Transparency and Search Markets

Daniel Stuart Savelle
Southbury, Connecticut

B.A. Economics, St. Mary's College of Maryland, 2013
M.A. Economics, University of Virginia, 2014

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia
May 2019

Abstract

In search markets where the choices of firms (such as prices) may or may not be observed by consumers prior to search, the set of classic discrete-choice models and the set of ordered-search models generate the same set of equilibrium demand structures. This set equality extends to the supply side of these models if consumers know all prices prior to search (full transparency). However, a firm's incentive to raise prices is increasing in the firm's relative opacity to consumers. In search markets with strategic complementarities, if any firm faces increased opacity, all equilibrium prices increase. Moreover, if an ordered-search market is incorrectly modeled with the equivalent classic discrete-choice model, empirical estimates of firm profit margins and theoretical predictions of market prices are at the lower bound corresponding to full transparency. For data sets that include information about aggregate (or individual) search and selection, I outline an identification strategy that can be used to partially identify the underlying supply-side parameters of the search market without making assumptions about the opacity of the market.

JEL Codes: D43, D83, L13.

Keywords: Ordered Search; Discrete Choice; Opacity;

Acknowledgements

I am grateful for the guidance of Simon Anderson, Maxim Engers, Zachary Bethune, Eric Young and other faculty at the University of Virginia. I am also grateful for the feedback I received from discussants and attendees at the Annual International Industrial Organization Conference (IIOC), the ZEW ICT Conference and the Annual Conference of the European Association for Research in Industrial Economics (EARIE).

This dissertation is dedicated to my wife, Mai Savelle.

Contents

Introduction

This dissertation investigates the role of opacity in markets where agents (such as consumers) search among a set of options (such as products) prior to making selections. For simplicity, the model supposes that firms make simultaneous pricing decisions prior to consumer search.[1] Opacity varies across individual consumer-product pairs where the pair is *opaque* if the consumer does not know the price of the product until after paying the search cost for that product and *transparent* if the consumer knows the price of the product prior to search. A market is *fully transparent* if all pairings are transparent, *fully opaque* if all pairings are opaque and *partially opaque* if some pairs are transparent and others are opaque.

To represent consumer choices, I use an ordered-search model with heterogeneous search costs where a consumer selects from a list of products, but must pay a search cost to learn the actual value of each product. The setup and solution of the ordered-search model parallels other recent work including Armstrong (2017), Choi et al. (2018) and Moraga-González et al. (2018). Search costs make the ordered-search model different from the classic discrete-choice model where consumers costlessly learn match values for each possible product and select the product with the highest match value.[2] I consider equilibria where firms use pure strategies and consumers correctly anticipate the prices of firms.

[1] Similar results apply to labor markets where wages are chosen by firms and agents correspond to potential workers.

[2] Anderson et al. (1992) provides an in-depth analysis of classic discrete choice.

Chapter 1 demonstrates that, if consumers correctly anticipate the choices of firms in ordered-search markets, then there is a demand-side equivalence between ordered search and classic discrete choice. Proving this demand-side equivalence, I introduce a subset of ordered-search models, Geometric Search Models (GSMs), which can be used to match the demand and payoffs functions of any classic discrete-choice model given correct anticipation. Since consumers correctly anticipate the choices of firms in equilibrium, the GSM generates an equivalent equilibrium demand structure regardless of the opacity of a market. In fully transparent markets, consumers observe the choices of firms so consumers always correctly anticipate firm choices. Thus, there is a market-level equivalence between classic discrete-choice models and fully transparent ordered-search models. As a result, any empirical or theoretical results from the classic discrete-choice literature hold in corresponding fully transparent search markets.

Chapter 2 outlines and solves the overall model to demonstrate the role opacity plays in market outcomes. If a firm deviates from the price that consumers anticipate, then consumers who have an opaque pairing with the deviating firm do not adjust their search behavior until after searching that firm. Thus, increasing the portion of opaque consumer-firm pairs decreases the sensitivity of demand to price deviations. As a result, full transparency corresponds to the lower bound on the incentive for firms to increase prices and full opacity corresponds to the upper bound on the incentive for the firms to increase prices.

For Geometric Search Models, I provide sufficient conditions for the existence and uniqueness of an equilibrium. In contrast with most ordered-search papers, these conditions guarantee the existence of a unique pure-strategy equilibrium in fully opaque markets without limiting the number of firms to two, without requiring firm symmetry and without restricting the equilibrium search order of consumers. Moreover, prices are increasing in the opacity of a market. The increase in prices occurs even if the opacity of only a single firm increases. Depending on the search-specific parameters in the GSM which preserve the equilibrium demand structure, opaque markets are able to generate arbitrarily high prices while preserving the prices in the corresponding fully transparent market. Thus, the difference between prices with full transparency and prices with full opacity can be arbitrarily large depending on the search parameters of the model. I also outline a parametrized version of the GSM, the Reverse Extreme Value Search Model, to demonstrate that equilibrium consumer payoffs are decreasing in the level of search costs. Alternatively, prices may be increasing or decreasing in the level of search costs depending on the opacity of the market.

Chapter 3 considers markets that are incorrectly modeled with classic discrete choice, but where the true underlying market is a search market with an equivalent equilibrium demand structure. For the demand side of the market, the GSM can be used to estimate search parameters with search data while preserving the original empirical methodology for estimating demand parameters. The aim is to provide a framework to implement search estimation without requiring changes in the demand estimation strategy. In particular, the Reverse Extreme Value Search Model (REVSM) is a tractable model for analyzing search and selection data, both for the partial identification of cost parameters and for questions directly related to the search process itself. To demonstrate the robustness of the REVSM, I estimate the position-specific mean match value and search cost parameters of options on a computer screen using experimental data from Reutskaja et al. (2011). The estimates are consistent with the experimental feature that objects are placed randomly on the screen, but objects in some locations are searched and selected more frequently (e.g. top left of the screen).

For supply-side analysis, the predictions from theoretical models that use classic discrete choice understate equilibrium profit margins for firms and predict lower prices, even though the two market specifications have the same firm costs and equilibrium demand functions. Similarly, empirical models that estimate demands, then invert first-order conditions to estimate the vector of unit costs, yield the highest possible estimate of unit costs, corresponding to the case where the underlying search market is fully transparent. Thus, estimates of market power in search markets are biased downward if a classic discrete-choice model is applied in partially or fully opaque markets. If the search parameters of consumers in a market are identified, then unit costs and markups are partially identified without imposing any requirements on the opacity of the market.

This dissertation aims to relate ordered search and classic discrete choice. Chapter 1 establishes an equivalence between these two types of models. Chapter 2 establishes differences between these two models in partially and fully opaque markets. Finally, Chapter 3 extends these results for demand-side identification and supply-side partial identification.

Chapter 1: Demand Equivalence

This chapter outlines the general ordered-search model and reviews key results about ordered search that have previously been established. Then it introduces and analyzes a new type of ordered-search model, the Geometric Search Model (GSM) and shows demand equivalence between the set of ordered-search models and the set of classic discrete-choice models. Finally, this chapter addresses the role of heterogeneous search costs in the GSM and the equivalence result.

I. Ordered Search

The selection choices of a consumer looking for a car to purchase, a worker looking for a job, or a family looking to buy a house, among others, are often represented by the classic discrete-choice model of decision making. The classic model posits an agent who receives a match value for each option and selects the option with the highest match value. However, in many important cases, agents may not know their values and may have to incur a cost to learn them. For example, consumers searching online for a product likely incur costs in terms of time and effort. In this context, the agent making a decision starts with information that can be updated through some costly process.

While there are many ways to model discrete choices with costs for acquiring information, I analyze a discrete-choice model with ordered search where agents are able to choose their own search order and can return to any previously searched option without incurring an additional search cost (costless recall).[3] In this setting, agents' preconceived notions about their options and the corresponding cost of searching each option drive the order of search, whereas the additional knowledge gained from search dictates when agents stop and what agents select.[4]

The ordered-search model in this dissertation builds on the search model in Weitzman (1979). In his model, optimal search is determined by the match value (V_{in}) and search threshold ($\bar{V}_{in}$) for each option i and agent n. Each search threshold $\bar{V}_{in}$ solves a corresponding myopic search problem where agent n has a currently selectable match value and is deciding whether to search option i. If the currently held match value for agent n is equal to $\bar{V}_{in}$, then the agent is indifferent about searching option i.

Weitzman (1979) proves that agents search in decreasing order of search thresholds and stop search if they have observed a match value above all remaining search thresholds. Building on these results, Armstrong (2017) presents a novel transformation of the choices arising from search by showing that agents always select the option i with the highest value of V^*_{in} where $V^*_{in} = min\{V_{in}, \bar{V}_{in}\}$. An

[3]An alternative model is the rational inattention model. Fosgerau et al. (2017) provides an equivalence result between models of rational inattention and classic discrete choice which generalizes an earlier result in Matejka and McKay (2015). I present a similar equivalence result for ordered search.

[4]There is a rich literature in marketing and economics that outlines and analyzes "simultaneous" search models where search occurs in two stages. Consumers in the first stage choose a set of options to search (consideration set). In second stage, the consumer observes all match values in the consideration set and is able to select options from that set (or the outside option). There is no search order in this type of search model and no concept of continuing to search. See Hauser and Wernerfelt (1990) for the theory behind this type of model. Honka and Chintagunta (2016) compare this style of search (simultaneous) to ordered search using auto insurance data.

option is selected only if it is the highest valued option that is searched, so either a low search threshold or a low match value prevents an option from being selected. While the results of these two papers create the overarching search structure on which I build, I adapt their generalized analysis to a flexible, but more structured search model.

Armstrong (2017) demonstrates that for any ordered-search model, there exists an equivalent classic discrete-choice model that generates the same demands and expected payoffs. This relationship means that we can apply solution methods from the classic discrete-choice literature to ordered search using the equivalent classic discrete-choice model. I constructively prove the other direction: For any classic discrete-choice model, there exists an equivalent ordered-search model that generates the same demands and expected payoffs. Combining these two directions, I prove that ordered search is equivalent to classic discrete choice in terms of the final choices of agents.[5] This equivalence hinges on the assumption that consumers correctly anticipate the choices of firms in search markets.

An implication of this direction is that all theoretical or empirical results that rely on a classic discrete-choice model hold for a family of corresponding search models where the choices of firms are always known to consumers prior to search. For example, the coefficients that result from any structural estimation using the classic version of discrete choice are exactly the same coefficients that would result

[5]Similar to proving concepts like set equality, both directions are required to show that search and classic discrete choice are parallel families of models in terms of the possible final choices of agents.

from the choice-equivalent search model. This includes papers on the automobile industry like Berry et al. (1995) and many other papers that analyze data with discrete options. While these papers do not explicitly model search, my equivalence result demonstrates that their work is consistent with fully transparent ordered search. Thus, in fully transparent search markets, there is no added bias in demand and welfare estimates when using a classic discrete-choice model in empirical analysis if agents are searching for cars, housing, or some other product using a process consistent with ordered search.[6]

Apart from its empirical applications, the equivalence property also sheds light on theoretical models. For example, Quint (2014) analyzes products where the individual components of each product are supplied by different firms. It follows from the equivalence that all of the results of his paper extend to corresponding models with consumer search including his results related to the existence and uniqueness of a pricing equilibrium. In this dissertation, I apply one of the results in Quint (2014) to analyze search markets with the GSM that have a unique pricing equilibrium and strategic complementarities. Explicitly modeling search shows how factors affecting search can affect equilibrium outcomes in these markets.

[6]This consistency is particularly relevant because I know of only one paper that estimates a BLP-style model with ordered search. Moraga-González et al. (2018) estimates a search model that is similar to the theoretical model in this dissertation. Their model assumes that match values are Type I Extreme Value. They then estimate and compare a model without search costs and a model that includes heterogeneous search costs. Similar to the work later in this dissertation, Moraga-González et al. (2018) also conceptualize the search problem by working backwards from the equivalent frictionless model to the search model. They use this approach to define a tractable search specification for estimation. As an alternate option to the method in their paper, I provide an identification strategy for BLP-style analysis that uses the Geometric Search Model (GSM) by taking advantage of the GSM's tractability.

As an example, a parametrized version of the model, the Reverse Extreme Value Search Model (REVSM), determines the general equilibrium effects of increases or decreases in the distribution of search costs on prices.

I know of one direct limitation of the demand-side equivalence that should be noted. Search provides a rich framework because it structures the search behaviors of consumers for models where consumer search choices affect the payoffs of firms regardless of the final selections made by consumers. For example, a firm that advertises online often pays a per-click fee to search engines for each consumer that clicks on the associated link to the firm's website. This cost depends on the number of consumers that click the ad, but not the number or the value of purchases made by those consumers. For these types of markets, a model of search is required because a firm's payoff is directly dependent on search probabilities.

Other contexts also require search for a more detailed analysis. For example, consumer search papers like Moraga-González and Petrikaitė (2013) analyze mergers in search markets. If a merger of two firms only implies coordination in prices, then the results from the classic discrete-choice literature can be applied using the equivalence in fully transparent markets. However, if mergers also include cross listing products or locating together in the search process, then an explicit search model is necessary to analyze mergers. As shown in Moraga-González and Petrikaitė (2013), interacting search with mergers can lead to cases where mergers improve welfare.

The next section formally outlines the ordered-search model used throughout the rest of this dissertation.

II. The Search Model

In the ordered-search model, there are N agents each choosing among I options where $I, N < \infty$. For agent n, the match value with option i is

$$V_{in} = x_i + \alpha_{in} + \epsilon_{in}$$

where the match value (V_{in}) is the value the agent receives conditional on selecting option i. For now, each option i has an exogenously determined x_i that represents a shift in the option's mean match value with agents. In this model, α_{in} and ϵ_{in} represent the heterogeneous aspects of the match value between agent n and option i. Ex-ante, each α_{in} is distributed iid F_i^{α} and each ϵ_{in} is distributed iid F_i^{ϵ} for each agent n and a specific option i.[7]

Before search starts, each agent n learns an outside option value V_{0n} which is independently distributed with CDF F_0 and learns α_{in} for every option i. For now, I assume that each agent n also knows x_i for every option i prior to search. This assumption is relaxed in the Chapter 2 and 3. After observing these values, agent

[7]For simplicity, I assume that all CDFs of heterogeneity in this dissertation are differentiable and have a support over a convex set. This assumption is not necessary for most of the results, but it simplifies the notation and arguments. Each option can have different CDFs for agent-option level heterogeneity, so independence holds across any agent-option pairings, but distributions need not be identically distributed across different option-agent combinations. All distributions are assumed to have a finite mean.

n has the choice whether to pay a search cost of κ_{in} to observe ϵ_{in} for any chosen option i, where each κ_{in} is also known prior to initiating search. Ex-ante, each κ_{in} is distributed iid F_i^κ for each agent n and a specific option i. Each F_i^κ has a support bounded below by 0.

Agents want to maximize their expected payoff where each agent's payoff from search is the match value of their final selection minus the sum of all search costs that are incurred throughout the decision-making process. At each stage of the choice structure, agent n can choose to search a previously unsearched option or to terminate search. When agent n terminates search, the agent must select a searched option or the outside option. There is no additional search cost if an agent chooses an option that was searched earlier in the process (costless recall). If agent n searches until all options have been observed, then the agent must terminate search and choose an option. The search process plus the final selection choice an agent makes is called an agent's *search path*.

I consider the primitives of this choice model to be $(\mathbf{x}, F_0, [F_i^\alpha, F_i^\epsilon, F_i^\kappa]_1^I)$. However, I do not include a specific $\mathbf{x}$ as a primitive when choice models are used to generate demands, payoffs and search volumes as a function of $\mathbf{x}$.

III. Optimal Search

This section describes agent behavior in this model given correct anticipation of the vector $\mathbf{x}$. I introduce opacity and the supply side of the model after demonstrating demand-side equivalence. Before delving into the full analysis of search, I first solve a simplified version of search where agent n is allowed to search only a single option. Agent n starts with a match value $\hat{V}_n$ and decides whether to search option i where the match value with option i is as previously outlined. In this simplified structure, if agent n searches option i, then search terminates and the agent receives the max of $\hat{V}_n$ and V_{in}. Agent n must pay κ_{in} to search option i. If agent n does not search option i, then agent n receives $\hat{V}_n$. In this scenario, agent n strictly prefers to search option i if and only if the expected improvement in their highest match value is greater than the search cost that must be paid, or equivalently,

$$E[max\{0, V_{in} - \hat{V}_n\}|\alpha_{in}] > \kappa_{in}$$

This inequality can be used to define a threshold $\bar{V}_{in}$ where searching option i is preferable when $\hat{V}_n < \bar{V}_{in}$. To determine the corresponding threshold, it is helpful to define a function $m_i(\cdot)$ and its inverse $m_i^{-1}(\cdot)$. Let $\bar{\epsilon}_i$ be the supremum of the support of F_i^{ϵ} so that $z < \bar{\epsilon}_i \Leftrightarrow F_i^{\epsilon}(z) < 1$. If the support is not bounded above, $\bar{\epsilon}_i = \infty$. For a given option i, let $m_i^{-1} : (-\infty, \bar{\epsilon}_i) \to (0, \infty)$ where for any $r \in (-\infty, \bar{\epsilon}_i)$,

$$m_i^{-1}(r) = E[max\{\epsilon_{in} - r, 0\}].$$

Lemma 1. *For each option i, m_i is a strictly decreasing bijection from $(0, \infty)$ to $(-\infty, \bar{\epsilon}_i)$ and m_i^{-1} is a strictly decreasing bijection from $(-\infty, \bar{\epsilon}_i)$ to $(0, \infty)$. Additionally,*

$$E[max\{\hat{V}_n, V_{in}\}|\alpha_{in}] - \hat{V}_n > \kappa_{in}$$

$$\Leftrightarrow$$

$$\bar{V}_{in} = x_i + \alpha_{in} + m_i(\kappa_{in}) > \hat{V}_n.$$

This approach using search thresholds appears in various forms throughout the search literature.[8] I refer to $\bar{V}_{in}$ as a search threshold because agent n searches option i if their current observed best match is less than $\bar{V}_{in}$. Intuitively, higher search thresholds correspond to higher expected improvements (net of the search cost) for agent n searching option i. Note that search costs can be mapped to an additive increment in search thresholds in the form of $m_i(\kappa_{in})$. The same search threshold which holds in this restricted version of search plays a direct role in the general search process where each agent can search over the entire set of options, which is considered next.

In the more complex search model with all I options, a rational agent is best off searching through options in decreasing order of $\bar{V}_{in}$ with search being terminated if and only if the agent has found a match value above the highest $\bar{V}_{in}$ over all options that have yet to be searched. As a result of costless recall, each agent selects the option with the highest observed value when the agent ends the search process. This solution to the search model follows from Weitzman (1979). See Armstrong (2017) for a recent explanation with terminology that more closely parallels the work here. The search thresholds in the general search model are the same as the thresholds established in Lemma 1 because an agent who is exactly indifferent about searching the next option will never search the option after it.

[8] All proofs are in the appendix.

This relationship between generalized search and the more restricted version of search implies that the thresholds in Lemma 1 extend to the generalized search model.

During optimal search, if option i has a search threshold that lies above either the match value or threshold for every other option, then option i must be searched before any other option is selected. Alternatively, if there exists an option j with both a match value and search threshold above the search threshold for option i, then option j is searched before option i and presents a match value that prevents option i from ever being searched. The following lemma is implied by these properties combined with the fact that costless recall ensures that the final selection choice of an agent is the highest observed match value.

Lemma 2. *Let* $V_{in}^* = min\{\bar{V}_{in}, V_{in}\}$ *for* $i = 1, ..., I$ *and* $V_{0n}^* = V_{0n}$*. Under optimal search, agent n selects option i if*

$$V_{in}^* > \max_{j \neq i} V_{jn}^*$$

and agent n does not select option i if

$$V_{in}^* < \max_{j \neq i} V_{jn}^*.$$

Additionally, agent n searches option i if

$$\bar{V}_{in} > \max_{j \neq i} V_{jn}^*$$

and agent n does not search option i if

$$\bar{V}_{in} < \underset{j \neq i}{max}\ V^*_{jn}.$$

Due to the continuous distribution of search costs, ties occur with 0 probability, so I ignore ties and do not specify a tie-breaking rule. For the rest of this dissertation, I assume that agents optimize, so Lemma 2 holds in all sections below. Lemma 2 has already been shown in Armstrong (2017) and Choi et al. (2018).

