Adverse Selection and Endogenous Information*

João Thereze †

July 28, 2023

[Click here for the latest version]

Abstract

This paper studies adverse selection markets in which consumers can choose to learn how much they value a product. Information is acquired after observing prices, so it is endogenous. This presents a trade-off: information increases the quality of consumers' choices but worsens selection. We characterize how this trade-off translates into distortions of consumers' demand and producers' cost. We show that information decisions produce a negative externality, may decrease welfare, and may lead to new forms of market breakdown. Moreover, efficiency is typically non-monotone in information costs. Two implications are that (1) standard measures underestimate the welfare costs of adverse selection; and (2) information policies can help correct its inefficiencies. Finally, we construct an empirical test to detect endogenous information in the data, and develop a framework for counterfactual policy analysis.

*I am indebted to Pietro Ortoleva, Alessandro Lizzeri and Sylvain Chassang for their continuous guidance throughout this project. I am grateful to Roland Bénabou, Kate Ho, Wolfgang Pesendorfer, Leeat Yariv, Eduard Boehm, Francesco Fabbri, Andrew Ferdowsian, Sam Kapon and Rafael Parente for valuable comments.

†The Fuqua School of Business, Duke University, Durham, NC 27701 — joao.thereze@duke.edu.

1 Introduction

This paper studies adverse selection when consumers can exert effort to learn how much they value a product. As is often the case, consumers decide what information to acquire *after* observing prices. In this setting, information is endogenous: by choosing prices, firms affect consumers' incentives to learn. For example, depending on how much a good costs, buyers may spend more or less time inspecting its details. We show the effects of endogenous information are substantial: endogeneity adds to the usual distortions due to adverse selection, thus harming consumers. On the other hand, information policies may help to mitigate inefficiencies. Moreover, we provide a set of tools for bridging the theoretical analysis to data.

Endogenous information acquisition affects a range of economic decisions. Indeed, evidence from several markets suggests that endogenous information is important in explaining individual behavior (Brown and Jeon, 2020; Kacperczyk et al., 2016; Taubinsky and Rees-Jones, 2018). Consider consumers deciding whether or not to purchase a health insurance plan. They first observe its premium, which is readily available, and then choose how much information to gather about its value. Information acquisition may take several forms: a patient with a heart condition may spend time investigating whether specific medications are covered by the health-care plan, whereas another individual may prefer to collect family history data before signing the contract. In either case, consumers' information determines who buys the insurance product and, consequently, insurers' costs. This interaction between endogenous information and the degree of selection is at the core of this paper.

We develop a tractable framework to combine endogenous information and adverse selection. Following typical assumptions, also found in the empirical literature, we consider suppliers who offer a homogeneous good and compete by setting prices (Einav et al., 2010). The market is adversely selected, in that buyers with higher valuation impose higher costs to the firm. This is the case in insurance markets, where the riskiest consumers are those who value coverage the most. In keeping with the insurance example, for most of the paper, we maintain that it is efficient for providers to serve all individuals. We depart from traditional models by allowing consumers to flexibly and costly acquire further information about their valuation after observing prices. To the extent that consumers' information is correlated with providers' costs, acquisition will affect selection as well as demand. We show this has significant implications on market efficiency.

Results. We start by characterizing how information acquisition affects aggregate market behavior. Consumers' decisions depend on the details of acquisition costs and can be difficult to study. Our approach circumvents this difficulty by focusing on the properties of demand and firms' cost curves that are robust to all typical costs of information. First, we note that these curves depend on the distribution of information in the market. While this distribution is exogenous in standard models of adverse selection, here it endogenously changes with prices. This distinction plays a key role in understanding the effect of endogenous information. Second, we show how changes in information costs lead to "rotations" of demand and firms' costs. As information costs increase, demand falls for high prices because less informed consumers are not willing to purchase an expensive product, and vice versa: demand grows for low prices. By contrast, firms' costs always decrease when information becomes more expensive, as some less informed consumers end up purchasing the product even when their valuation is low. Because demand and firms' costs alone identify the equilibrium, this characterization allows us to apply familiar tools to analyze the model.

With this characterization, we prove three results. First, we show efficiency often varies with information costs non-monotonically: as costs grow, efficiency first decreases, but it increases for high costs. In some situations, the non-monotonicity is so severe that an initial increase in information costs can induce a complete market breakdown, but a further increase will bring the economy to full efficiency. This result reflects a *trade-off*: on the one hand, information costs lower the quality of consumers' choices; on the other hand, they reduce the amount of private information in the market, thus alleviating selection. For a large class of information costs, the first effect dominates when acquisition is cheap, as small costs deteriorate consumers' decision-making abilities, but have little effect on sorting. As acquisition becomes expensive, no private information is acquired and the equilibrium approaches efficiency.

Our second result is that endogenous information hurts consumers. Under costly acquisition, equilibrium prices and quantities are further away from the efficient ones, when compared with exogenous information. Intuitively, by fine-tuning their information to the firms' offers, consumers improve the sorting of buyers to the product, thus worsening selection. As a consequence, there often exist prices that would benefit both buyers and sellers if information was fixed, but that become unprofitable for firms if consumers adjust the information they acquire. Because firms anticipate consumer behavior, those prices are ruled out in equilibrium. In other words, buyers would like to commit to *not* calibrating their information to market conditions, but they cannot. This worsening of selection has additional consequences: when information is endogenous, the market may unravel completely, even though trade is guaranteed to happen when information is exogenous, regardless of the distribution of consumers' beliefs. These results shed light on the substantial distinction between models of exogenous and endogenous information.

Third, we show too much information is acquired in equilibrium. For any level of demand, a benevolent planner would prefer consumers to learn less than they do in the decentralized market. This is because information decisions produce an externality in the presence of adverse selection. While consumers do internalize the benefit of information in decision-making, they do not acknowledge that their information affects sorting and, therefore, firms' costs. This information externality adds to the welfare losses from adverse selection. When private information is exogenous, subsidies can restore market efficiency by correcting the allocative distortion caused by selection. By contrast, subsidies alone cannot restore efficiency when information is endogenously acquired, because information decisions are socially sub-optimal as well. As a consequence, policies that target acquisition incentives have a role in mitigating the effects of adverse selection.

Implications. Our results have several implications for empirical work and policy analysis. First, we show endogenous information can be detected empirically even when no data is available about beliefs or information costs. We illustrate this by proposing a concrete test. Consider a researcher who observes the consumer demand and firms' cost curves, but nothing else — this is common in the empirical industrial organization, where these curves are estimated from individual-level data (Einav et al., 2021). Under the null hypothesis that there is no endogenous information acquisition, the researcher can construct a counterfactual curve for firm's costs from the demand curve. Our test then consists of comparing this counterfactual cost curve with the observed one. We prove that if endogenous information is present, the counterfactual curve must be lower than the observed one. This provides a simple test to estimate the presence of endogenous information acquisition and its severity, which can be implemented on existing datasets. The test exploits the worsening of adverse selection under endogenous information. Indeed, the same demand curve implies different levels of sorting depending on whether information is exogenously available or endogenously acquired.

Second, we provide tools for evaluating possible policy interventions. Governments may want to help or hinder information acquisition, however, our non-monotonicity results suggest the effect of such policies is unclear: they may increase or reduce consumer welfare. In general, addressing this question requires knowledge on how information costs affect both consumer demand and firms' costs, which may be unobservable. To circumvent this issue, we show that one can evaluate these welfare consequences using a simple sufficient statistic. The formula we derive connects the observable elasticity of firms' costs to the unobservable elasticity of demand with respect to information costs. This formula bounds the welfare effects of policies that manipulate consumer acquisition, and can be computed using the demand and cost curves. Both this formula and the detection tests require knowledge of how consumer valuations map into firms' costs. We briefly discuss how to identify these parameters in standard individual-level data.

Third, we show how the usual empirical estimates may fail to account for the welfare costs of adverse selection. In standard models, in which information is exogenous, all inefficiency can be corrected by subsidies, that move allocations along the demand curve. Following Einav et al. (2010), a large literature quantifies welfare costs in this way, using only the demand and firms' costs. In our model, this approach fails because of the information externality. Because the socially optimal information does not coincide with that acquired in the market, obtaining the optimal allocation would require not only subsidies, but changes in information decisions. Consequently, the observed curves are not sufficient for welfare analysis. In fact, we prove that the standard measurement never exceeds, and often underestimates, the real welfare costs.

We conclude by applying our model to study insurance take-up. In the context of the Massachusetts subsidized insurance exchange, Finkelstein et al. (2019) find that low insurance take-up is mostly explained by consumers' valuations being far lower than the cost they impose on insurers, rather than by adverse selection. This finding is at odds with textbook models of insurance, in which consumer valuation always exceeds firms' marginal costs, and full market coverage is efficient. We argue that costly information acquisition may partially rationalize this low willingness-to-pay. Indeed, when information is costly, the revealed willingness-to-pay of a large number of consumers can be far below their real valuation.

We then discuss how government policies based on information can increase insurance coverage. When the observed willingness-to-pay is low, the cost of subsidies designed to expand take-up can be substantial. However, if endogenous information is a large part of the inefficiency, the government may be able to improve welfare by providing information instead of subsidies. We show this is the case. Concretely, we prove that, by first providing information to consumers, a planner can reduce the subsidies required to attain a given market coverage, thus reducing the total cost for buyers. Intuitively, the planner can dampen consumers' incentives to acquire information that would worsen selection by providing them with *different* information. The example highlights how acknowledging endogenous information provides a new dimension of policies to mitigate inefficiencies in markets with adverse selection.

Related Literature. This paper relates to several strands of the literature. First, it contributes to applications of rational inattention in market contexts (Cusumano et al., 2022; Hefti, 2018; Martin, 2017; Matějka and McKay, 2012). We complement this literature by showing the impact of endogenous information acquisition in markets with adverse selection. The paper is close to both Mensch and Ravid (2022) and Thereze (2022), which study a monopolist designing a product line when consumers can learn their valuations after observing offers. By contrast, we investigate the interaction between costly acquisition and adverse selection, and focus on a competitive setting. Thereze (2022) shows that consumers may obtain a higher surplus in equilibrium due to the threat of acquiring information that is not optimal for the seller. In Section 8, we discuss how our model extends to a monopoly market, and argue that this information threat plays a similar role in this less complex environment.

Our model adds to a literature that studies endogenous information in the market for lemons (Dang, 2008; Pavan and Tirole, 2022). In Pavan and Tirole (2022), the potentially informed agent acquires information *before* deciding whether to interact with the uninformed agent, rather than after. Their key contribution is to provide conditions for expectation conformity in this class of games. While they focus on the effects of signaling, we study how the lack of commitment exacerbates adverse selection. Dang (2008) studies a take-it-or-leave-it bargaining model in which both agents in the negotiation may acquire information about the value of the good. The main finding is that the subject responding to the offer may accrue positive rents. Instead, we investigate the case in which information acquisition is one-sided and the product is traded in a competitive market, and show that the endogeneity of information hurts consumers.

A vast literature in industrial organization also studies adverse selection in models of willingness-to-pay and price competition (Einav et al., 2010; Finkelstein et al., 2019; Mahoney and Weyl, 2017; Panhans, 2019). Einav et al. (2010) develop a framework suitable to detect and measure the effects of adverse selection using the demand curve and firms' costs. Our model extends their framework to settings with endogenous information, allowing for policy and welfare evaluation in the information dimension, and suggesting areas where traditional methods may fall short when consumers can acquire information. Our paper is also connected to the literature that follows Johnson and Myatt (2006) and applies demand rotations to a range of problems from advertising to product design, and intermediation (Bar-Isaac et al., 2012; Bergemann and Bonatti, 2011; Inderst and Ottaviani, 2012; Johnson, 2013). We prove changes in information costs microfound the rotation of demand costs, and link this literature with adverse selection.

This paper is also related to the growing body of applied work that studies how behavioral frictions affect insurance markets (Brown and Jeon, 2020; Handel, 2013; Handel and Kolstad, 2015; Landais and Spinnewijn, 2021). We differ from that literature in that consumers make optimal choices subject to costly information acquisition, endogenizing information frictions and disciplining its effect in equilibrium. While that literature has acknowledged, long before us, that the presence of behavioral frictions invalidates the traditional approach for welfare measurement (Handel et al., 2019; Spinnewijn, 2017), our findings are in contrast. When consumers optimally choose what information to acquire, as in this paper, the demand curve still quantifies consumer surplus given information decisions, and is thus sufficient to evaluate the welfare effects of subsidy policies.

2 A Model of Adverse Selection with Endogenous Information

Two competitive firms aim to sell goods to a continuum of potential buyers with unit demand. Products are indivisible and homogeneous. It is common knowledge that consumers' valuations, denoted by ω, are drawn independently, according to a continuous distribution $F_o \in \Delta[0,1]$ with mean μ, and a log-concave and continuously differentiable density f_o.[1] Firms face a cost $\chi(\omega) = \alpha \cdot \omega$ to provide one unit of the product to a consumer with valuation ω, with $\alpha \in (0,1)$. Note that serving consumers with a higher willingness-to-pay is more expensive to the firm, which introduces adverse selection into the model. Moreover, $\alpha < 1$ guarantees gains from trade with all buyers, which is the usual assumption in insurance markets, for example. Our framework can be readily extended to cases in which there is no gain in trading with some buyers, and we discuss how to generalize firms' costs, χ, to arbitrary increasing functions, in Section 8. A consumer with valuation ω who purchases the good at price p receives utility $\omega - p$.

At the beginning of the game, each consumers' valuation is unknown, both to consumers and to the firms. After firms post their prices, buyers can choose to privately acquire information about their valuations. Formally, each individual can choose an information structure $(\mathcal{S}, \mathbf{P})$, which consists of a set of signal realizations $\mathcal{S}$ and a function $\mathbf{P} : [0,1] \to \Delta(\mathcal{S})$ that assigns to each valuation $\omega \in [0,1]$ a distribution over signal realizations. The agent is free to choose any information, but learning is costly, as described below. The assumption of symmetric information at the start of the problem is for ease of exposition: in Section 8, we show that our results can be extended to accommodate buyers who have some private information from the start.

The game unfolds as follows: first, each consumer's willingness-to-pay is drawn from F_o and then firms post prices. After observing prices, buyers decide what information to acquire. Upon observing a signal from their information structures, consumers make purchasing decisions. We study symmetric, Subgame Perfect Nash Equilibria in pure strategy — henceforth, denoted as 'equilibria'.

Example (An insurance market). The setup above is close to a standard insurance setting. To make the mapping explicit, consider an insurance market that develops in two periods. In period one, consumers are interested in insuring against a loss of $\ell > 0$ monetary units that may happen in period two. For each individual, the probability of the loss is $r \in [0,1]$, drawn independently from some common knowledge distribution. As in a standard insurance setting, consumers are identical, except for their risk. We assume they have common wealth w and share a utility u over second period consumption, but have quasilinear preferences with respect to first period consumption. There is no discounting.

In the first period, firms offer an insurance contract, covering a fraction $c \in (0,1)$ of the loss in case of accident and charging an upfront premium of p dollars. Thus, a consumer with loss probability r is willing to pay a premium for the insurance plan that equates their second period expected utility *with* coverage to their second period expected utility *without* coverage. For a consumer with loss probability r,

[1] Recall that a positive function f is log-concave if $\log f$ is concave.

the willingness-to-pay for insurance is then given by:

$$\omega = [u(w - c\ell) - u(w - \ell)] \cdot r.$$

Firms face cost $c\ell \cdot r$ to serve this consumer. By defining $\alpha = \frac{c\ell}{[u(w-c\ell)-u(w-\ell)]}$, the two setups match.

2.1 Information and Acquisition Costs.

Next, we describe how information is acquired. Because a consumer's utility is linear, it depends on information only through the mean of posterior beliefs. Following standard practice, we associate each signal realization to the posterior mean it generates (Dworczak and Martini, 2019; Gentzkow and Kamenica, 2016). Formally, a signal realization is denoted by $\theta \in [0,1]$, and the buyer's expected payoff after observing this realization is $\theta - p$, if the good is purchased. We refer to the realization of a consumer's private information, θ, as their type or, abusing notation, their posterior. By Strassen's theorem (Gentzkow and Kamenica, 2016; Kolotilin, 2018), a distribution over posterior means, represented by its cumulative distribution function F, is generated by some information structure if and only if F is a mean-preserving contraction of the prior:

$$F \preceq_{\text{m.p.s.}} F_o. \tag{BC}$$

For a fixed prior, let $\mathcal{A} = \{F : F \preceq_{\text{m.p.s.}} F_o\}$. Following Ravid et al. (2022), we define costs directly in terms of distributions over posterior means. Concretely, to acquire information F, consumers must spend $\kappa C(F)$, with $\kappa > 0$. We assume $C : \mathcal{A} \to \mathbb{R}_+$ is strictly increasing in the mean-preserving-spread order, convex, and lower-semicontinuous in $\mathcal{L}^1$.[2] The cost of acquiring no information is zero: $C(\delta_\mu) = 0$, where δ_x is the Dirac measure on x. This formulation encompasses the majority of cost functions in the literature, including posterior-separable information costs, in general, and the entropy-based costs applied in the rational inattention literature, in particular.[3]

Example (Uniform-Quadratic). Throughout the paper, we use the following example to illustrate the general results. Let $\omega \sim U[0,1]$, so that F_o is the uniform CDF, and assume $C(F) = \frac{1}{2}\mathbb{E}_F\left[(\theta - \frac{1}{2})^2\right]$. These quadratic costs measure the expected reduction in the prior variance obtained by observing information F.

