When Does Information Determine Market Size? Search and Rational Inattention*

Costas Cavounidis†

December 5, 2016

Abstract

I develop a model in which optimal costly information acquisition by individual firms causes adverse selection in the market as a whole. Each firm's information acquisition policy determines which customers they provide to, and that in turn affects the distribution of customers remaining in the market and hence other firms' incentives. I show that if firms possess the ability to choose any signal of the customer's type, in equilibrium all firms in the market will profit. By contrast, with restricted signal choice, only a limited number of firms can be profitable. In such a setting, the maximum number of profitable firms fails to increase with the number of potential customers. Smooth information acquisition dampens the adverse selection externality due to each firm, while lumpy information acquisition does not. I establish that my results apply to a broad class of continuous-time information acquisition processes.

*I am grateful to Kevin Cooke, Sambuddha Ghosh, Kevin Lang, Bart Lipman, Calvin Luscombe, Albert Ma, Dilip Mookherjee, Benjamin Ogden, Juan Ortner, Juuso Toikka, and audiences at Boston University. For the newest version of the paper, visit http://blogs.bu.edu/ccavouni/research/

†Boston University, ccavouni@bu.edu, http://blogs.bu.edu/ccavouni

1 Introduction

Consider a venture capitalist inspecting a start-up soliciting an investment. Should she invest in the start-up, she will get a payoff that depends on its profitability. Ideally, she invests if the return on investment clears some benchmark. However, start-ups vary in type, and this variation is not immediately observable to the venture capitalist. She may aid her decision by using an information acquisition technology to learn about the start-up's expected profitability, at a cost. The venture capitalist must, naturally, keep in mind the prior distribution over start-up types before deciding whether and how to use this technology.

However, she must also consider that she may not be the first potential investor solicited. Could the start-up have been rejected for investment by other venture capitalists? What would those other venture capitalists have learned about the start-up, and how would this information influence rejections? The answers to these questions affect the venture capitalist's beliefs about the start-up, and therefore the way the technology will be used. There are clearly spill-over effects in information acquisition decisions. However, an intriguing possibility arises: could it be that the induced update on beliefs by others' decisions is so pessimistic that the start-up is immediately rejected, without consideration? The answer to this question will turn on the nature of the information acquisition technology.

I develop a framework that allows exploration of the link between information acquisition technology and the limitations it imposes on the number of its users - the 'market size'. The framework is simple yet flexible, with the emphasis on allowing for variation of the information acquisition technology. I obtain precise market size predictions in the two most natural instances of the model, which are then augmented with a generalization to a large class of continuous-time information acquisition processes.

This paper studies a model of sequential search with asymmetric information regarding the value of transacting. A 'firm', faced with uncertainty, may choose whether, and possibly additionally what, to learn (at a cost) about the 'applicant' before deciding to accept or reject her. Each firm will choose what to learn optimally, given other firms' behavior.

Each firm does not know, when visited, if the applicant has visited other firms. As the applicant leaves the market when accepted and the same information acquisition technology is available to every firm, in equilibrium there is a form of adverse selection in the unmatched that is directly related to the technology's characteristics. Therefore, learning is interactive not because a firm is directly affected by its opponents' actions, but rather because these actions determine the level of adverse selection in the market. This endogenous adverse selection in turn affects what information the firm must acquire to protect itself. I study how the information acquisition technology, through this channel, will affect equilibrium payoffs.

An applicant sequentially visits firms until either she is accepted, or she has visited every one of the finitely many firms. The applicant is characterized by one of finitely many types, the value common to all firms of a transaction with her. A firm, upon being visited, may choose to acquire a signal whose distribution varies with the applicant's type, bearing a cost that depends on the signal's informativeness. The choice set of such signals is a parameter of interest I will vary, but is common to all firms. Conditional on the applicant's type, signals acquired by different firms are independent. Once the cost is paid and the signal observed, the firm may reject the applicant, getting no further payoff, or accept her, getting a payoff based on her type. An accepted applicant exits the model.

If the information acquisition technology is *smooth*, it allows the firm to decide what to learn about the applicant flexibly. On the other hand, if the technology is *lumpy*, it allows a firm to either become fully informed or stay uninformed (or either with some probability, via mixing).

The main insight in this paper is that when firms possess a lumpy technology, the adverse selection externality they collectively impose on one another grows with the number of firms; whereas if firms use a smooth technology, they adjust to increased competition in a way that reduces their individual externalities, keeping adverse selection in check. As a result, with restricted information acquisition only a limited number of firms can attain a strictly positive payoff, but with flexible information acquisition all firms will be strictly profitable. Therefore, the number of firms in a market is affected by the available information acquisition technology.

The intuition for these results lies in the distribution of types rejected by each firm. When the technology is smooth, the marginal cost of additional information acquired about rejected applicants is equal to the marginal benefit. When a firm is uncertain whether a visiting applicant has been rejected at another firm or not, its beliefs about her are therefore at least a bit better than about the rejected; it is therefore beneficial to acquire at least some information. This will in equilibrium imply strictly positive profits. When the technology is lumpy, with enough firms, beliefs about rejected applicants are not marginal but rather excessively bad, so that the number of firms that can profitably become informed is limited.

If the cost of the fully informative signal was fixed, one would be concerned that the result is driven by the fact that, as adverse selection grows worse with the number of firms, the information provided by the signal is diminishing but the cost is not. To avoid this trap, I allow the cost of the signal to depend continuously on the informational content, which in turn depends on the prior. With information costs based on Shannon's mutual information, firms only 'pay for what they get' - as the ex-ante uncertainty goes to zero, so does the signal cost.[1] The result is also not driven by discontinuity in the lumpy case either, as mixing allows a firm to produce a signal with any probability of being fully informative. The 'lumpiness' of the information comes not from an all-or-nothing choice of signal, then, but rather from the extreme nature of the informed posteriors.

In a sense, flexible learning, combined with continuous costs based on mutual information, 'dampens' adverse selection. The marginality of a rejected applicant ensures that the updating on being visited is not as pessimistic as in the lumpy case. As the number of firms expands in the smooth learning model, they each learn less, effectively always leaving surplus for the rest. However, the main results are not dependent on the mutual information cost function, and a generalization of the 'smooth' case to a very broad setting is considered in section 5.

The model in this paper can apply to a variety of contexts. Other than investment in start-ups, applications include hiring in labor markets with non-negotiable (e.g. minimum) wages, insurance markets, and even human mating. It can be used, for instance, to think about the effects of the prohibition of health insurance rejections in the Affordable Care Act. While one may think the prevention of screening by insurers would cause adverse selection, the results in this paper show that screening creates a different, potentially worse, kind of adverse selection. The model does not address price-setting, but it can accommodate pricing based on variable observable flow outside options - as long as it is not determined after information acquisition. One can show the paper's main results continue to hold, mutatis mutandis, even when

[1]This is going to require a somewhat non-standard formulation of the game, as beliefs do not typically enter payoffs directly.

market size affects such prices.

The present paper builds heavily on work on search with adverse selection. Inderst (2005) introduces a search model where contracts are used to separate types in equilibrium. Lauermann and Wolinsky's (2016) search model determines the level of information aggregation when exogenous, conditionally i.i.d. signals of a buyer's type are available to sellers. Zhu (2012) uses a decentralized search market with a similar signaling technology to model exploding offers in over-the-counter markets. These last two models strongly develop the intuition for the *solicitation effect*, the fact that being visited at all can be a negative signal, as it speaks of possible rejection from other potential transactions. However, the information acquisition process is taken to be both free and exogenous, with the focus being on prices.

The model in this paper has certain parallels with common values auctions, in particular in the existence of a probabilistic version of the Winner's Curse; in equilibrium, an accepted applicant will have had worse signal realizations at all firms visited previously. Persico (2000) examines optimal information acquisition in auctions, finding that different auction procedures induce different levels of information acquisition. Bergemann, Shi and Välimäki (2009) study auctions with interdependent values and a binary information choice. In that context they model the number of bidders who choose to become informed much in the same spirit as this paper addresses market size.