Let D_i be the realized demand for option i where D_i is the number of agents $n = 1, ..., N$ for whom $V^*_{in} > \underset{j \neq i}{max}\ V^*_{jn}$. Taking the expectation over all possible realizations of agent heterogeneity,

$$E[D_i] = \sum_{n=1}^{N} Prob\left[V^*_{in} > \underset{j \neq i}{max}\ V^*_{jn}\right].$$

Let S_i be the realized search volume for each option i where S_i is the number of agents $n = 1, ..., N$ for whom $\bar{V}_{in} > \underset{j \neq i}{max}\ V^*_{jn}$. Taking the expectation over all possible realizations of agent heterogeneity,

$$E[S_i] = \sum_{n=1}^{N} Prob\left[\bar{V}_{in} > \underset{j \neq i}{max}\ V^*_{jn}\right]$$

Let P_n be the realized payoff of agent n in this model. Recall that an agent's realized payoff is the match value of the final choice minus all search costs incurred. For each option i and agent n, let $\mu_{in} = m_i(\kappa_{in})$ be the threshold heterogeneity resulting from heterogeneity in search costs where F_i^{μ} is the corresponding CDF. To simplify the equations in the following lemma, let $\delta_{in} = \alpha_{in} + min\{\mu_{in}, \epsilon_{in}\}$ where F_i^{δ} denotes the resulting CDF and let $\gamma_{in} = \alpha_{in} + \mu_{in}$ where F_i^{γ} denotes the resulting CDF.

Lemma 3. *Expected payoffs, expected demands and expected search volumes are given by the following integrals:*

$$E[P] = \sum_{i=0}^{I} \int_{-\infty}^{\infty} v \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\delta}(v - x_i) dv$$

$$E[D_i] = N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\delta}(v - x_i) dv$$

$$E[S_i] = N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\gamma}(v - x_i) dv$$

The formulas for the expected search volumes and demands rely on the property that $V_{in}^* = \min\{V_{in}, \bar{V}_{in}\} = x_i + \delta_{in}$ and $\bar{V}_{in} = x_i + \gamma_{in}$. The part of this lemma that pertains to agent payoffs is less obvious and is the result of two counteracting effects. Replacing the true match value V_{in} with V_{in}^* decreases the payoff of agents if $V_{in} > V_{in}^*$. Alternatively, removing the search costs from the problem increases the payoff for agents. Because the optimal search process sets the expected undervaluation of match values ($V_{in} - V_{in}^*$) equal to the expected search costs, the two effects exactly counterbalance each other. This lemma provides a convenient formula for the expected payoff each agent receives which allows for statements about payoff equivalence later in this chapter.[9]

Throughout all of the results in this section, the only role search costs play in an agent's behavior and payoffs is through the value of $\mu_{in} = m_i(\kappa_{in})$ and its additive effect on $\bar{V}_{in}$. The following corollary demonstrates that the search model can be specified in terms of search threshold heterogeneity (μ).

Corollary 1. *Given a specific F_i^ϵ, each F_i^μ with a support that is bounded above by $\bar{\epsilon}_i$ corresponds to a unique F_i^κ and each F_i^κ corresponds to a unique F_i^μ with a support that is bounded above by $\bar{\epsilon}_i$ where all CDFs are continuous and differentiable over a convex support. As a result, search models can be specified by $(\mathbf{x}, F_0, [F_i^\alpha, F_i^\epsilon, F_i^\mu]_1^I)$ instead of $(\mathbf{x}, F_0, [F_i^\alpha, F_i^\epsilon, F_i^\kappa]_1^I)$ without imposing extra restrictions or changing the set of possible search models.*

The proof of this corollary is an extension of the bijective properties of m_i as outlined in Lemma 1. While each F_i^μ is not exogenous in the initial search structure, all future sections treat F_i^μ as a primitive by requiring that agents optimize. In order to relate the search model to a model with costless information, the next section formally defines the Classic Discrete-Choice Model.

IV. Classic Discrete-Choice Model

In the Classic Discrete-Choice Model (CDM), there are N agents each choosing among I options where $I, N < \infty$. For agent n, the match value with option i is

[9] There is a similar result for expected payoffs with constant search costs in the appendix of Choi et al. (2018) and the appendix of Armstrong (2017).

$$V_{in} = x_i + \beta_{in}.$$

An agent's match value with a given option i is the value the agent receives conditional on selecting option i. As in the search model, each option i has an exogenously determined x_i that represents a shift in the option's mean match value with agents. In this structure, β_{in} represents the heterogeneous aspects of the match value between agent n and option i. Ex-ante, each β_{in} is distributed iid F_i^{β} for each agent n given a fixed option i.

Each agent n costlessly observes all match values and chooses the highest valued match. This structure is consistent with the types of discrete-choice models outlined in Anderson et al. (1992) and is a popular way to model agent choices. A key result of this model is that the expected demand for option i can be expressed as

$$E[D_i^*] = N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\beta}(v - x_j) f_i^{\beta}(v - x_i) dv.$$

and the expected payoff for each agent can be expressed as

$$E[P^*] = \sum_{i=0}^{I} \int_{-\infty}^{\infty} v \prod_{j \neq i} F_j^{\beta}(v - x_j) f_i^{\beta}(v - x_i) dv.$$

Note the assumption of ex-ante independence for all β_{in} for a given agent, which simplifies the work in the following sections.[10]

As seen in Lemma 3, these formulas parallel the results from the search model

[10] Similar to other settings, this assumption can be relaxed to type-specific independence (e.g. BLP) or within-category independence (e.g. nested logit).

with β replacing δ. Recall that $\delta_{in} = \alpha_{in} + min\{\mu_{in}, \epsilon_{in}\}$. While each F_i^{β} is a primitive of the CDM, each F_i^{δ} results from the combination of three underlying distributions of heterogeneity corresponding to the two aspects of match value heterogeneity for option i, α_{in} and ϵ_{in}, as well the threshold heterogeneity for option i, μ_{in}, which is generated by the search cost heterogeneity. In the next section, I introduce a set of ordered-search models which are able to generate F_i^{δ} to equal any specified F_i^{β} for each option i.

V. The Geometric Search Model (GSM)

Now that I have outlined ordered search and classic discrete choice, I introduce a specific version of the search model called the Geometric Search Model (GSM). For each option i, the underlying heterogeneity of ϵ and μ is given by a CDF F_i^* and two parameters $a_i, b_i \in (0, \infty)$ where:

$$F_i^{\epsilon} = 1 - (1 - F_i^*)^{a_i} \qquad \text{and} \qquad F_i^{\mu} = 1 - (1 - F_i^*)^{b_i}.$$

All other features of the search model remain the same. Basically, F_i^* is the CDF used to specify the distributions of ϵ_{in} and μ_{in} where a_i and b_i are scaling parameters. As with all other CDFs, I impose that each F_i^* is differentiable over a convex support. Since the functions $1 - (1 - z)^{a_i}$ and $1 - (1 - z)^{b_i}$ are strictly increasing bijections from $[0, 1]$ to $[0, 1]$, F_i^{ϵ} and F_i^{μ} are both well-defined, continuous

CDFs that are differentiable over the same convex support.[11] With this specification, F_i^ϵ, F_i^μ and F_i^* all share the same support because if one of the functions is 0 or 1, the others functions must also be 0 or 1 respectively. Thus, the GSM is fully specified by $(\mathbf{x}, F_0, [F_i^\alpha, F_i^*, a_i, b_i]_1^I)$.

For simplicity, let $\bar{F}$ denote $\bar{F} = 1 - F$ for any specified CDF. Using this notation, $\bar{F}_i^\epsilon = (\bar{F}_i^*)^{a_i}$ and $\bar{F}_i^\mu = (\bar{F}_i^*)^{b_i}$. Additionally,

$$\bar{F}_i^{min} = (\bar{F}_i^\mu)(\bar{F}_i^\epsilon) = (\bar{F}_i^*)^{a_i+b_i} \quad \forall i = 1, ..., I.$$

The parameters a_i and b_i are of particular economic importance due to the direct influence these parameters have on the distributions of match values and search thresholds. Increasing a_i results in a strict increase in F_i^ϵ at any point where F_i^ϵ is not 0 or 1. Similarly, increasing b_i results in a strict increase in F_i^μ at any point where F_i^μ is not 0 or 1. Basically, search thresholds are decreasing in b_i, match values are decreasing in a_i and the minimum of these values is decreasing in both a_i and b_i. Additionally, as b_i approaches 0, the underlying search cost distribution limits to a mass point at 0 and the distribution of μ_{in} limits to a mass point at $\bar{\epsilon}_i$, which is the supremum of the support of F_i^ϵ, F_i^μ and F_i^*.

[11]As discussed earlier, I can define F_i^μ as a primitive given optimal search behavior. Alternatively, given F_i^ϵ, m_i is defined for each option i with the properties outlined in Lemma 1. Thus, $F_i^\kappa(\kappa) = (1 - F_i^*(m_i(\kappa)))^{b_i}$ is a CDF which results in the same F_i^μ imposed in the GSM. Changes in b_i relate to changes in the underlying search cost distribution. Alternatively, a_i changes the distribution of ϵ_{in}, which in turn changes the function m_i. Thus, the distribution of search costs also changes with a_i in order to preserve the resulting distribution of thresholds. For simplicity, I refer to changes in a_i as changes in match values (holding thresholds fixed) and changes in b_i as changes in thresholds and search costs.

The rest of this section sets $b_i = 1 - a_i$. This added assumption changes the way the model is specified, but does not limit further the family of models that can be specified.[12] A direct implication of this added structure on each F_i^ϵ and F_i^μ is that

$$F_i^{min} = 1 - (1 - F_i^\mu)(1 - F_i^\epsilon) = F_i^* \qquad \forall i = 1, ..., I.$$

This equality holds for any option i regardless of the value of a_i. Since F_i^{min} does not depend on a_i, the resulting F_i^δ does not depend on a_i. Now let $a_i^{'}$ and $a_i^{''}$ be two different possible values of a_i where $a_i^{'} < a_i^{''}$. From the CDFs above, increasing a_i increases F_i^ϵ and decreases F_i^μ over the interior of each function's support. Thus, F_i^ϵ when $a_i = a_i^{'}$ first order stochastically dominates F_i^ϵ when $a_i = a_i^{''}$. The opposite is true for F_i^μ. Effectively, a_i determines the relative importance of search threshold heterogeneity (μ_{in}) compared to the post-search match value heterogeneity (ϵ_{in}) while still preserving the resulting distribution of δ_{in}. The following Lemma builds on these underlying properties.

Lemma 4. *Consider a GSM* $(\mathbf{x}, F_0, [F_i^\alpha, F_i^*, a_i]_1^I)$ *where* $E[D_i] > 0 \;\; \forall i$. *For each option* i*, let* $\bar{D}_i = E[D_i | \epsilon_{in} = \bar{\epsilon}_i]$. *The following statements hold true:*

1) The ex-ante probability that an agent follows each possible search path is strictly greater than 0.

2) For any option i *and* j, $\frac{\partial E[D_i]}{\partial a_j} = 0$. *For any option* j *where* $j \neq i$, $\frac{\partial E[S_i]}{\partial a_j} = 0$.

[12] This new specification does not make a change to the family of GSMs because $\bar{F}_i^\epsilon = (\bar{F}_i^*)^{a_i} = ((\bar{F}_i^*)^{a_i + b_i})^{\frac{a_i}{a_i + b_i}}$ and $\bar{F}_i^\mu = (\bar{F}_i^*)^{b_i} = ((\bar{F}_i^*)^{a_i + b_i})^{\frac{b_i}{a_i + b_i}}$ for any option i.

3) For any option i, $E[S_i]$ as a function of a_i is a strictly increasing bijection from $(0,1)$ to $(E[D_i], \bar{D}_i)$.

4) For any I-dimensional vector $\mathbf{s}$ where $E[\mathbf{D}] < \mathbf{s} < \bar{\mathbf{D}}$, there exists a unique vector of $\mathbf{a}$ values such that $E[\mathbf{S}] = \mathbf{s}$.

Note that for a given option i, $\bar{D}_i = N$ if $\bar{\epsilon}_i = \infty$. For the analysis below, there are two implications of this lemma that play particularly important roles. First, the vector of search parameters a_i in the GSM can be uniquely mapped to a corresponding vector of expected search volumes. This mapping provides the core argument for identification of a structural model in Chapter 3. Lemma 4 also shows that the expected demand for any option and the expected search volumes for other options do not depend on a specific option $i's$ value of a_i. This property follows from Lemma 3 combined with the property that F_i^{δ} does not vary with a_i.

Parts 1 and 4 of this lemma demonstrate that the search structure defined here is relatively flexible compared to most search models that include endogenous firm choices (such as pricing games). While many theoretical assumptions can be imposed in search models, none of the assumptions here exclude the possibility of any specific agent search paths in equilibrium.[13] Additionally, a GSM where the support of each ϵ is not bounded above ($\bar{\epsilon}_i = \infty \;\; \forall i$) can generate any reasonable vector of expected search volumes within the limitations that an agent must search an option before selecting it ($E[\mathbf{D}] \leq E[\mathbf{S}]$) and that search volumes cannot exceed

[13] Recall that a search path is a full description of all choices an agent makes during their search process.

the number of agents ($E[\mathbf{S}] \leq N$).[14] These properties of the GSM do not directly improve on search with exogenous firm choices, but do improve on the ordered search models used in the literature on pricing games with ordered consumer search. In many papers that structure ordered search to analyze market pricing, agents search in the same order and/or do not return to options search earlier. For example, Anderson and Renault (2018) relies on a fixed order where agents do not return to previously searched options.

VI. The Equivalence

Now that the GSM has been outlined, I discuss the equivalence of ordered search and classic discrete choice. For the rest of this dissertation, I consider choice models to be defined over all possible **x** vectors. Thus, in the work below, I specify choice models without including a specific **x** vector. Moreover, choice models generate expected demands and search volumes that are functions of **x**.

Recall that Armstrong (2017) demonstrated that any ordered-search model has a corresponding CDM that generates the same expected demands and payoffs. Also, recall that the work thus far has relied on the assumption that consumers correctly anticipate the x_i choices of firms. Given any specified CDM $(F_0, [F_i^\beta]_1^I)$, the GSM with $\alpha_{in} = 0$ for each option i and agent n and $F_i^* = F_i^\beta$ for each option i generates the same expected demands and payoffs for all **x**. Furthermore, expected

[14]While expected search volumes in the GSM are strictly great than expected demands and strictly less than N, limit case arguments can be made for $E[\mathbf{D}] = E[\mathbf{S}]$ and $E[\mathbf{S}] = N$.

search volumes can be included in this equivalence by selecting a_i's appropriately. As previous ordered-search papers have already stated, the CDM is equivalent to a search model without positive search costs. However, the following theorem establishes this property for any CDM paired with a vector of potential search volumes.[15]

Theorem 1.

1) For any search model $(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$*, there exists a CDM* $(F_0, [F_i^{\beta}]_1^I)$ *that generates the same expected demands and expected payoffs for all* $\mathbf{x}$*.*

2) For any CDM $(F_0, [F_i^{\beta}]_1^I)$*, there exists a search model* $(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$ *that generates the same expected demands and expected payoffs for all* $\mathbf{x}$*.*[16]

3) For any CDM $(F_0, [F_i^{\beta}]_1^I)$ *and I dimensional function* $\mathbf{s}(\mathbf{x})$ *where* F_i^{β} *has support that is unbounded above and* $E[\mathbf{D}] < \mathbf{s}(\mathbf{x}) < \bar{\mathbf{D}}$ *for all* $\mathbf{x}$*, there exists a GSM* $(F_0, [F_i^{\alpha}, F_i^{*}, a_i(\mathbf{x})]_1^I)$ *that generates the same expected demands, expected payoffs and has the property that* $E[\mathbf{S}] = \mathbf{s}(\mathbf{x})$ *for all* $\mathbf{x}$*.*

4) For any search model $(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$*, there exists a GSM* $(F_0, [F_i^{\alpha}, F_i^{*}, a_i]_1^I)$ *that generates the same expected demands, expected payoffs and expected search volumes for all* $\mathbf{x}$*.*

This theorem establishes strong results with larger implications for the literature on both ordered search and classic discrete choice. Consider a market where each

[15] For the GSM in this section, a_i is a function of x_i and $b_i = 1 - a_i$.

[16] There is a continuum of models that match any specified CDM because the proof does not require a specific value of $\mathbf{a}$ in the GSM.

firm's profit depends on the actions of consumers only through the expected demands. For the ordered-search market, suppose that consumers always know the choices of firms so that the equivalence result holds in all possible states of the world. Under these assumptions, ordered-search models and classic discrete-choice models are interchangeable as long as the models generate the same demand functions. Moreover, for any DCM, an equivalent ordered-search model always exists. Thus, every theoretical or empirical model defined with classic discrete choice is equivalent to a corresponding model of ordered search. In other words, all of the results in papers that use a classic discrete-choice model hold also for a corresponding fully transparent search model, including results pertaining to pricing games, structural estimation, mergers, advertising, taxation and more.

This theorem also demonstrates the flexibility of the GSM as a model of consumer search. The GSM can match classic discrete choice augmented with a plausible function for expected search volumes.[17] Thus, even though the GSM imposes added structure on the underlying distributions of the search model, GSM is not restrictive at the aggregate level.

VII. The Importance of Search-Cost Heterogeneity

Before concluding this chapter, I discuss the role of heterogeneous search costs for my results throughout this chapter and the rest of this dissertation. This section

[17]For this part of the theorem, I require GSMs with values of $\mathbf{a}$ that vary with $\mathbf{x}$. However, this addition is necessary to avoid further restrictions on expected search volumes as a function of $\mathbf{x}$. For the rest of this dissertation, I assume that $\mathbf{a}$ and $\mathbf{b}$ are parameter vectors that do not depend on $\mathbf{x}$.

addresses one of the main tractability issues with specifications of ordered-search models. If search costs are homogeneous (i.e. $\kappa_{in} = \kappa_i \ \ \forall i$), then $min\{m_i(\kappa_i), \epsilon_{in}\}$ is a random variable with a mass point at $m_i(\kappa_i)$. These mass points have previously been addressed with strong assumptions because standard methods of showing the existence (and uniqueness) of a pricing equilibrium in classic discrete-choice models require a continuous match-value distribution. For example, the mass-point problem implies that the CDF in the corresponding CDM is not log-concave, so the results from Quint (2014) would not hold.

Choi et al. (2018) and Anderson and Renault (2018) use support restrictions to address this issue. In Anderson and Renault (2018), this assumption is limiting because it implies that no agent ever returns to a previously searched option. As a result, options are selected only if search terminates immediately after that option is observed. Ideally, a search model should have agents who sometimes return to previous observations when behaving optimally because this behavior occurs in reality.

In Choi et al. (2018), the support restriction is used to guarantee that a firm can make a sale only in a region where the density of V^*_{in} is log-concave.[18] An alternative option proposed in Choi et al. (2018) is to assume that $\alpha_{in} + min\{m_i(\kappa_i), \epsilon_{in}\}$ has desirable properties given a sufficiently large variance in α_{in} relative to the variance of $min\{m_i(\kappa_i), \epsilon_{in}\}$. By imposing a high variance on α (the aspect of a match value

[18]See Proposition 2 in Choi et al. (2018) for more details. Choi et al. (2018) relies on heterogeneity in α to provide new results for markets with search.

known prior to search), one imposes on the model that α drives variation in the final selection choices of agents with some sufficiently small aspect of the variation in decision-making process coming from actual search. This assumption yields the existence and uniqueness of a pricing equilibrium at the cost of restricting the role of search in an agent's decision-making process. To address the same underlying mass-point problem, Zhou (2014) and Haan et al. (2015) impose a duopoly structure in search markets and Zhou (2011) assume a uniform distribution for match values.