[2] We consider the space of functions from $[0,1]$ to $\mathbb{R}$. Thus, the $\mathcal{L}^1$ norm of a function F is defined as:

$$||F|| = \int_0^1 |F(x)| dx.$$

[3] See Ravid et al. (2022) for a proof.

3 How Endogenous Information Affects Demand and Firms' Costs

Our first step to characterize the equilibrium of the model is to analyze how endogenous information affects the demand curve and firms' average cost curves. Naturally, the demand curve reacts to information costs: as information costs increase, buyers acquire less information and make purchasing mistakes. Because of imperfect information, some consumers buy the good even though prices exceed their unknown valuation, ω, for the product; similarly, others will mistakenly forego the purchase. All in all, demand will be affected by the information consumers acquire. In markets with selection, consumer information also affects the supply side, because willingness-to-pay is correlated with firms' costs. As consumers acquire imperfect information about their willingness-to-pay, information costs affect selection in purchasing the good, thus, moving average costs. In what follows, we formalize how demand and average costs respond to changes in information costs.

Consider the consumer's problem. Buyers first observe prices and then choose what information to acquire, when anticipating their optimal purchasing decision for any possible signal realization, θ. Formally, for each p, they solve:

$$V_\kappa(p) = \max_{F \preceq_{\text{m.p.s.}} F_o} \mathbb{E}_F\left[\max\{\theta - p, 0\}\right] - \kappa C(F). \tag{1}$$

3.1 Simplifying the Consumer's Problem

As a first step, we show the consumer's problem in 1 can be simplified to a convex, two-dimensional optimization. Note that consumers make two choices: they first acquire information, and then decide whether to buy the product. The key insight is that these decisions can be subsumed in the choice of two variables, interpreted as the probability of purchase and the expected utility of consumption.

In what follows, we consider binary information structures; that is, information structures with two signal realizations in the support: $F \in \mathcal{A}$, with $|\operatorname{supp} F| = 2$. Each of the possible signal realizations works as a recommendation for the consumer: "buy", if the signal induces a high enough posterior, and "not buy" otherwise. With any binary structure F, we can associate the pair (U, Q), defined by $Q = 1 - F(\mu)$, and $U = \mathbb{E}_F[\theta \mathbb{1}_{\theta \geq \mu}]$. This association is one-to-one. Note that Q is the probability of the "buy" signal, so it can be interpreted as the probability of a purchase recommendation. Moreover, U is the expected utility obtained when the consumer follows F's recommendation. We can extend these definitions for the trivial information structure: no information acquired, that is $\operatorname{supp} F = \{\mu\}$. In that case, we associate F with multiple pairs. Concretely, any (U, Q) such that $Q \in [0,1]$ and $U = Q \cdot \mu$.

We denote information costs in (U, Q)-space as c, defined in the natural way. That is, if (U, Q) is associated to F:

$$c(U,Q) = C(F).$$

Proposition OA 1 in the Online Appendix A, proves c is a convex function. Before we state the result, we need two definitions. First, for any CDF G, let I_G be its integral CDF: $I_G(x) = \int_0^x G(v)dv$, for all $x \in [0,1]$. Finally, for a function f, we let f^* be its Legendre-Fenchel transform.[4]

Lemma 1. *Let the level of information costs be $\kappa > 0$. For all $p \in [0,1]$, the consumer's problem in 1 has a unique, at-most-binary optimal solution, F_κ^p with $|\operatorname{supp} F_\kappa^p| \leq 2$. Moreover, problem 1 is equivalent to the convex optimization:*

$$\max_{(U,Q)\in\mathcal{C}} \quad U - Q\cdot p - \kappa c(U,Q): \tag{CP}$$

$$s.t. \quad I_{F_o}^*(1-Q) \leq \mu - U, \tag{MPC}$$

where $\mathcal{C} = \left\{(U,Q) \in [0,1]^2 : \ U \in [Q\cdot\mu, \min\{Q,\mu\}]\right\}$.

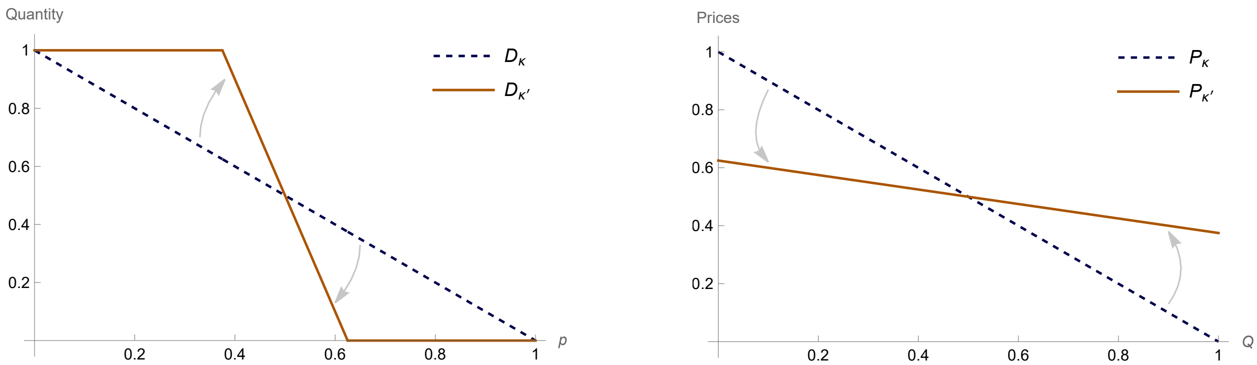
The objective function in the optimization above rewrites the buyer's problem in terms of the two variables of interest: Q, the probability of purchase, and U, the expected utility. It can be checked that $(U,Q) \in \mathcal{C}$ guarantees that the pair can be associated with a distribution F with mean μ. Recall that, in problem 1, the consumer is constrained to choose from the set $\mathcal{A}$ of mean-preserving contractions of the prior. Using integral CDFs, that constraint can be rewritten as the functional inequality: $I_F \leq I_{F_o}$. MPC reduces this infinite-dimensional constraint into a single, one-dimensional inequality. Intuitively, this works because of the convexity of I_{F_o} and the fact that F is binary.[5] Problem CP is particularly helpful because it is a finite, convex optimization and, therefore, can be solved by the usual methods.

Finally, note how Lemma 1 offers two equivalent ways of understanding the optimal information structure. One view is to see it as a distribution of posteriors for the product, F_κ^p. The other is to interpret it as a pair of variables, (U,Q), that reflect purchasing probabilities and expected utilities. In what follows, we use both views interchangeably, adopting the most convenient view for studying each concept.

Example (Uniform-Quadratic) We can rewrite the quadratic example in terms of (U,Q). Essentially, for $Q \in (0,1)$, the cost function satisfies:

[4] Recall that the Legendre-Fenchel transform of f is defined as $f^*(x^*) = \sup_x\{x\cdot x^* - f(x)\}$.

[5] Because I_F is the integral CDF of a binary distribution, we show it can be exchanged by an affine function. Thus, by convexity of I_{F_o}, the mean-preserving constraint either slacks or binds at exactly one point. One can then use this observation to derive MPC.



(a) Demand Rotation

(b) Inverse Demand Rotation

Figure 1: How Information Costs Affect Demand: $\kappa < \kappa'$.

Notes: Panels (a) and (b) show how demand and inverse demand change from information costs grow from κ to κ' in the uniform-quadratic example. For the figures, $\kappa = 0$ and $\kappa' = 4$.

$$c(U,Q) = \frac{1}{2}\frac{1}{Q(1-Q)}\left(U - \frac{Q}{2}\right)^2.$$

Using the integral CDF of the uniform distribution, one can calculate $I^*_{F_0}(1-Q) = \frac{(1-Q)^2}{2}$. It is easy to see, in this example, that solving for the optimal (U,Q) in CP is straightforward, whereas addressing problem 1 directly is considerably harder.

3.2 How Information Costs Affect Demand

We now study the implications of information costs on the demand curve. Because there is a continuum of consumers, the demand curve can be defined using the probability of purchase, Q, under the optimal information structures F^p_κ. In words, the demand at price p is the proportion of individuals whose expected valuation is equal or above p, given the information they acquire after observing that price.[6] Formally, for any $p \in [0,1]$, define demand as:

$$D_\kappa(p) = 1 - F^p_\kappa(p-),$$

where $F^p_\kappa(p-) = \lim_{x \uparrow p} F^p_\kappa(x)$. Naturally, we can readily associate an inverse demand curve to D_κ. For any level of demand, $Q > 0$, let $P_\kappa(Q) = \min\{p : D_\kappa(p) \geq Q\}$, and define $P_\kappa(0)$ by continuity. The next result shows how information costs affect the demand curve.

Theorem 1. *For any level of information cost, $\kappa \geq 0$, the following holds for the demand curve, D_κ:*

[6]This definition relies on a tie-breaking rule forcing individuals who are indifferent between purchasing or not to buy the good. This tie-breaking rule is meaningful in only one scenario: when consumers find it optimal to acquire no information regardless of prices. In that case, it selects for the equilibrium with the highest possible consumer surplus.

1. *Demand is a complementary CDF: There exists a CDF F_κ such that $D_\kappa(p) = 1 - F_\kappa(p-)$; and*

2. *Mean-preserving contraction: If $\kappa' > \kappa$, $F_{\kappa'} \preceq_{m.p.s.} F_\kappa$.*

The first part of Theorem 1 establishes that D_κ is a complementary CDF; that is, $D_\kappa = 1 - F_\kappa(p-)$ for some distribution F_κ. The second part states that, as information costs increase, this distribution becomes less informative. The first result is a law of demand: higher prices increase the marginal cost of buying the product, but the marginal benefit is unchanged, so demand falls. The mean-preserving contraction result is a consequence of the worsening of consumers' decisions as information costs increase. For higher costs, consumers acquire worse information and make more mistakes when choosing whether to buy the product. These mistakes arise because consumers fail to adjust their choices to their real, unknown, valuations and, as a consequence, higher costs imply a less-informative distribution of decisions. As Figure 1 illustrates, using the uniform-quadratic example, these results naturally carry over to the inverse demand curves.

For intuition on the second part of Theorem 1, consider how the demand changes for prices that are very high or very low. For simplicity, compare the demand at some positive κ' with the full information demand, obtained when $\kappa = 0$. When information is free, consumers can learn their valuation at no cost, and choose based on it. For a positive κ', however, buyers decide based on less precise information, incurring purchasing mistakes. When prices are very low, individuals know buying is disproportionately more likely to be the correct decision, so they will rationally err on the side of over-purchasing, causing demand to be higher than it would be under perfect information. Symmetrically, when prices are sufficiently high the opposite holds, and demand is lower than the perfect information one.

The intuition above suggests that over-purchasing patterns are monotone with prices. When that is the case, two demand curves, say for cost levels $\kappa'' < \kappa'$, should cross only once. In particular, the demand at κ' should be higher for low prices, representing an increase in over-purchasing mistakes, and lower for high prices. We say that such demand curves are rotation-ordered by κ. Rotation-ordered demand curves are well understood and studied in many applications in industrial organization, since Johnson and Myatt (2006).[7] In the Online Appendix A, we establish sufficient conditions on costs and priors for demands to be rotation-ordered, which is important for two reasons. First, it shows information costs provide a natural micro-foundation for rotation-ordered demand curves. Second, when demand curves are rotation-ordered, one can import results from that literature to study markets with endogenous information.

[7] For applications, see Mahoney and Weyl (2017).

3.3 How Information Costs Affect Firms' Average Costs

The next result explains how average costs change as information costs vary. The average cost curve captures how prices (or quantities) determine buyer selection. Prices affect information acquisition differently for different costs of information, so firms' costs vary with information costs. We are primarily interested in the average cost curve in quantity space, but it is easier to first define it using prices. Let $R_\kappa(p)$ be the expected firm's cost of posting price p:

$$R_\kappa(p) = \alpha \cdot \mathbb{E}_{F_\kappa^p}[\theta | \theta \geq p].$$

The average cost at price p is the expected value of buyers who purchase the good, computed given the endogenously determined information structure, F_κ^p, which itself depends on the price. We can then simply apply the inverse demand, $P_\kappa(Q)$, to obtain the average cost curve for a quantity Q as $AC_\kappa(Q) = R_\kappa(P_\kappa(Q))$.

Theorem 2. *Let the level of information costs be $\kappa' > \kappa \geq 0$. For any quantity $Q \in [0,1]$, the average cost curves satisfy:*

$$AC_{\kappa'}(Q) \leq AC_\kappa(Q).$$

Moreover, equality holds for $Q = 1$.

Theorem 2 states that average costs, as a function of quantities, weakly decrease as information costs increase. The result implies that selection decreases as information costs increase, because it becomes cheaper for the firm to cover the same fraction of consumers. For intuition, recall that we can represent consumer's information by two variables, (U, Q): the expected utility it provides to the agent, and the probability it recommends purchasing. For a fixed probability of purchase, Q, consumers with a higher expected utility U are costlier for the firm: indeed, in this case, a higher U indicates a higher valuation conditional on buying. At the same time, given the same probability of purchase, an information structure that guarantees a higher expected utility, and thus higher average costs, must be more informative. It is intuitive that consumers will acquire more information when costs are lower, implying a higher expected utility for a given probability of purchase, so that selection is worse when information is cheap. The proof of the theorem formally verifies this intuition.

It is relevant to identify the cases in which average costs are not changed with information costs. Those are cases in which selection is not eased, despite information becoming more expensive. We show that this requires the optimal information structures to be of a particular type, and we formalize when such information structures arise.

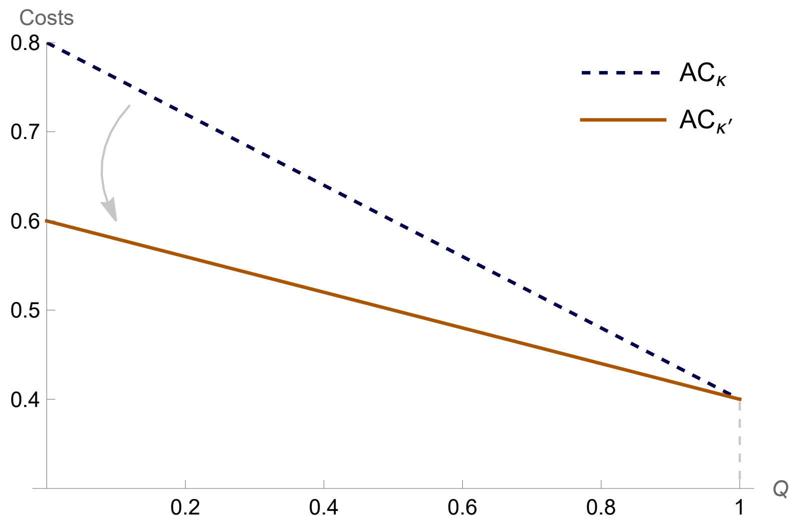


Figure 2: Average Costs Rotation: $\kappa < \kappa'$.

Notes: This figure plots a rotation of the average cost curve for the uniform-quadratic example, with $\alpha = .8$, $\kappa = 0$ and $\kappa' = 4$.

Definition 1. *A binary information structure $F \in \mathcal{A}$ is monotone if there exists $x \in (0,1)$ such that:*

$$\operatorname{supp} F = \left\{\mathbb{E}_{F_o}[\theta|\theta \leq x], \mathbb{E}_{F_o}[\theta|\theta \geq x]\right\}.$$

Monotone information structures have an intuitive interpretation: consumers receive a “buy" signal if and only if their real valuation, ω, is above a certain threshold. They are monotone in that, given two individuals who received different signals, it is known for sure that the one who was recommended to ”buy" has a higher valuation than the other.

