search with learning for differentiated products: Evidence from e-commerce’, *Journal of Business & Economic Statistics*, vol. 35(4), pp. 626–641.

Del Prete, S., Demma, C. and Rossi, P. (2017). ‘From few to many: Product differentiation in the italian mortgage market’, .

Diamond, P. (1971). ‘A model of price adjustment’, *Journal of Economic Theory*, vol. 3, pp. 156–168.

Duffie, D. and Dworczak, P. (2014). ‘Robust benchmark design’, National Bureau of Economic Research.

Duffie, D., Dworczak, P. and Zhu, H. (2017). ‘Benchmarks in search markets’, *The Journal of Finance*.

Gal-Or, E. (1985). ‘Information sharing in oligopoly’, *Econometrica: Journal of the Econometric Society*, pp. 329–343.

Gal-Or, E. (1986). ‘Information transmission—cournot and bertrand equilibria’, *The Review of Economic Studies*, vol. 53(1), pp. 85–92.

Gastwirth, J.L. (1976). ‘On probabilistic models of consumer search for information’, *The Quarterly Journal of Economics*, vol. 90(1), pp. 38–50.

Gilboa, I. and Schmeidler, D. (1989). ‘Maxmin expected utility with non-unique prior’, *Journal of Mathematical Economics*, vol. 18(2), pp. 141–153.

Hayashi, T. (2008). 'Regret aversion and opportunity dependence', *Journal of Economic Theory*, vol. 139, pp. 242–268.

Hayashi, T. (2009). 'Stopping with anticipated regret', *Journal of Mathematical Economics*, vol. 45, pp. 479–490.

Janssen, M. and Parakhonyak, A. (2013). 'Price matching guarantees and consumer search', *International Journal of Industrial Organization*, vol. 31, pp. 1–11.

Janssen, M. and Parakhonyak, A. (2014). 'Consumer search markets with costly revisits', *Economic Theory*, vol. 55, pp. 481–514.

Janssen, M., Pichler, P. and Weidenholzer, S. (2011). 'Oligopolistic markets with sequential search and production cost uncertainty', *The RAND Journal of Economics*, vol. 42(3), pp. 444–470.

Janssen, M. and Shelegia, S. (2015). 'Consumer search and double marginalization', *The American Economic Review*, vol. 105(6), pp. 1683–1710.

Janssen, M.C., Parakhonyak, A. and Parakhonyak, A. (2017). 'Non-reservation price equilibria and consumer search', *Journal of Economic Theory*, vol. 172, pp. 120–162.

Janssen, M.C.W., Moraga-González, J.L. and Wildenbeest, M. (2005). 'Truly costly sequential search and oligopolistic pricing', *International Journal of Industiral Organization*, vol. 23, pp. 451–466.

Kim, M., Kristiansen, E.G. and Vale, B. (2005). 'Endogenous product differentiation in credit markets: What do borrowers pay for?', *Journal of Banking & Finance*, vol. 29(3), pp. 681–699.

Koulayev, S. (2013). 'Search with dirichlet priors: estimation and implications for consumer demand', *Journal of Business & Economic Statistics*, vol. 31(2), pp. 226–239.

Li, L. (1985). 'Cournot oligopoly with information sharing', *The RAND Journal of Economics*, pp. 521–536.

Mailath, G.J., Postlewaite, A. and Samuelson, L. (2005). 'Contemporaneous perfect epsilon-equilibria', *Games and Economic Behavior*, vol. 53(1), pp. 126–140.

Manski, C. (2007). 'Minimax-regret treatment choice with missing outcome data', *Journal of Econometrics*, vol. 139, pp. 105–115.

Manski, C.F. (2004). 'Statistical treatment rules for heterogeneous populations', *Econometrica*, vol. 72(4), pp. 1221–1246.

Matutes, C. and Vives, X. (1996). 'Competition for deposits, fragility, and insurance', *Journal of Financial intermediation*, vol. 5(2), pp. 184–216.

Milnor, J.W. (1951). 'Games against nature', .

Moraga-González, J.L. and Petrikaitė, V. (2013). 'Search costs, demand-side economies, and the incentives to merge under bertrand competition', *The RAND Journal of Economics*, vol. 44(3), pp. 391–424.

Morgan, P. and Manning, R. (1985). 'Optimal search', *Econometrica: Journal of the Econometric Society*, pp. 923–944.