This dissertation avoids this mass-point issue by allowing for heterogeneity of search costs. Since there is no reason to believe that all consumers have exactly the same search costs for all options, the mass point issue is an artificial construct resulting from the specific assumptions in other models and is likely not an issue in the real world. This dissertation is not the first analysis that uses heterogeneity in search costs, but this model is one of the few that pairs search cost heterogeneity with this type of search model.[19] Basically, search cost heterogeneity adds a complexity to the model that balances the heterogeneity in match values an agent discovers after searching a given option.

However, a reasonable concern with the search model in this dissertation is that search costs are likely correlated in the real world. For example, faster internet, more free time, the use of a car, and other characteristics of agents will likely cause correlations in search costs between different options (websites, stores, etc) for the

[19] Recall that Moraga-González et al. (2018) use search cost heterogeneity, but focus primarily on empirical work.

same agent. In many choice models including the GSM that assume independence in heterogeneity, the independence assumption can be replaced with conditional independence. In the GSM, a_i and b_i can be agent-type specific values that change depending on relevant characteristics of agents. This introduces correlations in search costs because b_i affects the underlying distributions of search costs for a given agent-type. Unconditional on the type of agent, search costs could then be correlated due to the different specifications of b_i based on the heterogeneity over agent types. Conveniently, this added level of complexity preserves the underlying conditional independence assumption in the model for any fixed type of agent. For a related discussion that focuses on structural models of classic discrete choice with random coefficients, see Berry (1994). Note that even with correlated search costs, the variation in search costs all that is necessary to avoid mass-point issues.

With heterogeneous search costs, this chapter links ordered search and classic discrete choice by equating the set of possible demand systems generated by these two types of models. However, this equivalence hinges on consumers correctly anticipating the choices of firms. The next chapter considers markets with opacity to establish differences between search and classic discrete choice. The GSM provides the underlying structure for the results in the next two chapters.

Chapter 2: Pricing Games with Opacity

This chapter introduces opacity in firm choices and provides partial equilibrium properties for firms. After this, it establishes conditions for the existence and uniqueness of pricing equilibria and outlines general equilibrium properties. Finally, a parameterized search model, the Reverse Extreme Value Search Model, is provided to track the effect of changes to the search cost distribution on equilibrium outcomes in the model.

I. Introduction to Transparency, Opacity and Beliefs

Search markets with opacity provide a context in which the supply side of the market is not equivalent to a similar market with classic discrete choice. Models of markets with search must take a stance on the interaction between firm choices and consumer beliefs prior to search. In many models, a consumer knows or observes the relevant choices of all firms prior to paying a cost to acquire further information. As stated earlier, I refer to this type of market as fully transparent. In the pricing literature, this concept is sometimes referred to as price commitment or price advertising. An alternative is to require that no consumers observe the choices of firms before paying relevant search costs. Instead, consumers anticipate the choices of firms with the added requirement that, in equilibrium, consumers correctly anticipate the choices of firms. This is the more prevalent pricing model in the ordered-search literature as in Zhou (2011) and Anderson and Renault (2018). As stated earlier, I refer to this type of market as fully opaque.

I model opacity and transparency similarly, but I allow transparency and opacity to vary across consumer-firm pairings where a pairing is transparent if the consumer observes the firm choice prior to search and opaque if not. Modeling transparency and opacity at this level includes markets where transparency or opacity is not absolute. Markets are *fully opaque* if the probability of transparent consumer-firm pairs is 0, *fully transparent* if the probability of opaque consumer-firm pairs is 0 and *partially opaque* if the probability of opaque and transparent consumer-firm pairings are both positive. A firm is *more opaque* (*less opaque*) if there is an increase (decrease) in the probability of opaque consumer-firm pairs with that specific firm. Similar terminology applies at the market-level where more opaque firms, all else equal, correspond to more opaque markets.

This dissertation assumes that firms make their choices prior to consumer search. Thus, transparency and opacity determine the timing of consumer observations of firm choices. This model is analogous to a similar model where firms can signal their choices prior to search, but may or may not be able to commit to their signaled choice. In this setting, commitment corresponds to transparency and an inability to commit corresponds to opacity.

The market-level equivalence between classic discrete choice and ordered search relies on full transparency which seems a prohibitively strong assumption in many markets. Markets may suffer from a lack of transparency and/or a lack of commitment. If firms are able to add hidden fees after search, adjust prices before checkout or adjust the endogenous quality of the product in a way that is revealed only after search, then the market is partially or fully opaque.[20] Similarly, if some firms cannot communicate with some consumers prior to search, then the market is partially or fully opaque. The legal system and platforms often attempt to improve the transparency in search markets. For example, platforms like Google shopping and laws like the credit card bill of rights in the U.S. prohibit the inclusion of hidden fees. Similarly, consumers sometimes have methods to figure out prices and fees prior to actually searching a firm, including price comparison apps like Honey. In future work, I plan to analyze markets where transparency and opacity are endogenous.

While the terms transparency and opacity occur throughout this dissertation, transparency and opacity can refer to different versions of information transmission depending on the model. In this dissertation, I assume that every firm plays a pure strategy in any equilibrium. I also assume that consumers either observe the true choice or anticipate a pure strategy by each firm. Under these assumptions, transparency transmits information about firm cheating, not about any underlying heterogeneous values.[21]

[20]Opacity is similar to the concept of informed and uniformed consumers discussed in papers like Armstrong (2015).

[21]If firms were playing mixed strategies in equilibrium, then transparency would provide relevant information about the endogenous choices of firms even if all firms follow their equilibrium strategy. Transparency could also refer to the transmission of additional information about the exogenous characteristics of products, jobs, etc. However, transparency in this dissertation relates only to firm

II. Agent-Option Opacity

I add two additional features to the underlying search model from Chapter 1. First, the model now includes heterogeneity in the transparency and opacity of agent-option pairs, and second, agents now anticipate, possibly incorrectly, x_i for each option i. Recall that the work in Chapter 1 assumed correct anticipation.

Let σ_{in} be a new type of agent-option heterogeneity where $\sigma_{in} \in \{0, 1\}$. If $\sigma_{in} = 0$, then agent n observes x_i prior to search. If $\sigma_{in} = 1$, then agent n observes the value of x_i only after searching option i. An agent-option pair is *transparent* if $\sigma_{in} = 0$ and *opaque* if $\sigma_{in} = 1$. Let $\sigma_i^* = \frac{\sum_{n=1}^{N} E[\sigma_{in}]}{N}$ be the ex-ante probability that $\sigma_{in} = 1$ averaged over all agents. There are no additional assumptions on the distribution of σ values except that this new type of heterogeneity is distributed independently of all other types of agent-option heterogeneity (α, κ and ϵ). This requirement does not rule out correlations in σ values. By construction, $\sigma_i^* \in [0, 1]$ where $\sigma_i^* = 0$ corresponds a fully transparent option i and $\sigma_i^* = 1$ corresponds to a fully opaque option i.

Let $\hat{x}_i$ denote the value agent n anticipates for x_i if $\sigma_{in} = 1$. To simplify notation, x_{in}^e is agent $n's$ anticipated value where $x_{in}^e = x_i$ if $\sigma_{in} = 0$ and $x_{in}^e = \hat{x}_i$ if $\sigma_{in} = 1$.[22] I impose passive beliefs where an agent does not update x_{in}^e after observing x_j for

choices, not exogenous characteristics.

[22]Since I am introducing opacity and transparency to model agent choices for equilibria where firms play pure strategies, I do not address ordered search where opaque agents have different priors or prior beliefs that x values are drawn from a non-degenerate distribution. While ordered search can include these possibilities, for example if firms play mixed strategies in equilibrium, these topics are outside the scope of this dissertation.

any $j \neq i$. Passive beliefs imply that a consumer who observes one firm deviating from equilibrium continues with the belief that no other firms have deviated.[23]

At any point in the search process, agents still aim to maximize their expected payoff given their currently held beliefs about the options. As before, agent n selects option i if $V^*_{in} > \underset{j \neq i}{max}\ V^*_{jn}$ and searches option i if $\bar{V}_{in} > \underset{j \neq i}{max}\ V^*_{jn}$. However, since $\bar{V}_{in}$ depends on expected improvement of searching option i given beliefs prior to searching option i, $\bar{V}_{in} = x^e_{in} + \alpha_{in} + \mu_{in}$. As a result, $V^*_{in} = min\{V_{in}, \bar{V}_{in}\}$ depends on the opacity of an agent-option pair. The following theorem outlines properties of opacity for consumer choices.

Lemma 5. *Suppose that F_0 has a support that is not bounded above. Also suppose that $\forall i$, F^ϵ_i and F^μ_i have supports that are not bounded above.*[24] *Given two different options $i \neq j$, the following properties hold.*

1) If $\sigma^*_i, \sigma^*_j \in (0,1)$, $E[D_i]$ and $E[S_i]$ are strictly increasing in both x_i and $\hat{x}_i$ and strictly decreasing in both x_j and $\hat{x}_j$.

Suppose $\hat{x}_k = x_k \quad \forall k \neq i$.

2) If $\hat{x}_i < x_i$, $E[D_i]$ and $E[S_i]$ are strictly decreasing in σ^*_i and $E[D_j]$ and $E[S_j]$ are strictly increasing in σ^*_i.

3) If $\hat{x}_i > x_i$, $E[D_i]$ and $E[S_i]$ are strictly increasing in σ^*_i and $E[D_j]$ and $E[S_j]$ are strictly decreasing in σ^*_i.

[23]Assuming passive beliefs is standard in search models. As discussed in Janssen and Shelegia (2018), this assumption about beliefs can be restrictive.

[24]Without this assumption or a similar support assumption, the conclusions of the lemma would switch from strictly increasing/decreasing to weakly increasing/decreasing.

4) $\frac{\partial^2 E[D_i]}{\partial \sigma_i^* \partial x_i} < 0$ *and* $\frac{\partial^2 E[S_i]}{\partial \sigma_i^* \partial x_i} < 0$

5) If $\hat{x}_i = x_i$*,* σ_i^* *does not affect the choices of agents.*

6) Conditional on σ_i^**, no feature of the distribution of* σ *values across agent-option pairs affects selection or search probabilities.*

Property 1 states that increases in x_i or the anticipated value of x_i for opaque pairs increases expected demands and search volumes for option i and decreases expected demands and search volumes of all other options. This property follows from the fact that $\bar{V}_{in}$ and V_{in}^* are increasing in x_i and $\hat{x}_i$. Property 2 and property 3 follow from a similar argument about V_{in}^*, which is increasing in the relative opacity of option i if the anticipated x_i value is above x_i and decreasing if the anticipated x_i value is below x_i.

Property 4 implies that the effect on $E[D_i]$ and $E[S_i]$ from increasing x_i is decreasing in the opacity of option i with an upper bound corresponding to full transparency ($\sigma^* = 0$) and a lower bound corresponding to full opacity ($\sigma^* = 1$). This property arises because V_{in}^* is increasing in both x_i and x_{in}^a, which implies that changes in x_i have a larger effect for transparent pairs ($x_{in}^e = x_i$) than for opaque pairs ($x_{in}^e = \hat{x}_i$). In the next section, property 4 provides a bound on firm incentives in markets. Property 5 states that, if agents correctly anticipate x_i, then the relative opacity of product i is irrelevant because $x_i = \hat{x}_i = x_{in}^e$. Property 6 states that, if agents correctly anticipate the x values for all other options, then the mean opacity of option i is the only feature of the distribution of opacities that affects the actions of agents. This property is important in a model with single-product firms because a deviating firm assumes all other firms follow the anticipated equilibrium strategy. Thus, σ_i^* is the only feature of the distribution of opacities that directly affects the equilibrium deviation payoff of firm i.

Let $E[D_i^{trans}] = E[D_i|\sigma_{in} = 0\ \forall n]$ denote the expected demand for product i conditional on transparent pairs for product i and let $E[S_i^{trans}] = E[S_i|\sigma_{in} = 0\ \forall n]$ denote the expected search volume for product i conditional on transparent pairs for product i. Now consider the GSM. I do not impose that a_i and b_i sum to 1 in this chapter. To simplify notation in the lemma, let $a_i^* = \frac{a_i}{a_i+b_i}$. Changes in a_i^* indicate changes in the relative values of a_i and b_i while preserving the sum $a_i + b_i$.

Lemma 6. *In the GSM, suppose that $\hat{x}_i = x_i$ for all options and that $i \neq j$.*

7) $\frac{\partial E[D_i]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[D_i^{trans}]}{\partial x_i}$

8) $\frac{\partial E[S_i]}{\partial x_i} = (1 - \sigma_i^*)\frac{\partial E[S_i^{trans}]}{\partial x_i}$

9) $\frac{\partial E[D_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[D_j^{trans}]}{\partial x_i}$

10) $\frac{\partial E[S_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[S_j^{trans}]}{\partial x_i}$

11) $(1 - \sigma_j^* + \sigma_j^* a_j^*)\frac{\partial E[D_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[D_i]}{\partial x_j}$

Note: partial derivatives are taken only with respect to x_i or x_j, not $\hat{x}_j$ and $\hat{x}_j$.

Properties 7-10 relate the partial derivatives of selection and search probabilities in the general ordered-search model to the corresponding partial derivatives in the related fully transparent search model (which is consistent with classic discrete choice). Property 8 holds for any ordered-search model because $\bar{V}_{in}$ only adjusts with x_i if agent n has an opaque pair with firm i. However, properties 7, 9 and 10 rely on ordered search that is consistent with the GSM.

Recall that the GSM defines F_i^{ϵ} and F_i^{μ} in terms of F_i^*. With correct anticipation, $V_{in}^* - \alpha_{in} \sim F(v) = 1 - (\bar{F}_i^*(v-x))^{a_i+b_i}$. Equivalently, for transparent pairs, $V_{in}^* - \alpha_{in} \sim F(v) = 1 - (\bar{F}_i^*(v-x))^{a_i+b_i}$. However, if an agent faces an opaque pairing with firm i where $x_i \neq \hat{x}_i$, then

$$\begin{aligned} V_{in}^* - \alpha_{in} \sim F(v) &= 1 - (\bar{F}_i^*(v-x_i))^{a_i}(\bar{F}_i^*(v-\hat{x}_i))^{b_i} \\ &= 1 - ((\bar{F}_i^*(v-x_i))^{\frac{1}{a_i+b_i}})^{a_i^*}((\bar{F}_i^*(v-\hat{x}_i))^{\frac{1}{a_i+b_i}})^{1-a_i^*} \end{aligned}$$

Thus, for opaque pairs, the CDF of $V_{in}^* - \alpha_{in}$ is one minus the weighted geometric mean of $(\bar{F}_i^*(v-x_i))^{\frac{1}{a_i+b_i}}$ and $(\bar{F}_i^*(v-\hat{x}_i))^{\frac{1}{a_i+b_i}}$. As a result of this type of CDF, any probability that depends on V_{in}^* has the property that the partial derivative with respect to x_i is the same as the corresponding transparent partial derivative with respect of x_i multiplied by a scale factor of $(1 - \sigma_i^* + \sigma_i^* a_i^*)$. For consumers with an opaque pair with option i, a_i^* weights the relative importance of x_i in determining consumer selections.

III. Search Markets

This section discusses firm choices and payoffs in a market. Suppose that firms make choices prior to search.[25] The agents are now consumers and the options are products.

Consider a single firm m. Suppose firm m owns products $i \in I_m$ where I_m is a subset of the options in market. Firm m wants to maximize expected profits $E[\Pi_m] = \sum_{i \in I_m} g^i(x_i, E[D_i], E[S_i])$ where each g^i is strictly decreasing in the first argument, strictly increasing in the second argument and weakly increasing in the third argument.[26] Firm m chooses x_i for all products $i \in I_m$ where x_i is a general term which can be considered the endogenous quality of a product or the negative of the price. Basically, holding the other arguments constant, profits are increasing in demands and search volumes and decreasing in the endogenous quality. This generalized specification is consistent with a range of models including pricing games, wage games, endogenous quality choice models and more.

The following corollary applies the properties of Theorem 2 to analyze the role of opacity in firm m's payoffs.

Corollary 2. *Suppose consumers correctly anticipate the vector of x values* ($\hat{\mathbf{x}} = \mathbf{x}$).

1) Suppose firm m owns exactly one product $(I_m = \{i\})$. *It follows that* $\frac{\partial^2 E[\pi_m]}{\partial \sigma_i^* \partial x_i} < 0$

[25] An equivalent model can be outlined where firms signal their choice to a consumer with commitment if the pairing is transparent and cannot communicate or cannot commit if the pairing is opaque. I treat transparency and/or commitment as exogenous, not as a firm choice.

[26] I denote the partial derivative of g^i with respect to argument 1, 2 and 3 by g^i_1, g^i_2 and g^i_3 respectively.

and

$$\frac{\partial E[\pi_m|\sigma_{in}=1\ \forall n]}{\partial x_i} \leq \frac{\partial E[\pi_m]}{\partial x_i} \leq \frac{\partial E[\pi_m|\sigma_{in}=0\ \forall n]}{\partial x_i}$$

2) Suppose $\frac{\partial E[\pi_m]}{\partial x_i} - g_1^i > 0$ *and* $g_3^i = 0$*. It follows that* $\frac{\partial^2 E[\pi_m]}{\partial \sigma_i^* \partial x_i} < 0$ *and*

$$\frac{\partial E[\pi_m|\sigma_{in}=1\ \forall n]}{\partial x_i} \leq \frac{\partial E[\pi_m]}{\partial x_i} \leq \frac{\partial E[\pi_m|\sigma_{in}=0\ \forall n]}{\partial x_i}$$

Note: As before, the partial derivatives are taken with respect to x_i *holding* $\hat{x}_i$ *constant.*

This corollary states an important property of this model which follows from Property 4 from Lemma 5. If the conditions of either 1 or 2 hold, then firm m's incentives to increase x_i are strictly increasing with respect to average transparency. Thus, firm m's incentive is bounded above by full transparency and below by full opacity. While I have not outlined the market yet, I impose that any pure strategy equilibrium in a search market has the property that $\hat{\mathbf{x}} = \mathbf{x}$. This is a common assumption in the search literature and elsewhere. Thus, if a firm owns only one product, then these results hold. Similarly, this result holds for multi-product firms if search is consistent with the GSM as long as search does not directly affect payoffs.[27]

While these results pertain to the choices of one firm, the results are strong enough to make statements about the incentives of all firms in a market. Consider a

[27]The other condition, $\frac{\partial E[\pi_m]}{\partial x_i} - g_1^i > 0$ holds locally in a neighborhood of any potential interior maximum since $g_1^i < 0$ and thus $\frac{\partial E[\pi_m]}{\partial x_i} - g_1^i > 0$ at any point that a firm would potentially choose.

retail market where x_i is a decreasing function of the price of product i (p_i). In this setting, a firm's incentive to increase prices is increasing in opacity, so its incentive is bounded above by full opacity and below by full transparency. In markets where firm's make wage offers, firm m's incentive to increase wages is strictly decreasing in opacity, so its incentive is bounded above by full transparency and below by full opacity.