Remark 1. *Fix a quantity $Q \in (0,1)$, and assume consumers acquire some information for levels of information costs $\kappa' > \kappa \geq 0$. Then, the following are equivakent:*

1. *Average costs coincide for κ, κ': $AC_{\kappa'}(Q) = AC_\kappa(Q)$;*
2. *Average costs coincide up to κ': $AC_o(Q) = AC_\kappa(Q)$;*
3. *Optimal information structures are monotone for all $0 < \kappa'' \leq \kappa'$.*

This result makes two statements. First, if making information harder to acquire at a certain level of information costs, κ, does not ease selection, then it also would not ease it for any lower level of information costs. In other words, the average cost of the firm is the same for any information costs smaller than κ. Second, making information harder to acquire does not ease selection if and only if information structures are monotone. We focus on the second part. To see why monotone information implies that selection is not eased by information costs, consider a firm trying to sell Q units of the good. If information is free, consumers know their willingness-to-pay, and the firm will sell to the Q consumers with the highest valuation. What if information is costly? In that case, if the optimal information structure is monotone, the firm will serve the same consumers as under free information. As a consequence, the firm will have the same cost whether information is free or costly. The result shows the converse is true as well: insofar as average costs do not decrease, information structures must be monotone.[8]

4 The Equilibrium Effect of Endogenous Information

We now apply the previous results to study equilibrium. Note that a Subgame Perfect Nash equilibrium in this market requires firms to make zero profits. Additionally, it follows from Bertrand competition that there cannot be a price lower than the equilibrium price that is profitable for firms — otherwise, one firm could undercut its competitor and make positive profits, serving all their customers. These properties, together with consumer demand and firms' costs, characterize the equilibrium set. Formally, let $\mathcal{E}$ be the set of equilibrium prices and quantities, and recall $R_\kappa(p)$ represents firms' average costs for price p. Then:

$$\mathcal{E} = \{(p, Q) : P_\kappa(Q) = p,\ R_\kappa(p) = p \text{ and } R_\kappa(p') \geq p' \text{ for all } p' < p\},$$

where the first condition guarantees that prices and quantities are compatible with consumer demand, the second condition requires firms to make zero profits, and the third condition guarantees that there is no price below the equilibrium price that is profitable for firms. We say that the equilibrium with the highest price is the highest equilibrium.

Although there can be multiple equilibria, all but the highest equilibrium are trivial. In particular, they are not the product of firms' undercutting behavior, and thus can be readily ruled out by a natural equilibrium refinement.[9] Indeed, by the last condition in $\mathcal{E}$, any equilibrium price that is not the highest must be isolated from the set of prices that give positive profits for the firms. As a consequence, no firm would ever choose such a price in order to attract their competitor's customers: this price is not the limit of any reasonable competitive dynamics. Henceforth, we consider only the highest equilibrium. We denote (p^e, Q^e) equilibrium prices and quantities with $P_\kappa(Q^e) = p^e$, and:

$$p^e = \inf_{p \in [0,1]} \{p : R_\kappa(p) < p\}.$$

[8] The restriction to the case in which the consumer acquires information is to avoid a trivial case. Of course, when no information is acquired at κ', $AC_{\kappa'}(Q) = AC_{\kappa'}(1) = \alpha\mu$. If no information is acquired at κ', then no information is acquired at κ, and $AC_\kappa(Q) = AC'_\kappa(Q)$.

[9] See the Online Appendix D for a formal proof of this claim. In our refinement, we discretize firms' choice space (the set of prices a firm can choose from) and consider a sequence of markets in which this grid becomes progressively finer, approaching the continuum. The equilibrium with the highest price is the only one that is a limit of equilibria from the discretized markets.

Finally, because firms make zero profits in equilibrium, equilibrium welfare, W^e, coincides with consumer's ex-ante utility: $W^e = V_\kappa(p^e)$.

4.1 Efficiency is Non-monotonic with Acquisition Costs

Our first goal is to show how efficiency varies with information costs. Consider an economy in equilibrium, in which the product is traded at a high price and suppose that the information cost grows. Recall, from Theorem 1, that an increase in information costs causes consumers to make more mistakes, which, when prices are high, dampens demand. This channel suggests that higher information costs lead to even higher prices and a smaller surplus. On the other hand, Theorem 2 states that the cost of serving any given fraction of buyers decreases with information costs, lowering equilibrium prices and increasing surplus. We next study which of the two forces dominates.

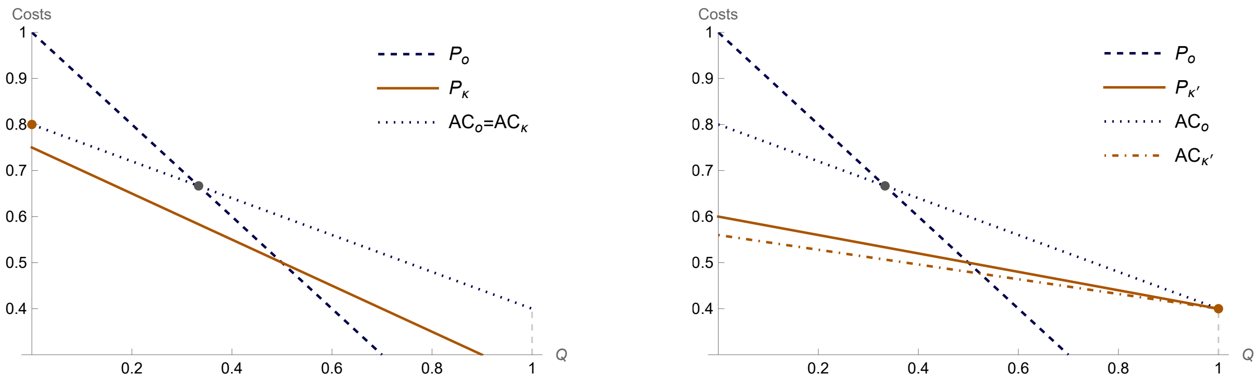
Proposition 1. *Let information costs C be Lipschitz continuous.*[10] *Then, there exist two thresholds for selection, $\overline{\alpha} < \tilde{\alpha} < 1$ such that:*

1. *If $\alpha > \overline{\alpha}$, welfare is non-monotonic: it first decreases when κ is low and it increases when κ is sufficiently high; and,*
2. *For all $\alpha > \tilde{\alpha}$, there is an interval of information cost levels, $\mathcal{K}_\alpha \subset (0, \infty)$ such that, for $\kappa \in \mathcal{K}_\alpha$, the market unravels — welfare and traded quantities are zero.*

Proposition 1 relates changes in equilibrium prices to changes of information costs. The first part of the result states that, when selection is severe (that is, α is large), efficiency is non-monotonic. In particular, welfare decreases when information costs rise from zero, but increases when information costs are higher. This is a consequence of the trade-off identified in the last paragraph: expensive information worsens decision-making, dampening demand, but it alleviates selection. Recall, from Remark 1, that average costs do not change when the optimal information structures are monotone. We prove that consumers acquire monotone information structures when κ is low, so that selection is not affected by information costs when acquisition is cheap. Therefore, small information costs can only decrease welfare. On the other extreme, when information costs are high, the easing of selection dominates, and efficiency increases.

Part 2 in the proposition states that this non-monotonicity result can be severe. The initial dampening in demand may be enough to move the economy all the way to market unraveling. By contrast, a subsequent increase in information costs can bring the equilibrium back to a higher level of efficiency — in fact, to

[10] That is, there exists $M > 0$ such that:
$$|C(F) - C(F')| \leq M \|F - F'\|, \text{ for all } F, F' \in \mathcal{A}.$$



(a) Equilibrium for low κ: Market Unravels

(b) Equilibrium for high κ': Full Efficiency

Figure 3: Equilibrium comparison: $\kappa < \kappa'$

Notes: In both figures, the dashed curve represents inverse demand and the dotted represents average costs when information is free. The gray dot, then, is the equilibrium for free information. Panel (a) compares the free information equilibrium with the equilibrium when $\kappa = 2$; that is, the orange dot. The average cost curve is the same in the two cases, but the inverse demand rotates to the full orange curve, being below cost for all quantities. Thus, equilibrium has zero trade: the market unravels. Panel (b) compares the free information equilibrium with the equilibrium when $\kappa' = 4$, again, the orange dot. In this case, the average cost curve, dot-dashed, is always below the inverse demand, and equilibrium is efficient.

full efficiency for high enough κ. Figure 3 illustrates this result in the uniform-quadratic example. In both panels, the inverse demand and average cost curves for free information are plotted in dashes and dots, respectively. The intersection of these two curves is the equilibrium when information is free, and the level of trade is intermediate. In Figure 3a, the full curve is the inverse demand for positive information costs, $\kappa = 2$. In the uniform-quadratic example, consumers choose monotone information structures for all prices when $\kappa \leq 2$, so the average cost curve does not decrease in that interval. The equilibrium for $\kappa = 2$, the orange point, has no trade because the average cost curve exceeds the demand for all quantities. Figure 3b shows this result is reversed for $\kappa' = 4$, where the inverse demand curve exceeds the dot-dashed average cost and equilibrium is efficient.

Proposition 1 has implications for policy. It shows that both aiding and hindering consumer learning can lead the market to breakdown, with substantial consequences for welfare. Whether such policies are welfare-improving is then, ultimately, an empirical question. In Section 5.2, we develop an approach to help researchers identify, using observable data, whether increasing or decreasing information costs would benefit consumers.

Lipschitz continuity. The assumption of Lipschitz continuity is sufficient for the result, but not necessary. When information costs are Lipschitz continuous, it is possible for consumers to acquire information that rules out certain valuations ω, at a finite marginal cost. This allows us to prove that when the level of information costs, κ, is low, consumers find it optimal to acquire monotone information structures and, thus, that selection is not eased for low κ. While important cost functions, such as quadratic costs, satisfy this assumption, mutual information does not. Indeed, for entropy-based costs, it is never optimal for consumers to acquire information structures that are monotone, as ruling out any possible valuation would imply an arbitrarily high marginal cost. Yet, we show in the Online Appendix A that the result of Proposition 1 extends to mutual information with a uniform prior.

4.2 Endogenous Information Worsens Selection

We now turn examine how information that is endogenously acquired differs from information that is exogenously distributed. To start, we discuss the possibility of unraveling: that is, of no trade occurring in equilibrium. Proposition 1 proves that, when information is endogenous, the market breaks down for a wide range of information costs. The next result shows this is in direct contrast with what happens when information is exogenous. To formalize that, for a given distribution F, let $(p^s(F), Q^s(F))$ be the equilibrium prices and quantities when information is exogenous, and valuations are distributed according to F. Moreover, define as $W^s(F)$ the consumer surplus obtained in that equilibrium, discounted of information costs. That is:

$$W^s(F) = \mathbb{E}_F[\theta|\theta \geq p^s(F)] - p^s(F)Q^s(F) - \kappa C(F).$$

Remark 2. *For all distributions of valuation F, equilibrium under exogenous information has positive trade:*

$$Q^s(F) > 0.$$

While markets often break down when information is endogenous, this will never be the case when information is exogenous, regardless of the distribution of valuations in the market. This result suggests that all the unraveling in this model comes from endogenous information. Endogenous information worsens selection, making it unprofitable for firms to sell the product at certain prices. Intuitively, at each price, consumers acquire information that allows them to make better purchasing decisions. This information will affect the pool of consumers who buy the good, increasing firms' costs. For some parameters, the higher firm costs induced by information acquisition will be so extreme that no price is profitable for firms, and there cannot be any trade.

Our next result shows how, by worsening selection, endogenous information hurts consumers. We compare consumer surplus when information is endogenous and exogenous. To obtain a meaningful comparison, we need to keep constant the information between the two models. We proceed as follows: first, we solve for equilibrium in the endogenous information model. At equilibrium, consumers acquire a certain information structure, F. We then take this information structure as the exogenous distribution of valuations in a standard adverse selection problem, solve for the equilibrium under exogenous information, and compare the two outcomes. This guarantees that, in equilibrium, the acquired information in the two models is the same.

Proposition 2. *Fix a level of information costs, κ, and let (p^e, Q^e) be equilibrium prices and quantities, and $F_\kappa^{p^e}$ be the information acquired in equilibrium, generating consumer surplus W^e. Then:*

1. *Equilibrium prices are larger and equilibrium quantities are lower under endogenous information: $p^e \geq p^s\left(F_\kappa^{p^e}\right)$ and $Q^e \leq Q^s\left(F_\kappa^{p^e}\right)$*
2. *Endogeneity hurts the consumer: $W^e \leq W^s\left(F_\kappa^{p^e}\right)$.*

Proposition 2 shows that efficiency is worsened under endogenous information. Indeed, consumers will be worse off. The intuition for this result is, again, that buyers' ability to adjust information to prices worsens selection — but, this time, off the equilibrium path. Consider a simple case in which the market unravels under endogenous information. Because there is no trade, consumers acquire no information in equilibrium, so buyers and sellers are symmetrically informed. If information were exogenously distributed, this symmetry would lead to efficiency: competition induces firms to serve all consumers at price $\alpha\mu$. Why is this not an equilibrium when information is endogenous? If firms charged $\alpha\mu$, consumers in our model would adjust their information strategy, acquiring further information and worsening selection at that price. This worsening of selection renders price $\alpha\mu$ unprofitable for firms, ruling it out as an equilibrium. This example shows that the loss of welfare induced by endogenous information can be considerable: while the exogenous equilibrium is welfare-maximizing, the endogenous one has no trade. An implication of this result is that consumers would be better off if they could commit *not* to adjust their strategies to prices.

Figure 4 shows an example in which endogenous information is strictly worse than exogenous information, but where markets do not unravel. It plots prices in the horizontal axis, and three curves. The endogenous information equilibrium is the intersection between average costs in price space, R_κ, and the 45-degree line, marking the zero profits condition. In this example, the endogenous equilibrium price is around .77. The dashed curve, R plots the average cost if the information endogenously acquired at .77 was exogenously distributed. In that case, equilibrium price drops to .4, and all consumers are covered, leading to a strict increase in welfare.

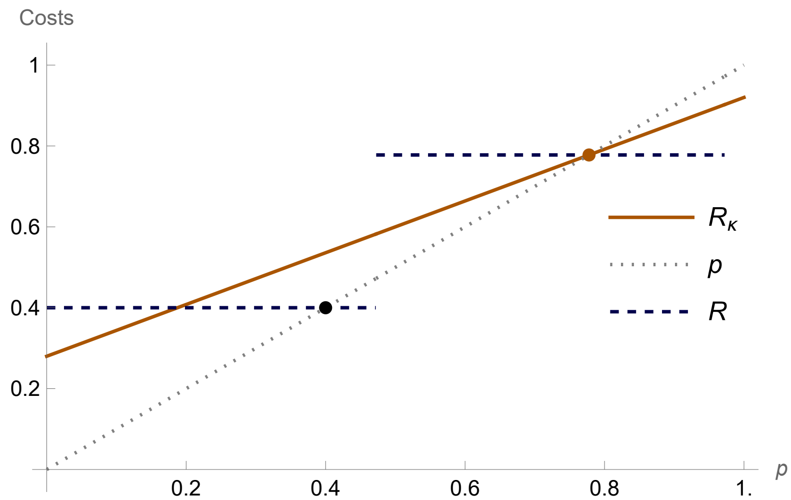


Figure 4: Endogeneity worsens selection

Notes: This figure plots two equilibria: one when information is endogenous and the other when information is exogenous. The full line is the average cost curve under endogenous information. The equilibrium in this case is at the intersection of that curve with the dotted 45-degree line: the point in orange. The dashed curve, R, is the average cost for the fixed information structure acquired at equilibrium. The equilibrium when that information is exogenous is the lower point in the figure, in black. The figure was produced using the quadratic-uniform example, with $\alpha = .8$ and $\kappa = 1.5$.

5 Identifying and Quantifying Endogenous Information

In this section, we provide tools to bring the theory to data. While data on consumers' decisions and prices are readily available in many markets, information costs and individuals' beliefs are considerably harder to observe. Our results take this restriction seriously by assuming that nothing is known about prior information or information costs. Despite these constraints, the model has clear empirical consequences. In what follows, we (1) provide a test for the presence of endogenous information in market data; and (2) develop tools for the analysis of policies that affect the incentives for information acquisition. These results not only help bridge the gap between the model and applied work, but also clarify the empirical content of the theory.

5.1 Detecting Endogenous Information

We provide a test to detect and quantify endogenous information acquisition in a dataset. A sizable literature in the industrial organization of insurance markets uses individual claims data to estimate demand and firms' cost curves.[11] Borrowing from that literature, we assume that these curves are observed, and nothing else. That is, the set of observables is the pair of curves $\mathcal{O} = \{P, AC\}$, where P represents inverse demand, and AC average costs. The goal of the following result is to find necessary conditions for a set of observables to be consistent with information acquisition. The test explores the key insight that selection

[11] See Einav et al. (2021) for an overview.

is worsened under endogenous information. As a consequence, endogenous information imposes a relationship between demand and firms' costs that is qualitatively different from exogenous information. We formalize the procedure below.[12]

Definition 2. *An inverse demand and cost $\mathcal{O}$ are consistent with endogenous information acquisition if there exists a distribution F_o, and information costs C, $\kappa > 0$, such that they can be generated by optimal information acquisition when F_o is the prior: that is, $\mathcal{O} = \{P_\kappa, AC_\kappa\}$. They are consistent with exogenous information when $\kappa = 0$ in the above, and F_o is then interpreted as the exogenous distribution of information in the economy.*

For simplicity, in what follows, we assume P is strictly decreasing and the parameter α is known. We revisit the latter assumption at the end of this section. Consider a researcher trying to identify whether or not buyers are acquiring price-dependent information in this market. The null hypothesis is that all information is exogenous. Under this hypothesis, there exists a distribution of information in the economy, F_o, and the demand function reflects this distribution of information: $P(Q) = {F_o}^{-1}(1-Q)$. In other words, under the assumption of endogenous information, one can reconstruct F_o by observing P.