Finally, the present model is heavily influenced by the literature on rational inattention. The model uses as a measure of informativeness mutual information, built on Shannon (1948)'s notion of entropy, used in economics as a cost of information measure since Sims (2003). Strategic interaction with rational inattention has been studied by both Yang (2015) and Denti (2016) in the context of coordination games. Gentzkow and Kamenica (2014) write a model in which the informed party must pay to disclose information, rather than to acquire it. Ravid (2016) examines rational inattention in a bargaining model with one-sided offers. Additionally, Matějka and McKay (2015) study discrete choices with rational inattention, and special attention is paid to binary choice by Woodford (2008).

The rest of this paper proceeds as follows. Section 2 introduces the main model. Section 3 contains existence theorems for the model's main cases, while section 4 provides the main market size results. Section 5 generalizes the main results to a broad class of settings with continuous-time learning, and section 6 concludes.

2 Model

There exist N identical firms and a single applicant. The applicant is characterized by a private type, θ, which is distributed according to prior $p_0(\cdot)$ with full support on finite, non-singleton $\Theta \subset \mathbb{R} \setminus \{0\}$. θ is the net benefit to a firm of accepting the applicant. Crucially, I assume $\min \Theta < 0 < \max \Theta$; firms want to accept some, but not all, applicants.

2.1 Timing

Nature chooses the type θ of the applicant according to p_0, and a *visit order* σ from the set of permutations of the N firms equiprobably. Then, the game proceeds in up to N stages, starting with stage 1.

In stage n, firm $\sigma^{-1}(n)$ is visited. The visited firm may then choose to acquire information about θ using the available technology. Once information is acquired, the firm may choose to accept or reject the applicant. If the applicant is accepted or $n = N$, the game ends. If the applicant is rejected and $n < N$, stage $n + 1$ follows.

That is, the applicant visits firms in a (uniform) random order, until either he is accepted, or he has been rejected by every firm. When the applicant visits a firm, the firm sees only that an applicant has arrived, not her history, the visit order or (equivalently) the date.

2.2 Beliefs

Each firm acts at a single information set. That is to say, firms are not aware of the applicant's history of visiting other firms. As a history with rejections is bad information for the applicant, the applicant would not disclose her history even if the model were augmented with a cheap-talk stage.

In equilibrium, each firm n has beliefs over the nodes in its information set whose marginal over θ is not, in general, $p_0(\cdot)$ but rather the *equilibrim belief* p_n. The equilibrium belief is computable from the prior, the applicant's random visit path, and the equilibrium acceptance probability of each type of applicant at each rival firm via Bayes' rule. These distributions are endogenous variables of the model, as they are produced by firms' strategies.

2.3 Information Acquisition

When the applicant visits a firm, that firm can then acquire information about the applicant's type. A *signal structure* is a set of conditional distributions $\{g(\cdot|\theta)\}_{\theta\in\Theta}$ for a signal $s \in S$, where S is a finite alphabet.[2] A collection of signal structures $\mathcal{G} = \{\{g_i(\cdot|\theta)\}_{\theta\in\Theta}\}_{i\in I} = \{g_i(\cdot|\cdot)\}_{i\in I}$ is called an *information menu.* Signal structures will vary in cost, so a more informative signal will not always be preferable.

I consider two different information menus. The first is the *unrestricted* information menu, $\mathcal{G}_U = (\Delta S)^\Theta$. As it is comprised of all conditional distributions for the signal, it is the largest possible set of signal structures. The second is the *restricted* information menu, $\mathcal{G}_R = \{g_{no}, g_{all}\}$, comprised of g_{no} with $g_{no}(s|\theta) = g(s)$, a completely uninformative signal structure, and the completely informative signal structure g_{all}, with $g_{all}(s|\theta) = \delta_\theta$, so that the signal is always the same as the applicant's type.

2.4 Strategies

A *single-structure* strategy for firm n is a pair (g, a) where

- $g \in \mathcal{G}$ is a signal structure; and
- $a : S \to [0, 1]$ maps the signal to a probability of accepting the applicant.

The information menu $\mathcal{G}$ is common to all firms. As the firm's problem may have multiple solutions, the firm may pursue a strategy that mixes over signal structures. As the posterior distribution of types conditional on a signal realization will depend on the signal structure that produced the signal in question, the conditional acceptance probability must be free to vary with the chosen signal structure. Therefore the set of mixed strategies is defined as the set of distributions over single-structure strategies, $\Delta(\mathcal{G} \times [0, 1]^S)$, rather than $\Delta\mathcal{G} \times \Delta([0, 1]^S)$.

2.5 Mutual Information

Shannon (1948) lays out a particular measure of uncertainty blind to economic consequences. For an arbitrary distribution P with finite support X, the Shannon entropy

[2] In general, a signal alphabet of cardinality equal to that of the set of resulting actions is sufficient in unrestricted models of rational inattention. I assume $|S| \geq |\Theta|$.

is defined as

$$H(P) \equiv -\sum_{x \in X} P(x) \ln P(x) \tag{1}$$

and can be interpreted, if $-\ln P(x)$ is thought of as the *surprisal*[3] in observing the realization x, as the *average surprisal* in P.[4]

If a signal distributed by $g(\cdot|\cdot)$ provides a lot of information about θ, it reduces the uncertainty about θ by on average concentrating its posterior distribution $p_n(\cdot|\cdot)$. Our measure of the informativeness of a signal structure is the expected reduction in the average surprisal achieved by updating the equilibrium belief p_n using the signal. This expected reduction in the entropy of p_n by observing a draw from g is known as the *mutual information* between p_n and g and is computed as

$$M(p_n, g) = H(p_n) - E_s[H(p_n(\cdot|s))] \tag{2}$$

where the expectation is taken according to g.

Alternatively, mutual information can be defined as the expected Kullback-Leibler divergence between the prior and the posterior distributions of θ:

$$M(p_n, g) = \sum_{s \in S} \left(\sum_{\theta \in \Theta} p_n(\theta) g(s|\theta) \right) D_{KL}(p_n(\cdot|s) || p_n(\cdot)). \tag{3}$$

This defines mutual information in terms of the extent to which it will in expectation shift the posterior, where the notion of posterior-shifting is given by the divergence.

Starting with Sims (2003), it is common for work on rational inattention to use mutual information as the cost of information. The main advantages are the treatment of uncertainty as generic and separable from economic consequences, the agreement with Blackwell-informativeness, the logit-like discrete choice probabilities as highlighted by Matěka and McKay (2015) and the ability to price all signal structures.

[3]The notion of surpisal originates in the information theory literature. Intuitively, observing an ex-ante unlikely event conveys more information than observing one initially thought to be likely. The log functional form (uniquely) allows additivity over intersections of independent events as $\ln(P(A \cap B)) = \ln(P(A) \cdot P(B)) = \ln P(A) + \ln P(B)$.

[4]For this and all other purposes, this paper uses the common convention $0 \ln 0 = 0$ in accordance with the limit.

2.6 Payoffs

A firm n pays for signal structure g a cost equal to a multiple k of the mutual information $M(\cdot,\cdot)$ between its equilibrium belief of the applicant's type p_n and that of the signal g.[5]

A firm that accepts an applicant with type θ gains utility θ for doing so, whereas rejecting any applicant gives a payoff of 0. Therefore, a firm n with beliefs p_n about a new applicant will choose what information to acquire via g and how to respond to it via the conditional acceptance probability a so as to maximize profits. I normalize payoffs so that they are ex-post of applicant arrival. Firm n's payoff, given equilibrium belief p_n and single-structure strategy $(g,a) \in \mathcal{G} \times [0,1]^S$

$$a(s)\theta - kM(p_n, g). \tag{4}$$

As this quantity depends directly on beliefs, the game defined above is a Psychological Game as described in Geanakoplos, Pearce and Stacchetti (1989). In effect, the requirement that firms only pay for the amount of information they acquire means that the signal's cost must depend on how much it shifts the firm's beliefs from the equilibrium beliefs.

3 Equilibrium

An equilibrium for the N-firm model is given by mixed strategies for all firms, with firm n playing $F_n \in \Delta(\mathcal{G} \times [0,1]^S)$, that form a Nash Equilibrium, given that each firm's beliefs about arriving applicants follow Bayesian inference.

In equilibrium, each firm n will therefore choose a strategy that solves

$$\max_{F_n \in \Delta(\mathcal{G}\times[0,1]^S)} \int_{\mathcal{G}\times[0,1]^S} \left[\sum_{\theta\in\Theta} p_n(\theta) \sum_{s\in S} g(s|\theta)a(s)\theta - kM(p_n, g) \right] dF_n(g,a). \tag{5}$$

[5] An abstract generalization of the model's main results to other information acquisition technologies can be found in section 5.