I now define the entire market to discuss equilibrium outcomes in search markets. The rest of this chapter makes the following assumptions. First, each firm i owns exactly one option i and chooses x_i. Second, each firm i maximizes their expected profit function $E[\pi_i] = h_i(x_i)E[D_i]$ where $h_i(x_i)$ is a strictly decreasing, differentiable, log-concave function with a support over $(-\infty, \bar{x}_i]$ for some $\bar{x}_i \in \mathbb{R}$. Additionally, $h_i(\bar{x}_i) = 0$. With profits specified this way, the actual choices of the agent enter into the payoffs for firms only through the agent's final selection choice (demand), not through search volumes. Each firm and consumer knows h_i for each firm i and firms know the distributions of all agent heterogeneity. Finally, I assume that all firms make choices simultaneously prior to the search process of agents. This framework is consistent with the following models:

Pricing: *Options correspond to products for sale. The value of x_i equals $q_i - p_i$ where q_i is an exogenous quality shifter and p_i is the price of product i. Furthermore, $h_i(q_i - p_i) = p_i - c_i$ where c_i is the unit cost of product i. Other cost structures that are consistent with the general model can also be used.*

Wage Game: *Options correspond to job openings with set wages. The value* x_i *equals* $w_i + \gamma_i$ *where* w_i *is the wage including all transfers from the potential employer to the potential worker. Consider* γ_i *to be any observable heterogeneity that is known to all potential employers including preferences for the job. Furthermore,* $h_i(w_i + \gamma_i) = \beta_i - w_i - \pi_i^*$ *where* β_i *is the monetary value of the worker and* π_i^* *is the expected profit of not filling the job with the currently available worker. Firms and the potential worker know* γ_i, β_i *and* π_i^* *which are exogenous values in this framework.*

Endogenous Quality Choice: *Options correspond to products for sale. Firms choose the endogenous quality of their product* (x_i)*. While endogenous quality may include the disutility from prices, it also includes other possible choices that increase the average effective value to consumers and the costs of firms to sell the product.*[28]

For clarity, the general results are stated in terms of the endogenous quality choice model with an emphasis on pricing to discuss the relationship between equilibrium prices and unit costs.

A pure strategy equilibrium in this model is a pair of vectors ($\mathbf{x}^*$, $\hat{\mathbf{x}}^*$) where consumers correctly anticipate the choices of firms ($\hat{\mathbf{x}}^* = \mathbf{x}^*$) and each x_i^* maximizes the expected profits of firm i given that all firms $j \neq i$ choose $x_j = x_j^*$ and consumers search with prior $\hat{\mathbf{x}} = \hat{\mathbf{x}}^*$. Since ($\hat{\mathbf{x}}^* = \mathbf{x}^*$) is a equilibrium condition, I refer to a

[28]Examples of increasing endogenous quality include adding accessories for products (e.g. a lens with a camera), adding extra content (e.g. commentary with a movie), paying more for a trade-in (e.g. cars, electronics), reducing other implicit or explicit transfers from the agent to the firm (e.g. reduced fees, free shipping) or any other type of quality adjustment. Thus, h_i is a composite unit profit margin corresponding to a specific level of endogenous quality. Note: Some of these adjustments would likely change fixed costs, not unit costs. I solve the model for unit costs, but similar results follow with different cost structures.

potential equilibrium in this model as a vector $\mathbf{x}^*$. The following theorem outlines existence, uniqueness and comparative statics for equilibrium outcomes in ordered search with the GSM.

Theorem 2. *Assume for each firm i that $h_i = h_i^*$ where h_i^* is log-concave for $x_i < \bar{x}_i$. Also, assume that search is defined by a GSM where f_i^* and f_i^α are log-concave with a support that is not bounded above for each firm i and f_0 is log-concave with a support that is not bounded above. Furthermore, assume that $a_i + b_i \geq 1 \quad \forall i$.*

If there exists a pure strategy equilibrium $\mathbf{x}^*$ in this partially transparent market, then it is unique and $\mathbf{x}^*$ is also the unique equilibrium vector in a fully transparent model with the same GSM where $h_i = (h_i^*)^{\frac{1}{1-\sigma_i^*+\sigma_i^* a_i^*}}$ for each firm i.

Assuming an equilibrium $\mathbf{x}^*$ exists:

1) For any i and j, x_i^* is strictly decreasing in σ_j^* and strictly increasing in a_j^* for $\sigma_i^* > 0$.

2) Expected consumer payoffs are strictly decreasing in σ_j^* and strictly increasing in a_j^* for $\sigma_i^* > 0$.

3) For $j \neq i$, expected profits for firm i are strictly increasing in σ_j^* and strictly decreasing in a_j^* for $\sigma_i^* > 0$.

Additionally, suppose firms are playing the pricing game.

4) p_i^* is strictly increasing in c_j and strictly decreasing in q_j for any firm $j \neq i$.

5) p_i^* is strictly increasing in c_i and strictly increasing in q_i.

6) $x_i^* = q_i - p_i^*$ is strictly decreasing in c_j and strictly increasing in q_j for any firm j.

7) Total Welfare (the sum of all expected firm profits and expected consumer payoffs) is increasing in a_j^* and decreasing in σ_j^*.

This theorem provides economic intuition in partially and fully opaque markets with the GSM assuming that an equilibrium exists at the relevant parameter values. Properties 2 and 7 show that more transparency in markets (lower values of σ_j^*) corresponds to better outcomes on average for consumers and higher overall welfare in the market. In the context of firm pricing, more transparency corresponds to lower prices for consumers. Alternatively, more opacity in markets harms consumers and leads to higher prices.

Importantly, more opacity for a single firm results in higher prices for all firms, with lower a_i^* values exacerbating this effect. With the GSM, a_i^* governs the relative importance of anticipated firm choices versus actual firm choices for consumer purchases. Effectively, higher a_i^* values put more weight on the actual choice of a firm, mitigating the effects of opacity if a firm chooses to deviate from a potential market equilibrium.

This theorem does not show that an equilibrium always exists in partially transparent markets because uniqueness is easier to prove in this model. The proof of uniqueness conditional on existence relies on the shared first-order conditions between firms in the market and firms in a corresponding fully transparent model which has a unique equilibrium. The existence and uniqueness proof for the corresponding fully transparent model follows from Quint (2014).[29]

[29] Another search paper, Choi et al. (2018), also uses the result from Quint (2014) to prove existence

While Quint (2014) proves that the second-order conditions of firms hold in the corresponding fully transparent market, a firm's expected profits before imposing correct anticipation may not be log-concave. Without log-concavity or a similar property, a firm's first-order condition may hold at a point which is not a global maximum for that firm. The following lemma establishes existence given the other assumptions in Theorem 2.

Lemma 7. *Assume for each firm i that $h_i = h_i^*$ where h_i^* is log-concave for $x_i < \bar{x}_i$. Also, assume that search is defined by a GSM where f_i^* and f_i^α are log-concave with a support that is unbounded above for each firm i and f_0 is log-concave with a support that is unbounded above. Furthermore, assume that $a_i + b_i \geq 1 \quad \forall i$.*

Any of the following is a sufficient condition for the existence of the unique pure strategy equilibrium:

(i) The market is fully transparent ($\sigma_i^* = 0 \ \forall i$).

(ii) All options are either fully transparent or fully opaque ($\sigma_i^* \in \{0,1\} \quad \forall i$) and $a_i, b_i \geq 1 \ \forall i$.

(iii) All options are either fully transparent or fully opaque ($\sigma_i^* \in \{0,1\} \ \forall i$) and $\frac{f_i^*}{1-F_i^*}$ is log-concave $\quad \forall i$.

Condition (i) is a direct application of the existence and uniqueness result in Quint (2014) because fully transparent models are equivalent to classic discrete-choice models. Condition (ii) and (iii) provide existence, but require firms to be

and uniqueness in a fully transparent model with homogeneous search costs.

either fully opaque or fully transparent where some firms can be fully opaque and others fully transparent. Note that either $a_i, b_i \geq 1$ or $\frac{f_i^*}{1-F_i^*}$ being log-concave guarantee that f_i^ϵ and f_i^μ are log-concave for each option i.[30] The log-concavity of f_i^ϵ and f_i^μ guarantee a log-concave profit function in these markets for firms that are opaque.

From the earlier theorem, we know that the equilibrium endogenous quality of all options is increasing in a_i^* and decreasing in $1 - a_i^*$. In the context of prices, equilibrium prices are increasing in $1 - a_i^*$ and decreasing in a_i^*. The following lemma provides limit arguments to show that for fully opaque options, the price for option i approaches infinity as a_i^* approaches 0.

Lemma 8. *Under the same assumptions as in theorem* 3, *suppose also that either condition* (ii) *or* (iii) *holds. Suppose* i *is an option where* $\sigma_i^* = 1$. *Also assume that a market equilibrium exists at all relevant parameter values.*

1) $\lim_{a_i^* \to 0} x_i^* = -\infty$ *and* $\lim_{a_i^* \to 0} x_j^* = x_j^{-i}$ *where* x_j^{-i} *is the equilibrium choice of firm* j *if option* i *is removed from the market.*

2) In a pricing game, $\lim_{a_i^* \to 0} p_i^* = \infty$ *and* $\lim_{a_i^* \to 0} p_j^* = p_j^{-i}$ *where* p_j^{-i} *is the equilibrium price of firm* j *if option* i *is removed from the market.*

A firm with a very small a_i^* value effectively prices itself out of a fully opaque market because the firm cannot influence the value that consumers anticipate. In equilibrium, the price of the product is high and consumers correctly anticipate

[30] Many commonly used distributions satisfy the requirement that $\frac{f_i^*}{1-F_i^*}$ is log-concave including TIEV, reverse TIEV and the exponential distribution.

the high price. Arbitrarily high prices can be obtained in search markets with full opacity by setting a_i^* sufficiently small for all firms. Importantly, a_i^* approaching 0 is an extreme case where search threshold heterogeneity in the model is the only heterogeneity that determines the value of $min\{\epsilon_{in}, \mu_{in}\}$. However, as discussed in the section where the GSM is introduced, a_i^* ties threshold and match value heterogeneity together to preserve the distribution of each V^*. This mathematical relationship makes further economic interpretation of a_i^* problematic. The next section outlines a parameterized version of the GSM which allows for a clear economic interpretation of the model parameters.

IV. The Reverse Extreme Value Search Model (REVSM)

This section presents a specific family of search models referred to as Reverse Extreme Value Search Models (REVSMs). A search model $(F_0, [F_i^\alpha, F_i^\epsilon, F_i^\mu]_1^I)$ is a REVSM if:

$$F_i^\epsilon(\epsilon) = 1 - e^{-e^{A_i \epsilon}} \qquad \text{and} \qquad F_i^\mu(\mu) = 1 - e^{-e^{A_i(\mu - y_i)}}$$

where $y_i, A_i \in \mathbb{R}$ and $A_i > 0$ for $i = 1, ..., I$. I do not include an option specific mean shift in each F_i^ϵ because x_i is already included in V_{in}, so an additional shift would be redundant.

The REVSM model is based on extreme value distributions that are the reverse version of the more common Type I Extreme Value (TIEV) distribution from classic discrete choice. The family of reverse TIEV distributions has been discussed in theoretical papers like Anderson and de Palma (1999) as well as empirical papers like Misra (2005). The key property of this family of distributions is that the family is preserved under minimization. For example, because F_i^{ϵ} and F_i^{μ} are both reverse TIEV, F_i^{min} is also reverse TIEV. To be exact, $F_i^{min} = 1 - e^{-e^{A_i(v)}(1+e^{-A_i y_i})} = 1 - e^{-e^{A_i(v - \frac{1}{A_i} g(A_i y_i))}}$ where $g(z) = -ln(1 + e^{-z})$ for each option i. Thus, in the REVSM, $E[V_{in}^*] = E[x_i + \alpha_{in} + min\{\epsilon_{in}, \mu_{in}\}]$ is increasing and concave in y_i because $\frac{\partial E[V_{in}^*]}{\partial y_i} = \frac{1}{e^{A_i y_i}+1}$.

The following lemma establishes equivalent specifications for a REVSM.

Lemma 9. *The following specifications are all equivalent given* $(\mathbf{x}, \mathbf{y}, \mathbf{A}, F_0, [F_i^{\alpha}]_{i=1}^{I})$.

1) Search defined with threshold heterogeneity (μ):

$(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$ *where* $F_i^{\epsilon}(v) = 1 - e^{-e^{A_i v}}$ *and* $F_i^{\mu}(v) = 1 - e^{-e^{A_i(v - y_i)}}$ *for* $i = 1, ..., I$.

2) Search defined with search cost heterogeneity (κ):

$(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\kappa}]_1^I)$ *where* $F_i^{\epsilon}(v) = 1 - e^{-e^{A_i v}}$ *and* $F_i^{\kappa}(v) = e^{-e^{A_i(m_i(v) - y_i)}}$ *for* $i = 1, ..., I$.

3) Geometric Search:

$(F_0, [F_i^{\alpha}, F_i^*, a_i, b_i]_1^I)$ *where* $a_i = 1$, $b_i = e^{-A_i y_i}$ *and* $F_i^*(v) = 1 - e^{-e^{A_i v}}$ *for* $i = 1, ..., I$.

4) Geometric Search where $b_i = 1 - a_i$ *for each i:*

$(F_0, [F_i^{\alpha}, F_i^*, a_i]_1^I)$ *where* $a_i = \frac{1}{1+e^{-A_i y_i}}$ *and* $F_i^*(v) = 1 - e^{-e^{A_i(v - \frac{1}{A_i} g(A_i y_i))}}$ *for* $i = 1, ..., I$.

This lemma demonstrates that the REVSM can be expressed as a Geometric Search Model (specification 3 and 4) and that higher y_i values correspond to first-order stochastically dominant decreases in the distribution of search costs for option i (specification 2). In this model, $a_i^* = \frac{1}{1+e^{-A_i y_i}}$ depends on y_i and A_i only. Moreover, specification 4 shows that changes in y_i propagate through as mean shifts in the distribution of V_{in}^* where $g(A_i y_i) = -ln(1 + e^{-A_i y_i}) = ln(a_i^*)$.

In the context of the overall model, increases in y_i generate an increase in the mean of V_{in}^* and an increase in a_i^*. y_i, which relates to search costs, and q_i, which represents exogenous quality, have similar effects in the model holding a_i^* constant. However, q_i shifts thresholds and match values whereas y_i shifts thresholds only. Since y_i is threshold-specific, an increase in y_i also corresponds to an increase in a_i^*.

Let $z_i = x_i + \frac{1}{A_i} g(A_i y_i)$ where z_i can be interpreted as the effective quality of option i including the distribution of search costs associated with option i given correct anticipation. The following market results for the REVSM build on the fact that the REVSM is consistent with the GSM.

Corollary 3. *Assume the market structure outlined in the previous section, but with search defined by a REVSM. Assume for each firm i that $h_i = h_i^*$ where h_i^* is log-concave for $x_i < \bar{x}_i$. Also, assume that f_i^* and f_i^α are log-concave with a support that is not bounded above for each firm i and f_0 is log-concave with a support that is not bounded above.*

All of the results of Theorem 3 hold in this market.

1) $z_i^* = x_i^* + \frac{1}{A_i} g(A_i y_i)$ is strictly increasing in y_j for any firm j.

2) For $j \neq i$, x_i^* is strictly increasing in y_j. In the pricing game, p_i^* is strictly decreasing in y_j.

3) If firm i is fully transparent, x_i^* is strictly decreasing in y_i. In the pricing game, p_i^* is strictly increasing in y_i.

4) If firm i is partially or fully opaque, x_i^* may be increasing or decreasing in y_i. In the pricing game, p_i^* may be increasing or decreasing in y_i.

5) A unique equilibrium exists if options in the market are either transparent or fully opaque ($\sigma_i^* \in \{0, 1\} \;\; \forall i$).

This corollary establishes the role of y_i in terms of the outcomes of the model. Property 1 follows from the fact that higher mean values of V_{in}^* and higher a_i^* values due to an increase in y_i result in a higher overall effective quality including the relative search costs of the option i. Since firm choices are strategic complements, all other firms respond with higher endogenous quality, so x_i^* and z_i^* are increasing in y_j for $j \neq i$. Property 2 follows from this argument. However, property 3 and property 4 relate firm i's endogenous quality (not including the change in search costs) and firms i's price to changes in y_i. In the fully transparent setting, higher y_i values result in higher prices for firm i and lower endogenous quality. This parallels the effect of a change in q_i on p_i in the previous section because y_i increases the mean V_{in}^* but does not change $g_i(x_i)$.

To better understand property 4, consider the model for a monopolist. Suppose $I = 1$, $F_0(v) = 1 - e^{-e^v}$, $\alpha_{1n} = 0 \;\; \forall n$ and $A_1 = 1$. Then,

$$E[D_1^{trans}] = N\int_{-\infty}^{\infty}(1-e^{-e^v})e^{-e^{v-x_1-g(y_1)}+v-x_1-g(y_1)} = 1-\frac{e^{-x_1-g(y_1)}}{1+e^{-x_1-g(y_1)}} = \frac{1}{1+e^{-x_1-g(y_1)}}.$$

From lemma 6, at points where $\hat{x}_i = x_i$

$$\frac{\partial ln(E[\pi_1])}{\partial x_i} = \frac{h_1'(x_1)}{h_1(x_1)} + (1-\sigma_1^* + \sigma_1^* a_1^*)\frac{\partial ln(E[D_1^{trans}])}{\partial x_i}$$

$$= \frac{h_1'(x_1)}{h_1(x_1)} + (1-\sigma_1^* + \frac{\sigma_1^*}{1+e^{-y_1}})\frac{e^{-x_1}(1+e^{-y_1})}{1+e^{-x_1-g(y_1)}}$$

If $\sigma_i^* = 0$, then $\frac{\partial ln(E[\pi_1])}{\partial x_i}$ is strictly decreasing in y_i. Alternatively, if $\sigma_i^* = 1$, then $\frac{\partial ln(E[\pi_1])}{\partial x_1}$ is strictly increasing in y_1. Thus, x^* is decreasing in y_1 for a fully transparent monopolist and increasing in y_1 for a fully opaque monopolist. Similar arguments can be made for prices where the fully opaque monopolist decreases the price if y_1 increases. In this fully opaque setting, the effect that y_1 has on the partial derivative through $a_1^*(y_1)$ dominates the countervailing effect from $g_1(y_1)$.

In contrast with many past papers on search markets, the results in this section pertain to changes in the underlying distribution of search costs instead of changes to a homogeneous search cost for each option.[31] Similarly, while I discuss increases and decreases in search costs, the distribution is changing by more than a simple mean shift.

Now consider a specific version of the REVSM, referred to as the standard

[31]Moraga-González et al. (2017b) and Moraga-González et al. (2017a) outline and solve search markets with heterogeneous search costs, but with a different type of search model.

REVSM where $F_0(v) = 1 - e^{-e^v}$, $\alpha_{in} = 0 \ \ \forall i, n$ and $A_i = 1 \ \ \forall i$. In this model, the only primitives are the two parameter vectors, $\mathbf{x}$ and $\mathbf{y}$. As the following lemma shows, a major advantage of the standard REVSM is that both expected demands and expected search volumes can be expressed as closed-form expressions that resemble multinomial logit.[32]

Lemma 10. *Assume a standard REVSM. Let $\mathcal{P}_{-i}$ denote the power set of the set of options excluding i. Let $z_i = x_i + g(y_i)$.*

$$E[D_i] = N \ Prob[V^*_{in} > \underset{j \neq i}{max} V^*_{jn}] = N \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \frac{e^{-z_i}}{e^{-z_i} + \sum\limits_{j \in \theta} e^{-z_j}}$$

$$E[S_i] = N \ Prob[\bar{V}_{in} > \underset{j \neq i}{max} V^*_{jn}] = N \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \frac{e^{-x_i - y_i}}{e^{-x_i - y_i} + \sum\limits_{j \in \theta} e^{-z_j}}$$

Let I^s denote the realized set of options which are searched (consideration set). Let i^* be the realized selection where $i^* \in I^s$.

$$Prob[i^*, I^s] = \sum_{\theta \in \mathcal{P}_{-i^*}} (-1)^{|\theta|} \frac{e^{-z_i}}{e^{-z_i} + \sum\limits_{j \in (\theta \cup I^s)} e^{-x_j - y_j} + \sum\limits_{j \in (\theta \cap I^s)} e^{-x_j}}$$

Moreover, $ln(E[D_i])$, $ln(E[S_i])$ and $ln(Prob[i^*, I^s])$ are concave in terms of $\mathbf{x}$ and $\mathbf{y}$.

This lemma establishes the closed-form result of the standard REVSM which is applied in the next chapter for empirical estimation. The expressions above resemble the expected demands from TIEV but require summation over $\mathcal{P}_{-i}$. This key difference arises because reverse TIEV is preserved over minima whereas

[32]The requirement that $A_i = 1$ is included to simplify the equations. This requirement is not necessary for closed form expressions.