The next step is to construct a counterfactual average cost curve, $\overline{AC}$. Under the null hypothesis, the average cost curve is given by the expected cost of consumer who purchase the good, when their beliefs are given by the exogenous distribution F_o:

$$\overline{AC}(Q) = \alpha \cdot \mathbb{E}_{F_o}[\theta | \theta \geq P(Q)].$$

As a consequence, $\mathcal{O}$ is consistent with exogenous information if and only if this counterfactual curve coincides with the observed one: $\overline{AC} = AC$. The next proposition presents a necessary condition for $\mathcal{O}$ to be consistent with information acquisition.

Proposition 3. *Let a demand curve and an average cost curve, $\mathcal{O}$, be consistent with endogenous information acquisition. Then, for all levels of demand, $Q \in (0,1)$, the observed average cost exceeds the counterfactual average cost:*

$$\overline{AC}(Q) < AC(Q).$$

Moreover, the distance between the two curves quantifies information acquisition. That is, for $p = P(Q)$:

[12] The procedure we formalize below can be applied when firm's costs are affine in consumers' types — that is, when consumers' valuations are perfectly correlated with firms' costs. In the Online Appendix C, we extend the test to arbitrary, strictly increasing, firms' cost functions, χ. For that, we assume consumers observe individual-level data on costs, which is compatible with the empirical literature.

$$AC(Q) - \overline{AC}(Q) = \frac{\alpha}{Q}\kappa C\left(F_{\kappa}^{p}\right)$$

Proposition 3 shows that, when the observed curves are a product of information acquisition, the observed average cost curve is strictly higher than the counterfactual one. For the intuition, note that, under exogenous information, the demand curve reflects a unique information structure, describing the distribution of information in the market. Under endogenous information, that curves traces out the probability of trade at each price, reflecting a continuum of information structures; one for each price. Intuitively, for a fixed traded quantity, the exogenous information structure is less informative than the endogenous one, because, under endogenous information, consumers can adjust the quality of their information to make better choices. This implies that selection will be worse under endogenous information, and the result follows.

In order to distinguish endogenous information from exogenous, both curves must be observed. For example, by observing demand alone, one cannot determine whether or not it came from endogenous information. In particular, the first part of Theorem 1 shows that demand is a complementary CDF, and, thus, could be generated by a certain exogenous distribution of valuations in the market.[13] Rather, it is the relationship between demand and cost curves that allows for testable implications.

5.2 An Approach for Policy Analysis

We now provide tools for evaluating policies that change information costs. Our model suggests that such policies can be most consequential. In particular, Proposition 1 shows that either increasing or decreasing information costs can lead to a complete market breakdown. Thus, we believe our result can be valuable for policy-makers in markets where endogenous information is a concern. Policy-makers may know little about the costs consumers face, but they may still be able to change them. For example, regulation that eases the comparison of plan alternatives, or lowers the cost of tests that would help individuals assess their valuation for a product, are examples of interventions that reduce the difficulty of information acquisition. Our results help to analyze these policies.

To evaluate the consequences of easing consumer learning, it is, in general, necessary to measure how information costs change both demand and firms' costs. In particular, assuming smoothness, welfare analysis depends on the two unobservable elasticities:[14]

[13]The impossibility is, in fact, stronger. Part 2 of Theorem 1 proves that endogenous information generates demands represented by mean-preserving contractions of the prior. This is also true if consumers have the same prior beliefs, and receive signals exogenously. Thus, even if the researcher observes the prior and the demand curve, it is impossible to distinguish the two types of information.

[14]One can express all these definitions in terms of demand, as opposed to inverse demand. However, not only are the expressions below more consistent with our choice of observables, but they are also is shorter and clearer in quantity space.

$$\varepsilon_{AC,\kappa} = \frac{\partial AC(Q)}{\partial \kappa}\frac{Q}{AC(Q)}, \qquad \varepsilon_{P,\kappa} = \frac{\partial P(Q)}{\partial \kappa}\frac{\kappa}{P(Q)}.$$

To proceed, we define the elasticity of the average cost curve with respect to quantities, which is observable:

$$\varepsilon_{AC,Q} = -\frac{\partial AC(Q)}{\partial Q}\frac{Q}{AC(Q)}.$$

Note that $\varepsilon_{AC,Q}$ includes a minus sign because the average cost curve is decreasing in Q in adversely selected markets. Our next result uses this observable elasticity to compute an unobservable one, $\varepsilon_{P,\kappa}$, and provides bounds for the other, $\varepsilon_{AC,\kappa}$.

Proposition 4. *For any observable inverse demand and average cost:*

1. *The elasticity of firm's costs is negative:*

$$\varepsilon_{AC,\kappa} \leq 0.$$

2. *The elasticity of demand can be computed from data:*

$$\varepsilon_{P,\kappa} = 1 + (\varepsilon_{AC,Q} - 1)\frac{AC(Q)}{\alpha \cdot P(Q)}.$$

Proposition 4 provides a bound on one of the unobservable elasticities, and precisely calculates the other using observables. The lower bound result is a direct consequence of Theorem 2. Indeed, part 1 simply says average costs decrease when information costs increase. The result for the elasticity of demand is more interesting and further exploits the relationship between demand and average cost imposed by information acquisition. A rough intuition for this result is that price variation is sufficiently rich to identify one dimension of information costs. Note that problem CP is similar to a quasilinear optimization, in which a consumer chooses consumption levels for two goods: (U, Q). Varying prices affects the value of purchasing the good Q relative to information costs, κ. Thus, one can infer how Q changes with κ by observing how consumers' decisions vary with prices. By contrast, one cannot obtain a similar expression for the elasticity of the average cost with respect to information costs, $\varepsilon_{AC,\kappa}$: price variation alone cannot provide information about the value of U relative to information costs for the consumer. The same argument carries over to average costs, which are a function of U.

The result in Proposition 4 provides a lower bound for the effect of an increase in information costs on equilibrium. Because equilibrium is determined by the intersection of the demand and average cost curves, the elasticities $(\varepsilon_{AC,\kappa}, \varepsilon_{P,\kappa})$ are sufficient to determine how traded quantities or prices change with information costs. The lower bound makes this inference harder, but still possible: whenever the elasticity of the demand curve is positive, equilibrium quantities must increase with information costs. A similar lower bound can be computed directly for welfare.

Identifying firms' costs. Throughout the previous discussion, we assumed the severity of selection, α, is known. Indeed, in our baseline model, α can be identified directly in the data: in particular, consider the equation $AC(1) = \alpha\mu$, which must hold under both endogenous and exogenous information. Because μ can be calculated as the mean of the demand function, and AC is observable, this equation identifies α.

However, our results extend to more realistic settings, where firms' costs depend on multiple parameters that cannot be easily identified without additional data. For example, assume firms' cost for serving a consumer with valuation ω is an affine function: $\alpha_o + \alpha\omega$. Now, there is only one equation to identify the two parameters:

$$AC(1) = \alpha_o + \alpha\mu.$$

This equation restricts a set of plausible (α_o, α), that are compatible with the observed data. An extra equation is required to determine both constants. This can be obtained if the researcher observes demand and costs for two groups who vary in their expected valuation. A concrete example is that of prescription medicine insurance. Young and old consumers might have a different distribution of probabilities in relation to their medication needs, thus affecting their willingness-to-pay. However, conditional on having the same probability of purchase, the costs they impose on the insurance company should be the same, justifying that (α_o, α) are the same between the two groups. Provided that expected risks are different between these two populations, one could identify these two constants using the two observed average costs curves; that is, one for each group.

6 The Welfare Cost of Selection

Next, we show that information decisions impose an externality in the market. Consumers do not internalize that their learning affects selection, causing excess acquisition. This externality extends the costs of selection beyond the consumption allocation: the inefficiency of information decision compounds the traditional welfare losses. Due to the information externality, the usual techniques for measuring welfare costs will often be inadequate in the presence of information acquisition. In particular, we show that such measures are a lower bound on the losses due to selection.

For this section, we extend the model so that firms have no gains from trading with some consumers. Concretely, firms' costs follow:

$$\chi(\omega) = \alpha_o + \alpha \cdot \omega,$$

with $\alpha_o > 0$.[15]

6.1 The Information Externality

We now formalize the externality imposed by information. We represent allocations by the pair (U, Q), where U is the expected utility of consumption for consumers, and Q is the demand or probability of purchase. For any allocation, welfare is defined as the total surplus in the economy; that is, the sum of consumers' utility and firms' profits:

$$W(U, Q) = (1 - \alpha)U - \alpha_o Q - \kappa c(U, Q).$$

We consider a planner who maximizes welfare by choosing any pair (U, Q), consistent with the consumer's prior information. It is useful to separate the planner's problem into two parts: the choice of information and the decision of a trade level. Conditional on the amount of trade, Q, the planner chooses information that is captured in the expected utility U, subject to the constraint that (U, Q) must be feasible for consumers. This requires the information structure associated with (U, Q) to be a mean-preserving contraction of the prior, which is equivalent to (U, Q) satisfying the MPC constraint in CP. Thus, for fixed Q, the planner chooses information:

$$U^*(Q) = \arg\max_U \{(1 - \alpha)U - \kappa c(U, Q) : MPC\}. \tag{2}$$

Denote the function U^* as the planner's information strategy. We are interested in comparing U^* with the consumers' information strategy: the function $U(\cdot)$ that describes the consumer's expected utility for each probability of purchase, given by their optimal information structure.

Remark 3. *The consumers' information strategy is always more informative than the welfare maximizing one:*

$$U(Q) \geq U^*(Q)$$

Moreover, equality holds if and only if the individual acquires no information (that is, $U(Q) = \mu Q$), or the planner's information structure, $(U^*(Q), Q)$, is monotone.

[15] All the results carry through to this extension without modification, except for Proposition 1. The results of that proposition still hold qualitatively: there will be parameters such that welfare is non-monotonic and markets unravel.

This result shows that, for any fixed probability of trade, consumers acquire more information individually than a planner would find optimal. Intuitively, because of adverse selection, individual information choices impose an externality on the rest of the economy, creating a wedge between the social and individual values of information. While consumers internalize that information improves the quality of their choices, they disregard that acquiring information worsens selection. Adverse selection is the cause of the gap between the strategies of an individual and the planner: in the absence of selection — $\alpha = 0$ — the strategies coincide. This result implies that the usual subsidies and price policies cannot fully mitigate the costs of adverse selection. This is because pricing policies alone can correct the adverse selection problem given information decisions but cannot align the social and individual values of information.

6.2 Measuring Welfare Costs

This section shows that a common measure of the welfare costs of selection falls short when information is endogenously acquired. Quantifying the inefficiencies due to adverse selection is the central theme of a large literature in industrial organization (Einav et al., 2010; Finkelstein et al., 2019; Panhans, 2019). Much of that literature measures the amount of welfare lost due to adverse selection using a geometric analysis, based on consumers' demand and firms' costs.

When information is exogenous, adverse selection imposes an allocative inefficiency that decreases the traded quantity with respect to what is socially optimal. The equilibrium quantity, Q^e, is such that inverse demand intersects average costs, whereas it is well known that the larger socially-efficient quantity, Q^{MC}, is such that the inverse demand curve equals marginal costs. In that case, the welfare loss due to adverse selection is quantified by the area between the inverse demand curve and the marginal cost of the firm, on the vertical axis, and between equilibrium and efficient pricing, on the horizontal one. Formally:

$$L\left(Q^e, Q^{MC}\right) = \int_{Q^e}^{Q^{MC}} [P(\nu) - MC(\nu)]\, d\nu,$$

where MC is the marginal cost curve.

When information is endogenously acquired, Remark 3 shows that information decisions are distorted, in addition to consumption decisions. We let Q^* be the socially-efficient level of trade when information is endogenous. Thus, the welfare loss due to selection can be decomposed as follows:

$$W(U^*(Q^*), Q^*) - W(U(Q^e), Q^e) =$$

$$\underbrace{W(U^*(Q^*), Q^*) - W\left(U(Q^{MC}), Q^{MC}\right)}_{\text{Information loss}} + \underbrace{W\left(U(Q^{MC}), Q^{MC}\right) - W(U(Q^e), Q^e)}_{\text{Loss given information choices}}$$

In words, the welfare loss consists of two terms. The second term captures the change in welfare from setting the level of trade at Q^{MC}, while maintaining individual information choices. This reflects what the planner can obtain by controlling prices or designing subsidies for this market. Indeed, the quantity is Q^{MC} is welfare-maximizing when the planner is restricted to choose the same information strategy as the buyer. By contrast, the first term captures the gain from allowing the planner to re-optimize the information choice. The next proposition proves that the second term is equivalent to the quantity $L(Q^e, Q^{MC})$, and then uses this result to conclude that $L(Q^e, Q^{MC})$ underestimates welfare losses.

Proposition 5. *The welfare loss is larger than the standard measure L:*

$$W(U^*(Q^*), Q^*) - W(U(Q^e), Q^e) \geq L(Q^e, Q^{MC}).$$

Equality holds if and only if $Q^* \in \{0, 1\}$ or $\left(U^*(Q^{MC}), Q^{MC}\right)$ is a monotone information structure.

Proposition 5 states that the commonly used measure, $L(Q^e, Q^{MC})$, is a lower bound for welfare loss in the presence of information acquisition. Intuitively, the observed demand and cost curves quantify the welfare costs conditional on the individual information strategies. Indeed, this part of the effect can be measured by the area between these curves. However, as shown in Remark 3, information decisions are also distorted by adverse selection. The demand and cost curves do not provide information about the loss arising from the socially sub-optimal information choices of consumers, and it therefore underestimates welfare costs. There are two cases in which the loss due to endogenous information does not matter, and $L(Q^e, Q^{MC})$ fully quantifies the welfare cost of selection. The first case is when the welfare-optimal level of trade is extreme: either the planner cannot prevent unraveling, or they would like to serve all consumers, as in our baseline model. The second case is when, at the observed quantity Q^{MC}, the planner chooses a monotone information structure. Note that, in the first case, the consumer acquires no information at all, whereas in the second, the planner acquires as much information as possible. Then, Remark 3 guarantees that the information choices of the planner and consumers coincide in both cases, and the sub-optimal information strategies lead to no welfare loss.

7 An Application: Increasing Insurance Take-up

This section illustrates that combining information design with subsidies can decrease the social cost of expanding insurance coverage. There exists ample evidence of low insurance take-up in subsidized markets, and surveys document the cost of insurance as the leading reason for non-enrollment (Assistant Secretary for Planning and Evaluation, 2021; Garfield et al., 2019). In the Massachusetts subsidized insurance exchange, Finkelstein et al. (2019) find that while adverse selection plays a role in explaining take-up, most of the low demand is due to a different fact: consumers' observed willingness-to-pay are far lower than the cost they impose on insurers. This finding is at odds with the textbook models of insurance, in which consumer valuation always exceeds firms' marginal costs, and it is socially optimal to serve all consumers.

Information acquisition can partially rationalize the observed mismatch between consumer value and firm's costs. Indeed, consumers' estimated willingness-to-pay may be much lower than their real willingness-to-pay when information is costly. When prices are relatively high, consumers tend to under-purchase the good, dampening demand. Because willingness-to-pay is inferred through the inverse demand, this dampening leads to an underestimation of consumers' valuation. This effect can be strong enough to bring much of the inverse demand curve under the marginal cost curve, even in a pure insurance economy. Figure 5a illustrates this case with the quadratic example in a pure adverse selection economy. It plots consumers' real willingness-to-pay (dashed curve), marginal costs (blue full curve) and their observed willingness-to-pay, when inferred through demand. While each consumer's real willingness-to-pay exceeds the cost they impose on insurers, the observed willingness-to-pay is below marginal costs for about 40% of the population.

If endogenous information contributes to a low take-up, information policies may help mitigate it. We show that in an example. Consider a planner controlling two dimensions of policy: (1) she decides how much information to disclose to consumers, while taking into account that buyers can choose to acquire additional information on their own; and (2) she chooses the level of product subsidies. The planner knows that the final allocation is determined by competitive equilibrium. For simplicity, we restrict our attention to a subset of information policies that we call 'partitional': the planner subdivides the state space in any number of intervals and reveals to each consumer the interval in which their real willingness-to-pay is. For example, in insurance markets governments could subsidize medical examinations for prospective insurees, categorizing consumers into risk classes. When a partitional information policy is binary (that is, it subdivides the state space in two intervals), we say it is 'monotone'.[16] Each realization of the planner's information policy generates a belief to the buyer, so the principal's information policy, τ, is a distribution over the consumers beliefs. We let $\mathcal{P} \subset \Delta\Delta[0,1]$ be the space of partitional policies, and we denote subsidies as s. Finally, let $C(G;\pi)$ the cost of acquiring information structure G conditional on the belief π. In our example, this cost is quadratic: if π has mean μ_π, $C(G;\pi) = \frac{1}{2}\mathbb{E}_G\left[(\theta - \mu_\pi)^2\right]$.