3.1 Inference

Consistency requires firms' beliefs $(p_n)_{n\leq N}$ are derived from firms' strategies $(F_n)_{n\leq N}$ and the prior p_0 via Bayes' Rule. A type θ applicant has a probability

$$A_m(\theta) \equiv \int_{\mathcal{G}\times[0,1]^S} \sum_{s\in S} g(s|\theta)a(s)dF_m(g,a) \tag{6}$$

of being accepted by firm m. The applicant's visiting sequence is a uniform draw from the permutations of N, with each permutation getting probability $1/N!$. Therefore, the probability a type θ applicant ever visits firm n is $1/N!$ times the sum over these permutations of the probability that the applicant is rejected at every previously visited firm:

$$\frac{1}{N!} \sum_{\sigma\in perm(N)} \prod_{m:\sigma(m)<\sigma(n)} (1 - A_m(\theta)). \tag{7}$$

Thus the posterior probability that an applicant arriving at firm n is of type θ is the prior $p_0(\theta)$ times the sum over all visiting sequences of the probability that θ visits firm n, over the same for all types:

$$p_n(\theta) = \frac{p_0(\theta) \sum\limits_{\sigma\in perm(N)} \prod\limits_{m:\sigma(m)<\sigma(n)} (1 - A_m(\theta))}{\sum\limits_{\theta'\in\Theta} p_0(\theta') \sum\limits_{\sigma\in perm(N)} \prod\limits_{m:\sigma(m)<\sigma(n)} (1 - A_m(\theta'))}. \tag{8}$$

I denote by $\hat{p}_n : [0,1]^{(N-1)|\Theta|} \to \Delta\Theta$ the *belief function for firm* n taking all other firms' acceptance probability vectors $(A_m)_{m\neq n}$ into firm n's posterior distribution for θ using Bayes' rule:

$$\hat{p}_n((A_m)_{m\neq n}) = \left(\frac{p_0(\theta) \sum\limits_{\sigma\in perm(N)} \prod\limits_{m:\sigma(m)<\sigma(n)} (1 - A_m(\theta))}{\sum\limits_{\theta'\in\Theta} p_0(\theta') \sum\limits_{\sigma\in perm(N)} \prod\limits_{m:\sigma(m)<\sigma(n)} (1 - A_m(\theta'))} \right)_{\theta\in\Theta}. \tag{9}$$

As the expression in the numerator is a polynomial in the $A_m(\theta)$s and the denominator is a sum of such polynomials, each of which is bounded below by $\frac{1}{N}p_0(\theta)$, $\hat{p}_n$ is a continuous function. Continuity of the belief functions will be useful in showing existence of equilibria.

3.2 Equilibrium with unrestricted information acquisition

I first turn to the case when $\mathcal{G} = \mathcal{G}_U$. Let two signal structures g, g' be *equivalent* if they induce the same distribution over posteriors on θ. It is useful here to use the characterization of binary choice with mutual information costs given by Woodford(2008). This will provide both uniqueness and continuity of best responses in p_n.

Theorem 1 (Woodford 2008) *In the model with unrestricted information acquisition, for a firm holding beliefs p_n, there is a unique up to equivalence optimal signal structure. Each signal in the support of an optimal signal structure leads to a pure choice, that is $a(s) \in \{0, 1\}$. Also, the induced optimal acceptance probabilities $A_n(\theta)$ satisfy*

- *if* $\sum p_n(\theta)e^{-\frac{\theta}{k}} \leq 1, \forall\theta\ A_n(\theta) = 1$
- *if* $\sum p_n(\theta)e^{\frac{\theta}{k}} \leq 1, \forall\theta\ A_n(\theta) = 0$
- *otherwise,* $A_n(\theta)$ *is given by*

$$\frac{A_n(\theta)}{1 - A_n(\theta)} = \frac{\bar{A}}{1 - \bar{A}} e^{\frac{\theta}{k}}, \text{ where} \tag{10}$$

$$\bar{A} = \sum p_n(\theta) A_n(\theta). \tag{11}$$

Importantly, we can use the theorem to exclude mixtures over non-equivalent signal structures (as they would each be optimal but non-equivalent). Mixing over equivalent signal structures merely amounts to randomizing *how* information is acquired, but not *what* is learned; the induced distributions over posteriors are the same. Effectively, equivalent signal structures are mere relabellings.

Furthermore, Woodford (2008) shows that optimal actions can be summarized as $(A_n(\theta))_{\theta\in\Theta}$, the optimal acceptance probability for each type; and that those are unique. In effect, the theorem allows us to restrict to a much smaller strategy space when looking for best responses. As this space is compact, equilibrium existence will be easy to show.

It is useful at this point to define the function that takes firm n's beliefs p_n into optimal acceptance probabilities as $\hat{A}_n : \Delta\Theta \to [0, 1]^{|\Theta|}$ using the rule given in (10) and (11).

Theorem 2 *For any $N \in \mathbb{N}$, there exists an equilibrium in the model with N firms and unrestricted information acquisition.*

All proofs are in the appendix. As I have shown that we only need tro consider a compact and convex action space, and Theorem 1 gives a continuous best response, the proof of Theorem 2 is a straightforward application of Brouwer's Fixed Point Theorem.

3.3 Equilibrium with restricted information acquisition

When $\mathcal{G} = \mathcal{G}_R$, when the firm becomes informed, it is optimal to accept all $\theta \geq 0$. Thus the only relevant choice is whether to acquire full or no information, with each firm choosing a probability of playing each of these two actions. However, (9) implies that the cost of becoming informed depends non-linearly on opponent strategies, the Nash Existence Theorem does not apply. However, the restricted action space is compact and allows an application of Kakutani's Fixed Point Theorem once the best response correspondences are shown to be closed-graph.

Theorem 3 *For any $N \in \mathbb{N}$, there exists an equilibrium in the model with N firms and restricted information acquisition.*

Neither of the existence theorems presented requires that the sequence in which the applicant visits firms is uniformly distributed. What is required is that there is a strictly positive probability for each firm to be the first visited, so that Bayes' rule is well-defined for every strategy profile[6]

4 Market Size

So far, I have kept the number of firms in the market fixed. This paper's main results pertain to the number of firms that can profit in the market.

4.1 Market size with unrestricted information acquisition

For the market with unrestricted information acquisition and N firms, two possibilities arise. If the proportions and relative value of the types in the prior are too low, low

[6] That is, (9) must have a positive denominator.

enough that even at the prior a monopolist is unwilling to acquire any information or ever accept applicants, then the only equilibrium will be one in which no firms profit. This will occur iff $\sum_\theta p_0(\theta)e^{\frac{\theta}{k}} \leq 1$. Firms that have no profits acquire no information[7]. If a monopolist cannot profit, it is quite clear that neither can more firms, with the associated negative externality each imposes on the others. This rather trivial case aside, Theorem 4 will show that as long as a monopolist would be profitable, any number of firms are profitable, as equilibrium beliefs p_n for any firm n will also satisfy $\sum_\theta q_m(\theta)e^{\frac{\theta}{k}} > 1$.

This guaranteed profitability occurs because in any equilibrium, any profitable firm's *rejected* applicants will have a probability distribution q_m so that $\sum_\theta p_n(\theta)e^{\frac{\theta}{k}} = 1$ holds exactly. Then, assuming a firm n gets a payoff of 0 and thus rejects all applicants, the pool of potential applicants it receives is comprised of rejects from some other firm m with type distribution q_m as well as first-time applicants, with a type distribution p_0; thus, the average applicant p_n will be a convex combination of q_m's and p_0 and as such will satisfy $\sum_\theta p_n(\theta)e^{\frac{\theta}{k}} > 1$. Therefore, the assumption that firm n rejects all applicants and gets a payoff of 0 is contradicted.

This surprising result occurs because rejected applicants have been *marginally* learned about. That is, optimal learning in the unrestricted information acquisition setting requires that at the time of rejection, the rejecting firm has beliefs about the applicant on the boundary of the set of profitable beliefs. If a firm is uncertain whether an applicant has just been rejected at some other firm or is a new draw, any interior probability weights on these two cases will produce beliefs about the applicant's type that allow the firm to make a profit.