TIEV is preserved over maxima. Ordered search inherently generates maxima over minima because the option that is selected corresponds to the maximum $V_i^* = min\{V_i, \bar{V}_i\}$ value.[33]

The expression of $Prob[i^*, I^s]$ is best understood by relying on the earlier results for optimal search. If option j is searched, but not selected by agent n, then $\bar{V}_{jn} > V_{i^*n}^* > V_{jn}$. If option j is not searched by agent n, then $V_{i^*n}^* > \bar{V}_{jn}$. Thus for ordered search where each $\alpha_{in} = 0$,

$$Prob[i^*, I^s] = \int_{-\infty}^{\infty} \prod_{j \in I^s - \{i^*\}} \left(F_j^{\epsilon}(v - x_j)(1 - F_i^{\mu}(v - x_j)) \right) \prod_{j \notin I^s} \left(F_j^{\mu}(v - x_j) \right) dF_i^{min}(v - x_i)$$

For the standard REVSM, this integral can be expressed as the earlier summation in lemma 6. I now move on to demand-side identification before returning to the standard REVSM for an empirical application.

[33] While the GSM can be used to define a search model with demands that exactly match multinomial logit, the REVSM has the added advantage of closed-form expressions of search volumes and $Prob[i^*, I^s]$. Additionally, there is a direct economic interpretation for y_i and z_i.

Chapter 3: Identifying and Estimating the GSM

This chapter first demonstrates demand-side identification for the GSM. Then it estimates the REVSM with experimental data as an illustration of the robustness of the REVSM. Finally, the chapter addresses the importance of opacity when partially identifying supply-side parameters in search markets.

I. Demand-side Identification

My work here presents a proof of identification for the demand-side of the GSM model. Suppose the choices of consumers arise from a GSM where $\alpha_{in} = 0$ and $b_i = 1 - a_i \;\; \forall i, n$. I consider two different scenarios for model identification from the perspective of an econometrican.

Scenario 1: The econometrician observes a distribution of markets with realized aggregate demands and search volumes. The econometrician knowns the number of consumers in the markets (N). The econometrician also knows that the aggregate demands and search volumes in each realized market are generated by the GSM with CDFs F_0 and F_i^* for each product i, where aggregate demands and search volumes are the same as expected demands and search volumes in the GSM. For simplicity, N, F_0 and F_i^* for each product i are the same across all possible realized markets. However, the econometrician does not observe the distribution of $(\mathbf{x}, \mathbf{a})$ which generates differences in the realized markets.

Scenario 2: The econometrician observes a distribution over the set of possible realized search and selection choices for a consumer ($N = 1$). The econometrician knows that the distribution of realized actions by the consumer is generated by the GSM with CDFs F_0 and F_i^* for each product i. However, the econometrician does not observe the value of $(\mathbf{x}, \mathbf{a})$ which generates the distribution of consumer actions.

Proving identification in both of these scenarios relies on showing that the GSM maps each possible pair of parameter vectors $(\mathbf{x}, \mathbf{a})$ to a unique pair of expected demands and search volume vectors ($E[\mathbf{D}], E[\mathbf{S}]$), where each possible pair of expected demand and search volume vectors ($E[\mathbf{D}], E[\mathbf{S}]$) is mapped by some pair of model parameters. This approach is similar to that of Berry (1994). The simplest level of Berry's demand-side identification strategy is the inversion of market shares to mean match values through the lens of a classic discrete-choice model.[34]

For scenario 1, this bijective mapping implies that a unique distribution of $(\mathbf{x}, \mathbf{a})$ can be inferred from the observed distribution of realized markets with aggregate demands and search volumes. For scenario 2, this bijective mapping implies that the value of the parameter vectors $(\mathbf{x}, \mathbf{a})$ can be inferred from the search and selection probabilities over the distribution of consumer choices. Note that expected demands and search volumes are equivalent to selection and search probabilities in the mapping because the population size N is known. Thus, scenario 2 is

[34]See Berry (1994) for a more detailed analysis of demand-side identification with classic discrete choice.

technically over-identified because the econometrician has extra information about the distribution of realized search paths and selections. By restricting the possible distributions in scenario 2 to the distributions that can arise from the specified GSM model, scenario 2 is identified. An important aspect of the GSM is that it can explain any realized search path combined with a final selection from the set of searched options. The GSM is flexible in this regard, but it is not flexible enough to match any distribution of search paths and selections without adding more parameters to the model.

Let $\mathbf{Q} = \mathbb{R}^I \times [0,1]^I$ where $\mathbf{Q}$ is the space of possible parameters. I define the space of parameters so that each a_i can equal 0 or 1. The case where $a_i = 1$ corresponds to i being a zero search cost option which is always searched ($E[S_i] = N$). However, $a_i = 0$ is not well defined because the limiting case corresponds to a mass point in match values at the upper bound of the support. For this section, consider $a_i = 1$ to be the limiting case where search costs (threshold) heterogeneity is the dominant factor in selections so $E[D_i] = E[S_i]$.[35]

Let $\mathbf{W} = \{(\mathbf{d}, \mathbf{s}) \in (0, N)^{2I} : \mathbf{d} \leq \mathbf{s}, \sum_{i=1}^{I} d_i \leq N\}$ where $\mathbf{W}$ is the set of all possible vectors of expected demands and search volumes. The restrictions on $\mathbf{W}$ follow from the fact that an option cannot be searched more times than there are consumers and the total number of selections in the system (excluding the outside option) cannot exceed the number of consumers. Additionally, $\mathbf{W}$ needs to

[35] $a_i = 1$ is similar to search where demands and search volumes are equal because a consumer knows their value prior to search. In this setting, search costs can be reinterpreted as a fixed cost of accessing the option.

be restricted because an option must be searched before it can be purchased, so expected demand cannot exceed expected search volumes.

Let $\mathbf{H} : \mathbf{Q} \to \mathbf{W}$ be the function defined by the optimal decision-making process of agents that maps a given pair of vectors $(\mathbf{x}, \mathbf{a})$ to the vector of expected demands and expected search volumes for options across all agents. More specifically, for any $(\mathbf{x}, \mathbf{a}) \in \mathbf{X}$, $\mathbf{H}(\mathbf{x}, \mathbf{a}) = (\mathbf{E}[\mathbf{D}], \mathbf{E}[\mathbf{S}])$ where $E[\mathbf{D}]$ and $E[\mathbf{S}]$ are the resulting vectors shown in the earlier section on optimal search behavior.

Corollary 4. *Assuming that F_0 and F_i^* for each option i have a support over $\mathbb{R}$, $\mathbf{H}$ is a continuous bijection from $\mathbf{Q}$ to $\mathbf{W}$.*

The proof of this corollary is comprised of two parts due to the invariance of expected demands with respect to a_i for any option i in the search model. Using the same proof provided in the appendix of Berry (1994), the optimal decision-making process of agents in the model results in a bijection from the space of possible $\mathbf{x}$ vectors to the space of possible observed demand vectors $\mathbf{d}$. By result 4 in Lemma 4, there exists a unique parameter vector $\mathbf{a}$ determined by the observed search volume vector $\mathbf{s}$ and the unique vector $\mathbf{x}^*$ that corresponds to the observed demands.

Given Lemma 4 and the fact that this model has the same inherent demand properties of the parallel CDM where $F_i^\beta = F_i^*$ for each option i, the GSM provides an easy method for extending structural models that were originally designed for classic discrete choice because the underlying identification strategy for $\mathbf{x}$ is preserved.

An important aspect of this section is that it is completely agnostic about the interpretation of the parameters as defined in the model. To be clear, identification and estimation in BLP relies on this type of inversion, but adds further structure to the interpretation of mean match values given other observables like prices and agent characteristics. Future papers can take the work presented here and use it as an identification technique. How the distribution of parameters is then interpreted and structured (eg. random coefficients) given observables that extend past $E[\mathbf{D}]$ and $E[\mathbf{S}]$ will depend on specific data sets and the goals of the empirical analysis.

The work in this section relies on the GSM where $b_i = 1 - a_i$ for each option i. This specification links changes in the CDF of search costs to changes in the CDF of match values, which clouds the economic interpretation of a_i. In the next section, I now return to the REVSM to estimate y_i instead of a_i because y_i determines only the distribution of search costs and does not affect the distribution of match values.

II. Estimating the REVSM

This section uses the standard REVSM to analyze search data that includes each individual agent's selection and consideration set. In this model, the primitives are the two parameter vectors $\mathbf{x}$ and $\mathbf{y}$. By the results in the previous section, $\mathbf{x}$ and $\mathbf{y}$ are identified by the selection and search probabilities for each option. I extend this identification argument to search data that includes each individual agent's selection and consideration set.[36] As earlier, let I_n^s denote the consideration set for a given consumer n, where a consideration set is the set of all options that are searched by the agent. Additionally, let i_n^* denote the option that agent n selects and let $[i_n^*, I_n^s]_{n=1}^N$ be a data set of consumer choices that are independently drawn from the population of possible choices.[37]

I now apply the empirical model to a dataset that was created for and analyzed in Reutskaja et al. (2011). The data were generated by an experiment in which people could select options on a screen. Options represented choices of snack foods, and the participants received a snack at the end that was randomly drawn from the participant's selections throughout the experiment. Each participant rated options before any selections were made. These options were then randomly placed on the screen with 4, 9 or 16 posted at a time. Participants had 3 seconds to choose, and all eye movements were monitored to infer which options were searched during the process. A total of 41 participants made 75 choices each (25 choices of each size).[38] While this data set is rich, for simplicity, I use only the set of searched options and the final selection.[39] I am ignoring the extra details of the data to focus on one feature of the experiment.

[36]Since $(\mathbf{x}, \mathbf{y})$ are identified by selection and search probabilities, the model is over-identified with the extra level of detail provided by consideration sets.

[37]$Prob[i_n^*, I_n^s]$ is log-concave in $\mathbf{x}$ and $\mathbf{y}$ so the log-maximum likelihood problem for any $[i_n^*, I_n^s]_{n=1}^N$ is concave.

[38]See Reutskaja et al. (2011) for a clear outline and analysis of the general patterns in the data. Their paper also estimates search models using the data. I provide this estimation of the standard REVSM as a validation exercise and initial application of my work.

[39]This removes any order from the data, any measurement of time spent looking at each option and the rating of each randomly placed option.

By design, options in the experiment were randomly placed on the screen. However, in the data, options placed in some positions on the screen were more likely to be searched and more likely to be selected. Due to the random placement of options, the differences in selection probabilities should arise from differences in search characteristics, not differences in match values. I test this hypothesis using the standard REVSM, where x_i represents a shift in the match value of all options in a given position and y_i represents a shift in the thresholds (and search costs) of all options in a given position. Thus, x_i and y_i are characteristics of each position on the screen across all options displayed in these positions. For simplicity, the analysis is only over the data with exactly 4 options. Top left is called option 1, top right is called option 2, bottom left is called option 3 and bottom right is called option 4.

The hypothesis that objects in each position have the same average match value is equivalent to $x_1 = x_2 = x_3 = x_4$. To test this, I run two regressions on the data. The first estimates the standard REVSM model. Since a selection must always be made from the set of 4 options, I normalize the value of x_3 to 0 in this regression. I chose x_3 because it was closest to the average. The second regression estimates the standard REVSM, but imposes the hypothesis that $x_1 = x_2 = x_3 = x_4 = 0$. These results are displayed in Table 1 in the appendix with the last column displaying the differences in y estimates between these two regressions.

As noted earlier, each observed consideration set and selection is assumed to be independently drawn from the population of possible search and selection choices. This can be relaxed for a more robust estimation, but the goal is to illustrate a trend without focusing on fixed effects and similar tools. As is seen in Table 1, the estimated values for x_1, x_2 and x_4 are all within a standard deviation of 0. However, the estimates for y_1, y_2, y_3, and y_4 differ from each other at a more significant level. Moreover, there is relatively little change in the estimates of the y parameters between regression 1, where x values can vary, and regression 2, where all options are assumed to have the exact same x values.

This data analysis serves two purposes. First, it shows how simple the model is to apply and estimate on data. Second, it shows that the model can recover aspects of search in the experiment where the features of the data are defined by the particular setting. This model can be used in a similar fashion to separately identify value and search cost parameters in settings where a platform observes search and selection behavior.

III. Supply-side Inversion and Partial Identification

This section outlines the results in the pricing model only, though similar results hold for other versions of the model. Consider a scenario where the econometrician knows all features of the demand side of the market that is defined by the GSM including including prices, expected demands as a function of prices and the vector of $\mathbf{a}$ values. The econometrician does not observe the vector of unit costs $\mathbf{c}$ for firms in the market.

In this setup, common supply-side estimation methods can be used with the GSM for fully transparent models. However, the first-order conditions of firms must be adapted to estimate unit costs in partially or fully opaque markets. Consider a pricing game with multiproduct firms. Let i be a product owned by firm m. It follows that

$$\frac{\partial E[\pi_m]}{\partial p_i} = E[D_i] + (1 - \sigma_i^* + \sigma_i^* a_i) \sum_{j \in I_m} (p_j - c_j) \frac{\partial E[D_j^{trans}]}{\partial p_i}.$$

All derivatives are evaluated at values where correct anticipation holds because the econometrician observes the equilibrium outcomes. In the fully transparent case ($\sigma_i^* = 0$), this derivative is the same as the partial derivative with the corresponding classic discrete-choice model.

$$\frac{\partial E[\pi_m]}{\partial p_i} = E[D_i] + \sum_{j \in I_m} (p_j - c_j) \frac{\partial E[D_j^{trans}]}{\partial p_i}$$

However, in the fully opaque case ($\sigma_i^* = 1$), the partial derivative is

$$\frac{\partial E[\pi_m]}{\partial p_i} = E[D_i] + a_i \sum_{j \in I_m} (p_j - c_j) \frac{\partial E[D_j^{trans}]}{\partial p_i}.$$

Using the first-order conditions for each option, a supply-side inversion can be applied to estimate the vector of unit costs in the market for both the fully transparent case and the fully opaque case.[40] For simplicity, I focus on the scenario where each firm i owns one product i.

In this setting, the firm's first-order condition is

$$\frac{\partial E[\pi_i]}{\partial p_i} = E[D_i] + (1 - \sigma_i^* + \sigma_i^* a_i)(p_i - c_i)\frac{\partial E[D_i^{trans}]}{\partial p_i} = 0$$

The values of p_i, $E[D_i]$ and $\frac{\partial E[D_i^{trans}]}{\partial p_i}$ for each option i are held fixed because these values do not directly depend on the relative opacity of the market. This first-order condition can be manipulated to express the unit cost parameter as a function of $(1 - \sigma_i^* + \sigma_i^* a_i)$ while preserving the price at which the first-order condition holds.

$$c_i = p_i + \frac{E[D_i]}{(1 - \sigma_i^* + \sigma_i^* a_i)\frac{\partial E[D_i]}{\partial p_i}}$$

This equation is the firm specific supply inversion that identifies unit costs given demand-side estimates, prices and a fixed level of opacity for each firm. Since $\frac{\partial E[D_i]}{\partial p_i}$ is negative, the value of c_i is decreasing in σ_i^* and increasing in a_i. The following lemma relies on this property for partial identification of unit costs.

Lemma 11. *Suppose equilibrium prices are known to the econometrician. Suppose the demand and search parameters are known to the econometrician* $(\mathbf{x}, \mathbf{a})$*.*

Let $\mathbf{c}^{trans}$ *be the identified unit cost vector assuming transparency* $(\sigma_i^* = 0 \quad \forall i)$.

[40] As is common in methodologies like BLP, these formulas can be extended to include heterogeneous a_i values that depend on agent types. However, a_i values would then remain inside the expectations of firm demands when calculating opaque prices. Additionally, for many applications of the identification in this section, endogeneity concerns would need to be addressed using instruments or context specific arguments.

Let $\mathbf{c}^{op}$ *be the identified unit cost vector assuming full opacity* ($\sigma_i^* = 1 \quad \forall i$).

For any vector of firm opacities in the market, the identified vector of unit costs $\mathbf{c}^*$ *satisfies:*

$$\mathbf{c}^{op} \leq \mathbf{c}^* \leq \mathbf{c}^{trans}$$

Now suppose the demand parameters ($\mathbf{x}$) *are known, but* $\mathbf{a}$ *is not known.*

For any vector of relative firm opacities in the market, the identified vector of unit costs $\mathbf{c}^*$*:*

$$\mathbf{c}^* \leq \mathbf{c}^{trans}$$

However, nothing more is known without making further assumptions about opacity (or imposing non-negativity of unit costs).

This lemma bounds the supply-side parameters that can arise in a search market if both mean value parameters and search parameters have been identified. However, if the search parameter vector $\mathbf{a}$ cannot be identified, then unit costs can only be bounded above because unit cost estimates from classic discrete-choice models only provide an upper bound of $\mathbf{c}^{trans}$. In this setting, the econometrician can identify only a lower bound on the markups of firms. If the upper bound is assumed to be the true vector of cost parameters when the market is not fully transparent, then the estimates of markups and the market power of firms (e.g. the Lerner index) will be biased downward.

However, even in the absence of search data, market characteristics can be informative about the magnitude of the bias. Taking the FOC for a single product firm i, the identified markup and the identified Lerner Index (L_i) are respectively:

$$p_i - c_i = \frac{-E[D_i]}{(1-\sigma_i^* + \sigma_i^* a_i)\frac{\partial E[D_i^{trans}]}{\partial p_i}} \quad \text{and} \quad L_i = \frac{p_i - c_i}{p_i} = \frac{-1}{(1-\sigma_i^* + \sigma_i^* a_i)M_i}$$

where M_i is the elasticity of demand for product i with full transparency. Thus, the identified markup and Lerner index from the classic discrete-choice model is off by a factor of $(1-\sigma_i^* + \sigma_i^* a_i)$ from the identified markup with partial or full opacity. For empirical work that uses classic discrete choice, lower estimated markups for firms implies a lower bias when not formally modeling opacity. Similarly, lower search costs relative to match values (a_i close to 1) and more transparency (σ_i^* close to 0) correspond to a relatively small bias when applying classic discrete choice. These properties allow researchers to make informed decisions with regard to the bias resulting from applying classic discrete-choice methods to identify supply-side parameters in search markets.

As demonstrated throughout this chapter, the GSM can be used for both demand-side and supply-side estimation of model parameters in search markets.

Conclusion

This dissertation introduces results for ordered-search models following three different themes. First, it establishes demand-level equivalence between ordered search and classic discrete choice. It provides a new family of search models (GSMs and REVSMs) which are both tractable and applicable as shown by my results for markets, demand-side identification and supply-side identification. Additionally, it establishes the role opacity plays in search markets to demonstrate the resulting bias from applying classic discrete choice to analyze partially or fully opaque markets.

In part due to the ever-increasing role that the internet plays in our lives, search markets are becoming more and more relevant in economics. Progress in the analysis and application of search models is important to analyze the underlying properties of massive markets which are continually growing and changing. For example, position auctions like the mechanism used for Google Adwords generate billions of dollars in revenue each year. Amazon continues to disrupt the retail industry, forcing many competitors to shut down their physical stores. Internet service providers have been the focus of many headline news stories due to the argument over whether access to the internet should be regulated or whether websites should be advantaged based on side payments. Moreover, many policies established by platforms and governments, both online and offline, have attempted to mitigate problems with opacity in search markets.

Proper investigations of important policy questions in search markets require tractable models of informational frictions. As demonstrated in this dissertation, the GSM is a tractable model of search for theoretical and empirical applications. In future work, I plan to use the GSM to analyze position auctions, entry games, advertising and switching costs in the context of search markets.

Appendix

Table 1

	Regression 1	Regression 2	Difference in y estimates
x_1	-0.056 (0.070)	0	
x_2	0.033 (0.072)	0	
x_4	-0.031 (0.074)	0	
y_1	2.213 (0.093)	2.166 (0.083)	0.047
y_2	1.725 (0.085)	1.777 (0.071)	-0.052
y_3	1.413 (0.079)	1.428 (0.062)	-0.016
y_4	1.533 (0.080)	1.514 (0.064)	0.019

As stated in the main body of the paper, regression 1 estimates the standard REVSM model using the data set. There are 1019 observations. Regression 2 estimates the standard REVSM, but imposes the hypothesis that $x_1 = x_2 = x_3 = x_4 = 0$. In regression 1, all x estimates are within 1 standard deviation of 0. The estimates in regression 2 for the y values are also within 1 standard deviation of the estimates from regression 1.