A buyer with belief π solves:

[16] A partitional, binary, information policy induces a monotone information structure.

(a) Low Observed Values

(b) Subsidies Comparison

Figure 5: Low Take-up and Information Design.

Notes: Panel (a) shows that the observed valuation can be lower than the marginal cost when information is endogenous. Although real consumer willingness-to-pay (dashed curve) exceeds the marginal costs (blue curve), the observed willingness-to-pay is below marginal costs for about 40% of consumers. Panel (b) compares the subsidies required to achieve each level of market coverage, Q. s_0 are the subsidies required without any information intervention, and s when the planner also implements an information policy. Both figures are based on the uniform-quadratic example, with $\alpha = .8$ and $\kappa = 2$.

$$V((1-s)\cdot p;\pi) = \max_{G \preceq_{\text{m.p.s.}} \pi} \mathbb{E}_G\left[\max\{\theta - (1-s)\cdot p, 0\}\right] - C(G;\pi).$$

This is the same as the original consumers' problem except that prices are subsidized, so consumers pay $(1-s)\cdot p$, and the cost reflects consumers' beliefs after observing the information disclosed by the principal. Let the solution to the problem above be by $F^{(1-s)p}(\,\cdot\,;\pi)$.

Assume the planner aims to cover a fraction Q of the population and pays for information using the same acquisition cost as consumers. With an abuse of notation, the cost of a partition τ is denoted $C(\tau;F_o)$. The problem of the planner is then to choose a partitional information structure and a subsidy, such that the induced equilibrium in the product market has quantity Q, while anticipating that a buyer who receives belief π will optimize their information choice with that belief as their prior:

$$W(\tau,s) = \max_{\tau\in\mathcal{P}, s\in[0,1]} \quad \mathbb{E}_\tau\left[V((1-s)\cdot p;\pi)\right] - C(\tau;F_o) - s\cdot p\cdot Q :$$

$$\text{s.t.} \quad (Q,p) \text{ is an equilibrium.}$$

The objective function, which we call 'welfare' in this section, represents the total surplus after discounting the social cost of the two policies. It can be interpreted as the social welfare when consumers fund the policies with a lump-sum tax. The main result of this section compares the solution of this problem with the traditional subsidy, s_0, required to obtain the same coverage in the absence of an information policy, and the welfare induced by this subsidy, W_0.

Proposition 6. *Consider the uniform-quadratic example with $0 < \kappa \leq 2$, and assume an information policy and a subsidy, (τ, s) solve the planner's problem for a fixed coverage $Q \in (0,1)$.*[17] *Then:*

1. *The optimal information policy, τ, is monotone: $\operatorname{supp}\tau = \{\pi_L, \pi_H\}$, and a buyer with belief π_H (π_L) buys (does not buy) with probability 1;*
2. *Consumers acquire no further information: for all $\pi \in \operatorname{supp}\tau$, $F^{(1-s)\cdot p}(p;\pi) = \delta_{\mu_\pi}$;*
3. *Subsidies are strictly lower with an information policy: $s < s_0$;*
4. *Welfare is strictly higher with an information policy: $W_0 < W(\tau, s)$.*

The content of the above is two-fold. Parts 1 and 2 show that the principal can choose an information policy that achieves three goals. First, it is monotone: that is, it partitions the state space into two. Equivalently, consumers are informed whether they are high or low risk. Second, regardless of their beliefs, consumers do not acquire further information. Finally, a consumer who learns his valuation is high, chooses to purchase the good for sure, whereas a consumer informed otherwise does not buy. Parts 1 and 2 can be seen as a revelation principle: for any of the possible information choices agents could make, the principal can anticipate that choice using information provision. This is possible, despite the constraint that the planner can only offer partitional policies.

Parts 3 and 4 show that, for any interior coverage, this combination of information policy and subsidies is strictly better than subsidies alone. Intuitively, by only subsidizing the market, the planner can counteract the distortion from adverse selection by affecting the level of demand, but not its shape. On the other hand, by first providing private information, the planner can manipulate the shape of demand. Although providing information increases the average cost curve, it also increases demand for consumers whose expected value is higher than the price. By making high-risk consumers more informed, the planner is also making them more optimistic about the value of insurance, increasing their willingness-to-pay. In our example, there is always a monotone information policy that makes this demand effect strong enough to improve welfare.

An implication of this result is that the social cost of achieving a certain level of insurance coverage is reduced by information policies. This is particularly important because it has been suggested that the cost of subsidies to increase coverage in some insurance markets may be substantial (Finkelstein et al., 2019). We illustrate the cost reduction in Figure 5b, which plots, for each targeted level of trade, Q, the subsidies required to achieve it, with and without an information policy — s and s_0, respectively. First, note that,

[17] The assumption of $\kappa \leq 2$ guarantees that if consumers have a private belief, π, from a partitional information policy, any information structure they might decide to acquire is monotone. This allows for the 'revelation-principle-like' result of this proposition.

in the absence of any intervention, the market would unravel. Indeed, without the information policy, the planner needs to provide large subsidies to generate any coverage. Furthermore, in that economy, a subsidy of 14% would be required to make sure 40% of the market is covered, at price .56. By contrast, the same 40% of market coverage can be attained with no additional subsidies if the planner offers an optimal information policy.

8 Discussion

In this section, we discuss how the results generalize beyond the main assumptions of the model.

Heterogeneous consumers. Consumers often have private information ex-ante, and their acquisition costs may be heterogeneous. In the Online Appendix B, we show how to extend our main results, in particular, Theorem 1 and Theorem 2, to a setting with heterogeneity. There are two difficulties. First, our approach applies the law of large numbers to obtain aggregate behavior, which naturally constrains heterogeneity. Second, on the converse, we use aggregate curves to infer individual consumers' decisions, which may be hindered by the presence of different individuals. To solve the first problem, we consider an arbitrary but finite number of groups of buyers. Within each group, consumers share the same information and acquisition costs, but those can vary arbitrarily across groups. We can then apply the law of large numbers within groups and combine the outcomes across groups to obtain market-level behavior.

To address the second issue, we prove a partial aggregation result. Although buyers are heterogeneous across groups, their optimal decisions maximize the utility of a fictitious consumer, subject to an array of information constraints. Moreover, this fictitious agent behaves similarly to the representative consumer in our baseline model. Thus, using aggregate curves to infer the fictitious buyers' behavior allows us to extend all the main results to this setting.

Firms' costs. The linearity of firms' costs substantially simplifies the previous analysis. In that case, both buyers' and firms' payoffs depend exclusively on the *mean* of consumers' beliefs about their own valuations. We show, in the Online Appendix C, that our results can be extended to general firms' costs by strengthening our assumptions about information costs. Concretely, we assume firms' costs are given by $\chi(\omega)$, where χ is a continuous and strictly increasing function. We show that all our results go through for any such χ, as long as information costs belong to a certain family of functions that include typical examples, such as mutual information and quadratic costs.

Because firms' costs are no longer linear, firms' payoffs now depend on the whole distribution of consumers' information, not only on its mean. As in the baseline model, consumers will acquire information that gives them one of two beliefs; recommending them to buy the good or not. In particular, if at price p consumers with belief $\pi \in \Delta[0,1]$ purchase the good, the average cost for the firm is $\mathbb{E}_{\pi}[\alpha(\omega)]$. The key difficulty in this setting is proving Theorem 2. To guarantee that more information induces higher costs for the firm, certain comparative statics must hold for beliefs, π, that recommend buying the product: π must be increasing in the first order stochastic dominance order with respect to the posterior mean it induces. By restricting attention to the family of information costs mentioned above, we guarantee this comparative statics result to be true.

The information threat guarantees rents for consumers, marking a benefit of endogenous information. In the presence of selection, this force competes with the worsening of sorting due to information acquisition, as demonstrated in Proposition 2. The net effect between these two forces determines whether consumers benefit or lose from endogenous information. We show that the force that dominates depends on the model parameters, so that the endogeneity of information can help or harm consumers, but the information threat is always dominant for large acquisition costs.

9 Conclusion

This paper studies the effects of endogenous information in adversely selected markets, and proposes methods for bringing the theory to data. Compared to traditional models of adverse selection, where information is exogenous, endogeneity hurts consumers by worsening selection. Exploiting this worsening of selection, we develop a test to detect endogenous information using aggregate data. Furthermore, we show that information decisions impose an externality in these markets, and information costs affect welfare non-monotonically. These results suggest that non-traditional policies, targeting consumers information choices, may worsen or help mitigate the inefficiencies from adverse selection, depending on the market. We propose an approach to evaluate some of these policies using observable data, and we show that information interventions can be welfare improving.

These results suggest several avenues for further research. Empirically, we believe that our framework will assist applied economists to quantify the presence of endogenous information in the data, as well as measure its equilibrium effects. This is a promising area of research, given recent evidence that endogenous information affects consumers' decisions (Brown and Jeon, 2020). Theoretically, the effects of endogenous information found in this and other recent papers suggests that endogeneity may be relevant for understanding a number economic environments (Mensch, 2022; Mensch and Ravid, 2022; Thereze, 2022). For example, endogenous information is likely to have important consequences in settings with product heterogeneity and imperfect competition.

Appendix

Proof of Lemma 1

Throughout, we fix $p \in [0,1]$ and $\kappa \geq 0$. We say a CDF F is admissible in the consumer's problem if it is in the choice set of consumers, that is $F \preceq_{\text{m.p.s.}} F_o$. Similarly, we say (U,Q) is admissible in CP if $(U,Q) \in \mathcal{C}$ and satisfies MPC.

Binary. Take any admissible F in the consumer's problem, and consider the information structure:

$$\hat{F} = F(p)\delta_{\mathbb{E}_F[\theta|\theta \leq p]} + (1 - F(p))\delta_{\mathbb{E}_F[\theta|\theta > p]}.$$

Clearly, $\hat{F} \preceq_{\text{m.p.s.}} F$. Because F was admissible, then so is $\hat{F}$ by transitivity: $\hat{F} \preceq_{\text{m.p.s.}} F_o$. We have:

$$\mathbb{E}_F[\max\{\theta - p, 0\}] - C(F) = (1 - F(p))\mathbb{E}_F[\theta|\theta > p] - C(F)$$
$$< (1 - F(p))\mathbb{E}_F[\theta|\theta > p] - C(\hat{F}) = \mathbb{E}_{\hat{F}}[\max\{\theta - p, 0\}] - C(\hat{F})$$

where the inequality comes from the strict monotonicity of F in the mean preserving spread order, and the last equality follows by definition of $\hat{F}$. Thus, we proved that any admissible information structure is dominated by a binary one, so the solution for the consumers' problem must be binary.

Uniqueness. Fix two admissible information structures, G and G' between which consumers are indifferent. Let $\alpha \in (0,1)$, and define $F = \alpha G + (1-\alpha)G'$. F is admissible, as the set of mean-preserving contractions of F_o is convex. Finally, define:

$$\hat{F} = F(p)\delta_{\mathbb{E}_F[\theta|\theta \leq p]} + (1 - F(p))\delta_{\mathbb{E}_F[\theta|\theta > p]}.$$

Clearly, $\hat{F}$ is admissible and, by the argument in the previous section, the consumer strictly prefers $\hat{F}$ to F. We now prove consumers prefer F to both G and G'.

$$\begin{aligned}
\mathbb{E}_F[\max\{\theta - p, 0\}] - C(F) &= \int_p^1 (\theta - p)\left(\alpha dG + (1-\alpha)dG'\right) - C\left(\alpha G + (1-\alpha)G'\right) \\
&\geq \alpha \int_p^1 (\theta - p)dG + (1-\alpha)\int_p^1 (\theta - p)dG' - \alpha C(G) - (1-\alpha)C(G') \\
&= \alpha\left\{\mathbb{E}_{\hat{G}}[\max\{\theta - p, 0\}] - C(\hat{G})\right\} + (1-\alpha)\left\{\mathbb{E}_{\hat{G}'}[\max\{\theta - p, 0\}] - C(\hat{G}')\right\} \\
&= \mathbb{E}_{\hat{G}}[\max\{\theta - p, 0\}] - C(\hat{G})
\end{aligned}$$

where the first equality used the definition of F, the inequality uses convexity of C, and the last equality uses the fact that the consumer is indifferent between G and G'. That implies the consumer weakly prefers F to both G and G'. Because the consumer strictly prefers $\hat{F}$ to F, there cannot be two optimal solutions.

Equivalence to CP. Start with an F admissible in the consumer problem, and we prove it induces an admissible pair (U,Q) in CP with the same value for the objective. If F is no-information, set $Q = \mathbb{1}_{\mu \geq p}$, $U = \mu Q$, it clearly obtains the same value for the objective function. It is also easy to check this is admissible in the original problem, so that proves the case of $|\operatorname{supp} F| = 1$. No assume $\operatorname{supp} F = \{\theta_L, \theta_H\}$, with $\theta_L < \theta_H$, and define (U,Q) as in text. That obtains the same value for the objective function. Next we prove (U,Q) is admissible. Recall $U = Q\theta_H$, which implies $\mu Q \geq U \leq Q$ — establishing two inequalities in $\mathcal{C}$. Finally, notice that, by Bayesian consistency:

$$U + (1-Q)\theta_L = \mu.$$

Thus, because $\theta_L \in [0,\mu]$, $\mu Q \leq U \leq \mu$, which establishes the final inequality in $\mathcal{C}$. Implying $(U,Q) \in \mathcal{C}$. By the integral form of mean-preserving contraction, because F is admissible:

$$I_F(x) \leq I_{F_o}(x).$$

Because F is binary, I_F can be written as:

$$I_F(x) = \begin{cases} 0 & \text{if } x \leq \theta_L \\ (1-Q)(x-\theta_L) & \text{if } \theta_L < x \leq \theta_H \\ x-\mu & \text{otherwise} \end{cases} \tag{3}$$

Notice that $(1-Q)(x-\theta_L) = (1-Q)x - \mu + U \geq I_F(x)$ for all x. Thus, by reorganizing the mean-preserving contraction inequalities in integral form:

$$\mu - U \geq \max_x \left\{(1-Q)x - I_{F_o}(x)\right\} = I^*_{F_o}(1-Q).$$

Thus, the F satisfies MPC, and is admissible in CP.

For the converse, take (U,Q) admissible in CP, with $U > \mu Q$. Note this implies $Q \in (0,1)$ because $(u,Q) \in \mathcal{C}$. Define $\theta_H = \frac{U}{Q}$ and $\theta_L = \frac{\mu - U}{1-Q}$. It is clear to see that, because $(U,Q) \in \mathcal{C}$, F has mean μ. Further:

$$I^*_{F_o}(1-Q) = \max_x \left\{(1-Q)x - I_{F_o}(x)\right\} \leq \mu - U \geq$$

which implies $(1-Q)x - I_{F_o}(x) \leq \mu - Q$ for all $x \in [\theta_L, \theta_H]$. By using the definition of I_F above, we obtain $I_F \leq I_{F_o}$, proving (U,Q) induces an admissible F in the consumer's problem. Furthermore:

$$U - pQ - \kappa c(U,Q) = Q(\theta_H - p) - \kappa C(F) = \mathbb{E}_F[\max\{\theta - p, 0\}] - \kappa C(F),$$

so the objectives coincide.

Finally, we deal with the case $U = \mu Q$. This case is not as straightforward, because, when $U = \mu Q$, (U,Q) maps into no-information, but it can be that no-information obtains a higher value in the consumer's problem than (U,Q) obtains in CP. However, this is not an issue: when that is the case, there exists (U', Q') that maps to no information — that is, $Q' \in \{0,1\}$ — which increases the objective in CP. Formally:

Take any (U,Q) with $U = \mu Q$. Recall that, for $U = \mu Q$, $c(U,Q) = 0$. So consider any $(U',Q') \in \arg\max_{q \in [0,1]}(\mu - p)q$. By definition, (U',Q') is better than (U,Q) in CP. Associate (U',Q') with no information: $F = \delta_\mu$, which is admissible in the consumer's problem. Then:

$$U' - pQ' = \max_{q \in [0,1]} (\mu - p)q = \max \mu - p, 0 = \mathbb{E}_F[\max\{\theta - p\}, 0]$$

Proving the two problems coincide again. This concludes the proof of equivalence.

Existence. The equivalence between problems shows we need only to prove existence for CP. Because $\mathcal{C}$ is compact and MPC is closed, we have the constraint set is compact. It suffices to prove $c(U,Q)$ is lower semi-continuous. Consider an admissible sequence $\{(U^n,Q^n)\}$ converging to (U,Q). Consider the associated sequence and limit information structures $\{F^n\}$, and F. Because C is lower semi-continuous in $\mathcal{L}^1$, it is sufficient to prove that convergence of (U^n,Q^n) to (U,Q) in $\mathbb{R}^2$ implies convergence of F^n to F in $\mathcal{L}^1$ in our admissible set. This is what we do next.