Theorem 4 *If a monopolist would be profitable, for every $N \in \mathbb{N}$, in every equilibrium of the model with unrestricted information acquisition and N firms each firm gets a strictly positive payoff.*

The theorem implies that if there exist a finite number of potential firms, each of which can decide to enter at sufficiently low cost, in equilibrium, they all will. If the choice of information structure is unrestricted, the information acquisition technology does not erect a barrier to entry.

Figure 1 shows how acceptance rates react to beliefs p_n in a two-type market. As the number of firms expands, the average acceptance probability for each firm $\bar{A}_n$

[7]The converse is not true; it is possible that a firm acquires no information but profitably accepts the applicant.

will approach 0, but reach it only in the limit; for any finite number of firms, information will be worth acquiring for each firm. The figure illustrates how a lower prior forces a firm to be more cautious, decreasing the acceptance rates for both types, while increasing the ratio of the probability a good type is accepted to that a bad type is accepted. As the number of firms increases, each firm moderates its acceptance probabilities, reducing its individual contribution to the market-level adverse selection.

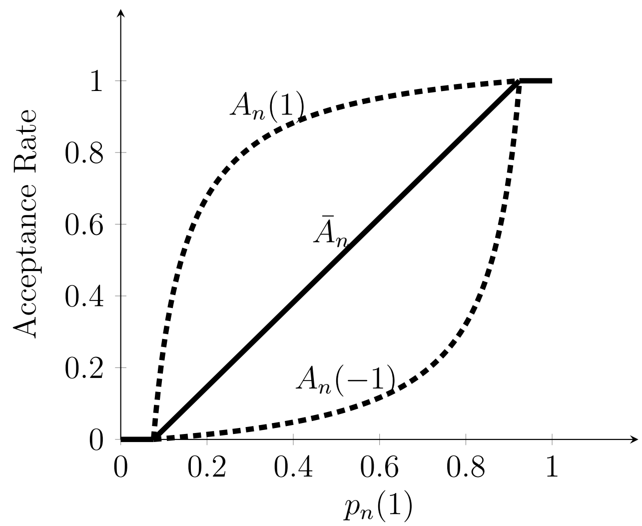

Figure 1: Optimal acceptance rates in the unrestricted information acquisition market with $\Theta = \{-1, 1\}, k = .4$ as a function of the probability that $\theta = 1$.

4.2 Market size with restricted information acquisition

I now turn to the case when $\mathcal{G} = \mathcal{G}_R$, when firms can choose to acquire either all or no information. The main assumption I make is $\Pi = \sum_{\theta>0} p_0(\theta)\theta + k \sum_{\theta\in\Theta} p_0(\theta) \ln p_0(\theta)$, which means that a monopolist would choose to become informed by purchasing the perfectly informative signal structure. As with more firms the equilibrium beliefs are worse than the monopolist's, a profitable firm in this market must choose to become informed.

As the number of profitable firms in the market grows, equilibrium beliefs get worse. However, this may be uncertainty-reducing, as it increasingly concentrates the prior on the negative types; therefore, the mutual information-based costs of acquiring information may be decreasing.

Theorem 5 shows that the expected net benefit to acquiring information (rather than rejecting all applicants) as a function of the number of firms that acquire information is decreasing and eventually becomes negative.

It therefore shows that in a market with few firms, all will be profitable. For a larger market, it provides an upper bound on the number of profitable firms, though equilibria exist where mixed strategies make even fewer firms profitable. However, if firms only enter if they expect to make strictly positive profits, Theorem 5 provides a precise prediction about the number of firms.

Theorem 5 *When* $k\sum_{\theta\in\Theta} p_0(\theta)\ln p_0(\theta) > \sum_{\theta<0}\theta p_0(\theta)$, *there exists an* $\bar{N}\in\mathbb{R}$ *such that*

(a) in the unique equilibrium of the model with restricted information acquisition and $N<\bar{N}$ *firms, every firm gets a strictly positive payoff and*

(b) in every equilibrium of the model with restricted information acquisition and $N>\bar{N}$ *firms, at most* $\bar{N}$ *firms get a strictly positive payoff.*

Figure 2 displays, on the left, the payoffs to the strategies of rejecting the applicant without information, accepting the applicant without information, and becoming informed in order to accept if $\theta>0$. As the probability of the 'good' type varies, the optimal strategy changes. If the prior is concentrated, whether it is good or bad, it is not worth becoming informed; it is in the middle region where gathering information is optimal. The right-hand panel shows how choosing the optimal strategy as displayed on the left results in acceptance probabilities for each type, and the average acceptance probability. As the equilibrium belief becomes worse with the addition of more firms, so long as firms are profitable, their behavior does not change; therefore, the effect on adverse selection is not dampened by changes in equilibrium information acquisition.

4.3 Discussion

4.3.1 Comparison

The theorems in this section have shown that the information acquisition technology limits the number of profitable firms in the restricted information market, but not

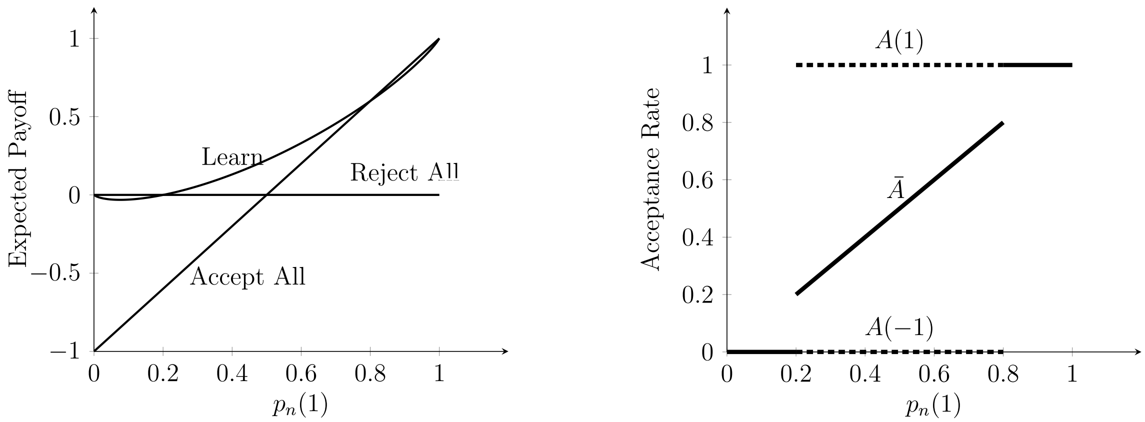

Figure 2: Restricted information acquisition model when $\Theta = \{-1, 1\}$. Payoff to each strategy as a function of probability that $\theta = 1$ (left); Optimal type-wise and average acceptance probability as a function of probability that $\theta = 1$ (right).

in the unrestricted information market. This results from the fact the restricted information market forces firms to buy 'chunky' information, which produces a larger negative externality on other firms. By contrast, the unrestricted information acquisition market allows for such fine-tuning that rejected applicants are not excessively bad draws - merely *marginally* so. This sort of fine-tuning is a property shared by a variety of continuous-time learning procedures, detailed in section 5, and is by no means an artefact of mutual information costs.

I assumed only a single applicant. It is useful to think of the applicant side of the market changing in size, and tracking the effect on the 'market' size. Crucially, both market size theorems address the profitability of firms *per applicant*; there are no fixed costs. Therefore, surprisingly, the market size prediction is invariant to the number of applicants when entry is free. If we consider costly entry, as the number of applicants grows, the restricted information market will see an increase in the number of active firms if initially less than $\lfloor \bar{N} \rfloor$; but once there, it remains there. The unrestricted information acquisition market with entry costs, on the other hand (considering the symmetric equilibrium at each N) will grow without bounds as the number of applicants increases.

4.3.2 Implicit Information Cascades

In the restricted information acquisition market, with too many (more than $\bar{N}$) firms, for some firms a visit leads to inference that the applicant is too likely to have been rejected elsewhere to be worth considering. Although the actions of other firms are not directly observable, information is carried by their strategies and Bayes' Rule. Furthermore, as the inferred action is rejection, and that action is then copied by the inactive firm, the effect is similar to an information cascade, as in Bikhchandani, Hirshleifer and Welch (1992). The inference, rather than the direct observation, of the action being copied gives rise to what might be called an *implicit information cascade*. Firms other than the $\bar{N}$ allowed by Theorem 5 'copy' the action they believe has been taken with high probability by other firms, without even paying to view their signal, in the same way that agents ignore their signals in BHW.