Lemma 1 (Extended)**.** *For each option i, m_i is a strictly decreasing bijection from $(0, \infty)$ to $(-\infty, \bar{\epsilon}_i)$ and m_i^{-1} is a strictly decreasing bijection from $(-\infty, \bar{\epsilon}_i)$ to $(0, \infty)$. Additionally,*

$$E[max\{\hat{V}_n, V_{in}\}|\alpha_{in}] - \hat{V}_n > \kappa_{in}$$

$$\Leftrightarrow$$

$$\bar{V}_{in} = \alpha_{in} + x_i + m_i(\kappa_{in}) > \hat{V}_n.$$

Furthermore,

$$m_i^{-1}(r) = E[(\epsilon_{in} - r)1[\epsilon_{in} - r > 0]] = \int_{-\infty}^{\infty} (v - r)1[v - r > 0]\ f_i^{\epsilon}(v)\ dv = \int_{r}^{\infty} (1 - F_i^{\epsilon}(v))\ dv.$$

By extension, $m_i'^{-1}(r) = -(1 - F_i^{\epsilon}(r))$, $m_i''^{-1}(r) = f_i^{\epsilon}(r)$, $\lim_{r \to -\infty} m_i^{-1}(r) = \infty$ and $\lim_{r \to \infty} m_i^{-1}(r) = 0$. Additionally, $r \geq \bar{\epsilon}_i \Leftrightarrow m_i^{-1}(r) = 0$.

This lemma is the direct result of the properties inherent to expectations over non-negative random variables with a finite expected value. Since $E[\epsilon_{in}]$ is finite, $E[\epsilon_{in} - r]$ is a finite. By the requirements of Lebesgue integration, $E[(\epsilon_{in} - r)1[(\epsilon_{in} - r) > 0]]$ is a finite expectation for a non-negative random variable. For any non-negative random variable X with a finite expected value, $E[X] = \int_0^{\infty} Prob[X > x]dx$. Thus,

$$m_i^{-1}(r) = \int_{0}^{\infty} (1 - F_i^{\epsilon}(v + r))\ dv = \int_{r}^{\infty} (1 - F_i^{\epsilon}(v))\ dv$$

The arguments that $m_i'^{-1}(r) = -(1 - F_i^{\epsilon}(r))$, $m_i''^{-1}(r) = f_i^{\epsilon}(r)$ and $\lim_{r \to -\infty} m_i^{-1}(r) = \infty$ follow directly from this specification of m_i^{-1}.

The argument that $\lim_{r \to \infty} m_i^{-1}(r) = 0$ follows from Lebesgue's Dominated Convergence Theorem because $(v-r)1[v-r>0]\ f_i^{\epsilon}(v)$ is weakly decreasing in r and converges to 0 as $r \to \infty$ for any fixed v. Since $m_i^{-1}(r)$ is finite and decreasing in r, $\underset{r\to\infty}{lim}\, m_i^{-1}(r) = E[0] = 0$

Thus, m_i is a strictly decreasing bijection from $(0,\infty)$ to $(-\infty,\bar{\epsilon}_i)$ and m_i^{-1} is a strictly decreasing bijection from $(-\infty,\bar{\epsilon}_i)$ to $(0,\infty)$.

Finally, the statement that

$$E[max\{\hat{V}_n, V_{in}\}|\alpha_{in}] - \hat{V}_n > \kappa_{in}$$

$$\Leftrightarrow$$

$$\bar{V}_{in} = \alpha_{in} + x_i + m_i(\kappa_{in}) > \hat{V}_n$$

follows from:

$$E[max\{\hat{V}_n, V_{in}\}|\alpha_{in}] - \kappa_{in} > \hat{V}_n$$

$$\Leftrightarrow$$

$$E[max\{\hat{V}_n, x_i + \alpha_{in} + \epsilon_{in}\} - \hat{V}_n|\alpha_{in}] > \kappa_{in}$$

$$\Leftrightarrow$$

$$E[max\{\hat{V}_n - x_i - \alpha_{in}, \epsilon_{in}\} - (\hat{V}_n - x_i - \alpha_{in})|\alpha_{in}] - \kappa_{in} > \hat{V}_n$$

$$\Leftrightarrow$$

$$m_i^{-1}(\hat{V}_n - \alpha_{in} - x_i) > \kappa_{in}$$

$$\Leftrightarrow$$

$$\bar{V}_{in} = \alpha_{in} + x_i + m_i(\kappa_{in}) > \hat{V}_n$$

The math here consists of manipulations of the expectations to rewrite the problem using the m_i^{-1} function to introduce the additive property of $m_i(\kappa_{in})$.

Lemma 2. *Let* $V_{in}^{*} = min\{\bar{V}_{in}, V_{in}\}$ *for* $i = 1, ..., I$ *and* $V_{0n}^{*} = V_{0n}$. *Under optimal search, agent n selects option i if*

$$V_{in}^{*} > \max_{j \neq i} V_{jn}^{*}$$

and agent n does not select option i if

$$V_{in}^{*} < \max_{j \neq i} V_{jn}^{*}.$$

Additionally, agent n searches option i if

$$\bar{V}_{in} > \max_{j \neq i} V_{jn}^{*}$$

and agent n does not search option i if

$$\bar{V}_{in} < \max_{j \neq i} V_{jn}^{*}.$$

See Armstrong (2017) for a related proof of this lemma. Let V_{max}^{*} be the the highest V^{*} excluding option i. Assume $min\{V_{in}, \bar{V}_{in}\} > V_{max}^{*}$. Then for any option $j \neq i$, $min\{V_{in}, \bar{V}_{in}\} > min\{V_{jn}, \bar{V}_{jn}\}$. Thus, either V_{jn} or $\bar{V}_{jn}$ (or both) are less than $min\{V_{in}, \bar{V}_{in}\}$. $V_{jn} < min\{V_{in}, \bar{V}_{in}\}$ implies that option j is not selected because option i would be searched before a match value of V_{jn} would be chosen and

$V_{in} > V_{jn}$. Alternatively, $\bar{V}_{jn} < min\{V_{in}, \bar{V}_{in}\}$ implies that option i is searched prior to option j and option i's match value is high enough that option j would not be searched. Thus, any option that is not option i will not be selected. Additionally, $min\{V_{in}, \bar{V}_{in}\} > V^*_{0n} = V_{0n}$, so the outside option is not selected because option i has a higher match value and search threshold. Combined with the fact that the agent must pick an option, this implies that option i is selected. By a parallel argument, any option that has the highest V^* is selected. A similar argument can be used for the outside option being selected where one considers $\bar{V}_{in}$ as effectively equal to ∞.

The second part of the theorem is a fairly simple result given the first part. Assume $\bar{V}_{in} > V^*_{max}$. If $V_{in} > V^*_{max}$, then $V^*_{in} > V^*_{max}$ so option i must be searched because it is selected. Alternatively, $V_{in} < V^*_{max}$ and $\bar{V}_{in} > V^*_{max}$ implies that $\bar{V}_{in}$ is either greater than the value of the selected option or the search threshold for the selected option. Thus, either option i is searched prior to when the other option is selected or option i is searched after the other option, but before search terminates. Figure 1 illustrates when option i is searched and when option i is selected for a fixed V^*_{max}.

To simplify the equations in the following lemma, let $\delta_{in} = \alpha_{in} + min\{\mu_{in}, \epsilon_{in}\}$ where F^{δ}_i denotes the resulting CDF and let $\gamma_{in} = \alpha_{in} + \mu_{in}$ where F^{γ}_i denotes the resulting CDF.

Lemma 3. *Expected payoffs, expected demands and expected search volumes are given by* *the following integrals:*

$$E[P] = \sum_{i=0}^{I} \int_{-\infty}^{\infty} v \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\delta}(v - x_i) dv$$

$$E[D_i] = N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\delta}(v - x_i) dv$$

$$E[S_i] = N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\gamma}(v - x_i) dv$$

From the definition of D_i, it follows that

$$E[D_i] = \sum_{n=1}^{N} \ Prob \left[V_{in}^* > \underset{j \neq i}{max} \ V_{jn}^* \right]$$

$$= N \ Prob \left[min\{\bar{V}_{in}, V_{in}\} > \underset{j \neq i}{max} \ \left[min\{\bar{V}_{jn}, V_{jn}\} \right] \right]$$

$$= N \ Prob \left[x_i + \alpha_{in} + min\{m(\kappa_{in}), \epsilon_{in}\} > \underset{j \neq i}{max} \ \left[x_j + \alpha_{jn} + min\{m_i(\kappa_{jn}), \epsilon_{jn}\} \right] \right]$$

$$= N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\delta}(v - x_i) dv$$

Similarly,

$$E[S_i] = N \ Prob \left[\bar{V}_{in} > \underset{j \neq i}{max} \ V_{jn}^* \right]$$

$$= N \ Prob \left[x_i + \alpha_{in} + m(\kappa_{in}) > \underset{j \neq i}{max} \ \left[x_j + \alpha_{jn} + min\{m_i(\kappa_{jn}), \epsilon_{jn}\} \right] \right]$$

$$= N \int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(v - x_j) f_i^{\gamma}(v - x_i) dv$$

Recall that P_n is the payoff of agent n in this model which includes all search costs that are incurred as well as the value of the selected option. Taking the expectation over all possible realizations of agent heterogeneity,

$$E[P_n] = \mathbf{V}_{0n} + \sum_{i=1}^{I} [\mathbf{V}_{in} - \mathbf{K}_{in}]$$

where

$$\mathbf{V}_{0n} = E[V_{0n}\ 1[V_{0n} > \underset{j\neq 0}{max}\ V^*_{jn}]]$$

$$\mathbf{V}_{in} = E[V_{in}\ 1[V^*_{in} > \underset{j\neq i}{max}\ V^*_{jn}]]$$

$$\mathbf{K}_{in} = E[\kappa_{in}\ 1[\bar{V}_{in} > \underset{j\neq i}{max}\ V^*_{jn}]]$$

The indicators inside each expectation weight match values that are not chosen with a value of 0 and search costs that are not paid with a value of 0. In essence, $\mathbf{V}_{0n}$ and $\mathbf{V}_{in} - \mathbf{K}_{in}$ are the expected gains from each option for a given agent n when searching optimally.

I write expectations that do not vary with n without the subscript. However, I will write V_{in}, V^*_{in} and $\bar{V}_{in}$ with n where agent n is effectively the representative agent.

$$\sum_{i=0}^{I} \mathbf{V}^*_i = \sum_{i=0}^{I} \int_{-\infty}^{\infty} v \prod_{j\neq i} F^{\delta}_j(v - x_j) f^{\delta}_i(v - x_i) dv.$$

For $i = 1, ..., I$

$$\mathbf{V}_i - \mathbf{V}^*_i$$

$$= E[V_{in} * 1[V_{in}^* > \max_{j \neq i} V_{jn}^*]] - E[V_{in}^* * 1[V_{in}^* > \max_{j \neq i} V_{jn}^*]]$$

$$= E[(V_{in} - \min\{V_{in}, \bar{V}_{in}\})1[V_{in}^* > \max_{j \neq i} V_{jn}^*]]$$

$$= E[(V_{in} - \bar{V}_{in})1[V_{in} > \bar{V}_{in}]1[V_{in}^* > \max_{j \neq i} V_{jn}^*]]$$

$$= E[(V_{in} - \bar{V}_{in})1[V_{in} > \bar{V}_{in}]1[\bar{V}_{in} > \max_{j \neq i} V_{jn}^*]]$$

$$= E[E[(V_{in} - \bar{V}_{in})1[V_{in} > \bar{V}_{in}]|\boldsymbol{\kappa}, \boldsymbol{\alpha}]1[\bar{V}_{in} > \max_{j \neq i} V_{jn}^*]]$$

$$= E[\kappa_{in}1[\bar{V}_{in} > \max_{j \neq i} V_{jn}^*]] = \mathbf{K}_i$$

While most of the steps above are easy to follow, it is important to note that $\kappa_{in} = E[(V_i - \bar{V}_i)1[V_i > \bar{V}_i]|\alpha_{in}, \kappa_{in}]$ is a restatement of the optimal search rule. The other steps use the properties of indicator functions, the min function and the law of iterated expectations to rewrite the difference between these two values. There is a different, but related proof of the expected payoff formula in the appendix of Armstrong (2017).

Corollary 1. *Given a specific F_i^ϵ, each F_i^μ with a support that is bounded above by $\bar{\epsilon}_i$ corresponds to a unique F_i^κ and each F_i^κ corresponds to a unique F_i^μ with a support that is bounded above by $\bar{\epsilon}_i$ where all CDFs are continuous and differentiable over a convex support. As a result, search models can be specified with $(F_0, [F_i^\alpha, F_i^\epsilon, F_i^\mu]_1^I)$ instead of $(F_0, [F_i^\alpha, F_i^\epsilon, F_i^\kappa]_1^I)$ without imposing extra restrictions or changing the set of possible search models.*

This corollary follows directly from the fact that m_i is a bijection from across the supports of the two distributions for any fixed i. Lemma 1 basically proves this, but I separate the results to introduce this property after the formal outline of a rational agent behavior.

Lemma 4. *Consider a GSM* $(\mathbf{x}, F_0, [F_i^{\alpha}, F_i^{*}, a_i]_1^I)$ *where* $E[D_i] > 0 \;\; \forall i$. *For each option i, let* $\bar{D}_i = E[D_i|\epsilon_{in} = \bar{\epsilon}_i]$. *The following statements hold true:*

1) The ex-ante probability that an agent follows each possible search path is strictly greater than 0.

2) For any option i and j, $\frac{\partial E[D_i]}{\partial a_j} = 0$. *For any option j where* $j \neq i$, $\frac{\partial E[S_i]}{\partial a_j} = 0$.

3) For any option i, $E[S_i]$ *as a function of* a_i *is a strictly increasing bijection from* $(0, 1)$ *to* $(E[D_i], \bar{D}_i)$.

4) For any I-dimensional vector $\mathbf{s}$ *where* $E[\mathbf{D}] < \mathbf{s} < \bar{\mathbf{D}}_i$, *there exists a unique vector of* $\mathbf{a}$ *values such that* $E[\mathbf{S}] = \mathbf{s}$.

I split this into parts to better address each statement.

1) Let ($\mathbf{I}^*$, $\{i_k\}$, i^*) be an arbitrary search path where $i^* \in \mathbf{I}^*$. A consumer follows this search path if and only if

i. $\forall j \notin \mathbf{I}^*, \bar{V}_{jn} < V^*_{i^*n}$

ii. $\bar{V}_{i_k n}$ is decreasing in k.

iii. $\forall j \neq i^*, V^*_{i^*n} > V^*_{jn}$

Recall that V_{in} and $\bar{V}_{in}$ are independent, continuous random variables with the same support. Since $E[D_i] > 0$ for any firm i, there must exist an interval $(\underline{v}, \bar{v})$ where $\underline{v} < \bar{v}$ that is a subset of the support of every V^*_{in} for each firm i (and the outside option). By the work earlier, we know that for each firm i, V_{in} and $\bar{V}_{in}$ have the same support as V^*_{in}.

Let ($\mathbf{I}^*$, $\{i_k\}$, i^*) be an arbitrary search path. Let's break the interval into $|\mathbf{I}^*| + 2$ smaller sub-interval of equal size where the highest sub-interval is considered the first sub-interval and the lowest sub-interval in considered the $|\mathbf{I}^*| + 2$th interval. For any $k \in \{1, ..., |\mathbf{I}^*|\}$, the probability $\bar{V}_{i_k n}$ is in the kth sub-interval is strictly greater than 0. For any $j \notin \mathbf{I}^*$, the probability that $\bar{V}_{jn}$ is in the $|\mathbf{I}^*| + 2$th sub-interval is strictly greater than 0. Similarly, for $j \neq i^*$, the probability that V_{jn} is in the $|\mathbf{I}^*| + 2$th sub-interval is strictly greater than 0. Finally, the probability that V_{in} is in the $|\mathbf{I}^*| + 1$th sub-interval is strictly greater than 0. By construction, if all these events occur at once, the conditions necessary for the search path defined by ($\mathbf{I}^*$, $\{i_k\}$, i^*) are all met. Since all of these events are independent, the probability they all occur at once is strictly greater than 0. Thus, the search path must occur with probability strictly greater than 0.

2) This follows from the fact that F_i^{δ} does not vary with a_j for any option j and F_i^{γ} does not vary with a_j for $j \neq i$.

3) The limit of F_i^{μ} as $a_i \to 0$ is F_i^{min}. Thus, the limit of $E[S_i]$ as $a_i \to 0$ is $E[D_i]$.

$\bar{F}_i^{\mu} = (\bar{F}_i^{*})^{1-a_i}$ is increasing in a_i so

$$E[S_i] = N\ Prob\left[\bar{V}_{in} > \max_{j \neq i} V^*_{jn}\right]$$

is increasing in a_i. Finally, as $a_i \to 1$, the distribution of μ_{in} limits to a mass point at $\bar{\mu}_i$ which is also the same value as $\bar{\epsilon}_i$, so $E[S_i]$ limits to $E[D_i|\epsilon_{in} = \bar{\epsilon}]$.

4) This statement follows directly from 2 and 3.

Theorem 1.

1) For any search model $(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$, there exists a CDM $(F_0, [F_i^{\beta}]_1^I)$ that generates the same expected demands and expected payoffs for all $\mathbf{x}$.

2) For any CDM $(F_0, [F_i^{\beta}]_1^I)$, there exists a search model $(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$ that generates the same expected demands and expected payoffs for all $\mathbf{x}$.[41]

3) For any CDM $(F_0, [F_i^{\beta}]_1^I)$ and I dimensional function $\mathbf{s}(\mathbf{x})$ where F_i^{β} has support that is unbounded above and $E[\mathbf{D}] < \mathbf{s}(\mathbf{x}) < \bar{\mathbf{D}}$ for all $\mathbf{x}$, there exists a GSM $(F_0, [F_i^{\alpha}, F_i^{*}, a_i(\mathbf{x})]_1^I)$ that generates the same expected demands, expected payoffs and has the property that $E[\mathbf{S}] = \mathbf{s}(\mathbf{x})$ for all $\mathbf{x}$.

4) For any search model $(F_0, [F_i^{\alpha}, F_i^{\epsilon}, F_i^{\mu}]_1^I)$, there exists a GSM $(F_0, [F_i^{\alpha}, F_i^{*}, a_i]_1^I)$ that generates the same expected demands, expected payoffs and expected search volumes for all $\mathbf{x}$.

[41] There is a continuum of models that match any specified CDM because the proof does not require a specific value of **a** in the GSM.

Explained in the main body of the dissertation. (1) is from Armstrong (2017). 2 follows from the GSM which can match any CDM. 3 and 4 follow from lemma 4 part 3.

Lemma 5. *Suppose that F_0 has a support that is not bounded above. Also suppose that $\forall i$, F_i^{ϵ} and F_i^{μ} have supports that are not bounded above.*[42] *Given two different options $i \neq j$, the following properties hold.*

1) If $\sigma_i^*, \sigma_j^* \in (0,1)$, $E[D_i]$ and $E[S_i]$ are strictly increasing in both x_i and $\hat{x}_i$ and strictly decreasing in both x_j and $\hat{x}_j$.

Suppose $\hat{x}_k = x_k \quad \forall k \neq i$.

2) If $\hat{x}_i < x_i$, $E[D_i]$ and $E[S_i]$ are strictly decreasing in σ_i^* and $E[D_j]$ and $E[S_j]$ are strictly increasing in σ_i^*.

3) If $\hat{x}_i > x_i$, $E[D_i]$ and $E[S_i]$ are strictly increasing in σ_i^* and $E[D_j]$ and $E[S_j]$ are strictly decreasing in σ_i^*.