Nashimura, K. and Ozaki, H. (2004). 'Search and knightian uncertainty', *Journal of Economic Theory*, vol. 119, pp. 299–333.

Okuno-Fujiwara, M., Postlewaite, A. and Suzumura, K. (1990). 'Strategic information revelation', *The Review of Economic Studies*, vol. 57(1), pp. 25–47.

Parakhonyak, A. (2014). 'Oligopolistic competition and search without priors', *The Economic Journal*, vol. 124, pp. 594–606.

Parakhonyak, A. and Sobolev, A. (2015). 'Non-reservation price equilibrium and search without priors', *The Economic Journal*, vol. 125(584), pp. 887–909.

Petrikaitė, V. (2016). 'Collusion with costly consumer search', *International Journal of Industrial Organization*, vol. 44, pp. 1–10.

Radner, R. (1982). 'Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives', *Noncooperative Approaches to the Theory of Perfect Competition*, pp. 17–35.

Raith, M. (1996). 'A general model of information sharing in oligopoly', *Journal of economic theory*, vol. 71(1), pp. 260–288.

Reinganum, J.F. (1979). 'A simple model of equilibrium price dispersion', *Journal of Political Economy*, vol. 87(4), pp. 851–858.

Rhodes, A. (2014). 'Multiproduct retailing', *Review of Economic Studies*, vol. forthcoming.

Riedel, F. (2009). 'Optimal stopping with multiple priors', *Econometrica*, vol. 77, pp. 857–908.

Rosenfield, D.B. and Shapiro, R.D. (1981). 'Optimal adaptive price search', *Journal of Economic Theory*, vol. 25(1), pp. 1–20.

Rothschild, M. (1978). 'Searching for the lowest price when the distribution of prices is unknown', *Journal of Political Economy*, pp. 425–454.

Salop, S. and Stiglitz, J. (1977). 'Bargains and ripoffs: A model of monopolistically competitive price dispersion', *The Review of Economic Studies*, pp. 493–510.

Schlag, K. and Zapechelnyuk, A. (2018). 'Robust sequential search', *Working Paper*.

Shapiro, C. (1986). 'Exchange of cost information in oligopoly', *The review of economic studies*, vol. 53(3), pp. 433–446.

Stahl, D. (1996). 'Oligopolistic pricing with heterogeneous consumer search', *International Journal of Industiral Organization*, vol. 14, pp. 243–268.

Stahl, D.O. (1989). 'Oligopolistic pricing with sequential consumer search', *The American Economic Review*, pp. 700–712.

Stigler, G.J. (1961). 'The economics of information', *Journal of political economy*, vol. 69(3), pp. 213–225.

Stiglitz, J. (1987). 'Competition and the nubmer of firms in a market: Are doupolies more competitive than atomistic markets?', *Journal of Political Economy*, vol. 95, pp. 1041–1061.

Stoye, J. (2011). 'Axioms for minimax regret choice correspondences', *Journal of Economic Theory*, vol. 146, pp. 2226–2251.

Stoye, J. (2012). 'Minimax regret treatment choice with covariates or with limited validity of experiments', *Journal of Econometrics*, vol. 166, pp. 138–156.

Tappata, M. (2009). 'Rockets and feathers: Understanding asymmetric pricing', *The RAND Journal of Economics*, vol. 40, pp. 673–687.

Varian, H.R. (1980). 'A model of sales', *The American Economic Review*, vol. 70(4), pp. 651–659.

Vives, X. (1990). 'Trade association disclosure rules, incentives to share information, and welfare', *the RAND Journal of Economics*, pp. 409–430.

Vives, X. (2000). *Oligopoly pricing: old ideas and new tools*, MIT press.

Vives, X. (2010). *Information and learning in markets: the impact of market microstructure*, Princeton University Press.

Wald, A. (1950). 'Statistical decision functions.', .

Wolinsky, A. (1986). 'True monopolistic competition as a result of imperfect information', *The Quarterly Journal of Economics*, vol. 101(3), pp. 493–511.

Zhou, J. (2014). 'Multiproduct search and the joint search effect', *American Economic Review*, vol. 104(9), pp. 2918–39.

Abstract

This dissertation consists of three papers which are aimed at extending consumer search theory by relaxing some of the informational assumptions and by inquiring into the incentives of firms to share information.