By contrast, such a cascade does not occur in the market with unrestricted information acquisition. Rejections are not as informative as in the restricted information market, and the additional type uncertainty generated by the visit order leaves some surplus for an additional firm.

Note that clearly, both technologies would lead to cascades if the visit order were known. In the restricted information case, an applicant known to have been rejected by another firm is known to have $\theta < 0$ and therefore not worth considering. Similarly, if a firm is observed to reject the applicant in the market with unrestricted information acquisition, the posterior $q(\cdot)$ for that applicant now has $\sum_{\theta \in \Theta} q(\theta) e^{\theta/k} = 1$ which means that the optimal action for all subsequent firms is to learn nothing and reject.

5 A theorem with continuous-time learning

Recent advances in the literature explore continuous-time learning for economic settings. Drift-diffusion models have been used to discuss two-alternative choice as in Fudenberg, Strack and Strzalecki (2015) and Hébert and Woodford (2016) consider a broad class of continuous time information acquisition technologies. I will present a theorem providing sufficient conditions for the outcome of Theorem 4 to apply in settings where firms have such a technology available. Moreover, by stating assumptions in terms of the properties of the solution to an unmodeled continuous-time information acquisition problem, I will arrive at a very general characterization, one that in fact implies Theorem 4.

5.1 Setting

In this section, I take a general statement of the outcome of a continuous time learning process without modeling it in detail. It may represent an exogenous process that the firm passively observes in preparation for a decision, or it may be driven by the firm optimally choosing how to learn. Details such as costs enter only implicitly. I denote by X the resulting stochastic process that describes in continuous time the firm's beliefs about θ given this learning. I give three conditions: that X's path is almost surely continuous with respect to time, that the firm makes its choice once X exits some convex continuation set that is not growing over time, and that a firm that engages in learning will in expectation make a profit. These conditions will be enough to ensure a general version of Theorem 4 holds.

The rather sparse assumptions on this continuous-time result allow for a variety of commonly used learning processes. Drift-diffusion models with constant (increasing) time costs, for instance, satisfy. The mutual information model described earlier in this paper can be sequentialized as well, in a way that fits these assumptions. These stipulations do not hold, however, if X is the result of a fully revealing Poisson signal - this latter case is the equivalent to the restricted information acquisition model.

5.2 Setup

Let $\{X_t : t \in \mathbb{R}_+\}$ be a continuous-time martingale process in the probability simplex over Θ denoting beliefs resulting from the firm's optimal acquisition process and the true value of θ. Its initial value is $x_0 = p_n$. The process results from some unmodeled policy that depends on the current value x_t, the time t, or both. Suppose furthermore that this optimal policy contains a stopping rule, given by a closed stopping region $\mathcal{S}(t) \subseteq \Delta\Theta$ that may vary with time. Assume that $\forall t, \mathcal{S}(0) \subseteq \mathcal{S}(t)$ and that $\mathcal{S}$ is a closed-graph correspondence of t. The stopping rule requires that the stopping time τ satisfies $\tau = \inf\{t | x_t \in \mathcal{S}(t)\}$. Once the stopping rule is triggered at some τ, the firm is assumed to make a decision to accept or reject the applicant based on the current beliefs x_τ. Assume furthermore that if $x_0 \notin \mathcal{S}(0)$, the firm's ex-ante expected payoff is positive.

5.3 Result

This section's result will rely, as did Theorem 4, on the intuition that rejections are in some sense *marginal*. To generate this property, assume that the path of X_t is almost surely a continuous function of t. Then, almost surely the posterior for rejections will be marginal, and a similar theorem holds.

Lemma 1 *If $p_n \in \Delta\Theta \setminus \mathcal{S}(0)$, if X_t is almost surely continuous in t, then almost surely for terminal τ, $x_\tau \in \overline{\Delta\Theta \setminus \mathcal{S}(0)}$.*

The optimal learning rule therefore implies that when learning terminates, the rejecting firm's beliefs about the applicant lie in the closure of the continuation set $\Delta\Theta \setminus \mathcal{S}(0)$.

Theorem 6 *If X_t is almost surely continuous, almost surely terminates and $\Delta\Theta \setminus \mathcal{S}(0)$ is convex, if $p_0 \in \Delta\Theta \setminus \mathcal{S}(0)$, for every $N \in \mathbb{N}$, in every equilibrium of the model with N firms, every firm gets a positive expected payoff.*

The theorem is proven by contradiction. If a n firm is inactive and gets 0 payoff, it must be that its initial beliefs p_n about applicants are in the stopping set. However, each firm's posterior about its rejected applicants will (almost surely) lie in the closure of the initial continuation set $\Delta\Theta \setminus \mathcal{S}(0)$ due to the continuity of the stochastic process. Then, p_n is just an appropriately-weighted convex combination of the prior p_0 which is in the interior of the initial continuation set, and points in the continuation set's closure. Given a convex initial continuation set $\Delta\Theta \setminus \mathcal{S}(0)$, p_n must lie in its interior, and by hypothesis a belief in the initial continuation set corresponds to a positive expected payoff. Figure 3 illustrates the theorem for an arbitrary process when $|\Theta| = 3$.

As the unrestricted information acquisition one-shot model with mutual information costs in this paper can be written as the output of optimal learning in a continuous-time model with mutual information-restricted learning per unit time and a constant time cost, Theorem 6 implies Theorem 4. The theorem also applies to a wide class of learning processes described in Hébert and Woodford (2016), drift diffusion models with non-decreasing time costs and other settings. To attain such generality, however, Theorem 6 must be written in terms of the solution, not the primitives, of some unspecified model. Nevertheless, an almost surely continuous

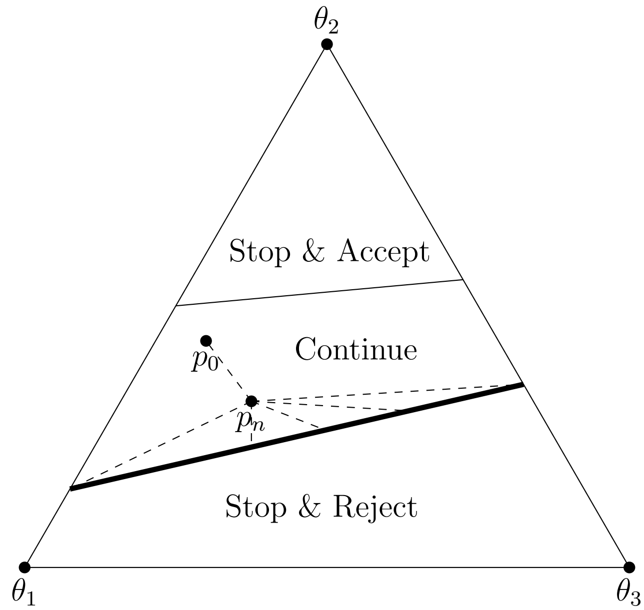

Figure 3: Equilibrium beliefs p_n for an inactive firm n are a convex combination of the prior p_0 and points on the marginal rejection line (bold).

path for the learning process and a convex continuation set are both easy to verify for a particular model. The market size result is therefore generalizable to a fairly broad setting and not particular to the unrestricted information acquisition process or a particular cost function.

Transporting the restricted information case into continuous time is also feasible, as it can be modeled by a fully revealing variable with a Poisson arrival process and constant time costs. Firms will choose to either accept the applicant without information, reject the applicant without information, or wait until the arrival of the fully informative signal[8], equivalently to the behavior in the restricted information one-shot case. Too many firms using the waiting strategy would push equilibrium beliefs for other firms into the region where profits cannot be attained; this limits the number of firms.

[8]Naturally, an indifferent firm could also employ a mixed strategy, or wait a finite amount of time, which is also equivalent to mixing.

6 Conclusion

In this paper I study a model of information acquisition in a search market. As each firm's actions affect the aggregate level of adverse selection, I explore how these externalities compound as the number of firms expands. Unexpectedly, the model shows that, with more firms, flexibility in information acquisition dampens the adverse selection externality each causes, as they will produce marginal posteriors. On the other hand, chunky information acquisition results in more pessimistic posteriors, so that with enough firms, adverse selection grows to the point of eradicating profits. I have argued that these results are relatively robust within a broad class of information acquisition technologies.