First, let $Q \in (0,1)$, and define $\theta_L^n = \frac{\mu - U_n}{1 - Q_n}$ and $\theta_H^n = \frac{U_n}{Q_n}$. Note the two must be well-defined for high enough n because $Q^n \to Q \in (0,1)$. Similarly, let $\theta_L = \frac{\mu - U}{1-Q}$ and $\theta_H = \frac{U}{Q}$. Additionally, for any two numbers, a, b we let $a \wedge b = \min\{a,b\}$, $a \vee b = \max\{a,b\}$. Moreover, for $i = \{L,H\}$, and $R \in \{\vee, \wedge\}$ we say $Q_{\theta_i^n R \theta_i} = Q\mathbb{1}_{\{\theta_i^n R\theta_i = \theta_i\}} + Q^n \mathbb{1}_{\{\theta_i^n R \theta_i \neq \theta_i^n\}}$. Thus:

$$\begin{aligned}\int_0^1 |F^n(x) - F(x)| dx = &\left(\theta_L^n \vee \theta_L - \theta_L^n \wedge \theta_L\right)(1 - Q_{\theta_L^n \wedge \theta_L}) \\ &+ (\theta_H^n \wedge \theta_H - \theta_L^n \vee \theta_L)\left|Q_{\theta_L^n \wedge \theta_L} - Q_{\theta_L^n \vee \theta_L}\right| \\ &+ \left(\theta_L^H \vee \theta_H - \theta_H^n \wedge \theta_H\right)\left|Q_{\theta_L^H \vee \theta_H} - Q_{\theta_H^n \wedge \theta_H}\right| \to 0\end{aligned} \tag{4}$$

where the first equality comes from the definition of the CDFs. The first and third terms converge to zero because the $\vee$ and $\wedge$ operators are continuous and $\theta_i^n \to \theta_i$, $i \in \{L, H\}$. The second term converges to zero because $Q^n \to Q$.

We finish by proving the result also holds for $Q = 0$. The proof is symmetric for $Q = 1$. When $Q = 0$, $(U, Q) \in \mathcal{C}$ implies $U = 0$, and the associated F is no information. If there is an N such that $(U^n, Q^n) = (U, Q)$ for all $n \geq N$, we are clearly done, because $F^n = F$ for $n \geq N$. Otherwise, there must exist a subsequence such that $Q^n \neq 0$, and because $Q^n \to 0$, this subsequence can be chosen with $Q^n < 1$. We focus on such subsequence from now on. For that, we can again define θ_L^n, θ_H^n as above. And

$$\begin{aligned}\int_0^1 |F^n(x) - F(x)| dx &= \left(\mu - \theta_L^n\right)(1 - Q^n) + (\theta_H^n - \mu) Q^n \\ &= \left(\mu - \frac{\mu - U^n}{1 - Q^n}\right)(1 - Q^n) + (U^n - \mu Q^n) \to 0\end{aligned} \tag{5}$$

where the second equality used the definitions of θ_i^n, and both terms clearly converge to zero by $(U^n, Q^n) \to (0, 0)$.

We have then proved c is lower semi-continuous, so the objective function in CP is upper semi-continuous. Because the constraint set is compact, existence is guaranteed.

■

Proof of Theorem 1

We start proving the first assertion. Notice that V_κ is naturally upper-semicontinuous in p by applying Berge's maximum theorem to problem CP. Moreover, the objective function is convex in p, so V_κ, being the maximum of convex functions is also convex in p, and thus must be continuous in $(0, 1)$. When $p = 0$, it is clear that one obtains no information and purchases the good with probability one, achieving the highest possible utility, $V_\kappa(0) = \mu$. Similarly, at $p = 1$ it is dominant not to acquire any information and purchase the good with probability 0, achieving $V_\kappa(1) = 0$. We use these facts to prove continuity at $p \in \{0, 1\}$.

For $p = 1$, upper semi-continuity and non-negativity of V_κ imply continuity — V_κ cannot jump downwards. For $p = 0$, take any sequence of prices $\{p^n\}$ with $p^n \to 0$, and consider the following strategy for the consumer: the consumer acquires no information and purchases the good with probability one. That strategy gives expected payoff $\mu - p^n$. We have:

$$V_\kappa(0) - p^n = \mu - p^n \leq V_\kappa(p^n) \leq V_\kappa(0),$$

where the first inequality holds because $V_\kappa(p^n)$ is the maximum payoff at p^n and the strategy of purchasing for sure is always feasible; and the second inequality holds because V_κ is decreasing in prices. By taking limits on the inequalities above, we obtain $V_\kappa(p_n) \to V_\kappa(0)$, proving V_κ is continuous at 1.

Because V_κ is continuous and convex in $[0,1]$, it follows that it is absolutely continuous and, by the envelope theorem:

$$V_\kappa(p) = V_\kappa(0) - \int_0^p (1 - F_\kappa^\nu(\nu))\,d\nu = V_\kappa(0) - \int_0^p D_\kappa(\nu)d\nu \tag{6}$$

By convexity of V_κ, it must be that D_κ is a decreasing function. By the argument of no information acquired for $p \in \{0,1\}$, $D_\kappa(0) = 1$ and $D_\kappa(1) = 0$. Finally, by 1, the optimal information structure is unique for each p.

We show this implies D_κ must be continuous in p, except possibly at $p = \mu$. To see that, assume for a contradiction that $p^n \to p$ but $D_\kappa(p^n) \to d \neq D_\kappa(p)$. Then, consider the associated sequence of optimal solutions for CP, (U^n, Q^n), and the solution (U, Q). Naturally, $Q^n = D_\kappa(p^n)$ and $Q = D_\kappa(p)$, so $Q^n \to d$. Because U^n is a bounded function, it has a convergent subsequence. Let u be the limit point of such subsequence. Finally, consider the strategy $\frac{1}{2}(u,d) + \frac{1}{2}(U,Q)$. It is clearly feasible and, by convexity of c must give at least the same utility to the agent at p then (U,Q). Thus, as long as one of (U,Q) and (u,d) is associated with some binary structure, this is a contradiction with optimality of (U,Q).

Therefore, any discontinuity in demand must happen at a point where the consumer acquires no information. Nevertheles, n that case, the consumer could only be indifferent between two trade probabilities if $p = \mu$. But note that, when $p = \mu$, $1 - D_\kappa(\mu) = 0$, so D_κ is a complementary CDF.

For the second assertion, note that, because at $p = 0$ no information is acquired, $V_\kappa(0) = V_{\kappa'}(0)$. Similarly, $V_\kappa(1) = V_{\kappa'}(1) = 0$. Then, applying 6, we have for $p \in [0,1]$:

$$0 \leq V_{\kappa'}(p) - V_\kappa(p) = \int_0^p D_\kappa(\nu)d\nu - \int_0^p D_{\kappa'}(\nu)d\nu$$

where the inequality stems from the problems being identical except for a higher information cost. Because D_κ is an inverse CDF, these inequalities imply that $1 - D_{\kappa'}$ second-order stochastically dominates $1 - D_\kappa$. The equality of value functions at $p \in \{0,1\}$ further implies the mean-preserving contraction relation.

■

Proof of Theorem 2

Fix some $Q \in [0,1]$, and let $p = P_\kappa(Q)$ and $p' = P_{\kappa'}(Q)$. At (κ, p), optimality of F_κ^p implies that utility must decrease if consumers, instead, act in the optimal way for (κ', p'). In other words, their utility decreases if they acquire $F_{\kappa'}^{p'}$ and purchase when their signal realization is above p':

$$\int_p^1 \theta dF_\kappa^p - \left(1 - F_\kappa^p(p)\right)p - \kappa C(F_\kappa^p) \geq \int_{p'}^1 \theta dF_{\kappa'}^{p'} - \left(1 - F_{\kappa'}^{p'}(p')\right)p - \kappa C(F_{\kappa'}^{p'})$$
$$\iff \kappa\left(C(F_{\kappa'}^{p'}) - C(F_\kappa^p)\right) \geq \int_p^1 \theta dF_\kappa^p - \int_{p'}^1 \theta dF_{\kappa'}^{p'}$$

where we used that, by definition, $Q = 1 - F_\kappa^p(p) = 1 - F_{\kappa'}^{p'}(p')$. Symmetrically, optimality at (κ, p) implies:

$$\kappa'\left(C(F_{\kappa'}^{p'}) - C(F_\kappa^p)\right) \leq \int_p^1 \theta dF_\kappa^p - \int_{p'}^1 \theta dF_{\kappa'}^{p'}.$$

Because $\kappa' > \kappa$, these inequalities imply: $C(F_{\kappa'}^{p'}) \leq C(F_\kappa^p)$. Now, by Lemma 1, information structures are at most binary binary. Thus, information structures can be represented by the probability of the purchase signal, and the expected utility under purchase. Let $F_\kappa^p = (Q, U)$ and $F_{\kappa'}^{p'} = (Q, U')$. It is easy to see that if $U' > U$, then $F_\kappa^p \preceq_{\text{m.p.s.}} F_{\kappa'}^{p'}$, which, by monotonicity of C, contradicts the inequality of costs above. Thus, it must be that $U' \leq U$, implying the result for average costs.

Finally, for $Q = 1$ all consumers must buy the good, and the average cost has to be $\mu \cdot \alpha$.

■

Proof of Remark 1

First, we prove a preliminary result relating monotone structures to the MPC constraint.

Monotone if and only if MPC binds. We start proving that if MPC binds, the structure is monotone. The converse is clear.

Let MPC bind for (U, Q), that is:

$$I_{F_o}^*(1-Q) = \max_x\{(1-Q)x - I_{F_o}(x)\} = \mu - U.$$

Because I_{F_o} is strictly convex, the optimality condition for the problem above is:

$$1 - Q = F_o(x),$$

so there is a solution x in $[0, 1]$ for the problem. Thus, for F associated with (U, Q):

$$I_{F_o}(x) = (1-Q)x + U - \mu = I_F(x)$$

Now, notice:

$$
\begin{aligned}
(1-F(x))\mathbb{E}_F[\theta|\theta \geq x] &= \mu - xF(x) + I_F(x) \\
&= \mu - xF_o(x) + I_{F_o}(x) = \\
&(1-F_o(x))\mathbb{E}_F[\theta|\theta \geq x]
\end{aligned}
\tag{7}
$$

where the first equality follows from integration by parts, and the second equality from the fact that $F_o(x) = 1 - Q = F(x)$, by the optimality condition, and $I_F(x) = I_{F_o}(x)$. That proves $\mathbb{E}_F[\theta|\theta \geq x] = \mathbb{E}_{F_o}[\theta|\theta \geq x]$. Thus, because the mean of the posterior means equals the prior, it must be that F is a monotone structure.

1 implies 3. Let $AC_\kappa(Q) = AC_{\kappa'}(Q)$. By assumption, consumers acquire some information in both cases. Now, using the notation of Lemma 1, associated, to the information structures at κ and κ', (U,Q) and (U',Q) respectively. By definition of average costs, it must be that:

$$U = U'$$

So the two information structures must be the same. For a contradiction, assume the information structure is not monotone. Thus, by our previous argument, it must be that MPC does not bind. As a consequence U can be determined by the following problem, which ignores that constraint:

$$U = \operatorname*{argmax}_u u - \kappa C(Q,u)$$

Because C is strictly increasing in U for fixed Q, this is a submodular optimization in (κ, u). Therefore, u is strictly decreasing in κ, which is in contradiction with $U = U'$. This proves the information structure is monotone for κ, κ'.

It is easy to see that if MPC holds for $\kappa' > 0$, it holds for all $\kappa'' < \kappa'$, which proves the monotonicity for all information costs smaller than κ'.

3 implies 2. We proved monotonicity if and only if MPC binds. Thus, 3 implies the same information structure is optimal for all $0 < \kappa'' < \kappa'$, which clearly implies the average cost is the same: $AC_{\kappa''} = AC_{\kappa'}$. We just need to extend the result for zero. For that, notice that when costs are zero, the consumer also chooses monotonically: the Q consumers purchasing the good are the ones with the highest valuation. Thus, the cost is also the same in that case.

2 implies 1. This is straightforward.

■

Proof of Proposition 1

$c(\cdot, Q)$ is Lipschitz for all $Q \in (0,1)$. We assumed C is Lipschitz continuous. Consider two pairs $(U, Q), (U', Q)$ with $Q \in (0,1)$, associated with F and F' respectively. Without loss of generality, let $U' > U$. Then:

$$F(x) = \begin{cases} 0, \text{ if } x < \frac{\mu - U}{1-Q} \\ 1-Q, \text{ if } x \in \left[\frac{\mu-U}{1-Q}, \frac{U}{Q}\right] \\ 1, \text{ otherwise} \end{cases}$$

and

$$F'(x) = \begin{cases} 0, \text{ if } x < \frac{\mu - U'}{1-Q} \\ 1-Q, \text{ if } x \in \left[\frac{\mu-U'}{1-Q}, \frac{U'}{Q}\right] \\ 1, \text{ otherwise} \end{cases}$$

We then have:

$$|c(U,Q) - c(U',Q)| = |C(F) - C(F')| \leq M\|F - F'\| = M\int |F(x) - F'(x)|dx =$$
$$M\left(\int_{\frac{\mu-U'}{1-Q}}^{\frac{\mu-U}{1-Q}} (1-Q)dx + \int_{\frac{U}{Q}}^{\frac{U'}{Q}} Qdx\right) = 2M(U' - U).$$

Where the inequality comes from Lipschitz continuity of C. So we proved Lipschitz continuity of $c(\cdot, Q)$. Furthermore, notice that the constant $\tilde{M} = 2M$ holds for all Q.

There exists $\underline{k}$ such that Information Structures are monotone for all $\kappa \leq \underline{k}$ and all $Q \in [0,1]$. We consider the choice of an information structure by the consumer for each Q. For $Q \in \{0,1\}$, the consumer acquires no information for all κ, so the information structure is always monotone. For $Q \in (0,1)$:

$$\max\left\{U - \kappa c(U,Q) : I^*_{F_o}(1-Q) \leq \mu - U\right\}$$

Recall c is a convex function, so this is a convex optimization. Assume MPC does not bind. Then, a solution to this problem must satisfy the first-order subgradient condition at the optimal U^*:

$$\frac{1}{\kappa} \in \partial c(U^*, Q)$$

But we proved $c(\cdot, Q)$ is Lipschitz continuous for $Q \in (0,1)$, accepting the same constant, $\tilde{M}$, for all Q. As a consequence, the subgradients of $c(\cdot, Q)$ are uniformly bounded by $\tilde{M}$. Thus, if $\frac{1}{\kappa} \in \partial c(U^*, Q)$ for any

Q, it must be that $\frac{1}{\kappa} \leq \tilde{M}$. Set $\overline{\kappa} = \frac{1}{\tilde{M}}$, and we have that, MPC does not bind only if $\kappa \geq \overline{\kappa}$. By Remark 1, this implies information structures are monotone for $\kappa < \overline{\kappa}$.

Welfare is non-monotonic. By Lemma OA1, P_o and AC_o intersect only once. Take κ small enough such that P_κ still crosses AC_o only once: that is possible by continuity of P_κ in κ. We proved above that there exists a $\underline{\kappa} > 0$ such that MPC holds for all Q and for all $\kappa < \underline{\kappa}$, so choose one such $\kappa < \underline{\kappa}$.

By Theorem 1, there exists $\underline{Q}$ such that $P_\kappa(Q) < P_o(Q)$ for all for $Q < \underline{Q}$. Pick α such that the equilibrium quantity at $\kappa = 0$ is smaller than $\underline{Q}$. To see this is possible, notice that, when $\alpha = 1$, the equilibrium quantity is zero. Equilibrium is continuous in α by Lemma OA1, as P_o and AC_o intersect only once for all α, and equilibrium quantities are clearly decreasing in α, so by choosing α close to 1, one would get the equilibrium quantity at zero, Q_o^e, such that $0 < Q_o^e < \underline{Q}$.

Because $P_\kappa(Q) < P_o(Q)$ and $AC_\kappa(Q) = AC_o(Q)$, the equilibrium quantity must be smaller. Because $AC_\kappa = AC_o$ and AC_o is decreasing, this implies equilibrium prices must be higher. Finally, because information is costly at κ, the consumer must have paid non-negative information costs. All these changes make the consumers worse off. Because firms make zero profits both at 0 and at κ, welfare decreases for small κ.

For the other side of non-monotonicity, note that market efficiency is achieved when consumers acquire no information in equilibrium. In that case, the price is also the lowest possible, as $p = \alpha\mu \leq AC_\kappa(Q)$ for all $\kappa \geq 0$ and $Q \in [0, 1]$. As $\kappa \to \infty$, we have that the solution to the consumer information problem is, for each $Q \in (0, 1)$:

$$0 \leftarrow \frac{1}{\kappa} \in \partial c(U_\kappa^*, Q).$$

Because c is strictly increasing in U, this implies U_κ^* converges to the lower bound $U_\kappa^* \to \mu Q$, so $AC_\kappa(Q) \to \alpha\mu$, for all $Q > 0$. It must then be that, for all p such that $D_\kappa(p) > 0$, $R_\kappa(p) \to \alpha\mu$, and, thus, equilibrium prices converge to $\alpha\mu$, implying equilibrium quantities converge to 1. Thus, as $\kappa \to \infty$, the equilibrium converges to the most efficient equilibrium, so welfare must increase and the equilibrium price must decrease.