4) $\frac{\partial^2 E[D_i]}{\partial \sigma_i^* \partial x_i} < 0$ and $\frac{\partial^2 E[S_i]}{\partial \sigma_i^* \partial x_i} < 0$

5) If $\hat{x}_i = x_i$, σ_i^* does not affect the choices of agents.

6) Conditional on σ_i^*, no feature of the distribution of σ values across agent-option pairs affects selection or search probabilities.

Suppose that F_0 has a support that is not bounded above. Also suppose that $\forall i$, F_i^{ϵ} and F_i^{μ} have supports that are not bounded above. This removes the possibility of products with $E[D_i] = 0$.

[42] Without this assumption or a similar support assumption, the conclusions of the lemma would switch from strictly increasing/decreasing to weakly increasing/decreasing.

For the moment, consider x_i and x_{in}^e to be unrelated values. Recall that $V_{in}^* = min\{x_i + \alpha_{in} + \epsilon_{in}, x_{in}^e + \alpha_{in} + \mu_{in}\}$. The CDF of V_{in}^* is weakly decreasing in both x_i and x_{in}^e and strictly decreasing in both x_i and x_{in}^e for sufficiently high values of V_{in}^* due to the unbounded supports of ϵ_{in} and μ_{in}.This implies that $E[D_i] = \sum_{n=1}^N \ Prob\left[V_{in}^* > \underset{j \neq i}{max}\ V_{jn}^*\right]$ is strictly increasing in x_i and ${x_{in}}^e$. Similarly, $E[D_j]$ and $E[S_j]$ are strictly decreasing in x_i and x_{in}^e. $\text{E}[S_i] = \sum_{n=1}^N \ Prob\left[\bar{V}_{in} > \underset{j \neq i}{max}\ V_{jn}^*\right]$ is strictly increasing in x_{in}^e and constant in terms of x_i.

With these properties established, now I consider the relationship between x_i and x_{in}^e as well as the relationship between $\hat{x}_i$ and x_{in}^e. (1) follows from the fact that x_{in}^e is strictly increasing in x_i for $\sigma_{in} = 0$ and strictly increasing in $\hat{x}_i$ for $\sigma_{in} = 1$. (2) follows from the property that $\hat{x}_i < x_i$ implies that x_{in}^e is higher for transparent pairs. (3) follows from the property that $\hat{x}_i > x_i$ implies that x_{in}^e is higher for opaque pairs. (4) follows from the property that $\frac{\partial x_{in}^e}{\partial x_i} = 1$ for transparent pairs and $\frac{\partial x_{in}^e}{\partial x_i} = 0$ for opaque pairs. (5) holds because $\hat{x}_i = x_i$ implies that x_{in}^e is the same for transparent and opaque pairs. (6) follows from the same argument, but with all options $j \neq i$

Lemma 6. *In the GSM, suppose that $\hat{x}_i = x_i$ for all options and that $i \neq j$.*

7) $\frac{\partial E[D_i]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[D_i^{trans}]}{\partial x_i}$

8) $\frac{\partial E[S_i]}{\partial x_i} = (1 - \sigma_i^*)\frac{\partial E[S_i^{trans}]}{\partial x_i}$

9) $\frac{\partial E[D_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[D_j^{trans}]}{\partial x_i}$

10) $\frac{\partial E[S_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial E[S_j^{trans}]}{\partial x_i}$

11) $(1-\sigma_j^*+\sigma_j^* a_j^*)\frac{\partial E[D_j]}{\partial x_i} = (1-\sigma_i^*+\sigma_i^* a_i^*)\frac{\partial E[D_i]}{\partial x_j}$

Note: partial derivatives are taken only with respect to x_i or x_j, not $\hat{x}_j$ and $\hat{x}_j$.

Let i be an arbitrary firm. For opaque pairs with firm i:

$$V_{in}^* - \alpha_{in} \sim F^{op}(v) = 1-(1-F_i^*(v-x_i))^{a_i}(1-F_i^*(v-\hat{x}_i))^{b_i}$$

$$= 1-((1-F_i^*(v-x_i))^{\frac{1}{a_i+b_i}})^{a_i^*}((1-F_i^*(v-\hat{x}_i))^{\frac{1}{a_i+b_i}})^{1-a_i^*}$$

For transparent pairs with firm i:

$$V_{in}^* - \alpha_{in} \sim F^{trans}(v) = 1-(1-F_i^*(v-x_i))^{a_i}(1-F_i^*(v-x_i))^{b_i}$$

$$= 1-((1-F_i^*(v-x_i))^{a_i+b_i}$$

$\frac{\partial F^{trans}(v)}{\partial x_i} = (a_i+b_i)f_i^*(v-x_i)(1-F_i^*(v-x_i))^{a_i+b_i-1}$

Thus,

$\frac{\partial F^{op}(v)}{\partial x_i} = a_i * f_i^*(v-x_i)(1-F_i^*(v-x_i))^{a_i-1}(1-F_i^*(v-x_i))^{b_i}$

When evaluated at $x_i = \hat{x}_i$, $\frac{\partial F^{op}(v)}{\partial x_i} = a_i * f_i^*(v-x_i)(1-F_i^*(v-x_i))^{a_i+b_i-1} = \frac{a_i}{a_i+b_i}\frac{\partial F^{trans}(v)}{\partial x_i}$

Let $j \neq i$. Since $\frac{\partial F^{op}(v)}{\partial x_i} = a_i^* \frac{\partial F^{trans}(v)}{\partial x_i}$ and $E[D_i]$, $E[D_j]$,$E[S_j]$ depend on the distribution of V_{in}^*,

7) $\frac{\partial E[D_i]}{\partial x_i} = (1-\sigma_i^*+\sigma_i^* a_i^*)\frac{\partial E[D_i^{trans}]}{\partial x_i}$

9) $\frac{\partial E[D_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*) \frac{\partial E[D_j^{trans}]}{\partial x_i}$

10) $\frac{\partial E[S_j]}{\partial x_i} = (1 - \sigma_i^* + \sigma_i^* a_i^*) \frac{\partial E[S_j^{trans}]}{\partial x_i}$

11 follows from 9.

8 does not rely on GSM. It follows from the property that $\frac{\partial x_{in}^e}{\partial x_i} = 1$ for transparent pairs and $\frac{\partial x_{in}^e}{\partial x_i} = 0$ for opaque pairs because $E[S_i]$ depends on the distribution of $\bar{V}_{in}$.

Corollary 2. *Suppose consumers correctly anticipate the vector of x values ($\hat{\mathbf{x}} = \mathbf{x}$).*

1) Suppose firm m owns exactly one product ($I_m = \{i\}$). It follows that $\frac{\partial^2 E[\pi_m]}{\partial \sigma_i^* \partial x_i} < 0$ and

$$\frac{\partial E[\pi_m | \sigma_{in} = 1 \ \forall n]}{\partial x_i} \leq \frac{\partial E[\pi_m]}{\partial x_i} \leq \frac{\partial E[\pi_m | \sigma_{in} = 0 \ \forall n]}{\partial x_i}$$

2) Suppose $\frac{\partial E[\pi_m]}{\partial x_i} - g_1^i > 0$ and $g_3^i = 0$. It follows that $\frac{\partial^2 E[\pi_m]}{\partial \sigma_i^* \partial x_i} < 0$ and

$$\frac{\partial E[\pi_m | \sigma_{in} = 1 \ \forall n]}{\partial x_i} \leq \frac{\partial E[\pi_m]}{\partial x_i} \leq \frac{\partial E[\pi_m | \sigma_{in} = 0 \ \forall n]}{\partial x_i}$$

Note: As before, the partial derivatives are taken with respect to x_i holding $\hat{x}_i$ constant.

Recall that $E[\Pi_m] = \sum_{j \in I_m} g^j(x_j, E[D_j], E[S_j])$. Thus,

$$\frac{\partial E[\Pi_m]}{\partial x_i} = \sum_{j \in I_m} \left[g_1^j(x_j, E[D_j], E[S_j]) + \frac{\partial E[D_j]}{\partial x_i} g_2^j(x_j, E[D_j], E[S_j]) + \frac{\partial E[S_j]}{\partial x_i} g_3^j(x_j, E[D_j], E[S_j]) \right]$$

Each $g_1^j(x_j, E[D_j], E[S_j])$, $g_2^j(x_j, E[D_j], E[S_j])$, $g_3^j(x_j, E[D_j], E[S_j])$ does not depend on σ_i^* and $g_2^j(x_j, E[D_j], E[S_j]) > 0$, $g_3^j(x_j, E[D_j], E[S_j]) \geq 0$. For a single product firm, $E[\Pi_m] = g^i(x_i, E[D_i], E[S_i])$.Thus, (1) follows from $\frac{\partial^2 E[D_i]}{\partial \sigma_i^* \partial x_i} < 0$ and $\frac{\partial^2 E[S_i]}{\partial \sigma_i^* \partial x_i} < 0$ which implies $\frac{\partial^2 E[\Pi_m]}{\partial \sigma_i^* \partial x_i} < 0$ because $g_2^i > 0, g_3^i \geq 0$ for a single product firm. For a single product firm, $E[\Pi_m] = g^i(x_i, E[D_i], E[S_i])$.

For a GSM where $g_3^i = 0$,

$$\frac{\partial E[\Pi_m]}{\partial x_i} = \sum_{j \in I_m} \left[g_1^j(x_j, E[D_j], E[S_j]) + \frac{\partial E[D_j]}{\partial x_i} g_2^j(x_j, E[D_j], E[S_j]) + \frac{\partial E[S_j]}{\partial x_i} g_3^j(x_j, E[D_j], E[S_j]) \right]$$

$$= (1 - \sigma_i^* + \sigma_i^* a_i^*) \sum_{j \in I_m} \left[g_1^j(x_j, E[D_j], E[S_j]) + \frac{\partial E[D_j^{trans}]}{\partial x_i} g_2^j(x_j, E[D_j], E[S_j]) + \frac{\partial E[S_j^{trans}]}{\partial x_i} g_3^j(x_j, E[D_j], E[S_j]) \right]$$

$$\Rightarrow$$

$$\frac{\partial^2 E[\Pi_m]}{\partial \sigma_i^* \partial x_i} < 0$$

Theorem 2. *Assume for each firm i that $h_i = h_i^*$ where h_i^* is log-concave for $x_i < \bar{x}_i$. Also, assume that search is defined by a GSM where f_i^* and f_i^α are log-concave with a support that is not bounded above for each firm i and f_0 is log-concave with a support that is not bounded above. Furthermore, assume that $a_i + b_i \geq 1 \quad \forall i$.*

If there exists a pure strategy equilibrium $\mathbf{x}^*$ in this partially transparent market, then it is unique and $\mathbf{x}^*$ is also the unique equilibrium vector in a fully transparent model with the same GSM where $h_i = (h_i^*)^{\frac{1}{1-\sigma_i^*+\sigma_i^* a_i^*}}$ for each firm i.

Assuming an equilibrium $\mathbf{x}^*$ exists:

1) For any i and j, x_i^* is strictly decreasing in σ_j^* and strictly increasing in a_j^* for $\sigma_i^* > 0$.

2) Expected consumer payoffs are strictly decreasing in σ_j^* and strictly increasing in a_j^* for $\sigma_i^* > 0$.

3) For $j \neq i$, expected profits for firm i are strictly increasing in σ_j^* and strictly decreasing in a_j^* for $\sigma_i^* > 0$.

Additionally, suppose firms are playing the pricing game.

4) p_i^* is strictly increasing in c_j and strictly decreasing in q_j for any firm $j \neq i$.

5) p_i^* is strictly increasing in c_i and strictly increasing in q_i.

6) $x_i^* = q_i - p_i^*$ is strictly decreasing in c_j and strictly increasing in q_j for any firm j.

7) Total Welfare (the sum of all expected firm profits and expected consumer payoffs) is increasing in a_j^* and decreasing in σ_j^*.

For the GSM with single product firms,

$$\frac{\partial ln(E[\pi_i])}{\partial x_i} = \frac{h^{*\prime}(x_i)}{h^*(x_i)} + (1 - \sigma_i^* + \sigma_i^* a_i^*) \frac{\partial ln(E[D_i^{trans}])}{\partial x_i}$$

when evaluated at $\mathbf{x} = \hat{\mathbf{x}}$. In comparison, in the fully transparent case where $h_i = (h_i^*)^{\frac{1}{1-\sigma_i^*+\sigma_i^* a_i^*}}$,

$$\frac{\partial ln(E[\pi_i])}{\partial x_i} = \frac{h^{*\prime}(x_i)}{h^*(x_i)(1 - \sigma_i^* + \sigma_i^* a_i)} + \frac{\partial ln(E[D_i^{trans}])}{\partial x_i}$$

Since h_i^* is log-concave, $(h_i^*)^{\frac{1}{1-\sigma_i^*+\sigma_i^* a_i^*}}$ is log-concave. Adapting the math in Quint (2014) in terms of $x_i = -p_i$ and $h_i(x_i)$ instead of $p_i - c_i$ for each firm i:

f_i^δ log-concave with a support that is unbounded above and $h_i(x_i)$ log-concave

$\forall i$ implies that a unique equilibrium exists for the fully transparent market. Moreover, under these conditions, $\frac{\partial^2 ln(E[D_i^{trans}])}{\partial x_i \partial x_i} < 0$, $\frac{\partial^2 ln(E[D_i^{trans}])}{\partial x_i \partial x_j} < 0$ for $j \neq i$. As a result, the choices of firms are strategic compliments (the best respond to increases in x_j values is to increase x_i).

Recall that $F_i^{min} = (1 - F_i^*)^{a_i + b_i}$ so $f_i^{min} = (a_i + b_i) f_i^* (1 - F_i^*)^{a_i + b_i - 1}$. Since f_i^* is log-concave, F_i^* and $\bar{F}_i^*$ are log-concave. By extension, $f_i^{min} = (a_i + b_i) f_i^* (1 - F_i^*)^{a_i + b_i - 1}$ is log-concave because $a_i + b_i - 1 \geq 0$. Since the sum of random variables with log-concave densities also has a log-concave density, f_i^δ is log-concave and unbounded above. Thus, the log-concave condition from Quint (2014) is met.

Since the model without imposing fully transparency shares first-order conditions with a fully transparent market with a unique equilibrium, the original model has a unique equilibrium if an equilibrium exists.

Assuming an equilibrium $\mathbf{x}^*$ exists:

1) For any i and j, x_i^* is strictly decreasing in σ_j^* and strictly increasing in a_j^* for $\sigma_i^* > 0$. This property holds because decreasing σ_j^* or increasing a_j^* for $\sigma_i^* > 0$ results in a higher value of $\frac{\partial ln(E[\pi_i])}{\partial x_i}$. Combined with strategic complements, x_i^* will increase as will x_j^* for $j \neq i$. Property 2 follows from the general property in discrete choice that higher mean values (V^* in search) result in higher consumer payoffs. (3) follows from the fact that x_i decreases result in increased profits of all other firms holding all else equal so a firm cannot be worse off (assuming rational behavior) if the x^* values of all other firms decrease.

(4), (5), (6) and (7) are common properties in pricing games with strategic compliments. These properties hold in this model because the math from CDMs holds in the equivalent fully transparent models.

Lemma 7. *Assume for each firm i that $h_i = h_i^*$ where h_i^* is log-concave for $x_i < \bar{x}_i$. Also, assume that search is defined by a GSM where f_i^* and f_i^α are log-concave with a support that is unbounded above for each firm i and f_0 is log-concave with a support that is unbounded above. Furthermore, assume that $a_i + b_i \geq 1 \quad \forall i$.*

Any of the following is a sufficient condition for the existence of the unique pure strategy equilibrium:

(i) The market is fully transparent ($\sigma_i^* = 0 \ \forall i$).

(ii) All options are either fully transparent or fully opaque ($\sigma_i^* \in \{0,1\} \quad \forall i$) and $a_i, b_i \geq 1 \ \forall i$.

(iii) All options are either fully transparent or fully opaque ($\sigma_i^* \in \{0,1\} \ \forall i$) and $\frac{f_i^*}{1-F_i^*}$ is log-concave $\quad \forall i$.

(i) is a sufficient condition as shown for the previous theorem. The previous work demonstrates that a unique candidate equilibrium exists where FOCs hold for all firms. I demonstrate that with (ii) or (iii), second order conditions hold so the candidate equilibrium is an equilibrium. For fully transparent firms, SOCs hold because $a_i + b_i \geq 1 \quad \forall i$.

Suppose correct anticipation for all options that are not option i. If $\sigma_i^* = 1$,

$$E[D_i] = \sum_{n=1}^{N} Prob\left[V_{in}^* > \max_{j \neq i} V_{jn}^*\right]$$

$$= N * \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \prod_{j \neq i} F_j^{\delta}(min\{\bar{v}, v\} - x_j) f_i^{\mu}(\bar{v} - \hat{x}_i - \alpha) f_i^{\epsilon}(v - x_i - \alpha) f_i^{\alpha}(\alpha) d\bar{v} dv d\alpha$$

If F_j^{δ} for $j \neq i$, f_i^{μ}, f_i^{ϵ}, f_i^{α} are log-concave, then $\prod_{j \neq i} F_j^{\delta}(min\{\bar{v}, v\} - x_j) f_i^{\mu}(\bar{v} - \hat{x}_i - \alpha) f_i^{\epsilon}(v - x_i - \alpha) f_i^{\alpha}(\alpha)$ is log-concave in $\bar{v}, v, \alpha, x_i, \hat{x}_i$ because $min\{\bar{v}, v\}$ is a concave function and log-concavity is preserved over multiplication. Since log-concavity is also preserved over integration, $E[D_i]$ is log-concave in x_i for any values of $\hat{x}_i$. Thus, for fully opaque firms, $ln(E[\pi_i]$ is concave in x_i. Thus, second order conditions hold for firm i. Recall that

$$f_i^{\epsilon}(v) = a_i f_i^*(v)(1 - F_i^*(v))^{a_i - 1} \quad \text{and} \quad f_i^{\mu}(v) = b_i f_i^*(v)(1 - F_i^*(v))^{b_i - 1}$$

If $a_i, b_i \geq 1$ then $a_i - 1, b_i - 1 \geq 0$ so both densities are log-concave. If $\frac{f_i^*}{1-F_i^*}$ is a log-concave, then each density is log-concave without making assumptions on a_i and b_i. Thus, either (ii) or (iii) is sufficient for the existence of a unique equilibrium. Note that if $\frac{f_i^*}{1-F_i^*}$ holds for all firms, a unique equilibrium exists even if $a_i + b_i < 1$. $\frac{f_i^*}{1-F_i^*}$ holds for many densities, most notably, TIEV and reverse TIEV.

Lemma 8. *Under the same assumptions as in theorem* 3*, suppose also that either condition* (ii) *or* (iii) *holds. Suppose* i *is an option where* $\sigma_i^* = 1$*. Also assume that a market* *equilibrium exists at all relevant parameter values.*

1) $\lim_{a_i^* \to 0} x_i^* = -\infty$ *and* $\lim_{a_i^* \to 0} x_j^* = x_j^{-i}$ *where* x_j^{-i} *is the equilibrium choice of firm j if option i is removed from the market.*

2) In a pricing game, $\lim_{a_i^* \to 0} p_i^* = \infty$ *and* $\lim_{a_i^* \to 0} p_j^* = p_j^{-i}$ *where* p_j^{-i} *is the equilibrium price of firm j if option i is removed from the market.*

For $\mathbf{x} = \hat{\mathbf{x}}$,

$$\lim_{a_i^* \to 0} \frac{\partial ln(E[\pi_i])}{\partial x_i} = \frac{h^{*\prime}(x_i)}{h^*(x_i)} < 0$$

Thus, $\lim_{a_i^* \to 0} x_i^* = -\infty$ and $\lim_{a_i^* \to 0} p_i^* = \infty$. The property for all other firms follows from this because V_{in}^* is limiting to a mass point at $-\infty$.