I have shown that as a result of this form of adverse selection, the information acquisition technology in a market can have strong implications about the number of firms that will be active. A rich information menu will allow all firms to be profitable, so with low enough entry costs many firms will enter. On the other hand, a chunky information acquisition technology will limit the number of potential entrants.

This research highlights the need to study the qualitative aspects of information acquisition as they appear in different contexts. It investigates what can loosely be considered a novel type of barrier to entry, and its determinants.

Finally, the model invites a broad class of applications in investment, labor, insurance, and other markets. As each market is ostensibly endowed with its own information acquisition technology, it is thus of paramount importance to consider the differential effects policies may have on otherwise similar markets.

References

Bergemann, Dirk, Xianwen Shi, and Juuso Välimäki. "Information acquisition in interdependent value auctions." Journal of the European Economic Association 7.1 (2009): 61-89.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. "A theory of fads, fashion, custom, and cultural change as informational cascades." Journal of political Economy (1992): 992-1026.

Denti, Tommaso. Unrestricted Information Acquisition. Working Paper, 2016.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki. Stochastic choice and optimal sequential sampling. Working Paper, 2015.

Geanakoplos, John, David Pearce, and Ennio Stacchetti. "Psychological games and sequential rationality." Games and economic Behavior 1.1 (1989): 60-79.

Gentzkow, Matthew, and Emir Kamenica. "Costly persuasion." The American Economic Review 104.5 (2014): 457-462.

Hébert, Benjamin, and Michael Woodford. Rational Inattention with Sequential Information Sampling. Working Paper, 2016.

Inderst, Roman. "Matching markets with adverse selection." Journal of Economic Theory 121.2 (2005): 145-166.

Lauermann, Stephan, and Asher Wolinsky. "Search with adverse selection." *Econometrica* 84.1 (2016): 243-315.

Matejka, Filip, and Alisdair McKay. "Rational inattention to discrete choices: A new foundation for the multinomial logit model." The American Economic Review 105.1 (2014): 272-298.

Persico, Nicola. "Information acquisition in auctions." Econometrica 68.1 (2000): 135-148.

Ravid, Doron. Bargaining with rational inattention. Working paper, 2014.

Shannon, C. E. "A Mathematical Theory of Communication." Bell Sys. Tech. J. 27 (1948): 379-423.

Sims, Christopher A. "Implications of rational inattention." Journal of Monetary Economics 50.3 (2003): 665-690.

Woodford, Michael. Inattention as a source of randomized discrete adjustment. Working Paper, 2008.

Yang, Ming. "Coordination with flexible information acquisition." Journal of Economic Theory 158 (2015): 721-738.

Zhu, Haoxiang. "Finding a good price in opaque over-the-counter markets." Review of Financial Studies 25.4 (2012): 1255-1285.

A Proofs

A.1 Proof of Theorem 2

Recall the *belief function* $\hat{p}_n$ that takes opponent type-wise acceptance probabilities $(A_m)_{m\neq n}$ into the Bayesian posterior for a visitng applicant at firm n and the function $\hat{A}_n$ that takes beliefs for firm n into optimal actions. The function $\hat{A}_n \circ \hat{p}_n$ thus takes opponent type-wise acceptance probabilities into (unique) optimal acceptance probabilities for firm n and is therefore a best-response function. As the composition of continuous functions, it is continuous. Forming the N-firm best response function $\hat{A} \circ \hat{p} : [0,1]^{|\Theta|N} \to [0,1]^{|\Theta|N}$ using each firm's best response functions, it inherits continuity and is defined over a compact[9], convex set. Therefore the Brouwer Fixed Point Theorem gives us the existence of a fixed point in $\hat{A} \circ \hat{p}$ and hence a Nash Equilibrium. ∎

A.2 Proof of Theorem 3

Once a firm acquires information, it will learn the applicant's type and accept iff $\theta > 0$. Therefore, there are only three pure strategies and their mixtures to consider:

[9]To be precise, the initial action space is not necessarily compact; however, best responses always fit Woodford (2008)'s description.

staying uninformed (choosing signal g_{no}, the completely uninformative signal structure) and rejecting all applicants, staying uninformed and accepting all applicants, and becoming informed to accept iff $\theta > 0$. Using formula (2) and noticing g_{no} offers no information and therefore $H(p_n) = H(p_n(\cdot|s))$ when s is drawn according to g_{no}, it is the case that $M(p_n, g_{no}) = 0$. Therefore, remaining uninformed and rejecting the applicant gives a payoff of 0. Accepting the applicant without information results in a payoff of

$$\sum_{\theta\in\Theta} \theta p_n(\theta) - kM(p_n, g_{no}) = \sum_{\theta\in\Theta} \theta p_n(\theta). \tag{12}$$

Becoming informed by choosing g_{all} (the completely informative signal structure) eliminates all uncertainty in the posterior, and therefore $H(p_n(\cdot|s)) = 0$; therefore $M(p_n, g) = H(p_n) = -\sum_{\theta\in\Theta} p_n(\theta) \ln p_n(\theta)$. Thus, becoming informed and accepting iff $\theta > 0$ results in a payoff of

$$\sum_{\theta>0} \theta p_n(\theta) - kM(p_n, g_{all}) = \sum_{\theta>0} \theta p_n(\theta) + k \sum_{\theta\in\Theta} p_n(\theta) \ln p_n(\theta). \tag{13}$$

First, suppose that the equilibrium belief for each firm n, p_n, is equal to the prior p_0. Then, if (12) is greater than (13), as (9) implies that if $\forall m \neq n, \forall\theta\ A_m(\theta) = 1$, and that in turn delivers that for each n, $\hat{p}_n((A_m)_{m\neq n}) = p_0$. Therefore, we have an equilibrium in pure strategies where $\forall n, g_n = g_{no}$ and $\forall s, a_n(s) = 1$. That is, it is an equilibrium that no firm becomes informed and all applicants are always accepted.

Otherwise, uninformed firms will reject the applicant. Each firm chooses a probability z_n of becoming informed. The set of optimal information acquisition strategies, as a function of p_n, is therefore given by

$$Z_n(p_n) = \begin{cases} \{0\} & \sum_{\theta>0} \theta p_n(\theta) + k \sum_{\theta\in\Theta} p_n(\theta) \ln p_n(\theta) < 0 \\ [0,1] & \sum_{\theta>0} \theta p_n(\theta) + k \sum_{\theta\in\Theta} p_n(\theta) \ln p_n(\theta) = 0 \\ \{1\} & \sum_{\theta>0} \theta p_n(\theta) + k \sum_{\theta\in\Theta} p_n(\theta) \ln p_n(\theta) > 0 \end{cases}$$

which is convex-valued and closed-graph.

Optimal actions following information acquisition for firm n satisfy $\forall\theta < 0, A_n(\theta) = 0$ and $\forall\theta > 0, A_n(\theta) = z_n$. Thus, using (A.2) we can write the set of optimal strategies in terms of $\hat{A}_n$, a convex-valued, closed-graph correspondence in p_n. Recall that $\hat{p}_n$ given by (9) is continuous. Therefore, $\hat{A}_n \circ \hat{p}_n : [0,1]^{N-1} \to 2^{[0,1]}$ mapping opponent acceptance rates into optimal own acceptance rates is a best-response correspondence for firm n and as a composition of a closed-graph, convex-valued correspondence on a continuous function, in turn convex-valued and closed-graph. This property is inherited by the N-firm best-response correspondence $\hat{A} \circ \hat{p} : [0,1]^N \to 2^{[0,1]^N}$ and hence as $[0,1]^N$ is a compact, convex subset of $\mathbb{R}^N$, the Kakutani Fixed Point Theorem applies, providing the existence of a fixed point and thereby an equilibrium. ■

A.3 Proof of Theorem 4

The proof proceeds inductively. Theorem (1) establishes that that profitability for firm n implies that $\sum_\theta p_n(\theta)e^{\frac{\theta}{k}} > 1$; a monopolist has $p_n = p_0$.