Unraveling. Start by noting that α does not affect the consumers' decision. Thus, if we define $AC_\kappa(Q, \alpha)$ as the average cost when information costs are κ and firms' costs are α, we have, for all $\alpha \in [0, 1]$, $AC_\kappa(Q, \alpha) = \alpha AC_\kappa(Q, 1)$.

For $\kappa = 0$ and $\alpha = 1$, we have, for all $Q > 0$

$$AC_o(Q, 1) = \mathbb{E}_{F_o}[\omega | \omega \geq F_o^{-1}(1 - Q)] > F_o^{-1}(1 - Q) = P_o(Q)$$

And

$$AC_o(0,1) = P_o(0).$$

For any $\kappa > 0$, by Theorem 1, $P_\kappa(0) < P_o(0)$. By continuity, there exists some $\tilde{\kappa} < \underline{\kappa}$ such that $P_\kappa(Q) < AC_o(Q,1) = AC_{\tilde{\kappa}}(Q,1)$. Because the inequality is strict and $AC_{\tilde{\kappa}}(Q,\cdot)$ is continuous, continuity again implies there exists $\tilde{\alpha}$ such that $AC_{\tilde{\kappa}}(Q,\alpha) > P_{\tilde{\kappa}}(Q)$ for all $\alpha > \tilde{\alpha}$ and Q.

Now, fix this $\tilde{\alpha}$. Let $Q^e_\kappa(\alpha)$ define equilibrium quantities for information cost κ and firm cost α. We proved $\tilde{\kappa} \in \{\kappa : Q^e_\kappa(\tilde{\alpha}) = 0\}$, so this set is non-empty. It is clear that 0 is not in this set by Remark 2. Because $\{\kappa : P_\kappa(Q) = AC_\kappa(Q) \text{ for some } Q\} \subset \{\kappa : Q^e_\kappa(\tilde{\alpha}) > 0\}$, and P_κ and AC_κ are continuous in κ, 0 must be in the interior of an open set. Thus, we can find an interval $\mathcal{K}_{\tilde{\alpha}}$, for $\tilde{\alpha}$ in the form of the statement.

Finally, notice that for $\alpha > \tilde{\alpha}$, $\mathcal{K}_{\tilde{\alpha}} \subset \{\kappa : Q^e_\kappa(\alpha) = 0\}$. To see that, take $\kappa \in \mathcal{K}_{\tilde{\alpha}}$. It must be that $AC_\kappa(Q,\tilde{\alpha}) > P_\kappa(Q)$ for all $Q > 0$. Thus, for all $Q > 0$:

$$AC_\kappa(Q,\alpha) > AC_\kappa(Q,\tilde{\alpha}) > P_\kappa(Q).$$

Thus, $\kappa \in \{\kappa : Q^e_\kappa(\alpha) = 0\}$. Then, by setting $\mathcal{K}_\alpha = \mathcal{K}_{\tilde{\alpha}}$ for all $\alpha > \tilde{\alpha}$, we have the final result.

■

Proof of Remark 2

Fix any distribution F, and $\alpha < 1$. Notice the highest equilibrium price is determined by:

$$p(F) = \inf_{p \in [0,1]} \{p : \alpha \mathbb{E}_F[\omega | \omega \geq p] < p\}.$$

We start showing the set is non-empty. For that, let $\max \operatorname{supp} F = \overline{\theta}$. Then, we must have:

$$\mathbb{E}_F[\omega | \omega \geq \overline{\theta}] = \alpha \overline{\theta} < \overline{\theta}$$

We then have that, for $p(F) \leq \overline{\theta}$. If $\overline{\theta}$ is a mass point, with mass m, we are done, because $Q(F) \geq m > 0$.

Assume F continuous at $\overline{\theta}$. As a consequence, the conditional expectation operator is continuous, and the inequality above implies there exists $\varepsilon > 0$ such that

$$\mathbb{E}_F[\omega | \omega \geq \overline{\theta} - \varepsilon] < \overline{\theta} - \varepsilon$$

which implies $p(F) \leq \overline{\theta} - \varepsilon$, and thus

$$Q(F) > 1 - F(\overline{\theta} - \varepsilon) > 0,$$

by continuity of F.

■

Proof of Proposition 2

We prove the more general result for when firms' costs are a strictly increasing function $\chi : [0,1] \to \mathbb{R}_+$. In that case, consumers acquire a distribution over posteriors τ in equilibrium, and we let π^e be the posterior under which consumers decide to purchase the product.

Part 1. By definition of equilibrium, we must have: $R_\kappa(p^e) = p^e$, which can be rewritten as:

$$\mathbb{E}_\tau[\chi(\omega)|\theta \geq p^e] = \mathbb{E}_{\pi^e}[\chi(\omega)] = p^e.$$

Because τ^e is binary, we have, for $\varepsilon > 0$ small enough, $\mathbb{E}_\tau[\chi(\omega)|\theta \geq p^e + \varepsilon] = \mathbb{E}_\tau[\chi(\omega)|\theta \geq p^e] < p^e + \varepsilon$. Therefore, the exogenous equilibrium price $p(\tau)$ is such that $p(\tau) \leq p^e$.

Finally, because the endogenous and exogenous demand coincide at Q^e, we must have $Q(\tau) \geq Q^e$.

Part 2. Assume now $(p(\tau), Q(\tau)) \neq (p^e, Q^e)$. If $Q^e \in \{0,1\}$, there is nothing to prove. Indeed, if $Q^e = 0$, welfare is minimal and thus cannot be smaller under exogenous information. Conversely, if $Q^e = 1$, we know that under exogenous information the quantity must also be one, and weakly lower prices imply consumers must be better off. Henceforth, we assume $Q^e \in (0,1)$.

Because τ is binary, the function

$$R(p) = \mathbb{E}_\tau[\chi(\omega)|\theta \geq p]$$

is a step-function. As a consequence, the set $\{p : R(p) = p\}$ has at most two points: one with full coverage, and one in which only consumers with signal π^e_H purchase. Because both p^e and $p(\tau) < p^e$ are in that set, it must have exactly two points. As a consequence, the market must be at full coverage under exogenous information. That is, $Q(\tau) = 1$ and:

$$p(\tau) = \mathbb{E}_{F_o}[\chi(\omega)].$$

Let F, with $\operatorname{supp} F = \{\theta_L, \theta_H\}$ be the distribution over posterior means implied by τ. By full coverage,

$p(\tau) < \theta_L$. We then have:

$$W(\tau) = \mu - \mathbb{E}_{F_o}[\chi(\omega)] \geq \mu - \theta_L$$

where the inequality comes from $R(p(\tau)) = p(\tau) < \theta_L$. Conversely:

$$W^e = (\theta_H - p^e)Q^e < (\theta_H - \theta_L)Q^e = \mu - \theta_L,$$

where the inequality follows from $R(p^e) = p^e > \theta_L$, because equilibrium prices excluded low-signal individuals. Moreover, the last equality follows from the fact that F has the same mean as the prior — the mean-preserving contraction constraint.

Comparing the two inequalities above we obtain the result. ■

Proof of Proposition 3

Start with 2 observations. First, for any $G \in \Delta([0,1])$:

$$\int_p^1 \theta dG(z) = 1 - pG(p) - \int_p^1 G(z)dz. \tag{a}$$

Second, when $\mathcal{O} = \{P_\kappa, AC_\kappa\}$:

$$D_\kappa(P(Q)) = Q = 1 - F_\kappa^{P(Q)}(P(Q)). \tag{b}$$

Fix $Q \in (0,1]$, let $p = P(Q)$ and assume $\mathcal{O}$ is consistent, so $AC_\kappa = AC$. Then:

$$\begin{aligned}\frac{Q}{\alpha}(AC_\kappa(Q) - \overline{AC}(Q)) &= 1 - pF_\kappa^p(p) - \int_p^1 F_\kappa^p(z)dz - 1 + p(1 - D_\kappa(p)) + \int_p^1 (1 - D_\kappa(p))dz \\ &= \int_p^1 F^z(z)dz - \int_p^1 F^p(z)dz\end{aligned}$$

where the first equality follows from observation (a) and the second from (b) and recalling the definition of D_κ. Thus, it is sufficient to prove the last term is non-negative.

Define $U(l) = \max_{d(\theta)\in\{0,1\}}\{\mathbb{E}_{F^p}[d(\theta)(\theta - l)] - \kappa C(F^p)\}$. We can interpret U as the utility of a buyer who is obliged to pay for and receive a signal from information structure F^p, but can freely decide whether to purchase or not the good. Note that $U(1) = -\kappa C(F^p)$, since at $p = 1$ not purchasing the good is an optimal decision for any information structure. Further, by the envelope theorem: $U(l) = U(0) + \int_0^l (F^p(z) - 1)dz$. Finally, $U(p) = V_\kappa(p)$, by construction. Thus:

$$\int_p^1 (F^p(z) - F^z(z))\,dz = U(1) - U(p) - V(1) + V(p)$$

$$= U(1) - V(1) = -\kappa C(F^p) \le 0$$

The inequality is strict as long as some information is acquired. It is easy to see that whenever the agent acquires no information he chooses $Q \in \{0,1\}$, proving the strict part of the result. ■

Proof of Proposition 4

Assuming differentiability, we can look at the first order conditions of problem CP.

Part 1. Follows as a corollary of Theorem 2.

Part 2. Start assuming that, at the observed κ, MPC does not bind. First order conditions then read:

$$\kappa c_U(U,Q) = 1 \qquad\qquad -\kappa c_Q(U,Q) = p. \tag{8}$$

Using the fact that $P(Q)$ must solve these equations for all Q and that $AC(Q) = \alpha\frac{U(Q)}{Q}$ we can differentiate the first equality above above with respect to Q to obtain:

$$\frac{dU}{dQ} = -\frac{c_{UQ}}{c_{UU}}. \tag{9}$$

Similarly, by differentiating the system in 8 with respect to κ:

$$\frac{dU}{d\kappa} = -\frac{1}{\kappa^2 c_{UU}}.$$

and

$$\frac{dp}{d\kappa} = \frac{1}{\kappa}\left(p + \frac{c_{QU}}{c_{UU}}\right) = \frac{1}{\kappa}\left(p - \frac{dU}{dQ}\right)$$

where the last equality used 9. Thus, by reorganizing this equation:

$$\varepsilon_{P,\kappa} = 1 - \frac{dU}{dQ}\frac{1}{P(Q)} \tag{10}$$

Now, note that $AC(Q) = \alpha\frac{U}{Q}$, so that:

$$\frac{dU}{dQ} = Q\frac{AC'(Q)}{\alpha} + \frac{AC}{\alpha}.$$

Plugging that back in 10

$$\varepsilon_{P,\kappa} = 1 - \frac{dU}{dQ}\frac{1}{P(Q)} = 1 - \left(\frac{QAC'}{\alpha P(Q)} + \frac{AC}{\alpha P(Q)}\right) = 1 + \left(\varepsilon_{AC,Q} + 1\right)\frac{AC(Q)}{\alpha P(Q)}.$$

Therefore, the Part 2 holds when MPC does not bind.

Now, assume instead MPC binds, so $U = \mu - I^*_{F_o}(1-Q)$. We can then rewrite the optimization CP as:

$$\max_{Q\in[0,1]} \mu - I^*_{F_o}(1-Q) - pQ - \kappa c\left(\mu - I^*_{F_o}(1-Q), Q\right),$$

and one can obtain first order conditions by differentiating with respect to Q:

$$F_o^{-1}(1-Q) - p = \kappa c_U F_o^{-1}(1-Q) + \kappa c_Q.$$

Differentiating the first order condition with respect to Q, we obtain:

$$\frac{dP(Q)}{d\kappa} = \frac{P(Q) - F_o^{-1}(Q)}{\kappa},$$

that can again be reorganized to $\varepsilon_{P,\kappa} = 1 - \frac{F_o^{-1}(Q)}{P(Q)}$.

Now, notice that, because the information structure is monotone, $AC(Q) = \frac{\alpha}{Q}\int_{F_o^{-1}(1-Q)}^{1} \theta f_o(\theta)d\theta$. Differentiating with respect to Q and reorganizing obtains:

$$F_o^{-1}(1-Q) = \frac{1}{\alpha}\left(QAC'(Q) + AC(Q)\right)$$

Finally, we can plug this expression back in the expression for $\varepsilon_{P,\kappa}$ above to recover the final result.

$$\varepsilon_{P,\kappa} = 1 + \left(\varepsilon_{AC,Q} + 1\right)\frac{AC(Q)}{\alpha P(Q)}$$

■

Proof of Remark 3

Fix $Q \in [0,1]$. Making the dependence on α explicit, the planner's information decision is defined by:

$$U^*_\alpha(Q) = \arg\max_U \{(1-\alpha)U - \kappa c(U,Q) : MPC\}.$$

Start considering the relaxed problem:

$$\tilde{U}^*_\alpha(Q) = \arg\max_U \{(1-\alpha)U - \kappa c(U,Q)\}. \tag{11}$$

Notice that $\tilde{U}^*_\alpha(Q) = U^*_\alpha(Q)$ if and only if: $I^*_{F_o}(1-Q) \leq \mu - \tilde{U}^*_\alpha(Q)$, that is, whenever the relaxed solution satisfied MPC. Otherwise:

$$\tilde{U}^*_\alpha(Q) > \mu - I^*_{F_o}(1-Q) \geq U^*_\alpha(Q)$$

Thus: with equality if and only if MPC holds in the original problem.

$$\tilde{U}^*_\alpha(Q) \geq U^*_\alpha(Q) \tag{12}$$

Note the objective function in the relaxed problem 11 is strictly submodular in (α, U). Therefore, $\tilde{U}^*_\alpha(Q)$ is strictly decreasing in α as long as $\tilde{U}^*_\alpha(Q) > \mu Q$.

Finally, by Lemma 1:

$$U(Q) = \arg\max_U \{U - \kappa c(U,Q) : MPC\} = U^*_0(Q).$$

We now consider three cases. First, assume $\tilde{U}^*_0(Q) = U^*_0(Q) > \mu Q$. Then:

$$U^*_0(Q) = \tilde{U}^*_0(Q) > \tilde{U}^*_\alpha(Q) = U^*_\alpha(Q),$$

where the last equality holds because $\tilde{U}^*_0(Q) > \tilde{U}^*_\alpha(Q)$, and $\tilde{U}^*_0(Q)$ satisfies MPC.

For the second case, assume $\tilde{U}^*_0(Q) = U^*_0(Q) = \mu Q$. In this case, the consumer acquires no information. Thus:

$$\begin{aligned} U^*_\alpha(Q) &\leq \tilde{U}^*_\alpha(Q) \\ &\leq \tilde{U}^*_0(Q) = U^*_0(Q) = \mu Q \leq U^*_\alpha(Q), \end{aligned} \tag{13}$$

where the first inequality follows from (12), the second second inequality by submodularity of the relaxed problem, the first equality by the fact that, at μQ, MPC always holds, and the last inequality because $(U^*_\alpha(Q), Q) \in \mathcal{C}$. Thus, for this case the equality in the problem equality holds.

Finally, consider the third case: $\tilde{U}^*_0(Q) > U^*_0(Q) \geq \mu Q$. That means, at zero, MPC binds. In that case:

$$U^*_\alpha(Q) \leq I^*_{F_o}(1-Q) - \mu = U^*_0(Q),$$

where the first inequality is from $U^*_\alpha(Q)$ satisfyind MPC. The equality in the remark then holds if and only if MPC binds for $U^*_\alpha(Q)$, concluding the proof.