The first paper analyses firm's incentives to share information in search markets. Many benchmarks in over-the-counter markets, such as LIBOR, are formed based on the submission of banks' interbank interest rates. Benchmarks are important, as they help consumers to search and provide firms with a basis for price setting. This paper explores firms' incentives to contribute information about their costs for the purpose of benchmark formation. We find that in order to deter consumer search, firms find it optimal to share their private costs if the search friction is sufficiently high. Moreover, the reduction of search in these markets can dramatically decrease social welfare.

The second analyses a model of oligopolistic competition in which consumers search without priors. Consumers do not have prior beliefs about the distribution of prices charged by firms and thus try to use a robust search procedure: they minimise the loss relative to the searcher, who knows the price distribution, in the worst case scenario. We show that the optimal stopping rule is stochastic and that for any distribution of search costs there is a unique market equilibrium which is characterised by price dispersion. Although listed prices approach the monopoly price as the number of firms increases, the effective price paid by consumers does not depend on the number of firms.

The third paper considers a problem of sequential search, in which a searcher relies on her initial information to make inferences about the unknown distribution of alternatives. Bayesian search rules are good only for specific environments and are often difficult to characterize. In this paper, we propose a robust search rule

that adapts to the information acquired while searching and performs well with respect to Bayesian optimal strategies for any priors. In a setting with binary-valued payoffs, we show that initial information and learning restrict the set of possible environments which gives rise to the adaptive rule with improved performance.

Zusammenfassung

Diese Dissertation besteht aus drei Teilen, welche zur Suchtheorie beitragen, indem sie Annahmen über die Verfügbarkeit und Verteilung von Informationen lockern und Anreize von Firmen erforschen, private Informationen freiwillig zu Verfügung zu stellen.

Der erste Artikel befasst sich mit den Anreizen von Firmen in Suchmärkten, Information mit anderen Marktteilnehmern zu teilen. Viele Benchmarks in sogenannten "over-the-counter" Märkten, wie beispielsweise der LIBOR, basieren auf der freiwilligen Zuerfügungstellung von privaten Informationen. Solche Benchmarks sind von Bedeutung um Konsumenten die Suche zu erleichtern und Firmen als Grundlage zur Preissezuung zu dienen. Dieser Artikel untersucht die Anreize von Firmen private Information über Produktionskosten zum Bilden einer Benchmark zu Verfügung zu stellen. Weiters zeigen wir, dass Benchmarkformation zum Beeinflussen des Suchverhaltens optimal für Firmen ist, wenn Informationskosten ausreichend hoch sind. Aufgrund der verringerten Suche sinkt in solchen Märkten die Wohlfahrt drastisch.

Der zweite Artikel analysiert ein oligopolistisches Wettbewerbsmodell, in dem Konsumenten nach Preisen suchen ohne a-priori Annahmen über die Verteilung der Preise zu haben. Sie versuchen daher robustes Suchverhalten anzuwenden, wobei das Verlustrisiko gegenüber einem Suchenden, welcher die Preisverteilung kennt, reduziert wird. Aus unserer Analyse geht hervor, dass optimales Suchverhalten ein stochastisches Abbruchkriterium aufweist. Für eine gegebene Verteilung von Informationskosten existiert ein eindeutiges Marktgleichgewicht, welches durch Preisdispersion charakterisiert ist. Obwohl Preise gegen den Monopolpreis konvergieren wenn die Anzahl der Firmen steigt, bleibt der Preis, den Konsumenten zu zahlen erwarten, davon unberührt.

Der dritte Artikel befasst sich mit dem Problem der sequentiellen Suche, in dem ein Konsument - abhängig von seinen ursprünglichen Annahmen - Schlüsse über die unbekannte Verteilung von Alternativen zieht. Bayesianische Suchalgorithmen eignen sich nur unter bestimmten Umständen und sind oft schwierig zu charakterisieren. Wir schlagen einen robusten Suchalgorithmus vor, der neue, durch Suche erworbene Informationen berücksichtigt und gegenüber Bayesianischen Algorithen unter beliebige Verteilungsannahmen vergleichsweise gute Ergebnisse liefert. Wenn Nutzwerte binär Verteilt sind, zeigen wir, dass anfängliche Informationen und Lernen die Menge an möglichen Umgebungen beschränken und einen adaptiven Suchalgorithmus ermöglichen, der bessere Ergebnisse liefert.