Let $\hat{p}_m^N$ be the belief function for firm m of N firms as defined in (9), and let $\hat{A}_m^N : \Delta\Theta \to [0,1]^{|\Theta|}$ be firm m of N's optimal acceptance probability function, given by Theorem 1, so that $\hat{A}_m^N \circ \hat{p}_m^N$ is firm m's best response function in the N-firm game. Let $\hat{A}^N \circ \hat{p}^N$ be the N-firm best response function, so that an equilibrium corresponds to a fixed point of the function. For each $N \in \mathbb{N}$ such a fixed point is guaranteed to exist by Theorem 2.

Suppose, for induction, that every fixed point of $\hat{A}^N \circ \hat{p}^N$ has $\forall \theta, n$, $A_n(\theta) > 0$ and therefore positive expected profits by Theorem 1. Recall that if in equilibrium $A_n(\theta) > 0$, firm n gets a strictly positive payoff. Assume, for contradiction, an equilibrium given by $(A_n, p_n)_{n \leq N+1}$ in the market with $N+1$ firms where firm $N+1$ gets a payoff of 0. Therefore $(A_n)_{n \leq N+1}$ is a fixed point of $\hat{A}^{N+1} \circ \hat{p}^{N+1}$ and $\forall \theta, A_{N+1}(\theta) = 0$. Thus, using (9), for each firm $m \leq N+1$, $\hat{p}_m^{N+1}((A_n)_{n \leq N+1}) = \hat{p}_m^N((A_n)_{n \leq N})$. Therefore, as $\hat{A}_n^N = \hat{A}_n^{N+1}$, $(A_n)_{n \leq N}$ is a fixed point of $\hat{A}^N \circ \hat{p}^N$. As by the inductive hypothesis every such fixed point has $\forall \theta \in \Theta, \forall n \leq N, A_n(\theta) > 0$, each firm $n \leq N$ is getting a strictly positive payoff in the $N+1$ firm equilibrium as well.

What remains to be shown is that firm $N+1$ has a profitable deviation. As each firm $n \leq N$ gets a positive payoff, from Theorem (1),

$$\sum_{\theta \in \Theta} p_n(\theta) e^{\frac{\theta}{k}} > 1 \tag{14}$$

and A_n, the acceptance probability vector for n, if not identically 1, is given by

$$\frac{A_n(\theta)}{1-A_n(\theta)} = \frac{\bar{A}_n}{1-\bar{A}_n} e^{\frac{\theta}{k}} \tag{15}$$

$$\bar{A}_n = \sum_{\theta\in\Theta} p_n(\theta) A_n(\theta). \tag{16}$$

Combining the above, and manipulating, we get

$$\bar{A}_n = \sum_{\theta\in\Theta} \frac{p_n(\theta)\frac{\bar{A}_n}{1-\bar{A}_n} e^{\frac{\theta}{k}}}{1+\frac{\bar{A}_n}{1-\bar{A}_n} e^{\frac{\theta}{k}}} \tag{17}$$

$$1 = \sum_{\theta\in\Theta} \frac{p_n(\theta) e^{\frac{\theta}{k}}}{1-\bar{A}_n + \bar{A}_n e^{\frac{\theta}{k}}} \tag{18}$$

The distribution of types rejected by firm $n \leq N$, q_n, is given by:

$$q_n(\theta) = \frac{p_n(\theta)(1-A_n(\theta))}{\sum_{\theta'\in\Theta} p_n(\theta')(1-A_n(\theta'))} = \frac{p_n(\theta)(1-A_n(\theta))}{1-\bar{A}_n}. \tag{19}$$

Substituting for $A_n(\theta)$ from (15) we get

$$q_n(\theta) = \frac{p_n(\theta)}{\left(1+\frac{\bar{A}_n}{1-\bar{A}_n} e^{\frac{\theta}{k}}\right)(1-\bar{A}_n)} = \frac{p_n(\theta)}{1-\bar{A}_n+\bar{A}_n e^{\frac{\theta}{k}}} \tag{20}$$

and using (18) we have

$$\sum_{\theta\in\Theta} q_n(\theta) e^{\frac{\theta}{k}} = \sum_{\theta\in\Theta} \frac{p_n(\theta) e^{\frac{\theta}{k}}}{1-\bar{A}_n+\bar{A}_n e^{\frac{\theta}{k}}} = 1 \tag{21}$$

As A_{N+1} is identically 0 by hypothesis, the distribution of rejected types for $n \leq N+1$ is independent of whether the applicant has visited n. Therefore, we can write the type distribution of applicants visiting firm $N+1$, p_{N+1} as a weighted sum of the distribution of rejected applicants at firms $n \leq N$ and the new applicant distribution p_0:

$$p_{N+1}(\theta) = \sum_{n\leq N} \phi_n q_n(\theta) + (1 - \sum_{n\leq N} \phi_n) p_0(\theta) \tag{22}$$

where ϕ_n is the probability an applicant visited firm n last before visiting firm $N+1$, conditional on firm $N+1$ being visited. If a firm sets $\bar{A}_n = 1$, it has no rejected applicants and thus for that n, $\phi_n = 0$. The remaining probability $1 - \sum_{n \leq N} \phi_n$, is the probability (conditional on $N+1$ being visited) that firm $N+1$ is the first firm visited; note $1 - \sum_{n \leq N} \phi_n \geq \frac{1}{N+1}$. Then,

$$\sum_{\theta \in \Theta} p_{N+1}(\theta) e^{\frac{\theta}{k}} = \sum_{n \leq N} \phi_n \sum_{\theta \in \Theta} q_n(\theta) e^{\frac{\theta}{k}} + (1 - \sum_{n \leq N} \phi_n) \sum_{\theta \in \Theta} p_0(\theta) e^{\frac{\theta}{k}} \tag{23}$$

so that substituting (19) we get

$$\sum_{\theta \in \Theta} p_{N+1}(\theta) e^{\frac{\theta}{k}} = \sum_{n \leq N} \phi_n \sum_{\theta \in \Theta} q_n(\theta) e^{\frac{\theta}{k}} + (1 - \sum_{n \leq N} \phi_n) \sum_{\theta \in \Theta} p_0(\theta) e^{\frac{\theta}{k}} \tag{24}$$

and using (21) we attain

$$\sum_{\theta \in \Theta} p_{N+1}(\theta) e^{\frac{\theta}{k}} = \sum_{n \leq N} \phi_n + (1 - \sum_{n \leq N} \phi_n) \sum_{\theta \in \Theta} p_0(\theta) e^{\frac{\theta}{k}}. \tag{25}$$

Now, since by hypothesis $\sum_{\theta \in \Theta} p_0(\theta) e^{\frac{\theta}{k}} > 1$, we have that

$$\sum_{\theta \in \Theta} p_{N+1}(\theta) e^{\frac{\theta}{k}} > 1 \tag{26}$$

but this is the condition under which Theorem 1 guarantees firm $N+1$ a positive payoff, a contradiction. Therefore, all firms in the $N+1$ firm market get positive payoff. By PMI, this holds for all $N \in \mathbb{N}$, and the theorem is proven. ∎

A.4 Proof of Theorem 5

First, suppose the payoff to becoming informed at the prior $\Pi = \sum_{\theta > 0} p_0(\theta)\theta + k \sum_{\theta \in \Theta} p_0(\theta) \ln p_0(\theta)$ is weakly negative. Then no firm (even a monopolist) can ever profit and $\bar{N} = 0$ trivially satisfies both sides. The rest of the proof assumes that the monopolist would make strictly positive profits..