■

Proof of Proposition 5

We start by decomposing the welfare loss in three terms:

$$\begin{aligned} W\left(U^*(Q^*),Q^*\right)-W\left(U(Q^e),Q^e\right) &= W\left(U^*(Q^*),Q^*\right)-W\left(U^*(Q^{MC}),Q^{MC}\right) \\ &+W\left(U^*(Q^{MC}),Q^{MC}\right)-W\left(U(Q^{MC}),Q^{MC}\right) \\ &+W\left(U(Q^{MC}),Q^{MC}\right)-W\left(U(Q^e),Q^e\right) \end{aligned} \tag{14}$$

Notice $W\left(U^*(Q^*),Q^*\right)-W\left(U^*(Q^{MC}),Q^{MC}\right)\geq 0$, because Q^* is welfare-maximizing, and $W\left(U^*(Q^{MC}),Q^{MC}\right)-W\left(U(Q^{MC}),Q^{MC}\right)\geq 0$ because, conditional on Q^{MC}, U^* is the welfare-optimal information decision. We prove the last term is equal to $L\left(Q^e,Q^{MC}\right)$. For that, notice:

$$\begin{aligned} W\left(U(Q^{MC}),Q^{MC}\right)-W\left(U(Q^e),Q^e\right) &= \underbrace{V_\kappa\left(P_\kappa\left(Q^{MC}\right)\right)}_{\text{Consumer surplus at } Q^{MC}} + \\ &\underbrace{P_\kappa(Q^{MC})Q^{MC}-\alpha U\left(Q^{MC}\right)}_{\text{Profits at } Q^{MC}} - \underbrace{V_\kappa\left(P_\kappa\left(Q^e\right)\right)}_{\text{Consumer surplus at } Q^e} \\ = -\underbrace{\int_{Q^{MC}}^{Q^e} Q\,dP(Q)}_{\text{Difference of Consumer Surplus}} &+ \underbrace{\int_{Q^{MC}}^{Q^e}\left(Q\,dP(Q)+P_\kappa(Q)dQ-\alpha\, dU(Q)\right)}_{\text{Difference in Profits}} \\ &= \int_{Q^{MC}}^{Q^e}\left(P_\kappa(Q)-MC_\kappa(Q)\right)dQ = L\left(Q^e,Q^{MC}\right). \end{aligned} \tag{15}$$

The first equality is by definition, recalling profits are zero at Q^e. For the second equality, we notice first that $V'_\kappa(p)=-Q$. Using the chain rule for V_κ and noticing U and P are continuous and monotone, and thus of bounded variation, the Riemmann-Stjeltjes integrals above exist. The second to last equality recognizes $\alpha\, dU(Q)$ is the marginal cost curve. Finally, the last equality is by definition of L — the are between inverse demand and marginal cost on the vertical axis, and Q^{MC} and Q^e, on the horizontal.

Therefore, we have: $W\left(U^*(Q^*),Q^*\right)-W\left(U(Q^e),Q^e\right)\geq L(Q^e,Q^{MC})$. We now prove then the equality holds. First, assuming $U^*(Q^{MC})\neq U(Q^{MC})$ guarantees inequality holds. Thus, by Remark 3, a necessary condition for equality is $U^*(Q^{MC})=U(Q^{MC})$. Again, by Remark 3, this necessary condition holds if and only if the planner acquires monotone information at Q^{MC}, or that the consumer acquires no information. We prove these are also sufficient for equality.

Start assuming the planner acquires monotone information at Q^{MC}. We prove $Q^*=Q^{MC}$ in that case.

To see that, recall Remark 3 implies $U(Q^{MC}) = U^*(Q^{MC})$, so that consumers also acquire monotone information. Consider the first order conditions for the planner's problem:

$$\alpha_o = -c_Q(U^*(Q), Q) + ((1-\alpha) - c_U(U^*(Q), Q)) F_o^{-1}(1-Q)$$

and for the consumer's problem:

$$p = -c_Q(U(Q), Q) + (1 - c_U(U(Q), Q)) F_o^{-1}(1-Q)$$

Now, in both conditions let $Q = Q^{MC}$. Because $U(Q^{MC}) = U^*(Q^{MC})$, and, by definition of Q^{MC} $P(Q^{MC}) = \alpha_o + \alpha F_o^{-1}(1-Q)$, we have that both conditions hold. Because first order conditions for the planner's problem are necessary and sufficient, $Q^{MC} = Q^*$. So we proved in this case $W(U^*(Q^*), Q^*) - W(U(Q^e), Q^e) = L(Q^e, Q^{MC})$.

Now, assume individuals acquire no information at Q^{MC}. That implies the planner acquires no information at Q^{MC}, and $U(Q^{MC}) = U^*(Q^{MC})$, so we can use first order conditions, once more, but with inequalities, to prove $Q^* = Q^{MC}$. But then, the principal acquires no information at Q^*, and optimality requires $Q^* \in \{0, 1$. We have thus proved the result.

■

Proof of Proposition 6

Step 1. For any partitional initial belief, consumers acquire a monotone structure. To see that, notice that any partitional belief must be of the form $F_i \sim U[\underline{\theta}, \overline{\theta}]$, where $0 \leq \underline{\theta} \leq \overline{\theta} \leq 1$. Thus, we have:

$$I^*_{F_i}(1-Q) = \frac{(\overline{\theta}(1-Q) + \underline{\theta}Q)^2}{2(\overline{\theta} - \underline{\theta})}$$

The consumers' problem is:

$$\max_{(U,Q) \in \mathcal{C}} \left\{ U - pQ - \frac{\kappa}{2} \frac{1}{Q(1-Q)} \left(U - \frac{Q(\overline{\theta} + \underline{\theta})}{2} \right)^2 : I^*_{F_i}(1-Q) \leq \frac{\overline{\theta} + \underline{\theta}}{2} - U \right\}$$

Consider the relaxed problem that does not impose MPC. We show that the relaxed solution does not satisfy the constraint, so MPC must bind. Fixing any Q and taking first order conditions, we obtain that the relaxed solution must satisfy:

$$\tilde{U}(Q) = \left(\frac{\overline{\theta} + \underline{\theta}}{2} + \frac{1-Q}{\kappa} \right) Q$$

We then have:

$$\frac{\overline{\theta}+\underline{\theta}}{2}-\tilde{U}(Q)-I^*_{F_i}(1-Q)=\frac{1}{2}\left(\overline{\theta}\left(-\frac{\underline{\theta}}{\overline{\theta}-\underline{\theta}}-Q^2+Q\right)+\frac{(Q-1)Q(k\underline{\theta}+2)}{k}+\underline{\theta}\right)$$

Notice this function is minus infinity at $\kappa=0$ and increasing in κ. Choosing $ka\tilde{p}pa$ to set it at zero, one obtains:

$$\tilde{\kappa}=2\frac{(\overline{\theta}-\underline{\theta})Q(1-Q)}{(\overline{\theta}-\underline{\theta})^2Q(1-Q)-\underline{\theta}^2}$$

Clearly, $\tilde{\kappa}>2$ or $\tilde{\kappa}<0$. The negative root is irrelevant, because this function becomes minus infinity at zero. Because the function is increasing, it means that it is always negative for $\kappa\leq 2$. Thus, MPC cannot hold for $\tilde{U}$. This proves the solution to the consumer problem must bind MPC. By 1, the consumer chooses a monotone information structure, and we conclude this part of the proof.

Step 2. Consumer information for a monotone partition. Consider a monotone policy: the monopolist chooses a threshold z, and tells consumers whether their valuation is higher or lower than z. That is, high-signal consumers know their value is in $[z,1]$. Assume the effective price $\tilde{p}=(1-s)p$ is higher than z. We start calculating the optimal information structure obtained by those that are informed their type is above z. We know MPC must bind. We then use the first order condition of the consumers' problem to solve for the optimal threshold $x\in[z,1]$ that maximizes the consumers' utility. That threshold satisfies:

$$x_\kappa(p,z)=\frac{\tilde{p}-(1-z^2)\frac{\kappa}{8}}{1-(1-z)\frac{\kappa}{4}}, \tag{16}$$

Step 3. Proof of Parts 1 and 2. Let information policy τ and a subsidy s implement an equilibrium (Q,p) with effective price $\tilde{p}=(1-s)p$. Assume τ is defined by a family of disjoint intervals $\{[\underline{\theta}_\gamma,\overline{\theta}_\gamma]\}_{\gamma\in 1,\dots,N}$, and $\{F_\gamma\}_{\gamma\in 1,\dots,N}=\operatorname{supp}\tau$ — that is, F_γ is uniform in the set $[\underline{\theta}_\gamma,\overline{\theta}_\gamma]$. Without loss, let $\overline{\theta}_\gamma<\underline{\theta}_{\gamma+1}$, for $\gamma<N$. Finally, let the optimal information structures for a price p be: $\operatorname{supp}F^p_\gamma=\{\theta_{L,\gamma},\theta_{H,\gamma}\}$, and let Q_γ be the probability of $\theta_{H,\gamma}$.

First, let us prove there is no additional cost in providing consumers directly the information $\{F^p_\gamma\}_\gamma$. To see that, define F_τ as the distribution over posterior means induced by τ. Then, define τ' as the information policy that delivers to consumers their ex-post beliefs. That is, if consumer would receive signal $\theta_{i,\gamma}$ after acquiring information, the information policy reveals this directly. Then, the total cost satisfies:

$$\mathbb{E}_{F_\tau}\left[(\theta-\mu)^2\right]+\mathbb{E}_\tau\left[\mathbb{E}_{F_\gamma^p}\left[(\theta-\mu_\gamma)^2\right]\right]=$$
$$\sum_\gamma \tau_\gamma\left\{(\mu_\gamma-\mu)^2+\mathbb{E}^p_{F\gamma}[(\theta-\mu_\gamma)]\right\} \tag{17}$$
$$=\sum_\gamma\left\{\tau_\gamma Q_\gamma\left(\theta_{H,\gamma}-\mu\right)^2+\tau_\gamma(1-Q_\gamma)\left(\theta_{L,\gamma}-\mu\right)^2\right\}=\mathbb{E}_{\tau'}\left[(\theta-\mu)\right],$$

where τ' is an information structure that is partitional and anticipates their information decisions. There is, of course, no guarantee that, upon offering τ' consumers will not acquire different information. We next prove that we can choose information so that this is the case.

Note that there exist a unique $\tilde{\gamma}$ such that $\tilde{p}\in[\underline{\theta}_{\tilde{\gamma}},\overline{\theta}_{\tilde{\gamma}}]$. For all $\gamma\neq\tilde{\gamma}$, information is irrelevant — they either purchase the good with certainty if $\gamma>\tilde{\gamma}$, or disregard the good with certainty. Denote F_γ^p as the optimal information structure at equilibrium price p for belief F_γ. Then:

$$\operatorname{supp} F_\gamma^p=\left\{\frac{\overline{\theta}_\gamma+\underline{\theta}_\gamma}{2}\right\},$$

for all $\gamma\neq\tilde{\gamma}$. Thus, the only non-trivial optimal information structure is for $\tilde{\gamma}$. By the previous result, $F_{\tilde{\gamma}}^p$ is monotone. By the monotonicity structure, it must be that the Q highest value consumers purchase the good. Consider the alternative, monotone, information policy, $\tilde{\tau}$, that partitions the state space in $[0,z),[z,1]$. Choose

$$z=\frac{-4+k+2\sqrt{4-2k(1-\tilde{p})}}{k}.$$

By inspection of the result in Step 2, we can see that $x_\kappa(\tilde{p},z)=z$, so this information policy satisfies the assumptions of 1 and 2. We now prove it dominates τ. Indeed, it is constructed so that (p,Q) is an equilibrium. Finally, it strictly saves in costs, because $\tilde{\tau}$ is less informative than τ', as it just reveals one threshold.

Step 4. Proof of Parts 3 and 4. To obtain this result, we denote the information policies above as 'monotone policies', and identify them with the threshold z. Given a level of subsidy, s, there is a unique 'monotone policy' z that guarantees equilibrium. This is because the monotone policy must have the property that everyone who receives the high signal buys — and no one who receives the low signal buys. That is, the equilibrium price for zero profits must be: $p=\alpha\frac{1+z}{2}$. The equation determining the threshold z solves:

$$x_\kappa((1-s)*(\alpha\frac{(1+z)}{2}),z)=z$$

That is, consumers optimally choose not to learn anything after receiving z. The solution to this equation is:

$$z_\kappa(s) = \frac{-4+\kappa+2\sqrt{4-2\kappa+\alpha\left(2\kappa-4+\alpha(1-s)\right)(1-s)}+2\alpha(1-s)}{k}$$

To match a quantity Q, we then have that subsidies must satisfy:

$$s_\kappa(Q) = \frac{4\alpha(2-Q)+Q(8-\kappa Q)-8}{4\alpha(2-Q)}$$

Now, to solve for $s_{o,\kappa}(Q)$, we simply solve for equilibrium with subsidies in the regular uniform quadratic example, obtaining:

$$s_{0,\kappa}(Q) = \frac{8\alpha-4\alpha Q-2\kappa Q+\kappa+8Q-8}{4\alpha(2-Q)}$$

The difference between the two levels of subisidies is then:

$$s_{0,\kappa}-s_\kappa(Q) = \frac{\kappa(1-Q)^2}{4\alpha(2-Q)}$$

This difference is strict for any $Q<1$, proving part 3. To prove part 4, just notice that (p,Q) are the same, so the only difference between the welfare in the two cases is in the subsidies, so that welfare must be higher when subsidies are smaller. The proof is then complete.

■

References

Assistant Secretary for Planning and Evaluation. (2021). Reaching the remaining uninsured: An evidence review on outreach & enrollment. *U.S. Department of Health and Human Services*, *1*.

Bar-Isaac, H., Caruana, G., and Cuñat, V. (2012). Search, design, and market structure. *American Economic Review*, *102*(2), 1140–60.

Bergemann, D., and Bonatti, A. (2011). Targeting in advertising markets: Implications for offline versus online media. *The RAND Journal of Economics*, *42*(3), 417–443.

Brown, Z. Y., and Jeon, J. (2020). *Endogenous information and simplifying insurance choice* (tech. rep.). Mimeo, University of Michigan.

Cusumano, C., Fabbri, F., and Pieroth, F. (2022). Competing to commit: Markets with rational inattention. *Available at SSRN 4138071*.

Dang, T. V. (2008). Bargaining with endogenous information. *Journal of Economic Theory*, *140*(1), 339–354. https://doi.org/https://doi.org/10.1016/j.jet.2007.09.006

Dworczak, P., and Martini, G. (2019). The simple economics of optimal persuasion. *Journal of Political Economy*, *127*(5), 1993–2048.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *The quarterly journal of economics*, *125*(3), 877–921.

Einav, L., Finkelstein, A., and Mahoney, N. (2021). The io of selection markets. In *Handbook of industrial organization* (pp. 389–426). Elsevier.

Finkelstein, A., Hendren, N., and Shepard, M. (2019). Subsidizing health insurance for low-income adults: Evidence from massachusetts. *American Economic Review*, *109*(4), 1530–67.

Garfield, R., Orgera, K., and Damico, A. (2019). The uninsured and the aca: A primer. *Kaiser Family Foundation. Retrieved November*, *1*, 2020.

Gentzkow, M., and Kamenica, E. (2016). A rothschild-stiglitz approach to bayesian persuasion. *American Economic Review*, *106*(5), 597–601.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, *103*(7), 2643–82.

Handel, B. R., and Kolstad, J. T. (2015). Health insurance for" humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, *105*(8), 2449–2500.

Handel, B. R., Kolstad, J. T., and Spinnewijn, J. (2019). Information frictions and adverse selection: Policy interventions in health insurance markets. *Review of Economics and Statistics*, *101*(2), 326–340.

Hefti, A. (2018). Limited attention, competition and welfare. *Journal of Economic Theory*, *178*, 318–359.

Inderst, R., and Ottaviani, M. (2012). Competition through commissions and kickbacks. *American Economic Review*, *102*(2), 780–809.

Johnson, J. P. (2013). Targeted advertising and advertising avoidance. *The RAND Journal of Economics*, *44*(1), 128–144.

Johnson, J. P., and Myatt, D. P. (2006). On the simple economics of advertising, marketing, and product design. *American Economic Review*, *96*(3), 756–784.

Kacperczyk, M., Van Nieuwerburgh, S., and Veldkamp, L. (2016). A rational theory of mutual funds' attention allocation. *Econometrica*, *84*(2), 571–626.

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

Landais, C., and Spinnewijn, J. (2021). The value of unemployment insurance. *The Review of Economic Studies*, *88*(6), 3041–3085.

Mahoney, N., and Weyl, E. G. (2017). Imperfect competition in selection markets. *Review of Economics and Statistics*, *99*(4), 637–651.

Martin, D. (2017). Strategic pricing with rational inattention to quality. *Games and Economic Behavior*, *104*, 131–145.

Matějka, F., and McKay, A. (2012). Simple market equilibria with rationally inattentive consumers. *American Economic Review*, *102*(3), 24–29.

Mensch, J. (2022). Screening inattentive buyers. *American Economic Review*, *112*(6), 1949–84.

Mensch, J., and Ravid, D. (2022). Monopoly, product quality, and flexible learning. *arXiv preprint arXiv:2202.09985*.

Panhans, M. (2019). Adverse selection in aca exchange markets: Evidence from colorado. *American Economic Journal: Applied Economics*, *11*(2), 1–36.

Pavan, A., and Tirole, J. (2022). Knowing your lemon before you dump it. *Working Paper*.

Ravid, D., Roesler, A.-K., and Szentes, B. (2022). Learning before trading: On the inefficiency of ignoring free information. *Journal of Political Economy*, *130*(2), 346–387.

Spinnewijn, J. (2017). Heterogeneity, demand for insurance, and adverse selection. *American Economic Journal: Economic Policy*, *9*(1), 308–43.

Taubinsky, D., and Rees-Jones, A. (2018). Attention variation and welfare: Theory and evidence from a tax salience experiment. *The Review of Economic Studies*, *85*(4), 2462–2496.

Thereze, J. (2022). Screening costly information. *Working Paper*.