Lemma 9. *The following specifications are all equivalent given* $(\mathbf{x}, \mathbf{y}, \mathbf{A}, F_0, [F_i^\alpha]_{i=1}^I)$.

1) Search defined with threshold heterogeneity (μ):

$(F_0, [F_i^\alpha, F_i^\epsilon, F_i^\mu]_1^I)$ *where* $F_i^\epsilon(v) = 1 - e^{-e^{A_i v}}$ *and* $F_i^\mu(v) = 1 - e^{-e^{A_i(v-y_i)}}$ *for* $i = 1, ..., I$.

2) Search defined with search cost heterogeneity (κ):

$(F_0, [F_i^\alpha, F_i^\epsilon, F_i^\kappa]_1^I)$ *where* $F_i^\epsilon(v) = 1 - e^{-e^{A_i v}}$ *and* $F_i^\kappa(v) = e^{-e^{A_i(m_i(v)-y_i)}}$ *for* $i = 1, ..., I$.

3) Geometric Search:

$(F_0, [F_i^\alpha, F_i^*, a_i, b_i]_1^I)$ *where* $a_i = 1$, $b_i = e^{-A_i y_i}$ *and* $F_i^*(v) = 1 - e^{-e^{A_i v}}$ *for* $i = 1, ..., I$.

4) Geometric Search where $b_i = 1 - a_i$ *for each i:*

$(F_0, [F_i^\alpha, F_i^*, a_i]_1^I)$ *where* $a_i = \frac{1}{1+e^{-A_i y_i}}$ *and* $F_i^*(v) = 1 - e^{-e^{A_i(v - \frac{1}{A_i} g(A_i y_i))}}$ *for* $i = 1, ..., I$.

(1) is the initial specification. (2) is from the connection between F_i^κ and F_i^μ

because $F_i^\kappa(v) = e^{-e^{A_i(m_i(v)-y_i)}}$ implies that $m_i(\kappa)$ has CDF $F_i^\mu(v) = 1 - e^{-e^{A_i(v-y_i)}}$. (3) follows from $F_i^\epsilon(v) = 1 - (e^{-e^{A_i v}})^1$ and $F_i^\mu(v) = 1 - (e^{-e^{A_i v}})^{e^{-A_i y_i}}$. Recall $g(A_i y_i) = -ln(1 + e^{-A_i y_i})$. (4) follows from this because $F_i^\epsilon(v) = 1 - (e^{-e^{A_i v}(1+e^{-A_i y_i})})^{\frac{1}{1+e^{-A_i y_i}}} = 1 - (e^{-e^{A_i v}})$ and $F_i^\mu(v) = 1 - (e^{-e^{A_i v}(1+e^{-A_i y_i})})^{\frac{e^{-A_i y_i}}{1+e^{-A_i y_i}}} = 1 - e^{-e^{A_i(v-y_i)}}$.

Corollary 3. *Assume the market structure outlined in the previous section, but with search defined by a REVSM. Assume for each firm i that $h_i = h_i^*$ where h_i^* is log-concave for $x_i < \bar{x}_i$. Also, assume that f_i^* and f_i^α are log-concave with a support that is not bounded above for each firm i and f_0 is log-concave with a support that is not bounded above.*

All of the results of Theorem 3 hold in this market.

1) $z_i^* = x_i^* + \frac{1}{A_i} g(A_i y_i)$ is strictly increasing in y_j for any firm j.

2) For $j \neq i$, x_i^* is strictly increasing in y_j. In the pricing game, p_i^* is strictly decreasing in y_j.

3) If firm i is fully transparent, x_i^* is strictly decreasing in y_i. In the pricing game, p_i^* is strictly increasing in y_i.

4) If firm i is partially or fully opaque, x_i^* may be increasing or decreasing in y_i. In the pricing game, p_i^* may be increasing or decreasing in y_i.

5) A unique equilibrium exists if options in the market are either transparent or fully opaque ($\sigma_i^* \in \{0, 1\} \;\; \forall i$).

(1) Reframe the problem as firms selecting $z_i = x_i + \frac{1}{A_i} g(A_i y_i)$ where

$$\frac{\partial ln(E[\pi_i])}{\partial x_i} = \frac{h^{*\prime}(z_i - \frac{1}{A_i} g(A_i y_i))}{h^*(z_i - \frac{1}{A_i} g(A_i y_i))} + (1 - \sigma_i^* + \sigma_i^* a_i^*) \frac{\partial ln(E[D_i^{trans}])}{\partial x_i}$$

Since h_i is log-concave, $\frac{h^{*\prime}(z_i - \frac{1}{A_i}g(A_i y_i))}{h^*(z_i - \frac{1}{A_i}g(A_i y_i))}$ is increasing in y_i. $(1 - \sigma_i^* + \sigma_i^* a_i^*)\frac{\partial ln(E[D_i^{trans}])}{\partial x_i}$ is also increasing in y_i because $\sigma_i^* a_i^*$ is increasing in y_i and $\frac{\partial ln(E[D_i^{trans}])}{\partial x_i} > 0$.

(2) follows from the strategic compliments property where other firm j$\neq$ i reacts to a higher z_i with higher z_j values. Since y_j doesn't change, z_j increasing corresponds to a higher x_j (p_j in the pricing game).

(3) In the fully transparent setting,

$$\frac{\partial ln(E[\pi_i])}{\partial x_i} = \frac{h^{*\prime}(z_i - \frac{1}{A_i}g(A_i y_i))}{h^*(z_i - \frac{1}{A_i}g(A_i y_i))} + \frac{\partial ln(E[D_i^{trans}])}{\partial x_i}$$

Since $\frac{\partial^2 ln(E[D_i^{trans}])}{\partial x_i \partial x_i} < 0$ and increasing y_i increases the effective mean value of the product z_i, $\frac{\partial^2 ln(E[D_i^{trans}])}{\partial x_i \partial y_i} < 0$. A similar argument holds for prices.

(4) is proved in the main body of the dissertation with the monopolist example. I include the example here as well with some extra work:

Suppose $I = 1$, $F_0(v) = 1 - e^{-e^v}$, $\alpha_{1n} = 0 \ \ \forall n$ and $A_1 = 1$. In this model,

$$E[D_1^{trans}] = N\int_{-\infty}^{\infty}(1 - e^{-e^v})e^{-e^{v-x_1-g(y_1)}+v-x_1-g(y_1)} = 1 - \frac{e^{-x_1-g(y_1)}}{1+e^{-x_1-g(y_1)}}.$$

From the lemma 6, at points where $\hat{x}_i = x_i$

$$\frac{\partial ln(E[\pi_1])}{\partial x_i} = \frac{h_1'(x_1)}{h_1(x_1)} + (1 - \sigma_1^* + \sigma_1^* a_1^*)\frac{\partial ln(E[D_1^{trans}])}{\partial x_i}$$

$$= \frac{h_1'(x_1)}{h_1(x_1)} + (1 - \sigma_1^* + \frac{\sigma_1^*}{1 + e^{-y_1}}) \frac{e^{-x_1}(1 + e^{-y_1})}{1 + e^{-x_1 - g(y_1)}}$$

If $\sigma_i^* = 0$, then $\frac{\partial ln(E[\pi_1])}{\partial x_i} = \frac{h_1'(x_1)}{h_1(x_1)} + \frac{e^{-x_1}(1+e^{-y_1})}{1+e^{-x_1-g(y_1)}}$ is strictly decreasing in y_i. Alternatively, if $\sigma_i^* = 1$, then $\frac{\partial ln(E[\pi_1])}{\partial x_1} = \frac{h_1'(x_1)}{h_1(x_1)} + \frac{1}{1+e^{-y_1}} \frac{e^{-x_1}(1+e^{-y_1})}{1+e^{-x_1-g(y_1)}} = \frac{h_1'(x_1)}{h_1(x_1)} + \frac{e^{-x_1}}{1+e^{-x_1-g(y_1)}}$ is strictly increasing in y_1. Thus, x^* is decreasing in y_1 for a fully transparent monopolist and increasing in y_1 for a fully opaque monopolist. Similar arguments can be made for prices where the fully opaque monopolist will decrease the price if y_1 increases. In this fully opaque setting, the effect that y_1 has on the partial derivative through $a_1^*(y_1)$ dominates the countervailing effect from $g_1(y_1)$.

(5) $F_i^*(v) = 1 - e^{-e^{A_i v}}$ and $f_i^*(v) = e^{-e^{A_i v} - A_i v}$ so

$$\frac{f_i^*(v)}{1 - F_i^*(v)} = e^{-A_i v}$$

which is log-linear (and weakly log-concave).

Lemma 10. *Assume a standard REVSM. Let $\mathcal{P}_{-i}$ denote the power set of the set of options excluding i. Let $z_i = x_i + g(y_i)$.*

$$E[D_i] = N\ Prob[V_{in}^* > \max_{j \neq i} V_{jn}^*] = N \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \frac{e^{-z_i}}{e^{-z_i} + \sum_{j \in \theta} e^{-z_j}}$$

$$E[S_i] = N\ Prob[\bar{V}_{in} > \max_{j \neq i} V_{jn}^*] = N \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \frac{e^{-x_i - y_i}}{e^{-x_i - y_i} + \sum_{j \in \theta} e^{-z_j}}$$

Let I^s denote the realized set of options which are searched (consideration set). Let i^* be the realized selection where $i^* \in I^s$.

$$Prob[i^*, I^s] = \sum_{\theta \in \mathcal{P}_{-i^*}} (-1)^{|\theta|} \frac{e^{-z_i}}{e^{-z_i} + \sum_{j \in (\theta \cup I^s)} e^{-x_j - y_j} + \sum_{j \in (\theta \cap I^s)} e^{-x_j}}$$

Moreover, $ln(E[D_i])$, $ln(E[S_i])$ *and* $ln(Prob[i^*, I^s])$ *are concave in terms of* $\mathbf{x}$ *and* $\mathbf{y}$*.*

$$E[D_i] = N * \int_{-\infty}^{\infty} \prod_{j \neq i} (1 - e^{-e^{v - z_j}}) e^{-e^{v - z_i} + v - z_i} dv$$

$$= N * \int_{-\infty}^{\infty} \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} (e^{-e^{v} (\sum_{j \in \theta} e^{-z_j})}) e^{-e^{v - z_i} + v - z_i} dv$$

$$= N * \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \int_{-\infty}^{\infty} (e^{-e^{v} (\sum_{j \in \theta} e^{-z_j} + e^{-z_i})}) e^{-e^{v} + v - z_i} dv$$

$$= N * \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \frac{e^{-z_i}}{e^{-z_i} + \sum_{j \in \theta} e^{-z_j}}$$

$$E[S_i] = N * \int_{-\infty}^{\infty} \prod_{j \neq i} (1 - e^{-e^{v - z_j}}) e^{-e^{v - x_i - y_i} + v - x_i - y_i} dv$$

$$= N * \int_{-\infty}^{\infty} \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} (e^{-e^{v} (\sum_{j \in \theta} e^{-z_j})}) e^{-e^{v - x_i - y_i} + v - x_i - y_i} dv$$

$$= N * \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \int_{-\infty}^{\infty} (e^{-e^{v} (\sum_{j \in \theta} e^{-z_j} + e^{-x_i - y_i})}) e^{-e^{v} + v - x_i - y_i} dv$$

$$= N * \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \frac{e^{-x_i - y_i}}{e^{-x_i - y_i} + \sum_{j \in \theta} e^{-z_j}}$$

To better understand the expression of $Prob[i^*, I^s]$, I rely on the earlier results for optimal search. If option j is searched, but not selected by agent n, then

$\bar{V}_{jn} > V^*_{i^*n} > V_{jn}$. If option j is not searched by agent n, then $V^*_{i^*n} > \bar{V}_{jn}$. Thus for ordered search where each $\alpha_{in} = 0$,

$$Prob[i^*, I^s] = \int_{-\infty}^{\infty} \prod_{j \in I^s - \{i^*\}} \left(F^{\epsilon}_j (v - x_j)(1 - F^{\mu}_i (v - x_j)) \right) \prod_{j \notin I^s} \left(F^{\mu}_j (v - x_j) \right) dF^{min}_i (v - x_i)$$

$$= \int_{-\infty}^{\infty} \prod_{j \in I^s - \{i^*\}} \left((1 - e^{-e^{v - x_j}})(e^{-e^{v - x_j - y_j}}) \right) \prod_{j \notin I^s} (1 - e^{-e^{v - x_j - y_j}}) e^{-e^{v - z_i} + v - z_i} dv$$

$$= \int_{-\infty}^{\infty} \prod_{j \in I^s - \{i^*\}} (1 - e^{-e^{v - x_j}}) \prod_{j \notin I^s} (1 - e^{-e^{v - x_j - y_j}}) \prod_{j \in I^s - \{i^*\}} (e^{-e^{v - x_j - y_j}}) e^{-e^{v - z_i} + v - z_i} dv$$

$$= \int_{-\infty}^{\infty} \prod_{j \in I^s - \{i^*\}} (1 - e^{-e^{v - x_j}}) \prod_{j \notin I^s} (1 - e^{-e^{v - x_j - y_j}}) e^{-e^{v}(e^{-z_i} + \sum_{j \in I^s} e^{-x_j - y_j}) + v - z_i} dv$$

$$= \int_{-\infty}^{\infty} \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} e^{-e^{v}(e^{-z_i} + \sum_{j \in I^s} e^{-x_j - y_j} + \sum_{j \in (\theta - I^s)} e^{-x_j - y_j} + \sum_{j \in (\theta \cap I^s)} e^{-x_j}) + v - z_i} dv$$

$$= \sum_{\theta \in \mathcal{P}_{-i}} (-1)^{|\theta|} \int_{-\infty}^{\infty} e^{-e^{v}(e^{-z_i} + \sum_{j \in (\theta \cup I^s)} e^{-x_j - y_j} + \sum_{j \in (\theta \cap I^s)} e^{-x_j}) + v - z_i} dv$$

$$\sum_{\theta \in \mathcal{P}_{-i^*}} (-1)^{|\theta|} \frac{e^{-z_i}}{e^{-z_i} + \sum_{j \in (\theta \cup I^s)} e^{-x_j - y_j} + \sum_{j \in (\theta \cap I^s)} e^{-x_j}}$$

Recall, log-concavity is preserved over multiplication and integration. Also, if a density f is log-concave, then $1 - F$ and F are log-concave. Since $\Pi(1 - e^{-e^{v - z_j}})e^{-e^{v - z_i} + v - z_i}$ is log-concave in ($\mathbf{z}$,v) and $z_i = x_i - ln(1 + e^{-y_i})$ is concave and increasing in (x_i, y_i) $\forall i$, $E[D_i]$ is log-concave in $\mathbf{x}$ and $\mathbf{y}$. By a similar argument, $E[S_i]$ and $Prob[i^*, I^s]$ are log-concave in $\mathbf{x}$ and $\mathbf{y}$.

Corollary 4. *Assuming that F_0 and F^*_i for each option i have a support over $\mathbb{R}$, $\mathbf{H}$ is a continuous bijection from $\mathbf{Q}$ to $\mathbf{W}$.*

This corollary relies on the bijection proof for classic discrete choice in the appendix of Berry (1994) which proves that there is a bijection from the space of vectors of $\mathbf{x}$ to expected demands. From lemma 4, expected demands do not change with $\mathbf{a}$ and :

for any I dimensional vector $\mathbf{s}$ where $E[\mathbf{D}] < \mathbf{s} < \bar{\mathbf{D}}$, there exists a unique vector of $\mathbf{a}$ values such that $E[\mathbf{S}] = \mathbf{s}$. Thus, $\mathbf{H}$ is a bijection.

Lemma 11. *Suppose equilibrium prices are known to the econometrician. Suppose the demand and search parameters are known to the econometrician* ($\mathbf{x}, \mathbf{a}$)*.*

Let $\mathbf{c}^{trans}$ *be the identified unit cost vector assuming transparency* ($\sigma_i^* = 0 \quad \forall i$)*.*

Let $\mathbf{c}^{op}$ *be the identified unit cost vector assuming full opacity* ($\sigma_i^* = 1 \quad \forall i$)*.*

For any vector of firm opacities in the market, the identified vector of unit costs $\mathbf{c}^*$ *satisfies:*

$$\mathbf{c}^{op} \leq \mathbf{c}^* \leq \mathbf{c}^{trans}$$

Now suppose the demand parameters ($\mathbf{x}$) *are known, but* $\mathbf{a}$ *is not known.*

For any vector of relative firm opacities in the market, the identified vector of unit costs $\mathbf{c}^*$*:*

$$\mathbf{c}^* \leq \mathbf{c}^{trans}$$

However, nothing more is known without making further assumptions about opacity (or imposing non-negativity of unit costs).

This follows directly from the fact that the inverted values for c_i is decreasing in σ_i^* because

$$c_i = p_i + \frac{E[D_i]}{((1 - \sigma_i^* + \sigma_i^* a_i))\frac{\partial E[D_i]}{\partial p_i}}$$

This holds for any firm i.

References

Anderson, Simon P, Andre De Palma, and Jacques François Thisse (1992) *Discrete choice theory of product differentiation*: MIT press.

Anderson, Simon P and André de Palma (1999) "Reverse discrete choice models," *Regional Science and Urban Economics*, Vol. 29, pp. 745–764.

Anderson, Simon and Regis Renault (2018) "Search direction: Position externalities and position bias," *mimeo*.

Armstrong, Mark (2015) "Search and ripoff externalities," *Review of Industrial Organization*, Vol. 47, pp. 273–302.

——— (2017) "Ordered consumer search," *Journal of the European Economic Association*, Vol. 15, pp. 989–1024.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile prices in market equilibrium," *Econometrica*, pp. 841–890.

Berry, Steven T (1994) "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, pp. 242–262.

Choi, Michael, Anovia Yifan Dai, and Kyungmin Kim (2018) "Consumer search and price competition," *Econometrica*, Vol. 86, pp. 1257–1281.

Fosgerau, Mogens, Emerson Melo, De Palma Andre, and Matthew Shum (2017) "Discrete choice and rational inattention: A general equivalence result," *mimeo*.

Haan, Marco, José Luis Moraga-González, and Vaiva Petrikaite (2015) "Price and match-value advertising with directed consumer search," *mimeo*.

Hauser, John R and Birger Wernerfelt (1990) "An evaluation cost model of consideration sets," *Journal of consumer research*, Vol. 16, pp. 393–408.

Honka, Elisabeth and Pradeep Chintagunta (2016) "Simultaneous or sequential? search strategies in the us auto insurance industry," *Marketing Science*, Vol. 36, pp. 21–42.

Janssen, Maarten and Sandro Shelegia (2018) "Beliefs and consumer search in a vertical industry," *mimeo*.

Matejka, Filip and Alisdair McKay (2015) "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, Vol. 105, pp. 272–98.

Misra, Sanjog (2005) "Generalized reverse discrete choice models," *Quantitative Marketing and Economics*, Vol. 3, pp. 175–200.

Moraga-González, José Luis and Vaiva Petrikaitė (2013) "Search costs, demand-side economies, and the incentives to merge under Bertrand competition," *The RAND Journal of Economics*, Vol. 44, pp. 391–424.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R Wildenbeest (2017a) "Nonsequential search equilibrium with search cost heterogeneity," *International Journal of Industrial Organization*, Vol. 50, pp. 392–414.

——— (2017b) "Prices and heterogeneous search costs," *The RAND Journal of Economics*, Vol. 48, pp. 125–146.

——— (2018) "Consumer search and prices in the automobile market," *mimeo*.

Quint, Daniel (2014) "Imperfect competition with complements and substitutes," *Journal of Economic Theory*, Vol. 152, pp. 266–290.

Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel (2011) "Search dynamics in consumer choice under time pressure: An eye-tracking study," *American Economic Review*, Vol. 101, pp. 900–926.

Weitzman, Martin L (1979) "Optimal search for the best alternative," *Econometrica*, pp. 641–654.

Zhou, Jidong (2011) "Ordered search in differentiated markets," *International Journal of Industrial Organization*, Vol. 29, pp. 253–262.

——— (2014) "Multiproduct search and the joint search effect," *The American Economic Review*, Vol. 104, pp. 2918–2939.