Firms always reject known $\theta < 0$, and by hypothesis $k \sum_{\theta \in \Theta} p_0(\theta) \ln p_0(\theta) > \sum_{\theta < 0} \theta p_0(\theta)$ so that acceptance without information is not optimal; therefore, no $\theta < 0$ is ever accepted. Thus, we can write the equilibrium posterior for firm n as a function of x, the total probability that a $\theta > 0$ applicant was not accepted at a previous firm:

$$p_n(\theta) = \begin{cases} \frac{x p_0(\theta)}{x\sum_{\theta>0} p_0(\theta)+\sum_{\theta<0} p_0(\theta)} & \theta > 0 \\ \frac{p_0(\theta)}{x\sum_{\theta>0} p_0(\theta)+\sum_{\theta<0} p_0(\theta)} & \theta < 0. \end{cases} \tag{27}$$

The payoff to becoming informed by choosing g_{all} (and accepting $\theta > 0$) at p_n is

$$x\sum_{\theta>0} \theta p_0(\theta) - kH(p_n(\theta)) \tag{28}$$

so that substituting in (27) and manipulating, we get

$$\Pi(x) = \left[x\sum_{\theta>0} p_0(\theta) + \sum_{\theta<0} p_0(\theta)\right]^{-1} \cdot$$
$$\left[x\sum_{\theta>0} p_0(\theta)\theta + xk\sum_{\theta>0} p_0(\theta)\left(\ln x + \ln p_0(\theta)\right) + k\sum_{\theta<0} p_0(\theta)\ln p_0(\theta)\right]$$
$$-k\ln\left(x\sum_{\theta>0} p_0(\theta) + \sum_{\theta<0} p_0(\theta)\right).$$

Notice $\Pi(0) \leq 0$ and that, as $x = 1$ corresponds to the prior and it is assumed that $\sum_{\theta>0} p_0(\theta)\theta + \sum_{\theta\in\Theta} p_0(\theta)\ln p_0(\theta) > 0$, we have $\Pi(1) > 0$. Now, consider the derivative of the payoff to becoming informed wrt x:

$$\Pi'(x) = \left[\left(\sum_{\theta<0} p_0(\theta)\right)\left(\sum_{\theta>0} p_0(\theta)(\theta + k\ln p_0(\theta))\right) - \left(\sum_{\theta>0} p_0(\theta)\right)\left(\sum_{\theta<0} p_0(\theta)k\ln p_0(\theta)\right)\right.$$
$$\left. + k\ln x\cdot(\sum_{\theta>0} p_0(\theta))(\sum_{\theta<0}(p_0\theta))\right]\left[x\sum_{\theta>0} p_0(\theta) + \sum_{\theta<0} p_0(\theta)\right]^{-2}$$

As $\Pi'(x)$ limits to $-\infty$ as $x \to 0$ from the right, as it has at most a single root in x, given that $\Pi(0) \leq 0$ and $\Pi(1) > 0$, the payoff to becoming informed has a single root in $(0, 1)$, at a point I denote $\bar{x}$.

Set $\bar{N} = \frac{1}{\bar{x}}$. For (a), suppose there are $N < \bar{N}$ firms in the market. Then, from the proof of Theorem 3, for $\theta < 0$, $A_m(\theta) = 0$ and for $\theta > 0$, $A_m(\theta) = z_m$. Beliefs for firm n are given therefore given by (27) where $x \geq \frac{1}{N}$ (the value taken when all other firms set $z_m = 1$). Therefore from $\frac{1}{N} > \frac{1}{\bar{N}} = \bar{x}$, we have $\Pi(\frac{1}{N}) > 0$ and thus firm n must necessarily profit.

For (b), suppose there are $N > \bar{N}$ firms getting positive payoff in an equilibrium. A firm making a profit uses a strategy with $z_m = 1$ (due to non-indifference) so that the equilibrium beliefs of firm n are given by (27) where $x \leq \frac{1}{N} < \frac{1}{\bar{N}} = \bar{x}$. Therefore, $\Pi(\frac{1}{N}) < 0$, so that firm n cannot get a positive profit, a contradiction. ■

A.5 Proof of Lemma 1

Without loss of generality, take a realization x of X_t, written as $x : \mathbb{R}_+ \to \Delta\Theta$ and denote its stopping time by τ. Supposing x is continuous, and since $\tau = \inf\{t|t \in \mathcal{S}(t)\}$, $x(0) = p_n \in \Delta\Theta \setminus \mathcal{S}(0)$ and $\mathcal{S}$ is closed-graph, we have $\tau \neq 0$. As $\forall t, \mathcal{S}(0) \subseteq \mathcal{S}(t)$, the image under x of $[0, \tau)$ satisfies $x([0, \tau)) \subseteq \Delta\Theta \setminus \mathcal{S}(0)$. If x is continuous, the image under x of a set's closure is contained in the closure of its image; $\overline{\Delta\Theta \setminus \mathcal{S}(0)} \supseteq x(\overline{[0, \tau)}) = x([0, \tau])$. Thus, as the realized path x is almost surely continuous, $x_\tau \in \overline{\Delta\Theta \setminus \mathcal{S}(0)}$ almost surely.

A.6 Proof of Theorem 6

Let $a_n : \Delta\Theta \to [0, 1]$ denote acceptance probability for each posterior once the process X has stopped. Let $\hat{A}_n^N : \Delta\Theta \to [0, 1]^{|\Theta|}$ be the function mapping initial beliefs p_n for firm n of N into an ex-ante vector of type-specific acceptance probabilities. Additionally use $\hat{p}_n^N$ as given by (9), the belief function for player n of N, to map opponents' acceptance probabilities into beliefs for firm n. Write the N-firm acceptance probability function as $\hat{A}^N : (\Delta\Theta)^N \to [0, 1]^{|\Theta|N}$ and the N-firm belief function as $\hat{p}^N : [0, 1]^{|\Theta|N} \to (\Delta\Theta)^N$; an equilibrium for N firms then corresponds to each fixed point of $\hat{A}^N \circ \hat{p}^N$.

For each firm n, if $p_n \in \Delta\Theta \setminus \mathcal{S}(0)$, write as q_n the average over X_t's paths posterior distribution for rejected applicants. As from Lemma (1) each posterior is almost surely in $\overline{\Delta\Theta \setminus \mathcal{S}(0)}$ and that set is convex as the closure of a convex set, $q_n \in \overline{\Delta\Theta \setminus \mathcal{S}(0)}$.

The proof will proceed with induction on the number of firms. With 1 firm in the market, its beliefs are $p_1 = p_0 \in \Delta\Theta \setminus \mathcal{S}(0)$ so the firm will get a positive payoff. Assume for induction that every equilibrium with N firms gives positive payoff to every firm; that is, $\forall n \leq N, p_n \in \Delta\Theta \setminus \mathcal{S}(0)$ or A_n identically 1.

Assume for contradiction that in some equilibrium described by $(A_n)_{n \leq N+1}$ of the market with $N + 1$ firms, firm $N + 1$ gets weakly negative expected payoff.

By equilibrium, $\hat{A}^{N+1} \circ \hat{p}^{N+1}((A_n)_{n \leq N+1}) = (A_n)_{n \leq N+1}$. As A_{N+1} is identically 0, $\hat{p}^{N+1}((A_n)_{n \leq N+1}) = \hat{p}^{N+1}((A_n)_{n \leq N}, \vec{0})$. Using (9), we have $\hat{p}^{N+1}((A_n)_{n \leq N}, \vec{0}) = (\hat{p}^N(A_n^N)_{n \leq N}, p_{N+1})$ for some p_{N+1}; so $(A_n)_{n \leq N}$ is a fixed point of $\hat{A}^N \circ \hat{p}^N$ and therefore defines beliefs in an N firm equilibrium.

From the inductive hypothesis, in each such N-firm equilibrium, each firm $n \leq N$ is getting a strictly positive expected payoff. Thus for each firm n either A_n is identically 1 or $p_n \in \Delta\Theta \setminus S(0)$ and therefore Lemma (1) applies to each $n \leq N$. Then, firm $N+1$ is either receiving an applicant last rejected by firm n where A_n is not identically 1, or an applicant that has never visited another firm. Therefore

$$p_{N+1} = \sum_{n \leq N : \bar{A}_n \neq 1} \phi_n q_n(\theta) + \left(1 - \sum_{n \leq N : \bar{A}_n \neq 1} \phi_n\right) p_0(\theta) \tag{29}$$

for some probability weights ϕ, where $\sum_{n \leq N : \bar{A}_n \neq 1} \phi_n \leq N/(N+1)$. As $p_0 \notin \mathcal{S}(0)$ and $\mathcal{S}(0)$ closed, $p_0 \in int(\Delta\Theta \setminus \mathcal{S}(0))$.

But as Lemma (1) gives that $q_n \in \overline{\Delta\Theta \setminus \mathcal{S}(0)}$, as $p_0 \in int(\Delta\Theta \setminus \mathcal{S}(0))$ and $\Delta\Theta \setminus \mathcal{S}(0)$ is convex,

$$p_{N+1} \in \Delta\Theta \setminus \mathcal{S}(0) \tag{30}$$

and therefore firm $N+1$ gets a strictly positive payoff, a contradiction. Thus, by PMI, for any $N \in \mathbb{N}$, in every equilibrium of the market with N firms, each firm gets strictly positive profit, proving the theorem. ■